ISO/IEC JTC 1/SC 42/WG 4

Use cases and applications

Convenorship: JISC (Japan)

Document type: Officer's Contribution

Title: TR 24030 working draft v8 clean

Status:

Date of document: 2019-12-09

Source: Project Editor

Expected action: INFO

Email of convenor: maruyama.f@fujitsu.com

Committee URL: https://isotc.iso.org/livelink/livelink/open/jtc1sc42wg4

ISO/IEC 24030:2019(E)

ISO/IEC JTC 1/SC 42/WG 4

Secretariat: ANSI

Information technology — Artificial Intelligence (AI) — Use cases

WD/CD/DIS/FDIS stage

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 42, *Artificial Intelligence*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document provides a collection of representative use cases of Artificial Intelligence (AI) applications in a variety of domains. The current document reflects contributions and discussions by ISO/IEC JTC 1 SC42 WG 4 experts and liaison members, and JTC 1 SC 42 national mirror committees.

In total 132 AI use cases were submitted by the end of August 2019. Experts from the following national committees and liaison organizations contributed use cases on AI: Austria, Canada, China, Germany, India, Ireland, Italy, Japan, Korea, Republic of, Russian Federation, Singapore, United Kingdom, JTC 1 SC 36, and JTC 1 SC 38.

The rationale for this document is as follows:

– Illustrating the applicability of the SC 42 program of work across a variety of application domains

– Input to and reference by SC 42 program of work

– Sharing the collected use cases in support of the SC 42 program of work with external organizations and internal entities to foster collaboration

– Reach out to new stakeholders interested in AI applicability

– Establishment of category C Liaisons to collect requirements for AI via use cases

To collect use cases, first step is to identify application domains of AI systems (described in clause 4) and to provide a use case template (described in sub-clause 5.2 and 5.3). Contributors were requested to submit use cases using the provided template.

For improving the quality of use case description, a guidance is provided for contributors. The guidance includes identified acceptable sources (described in sub-clause 5.4) and AI characteristics (described in sub-clause 5.4) for preparing use cases.

By investigating use cases, it is possible to find the new technical requirements (standardized demand) from the market, accelerating the transformation of science and technology achievements. In this document, sub-clause 5.6 includes basic statistics of use cases. Sub-clause 5.7 and sub-clause 5.8 describe the finding from use case analysis.

The use case template helped to group and categorize the use cases according to the identified application domains. In this document, use cases are summarized and grouped according to the application domains in clause 6. Readers of this document could find use cases that regard to desired application domain and could find original submissions of use cases in Annex A, which includes all submissions of use cases.

Title Information technology — Artificial Intelligence (AI) — Use cases

1 Scope

This document provides a collection of representative use cases of AI applications in a variety of domains.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 22989, *Artificial intelligence* -- *Concepts and terminology*

ISO/IEC 23053, *Artificial intelligence* -- *Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989, ISO/IEC 23053 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1 Terms defined elsewhere

None

3.2 Terms defined in this document

3.2.1

artificial intelligence (AI)

<system>capability of an engineered system to acquire, process and apply knowledge and skills

Note 1 to entry: knowledge are facts, information, and skills acquired through experience or education

[SOURCE: ISO/IEC 22989, 3.2.1.2]

<engineering discipline>discipline which studies the engineering of systems with the capability to acquire, process and apply knowledge and skills

Note 1 to entry: knowledge are facts, information, and skills acquired through experience or education

[SOURCE: ISO/IEC 22989, 3.2.1.3]

3.2.2

AI system

technical system that uses artificial intelligence to solve problems

[SOURCE: ISO/IEC 22989, 3.2.1.4]

3.2.3

anomaly detection

task of anomaly detection is to identify data instances that do not conform to an expected pattern, especially within data sets that appear to be homogeneous.

Note 1 to entry: Anomaly detection is useful for cases of fraud detection, detecting suspicious activities, etc.

Note 2 to entry: With anomaly detection, the training data is all of one class and the ML model predicts if a data point is typical for a given distribution or not.

Note 3 to entry: Anomaly detection typically employs unsupervised learning.

[Modified text based on: ISO/IEC 23053, 6.2.3.5]

3.2.4

application

software or a program that is specific to the solution of an application problem

[SOURCE: ISO/IEC 11801:2002, definition 3.1.2]

3.2.5

automation

Process that occurs when a machine does work that might previously have been done by a living being

Note 1 to entry: Automation relates to both physical work and mental or cognitive work.

[SOURCE: ISO/IEC 22989, 3.2.1.7]

3.2.6

bias

systematic difference between true (or accepted) value and measured value

[SOURCE: ISO 14488:2007(en), 3.1]

3.2.7

big data

extensive datasets (ISO/IEC 20546:2019(en), 3.1.11) — primarily in the data (ISO/IEC 20546:2019(en), 3.1.5) characteristics of volume, variety, velocity, and/or variability — that require a scalable technology for efficient storage, manipulation, management, and analysis

Note 1 to entry: Big data is commonly used in many different ways, for example as the name of the scalable technology used to handle big data extensive datasets.

[SOURCE: ISO/IEC 20546:2019(en), 3.1.2]

3.2.8

classification

task of assigning collected data to target categories or classes.

Note 1 to entry: Models can be created either for binary classification which is the prediction that data belongs to one of two different classes, or for multiclass classification where ML models learn to predict the category of an instance of data.

Note 2 to entry: An example of classification is to predict if a photograph of an animal is a cat or a dog or even a different species. Classification employs supervised learning.

[Modified text based on: ISO/IEC 23053, 6.2.3.3]

3.2.9

cloud

collection of networked remote servers

[SOURCE: ISO 20294:2018(en), 3.5.8]

3.2.10

computer vision

capability of a functional unit to acquire, process, and interpret visual data

Note 1 to entry: Computer vision involves the use of visual sensors to create an electronic or digital image of a visual scene.

Note 2 to entry: Not to be confused with machine vision.

Note 3 to entry: computer vision; artificial vision: terms and definition standardized by ISO/IEC [ISO/IEC 2382-28:1995].

Note 4 to entry: 28.01.19 (2382)

[SOURCE: ISO/IEC 2382:2015(en), 2123787]

3.2.11

data analysis

systematic investigation of the data and their flow in a real or planned system

[SOURCE: ISO/IEC 2382:2015(en), 2122686]

3.2.12

data set

identifiable collection of data (ISO/IEC 20546:2019(en), 3.1.5) available for access or download in one or more formats

[SOURCE: Adapted from ISO 19115-2:2009, 4.7]

[SOURCE: ISO/IEC 20546:2019(en), 3.1.11]

3.2.13

decision making

adoption and authorization of a project plan

[SOURCE: ISO/TR 21245:2018(en), 3.6]

3.2.14

deep learning

approach to creating rich hierarchical representations through the training of neural networks with many hidden layers

Note 1 to entry: In recent years, some of the most impressive advancements in machine learning have been in the subfield of deep learning, also known as deep neural network learning. Deep learning uses multi-layered networks of simple computing units (or “neurons”). In these neural networks each unit combines a set of input values to produce an output value, which in turn is passed on to other neurons downstream. Neural networks in Deep learning are composed of several hidden layers.

[SOURCE: ISO/IEC 23053, 3.13]

3.2.15

end user

individual person who ultimately benefits from the outcomes of the system

Note 1 to entry: The end user may be a regular operator of the software product or a casual user such as a member of the public.

[SOURCE: ISO/IEC 25000:2014(en), 4.7]

3.2.16

machine learning

process using computational techniques to enable systems to learn from data or experience

[SOURCE: ISO/IEC 23053, 3.16]

3.2.17

natural language processing

information processing based upon natural-language understanding

Note 1 to entry: NLP is a field of AI

Note 2 to entry: Natural language is any human language, such as English, Spanish, Arabic, or Japanese, to be distinguished from formal languages, such as Java, Fortran, C++, or First-Order Logic.

Note 3 to entry: examples of expression of natural language are text, speech, gestures and sign language

[SOURCE: ISO/IEC 22989, 3.2.1.22]

3.2.18

neural network

network of primitive processing elements connected by weighted links with adjustable weights, in which each element produces a value by applying a nonlinear function to its input values, and transmits it to other elements or presents it as an output value

Note 1 to entry: Whereas some neural networks are intended to simulate the functioning of neurons in the nervous system, most neural networks are used in artificial intelligence as realizations of the connectionist model.

Note 2 to entry: Examples of nonlinear functions are a threshold function, a sigmoid function, and a polynomial function.

Note 3 to entry: This entry is an improved version of the entry 28.01.22 in ISO/IEC 2382-28:1995.

Note 4 to entry: neural network; neural net; NN; artificial neural network; ANN: terms, abbreviations and definition standardized by ISO/IEC [ISO/IEC 2382-34:1999].

Note 5 to entry: 34.01.06 (2382)

[SOURCE: ISO/IEC 2382:2015(en)]

3.2.19

parameter

any characteristic that can help in defining or classifying a particular system

Note 1 to entry: i.e. a parameter is an element of a system that is useful or critical when identifying the system or when evaluating its performance, status or condition.

3.2.20

pattern recognition

identification, by a functional unit, of physical or abstract patterns, and of structures and configurations

Note 1 to entry: This is an improved version of the definition in ISO/IEC 2382-12:1988.

Note 2 to entry: pattern recognition: term and definition standardized by ISO/IEC [ISO/IEC 2382-28:1995].

Note 3 to entry: 28.01.13 (2382)

[SOURCE: ISO-IEC-2382-28 * 1995 * * *]

[SOURCE: ISO/IEC 2382:2015(en), 2123781]

3.2.21

quality

conformance to specified requirements

[SOURCE: ISO 13628-2:2006(en), 3.33]

3.2.22

retraining

generation of new trained parameters in a trained model through training by applying different training data

3.2.23

robot

programmed actuated mechanism with a degree of autonomy, moving within its environment, to perform intended tasks

Note 1 to entry: A robot includes the control system and interface of the control system.

Note 2 to entry: The classification of robot into industrial robot or service robot is done according to its intended application.

Note 3 to entry: In order to properly perform its tasks, a robot makes use of different kinds of sensors to confirm its current state and perceive the elements composing the environment in which it operates.

[Modified text based on ISO 18646-2:2019(en), 3.1]

[SOURCE: ISO/IEC 22989, 3.2.1.31]

3.2.24

service

performance of activities, work, or duties

Note 1 to entry: A service is self-contained, coherent, discrete, and can be composed of other services.

Note 2 to entry: A service is generally an intangible product.

[SOURCE: ISO/IEC/IEEE 12207:2017(en), 3.1.50]

3.2.25

task

activities required to achieve a goal

Note 1 to entry: These activities can be physical and/or cognitive.

[SOURCE: ISO 9241-11:1998, definition 3.9]

3.2.26

trained model

final deliverable generated by training process using training data

3.2.27

training data

subset of available data used to fit a machine learning model (ISO/IEC 23053, 3.6)

[SOURCE: ISO/IEC 23053, 3.8]

3.3 Abbreviated terms

2D two-Dimensional

3D three-Dimensional

5G 5th Generation

ACC Adaptive Cruise Control

ACU Air Control Unit

AF Atrial Fibrillation

AI Artificial Intelligence

AI/ML Artificial Intelligence/Machine Learning

AMI Advanced Metering Infrastructure

AMR Adaptive Mesh Refinement

ANN Artificial Neural Networks

API Application Programming Interface

AR Augmented Reality

AS Active substances

ATC Air Traffic Controllers

AUC Area Under the Curve

AUC Appropriate Use Criteria

AWS Amazon Web Services

BAS Building Automation System

BDEC Big Data and Extreme-scale Computing

BIOSIS BioSciences Information Service of Biological Abstracts

BMS Building Management System

BNN Binarized Neural Network

BOSS Business Operations Support System

CACC Cooperative Adaptive Cruise Control

CAPEX Capital Expenditure

CART Classification and regression trees

CDSS Clinical Decision Support System

CG Computer Graphics

CHD Coronary heart Disease

C-Lab Creative Lab

CNN Convolutional Neural Network

CPE Customer Premises Equipment

CPU Central Processing Unit

CR Clinical Recommendations

CR Checkpoint/Restart

CRWB Cooking Recipes without Border

CS-DC Complex Systems Digital Campus

CSE Computational Science & Engineering

CSP Cloud Service Provider

CT Computed Tomography

CV Computer Vision

CVD Cardiovascular Disease

DDA Data Driven Approach

DDC Direct Digital Control

DICOM Digital Imaging and COmmunications in Medicine

DL Deep Learning

DNN Deep Neural Network

DOP Department of Police

DOT Department of Transportation

DVB Digital Video Broadcasting

DVB-S Digital Video Broadcasting - Satellite

DVB-S2 Digital Video Broadcasting - Satellite - Second Generation

DVB-S2x DVB-S2 Extensions

DW Data Warehouses

E2E end to end

EC2 Amazon Elastic Compute Cloud

ECG Electrocardiogram

ECS Amazon Elastic Container Service

EDW Enterprise Data Warehouses

EMR Electronic Medical Record

EncDec-AD Encoder-Decoder scheme for Anomaly Detection

ENSEMBLE Expectation and Non-formal Skills to Empower Migrants and to Boost Local Economy

EO Electro-optical

EPP Environmentally Preferable Purchasing

ET Evolutionary Technology

FAR False Acceptance Rate

FBI Federal Bureau of Investigation

FCV Flow Control Valves

FFS Fast File System

FG Pharm group

FLAC Fourier Local Auto Correlation

FM Facilities Management

FMSI Facility Master System Integrator

FPGA Field-Programmable Gate Array

GAN Generative Adversarial Nets

GB Giga Byte

GDPR General Data Protection Regulation

GIS Geographic Information System

GLM Generalized Linear Model

GPCR G protein-coupled receptor

GPS Global Positioning System

GPU Graphics Processing Unit

GUI Graphical User Interface

HAN Hierarchical Attention Networks

HDD hard disk drive

hDDA Hierarchical Data Driven Approach

HLAC Higher-order Local Auto Correlation

HPC High performance computing

HTS High-throughput satellite

HV High-voltage

HVAC Heating, Ventilation, and Air Conditioning

ICD International Classification of Diseases

ICT Information and Communication Technology

IEC International Electrotechnical Commission

IIT-Delhi Indian Institute of Technology Delhi

IMEI International Mobile Equipment Identity

IMF International Monetary Fund

IMU Instructions for Medical Usage of Drugs

INSPEC Institute of Engineering and Technology

IoT Internet of Things

IP Internet Protocol

IR infrared

ISA International Society of Automation

ISO International Organization for Standardization

ITS Intelligent Transportation Systems

IUT Institute of Technology

JMA Japan Meteorological Agency

JSAI Japanese Society of Artificial Intelligence

KNN K-Nearest Neighbor

KPI Key Performance Indicator

KYC Know Your Customer

LASIK Laser-Assisted in SItu Keratomileusis

LDA Linear discriminant analysis

LIDAR Light Detection and Ranging; Laser Imaging Detection and Ranging

LMD Lift Monitoring Device

LSTM Long Short Term Memory Networks

LSTM-AD Long Short Term Memory Networks for Anomaly Detection

LVPEI L. V. Prasad Eye Institute

M.O.S Mean Opinion Score

MAE Mean Absolute Error

MEM Multi-electrode Mapping

MES Manufacturing Execution System

MIoU Mean Intersection over Union

ML Machine Learning

ML/DL Machine Learning and Deep Learning

MLC Multi-level Checkpoint

MND Mobile phone Network Data

MODLE Mobility on Demand Laboratory Environment

MOOCs Massive Open Online Courses

MOR Model Order Reduction

MPCA Multilinear Principal Component Analysis

MRI Magnetic Resonance Imaging

MRI Meteorological Research Institute

MTS Mobile TeleSystems

MW Mega Watt

NB Naïve Bayes algorithm

NDA	Non-disclosure agreement
NIED	National Research Institute for Earth Science and Disaster Resilience
NIOM	Near-infrared Optical Mapping
NIR	Near InfraRed
NLP	Natural Language Processing
NLU	Natural Language Understanding
NPU	Neural Network Processing Unit
NTPC	National Thermal Power Corporation
O&M	Operation & Maintenance
OLAP	Online Analytical Processing
Online-AD	Online Anomaly Detection
OPEX	Operating Expense
OWL	Web Ontology Language
PACS	Picture Archiving and Communication Systems
PC	Personal Computer
PCA	Principal Component Analysis
PII	Personally Identifiable Information
PoC	Proof of Concept
POI	Point of Interest
PROCAM	Prospective Cardiovascular Munster
PSNR	Peak Signal-to-noise Ratio
QA	Quality Assurance
QC	Quality Control
R&D	Research and development
RADAR	Radio Detection and Ranging
RAIMS	Research Association for Infrastructure Monitoring System
RAM	Random Access Memory
RDF	Resource Description Framework
RGB	Red Green Blue

RGB-D Red Green Blue Depth

RMSE Root Mean Square Error

RNN Recurrent neural networks

RNN-AD Recurrent neural networks for Anomaly Detection

RoI Return on Investment

RSU Reves de Scenes Urbaines

SaaS Software as a Service

SCADA Supervisory Control And Data Acquisition

SCORE strategies concentrating on risk evaluation

SDGs Sustainable Development Goals

SEM Scanning Electron Microscope

SFS Shared File System

SID Infant Death Syndrome

SIM Simulation Nodes

SIS Swarm Intelligence System

SLAs Service Level Agreement

SLR Single-lens reflex camera

SMS Short Message Service

SNR Signal-to-noise Ratio

SQL Structured Query Language

SRGAN Super-Resolution GAN

SSIM Structural Similarity

SVM Support Vector Machine

TB Tera Byte

TCO Total Cost of Ownership

TPA Third Party Administrator

TPU Tensor processing unit

t-SNE T-distributed Stochastic Neighbor Embedding

UAV Unmanned Aerial Vehicle

UCVA Uncorrected Visual Acuity

UI User Interface

UMTS Universal Mobile Telecommunications System

UQ Uncertainty Quantification

USB Universal Serial Bus

UT Ultrasonic Testing

UTO Unattended Train Operation

V2V Virtual to Virtual

VAE Variational Auto Encoder

VC Vital Characteristics

VFD Variable Frequency Device

WAN Wide Area Network

WTTx Wireless To The x

xAPI Experience API

xGBM Extreme Gradient Boosting Machine

XML Extensible Markup Language

4 Applications

4.1 General

While it started a bottom-up approach from collecting use cases, this document takes a top-down approach, to identify AI applications from the perspectives of their deployment models and application domains of their use, as well in parallel.

4.2 Application domains

This document considers the use of AI applications that are described in [5] and [6] to collect application domains. 24 application domains that are list as follows are considered as target domains to collect use cases:

Agriculture, Construction, Defence, Digital marketing, Education, Energy, Fintech, Healthcare, Home/Service Robotics, ICT, Knowledge management, Legal, Logistics, Low-resource Communities, Maintenance & support, Manufacturing, Media and Entertainment, Mobility, Public sector, Retail, Security, Social infrastructure, Transportation, Work & life

4.3 Deployment models

This document considers the use of AI applications ([5]) and list passible deployment models of AI applications as follows:

Cloud services, On-premise systems, Embedded systems, Cyber-physical systems, Social networks, Hybrid

4.4 Examples of AI Application

Table 1 —Examples of AI Application lists examples of AI applications. These application examples were derived from "Artificial Intelligence White Paper" ([5]). Each example in Table 1 —Examples of AI Application has application domain, deployment mode and short description.

The abbreviations used in Table 1 —Examples of AI Application are defined as follows:

— CL:Cloud

— ES:Embedded System

— OP:On-premise

— CP:Cyber-phyusical

— SN:Social Net

— HY:Hybrid or other

Table 1 —Examples of AI Application

Application domain	Deployment mode	Application	Short description
Manufacturing	CL, OP, HY	Development Design	CL: Accumulate Design Patterns to help Designer OP: Check design pattern with real constraints on premise
Manufacturing	CL, OP, HY	Production process	CL: Accumulate production quality actuation pattern and estimate the performance of quality OP: Accumulate production through put related parameters and estimate the output throughput
Manufacturing	On-premise systems	Product quality inspection	Inspect products by image recognition
Mobility	CL, ES, OP,	Autonomous driving	Mainly realized cars. CL: Update cruising control software dynamically ES: Enablize autonomous driving without any help from connected devices. CL: Accumulate road condition and disseminate them to autonomous agents

Mobility	CL, ES, OP,	Automatic cruise control	Mainly enablized at wheel chairs, ships, and autonomous robots CL: Update cruising control software dynamically ES: Enablize autonomous driving without any help from connected devices. CL: Accumulate road condition and disseminate them to autonomous agents
Mobility	CL, ES, OP,	Robot Taxi	CL: Pick-up arrangement system controls robot taxis effectively EM: Autonomously drive through the road with dynamic control of steering and acceleration and breaking. OP: Autonomously drive through the road with road map
Mobility	CL, OP, HY	Dynamic map for autonomous cruise control	CL: Create, maintain, and disseminate map information with semantic tags with real-time communication with mobile agents such as cars, wheelchairs, robots, and human beings. OP: Accumulate the real road situation and recognize the objects which is not involved in the map to be shared.
Mobility	Cyber-physical systems	City-wide traffic control	Optimize city-wide traffic by inspecting real-time traffic image and controlling traffic signals
Social infrastructure	CL, OP, HY	Abnormality or malfunction prediction	CL: Accumulate normal signal patterns to learn normal signals OP: Find out abnormal signal patterns on premise
Social infrastructure	CL, OP	Equipment operation	CL: Accumulate operational parameters and learn normal operations OP: Monitor operation and find out abnormal operation patterns
Agriculture	CL, OP	Agricultural automation	CL: Monitor the field condition and manage the field condition OP, ES: Accumulate weeds or insects pattern and eliminate them
Agriculture	CL, OP	Craftsmanship skill transfer	CL: Learn about best practices by craftsmen, and feedback them to others
Agriculture	OP	Cultivation management	OP: Monitor the field condition and manage irrigation condition
Healthcare	CL, OP	Diagnosis support	CL: Learn about normal condition OP: Find out abnormal condition compared with normal condition
Healthcare	CL	New drug development	Curation: CL: Find out the co-relation among submitted papers Molecular pattern: CL: Find out the effective coordination of target molecular

Healthcare	CL, OP, ES	Surgical automation	CL: Accumulate disease patterns and healthy patterns OP: Support identification of disease patterns on premise
Healthcare	CL, OP, ES, HY	Medical Platform	Accumulate and disseminate the learning patterns and assistants as an integrated medical support system.
Security	CL	Cyber Security	Monitoring transactions in cyber space and find out attacks through finding abnormal transaction patterns
Security	CL	Personal Information Management	Monitoring operations for GDPR conformance assurance
Security	CL	Video Surveillance & Crime risk prediction	Monitor the behavioral patterns in town and predict crime risk and find out criminal patterns.
Defense	CL	Cyber Security	Monitoring cyber transactions against important defense assets and find out attack patterns and prevent their intrusion.
Defense	CL, OP, ES	Electronic warfare	Autonomous pilot with cloud support to enablise the electronic warfare
Social infrastructure	CL, OP	Landslide, flood prediction	Monitoring weather and ground condition in realtime and predict the disaster such as landslide, flood etc.
Social infrastructure	Cloud services	Power demand forecasting	Learn about demand patterns with other significant parameters and forecast the future demand
Social infrastructure	Cloud services	Improving operational efficiency	Learn about the co-relation among significant parameters and manage to realize the most efficient operations. Traffic control, Electricity supply control, etc.
Education	On-premise systems	Adaptive learning	Through using learning model, provide personalized learning materials to achieve the efficient learning results.
Education	On-premise systems	Scoring	Through using the cognitive learning model, make the most effective feed back to the learners to achieve the most effective learning results.
Fintech	Cloud services	Stock exchange and trading	Accumulate the best practices and enablize the 7week 24 hours trading
Fintech	Cloud services	Asset management	Accumulate and learn about the best practices and apply them to realize the customer satisfaction

Fintech	Cloud services	Loan screening	Learn about the normal backgrounds of customers to find out the abnormal loan patterns
Fintech	Cloud services	Fraud identification	Identify fraud transactions and make warning to the managers
Fintech	Cloud services	Security assurance against cyber attacks	
Logistics	CL OP	Procurement logistics	Analyze the procurement context and propose the best procurement actions.
Logistics	CL OP	Logistics in the base	Coordinate the best logistic move in the local procurement base warehouse
Logistics	CL OP	Sales logistics	Analyze and learn about the best practices of sales logistics and provide the most effective routs to sales move.
Construction	Cloud services	Construction planning	Learn about the best practices and apply them to coming planning
Construction	CL OP	Robot construction	Provide autonomous construction robot to the construction sites
Retail	On-premise systems	Autonomous driving store	Provide autonomous driving sales robot
Retail	On-premise systems	Register less store	Monitor all the moves of each customers to realize cash register less retail shops.
Digital marketing	Cloud services	Online campaign performance optimization	As we have in GAFA
Legal	Cloud services	Early case assessment	AI support the work preps had been doing
Legal	Cloud services	Judicial recommendation	Judge support by using the previous judicial judgement cases
Public sector	CL OP	Public service matching	Optimize matching between residents and public services
Public sector	CL SN	Online service support	Provide residents with support for online services
Work & life	Embedded systems	Smart home appliances	Include robot vacuums and refrigerators and air conditioners with sophisticated control
Work & life	Social networks	Smart personal agent	Smart agents assist individual users

**Note: Collected use cases are not necessarily addressing all areas in this table.

5 Use cases

5.1 Introduction

This document collected 132use cases. Sub-clauses 5.2 Properties and 5.3 Template describes a template that is used for collecting use cases and show a blank template. Then this document give some basic statistics of collected 132 use cases in sub-clause 5.6 Basic statistics.

5.2 Properties

5.2.1 General

General information of the use case

— Use case name: Use case name provided by the use case contributor

— Application domain: Refers to 4.2 application domains

— Deployment models: Refers to 4.3 deployment models

— Status: The status of the use case, includes Prototype, PoC (Proof of Concept), or in-operation

— Scope: The scope defines the intended area of applicability, limits, and audience.

— Objective(s): The intention of the system; what is to be accomplished?; who/what will benefit?.

— Narrative: Descriptions(short and complete) of the use case

— Stakeholders: Stakeholder are those that can affect or be affected by the AI system in the scenario; e.g., organizations, customers, 3rd parties, end users, community, environment, negative influencers, bad actors, etc.

— Stakeholders' assets, values: Stakeholders' assets and values that are at stake with potential risk of being compromised by the AI system deployment – e.g., competitiveness, reputation, trustworthiness, fair treatment, safety, privacy, stability, etc.

— System's threats and vulnerabilities: Threats and vulnerabilities can compromise the assets and values above - e.g., different sources of bias, incorrect AI system use, new security threats, challenges to accountability, new privacy threats (hidden patterns), etc.

— Key performance indicators (KPIs): Descriptions of KPIs for evaluating the performance or usefulness of use cases. Descriptions include KPI's name, description of the KPI and reference to mentioned use case objectives

— AI features: Descriptions of features of use case in AI consideration. Descriptions include:

1) Task(s): The main task in use case. A pull-down list includes the following terms: Recognition, Natural language processing, Knowledge processing & discovery, Inference, Planning, Prediction, Optimization, Interactivity, Recommendation or Other

2) Method(s): AI method(s)/framework(s) used in development.

3) Hardware: Hardware system used in development and deployment.

4) Topology: Topology of the deployment network architecture.

5) Terms and concepts used: Terms and concepts used here should be consistent with those defined by ISO/IEC 22989 and ISO/IEC 23053 or to be recommended for inclusion.

— Standardization opportunities/requirements: Descriptions of Standardization opportunities/ requirements that are derived from the use case.

— Challenges and issues: Descriptions of challenges and issues in the use case

— Societal concerns:

1) Description: Description of societal concerns that are derived from the use case.

2) SDGs to be achieved: The Sustainable Development Goals (SDGs), otherwise known as the Global Goals, are a collection of 17 global goals set by the United Nations General Assembly. SDGs are a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity.

5.2.2 References

References related to the use case

— Type: Document type of the reference (e.q. standards, paper, patent, press release)

— Reference: Title of the reference

— Status: The status of the referenced document.

— Impact on use case: Where does the document influence the use case?

— Originator/organization: Who published the document?

— Link: If available, a public link can be provided.

5.3 Template

Table 2 — General part of use case template and Table 3 — Reference part of use case template are used for collecting use cases. The terms used in that template were defined in 5.2.

The template is based on:

— ISO/IEC 20547-2: Big data reference architecture – Part2

— IEC 62559: Use case methodology

— IEEE P7003: Use case template

It was intended to be augmented by "process" part, training, evaluation, execution, and refraining.

Table 2 — General part of use case template

ID	
Use case name	
Application domain	
Deployment	

<table>
<tr><td>Model</td><td colspan="4"></td></tr>
<tr><td>Status</td><td colspan="4"></td></tr>
<tr><td>Scope</td><td colspan="4"></td></tr>
<tr><td>Objective(s)</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3"></td></tr>
<tr><td>Complete description</td><td colspan="3"></td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4"></td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3"></td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3"></td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Societal concerns</td><td>Description</td><td colspan="3"></td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3"></td></tr>
</table>

Table 3 — Reference part of use case template

<table>
<tr><td colspan="7">References</td></tr>
<tr><td>No.</td><td>Type</td><td>Reference</td><td>Status</td><td>Impact on use case</td><td>Originator/organization</td><td>Link</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

5.4 Acceptable Sources of Use Case

For improving the quality of use case description, acceptable sources are:

— Peer-reviewed scientific/technical publications on AI applications (e.g. [1]).

— Patent documents describing AI solutions (e.g. [2], [3]).

— Technical reports or presentations by renowned AI experts (e.g. [4])

— High quality company whitepapers and presentations

— Publicly accessible sources in sufficient detail

This list is not exhaustive. Other credible sources may be acceptable as well.

5.5 Use Case Selection Guidance

For preparing use cases that cover both the most important application areas and the most relevant AI technologies, use case contributors can consider the following AI characteristics as useful selection guidance:

— Data Focus & Learning: Use Cases for AI system which utilizes Machine Learning, and those who use a fixed a-priory knowledge base.

— Level of Autonomy: Use cases demonstrating several degrees (dependent, autonomous, human/critic in the loop, etc.) of AI system autonomy.

— Verifiability & Transparency: Use cases demonstrating several types and levels of verifiability and transparency, including approaches for explainable AI, accountability, etc.

— Impact: Use cases demonstrating the impact of AI systems to society, environment, etc.

— Architecture: Use cases demonstrating several architectural paradigms for AI systems (cloud, distributed AI, Crowdsourcing, Swarm Intelligence)

5.6 Basic statistics

5.6.1 Application domain

Figure 1 — Distribution of use cases by application domains describes the percentage of use cases by application domain. This figure did not include the following application domains because these did not have any use cases:

— Construction, Defence, Knowledge management, Low-resource Communities

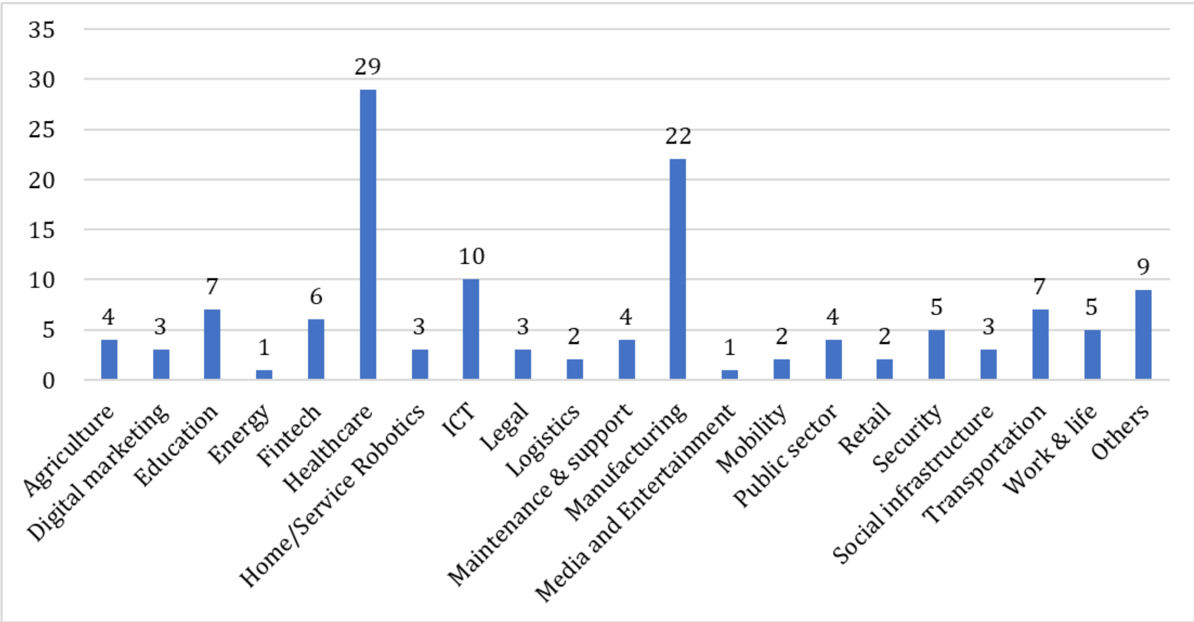


Figure 1 — Distribution of use cases by application domains

5.6.2 Status

Figure 2 — Distribution of use cases by status describes the percentage of use cases by status.

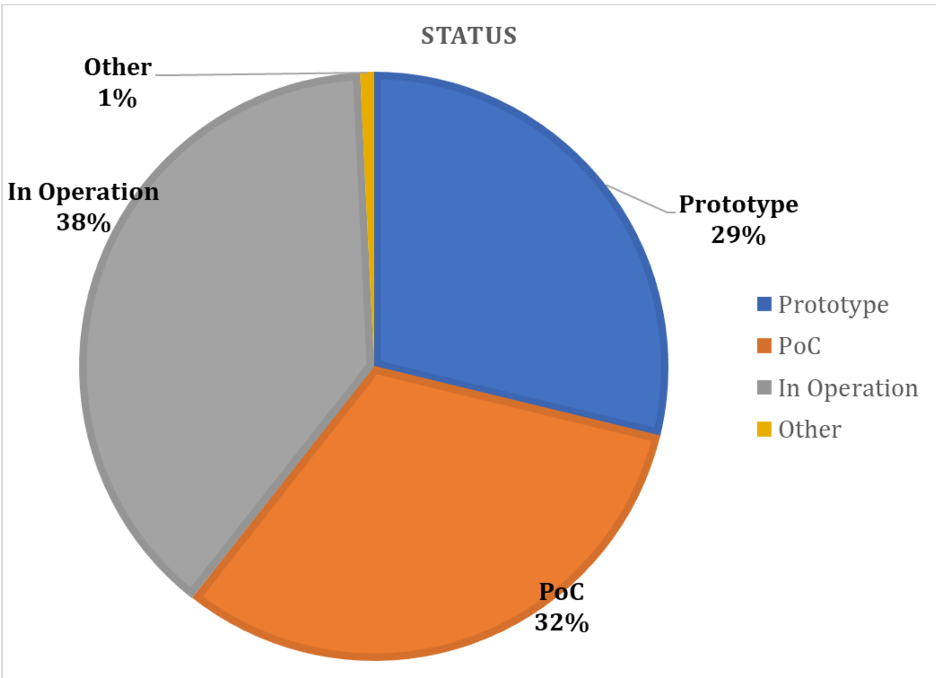


Figure 2 — Distribution of use cases by status

5.6.3 AI features (Task)

Figure 3 — Distribution of use cases by AI task describes the percentage of use cases by AI features (Task).

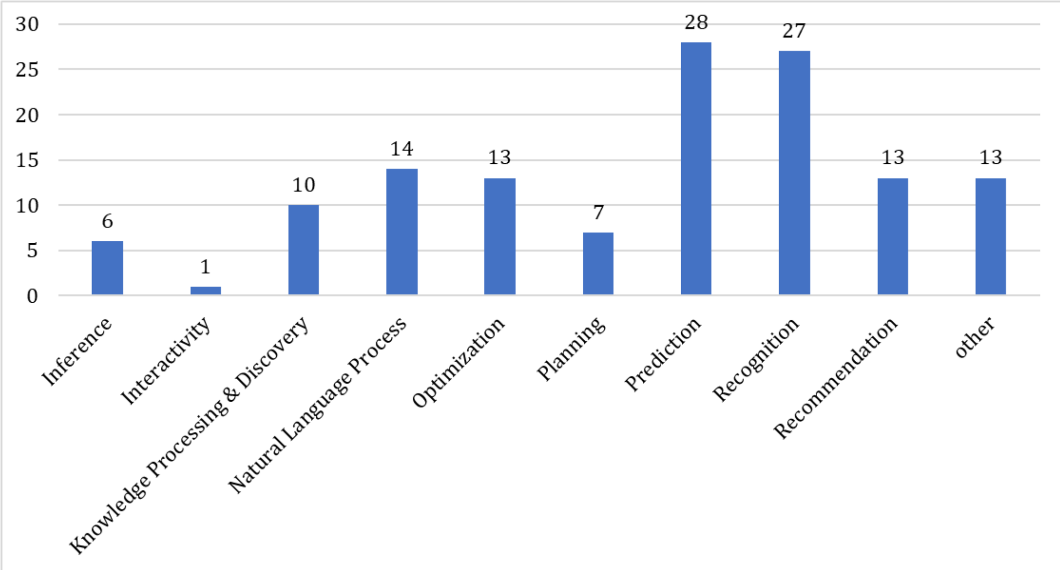


Figure 3 — Distribution of use cases by AI task

5.7 Societal concerns

5.7.1 Impact Analysis

AI systems can be compared based on the potential impact of failure or unexpected behaviour. AI systems fall on a spectrum of risk, determined by the severity of the potential impact of a failure or unexpected behaviour. Relevant aspects for assessing the level of risk include: (1) The type of action space the system is operating in (e.g. recommendations vs direct action in an environment); (2) The presence/absence of external supervision; (3) The type of external supervision (automated or manual); (4) The ethical relevance of the task and/or domain; (5) The level of transparency of decisions or processing steps; (6) The degree of system autonomy.

For example, a system that only gives recommendations and cannot act on its own, in a domain that has no ethical relevance, and has no autonomy, would be considered low-risk. Conversely, we would consider a system to be highly critical if its actions have direct impact on human lives, it operates autonomously and without external supervision, and its decision-making is opaque.

Case studies (Annex B) submitted covered numerous topics, from manufacturing, healthcare, to mobility and transportation. Few of the use cases include a description of any societal concern. The details in the use cases are varied making it difficult to fully analyze the application and its implications. The benefits are also rarely defined. There was a disproportionate number of cases in manufacturing and mobility. The earlier cases studies are generally more detailed and complete than later ones. Few of the case studies included identified of stakeholders. There are a few "use cases" that are not an application of an AI system, and thus, cannot be analyzed for societal impacts.

Societal impacts are generally not well defined in the use cases submitted. Vulnerable users are not generally identified. The explainability of the technology or use case is not present. The potential biases and equity issues are not presented. Data and privacy issues are not presented. Potential interactions and feedbacks that may have adverse impacts are not presented.

5.8 Findings

5.8.1 Use Case Analysis

5.8.1.1 Introduction

To reflect on the context of the work of SC 42 and determine its relevance to real-world AI applications, and that concerns/expectations of key stakeholders of AI standardization are brought up explicitly, one use case was analyzed considering a six-step process, and mapping the quality criteria to various aspects of use cases.

The quality criteria comprises of the use case selection guidance (refer to 5.5 Use Case Selection Guidance) and analyzing the inputs in the "Standardization Opportunities and Requirements" section to extract useful directions for SC 42. This is the area that differentiate this document from other documents in the open domain.

5.8.1.2 Approach to Use Case Analysis

This analysis considers a six-step process as follows:

Step 1: Use the AI definition from ISO/IEC 22989 (refers to 3.2.1). Identify the different components of this definition in each use case.

Identify other foundational, trustworthiness, societal concerns, and life cycle elements corresponding to acquire/process/apply aspects in the use cases. Identify the additional essential requirements for AI software products as compared to conventional non-AI software products.

To come up with these requirements, each AI application can be evaluated considering following three scenarios:

1) AI: Solution using existing AI approach

2) Human expertise: Solution using human expertise (but no AI software)

3) Non-AI: Solution which doesn't use any of the above two options

Each of the three options can be analyzed w.r.t. acquire, process and apply requirements to explicitly address differences in requirements for AI application.

Step 2: Do the missing data (not filled in) on the submission template affect the quality and understanding of the use case?

Step 3: Apply the quality criteria (refers to 5.8.1.1 Introduction) to the use cases.

Step 4: Identify use cases that have insufficient data or did not satisfy the selection criteria.

Step 5: Identify five use cases from the collection that are deemed illustrative and best exemplify AI applications. These could be used as samples for potential submitters.

Step 6: Analyze in detail w.r.t. scenarios mentioned in step 1 and categorize the requirements to be specifically directed to other existing working groups.

5.8.1.3 Use Case Analysis Result

The result of use case analysis is as follows:

Use Case: Machine Learning Tools in Support of Transformer Diagnostics (A.51)

Following three tables analyze the use case to bring out specific requirements w.r.t. AI standardization.

Table 4 — Use case scenarios as per step 1 (5.8.1.2) and selection guidance: Level of Autonomy

Scenario 1: AI	Scenario 2: Human Expert	Scenario 3: Non-AI
Continuous monitoring and control to improve transformer diagnostics based on insights from AI.	Operate as per best practices evolved over a period of time based on experience of an expert.	Maintenance as prescribed by OEM based on design and average operating conditions.

Table 5 — Analysis of Acquire-Process-Apply framework for different scenarios*

Scenario	Acquire	Process	Apply
1	Operation hours of transformer + information on critical operation parameters + **additional information on temperature, vibration, noise, and other operational details for multiple transformers.** **(selection guidance: Data Focus & Learning)**	**Artificial Neural Network** **(selection guidance: Data Focus & Learning)**	**Partly explainable and improved maintenance schedule (over and above scenario 2) that might have probability to lead to safety and operational issues leading to grid failure.** **(selection guidance: Verifiability & Transparency Data Focus & Learning)**
2	Operation hours of transformer + information on critical operation parameters	Reasoning based on expert analysis	Improved operation performance, based on maintenance trigger, that can be explained through expert reasoning
3	Operation hours of transformer	Simple logic as prescribed by OEM based on physics-based approach of transformer design and operation	Maintenance trigger based on the rules set by OEM

** AI specific aspects are marked in bold letters.*

Table 6 highlights important issues that are to be addressed by standardization and the closest reference to ISO/IEC 22989 and ISO/IEC 23053.

Table 6 —Issues to be addressed by standardization (for items highlighted in bold in Table 5) and reference in ISO/IEC 22989 and ISO/IEC 23053 [Document: Clause]

	Acquire	**Process**	**Apply**
General	Missing data treatment [ISO/IEC 23053: 3.23, 8.3(item “Imputation”)][1] [ISO/IEC 22989: 7.6][2]	Criteria for training [ISO/IEC 23053: 8.4(item “Model Evaluation”)][1]	Explainability of results [ISO/IEC 22989: 8.23.7][2]
	Choice of data frequency [ISO/IEC 23053:3.27][2]	Criteria for validation [ISO/IEC 23053:8.4(item “Model Evaluation”)][1]	Acceptable output for commercialization [ISO/IEC 23053:8.4(item “Model Evaluation”)][1]
	Establishing bias in input data [ISO/IEC 22989:5.2][2]	Criteria for retraining [ISO/IEC 22989:3.2.1.11(item “continuous learning”)][1]; [ISO/IEC 23053: 8.4(item “Model Evaluation”)][1]	Fail safe mode of operation against biases, safety, health and environment impact [ISO/IEC 22989:8.23][2]
	Minimum size for training, data sufficiency [ISO/IEC 23053:3.27][2]	Implementation in existing systems [ISO/IEC 23053:8.5, 8.6][1]	Risk-impact assessment considering all stakeholders [ISO/IEC 22989:8.23][2]
	Context definition [ISO/IEC 22989:: 5.1(item “Contextual”)][2]		
Specific	Capturing maintenance and event related data [ISO/IEC 22989:5.2][2]		
	Aligning inputs from multiple equipments [ISO/IEC 22989:5.2][2]		
	Definition of fault and error codes [ISO/IEC 22989:5.2][2]		
	Input characterization [ISO/IEC 22989:5.2][2]		

	[ISO/IEC 23053:8.4 (item “Feature Engineering”)] [1]		

([1]Aspects that are addressed and need validation; [2]Aspects that are partly addressed and need discussion)

Selection guidance: Impact (Societal Concerns):

— Safe and reliable power delivery

Selection guidance: Architecture:

— On premise system

6 Use cases summaries

6.1 Basic information of use cases

Table 7 — List of use cases describes basic information of use cases, which includes use case name, application domain, deployment mode and status.

Table 7 — List of use cases

Correspond ingsub-clause number	Use case name	Application domain	Deployment Model	Status
6.2.1	AI to understand adulteration in commonly used food items	Agriculture	Cloud services	PoC
6.2.2	bioBotGuard	Agriculture	Cloud services	PoC
6.2.3	Ecosystems management from causal relation inference from observational data	Agriculture, Knowledge management, ICT	Cloud Services, On-premise systems, Embedded systems, Hybrid	PoC
6.2.4	Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI	Agriculture	Embedded system	In operati on
6.3.1	Improving conversion rates and RoI (Return on Investment) with AI technologies	Digital marketing	On-premise systems	In operati on
6.3.2	Logo and Trademark Detection	Digital Marketing	Hybrid	PoC
6.3.3	Flavorlens	Digital Marketing	Cloud services	Prototy pe
6.4.1	VTrain recommendation engine	Education	On-premise systems	In operati on
6.4.2	RAVE	Education	Hybrid	PoC

6.4.3	IFLYTEK Intelligent marking system	Education	On-premise systems	In operation
6.4.4	Intelligent educational robot	Education	On-premise systems	In operation
6.4.5	AI solution to intelligence campus	Education	Cloud services	In operation
6.4.6	AI Adaptive Learning Platform for Personalized Learning	Education	Cloud services	In operation
6.4.7	AI Adaptive Learning Mobile App	Education	Hybrid	In operation
6.5.1	AI-dispatcher (operator) of large-scale distributed energy system infrastructure	Energy	On-premise systems	PoC
6.6.1	Detection of frauds based on collusions	Fintech	On-premise systems	In operation
6.6.2	Credit scoring using KYC data	Fintech	On-premise systems	PoC
6.6.3	Virtual Bank Assistant	Fintech	Cloud services	In operation
6.6.4	Forecasting prices of commodities	Fintech	On-premise systems	In operation
6.6.5	Finance Advising and Asset Management with AI	Fintech	Cloud service	In operation
6.6.6	Loan in 7 minutes	Banking and Financial Services	On-premise systems	In operation
6.7.1	Explainable artificial intelligence for Genomic Medicine	Healthcare	Cloud services	Prototype
6.7.2	Revolutionizing clinical decision-making using artificial intelligence	Healthcare	On-premise systems	PoC
6.7.3	Computer-aided diagnosis in medical imaging based on machine learning	Healthcare	On-premise systems	PoC
6.7.4	AI solution to predict Post-Operative Visual Acuity for LASIK Surgeries	Healthcare	Cloud services	In operation
6.7.5	Chromosome Segmentation and Deep Classification	Healthcare	Hybrid	PoC
6.7.6	AI solution to quality control of Electronic Medical Record(EMR) in real time	Healthcare	Cloud services	In operation
6.7.7	Dialogue-based social care services for people with mental illness, dementia and the elderly living alone	Healthcare	Hybrid	Prototype
6.7.8	Pre-screening of cavity and oral diseases based on 2D digital images	Healthcare	Hybrid	Prototype

6.7.9	Real-time patient support and medical information service applying spoken dialogue system	Healthcare	Hybrid	Prototy pe
6.7.10	Integrated recommendation solution for prosthodontic treatments	Healthcare	Hybrid	Prototy pe
6.7.11	Infant SID	Healthcare	Cloud services	Prototy pe
6.7.12	Discharge Summary Classifier	Healthcare	On-premise systems	In operati on
6.7.13	Generation of Clinical Pathways	Healthcare	On-premise systems	In operati on
6.7.14	Hospital Management Tools	Healthcare	On-premise systems	In operati on
6.7.15	Predicting relapse of a dialysis patient during treatment	Healthcare	Cloud services	In operati on
6.7.16	Instant triaging of wounds	Healthcare	Cloud services	In operati on
6.7.17	Accelerated acquisition of magnetic resonance images	Healthcare	Hybrid	Prototy pe
6.7.18	AI based text to speech services with personal voices for speech impaired people	Healthcare	On-premise systems	Prototy pe
6.7.19	AI Platform for Chest CT-Scan Analysis (early stage lung cancer detection)	Healthcare	Cloud services	In operati on
6.7.20	AI-based design of pharmacologically relevant targets with target properties	Healthcare	On-premise systems	Prototy pe
6.7.21	AI-based mapping of optical to multi-electrode catheter recordings for Atrial Fibrillation Treatment	Healthcare	Embedded systems	PoC
6.7.22	Generation of Computer Tomography scans from Magnetic Resonance Images	Healthcare	Embedded systems	PoC
6.7.23	Generation of Computer Tomography Scans from Magnetic Resonance Images	Healthcare	Embedded systems	PoC
6.7.24	Improving the knowledge base of prescriptions for drug and non-drug therapy and its use as a tool in support of medical professionals	Healthcare	Cloud services	Prototy pe
6.7.25	Neural Network Formation of 3D-models orthopedic insoles	Healthcare	Client and server systems	In operati on
6.7.26	Search of undiagnosed patients	Healthcare	Social networks	In operati on
6.7.27	Support system for optimization and personification of drug therapy	Healthcare	On premise system	PoC

6.7.28	Syntelly - computer aided organic synthesis	Drug design, digital pharma	System for the prediction of the properties of pharmaceutically relevant molecules	Prototype
6.7.29	WebioMed Clinical Decision Support System	Healthcare	Cloud services	In operation
6.8.1	Robot consciousness	Home/Service Robotics	Embedded systems	PoC
6.8.2	Social humanoid technology capable of multi-modal context recognition and expression	Home/Service Robotics	Embedded systems	Prototype
6.8.3	Application of Strong Artificial Intelligence	Hi-Tech Labor Market	Artificial Agents	In operation
6.9.1	Autonomous network and automation level definition	ICT	Cyber-physical systems	PoC
6.9.2	Autonomous network scenarios	ICT	Cyber-physical systems	PoC
6.9.3	A judging support system for gymnastics using 3D sensing	ICT	On-premise systems	PoC
6.9.4	Active Antenna Array Satellite	ICT	Cyber-physical systems	Prototype
6.9.5	Carrier interference detection and removal for satellite communication	ICT	On-premise systems	PoC
6.9.6	Ontologies for Smart Buildings	ICT	Hybrid	Prototype
6.9.7	Product failure prediction for critical IT infrastructure	ICT	On-premise systems	In operation
6.9.8	Data compression with AI techniques	ICT	On-premise systems	Prototype
6.9.9	Optimization of software configurations with AI techniques	ICT	On-premise systems	Prototype
6.9.10	Better human-computer interaction with advanced language models	ICT	Hybrid	Prototype
6.10.1	Tax Rules Updates and Classification	Legal	On-premise systems	PoC
6.10.2	AI Contract Management	Legal	On-premise systems	In operation
6.10.3	Semantic Analysis of Legal Documents	Legal	On-premise systems	Prototype
6.11.1	Improving Productivity for Warehouse Operation	Logistics	On-premise systems	PoC
6.11.2	AI based dynamic routing SaaS	Logistics	Cloud services	In operation
6.12.1	Anomaly Detection in Sensor Data Using Deep Learning techniques	Maintenance & support	Hybrid	PoC

6.12.2	Jet Engine Predictive Maintenance Service	Maintenance & support	Cloud services	Prototype
6.12.3	Detection of fraudulent medical claims	Maintenance & support	On-premise systems	In operation
6.12.4	AI Virtual Assistant for Customer Support and Service	Maintenance & support	Embedded systems	In operation
6.13.1	AI solution to calculate amount of contained material from mass spectrometry measurement data	Manufacturing	Embedded systems	PoC
6.13.2	AI solution to quickly identify defects during quality assurance process on wind turbine blades	Manufacturing	On-premise systems	In operation
6.13.3	Solution to detect signs of failures in wind power generation system	Manufacturing	On-premise systems	PoB
6.13.4	Generative design of mechanical parts	Manufacturing	On-premise systems	In operation
6.13.5	Information Extraction from Hand-marked Industrial Inspection Sheets	Manufacturing	Cloud services	PoC
6.13.6	Automated defect classification on product surfaces	Manufacturing	On-premise systems	PoC
6.13.7	Robotic task automation: Insertion	Manufacturing	Hybrid	PoC
6.13.8	Powering Remote Drilling Command Centre	Manufacturing	Cloud services	In operation
6.13.9	Leveraging AI to enhance adhesive quality	Manufacturing	On-premise systems	In operation
6.13.10	Machine learning driven approach to identify the weak spots in the manufacturing of the circuit breakers.	Manufacturing	On-premise systems	Prototype
6.13.11	Machine Learning Driven Analysis of Batch Process Operation Data to Identify Causes for Poor Batch Performance	Manufacturing	On-premise systems	Prototype
6.13.12	Empowering Autonomous Flow meter control- Reducing time taken to “proving of meters”	Manufacturing	Cloud services	In operation
6.13.13	Adaptable Factory	Manufacturing	Hybrid	PoC
6.13.14	Order-Controlled Production	Manufacturing	Cloud Services	Prototype
6.13.15	Value-based Service	Manufacturing	Hybrid	PoC
6.13.16	Improvement of productivity of semiconductor manufacturing	Manufacturing	On-premise systems	In operation
6.13.17	AI Decryption of Magnetograms	Manufacturing, Gas & Oil	Client and server systems	In operation
6.13.18	Analyzing and Predicting Acid Treatment Effectiveness of Bottom Hole Zone	Manufacturing, gas & oil	Client and server systems	In operation

6.13.19	Automatic Classification Tool for Full Size Core	Manufacturing, Gas & Oil	Client and server systems	In operation
6.13.20	Intelligent Technology to Control Manual Operations on Video — “Norma”	Manufacturing	On-premise systems	Prototype
6.13.21	Optimization of ferroalloy consumption for a steel production company	Manufacturing	Embedded systems	PoC
6.13.22	Device Control Using both cloud AI and embedded AI	Manufacturing	Learning in both Cloud and Embedded	In operation
6.14.1	Predictive analytics for the behavior and psycho-emotional conditions of eSports players using heterogeneous data and artificial intelligence	eSports	Cyber-physical systems	Prototype
6.15.1	autonomous apron truck	Mobility	Embedded systems	PoC
6.15.2	AI solution to help mobile phone to have better picture effect	Mobility	Hybrid	In operation
6.16.1	AI Ideally Matches Children to Daycare Centers	Public sector	On-premise systems	In operation
6.16.2	AI Sign Language Interpretation System for the Hearing-Impaired	Public sector	Embedded systems	Prototype
6.16.3	AI Situation Explanation Service for the Visually Impaired	Public sector	Hybrid	Prototype
6.16.4	Predictive maintenance of public housing lifts	Public sector	Embedded systems	PoC
6.17.1	Emotion-sensitive AI Customer Service	Retail	On-premise systems	In operation
6.17.2	Deep Learning Based User Intent Recognition	Retail	On-premise systems	In operation
6.18.1	Behavioural and sentiment analytics	Security	On-premise systems	PoC
6.18.2	AI (Swarm Intelligence) solution for Attack Detection in IoT Environment	Security	Hybrid	Prototype
6.18.3	Use of robotic solution for traffic policing and control	Security	On-premise systems	PoC
6.18.4	Robotic solution for replacing human labour in Hazardous condition	Security	On-premise systems	PoC
6.18.5	Non-intrusive detection of malware	Security	Cloud services	In operation
6.19.1	Deep Learning Technology Combined with Topological Data Analysis Successfully Estimates Degree of Internal Damage to Bridge Infrastructure	Social infrastructure	Cloud services	PoC
6.19.2	Water Crystal Mapping	Social infrastructure	Cloud services	Prototype

6.19.3	System for Real-Time Earthquake Simulation with Data Assimilation	Social infrastructure	On-premise systems	Prototy pe
6.20.1	AI Components for Vehicle Platooning on Public Roads	Transportation	Self-driving vehicles	Prototy pe
6.20.2	Self-Driving Aircraft Towing Vehicle	Transportation	Self-driving vehicles	Prototy pe
6.20.3	Unmanned Protective Vehicle for Road Works on Motorways	Transportation	Self-driving vehicles	Prototy pe
6.20.4	Enhancing traffic management efficiency and infraction detection accuracy with AI technologies	Transportation	Hybrid	In operati on
6.20.5	AI solution for traffic signal Optimization based on multi-source data fusion	Transportation	Cloud services	In operati on
6.20.6	Automated Travel Pattern Recognition using Mobile Network Data for Applications to Mobility as a Service	Transportation	Activity- based Modelling for New mobility Services	PoC
6.20.7	Autonomous Trains (Unattended Train Operation (UTO))	Transportation	Self-driving vehicles	prototy pe
6.21.1	Robotic prehension of objects	Work & life	Embedded systems	PoC
6.21.2	Robotic vision – scene awareness	Work & life	Embedded systems	PoC
6.21.3	Recommendation algorithm for improving member experience and discoverability of resorts in the booking portal of a hotel chain	Work & life	Cloud services	In operati on
6.21.4	CRWB Recommendation benchmark	Work & life	Cloud services	Prototy pe
6.21.5	Improving the quality of online interaction	Work & life	Cloud services	In operati on
6.22.1	AI solution to identify automatically false positives from a specific check for “untranslated target segments” from an automated quality assurance tool	This will be relevant for content from across any domains	Cloud services	PoC
6.22.2	AI solution for Car Damage Classification	Insurance	Cloud services	PoC
6.22.3	Causality-based Thermal Prediction for Data Center	data center	On-premise systems	Prototy pe
6.22.4	Machine Learning Tools in Support of Transformer Diagnostics	Performance evaluation and diagnostics	Prototype	Prototy pe
6.22.5	Video on Demand Publishing Intelligence Platform	TMT Industry, Technology Department	On-premise systems	In operati on
6.22.6	Predictive Testing	TMT Industry – Application development	On-premise systems	PoC
6.22.7	Predictive Data Quality	Data Management	Hybrid	PoC

6.22.8	Expansion of AI training dataset and contents using artificial intelligence techniques	IT, AI, Future services	Server system	Prototy pe
6.22.9	Open spatial dataset for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data	earth science, digital cartography	On-premise systems	In operati on

6.2 Agriculture

6.2.1 AI to Understand Adulteration in Commonly Used Food Items (A.19)

6.2.1.1 Scope

Understand the patterns in hyperspectral / NIR or visual imaging specifically for adulteration in milk, banana and mangoes.

6.2.1.2 Objective

To device a simple, cost effective tool to identify the adulteration in food items at point of purchase.

6.2.1.3 Narrative (Short description)

Food adulteration is one of the big evil of modern society. Adulterated milk is hazard for children, many aliments including cancer / kidney failures due to consumption of adulterated food. Hyperspectral technology was evaluated to find out adulteration in food items.

6.2.1.4 Challenges and issues

Large scale data collection, Miniaturization of frugal NIR / Hyperspectral sensor.

6.2.1.5 Societal concerns

If the AI system is rolled out and taken as reliable then it should be able to perform in all cases and scenarios. Incorrect classification can lead to false accusations.

SDGs to be achieved: Good health and well-being for people

6.2.2 bioBotGuard (A.54)

6.2.2.1 Scope

Use visual recognition to identify and help fight parasites attacking organic farms.

6.2.2.2 Objective

The use case shows how AI contributing to modernize Agriculture industry.

6.2.2.3 Narrative (Short description)

BioBotGuard defines itself as an initiative of Precision Farming as a Service. From an IT perspective it uses drones with GPS and high-resolution cameras to monitor the crops; the images are then processed by computer vision API in order to spot diseases and harmful insect attacks, building a georeferenced risk map of the crop. This can be used to send operational drones to put the treatment (or antagonist insects) only when and where it is needed.

6.2.2.4 Challenges and issues

Acquire filed as well as crop images at different distances and normalize image recognition and pattern detection.

6.2.2.5 Societal concerns

None identified.

6.2.3 Ecosystems management from causal relation inference from observational data (A.96)

6.2.3.1 Scope

Infer important latent variables to control whole ecosystem from database including human observation and sensor data.

6.2.3.2 Objective

To provide some suggestions for managing ecosystems and repeatedly improve it with the introduction of possibly latent variables and new data.

6.2.3.3 Narrative (Short description)

We can find diverse relations between climate, animals and plants that infer ecologically consistent structure.

To determine the factors that support a species niche is necessary to diversify the polyculture in ecological optimum, which is a complex entanglement that depends on environmental condition, associated biodiversity, farming option, etc.

In our Synecoculture project, polyculture with ecological optimum requires a huge amount of information on biodiversity, interactions, and vegetation succession parameters, generally sparse possibly biased, open-ended, etc., because it relies on human observation. Still, it can bring useful information and intriguing insight on the management if powerful algorithmic analysis is combined with appropriate human evaluation.

6.2.3.4 Challenges and issues

None identified.

6.2.3.5 Societal concerns

SDGs to be achieved: No poverty; Zero Hanger; Good health and well-being; Clean water and Sanitation; Decent work and economic growth; Industry, innovation and infrastructure; Reduce inequalities; Responsible consumption and production; Climate action; Life on land; Partnerships for the goals

6.2.4 Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI (A.126)

6.2.4.1 Scope

The project is devoted to the development of a low-power embedded system and AI algorithm for real-time plant segmentation and prediction of its growth. The proposed distributed system is aimed for use in greenhouses and remote areas, where edge-computing autonomous systems are in demand. A branch of this project also aims to develop the payload for drones for the segmentation of harmful plants in real-time.

6.2.4.2 Objective

Prediction of harvest, biomass/leaf area dynamics, leaf index, parameters describing the quality of produced food, consumption of resources from sequences of images of plant growth (including multispectral), data from sensors that describe environmental conditions and artificial growing system parameters representing the state of the growing system.

6.2.4.3 Narrative (Short description)

Research efforts towards low-power sensing devices with fully-functional AI on board are still fragmented. In our project, we present an embedded system enriched with AI that ensures the continuous analysis and in-situ prediction of the plant leaf growth dynamics and other important growth parameters. The embedded solutions grounded on a low-power embedded sensing system with a Graphics Processing Unit (GPU) are able to run the neural networks-based AI on board. Advantages of the proposed system include portability and ease of deployment. The proposed approach guarantees the system autonomous operation for 180 days using a standard Li-ion battery. We rely on state-of-the-art mobile graphic chips for smart analysis and control of autonomous devices. The data was used for training and testing the Recurrent Neural Network, Convolutional Neural Network algorithms, and the segmentation algorithms. All this allows for high performance in-situ optimization of plant growth dynamics and resource consumption.

6.2.4.4 Challenges and issues

— 1) The plant growth data significantly depends on multiple factors, including used solutions, illumination characteristics (for greenhouses), weather and seasonal conditions (for outdoors).

— 2) The architecture of the neural network should have both high accuracy, high framerate, but low amount of layers and trained parameters for further inference on low-power embedded systems. These controversial factors should be met since embedded systems have limited processing capabilities.

— 3) high diversity of data types and no standardization of data obtained by farmers.

6.2.4.5 Societal concerns

Good health and well-being for people; elimination of hunger; availability of cheap and healthy food for everyone; colonization of harsh environments on Earth and in space exploration.

SDGs to be achieved: Good health and well-being; Zero Hanger

6.3 Digital marketing

6.3.1 Improving conversion rates and RoI (Return on Investment) with AI technologies (A.53)

6.3.1.1 Scope

Utilizing AI technologies in digital marketing.

6.3.1.2 Objective

1) Help the operation team identify new business scenarios and seize more market opportunities,

2) Increase conversion rate and marketing effectiveness,

3) Improve user experience by providing individually customized services

6.3.1.3 Narrative (Short description)

Personalized digital marketing has become increasingly important in response to the needs of providing different services to different consumers. The combination of big data and AI algorithms is the core of personalized digital marketing. By modelling user preferences, we can predict the services that users may be interested in, improve marketing effectiveness and enhance user experience.

6.3.1.4 Challenges and issues

How to collect, utilize and protect user information within the scope of what is permitted by relevant national and regional legislation and regulations.

How to let the system evolve and improve continuously with applying new AI models and algorithms.

6.3.1.5 Societal concerns

For Users: enjoy better service at a lower cost

For Merchants: Increase profits and decrease costs

For Cities and communities: Promote economic prosperity and develop green economy

SDGs to be achieved: Sustainable cities and communities

6.3.2 Logo and Trademark Detection (A.56)

6.3.2.1 Scope

Identification of logos / trademarks in pictures, optionally performing sentiment analysis associated to the product

6.3.2.2 Objective

Understand usage of retail or fashion products and optionally sentiment associated to it, according to pictures posted on the internet or social networks by customers

6.3.2.3 Narrative (Short description)

The case is about being able to identify logos and trademarks in pictures provided to the AI systems, and optionally derive a positive or negative sentiment for the product based on the written context that was provided with the picture.

6.3.2.4 Challenges and issues

The primary challenge is to be able to correctly identify trademarks in all situations (with bad lighting, image distortions, dirt, etc.) and interpret the sentiment and tone in different countries and languages, as people might use slang and irony.

6.3.2.5 Societal concerns

Automated analysis of public posts on social networks might be seen unethical in certain cultures.

6.3.3 Flavorlens (A.76)

6.3.3.1 Scope

Multi-sensing Dish tasting experience sharing in a social media ecosystem

6.3.3.2 Objective

Users share their experiences and dish recommendation

6.3.3.3 Narrative (Short description)

Social network to enable dish tasting experiences

6.3.3.4 Challenges and issues

Personal expectation related to flavor, taste and texture

6.3.3.5 Societal concerns

Local healthy dish for user satisfaction and preference

SDGs to be achieved: Good health and well-being for people

6.4 Education

6.4.1 VTrain Recommendation Engine (A.23)

6.4.1.1 Scope

Based on an employee's career objectives find skill requirements and its training.

6.4.1.2 Objective

Recommend a personalised list of "best" training courses to an employee, which will help him/her meet his/her career objectives.

6.4.1.3 Narrative (Short description)

The vTrain system helps employees improve their skills by recommending appropriate training courses from a given list and historical data.

6.4.1.4 Challenges and issues

Need large amounts of training data; predicting human behaviour is tricky.

6.4.1.5 Societal concerns

Employees may feel challenged or demoralized.

SDGs to be achieved: Decent work and economic growth

6.4.2 RAVE (A.55)

6.4.2.1 Scope

Use of advanced an multimodal sensing ability to facilitate a complex task

6.4.2.2 Objective

Avatar and social robot interact with deaf babies for facilitating language learning.

6.4.2.3 Narrative (Short description)

RAVE system is an integrated multi-agent system involving a robot and virtual human designed to augment language exposure for 6-12 month old infants. The system is an engineered robot and avatar to provide visual language to effect socially contingent human conversational exchange. The team demonstrated the successful engagement of our technology through case studies of deaf and hearing infants.

6.4.2.4 Challenges and issues

Ability to decode a learner cognitive status and his attention level.

6.4.2.5 Societal concerns

None identified.

6.4.3 IFLYTEK Intelligent marking system (A.83)

6.4.3.1 Scope

It can realize intelligent detection and grading of all subjective questions.

6.4.3.2 Objective

To reduce a lot of labor and organizational costs.

6.4.3.3 Narrative (Short description)

IFLYTEK intelligent marking system is based on the core technology design research, including IFLYTEK independent intellectual property rights handwritten recognition, natural language understanding, intelligent evaluation and other artificial intelligence and so on. It can realize the detection of blank questions for all types of questions except multiple choice questions, and the computer intelligent evaluation of Chinese, English composition, English translation, Literature synthesis category short answer questions and English blank questions. At the same time, for Chinese composition and English composition, it can also effectively detect the abnormal answer papers which are highly similar to the dry content of the test paper or the content of the external model text.

6.4.3.4 Challenges and issues

The accuracy of marking paper needs to be further improved.

6.4.3.5 Societal concerns

There is a scientific and unified scoring standard, which can ensure the fairness of the marking results.

Reduced a lot of labor and organizational costs

SDGs to be achieved: Quality education

6.4.4 Intelligent educational robot (A.84)

6.4.4.1 Scope

It's the best partner of a child, and make the child learn in play.

6.4.4.2 Objective

To improve the pleasure of learning.

6.4.4.3 Narrative (Short description)

Educational robot is a new teaching tool to cultivate students' comprehensive ability. It mainly uses artificial intelligence technology, speech recognition technology and bionic technology to cultivate students' various abilities. Educational robots have hearing, vision, oral skills, recognition, emotional detection and the ability to interact for a long time.

6.4.4.4 Challenges and issues

Be able to sense students' emotions like teachers.

Accurately capture students' gestures, postures, face information, etc.

6.4.4.5 Societal concerns

To give students emotional support.

Stimulate students' interest in learning.

SDGs to be achieved: Quality education

6.4.5 AI solution to intelligence campus (A.85)

6.4.5.1 Scope

It is a full range of products and integrated solutions for teaching, examination, evaluation, management, learning.

6.4.5.2 Objective

This scheme provides a comprehensive intelligent sensing environment and comprehensive information service platform for teachers and students, so as to realize the integration of human and business information.

6.4.5.3 Narrative (Short description)

Based on big data and artificial intelligence technology, the scheme brings teaching, examination, learning and management into the integrated system of mutual cooperation, based on accompanying data acquisition and dynamic big data analysis, combined with process evaluation, to help teachers and students to realize teaching according to their aptitude and individualized learning, to help managers to supervise and assist decision-making, and to greatly promote the transformation of education, learning and management to intelligence.

6.4.5.4 Challenges and issues

The implementation of intelligent campus makes the data of students and teachers be collected and processed in large quantities, which is likely to lead to the disclosure of private data. Therefore, the establishment of data privacy protection mechanism should be strengthened in intelligent platform.

6.4.5.5 Societal concerns

Intelligent campus solution leads artificial intelligence technology into the campus, into the classroom, promotes students' learning and teachers' teaching, and facilitates teaching management.

SDGs to be achieved: Quality education

6.4.6 AI Adaptive Learning Platform for Personalized Learning (A.102)

6.4.6.1 Scope

2,5 million users.

6.4.6.2 Objective

Open access, Interactive tasks, Personalization, User-generated content, Learning graph. Summarizing - equal access to high-quality education.

6.4.6.3 Narrative (Short description)

Adaptive learning platform (AiEd platform) is an elearning platform and course-builder which uses AI for forming adaptive learning paths.

6.4.6.4 Challenges and issues

Edstories (micro-learning video stories) should be included to satisfy the pedagogical model of movement-based learning.

6.4.6.5 Societal concerns

The system should be integrated into secondary and tertiary school-systems that still face legal boundaries and limitations for scaling

SDGs to be achieved: Quality education

6.4.7 AI Adaptive Learning Mobile App (A.124)

6.4.7.1 Scope

None identified.

6.4.7.2 Objective

Providing easy, convenient and adaptive learning of English with the help of a virtual teacher based on artificial intelligence.

6.4.7.3 Narrative (Short description)

A mobile application for learning English, which is based on a program that adapts content to the student and learns with them. During registration, the program analyzes the user's account on a social network and draws up an individual training plan based on the student’s interests.

6.4.7.4 Challenges and issues

The development of a personalized approach to learning.

6.4.7.5 Societal concerns

This case of the use of artificial intelligence in the educational process can complement teachers as knowledge transmitters and make education accessible to everyone. At the same time, artificial intelligence, performing the functions of analytics, packaging and personalization of educational content, is much more effective than a person in the role of an assistant to a teacher and shifts the role of a classical teacher towards mentoring.

SDGs to be achieved: Quality education

6.5 Energy

6.5.1 AI-dispatcher (operator) of large-scale distributed energy system infrastructure (A.109)

6.5.1.1 Scope

Monitoring, optimization and control of large scale distributed energy systems using Deep Reinforcement Learning (gas, oil, power, heat, water transmission and distribution infrastructure systems).

6.5.1.2 Objective

To develop an effective industrial AI solution which is able to recommend the optimal control of energy infrastructure systems in real-time in order to:

— satisfy the energy demand of consumers.

— minimize possible negative impacts on the environment.

— reduce operational costs through systems’ real-time continuous optimization in self-adaptive manner.

6.5.1.3 Narrative (Short description)

An AI solution is currently in development that uses hybrid models (based on both traditional physics models and artificial neural networks), “digital twins,” and deep reinforcement learning to support decision making and control of energy infrastructure systems in real-time.

6.5.1.4 Challenges and issues

— To achieve a high level of efficiency of complex energy system’s optimization and dispatching control.

— To learn from human-beings, including machine teaching techniques.

— To employ meta-learning techniques in real industrial environments, which can help AI-agents to adopt efficiently to different systems (for example, from small scale to large scale industrial systems, from gas to oil transmission system, from power to heat infrastructure systems, and vice versa).

— To deal effectively with partially observed systems.

— To develop an AI-solution which reacts reliably to rare events.

6.5.1.5 Societal concerns

Safety, security and reliability of AI solutions that are used in energy infrastructure management.

SDGs to be achieved: Affordable and clean energy

6.6 Fintech

6.6.1 Detection of Frauds based on Collusions (A.20)

6.6.1.1 Scope

Validating the predicted collusion set is effort-intensive and needs investigative and legal expertise.

6.6.1.2 Objective

Automatic unsupervised detection of frauds based on collusions.

6.6.1.3 Narrative (Short description)

A set of unsupervised machine learning algorithms to detect collusion-based frauds, particularly, circular trading and price manipulation in stock market trading.

6.6.1.4 Challenges and issues

Actual examples of collusion-based frauds may not be available easily, even for evaluation and testing.

6.6.1.5 Societal concerns

Incorrect detection of Collusions and frauds may cause unnecessary stress in stock traders.

SDGs to be achieved: Decent work and economic growth

6.6.2 Credit Scoring using KYC Data (A.27)

6.6.2.1 Scope

Building a risk scorecard for loan applicants using KYC data for better risk management and high population coverage.

6.6.2.2 Objective

Assigning a risk score to every loan applicant in real time, using just KYC data, which will ensure both new-to-credit and mature customers can be assessed for their creditworthiness, and offered loans on appropriate terms.

6.6.2.3 Narrative (Short description)

It can be often difficult to build a risk scorecard using only KYC data, which often has noisiness and incompleteness issues. However if realized, it can be used to provide an objective score to all loan applicants, even the new-to-credit ones. Non-linear classification algorithms are suitable for this purpose.

Several variables are collected from the customer during the KYC process such as Age of customer, Self-reported income, Type of Occupation, Purpose of loan, etc. All these features can be added to a non-linear risk model and their complex interactions allowed to take place.

6.6.2.4 Challenges and issues

— KYC data obtained from extreme rural areas can be noisy, may have several missing values, and needs appropriate preprocessing and treatment before feeding to the model algorithm.

— Non-linear models like Random Forest and XGBoost need significant computational power during the training phase.

6.6.2.5 Societal concerns

We don’t see any societal concerns if it is used.

6.6.3 Virtual Bank Assistant (A.57)

6.6.3.1 Scope

Use of advanced chatbots and dialogue systems to automatize part of the call center activities.

6.6.3.2 Objective

Provide better quality help desk support to employees.

6.6.3.3 Narrative (Short description)

The Virtual Assistant of the Bank is the first point of contact for branch operators, who receive immediate answers at any time - it allows to optimize the time of the "human operators" of the Service Desk, which they are dedicated to activities of greater value.

6.6.3.4 Challenges and issues

Provide a natural and consistent interaction with users from different levels of experience (and thus terminology) and background.

6.6.3.5 Societal concerns

None identified.

6.6.4 Forecasting prices of commodities (A.91)

6.6.4.1 Scope

Build a neural network to forecast the price of base metal commodities.

6.6.4.2 Objective

Use forecasted prices to interpret trading trends.

6.6.4.3 Narrative (Short description)

A trading company needed to improve the forecast accuracy of price points for specific commodities.

6.6.4.4 Challenges and issues

Challenge in modelling a neural network model that ingest large and wide array of data, while calibrating for variables that have short term versus long term impact.

6.6.4.5 Societal concerns

Unpredictable flow of materials and commodities due to price shocks.

SDGs to be achieved: Reducing inequalities

6.6.5 Finance Advising and Asset Management with AI (A.114)

6.6.5.1 Scope

Financial advising and portfolio management for financial institutions and consumers.

6.6.5.2 Objective

Designed to manage exchange-traded securities portfolios of conservative investors in real time, using asset price data and macroeconomic data, to make the most accurate decisions at a given yield and moderate risk. Prediction of significant depreciation of exchange-traded asset prices as a result of a sharp monetary contraction called financial crises.

6.6.5.3 Narrative (Short description)

The core of the system carries out a structured collection from open sources and multi-threaded parallel analysis of information; it regulates the application of basic algorithms and rules for changing these algorithms that change the purpose of the task. (Intermediate goal setting is one of the elements of "Strong AI”). One of the tasks is to assess market trends, as well as market and interest rate risk. Changes in the algorithm of actions depend on the macroeconomic information received from the outside. It translates notoriously weakly formalized parameters into specific decisions on the formation of investment portfolios and issues orders to brokers to purchase, rebalance, or sell assets in stock exchanges.

The macroeconomics unit is an autonomous system that generates indicators of time periods and geographical areas with different weights of investment potential.

6.6.5.4 Challenges and issues

— Data can be noisy, may have several missing values, and needs appropriate pre-processing and treatment before feeding to the model algorithm.

— Working with financial assets requires high reliability of computing systems and replication systems.

6.6.5.5 Societal concerns

SDGs to be achieved: No poverty

6.6.6 Loan in 7 minutes (A.119)

6.6.6.1 Scope

A completely automated solution which analyzes customer behavior and makes loan offers best for the customer.

6.6.6.2 Objective

Create lending product for clients of medium and large businesses (LMB) with the shortest delivery time possible taking into account the extremely detailed customer profile.

6.6.6.3 Narrative (Short description)

Loan in 7 minutes is the first solution in the world where the credit decision is made by artificial intelligence without human participation in just a few minutes.

A complex machine learning settlement system was implemented on one of the largest Hadoop-cluster in Eastern Europe (tens of petabytes of data) and integrated into the business process of corporate lending of the Bank.

The new project has significantly improved customer experience:

– Eliminated the need for the client to contact the Bank in person for a loan.

– Requires no additional documents from the client to get a decision.

– Bank’s automated systems were improved in terms of automatic transaction creation.

– Substantially simplified the process of issuing a loan.

6.6.6.4 Challenges and issues

Non-linear models based on big data need significant computational power during the training phase.

6.6.6.5 Societal concerns

Investment in technological innovation and infrastructure are crucial drivers of higher levels of productivity and economic growth.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.7 Healthcare

6.7.1 Explainable Artificial Intelligence for Genomic Medicine (A.1)

6.7.1.1 Scope

To explain reason and basis behind AI-generated findings in genomic medicine.

6.7.1.2 Objective

To improve the efficiency of investigatory work for experts in genomic medicine.

6.7.1.3 Narrative (Short description)

This technology was deployed to improve the efficiency of investigatory work for experts in genomic medicine, utilizing training data and a knowledge graph that made use of public databases and medical literature databases in the field of bioinformatics. It was then evaluated to validate that it was possible to find and link the basis supporting findings with regard to phenomena whose interrelationships are only partially understood.

6.7.1.4 Challenges and issues

Challenges: To reduce experts' workloads, shortening determination periods in genomic medicine.

6.7.1.5 Societal concerns

— Accountability for using AI in medical examination.

— Incorrect explanation will cause the determination periods increasing.

SDGs to be achieved: Good health and well-being for people

6.7.2 Revolutionizing Clinical Decision-making using Artificial Intelligence (A.2)

6.7.2.1 Scope

To improve clinical decision-making and the accurate assessment of risks for individual patients of mental healthcare.

6.7.2.2 Objective

Halving the time to pre-screen patient records and giving more time for patient consultations.

6.7.2.3 Narrative (Short description)

The solution has halved the time for the preliminary assessment of patient records, increasing the time available for consultations.

6.7.2.4 Challenges and issues

The incorporation of many different types of data is revolutionizing the healthcare sector. The ability to apply semantic and analytic technologies to this heterogeneous mass of data, as well as traditional healthcare data, to discover hidden correlations, identify care patterns and support clinical decision-making is paving the way for a new generation of improved healthcare services.

6.7.2.5 Societal concerns

Incorrect decision and unexplainable result.

SDGs to be achieved: Good health and well-being for people

6.7.3 Computer-aided Diagnosis in Medical Imaging based on Machine Learning (A.6)

6.7.3.1 Scope

Detecting image anomality.

6.7.3.2 Objective

Provide AI method to alleviate growing burden of histopathological diagnosis by human.

6.7.3.3 Narrative (Short description)

The advances in image recognition technology enable the machine learning system to support diagnosis in medical imaging. This technology is expected to contribute the great reduction of the burden on doctors and the improvement of diagnostic accuracy when it is used for screening and double checking. Specifically, a support system is currently under development that analyzes histopathological images to automatically detect suspected lesion.

6.7.3.4 Challenges and issues

None identified.

6.7.3.5 Societal concerns

None identified.

6.7.4 AI Solution to Predict Post-Operative Visual Acuity for LASIK Surgeries (A.24)

6.7.4.1 Scope

Predicting Post-Operative Visual Acuity for LASIK Surgeries from retrospective LASIK surgery data with patient follow-ups.

6.7.4.2 Objective

Given: Pre-operative examination results and demography information about a patient. Predict: Post-operative UCVA after one day, one week and one month of the surgery.

6.7.4.3 Narrative (Short description)

LASIK (Laser-Assisted in SItu Keratomileusis) surgeries have been quite popular for treatment of myopia, hyperopia and astigmatism over the past two decades. In the past decade, over 10 million LASIK procedures had been performed in the United States alone with an average cost of approximately $2000 USD per surgery. While 99% of such surgeries are successful, the commonest side effect is a residual refractive error and poor uncorrected visual acuity (UCVA). In this work, we aim at predicting the UCVA post LASIK surgery. We model the task as a regression problem and use the patient demography and pre-operative examination details as features. To the best of our knowledge, this is the first work to systematically explore this critical problem using machine learning methods. Further, LASIK surgery settings are often determined by practitioners using manually designed rules. We explore the possibility of determining such settings automatically to optimize for the best post-operative UCVA by including such settings as features in our regression model. Our experiments on a dataset of 791 surgeries provides an RMSE (root mean square error) of 0.102, 0.094 and 0.074 for the predicted post-operative UCVA after one day, one week and one month of the surgery respectively.

6.7.4.4 Challenges and issues

The problem is challenging because: (1) large amount of data about such surgeries is not easily available; (2) there are a lot of pre-operative measurements that can be used as signals; and (3) data is sparse, i.e., there are a lot of missing values.

6.7.4.5 Societal concerns

SDGs to be achieved: Good health and well-being for people

6.7.5 Chromosome Segmentation and Deep Classification (A.44)

6.7.5.1 Scope

Karyotyping of the chromosomes is restricted to healthy patients.

6.7.5.2 Objective

— Automating Karyotyping of the chromosomes in cell spread images.

— Segmentation of chromosomes in the images using non expert crowd.

6.7.5.3 Narrative (Short description)

Karyotyping of the chromosomes micro-photographed under metaphase is done by characterizing the individual chromosomes in cell spread images. Currently, considerable effort and time is spent to manually segment out chromosomes from cell images, and classifying the segmented chromosomes. We proposed a method to segment out and classify chromosomes for healthy patients using a combination of crowdsourcing, preprocessing and deep learning, wherein the non-expert crowd from external crowdsourcing platform is utilized to segment out the chromosomes, which are then classified using deep neural network. Results are encouraging and promise to significantly reduce the cognitive burden of segmenting and karyotyping chromosomes.

6.7.5.4 Challenges and issues

— Crowd's job satisfaction.

— Spamming in annotated data.

6.7.5.5 Societal concerns

Inaccurate classification of chromosomes can lead to stress in patients in case the classification is not reviewed by expert doctors.

SDGs: Good health and well-being for people

6.7.6 AI Solution to Quality Control of Electronic Medical Record(EMR) in Real Time (A.50)

6.7.6.1 Scope

Detecting defects in EMR by inspecting unstructured data based on Natural Language Processing (NLP) ability.

6.7.6.2 Objective

To insure the completeness, consistency, punctuality and medical-compliance of EMR written by physicians.

6.7.6.3 Narrative (Short description)

This AI solution in ET Medical Brain Medical service support system was developed that could simultaneously detect mistakes while physicians wrote EMR(Electronic Medical Record).

Using NLP(Natural Language Processing) ability, it can process a large amount of unstructured text and judge the accuracy according to recognized medical reference.

It achieved 80% coverage of all the EMR quality control requirements issued by Chinese government, and human labour of EMR QC(Quality Control) was reduced 60%, which translated into cost savings, and enhanced physician education.

6.7.6.4 Challenges and issues

Challenges: Achieve all EMR QC requirements in different disease areas.

Issues: 1) Lack of medical reference data 2) Lack of medical knowledge graph

6.7.6.5 Societal concerns

Achieved 80% coverage of all the EMR quality control requirements issued by Chinese government, and human labour of EMR QC(Quality Control) was reduced 60%, which translated into cost savings, and enhanced physician education.

SDGs to be achieved: Good health and well-being for people

6.7.7 Dialogue-based social care services for people with mental illness, dementia and the elderly living alone (A.63)

6.7.7.1 Scope

Dialogue-based social care services for people with mental illness, dementia and the elderly living alone

6.7.7.2 Objective

Dialogue-based interaction between people and machines utilizing artificial intelligence technology helps people with accessibility issues to IT devices

6.7.7.3 Narrative (Short description)

Daily life support services based on artificial intelligence conversation technology that can perform information processing tasks through natural language conversation with users

6.7.7.4 Challenges and issues

Multimodal data handling based multimodal interaction

Multimodal data analysis

Multimodal data-based inferences

6.7.7.5 Societal concerns

Promoting welfare and supporting social activities for the inconvenient

SDGs to be achieved: Good health and well-being for people

6.7.8 Pre-screening of cavity and oral diseases based on 2D digital images (A.67)

6.7.8.1 Scope

Artificial intelligence-based oral examination platform

6.7.8.2 Objective

AI based oral disease self-examination solution

Cavity, periodontal disease, oral disease, tooth care and oral care self-care prevention management

6.7.8.3 Narrative (Short description)

This service utilizes artificial intelligence technology to analyze the oral condition by sending oral images to the diagnostic server without visiting the dentist.

6.7.8.4 Challenges and issues

Dental image processing using artificial intelligence

6.7.8.5 Societal concerns

Elimination of inequalities in regional health care services

SDGs to be achieved: Good health and well-being for people

6.7.9 Real-time patient support and medical information service applying spoken dialogue system (A.68)

6.7.9.1 Scope

Medical business support system using artificial intelligence based human computer interface technology

6.7.9.2 Objective

Acquisition, retrieval and provision of patients and related data needed by medical staffs in real time through a voice dialogue interface during medical treatment

6.7.9.3 Narrative (Short description)

The service is a medical system that provides patient information and related data for treatment in real time based on a voice dialogue interface to help medical hands-on medical activities, such as dental, first aid, and surgery.

6.7.9.4 Challenges and issues

Dialogue service in medical data and knowledge

Question and answering in a medical expert system

Multi-task handling in a dialogue-based interfacing environment

Remote speech recognition

6.7.9.5 Societal concerns

Improving medical service efficiency and patient satisfaction.

SDGs to be achieved: Good health and well-being for people

6.7.10 Integrated recommendation solution for prosthodontic treatments (A.69)

6.7.10.1 Scope

In order to support complicated prosthetic treatments according to the patient's condition, the artificial intelligence technology provides a comprehensive analysis of the given information and situations to recommend various prosthetic treatment methods and visualize them to support doctors and patients.

6.7.10.2 Objective

Various knowledge in dentistry and related patient data for prosthodontic treatment are collected in advance

Suggesting recommended cases and possible solutions for the prosthesis

6.7.10.3 Narrative (Short description)

This service includes sufficient dental knowledge and patient data for prosthodontic treatment, and uses a variety of artificial intelligence techniques to provide recommended practices and possible solutions for prosthodontics.

6.7.10.4 Challenges and issues

Discovery satisfied solutions based on medical knowledge and clinical data

Reasoning novel cases by combining expert knowledge and case studies

6.7.10.5 Societal concerns

Improving medical service efficiency and patient satisfaction.

SDGs to be achieved: Good health and well-being for people

6.7.11 Infant SID (A.74)

6.7.11.1 Scope

Use of facial recognition in healthcare

6.7.11.2 Objective

None identified.

6.7.11.3 Narrative (Short description)

ML-based facial recognition technology detects when infant is lying on her back or face down, alerting care taker to intervene when infant in on her stomach, hence lowering the statistical chance of infant death syndrome (SID)

6.7.11.4 Challenges and issues

— Explainability and transparency regarding the training data used, from the perspective of privacy concerns, and racial and ethnics biases which may be unintentionally built into the trained model.

— Need a structured, common and standardized way to describe the stages of the machine learning model training process, and the types and aspects of the data used in the various stages of the process so the stakeholders (policy makers, privacy advocates and customers) can build confidence and trust in such ML-based product or service. The various aspects of data are described in ISO/IEC 19944 and the new version of it.

6.7.11.5 Societal concerns

— Cost and availability of the ML-based service for low income populations who may not have access to high speed internet access or may not afford the ML-based cloud service

— Any unintentional bias built into the training data used which may hinder effectiveness of the algorithm when used with infants from other races or ethnic backgrounds

SDGs to be achieved: Good health and well-being for people

6.7.12 Discharge Summary Classifier (A.79)

6.7.12.1 Scope

Decision Tree, Random Forest, SVM, BNN, Deep Learning

6.7.12.2 Objective

Classification of Discharge Summaries

6.7.12.3 Narrative (Short description)

This system proposes a method for construction of classifiers for discharge summaries.

6.7.12.4 Challenges and issues

Computational Complexity

6.7.12.5 Societal concerns

- Refinement of Medical Texts
- Medical Hospital Management

SDGs to be achieved: Good health and well-being for people

6.7.13 Generation of Clinical Pathways (A.80)

6.7.13.1 Scope

Decision Tree, Clustering

6.7.13.2 Objective

Nursing clinical pathway

6.7.13.3 Narrative (Short description)

This system proposes a temporal data mining method to construct and maintain a clinical pathway used for schedule management of clinical care.

6.7.13.4 Challenges and issues

Computational Complexity

6.7.13.5 Societal concerns

Good Practice of Medical Services

SDGs to be achieved: Good health and well-being for people

6.7.14 Hospital Management Tools (A.81)

6.7.14.1 Scope

Temporal Data Mining, Visualization

6.7.14.2 Objective

Hospital Management

6.7.14.3 Narrative (Short description)

Temporal Data Mining Methods (Multi-scale comparison with clustering and Temporal Frequent Item Sets) is applied to Hospital Data.

6.7.14.4 Challenges and issues

Computational Complexity

6.7.14.5 Societal concerns

Good Practice of Medical Services

SDGs to be achieved: Good health and well-being for people

6.7.15 Predicting relapse of a dialysis patient during treatment (A.87)

6.7.15.1 Scope

Build an AI solution to augment dialysis nurses.

6.7.15.2 Objective

Use AI to predict if a patient may relapse during dialysis to reduce patient trauma.

6.7.15.3 Narrative (Short description)

A deep learning model to learn from historical and real-time parameters about a patient to identify the probability he or she may relapse during dialysis.

6.7.15.4 Challenges and issues

Challenges in feature engineering the scores of datasets into a logical format that allows the prediction model to retrain without need for high compute.

6.7.15.5 Societal concerns

Lack of reliable and accessible healthcare facilities.

SDGs to be achieved: Good health and well-being for people

6.7.16 Instant triaging of wounds (A.89)

6.7.16.1 Scope

Build an AI solution to augment triaging decisions of wound nurses.

6.7.16.2 Objective

Use AI to identify and classify the intensity of wounds.

6.7.16.3 Narrative (Short description)

A computer vision model able to use RGB and IR wavelengths to measure the size, depth and intensity of a wound.

6.7.16.4 Challenges and issues

Challenges in integrating RGB models and IR models into a single, interpretable visualization for the nurses.

6.7.16.5 Societal concerns

Shortfalls in access to trained nurses and medical imaging technology.

SDGs to be achieved: Good health and well-being for people

6.7.17 Accelerated acquisition of magnetic resonance images (A.101)

6.7.17.1 Scope

Innovations in MRI image formation.

6.7.17.2 Objective

Developing new approaches to MRI image formation aimed at reducing image acquisition time while maintaining the diagnostic image quality.

6.7.17.3 Narrative (Short description)

Magnetic resonance imaging (MRI) is an essential instrument in precision diagnostics of neurological, oncological, musculoskeletal and other diseases. However, long acquisition times combined with the requirement for patient stillness pose a challenge for both patient and the radiology department, leading to high exam costs. Recent advances in sparse raw signal acquisition and specific image reconstruction show that it is possible to significantly reduce the acquisition time.

6.7.17.4 Challenges and issues

— Image quality measurements shall correlate with the diagnostic value – extensive clinical validation and A/B testing is needed, but it is expensive.

— It is necessary to guarantee quality for all possible combinations of MRI sequence parameters, anatomical areas, patient cohorts, or to be very conservative in defining the limits of applicability.

6.7.17.5 Societal concerns

(If safety/quality is guaranteed), MRI imaging will be used more often, more images will be generated which will increase radiologists’ workloads. Development of AI-assisted image interpretation tools will be very much demanded.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.7.18 AI based text to speech services with personal voices for speech impaired people (A.103)

6.7.18.1 Scope

All people who has some sort of speech impairments including but not limited to three basic types: articulation disorders, fluency disorders, and voice disorders.

6.7.18.2 Objective

People with speech impairments will be fully integrated into social processes without communication restrictions.

6.7.18.3 Narrative (Short description)

Communication with other people can be difficult for those who have speech disorders. This seriously complicates communication with the surrounding domestic processes and the involvement of a person in society. A personal wearable device is capable of online synthesizing voice over text or correcting distorted speech. The voice can be fully synthesized with individually selected tone, timbre and pronunciation style settings.

6.7.18.4 Challenges and issues

— Minimization of source records to create a synthesized voice from tens of hours to several tens of minutes

— Hardware requirements for voices based on neural networks should be reduced to the level available on wearable devices.

— The ability to control intonations, speech style should be expanded for use in a natural dialogue between people.

6.7.18.5 Societal concerns

SDGs to be achieved: Good health and well-being for people

6.7.19 AI Platform for Chest CT-Scan Analysis (early stage lung cancer detection) (A.105)

6.7.19.1 Scope

Detecting malignant neoplasms (lungs) on chest CT-scans.

6.7.19.2 Objective

To facilitate early stage oncology chest CT-scans through the application of the Botkin.AI platform based on artificial intelligence.

6.7.19.3 Narrative (Short description)

"Botkin.AI" is a software platform for the diagnosis and assessment of pathology risks using artificial intelligence technologies. The product supports radiologists and oncologists, facilitating the analysis and recognition of diagnostic images of CT-scans, digital X-rays and mammography. The project aims to reduce costs and improve diagnostic accuracy, while detecting pathologies at early stages.

6.7.19.4 Challenges and issues

Challenges: Achieving a higher confirmed level than accredited radiologists in the detection of lung cancer.

6.7.19.5 Societal concerns

SDGs to be achieved: Good health and well-being for people

6.7.20 AI-based design of pharmacologically relevant targets with target properties (A.107)

6.7.20.1 Scope

AI-based engineering of G protein-coupled receptors with enhanced stability.

6.7.20.2 Objective

— Given: protein template in a form of a protein sequence or structure; target properties.

— Predict: protein sequence that satisfies target properties and has minimal differences from the given template.

6.7.20.3 Narrative (Short description)

Molecular design is one of the most important and rapidly developing fields in biotechnology, where the protein engineering plays a significant role in major topics. With an accumulation of biophysical data, AI-based approaches become beneficial in protein design for biotechnology. A particular case is to design stable forms of pharmacological targets, such as G protein-coupled receptors (GPCRs). Malfunctions of these receptors typically lead to various diseases: neurodegenerative, oncological and cardiovascular diseases, asthma, depression, obesity, drug dependence, etc. GPCR receptors are one of the main targets for pharmacological companies, and about 1/3 of all drugs produced in the world are oriented on GPCRs. Obtaining the spatial structure of a single receptor is an extremely difficult and resource-intensive task. We developed an innovative AI-based digital platform for GPCR design, which allowed for a technological breakthrough in obtaining spatial structures of GPCR for the rational development of a new generation drugs.

6.7.20.4 Challenges and issues

Biophysical data is typically very noisy, and the results critically depend on the used experimental assay and initial conditions. Therefore, the training data must be carefully processed with expert knowledge. Consequently, the derived prediction models must rigorously analyzed for robustness, domain applicability, and generalizing power.

6.7.20.5 Societal concerns

Discovery of more efficient, safer and personalized drugs.

SDGs to be achieved: Good health and well-being for people

6.7.21 AI-based mapping of optical to multi-electrode catheter recordings for Atrial Fibrillation Treatment (A.108)

6.7.21.1 Scope

Predicting possible targets for Atrial Fibrillation Ablation based on explanted human heart data of two modalities (multi-electrode mapping and near-infrared optical imaging).

6.7.21.2 Objective

— Given: Recordings from multi-electrode catheter grid, with ground-truth labels from near-infrared optical mapping, obtained from explanted hearts.

— Output: possibility of recordings to be from source (driver) region of atrial fibrillation.

6.7.21.3 Narrative (Short description)

Atrial fibrillation (AF) is the leading cause of stroke with low treatment rate maintained by micro-anatomic intramural re-entry called drivers. Unfortunately, the current clinical method to look for drivers (multi-electrode mapping, MEM) suffers from many limitations, including poor resolution and only-surface tissue mapping. On the other hand, near-infrared optical mapping (NIOM) has 1000 times higher resolution and records electrical activity from the depth of atrial tissue (up to 5 mm), but needs specific voltage-sensitive dye. For our research, we used simultaneous recordings of AF episodes from Ohio State University. We predicted the possibility of AF drivers to be visible in the MEM recording as trained by the Optical ex-vivo data. We created the machine learning classifier with ground-truth labels based on NIOM maps. As features, we used characteristics from the Fourier spectra of MEM recordings.

6.7.21.4 Challenges and issues

— There is only one laboratory in the world that provide the needed explanted human atria.

— The number of experiments is limited (approximately 20 atria per year), and collecting the data is difficult.

— Only a few experiments consist of two modalities recordings and are therefore suitable for this research.

6.7.21.5 Societal concerns

Better life quality for Atrial Fibrillation patients, diminishment of stroke accidents caused by Atrial Fibrillation genesis; as a result, decreased mortality of such patients.

SDGs to be achieved: Good health and well-being for people

6.7.22 Generation of Computer Tomography scans from Magnetic Resonance Images (A.115)

6.7.22.1 Scope

Restoration of naturally distorted microscopy images for following visualization and analysis of meaningful patterns of protein formation inside living cells.

6.7.22.2 Objective

Create a method for automatic analysis and clustering of cell microscopy images, including microscopy of multilayer 3D objects, and implement the developed method for processing of 2D/3D images of cultured human cell models and clustering based on protein modification patterns.

6.7.22.3 Narrative (Short description)

Patterns of protein modification inside cells play an important role in the regulation of gene expression. Here, we aim to develop a method allowing for a detailed analysis of the meaningful protein formation inside living cells with visualization and the processing of microscopy cell images. However, the observed microscopy images suffer from visible artifacts related to blurriness and noise. In this work, we aim to implement AI methods throughout the pipeline of microscopy cell image restoration and analysis. Thereafter, we plan to implement AI approaches for the extraction of meaningful patterns of protein modifications inside cells and use this information for effective cell clustering. Our experiments are on 2D images as well as multilayer 3D objects. To the best of the author's knowledge, this is the first work to apply AI for living cells featuring extraction and clustering.

6.7.22.4 Challenges and issues

— An effective localization of living cells without losing meaningful information must be done.

— Multilayer 3D objects require more computational time and resources, as well as slightly different restoration approaches, due to the 3D object formation model, compared to 2D images.

6.7.22.5 Societal concerns

The developed method of analysis of protein modifications inside living cells is applicable to a wide range of biological and biomedical tasks, far beyond the scope of this project.

SDGs to be achieved: Good health and well-being for people

6.7.23 Generation of Computer Tomography Scans from Magnetic Resonance Images (A.116)

6.7.23.1 Scope

Train a model that generates CT images from MRI scans. Synthetic CT image may be used for radiation dose calculation in radiation therapy.

6.7.23.2 Objective

Generation a CT image from a given MRI image.

6.7.23.3 Narrative (Short description)

Generating radiological scans has grown in popularity in recent years. Here, we generate synthetic Computed Tomography (CT) images from real Magnetic Resonance Imaging (MRI) data. Our architectures were trained on unpaired MRI-CT data and then evaluated on a paired brain dataset. The MRI-CT translation approach holds the potential to eliminate the need for the patients to undergo both examinations and to be clinically accepted as a new tool for radiotherapy planning.

6.7.23.4 Challenges and issues

— Large amounts of paired MRI-CT data is not easily available.

— Doctors are reluctant to accept synthetic CT scans.

6.7.23.5 Societal concerns

Savings for oncologic patients. Reduced radiation dosage.

SDGs to be achieved: Good health and well-being for people

6.7.24 Improving the knowledge base of prescriptions for drug and non-drug therapy and its use as a tool in support of medical professionals (A.117)

6.7.24.1 Scope

Providing the medical professional with methods and means that will allow, within the time allotted for the appointment of a patient with a known nosology, to make a high-quality choice of drugs and to formulate a prescription corresponding to “good medical practices”.

6.7.24.2 Objective

Helping a medical professional consider the influence of a selected drug therapy, as well as monitor the patient’s vital characteristics to reduce the risk of wrong prescriptions and to prevent negative consequences from the prescribed drugs.

6.7.24.3 Narrative (Short description)

Generating radiological scans has grown in popularity in recent years. Here, we generate synthetic Computed Tomography (CT) images from real Magnetic Resonance Imaging (MRI) data. Our architectures were trained on unpaired MRI-CT data and then evaluated on a paired brain dataset. The MRI-CT translation approach holds the potential to eliminate the need for the patients to undergo both examinations and to be clinically accepted as a new tool for radiotherapy planning. Services are developed designed to improve the efficiency and quality of medical care in third-level medical organizations, which have in their structure units providing high-tech medical care. A knowledge base of prescribed drug and non-drug therapy was formed based on the RLS® database. For its improvement and scaling throughout the industry, it is advisable to use AI methods.

6.7.24.4 Challenges and issues

— The existence in parallel of several CR used by doctors.

— The difference in the information of CR and IMU.

— The need for complementing the information of CR and IMU.

— The discrepancy between the information of CR and the real situation in the pharmaceutical market.

6.7.24.5 Societal concerns

The widespread use of the solution will allow the doctor:

— Develop competencies in the field of drug selection, considering VC and drug interactions when prescribing.

— Reduce the risks of erroneous prescriptions.

— Improve the quality of medical care.

In the end, this will allow:

- Preserve the health of the patient, and of their loved ones.
- Extend the quality of a full life.

SDGs to be achieved: Good health and well-being for people

6.7.25 Neural Network Formation of 3D-models orthopedic insoles (A.121)

6.7.25.1 Scope

Artificial intelligence methods using to construction of individual medical products to reduce the risk of developing diseases of the musculoskeletal system.

6.7.25.2 Objective

Development of comfortable, individualized, anatomically correct orthopedic 3D insoles for the treatment of flat feet.

6.7.25.3 Narrative (Short description)

Using artificial intelligence methods, the system converts a pre-scanned foot print into an innovative, medically-based 3D-insole. The AI-system will independently make a medical decision based on the collected medical history, and anthropometric data.

Initial training of the AI-system will take place together with the doctor. In the future, the system will begin by independently choosing the most suitable location options for a patient vaults and indentations and plan an anatomically correct and secure 3D-insole.

6.7.25.4 Challenges and issues

None identified.

6.7.25.5 Societal concerns

SDGs to be achieved: Good health and well-being for people

6.7.26 Search of undiagnosed patients (A.127)

6.7.26.1 Scope

Search of undiagnosed patients with orphan diseases, define patients' journey.

6.7.26.2 Objective

Deep semantic analysis of unstructured texts (based on meaning, rather than keywords, i.e. using natural language processing technology).

6.7.26.3 Narrative (Short description)

Knowledge extraction from the massif of user posts in patient forums, and physicians' professional networks, health-related portals, etc.

6.7.26.4 Challenges and issues

Personal data of the subjects planned to be identified, especially patients', i.e. special health information could potentially be in risk area.

6.7.26.5 Societal concerns

SDGs to be achieved: Good health and well-being for people

6.7.27 Support system for optimization and personification of drug therapy (A.129)

6.7.27.1 Scope

It is a full-range of integrated solutions for the selection of the optimal type of drug, its dose, and its combination with other drugs.

6.7.27.2 Objective

Support system for optimization of the medical therapy of the patient taking into account their individual physiological features, type, and disease severity.

6.7.27.3 Narrative (Short description)

Data from the laboratory and clinical examinations of a particular patient are displayed in a single integrative medical record.

There is currently a significant amount of patient data available electronically. Based on the pool of data of patients receiving a known drug, training is conducted in the recommendation system using AI, taking into account their individual physiological characteristics, type, and severity of the disease, as well as the particular drug's combined administration with other drugs.

When requesting recommendations for a patient, after entering information of their current condition, the system will give individualized recommendations for optimizing drug therapy. Furthermore, the system in the course of treatment, receiving fresh data, makes recommendations for the correction of therapy.

6.7.27.4 Challenges and issues

In addition to the classic data analysis with new technologies to find hidden patterns in relation to health care, the possibility of using methods and technologies to analyze a heterogeneous mass of data with a significant percentage of emissions and uneven distribution of data by classes and categories is a challenge. Of challenge is well is identifying hidden correlations and thereby improving the quality of medical services.

6.7.27.5 Societal concerns

— Incorrect decision.

— Unexplainable result.

— Improving the effectiveness of drug therapy.

SDGs to be achieved: Good health and well-being for people

6.7.28 Syntelly - computer aided organic synthesis (A.130)

6.7.28.1 Scope

Recent progress in deep learning has made a revolution in many areas of science and technology. However, the potential of this method in drug discovery has not yet been fully elaborated. The Syntelly project intends to close this gap. We are developing a web-based platform that helps chemists navigate through chemical space by predicting synthetic availability and ways of synthesis for new drug candidates that have not yet been studied; it also estimates the potential efficiency and safety of specific molecules. We hope that the successful implementation of our project will reduce drug discovery costs and related risks, which will stimulate pharmaceutical companies to search for unexplored molecules as a base for a new generation of drugs.

6.7.28.2 Objective

Exploration of chemical space is a very complicated task due to a large number of predicted chemical molecules. The number of described molecules is only several million compounds, but the estimated number of potentially synthetically accessible molecules is enormous: around 10^60, and neither man nor machine can directly process such a volume of data. The only hope is the development of methods and tools, based on deep learning, which will trigger a chemist-machine alliance to analyze chemical Big Data.

6.7.28.3 Narrative (Short description)

The Syntelly project is directed to help organic chemists in chemical space exploration. Due to high risks and cost of new molecule trials, pharmaceutical companies do not prefer to open new chemical space areas in an experimental way. Using deep learning based on the chemical reaction databases, we predict the best retrosynthesis pathway to achieve the easiest way to a molecule synthesis. The next task is the prediction of the toxicity and bioconcentration of the molecule.

6.7.28.4 Challenges and issues

- a) The large size of chemical space implies the development of machine learning algorithms in two directions: to generate molecules and estimate their parameters, and for chemical space customization for new synthetic pathways

- b) Characteristics of organic compounds are extremely diverse. They are collected from different sources and may be represented in many ways (i.e. toxicity can be measured on different animals).

- c) There are only two major players on the market of chemical and reaction data, and the possibilities to obtain the whole datasets required for deep learning are heavily restricted.

- d) Synthetic and medical chemists prefer to ignore computer-based approaches.

6.7.28.5 Societal concerns

Our primary goal is to make the drug discovery process easier and cheaper. It will stimulate pharmaceutical companies and academic researchers to study new compounds and new scaffolds. Finally, society will obtain new effective drugs against the most dangerous bacterial and viral diseases. Reducing risks will generate interest in developing drugs for orphan diseases, which is now one of the biggest problems for society.

SDGs to be achieved: Good health and well-being for people; responsible consumption and production

6.7.29 WebioMed Clinical Decision Support System (A.131)

6.7.29.1 Scope

Screening for cardiovascular disease risk prediction with machine and deep learning methods.

6.7.29.2 Objective

Advances in precision medicine will require an increasingly individualized prognostic evaluation of patients in order to provide the patient with appropriate therapy.

6.7.29.3 Narrative (Short description)

Cardiovascular disease (CVD) continues to be the most relevant health problem of most countries in the world, including the Russian Federation. According to the World Health Organization, more than 17 million people die each year from CVD worldwide, including more than 7 million from coronary heart disease (CHD).

The machine learning models outperformed traditional approaches for CVD risk prediction (such as SCORE, PROCAM, and Framingham equations). This approach was used to create a clinical decision support system (CDSS). It uses both traditional risk scales and models based on neural networks. Of notable importance is the fact that the system can calculate the risk of cardiovascular disease automatically and recalculate immediately after adding new information to the EHR. The results are delivered to the user's personal account.

6.7.29.4 Challenges and issues

Challenges: to provide physician tools to easily calculate cardiovascular risk anywhere in a world.

6.7.29.5 Societal concerns

One of the major concerns about AI-assisted CDSS is how the machines reach decisions, and whose decision should prevail when there is disagreement between the CDSS and the medical professional. This lack of transparency is referred to as the ‘black box’ of AI. In addition to the lack of transparency, the necessary use of large training data sets coupled with mathematical and statistical algorithms and sometimes neural networks, whether with or without full understanding of the internal workings, presents a challenge in educating doctors to use these tools in a clinically relevant way.

SDGs to be achieved: Good health and well-being for people

6.8 Home/Service Robotics

6.8.1 Robot consciousness (A.61)

6.8.1.1 Scope

A robot for museum tours equipped with the main capabilities of functional consciousness, accepted and transparent to untrained users.

6.8.1.2 Objective

The robot “CiceRobot” offering guided tours in indoor and outdoor museum and equipped with capabilities of functional consciousness, with no concern on the robot qualitative experience. The objective of case study is the acceptance and transparency of the autonomous behavior of the robot in an environment populated with untrained users as the museum visitors.

6.8.1.3 Narrative (Short description)

The “CiceRobot” is a robot with capabilities associated with functional aspects of consciousness. CiceRobot offered indoors guided tours and outdoors guided tours. The outcome of the project is the acceptance and transparency of the autonomous behavior of the robot towards untrained visitors.

6.8.1.4 Challenges and issues

The primary challenge of robot consciousness is the transparency and acceptance of robot operations, important in environments populated by untrained people as tourists in an archaeological museum.

6.8.1.5 Societal concerns

The main concern may be the capability of the robot to act in a way which may is considered unethical to humans.

6.8.2 Social humanoid technology capable of multi-modal context recognition and expression (A.65)

6.8.2.1 Scope

Human-AI sympathetic technology expressing dynamic immersive dialogue with humans through a combination of various artificial intelligence technologies.

6.8.2.2 Objective

— Sympathetic dialogue technology in order to understand socio-cultural consensus and emotions.

— Creation of para-verbal expressions to induce sympathy with a speaker.

— Representing non-verbal expressions reflecting the emphasis and intention of each utterance.

— Deep dialogue management and combination of multimodal expressions for in-depth sympathy while conversations.

6.8.2.3 Narrative (Short description)

A highly immersive sympathetic conversation technology based on artificial intelligence that includes integrated understanding and expression skills of verbal, nonverbal, and para-verbal information to derive complete communion with humans

6.8.2.4 Challenges and issues

Multimodal data understanding / inference / representation

6.8.2.5 Societal concerns

The increase in the elderly population and the decrease in the total population are increasing the inequality of social welfare and benefits according to generation, class and region.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.8.3 Application of Strong Artificial Intelligence (A.111)

6.8.3.1 Scope

Economic sectors and social services.

6.8.3.2 Objective

Find accurate and universal application of strong artificial intelligence.

6.8.3.3 Narrative (Short description)

Strong artificial intelligence is a digital twin of human intelligence, capable of learning, retraining, self-realization and development by improving functional activities through the mastery of creative and innovative high-tech professional and behavioral skills and competences according to a criteria of preferences with qualitative choices.

6.8.3.4 Challenges and issues

Qualitatively new type of thinking not available to humans.

6.8.3.5 Societal concerns

Security and ethical and legal aspects.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9 ICT

6.9.1 Autonomous Network and Automation Level Definition (A.30)

6.9.1.1 Scope

Communications network

6.9.1.2 Objective

To define autonomous network concept and automation level for the common understanding and consensus.

6.9.1.3 Narrative (Short description)

With the goal of providing common understanding and consensus for autonomous self-driving network, this use case delivers a harmonized classification system and supporting definitions that:

— Define the concept of autonomous network.

— Identify six levels of network automation from “no automation” to “full automation”.

— Base definitions and levels on functional aspects of technology.

— Describe categorical distinctions for a step-wise progression through the levels.

— Educate a wider community by clarifying for each level what role (if any) operators have in performing the dynamic network operations task while a network automation system is engaged.

6.9.1.4 Challenges and issues

Data usage and sharing, human expertise & competence

6.9.1.5 Societal concerns

None.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.2 Autonomous network scenarios (A.31)

6.9.2.1 Scope

Communications network.

6.9.2.2 Objective

Clarification and showcases of autonomous network usage.

6.9.2.3 Narrative (Short description)

Multiple scenarios of autonomous network enabled by AI is addressed for improving operational efficiency, customer experience and service innovation, including wireless network performance improvement, optical network failure prediction, data center energy saving etc.

6.9.2.4 Challenges and issues

Data usage and sharing, human expertise & competence.

6.9.2.5 Societal concerns

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.3 A judging support system for gymnastics using 3D sensing (A.70)

6.9.3.1 Scope

Skeleton recognition for gymnastics

6.9.3.2 Objective

To support judgement of difficult element by high-level and high-speed.

6.9.3.3 Narrative (Short description)

We have been developing a judging support system for artistic gymnastics to enhance accuracy and fairness in judging. We developed a skeleton recognition technique using the learned model that we trained using a large amount of depth images of gymnastics created from CG in advance. With this technology, it is possible to recognize a human 3D skeleton from depth image.

6.9.3.4 Challenges and issues

Challenges: Recognize skeleton of all gymnastics element.

Issues: Recognize 3D skeleton in gymnastics that are complex movements from depth image.

6.9.3.5 Societal concerns

Positive: Fairness of scoring, reducing burden of referee, and technical improvement of gymnast.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.4 Active Antenna Array Satellite (A.71)

6.9.4.1 Scope

Determine optimal spot beam patterns for communication satellites in order to react to changing geographic distribution and bandwidth requirements of terminals

6.9.4.2 Objective

Optimise service quality and bandwidth allocation for users of satellite system

6.9.4.3 Narrative (Short description)

Future high throughput satellites (HTS) will be equipped with an active antenna array instead of a fixed multiple spot beam pattern. This allows generating multiple spot beams with different number, size and shape. Moreover, the parameters, i.e. number, size and shape, can be adapted in a flexible way.

6.9.4.4 Challenges and issues

None identified.

6.9.4.5 Societal concerns

Potential to provide demand-adapted service coverage in sparsely populated areas that might not be well served in a fixed configuration scenario

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.5 Carrier interference detection and removal for satellite communication (A.72)

6.9.5.1 Scope

Machine-learning-based detection, classification and removal of interference signal for satellite communication systems

6.9.5.2 Objective

Detection (and possibly classification) of interfering signals in satellite communication systems (e.g., DVB-S2 or DVB-S2x), and removal of the interfering signal using the gained knowledge about the interfere characteristics, with the aim of reducing the error rate at the receiver.

6.9.5.3 Narrative (Short description)

In satellite communication systems, unintended or intended interferences are quite common. For instance, interferences might originate from a mis-pointed terminal antenna, a radar signal or from another terrestrial radio source. In this use-case, the intention is to detect the presence of an interferer in addition to a desired carrier and potentially classify it.

The setting for this use-case is as follows:

— The terminal receives a desired carrier.

— The details of the desired carrier are known, e.g. a DVB-S2x carrier with known symbol rate and modulation scheme.

— There might be an interferer present with unknown frequency, bandwidth and structure.

— The objective is to detect the presence of such an interferer and to classify the interferer, e.g. in terms of power, bandwidth and type.

— Additionally, it may be desired to remove the influence of the interferer from the signal.

6.9.5.4 Challenges and issues

Performance and robustness needs probably be defined w.r.t. a certain class of signals (e.g. DVB-S but not generally)

6.9.5.5 Societal concerns

None identified.

6.9.6 Ontologies for Smart Buildings (A.78)

6.9.6.1 Scope

Renovation of buildings, improve the life’s quality of residents - limited to data issues in a building, - Audience: citizen, public and private actors, companies involved in the ICT System managing the building. Building Management System (BMS) is not the limited scope, we would like to open it to data produced by residents, coupled with data coming from BMS.

6.9.6.2 Objective

None identified.

6.9.6.3 Narrative (Short description)

The general question is How to build and to standardize ontologies for data produced, in a broad sense, in a building. Data are coming both from the System managing the building but also from residents.

6.9.6.4 Challenges and issues

None identified.

6.9.6.5 Societal concerns

None identified.

6.9.7 Product failure prediction for critical IT infrastructure (A.86)

6.9.7.1 Scope

Building an AI solution to augment QA engineers.

6.9.7.2 Objective

Reduce the likelihood of releasing defective batches of hardware.

6.9.7.3 Narrative (Short description)

A deep learning model to learn from a visual representation of the number of items that failed in a specific batch of hardware as well as the type of defect.

6.9.7.4 Challenges and issues

Challenges in identifying which deep learning model gives the best performance output, and challenges in indexing raw flat files into visualization images.

6.9.7.5 Societal concerns

Address issues of sustainable manufacturing and high-value technical jobs.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.8 Data compression with AI techniques (A.98)

6.9.8.1 Scope

Data center/Supercomputing center.

6.9.8.2 Objective

Fast data transfer via WAN.

6.9.8.3 Narrative (Short description)

Improving Data Compression with Deep Predictive Neural Network for Time Evolutional Data.

6.9.8.4 Challenges and issues

More accurate prediction to data to be compressed.

6.9.8.5 Societal concerns

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.9 Optimization of software configurations with AI techniques (A.99)

6.9.9.1 Scope

Data center/Supercomputing center.

6.9.9.2 Objective

Optimization of software configurations.

6.9.9.3 Narrative (Short description)

Optimizing Asynchronous Multi-level Checkpoint/Restart Configurations with Machine Learning.

6.9.9.4 Challenges and issues

More accurate prediction for the optimization.

6.9.9.5 Societal concerns

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.9.10 Better human-computer interaction with advanced language models (A.100)

6.9.10.1 Scope

Human-computer interaction.

6.9.10.2 Objective

Improve quality of human-computer interaction.

6.9.10.3 Narrative (Short description)

Better language models are crucial for improving the quality of human-computer interaction, for example tasks like question answering, summarization etc. We use large-scale compute systems to develop better language models by exploiting neural architecture search, large datasets and holistic evaluation framework.

6.9.10.4 Challenges and issues

High computational costs.

6.9.10.5 Societal concerns

SDGs to be achieved: Partnerships for the goals

6.10 Legal

6.10.1 Tax Rules Updates and Classification (A.95)

6.10.1.1 Scope

Build an AI solution that identify updates on tax laws and classify them.

6.10.1.2 Objective

Use NLP to identify new tax laws from different countries and classify them.

6.10.1.3 Narrative (Short description)

An NLP model that helps an investment firm identify tax laws and trends that have an impact on their current and future portfolio.

6.10.1.4 Challenges and issues

The classes are pre-determined, and if these are changed, it will affect the ability of the model to re-classify.

6.10.1.5 Societal concerns

Erratic changes in local and cross-border tax rules which have repercussions on economic growth.

SDGs to be achieved: Decent work and economic growth

6.10.2 AI Contract Management (A.120)

6.10.2.1 Scope

Building an AI Contract Management solution for the business process of documents automation: data classification, automatic data extraction and contract monitoring.

6.10.2.2 Objective

Creating a solution that is able to standardize contract management process, improve quality of work on problematic contracts and claims and optimize lawyers' working process and relieve them from routine tasks.

6.10.2.3 Narrative (Short description)

MTS AI Contract Management solution is built on our AI legal core, which includes technology that enables to convert different types of documents into digital format, replicate the natural human-like text recognition and extract data to automate business tasks.

6.10.2.4 Challenges and issues

— Noisy data (different scans quality).

— Working with private data (information security).

— Non-linear models need significant computational power during the training phase.

6.10.2.5 Societal concerns

We create the helpful industrial solution that can optimize the current contract management process and assist to make easier the legal departments job.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.10.3 Semantic Analysis of Legal Documents (A.128)

6.10.3.1 Scope

Semantic analysis of legal documents in the course of its development, verification and improvement.

6.10.3.2 Objective

Machine understanding of the meaning of legal documents.

The obtaining of semantic visual images of documents; the detection of contradictions and inaccuracies in legal documents describing similar objects of law for the task of classifying documents; quick document comprehension; and analyzing the consistency of the legal base.

6.10.3.3 Narrative (Short description)

The software tool is oriented on the analysis and representation content of normative documents in the form of formal ontology (OWL ontology) and the construction of their visual images for the subsequent detection of inaccuracies and contradictions using logical inference and visual analysis methods.

6.10.3.4 Challenges and issues

Different levels of abstraction of concepts in documents.

6.10.3.5 Societal concerns

None identified.

6.11 Logistics

6.11.1 Improving Productivity for Warehouse Operation (A.41)

6.11.1.1 Scope

Big data analysis for enhancing productivity.

6.11.1.2 Objective

To improve productivity of warehouse operation by detecting and changing controllable factors.

6.11.1.3 Narrative (Short description)

AI-driven operating system that uses big data from work performance information to issue appropriate work instructions has been developed. In PoC, picking operation improvement was conducted in a distribution warehouse. As the result, 8% work reduction was performed.

6.11.1.4 Challenges and issues

Understanding of workers' human factors (privacy, additional work etc.)

6.11.1.5 Societal concerns

Solving labor shortage problem and improving labor related issues with aiming improving productivity.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.11.2 AI based dynamic routing SaaS (A.92)

6.11.2.1 Scope

Build an ML model that dynamically corrects routes.

6.11.2.2 Objective

Incorporate last minute human-driven factors into optimising delivery routes.

6.11.2.3 Narrative (Short description)

A machine learning model that dynamically corrects the delivery route and time to delivery.

6.11.2.4 Challenges and issues

Challenges in feature engineering static and dynamic variables, and over reliance on internet connectivity of the dynamic routing device.

6.11.2.5 Societal concerns

Over utilization of resources and emittance of greenhouse gases to fulfil the trend of e-commerce.

SDGs to be achieved: Climate action

6.12 Maintenance & support

6.12.1 Anomaly Detection in Sensor Data Using Deep Learning Techniques (A.45)

6.12.1.1 Scope

Temporal Data captured from sensors.

6.12.1.2 Objective

Identify Anomalies and Events by learning the temporal patterns of sensor data, based on Deep Learning techniques.

6.12.1.3 Narrative (Short description)

Mechanical devices such as engines, vehicles, aircrafts, etc., are typically instrumented with numerous sensors to capture the behaviour and health of the machine. The sensors temporal data has several complex patterns that are very hard to identify with traditional methods. We have proposed the use of Deep Learning algorithms for analysing such temporal patterns for anomaly/event detection, diagnosis, root cause analysis.

Algorithms proposed so far are LSTM-AD, EncDec-AD, online RNN-AD. We used industrial datasets wherever possible and publically available datasets in other scenarios. In most of the cases, our algorithms were significantly better than other methods.

6.12.1.4 Challenges and issues

- Noisy Data
- Data with missing temporal features
- Rarity of Anomalous Data

6.12.1.5 Societal concerns

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.12.2 Jet Engine Predictive Maintenance Service (A.73)

6.12.2.1 Scope

Use of jet engine telemetry data to train predictive maintenance algorithms

6.12.2.2 Objective

None identified.

6.12.2.3 Narrative (Short description)

ML-based jet-engine predictive maintenance technology predicts the next maintenance tasks proactively using machine learning model trained by jet engine telemetry data and maintenance history

6.12.2.4 Challenges and issues

- Explainability and transparency regarding the training data used, from the perspective of corporate confidentiality concerns,
- Need a structured, common and standardized way to describe the stages of the machine learning model training process, and the types and aspects of the data used in the various stages of the process so the stakeholders (policy makers, partners and customers) can build confidence and trust in such ML-based product or service, ensuring that their corporate trade secrets are not leaked when they contribute to shared pools of data used for model training. The various aspects of data are described in ISO/IEC 19944 and the new version of it.

6.12.2.5 Societal concerns

Ability for industry players to share their data with their partners to develop ML-based algorithms while protecting their IP and interest would allow for flourishing of commercial AI/ML applications and solutions.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.12.3 Detection of fraudulent medical claims (A.90)

6.12.3.1 Scope

Build a ML model to classify if a particular claim could be fraudulent.

6.12.3.2 Objective

Upgrade from an only-human-interpretation to an ML-assisted fraud detection.

6.12.3.3 Narrative (Short description)

A machine learning model to identify true anomalies and trends of fraudulent claims customized to the source of fraud.

6.12.3.4 Challenges and issues

The challenge was in building separate models for the each major sources of fraudulent claims.

6.12.3.5 Societal concerns

Unintended or unlawful use of funds that are meant for essential services to people.

SDGs to be achieved: Sustainable cities and communities

6.12.4 AI Virtual Assistant for Customer Support and Service (A.106)

6.12.4.1 Scope

— Customer support service, product and service consulting .

— Limitations - support for dialogs exclusively within MTS products.

— Target audience - b2b, b2c clients of MTS Russia.

6.12.4.2 Objective

Optimization of company resources for support and customer service by automating the customer service process. As a result of the implementation of the system, the company was able to cover a greater volume of customer requests without needing to increase its staff of operators. This allowed the prevention of an increase in the company's operating expenses.

6.12.4.3 Narrative (Short description)

The system automatically answers customer questions in the application and on the company website. At peak, service automation reaches 85%.

6.12.4.4 Challenges and issues

- The readiness of external systems' API for integration with the bot platform.

- Biased customer attitudes towards chatbots.

6.12.4.5 Societal concerns

SDGs to be achieved: Affordable and clean energy

6.13 Manufacturing

6.13.1 AI Solution to Calculate Amount of Contained Material from Mass Spectrometry Measurement Data (A.3)

6.13.1.1 Scope

Calculating amount of contained material from mass spectrometry measurement data using chromatography.

6.13.1.2 Objective

To find an accurate and efficient solution to calculating amount of contained material without dependence on individuals.

6.13.1.3 Narrative (Short description)

An AI solution was developed that could automatically pick the peak related to the contained material from measurement data through deep learning. Compared with manual results by an experienced operator, the automated peak picking results using AI had a false detection rate of 7% and an undetected rate of 9%. The peak picking operation time using AI was estimated to be about one fifth.

6.13.1.4 Challenges and issues

— Challenges: Achieve the same level as experienced operators for peak picking.

— Issues: 1) Lack of training data per contained material, 2) how to create good images for deep learning from mass spectrometry measurement data.

6.13.1.5 Societal concerns

None identified.

6.13.2 AI solution to quickly identify defects during quality assurance process on wind turbine blades (A.4)

6.13.2.1 Scope

Detecting defects in products by inspecting nondestructive testing scanning data.

6.13.2.2 Objective

To find an accurate and efficient solution to detect defects without compromising the detection of in-material damage and risking a loss in reputation.

6.13.2.3 Narrative (Short description)

An AI solution was developed that could automatically detect defects through deep learning together with what is called "imagification"; it achieved high coverage of various defects and evaluation of each nondestructive testing scanning was reduced by 80%, which translated into cost savings, reduced production lead times, and increased productivity.

6.13.2.4 Challenges and issues

Challenges: Achieve the same level as ultrasonic accredited engineers for detecting critical defects.

Issues: 1) Lack of defect data per defect type, 2) how to create good images for deep learning from UT raw data, and 3) back wall detection

6.13.2.5 Societal concerns

SDGs to be achieved: Affordable and clean energy

6.13.3 Solution to Detect Signs of Failures in Wind Power Generation System (A.5)

6.13.3.1 Scope

Detect signs of malfunction (failure) in wind power generators.

6.13.3.2 Objective

Detect signs of failure in wind power generation, earlier than human specialists.

6.13.3.3 Narrative (Short description)

A system is currently in development that uses machne learning to detect signs of equipment failure that would be difficult to detect from visual inspection. Currently, sensor data is being collected from 43 actual domestic large wind turbines, and large-scale verification testing is being conducted. The goal is for a paradigm shift from responding after the fact to maintenance that prevents problems and maintenans safety

6.13.3.4 Challenges and issues

None identified.

6.13.3.5 Societal concerns

None identified.

6.13.4 Generative Design of Mechanical Parts (A.15)

6.13.4.1 Scope

Help mechanical engineers design lighter, strong, better parts.

6.13.4.2 Objective

Create optimized parts following precise mechanical constraint while permitting cost savings by reducing the amount of material necessary to achieve goals.

6.13.4.3 Narrative (Short description)

From Wikipedia: Generative design is an iterative design process that involves a program that will generate a certain number of outputs that meet certain constraints, and a designer that will fine tune the feasible region by changing minimal and maximal values of an interval in which a variable of the program meets the set of constraints, in order to reduce or augment the number of outputs to choose from.

6.13.4.4 Challenges and issues

Challenges: Environment may be cluttered, occlusions of target might occur, objects may move around. Issues: For safety reasons, speed and force of robot need to be limited in assistive environment to avoid harm. Human intervention can happen at any time.

6.13.4.5 Societal concerns

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.5 Information Extraction from Hand-marked Industrial Inspection Sheets (A.21)

6.13.5.1 Scope

Localization and Mapping of machine zones, arrows and text, to extract information from manually tagged inspection sheets.

6.13.5.2 Objective

To create a pipeline to build an information extraction system for machine inspection sheets, by mapping the machine zones to the handwritten code using state-of-the-art deep learning and computer vision techniques.

6.13.5.3 Narrative (Short description)

Inspection Sheets are filled regularly to detect defects and maintain heavy machines. Sheets contains a lot of unstructured information and requires domain experts' intervention to read and digitize. We have proposed a novel pipeline to build an information extraction system for such machine inspection sheets, utilizing state-of-the-art deep learning and computer vision techniques.

6.13.5.4 Challenges and issues

Challenges:

- Quality of Images
- Structural deformities of individual components(arrows, handwritten code)
- Quantity of data
- Cascading effect of error at each stage of the pipeline

6.13.5.5 Societal concerns

Inspection engineers may have to develop other skills.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.6 Automated Defect Classification on Product Surfaces (A.33)

6.13.6.1 Scope

Image Analytics for water taps in sanitary industries.

6.13.6.2 Objective

Image analytics using a combination of feature extraction and classification of defects on shining surfaces in sanitary industries.

6.13.6.3 Narrative (Short description)

A vision system that inspects and identifies the defects on water taps in sanitary industries. The system uses a combination of features for an automatic defect classification on product surfaces. All defects (15 types are identified) are classified into two major categories, real-defects and pseudo-defects. The pseudo-defects cause no quality problem; while the real-defects are critical as they might malfunction the final products.

The AI system uses Support Vector Machine (SVM) classifier along with the combined features to identify the defect types. With the vision system in place, the quality control process is fully automated without any human intervention.

6.13.6.4 Challenges and issues

Real time implementation, accurately identify the nature of defects.

6.13.6.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.7 Robotic Task Automation: Insertion (A.34)

6.13.7.1 Scope

Robotic assembly.

6.13.7.2 Objective

— Simple programing/instruction and flexibility in usage

— Automation of tasks lacking analytic description

— Reliability and efficiency

6.13.7.3 Narrative (Short description)

Assembly process often includes steps where two parts need to be matched and connected to each other through force exertion. In an ideal case, perfectly formed parts can be matched and be assembled together with predefined amount of force. Due to imperfection of production steps, surface imperfection and other factors such as flexibility of parts, this procedure can become complex and unpredictable. In such cases, human operator can be instructed with simple terms and demonstrations and perform the task easily, while a robotic system will need very detailed and extensive program instructions to be able to perform the task including required adaptation to the physical world. The need for such a complex program instruction will make use of automation cumbersome or uneconomical. Control algorithm that are based on machine learning, especially those including reinforcement learning can become alternative solutions increasing and extending the level of automation in manufacturing.

6.13.7.4 Challenges and issues

- Complex and unpredictable assembly process due to imperfection of production steps, surface imperfection and other factors such as flexibility of parts.
- Accuracy of sensing
- Coworking with humans

6.13.7.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.8 Powering Remote Drilling Command Centre (A.36)

6.13.8.1 Scope

Oil and Gas Upstream (Deployed in 150 Oil Rigs and 2.5 Billion+ Data Points each).

6.13.8.2 Objective

Automatic generation of Daily Performance Report, reduction in overall drilling time, cut down Invisible Loss Time and improve rig asset management.

6.13.8.3 Narrative (Short description)

It is important for a drilling contractor to have real time monitoring of rig parameters to optimize operations. The customer lacked granular insights during drilling, could not ascertain the root cause of non-productive time, and manual interpretation of signals led to missing of anomalies further degrading performance.

6.13.8.4 Challenges and issues

Compliance of organizations.

6.13.8.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.9 Leveraging AI to Enhance Adhesive Quality (A.37)

6.13.9.1 Scope

Batch/Continuous/Discrete Manufacturing (Deployed in 75+ manufacturing lines in 10+ countries; Specifically identified the contributors to quality; predict potential quality failures).

6.13.9.2 Objective

Enhance Adhesive Quality, Performance Benchmarking.

6.13.9.3 Narrative (Short description)

Cerebra IOT signal intelligence platform provides the ability to have a holistic perspective and understanding of the sensitivity of the key parameters affecting output quality and ability to monitor and control the process in real-time. This will avoid variations in yields, build-up of inventories and missed customer deadlines.

6.13.9.4 Challenges and issues

Patented process if any, security restrictions.

6.13.9.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.10 Machine Learning Driven Approach to Identify the Weak Spots in the Manufacturing of the Circuit Breakers (A.38)

6.13.10.1 Scope

Detecting the issues in manufacturing process that leads to early failures of the circuit breakers through the data mining of the manufacturing process.

6.13.10.2 Objective

To generate actionable intelligence to improve the manufacturing process of circuit breakers through mining of manufacturing related data.

6.13.10.3 Narrative (Short description)

An approach was developed that can mine the manufacturing data of circuit breakers through multiple machine learning algorithms. The approach could successfully identify the weak spots in the manufacturing where failure rate jumped from 0.2% to 7% (35 fold more probability of failure) and hence candidates for improvement in the manufacturing process.

6.13.10.4 Challenges and issues

Discovering actionable insight with partial data set and managing bias in ML models due to limited number of failed cases.

6.13.10.5 Societal concerns

Safe and reliable power delivery.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.11 Machine Learning Driven Analysis of Batch Process Operation Data to Identify Causes for Poor Batch Performance (A.39)

6.13.11.1 Scope

Detecting the issues in batch manufacturing process that leads to bad quality products or longer cycle times of batch processing.

6.13.11.2 Objective

Provide insight to the operation team to improve the productivity of batch manufacturing through machine learning on historical operation data.

6.13.11.3 Narrative (Short description)

An approach was developed that can use machine learning models to identify issues in batch manufacturing.

6.13.11.4 Challenges and issues

Discovering actionable insight with limited industrial data set, handling dynamics in the process variables.

6.13.11.5 Societal concerns

Consistent batch operation lead to enhanced productivity.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.12 Empowering Autonomous Flow Meter Control- Reducing Time Taken to "Proving of Meters" (A.40)

6.13.12.1 Scope

Calibration of control devices.

6.13.12.2 Objective

Reduce the time taken for trial & error methods to set the VFD and FCV setpoints.

6.13.12.3 Narrative (Short description)

The customer had to set VFD and FCV % manually to achieve desired flowrate using trial & error methods, which could take about 3-4 hours. Efficiency for the proving of the meters was very less & improvement was needed to remove any aberration in reading as it was time consuming.

6.13.12.4 Challenges and issues

None identified.

6.13.12.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.13 Adaptable Factory (A.46)

6.13.13.1Scope

(Semi-)Automatic change of a production system's capacities and capabilities from a behavioral and physical point of view.

6.13.13.2Objective

The objective is to enable flexible production resources which enable fast reconfiguration and adaptation to changing situations, context, and requirements which facilitate optimized resource usage under uncertainty.

6.13.13.3Narrative (Short description)

Rapid, and in some cases completely automated, conversion of a manufacturing facility, by changing both production capacities and production capabilities. This use case describes the adaptability of an individual factory by (physical) conversion and/or adaption of a factory's and its machines behavior in order to adjust to changing situations like disruptions, material quality variation, production of new products, etc.

A prerequisite is a modular and thereby adaptable design for manufacturing within the factory. The result is a need for intelligent and interoperable modules that basically adapted to an altered configuration on their own, and standardized interfaces between these modules.

6.13.13.4Challenges and issues

None identified.

6.13.13.5Societal concerns

Enabling flexible and autonomously reconfigurable production systems ease human-machine configuration, facilitate optimized machine use, reduce failures through autonomous compensation, optimized product quality through prediction techniques.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.14 Order-Controlled Production (A.47)

6.13.14.1Scope

Automatic distribution of production jobs across dynamic supplier networks.

6.13.14.2Objective

The objective is to enable automatic supplier contracting for optimized utilization of manufacturing capabilities at suppliers, novel degrees of flexibility in contract manufacturing, and enable (mass) customized customer ordering.

6.13.14.3 Narrative (Short description)

A network of production capabilities and capacities that extend beyond factory and company boundaries allows for a quick order-controlled adaption to changing market and order conditions. The result is a largely fragmented and dynamic value chain network that change as required by the individual order, and thereby make the best use of capabilities and capacities of existing production facilities. The goal is to allow for automated order planning, allocation and execution, thereby considering all production steps and facilities required to facilitate linking external factories into a company's production process, as automated as possible.

6.13.14.4 Challenges and issues

None identified.

6.13.14.5 Societal concerns

Enabling mass-customized production in global dynamic supply chains, and by that, ease production of small lot sizes for customized products.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.15 Value-based Service (A.48)

6.13.15.1 Scope

Process and status data from production and product use sources are the raw materials for future business models and services.

6.13.15.2 Objective

The objective of this use case is the provision of remote services for product and production based on (generic) service platforms. This use case can be seen as a fundament for the deployment of arbitrary AI remote services.

6.13.15.3 Narrative (Short description)

Service platforms collects data from product use – for example machines or plants – and analyses and processes this data to provide tailor-made individualized services, e.g. optimized maintenance at the proper time, or the timely provision of the correct process parameters for a production task currently being requested. Companies offering these services (service providers) occupy the interface between the product provider and the user.

6.13.15.4 Challenges and issues

None identified.

6.13.15.5 Societal concerns

Increasing complexity of modern cyber-physical production systems cannot be managed by humans. AI technologies provide one solution in this context for more reliable, fault-tolerant, safe and secure production systems.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.16 Surgeries Improvement of productivity of semiconductor manufacturing (A.82)

6.13.16.1 Scope

Analysis of data taken from production equipment and improvement of productivity based on the analysis.

6.13.16.2 Objective

Cost reduction of semiconductor manufacturing.

6.13.16.3 Narrative (Short description)

In modern semiconductor manufacturing, huge amount of data are gathered and used to improve yields. However, it is difficult even for skilled engineers to promptly achieve the improvements by means of manual analysis because of the complexity of the production process and the scale of the data. In Yokkaichi operation, where more than 5,000 pieces of equipment are working and two billion records of data are daily created, it is difficult to secure enough engineers to resolve problems arise in the production. Toshiba Memory Corporation tackled the issue with AI technology including machine learning. The endeavor resulted in improvement of the productivity through the stable quality based on semi-automated data analysis.

6.13.16.4 Challenges and issues

Guarantee of correctness of analysis by AI.

Automatic physical model building for a failure.

6.13.16.5 Societal concerns

Hollowing out of analytic know-how.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.17 AI Decryption of Magnetograms (A.104)

6.13.17.1 Scope

Oil and gas transportation. AI solution to quickly identify defects during the quality assurance process on field pipeline.

6.13.17.2 Objective

— Detection of internal defects (pits, ulcers, etc.).

— Detection of structural elements (welds, bends, etc.).

6.13.17.3 Narrative (Short description)

A solution has been developed that allows for the detection of internal defects and structural elements.

6.13.17.4 Challenges and issues

To achieve high level accuracy recognizing defects and welds.

To reduce the processing time of magnetograms.

6.13.17.5 Societal concerns

Minimizing the risk of environmental disasters associated with oil spills.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.18 Analyzing and Predicting Acid Treatment Effectiveness of Bottom Hole Zone (A.110)

6.13.18.1 Scope

Mining of oil and gas; digital assistant for analyzing and predicting the effectiveness of acid treatments of the bottom hole zone

6.13.18.2 Objective

Predict the effectiveness of acid treatments of the bottom hole zone.

6.13.18.3 Narrative (Short description)

Predicting the technological and economic efficiency of acid treatments of the bottom-hole zone of the well.

6.13.18.4 Challenges and issues

Challenges: To achieve high level accuracy of prediction efficiency of acid treatments.

6.13.18.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are important for facilitating sustainable development.

SDGs to be achieved: Industry, innovation, and infrastructure

6.13.19 Automatic Classification Tool for Full Size Core (A.112)

6.13.19.1 Scope

Oil and Gas exploration, classification of rock types, oil saturation, carbonate and fracture according to core images.

6.13.19.2 Objective

— Classification of rock types.

— Classification of oil saturation.

— Classification of carbonate.

— Classification of fracture according of core.

6.13.19.3 Narrative (Short description)

A solution has been developed that allows for the classification of rock types into four classes. This resulted in an 80% reduction in core image analysis.

6.13.19.4 Challenges and issues

— To achieve the same level of accuracy of recognition of rock types as expert lithologists.

— To minimize the set of laboratory tests due to visual recognition of rock types and their parameters from core images.

6.13.19.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, is important for facilitating sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.20 Intelligent Technology to Control Manual Operations on Video — “Norma” (A.118)

6.13.20.1 Scope

Tooltip visualization technology (augmented reality) based on technological process and manual operations control in the assembly, maintenance, and repair of engineering products.

6.13.20.2 Objective

“Norma” technology will reduce the number of errors made by technical personnel during manual assembly of products to the lowest possible minimum. It visualizes the correct sequence of actions to the user-assembler on top of the parts through augmented reality glasses. Norma controls the correctness of manual operations and the tool used. It fixes the detected deviations in the electronic passport of the product. Additionally, Norma promptly reports identified violations of the process to the quality control department. Norma will provide a dramatic improvement in the quality of production and technological operations without the widespread use of industrial robotics, which will avoid the negative social consequences caused by automation of production.

6.13.20.3 Narrative (Short description)

The Norma technology is designed to control manual operations during assembly, maintenance, and repair of engineering products using video data.

6.13.20.4 Challenges and issues

— Small (or none) number of real photos for training — neural networks shall be trained on a synthetic data.

— Synthetic data shall be generated to cover all possible light conditions in which system can be used.

— System shall operate in real time.

6.13.20.5 Societal concerns

Norm technology will provide quality improvement in production without the use of robotic systems, which will not lead to a reduction in jobs and will therefore avoid negative social consequences.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.21 Optimization of ferroalloy consumption for a steel production company (A.123)

6.13.21.1 Scope

Recommendation for the optimal consumption of ferroalloys at ladle furnace treatment during secondary steelmaking.

6.13.21.2 Objective

Reducing the usage of ferroalloys in metallurgical plants while maintaining alloy quality standards for steel. Improving production efficiency.

6.13.21.3 Narrative (Short description)

Digital advisor in steel ladle treatment. Recommends the optimal consumption of ferroalloys at ladle furnace treatment during secondary steelmaking.

The solution is based on physico-chemical technological models and machine learning models.

Datana Smart uses historical data, different factors and correlations, with high accuracy based on real dependencies on the physical process.

6.13.21.4 Challenges and issues

— There is no data available for creating mathematical models.

— Incorrect/insufficient data; outliers, gaps, accumulated errors, and inaccurate measurements.

6.13.21.5 Societal concerns

Promoting sustainable industries, and investing in innovation, are important for facilitating sustainable development

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.13.22 Device Control Using both cloud AI and embedded AI (A.132)

6.13.22.1 Scope

Learn the user's preferred temperature for each situation for the control of home appliances (air conditioning equipment)

6.13.22.2 Objective

Keep comfortable room status by driving home appliances (air conditioning equipment) at the user's preferred temperature according to the situation.

6.13.22.3Narrative (Short description)

Because temperature that the user feels comfortable depending on the situation, such as the time of day and the day of the week, the user changes set temperature every time the user feels uncomfortable.

By Learning the user's preferred temperature for each situation, home appliances (air conditioning equipment) can keep room comfortable state automatically.

For the learning of the operation with long-term cycle, such as a fixed operation for each day of the week, it is effective learning from the accumulated operation history. So, A model is learning on the cloud.

For sudden operation pattern changes, e.g., when the temperature of the day rises suddenly and user react to it, high frequency online machine learning inside the equipment can adjust the model immediately.

The consistency between the model learned on the cloud and one adjusted inside the equipment should be kept.

6.13.22.4Challenges and issues

— During actual use, there is a possibility of significant difference between the model learned by cloud and the model adjusted in air-conditioner. It leads significant change of temperature setting when the model in the air conditioner is overridden by the model learned by the cloud.

— How and when to detect whether there has been a significant difference.

— How does air-conditioner explain a significant difference when it is detected. Criteria for determining whether or not to explain

6.13.22.5Societal concerns

By automatically adjusting the temperature so that the user feels comfortable, it can suppress unnecessary power due to overtemperature or overcool.

SDGs to be achieved: Affordable and clean energy

6.14 Media and Entertainment

6.14.1 Predictive analytics for the behavior and psycho-emotional conditions of eSports players using heterogeneous data and artificial intelligence (A.125)

6.14.1.1 Scope

Prediction of psycho-emotional conditions of eSports players. To form predictions, we collect the physiological data from wearables/video cameras/eye tracker, game telemetry data from keyboard/mouse/demo files, and environmental conditions followed by the application of machine learning methods for the analysis of the collected data.

6.14.1.2 Objective

Predict psycho-emotional conditions of eSports players in particular game scenarios based on collected heterogeneous data.

6.14.1.3 Narrative (Short description)

eSports is organized video gaming, where single players or teams compete against each other with the aim of achieving a specific goal by the end of the game. The eSports industry has progressed considerably within the last decade: a huge number of professional and amateur teams take part in numerous competitions where the prize pools amount to tens of millions of dollars USD. Its global audience has already reached 380 million in 2018 and is expected to reach more than 550 million in 2021. However, a lack of tools exists to help assess the physiological and psycho-emotional conditions of eSports players.

In this project, we collect three classes of data (physiological, game telemetry, and environmental conditions) followed by a data analysis using artificial intelligence based on machine learning algorithms. For example, we apply machine learning and recurrent neural networks with attention to assessing player performance dynamics.

6.14.1.4 Challenges and issues

The challenges are associated with data collection and data analysis. To create a reasonably large dataset, a high number of Pro eSports athletes is required. Moreover, it is not a trivial task to collect the data during competitions; the sensors must ensure unobtrusive sensing. At the same time, the collected data is truly heterogeneous, e.g. video/time-series/tests, requiring new methods of data storage and data analysis.

6.14.1.5 Societal concerns

Although eSports has evolved from amateur video gaming to a developing and innovative industry, there is a skeptical attitude to eSports in our society. A common understanding in particular communities is that eSports could be dangerous and cannot serve as a profession of the future.

SDGs to be achieved: Good health and well-being for people

6.15 Mobility

6.15.1 Autonomous Apron Truck (A.12)

6.15.1.1 Scope

Automated transportation of luggage (carts) to requested destinations on an airport apron while following local traffic rules and resolve unplanned conflicts.

6.15.1.2 Objective

Automate transport to increase reliability, precision, efficiency and safety.

6.15.1.3 Narrative (Short description)

An AI solution was planned that could operate a luggage truck on an airport apron where it interacts with aircrafts, other machines and humans. It prevents accidents with humans at all times and follows local traffic rules.

6.15.1.4 Challenges and issues

Challenges: Achieve at least the same level as human truck operators.

Issues: 1) detect other apron traffic participants (especially aircraft) including intentions 2) Multiplicity of various outside conditions (e.g. signs painted on road but ice and snow covering it), and 3) prediction of human behaviour (e.g. workers in reverse walk).

6.15.1.5 Societal concerns

Changed work environment for workers during loading/unloading with less interactions with co-workers but more non-social interactions (machines).

6.15.2 AI Solution to Help Mobile Phone to have Better Picture Effect (A.32)

6.15.2.1 Scope

Better understanding the image and improving image effect on smartphone by using DL model which is trained in the cloud or offline.

6.15.2.2 Objective

To find an efficient solution to Increase camera image quality on smartphone without Increasing too much operation and power burden for mobile phone.

6.15.2.3 Narrative (Short description)

An AI solution was developed that could increase smartphone camera image quality. Using deep learning, smartphone can identify more scenarios and objects than before. Based on the identified scenarios and objects, smartphone can better understand the image and improve image effect.

6.15.2.4 Challenges and issues

Challenges: Achieve the same level as professional SLR camera for pictures.

Issues:

— Lack of data for certain scene;

— Lack of computing ability on terminal side；

— Users can feel the improvement of image quality, but may not know that it is brought by AI.

6.15.2.5 Societal concerns

For the wrong object detection, it may lead to racial prejudice or privacy protection problems.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.16 Public sector

6.16.1 AI Ideally Matches Children to Daycare Centers (A.7)

6.16.1.1 Scope

Assignment pattern that satisfies complex applicants' requirements.

6.16.1.2 Objective

To determine the assignment pattern that will fulfill the preferences of as many applicants as possible automatically.

6.16.1.3 Narrative (Short description)

This AI technology automatically determines the assignment pattern while fulfilling as many applicants' preferences as possible by priority ranking by using game theory.

6.16.1.4 Challenges and issues

Challenges: Determine an optimal assignment pattern instantly and fairly depending on unique and complex rules in each local government.

Issues: Long calculation time is required in the case of a large number of children and siblings

6.16.1.5 Societal concerns

— Supporting working women

— Resolving the problem of children waiting for day care

SDGs to be achieved: Decent work and economic growth

6.16.2 AI Sign Language Interpretation System for the Hearing-Impaired (A.62)

6.16.2.1 Scope

Increase the convenience of public services to hearing-impaired people by providing a service to translate sign language image information into natural language

6.16.2.2 Objective

Supporting communication between hearing-impaired and non-disabled people

6.16.2.3 Narrative (Short description)

In this use case scenario, hearing impaired and non-disabled people are able to communicate each other through the AI sign language-natural language interpretation service.

6.16.2.4 Challenges and issues

Multimodal interactions

Translation from visual information to textual information

Translation from textual information to visual information

6.16.2.5 Societal concerns

Promoting welfare and supporting social activities for the disabled

SDGs to be achieved: Good health and well-being for people

6.16.3 AI Situation Explanation Service for the Visually Impaired (A.64)

6.16.3.1 Scope

A real-time situation explanation service through voice for the visually impaired.

6.16.3.2 Objective

- Recognizing Texts around the visually impaired
- Recognizing Faces around the visually impaired
- Recognizing Objects around the visually impaired
- Assisting the mobility of the visually impaired
- Describe scenes and photos for the visually impaired

6.16.3.3 Narrative (Short description)

A daily life support service, based on artificial intelligence technologies, that can explain the situation around visually impaired people while moving.

6.16.3.4 Challenges and issues

Vision

6.16.3.5 Societal concerns

Promoting welfare and supporting social activities for the blind.

SDGs to be achieved: Good health and well-being for people

6.16.4 Predictive maintenance of public housing lifts (A.94)

6.16.4.1 Scope

Build an AI solution that can predict malfunction in a lift.

6.16.4.2 Objective

Use RNN to predict possibility and type of malfunction in a lift

6.16.4.3 Narrative (Short description)

An AI model that helps the facilities management company of public housing to move from a reactive to predictive maintenance of lifts.

6.16.4.4 Challenges and issues

The model may at times predict false-positives which may lead to unnecessary deployment of repair & maintenance manpower.

6.16.4.5 Societal concerns

Disruptions to public due to breakdown of shared infrastructure.

SDGs to be achieved: Climate action

6.17 Retail

6.17.1 Emotion-sensitive AI Customer Service (A.42)

6.17.1.1 Scope

Extracting sentiment and its intensity from customers' input, and responding with appropriate attitude in order to improve the quality of customers' inquiry.

6.17.1.2 Objective

To design an efficient solution for customers' sentiment and intensity detection, especially in the situation of limited training dataset.

6.17.1.3 Narrative (Short description)

The emotion-sensitive AI customer service of JD.com Int., is supported by AI technology and deep learning method. It is developed for ameliorating accuracy of customer sentiment and intensity. In sentiment classification, it has achieved 74% accuracy and 90% recall score while in intensity detection, it has accomplished 85% accuracy and 85% recall. During the special sale of "618", it has increased customer satisfaction by 57%.

6.17.1.4 Challenges and issues

Challenge: the system's performance should be as good as the human customer server.

Issues: 1) limited training data; 2) sentiment classification among seven categories.

6.17.1.5 Societal concerns

Improving the corresponding efficiency of customer service, improving customer service experience ;

Reducing labor costs, and reducing operating costs.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.17.2 Deep Learning Based User Intent Recognition (A.43)

6.17.2.1 Scope

Recognizing users' intent to solve their problems in e-commerce fields.

6.17.2.2 Objective

To recognize and understand users' intent by AI and deep learning technologies and apply such technologies to build chat bot systems to further reduce labor cost and to be applied in various fields.

6.17.2.3 Narrative (Short description)

Intelligent customer service chat bot is mainly used to categorize users' questions, recognize users' intents and answer users' questions intelligently for different business jobs. Currently, this chat bot has been used to handle 90% of online customer service and has enabled JD.com to save over 100 million labor costs every year.

6.17.2.4 Challenges and issues

Current challenges of deep leaning and intent recognition:

— High semantic ambiguity, similar sentences can deliver different meanings.

— Unclear classification rules caused by complicated business logics

— Hard to answer reasoning questions

6.17.2.5 Societal concerns

— Solve problems intelligently to increase efficiency

— Free labors from repetitive work to save large amount of resources for the society

SDGs to be achieved: Decent work and economic growth

6.18 Security

6.18.1 Behavioural and Sentiment Analytics (A.14)

6.18.1.1 Scope

Derive emotional state and goal of person from their gestures, face, actions.

6.18.1.2 Objective

Determine if the movements, actions and general behaviour of a person is sign of malevolent intentions. Detect stealing of objects and other criminal behaviours. Prevent undesired behaviour (suicide), adapt narrative to state of person, provide dynamic content according to emotional responses.

6.18.1.3 Narrative (Short description)

None identified.

6.18.1.4 Challenges and issues

Challenges: Surveillance cameras often have low resolution, can be in poorly lit environment with bad top-down view angle. A lot of suspicious behaviour can be hidden by passer-by or large crowds. Issues: Unwanted behaviours is MUCH LESS frequent than normal behaviour and can take on various forms.

6.18.1.5 Societal concerns

Right to privacy.

SDGs to be achieved: Peace, justice and strong institutions

6.18.2 AI (Swarm Intelligence) Solution for Attack Detection in IoT Environment (A.22)

6.18.2.1 Scope

Anomaly Based Attack Detection in IoT environment using Swarm Intelligence.

6.18.2.2 Objective

Given: AMI (Advanced Metering Infrastructure – Smart Meters in Smart Buildings in Smart Cities.

Detect: Detect energy theft / meter tampering by consumer in AMI (Advanced Metering Infrastructure) or hacking attack by an external agent (man in the middle) for edge computing security scenarios with intermitted disconnection, near real-time response without using server or cloud-based analytics.

6.18.2.3 Narrative (Short description)

This is a unique approach to detect attacks in IoT environment using Anomaly Based Attack Detection using Swarm Intelligence methods. This is a key solution to detect energy theft scenario in Smart Metering. Energy Theft problem varies from 2% in developed countries to 35% in developing countries. This is complimentary to traditional AI or other static rule-based analysis which is heavily dependent on analysis of huge amounts of data on centralized cloud infrastructure. This solution is simple, nimble and can be run on low powered edge (IoT Nodes) for near real-time, low latency, low power, small compute, small storage Mist / Edge Computing Scenarios.

6.18.2.4 Challenges and issues

The problem is challenging because

— 1. Varied data set for different scenarios - large amount of data needs to be pre-processed to arrive at operation threshold parameters to be used for detection in real-time.

— 2. IoT (Edge) Nodes Configuration to suite specific environments The Swarm Intelligence System (SIS) involves a swarm of devices. It should be possible to easily configure the entire swarm for different network environments and locations.

— Solution: Many reusable modules for Logging, Debugging and configuration through XML has been developed which has enabled binary re-use without having to change any code to suit a new network environment.

— 3. Flexible to reuse / customize solution for different use-cases / scenarios and scalability

— The platform needs to be able to provide facilities for different algorithms for anomaly detection to be plugged in with minimum modification, recoding, recompilation.

— Solution: Completely dynamically pluggable Algorithm binaries can be developed that conforms to defined interface Specifications, which gives flexibility to try out new algorithms, without needing to change existing code or re-compile. Use of Swarm Intelligence ensures very less localized communication that is required. Furthermore, the Swarm Intelligence System communication capability also addresses throttling of network traffic because of multi-threading / queuing capability built in.

6.18.2.5 Societal concerns

Accuracy of Solution. Fraud (Anomaly Detection) usually incurs a false positive alarm issue.

SDGs to be achieved: Responsible consumption and production

6.18.3 Use of robotic solution for traffic policing and control (A.25)

6.18.3.1 Scope

Robotics based traffic policing system.

6.18.3.2 Objective

Efficient traffic control through use of Humanoid robots for traffic control.

6.18.3.3 Narrative (Short description)

Creation of a humanoid robot which can be deployed for traffic monitoring and control on roads. The solution will use computer vision and will be enabled with IOT for centralized control and data collection. This will relieve the human police from working in polluted environment.

6.18.3.4 Challenges and issues

The problem is challenging because accurate control instructions is crucial for proper traffic control.

6.18.3.5 Societal concerns

Addresses the pressing concern of effective traffic control.

SDGs to be achieved: Sustainable cities and communities

6.18.4 Robotic Solution for Replacing Human Labour in Hazardous Condition (A.26)

6.18.4.1 Scope

Building an AI based robotics solution for replacing Human Labour in Hazardous condition.

6.18.4.2 Objective

Offer AI based robotic solution which can be customized to work in different kind of Hazardous work environment such as Mines, Blast Furnaces, Boilers etc.

6.18.4.3 Narrative (Short description)

Building an AI based robotic solution enabled with computer vision and equipped with various sensors such as temperature, pressure, smoke detector etc which can effectively replace human labour in risky work environment.

6.18.4.4 Challenges and issues

The problem is challenging because

— Solution should be customizable for different work environments.

6.18.4.5 Societal concerns

Addresses the issue of accidents in Hazardous work environment.

SDGs to be achieved: Decent work and economic growth

6.18.5 Non-intrusive detection of malware (A.93)

6.18.5.1 Scope

Build an AI solution that detects malware activities.

6.18.5.2 Objective

User ML to flag out activities induced by malware without access to personal data on local devices.

6.18.5.3 Narrative (Short description)

A machine learning model that interprets phone activities like use of battery, data, location services or microphone to flag out possible malware in a local mobile device.

6.18.5.4 Challenges and issues

The model has limitations of the malware attacks are highly sophisticated and not easily detectable.

6.18.5.5 Societal concerns

Disparate non-institutional sources of cyber attacks.

SDGs to be achieved: Sustainable cities and communities

6.19 Social infrastructure

6.19.1 Deep Learning Technology Combined with Topological Data Analysis Successfully Estimates Degree of Internal Damage to Bridge Infrastructure (A.8)

6.19.1.1 Scope

Estimate and detect the risk of the catastrophic collapses of old bridges.

6.19.1.2 Objective

Enables estimation of failure, state of degradation with surface-mounted sensors.

6.19.1.3 Narrative (Short description)

Development of sensor data analysis technology that can aggregate vibration data from sensors attached to the surface of a bridge, and then estimate the degree of the bridge's internal damage.

6.19.1.4 Challenges and issues

Challenges: Detecting the occurrence of internal stress using this technology allows for the estimation of damage in its earliest stages, and can contribute to early countermeasures.

Issues: Conduct trials using vibration data from actual bridges, with the goal of real-world usage.

6.19.1.5 Societal concerns

None identified.

6.19.2 Water Crystal Mapping (A.77)

6.19.2.1 Scope

Increase citizen awareness on the quality of water

6.19.2.2 Objective

Map of the similarity of water crystals

6.19.2.3 Narrative (Short description)

Deep learning-based approach to automatically classify water crystals.

6.19.2.4 Challenges and issues

Water quality, ice memory

6.19.2.5 Societal concerns

Sustainable Development Goal 6 - UN Sustainable Development (water)

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.19.3 System for Real-Time Earthquake Simulation with Data Assimilation (A.97)

6.19.3.1 Scope

This system provides accurate information for evacuation in earthquake disaster.

6.19.3.2 Objective

The system conducts large-scale simulation of 3D Seismic Wave Propagation, and results are improved based on real-time data assimilation using observation and machine-learning.

6.19.3.3 Narrative (Short description)

This system provides accurate information for evacuation in earthquake disaster. The system integrates Simulation, Data Analytics and Learning (S+D+L) on the BDEC System with h3-Open-BDEC which will be introduced at the University of Tokyo in April 2021. It conducts large-scale simulation of 3D Seismic Wave Propagation, and results are improved based on real-time data assimilation using observation and machine-learning. Observations of seismic activities at more than 2,000 points in Japan are obtained by JDXnet developed by ERI/U.Tokyo through SINET in real-time manner. Construction of the detailed and accurate underground model is crucial for accurate simulations. Optimized underground model is also constructed by integration of (S+D+L). The BDEC system is 40+PF heterogeneous supercomputer system which includes Simulation Nodes for S, Data/Learning Nodes for D and L, and Integration Nodes. h3-Open-BDEC is a software infrastructure for application development towards integration of (S+D+L) supported by the Japanese Government (JSPS KAKENHI Kiban-S).

6.19.3.4 Challenges and issues

— Construction of reasonable and realistic underground model for simulation.

— Real-time earthquake simulation with data assimilation.

6.19.3.5 Societal concerns

Earthquake Disasters

SDGs to be achieved: Sustainable cities and communities

6.20 Transportation

6.20.1 AI Components for Vehicle Platooning on Public Roads (A.9)

6.20.1.1 Scope

Trains of vehicles that drive very close to each other at nearly equal speed (platoons) on public roads, in particular platooning trucks on motorways.

6.20.1.2 Objective

The objectives of truck automation are energy saving and enhanced transportation capacity by platooning, and eventually possible reduction of personnel cost by unmanned operation of following vehicles. In a variant of this concept, platoons of passenger cars follow a truck autonomously.

6.20.1.3 Narrative (Short description)

The overall concept of automated platooning is that the lead vehicle will be driven as normal by a trained (professional) driver, and the following vehicles will be driven fully automatically by the system, allowing the drivers to perform tasks other than driving their vehicles. The EU roadmap for truck platooning (EU project ENSEMBLE) envisions market introduction of multi-brand platooning by 2025. Several pilot projects have been carried out since about the year 2000. While a few AI components are already used in the pilot projects (e.g. lane keeping), future products are likely to incorporate AI solutions on several functional levels.

6.20.1.4 Challenges and issues

Highly unpredictable traffic environment, legislative situation, standardisation, stress and comfort of human drivers involved

6.20.1.5 Societal concerns

Stress or boredom for the drivers, Big Brother and constant monitoring, Safety, system security, and reliability, Risk of hacking and hijacking a long-haul freight truck poses great danger, Trust over system reliability when driving next to a computer-controlled platoon.

6.20.2 Self-Driving Aircraft Towing Vehicle (A.10)

6.20.2.1 Scope

Self-Driving towing vehicle for aircrafts, operating on an airfield autonomously.

6.20.2.2 Objective

A towing vehicle that will, on command, autonomously navigate to an assigned aircraft, attach itself, tow the aircraft to an assigned location (a runway for departures, a gate for arrivals), autonomously detach itself, and navigate to an assigned location, either a staging area or to service another aircraft.

6.20.2.3 Narrative (Short description)

Self-driving vehicle technology is applied to the problem of towing aircraft at busy airports from gate to runway and runway to gate. Autonomous aircraft towing can be supervised by human ramp controllers, by air traffic controllers (ATC), by pilots, or by ground crew. The controllers provide route information to the tugs, assisted by an automated route planning system. The planning system and tower and ground controllers work in conjunction with the tugs to make tactical decisions during operations to ensure safe and effective taxiing in a highly dynamic environment.

6.20.2.4 Challenges and issues

Safe operations in the airfield environment, minimal changes to the airport infrastructure, minimal impact of their incorporation into normal operations

6.20.2.5 Societal concerns

If labor replacements are involved, then the use of autonomy must provide an equivalent or greater benefit to some portion of the labor pool to offset the potential job loss; furthermore, they must operate in a way that feels common and familiar to humans, and must be perceived as completely safe, simple and non-intimidating.

6.20.3 Unmanned Protective Vehicle for Road Works on Motorways (A.11)

6.20.3.1 Scope

Unmanned operation of a protective vehicle in order to reduce the risk for road workers in short-time and mobile road works carried out in moving traffic.

6.20.3.2 Objective

A vehicle that is able to follow mobile road works automatically on the hard shoulder of a German motorway.

6.20.3.3 Narrative (Short description)

Mobile road works on the hard shoulder of German highways bear an increased accident risk for the crew of the protective vehicle safeguarding road works against moving traffic. The "Automated Unmanned Protective Vehicle for Highway Hard Shoulder Road Works” aims at the unmanned operation of the protective vehicle in order to reduce this risk. The vehicle has first been tested in a real operation on the German autobahn A3 in June 2018 [4]. It is actually the very first unmanned operation of a vehicle on German roads in public traffic. The scientific challenges of the project are strongly related to the general challenges in the field of automated driving.

6.20.3.4 Challenges and issues

Safe operations in public traffic, compliance with ISO 26262.

6.20.3.5 Societal concerns

None identified.

6.20.4 Enhancing traffic management efficiency and infraction detection accuracy with AI technologies (A.29)

6.20.4.1 Scope

Utilizing AI technologies in traffic monitoring and management.

6.20.4.2 Objective

To increase the accuracy and efficiency of infraction detection, traffic monitoring and flow analysis, while minimizing the human effort and the overall solution cost.

6.20.4.3 Narrative (Short description)

Big data enabled AI technologies are applied to monitoring and managing the traffic in a large municipality in China. Multi-sourced data (traffic flow, vehicle data, pedestrian movement, etc.) is monitored, from which illegal operation of vehicles, unexpected incidents, surge of traffic etc. are detected and analysed with machine learning (ML) methods. ML tasks (including training and deployment) are carried out on a platform supporting the integration of various ML frameworks, models and algorithms. The platform is based on heterogeneous computing resources. The efficiency and accuracy of infraction detection, and the effectiveness of traffic management are significantly improved, with much reduced human effort and overall solution cost.

6.20.4.4 Challenges and issues

— Constant improvement in hardware architecture to increase the performance and efficiency of running ML/DL tasks.

— Consistent interfaces between applications, ML engines and heterogeneous resource pools.

— Support of new models and emerging algorithms for growing functionalities.

6.20.4.5 Societal concerns

AI's application in urban transportation significantly improves the quality of life for urban citizens, reduces the time wasted in heavy traffic and the air pollution from vehicles.

SDGs to be achieved: Sustainable cities and communities

6.20.5 AI Solution for Traffic Signal Optimization based on Multi-source Data Fusion (A.49)

6.20.5.1 Scope

Generate traffic signal timing plans by analyzing traffic flow status and patterns based on fusing internet data, induction coils data and video data, and control the traffic signal with the generated timing plans in a real-time, self-adaptive and cooperative way.

6.20.5.2 Objective

To find an effective and efficient solution to improve the road utilization efficiency by increasing traffic flow speed and reducing traffic flow waiting time.

6.20.5.3 Narrative (Short description)

An AI solution was developed that could recognize real-time traffic flow status and abstract traffic flow patterns by fusing internet data, induction coils data and video data, and could generate optimized traffic signal timing plan by self-adaptively responding to real-time traffic flow fluctuation and with regards to traffic flow coordination among multiple intersections within a given region.

6.20.5.4 Challenges and issues

Challenges: Traffic signal self-adaptive and coordinated control for a large number of intersections. Issues: 1. Not all intersections are equipped with detectors such as induction coil or video. 2. The detectors may output abnormal values which need data clean processing.

6.20.5.5 Societal concerns

Relieve urban road congestion.

SDGs to be achieved: Sustainable cities and communities

6.20.6 Automated Travel Pattern Recognition using Mobile Network Data for Applications to Mobility as a Service (A.52)

6.20.6.1 Scope

Detect automatically travel pattern recognition from anonymized and aggregated Mobile phone Network Data.

6.20.6.2 Objective

Phase 1: Attribute trip purpose and mode of transport to multimodal door-to-door journeys from Mobile phone Network Dataset using AI and machine learning techniques (Activity based model)

Phase 2: Generate daily activities for static agents in the Agent Based Model

Phase 3: Optimisation of New Mobility services in integration with mass transit

6.20.6.3 Narrative (Short description)

Activity- based modelling has the capability to exploit big data source generated by smart cities to create a digital twin of urban environments to test Mobility as a Service schemes. MND data have been used to create activities for an Agent Based Model.

AI is used to automatically detect purpose and mode of transport in multimodal round trips, obtained by anonymized and aggregated MND trip-chains dataset. Data fusion techniques and SQL queries were also used to consider land use and facilities in the urban area of interest.

6.20.6.4 Challenges and issues

The use of Mobile Phone Network data is still not precise for shorter trips and internal trips which might be not detected. However, with the introduction of 5G, MND will be even more reliable and available to use in transport modelling.

6.20.6.5 Societal concerns

The use of anonymization techniques minimise the risk of disclosing personal information when analyzing location based data and Mobile phone Network Data.

6.20.7 Autonomous Trains (Unattended Train Operation (UTO)) (A.113)

6.20.7.1 Scope

Freight and passenger trains operate autonomously, excluding any crew presence on board, but with remote operator attention involved (GoA 4).

6.20.7.2 Objective

The critical objective of automation in trains is to provide extra reliability, safety and to prevent accidents on railways, which tend to be caused by human error. Moreover, the provided innovation leads to energy consumption optimization, transport capacity increases, and, eventually, possible reduction of personnel costs due to the autonomous operation.

6.20.7.3 Narrative (Short description)

Regarding passenger transportation, UTO enables unattended operation of trains according to schedule. The system is responsible for the train's acceleration, braking, speed control, station departure, doors opening and closing, obstacle detection, management of hazardous conditions, and emergency situations.

Autonomous trains obtain data from sensors (internal - GPS, various types of cameras, LIDARs, RADARs) and traffic control systems (train schedule, movement authority), in order to interact with passengers, other vehicles, and obstacles based on information about the environment.

6.20.7.4 Challenges and issues

None identified.

6.20.7.5 Societal concerns

Safety, reliability, security, (potential) job loss.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.21 Work & life

6.21.1 Robotic Prehension of Objects (A.16)

6.21.1.1 Scope

Outputting end effector velocity & rotation vector in response to view from RGB-D camera located on robot wrist.

6.21.1.2 Objective

Use reinforcement learning to train the robot to grasp misc. objects in simulation and transfer this learning to real-life robots.

6.21.1.3 Narrative (Short description)

It may be difficult and time-consuming for clients of assistive robotic arms to control them with the fine degree required for grasping household objects (such as in the context of having a meal). In order to improve their quality of life, we propose a method by which users can select the bounding box around the object they wish grasped, and the robot performs the grasping action. We use methods from reinforcement learning to train first in simulation, in order to reduce total training time and potential robot breakage, and then transfer this learning to real-life.

6.21.1.4 Challenges and issues

Challenges: The camera cannot have a bird's eye view and will instead move with the robot. Sparse rewards may complicate learning. Environment may be cluttered, occlusions of target might occur, objects may move around Issues: For safety reasons, speed and force of robot need to be limited in assistive environment to avoid harm. Human intervention can happen at any time.

6.21.1.5 Societal concerns

Prevent arm to people and animals near robot when it is performing a grasping task

SDGs to be achieved: Good health and well-being for people

6.21.2 Robotic Vision – Scene Awareness (A.17)

6.21.2.1 Scope

Determining in which environment the robot is and which actions are available to it.

6.21.2.2 Objective

Robustly identify the scene from video and depth sensors. From the scene and the seen objects, propose the actions to make to human collaborator .

6.21.2.3 Narrative (Short description)

Household robots need to navigate a very diverse set of environments and be able to accomplish different tasks depending on their position and action set. To meet these goals, the robots need to quickly and accurately identify the visual context in which they operate and derive the set of possible actions from this context. They can then propose relevant actions to the end user so that he does not have to define context himself and then sift through a long list of irrelevant actions.

6.21.2.4 Challenges and issues

Challenges: Environment can be poorly lit leading to difficult context recognition. Issue: Sensors degradation can occur.

6.21.2.5 Societal concerns

Privacy concerns (what data from sensors is kept, reviewed and used to improve models).

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.21.3 Recommendation Algorithm for Improving Member Experience and Discoverability of Resorts in the Booking Portal of a Hotel Chain (A.28)

6.21.3.1 Scope

Building a personalized recommendation algorithm to help members of the hotel chain to find their desirable hotel for the family holiday.

6.21.3.2 Objective

Offering personalized recommendations by understanding the member preferences from past holiday patterns and searches in the booking portal. Various member and hotel features were also considered for the model.

6.21.3.3 Narrative (Short description)

Refining existing system and implement a new model that can give personalized recommendations to members and improve bookings at the undiscoverable or not-so-popular hotels. The algorithm would help in reshaping the demand and increase the visibility of the hotels which are at the lower spectrum of demand.

We would include member and resort features along with interaction data like members visiting a hotel, and giving a rating to a resort visit etc.

6.21.3.4 Challenges and issues

— Cold Start Problem: Since the member has only visited certain hotels in the past, the interaction matrix is very sparse.

— The matrix computation at times is computational resource intensive causing system failures.

6.21.3.5 Societal concerns

We don’t see any societal concerns if it is used.

6.21.4 CRWB Recommendation benchmark (A.75)

6.21.4.1 Scope

Cooking recipe execution plan decision support and nutrition recommendation

6.21.4.2 Objective

Machine Data understandable

6.21.4.3 Narrative (Short description)

Recommendation benchmark based on a cooking recipe dataset of cooking recipe execution plans

6.21.4.4 Challenges and issues

Personal expectation related to flavor, taste and texture

6.21.4.5 Societal concerns

Local Production for Local Consumption

SDGs to be achieved: Responsible consumption and production

6.21.5 Improving the quality of online interaction (A.88)

6.21.5.1 Scope

Build an AI solution to recommend relevant ideas to users in a chat interface.

6.21.5.2 Objective

To improve the quality of conversations and translating online chat to meet ups.

6.21.5.3 Narrative (Short description)

A recommendation engine operating live in a chat interface to help both users decide on the next steps they can take of high interest to both.

6.21.5.4 Challenges and issues

Translating sociological theories, customized to Singapore's context, and translating then into data labelling for the first step of NLU.

6.21.5.5 Societal concerns

Improper use of online engagements that compromise on the culture of mutual respect and dignity.

SDGs to be achieved: Good health and well-being for people

6.22 Others

6.22.1 AI Solution to Identify Automatically False Positives from a Specific Check for "Untranslated Target Segments" from an Automated Quality Assurance Tool (A.13)

6.22.1.1 Scope

The scope of this use case is limited to automated linguistic quality assurance tools, but the outcome of this use case could be applicable to other areas, such as for example: Machine Translation, automated post-editing, Computer Aided Translation Analysis and pre-translation, etc.

6.22.1.2 Objective

To reduce the number of false positive issues for check for untranslated target segment for bilingual content with in-house automated quality assurance tool.

6.22.1.3 Narrative (Short description)

In the future, we aim to build an AI solution that could automatically identify likely false positives issues from the results of the "check for untranslated target segments" following an approach where we could use machine learning based on already identified false positives by our users.

The expected outcome would be to increase end user's productivity when reviewing automated quality assurance findings and to change user behaviour to pay more attention to this type of issues by reducing the number of false positives in 80%. In addition, we would like to reduce the amount of time, we spent on a yearly basis on refining this check manually based on users' feedback.

6.22.1.4 Challenges and issues

Challenges: Try to achieve eventually 80% of the accuracy of linguists when identifying false positives for untranslated target segments, preventing as much as possible false negatives.

Issues: segmentation of false positive data by Customer and Product profile could be challenging.

6.22.1.5 Societal concerns

None identified.

6.22.2 AI Solution for Car Damage Classification (A.18)

6.22.2.1 Scope

Car damage classification for common damage types such as bumper dent, door dent, glass shatter, head lamp broken, tail lamp broken, scratch and smash.

6.22.2.2 Objective

- To create an automated system for car damage classification using CNNs.
- Experiment using transfer and ensemble learning to find which is better for training a CNN for car damage classification.

6.22.2.3 Narrative (Short description)

Image based vehicle insurance processing is an important area with large scope for automation. We have considered the problem of Car damage classification. We explore deep learning based techniques for this purpose. Initially, we try directly training a CNN. However, due to small set of labeled data, it does not work well. Then, we explore the effect of domain-specific pre-training followed by fine-tuning. Finally, we experiment with transfer learning and ensemble learning. Experimental results show that transfer learning works better than domain specific fine-tuning. We achieve accuracy of 89.5% with combination of transfer and ensemble learning. We hosted the trained model on cloud that can be plugged into applications using API and can be used for automated first level assessment of the damage, in car insurance sector.

6.22.2.4 Challenges and issues

- Small size of the damages
- Less Quantity of data
- Ambiguity in damaged and non-damaged images

6.22.2.5 Societal concerns

Insurance agents may need to be re-skilled

SDGs to be achieved: Decent work and economic growth

6.22.3 Causality-based Thermal Prediction for Data Center (A.35)

6.22.3.1 Scope

Data center cooling control involving use of air cooling to control hot spots in data center.

6.22.3.2 Objective

Minimize energy usage in managing data center.

6.22.3.3 Narrative (Short description)

Data centers tend to be overcooled to prevent computing machines from failing due to heat. A reliable fine-grained control that could regulate air control unit (ACU) supply air temperature or flow is needed to avoid overcooling. Methods that are based on correlation-based techniques do not generalize well. Hence, we seek to uncover the causal relationship between ACUs supplying cool air and temperature at the cabinets to prioritize which ACUs should be regulated to control a hot-spot near a cabinet.

6.22.3.4 Challenges and issues

Data sufficiency.

6.22.3.5 Societal concerns

Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.22.4 Machine Learning Tools in Support of Transformer Diagnostics (A.51)

6.22.4.1 Scope

Power Transformers operation and maintenance

6.22.4.2 Objective

Use of Machine Learning (ML) algorithms as supporting tools for the automatic classification of power transformers operating condition

6.22.4.3 Narrative (Short description)

The successful use of ML tools may find multiple applications in the industry such as providing fast ways of analysing new data streaming from online sensors, evaluating the importance of individual variables in the context of transformer condition assessment and also the need or adequacy of data imputation in the so widely common problem of missing data

6.22.4.4 Challenges and issues

Data availability, missing data, imbalanced classes

6.22.4.5 Societal concerns

Safe and reliable power delivery

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.22.5 Video on Demand Publishing Intelligence Platform (A.58)

6.22.5.1 Scope

Predictive maintenance platform on a Video on Demand Content Preparation Process

6.22.5.2 Objective

The goals of the project are:

1. Process fault comprehension
2. Fault prediction
3. Fault recovery through a recommendation engine
4. Productive interaction between the fault prediction and recovery recommendation engines for a proactive process maintenance

6.22.5.3 Narrative (Short description)

An E2E platform was developed in order to achieve accurate fault prediction with Machine Learning and useful recovery action recommendation using Reinforcement Learning

6.22.5.4 Challenges and issues

The Machine Learning Engine processing time had to be very short

6.22.5.5 Societal concerns

None identified.

6.22.6 Predictive Testing (A.59)

6.22.6.1 Scope

Automatic detection of inaccurate test outcomes in an application development process

6.22.6.2 Objective

The goal of the project is the improvement of the automation level in the application testing process. This is achieved by the automatic identification of inaccurate test outcomes, reducing the number of failure alerts

6.22.6.3 Narrative (Short description)

The solution adopts machine learning to analyze event logs of test results in order to reduce the number of wrongly failed tests

6.22.6.4 Challenges and issues

Being able to manage and handle different types of data (including contextual information), integrating the solution in the processes and procedures of the company

6.22.6.5 Societal concerns

None identified.

6.22.7 Predictive Data Quality (A.60)

6.22.7.1 Scope

A solution for assessing Data Quality in data collection systems

6.22.7.2 Objective

Using machine learning techniques for identifying complex or unknown correlation among data in order to score its quality and enhance the confidence for data consumer in using data for the decision making processes

6.22.7.3 Narrative (Short description)

The solution adopt machine learning methods to analyze data collected in order to identify complex correlation on data (unknown at priori) and predict data quality issues.

6.22.7.4 Challenges and issues

Being able to manage and handle different type of data, link data to reference knowledge model, change management in the organization

6.22.7.5 Societal concerns

None identified.

6.22.8 Expansion of AI training dataset and contents using artificial intelligence techniques (A.66)

6.22.8.1 Scope

Data self-propagation and validation service for deep learning and contents services

6.22.8.2 Objective

Self-propagation of data to enhance the performance of application systems and to support the expansion of data for deep learning

Self-propagated data evaluation for qualitative verification

6.22.8.3 Narrative (Short description)

The service expands the data used for deep learning for rapid commercialization of artificial intelligence technologies. The service includes quantitative extensions of the amount of learning data for high-quality in-depth learning and qualitative verification of extended data applied to machine learning or commercial content services.

6.22.8.4 Challenges and issues

The optimized self-propagation techniques for various types of data

6.22.8.5 Societal concerns

The technology polarization in artificial intelligence technical area becomes serious more and more.

SDGs to be achieved: Industry, Innovation, and Infrastructure

6.22.9 Open spatial dataset for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data (A.122)

6.22.9.1 Scope

Analytical services for automatic detection of changes of the state of ground surface objects for administrative, government, and social purposes in different use-cases, such as:

— Urban monitoring: cadastral data, land management, estimation of the living population etc.

— Emergency mapping: estimation of disaster damages.

— Security and risk management monitoring of protected zones (powerlines, railroads, pipelines): detection of vegetation growth, control of the safety etc.

6.22.9.2 Objective

The growth of the Russian market of geo-analytical cloud-services based on remote sensing data and AI technologies; open benchmark datasets for the R&D community; and bringing the power of AI and the global coverage of remote sensing imagery closer to the people.

6.22.9.3 Narrative (Short description)

Despite the increasing number of datasets and competitions in remote sensing data science (e.g. Spacenet) there is still a lack of geographical diversity, of training classes, and of interoperability of datasets.

The proposed approach is to be extended to different types of remote sensing data and application domains based on classification of the natural and man-made objects that have a clear interpretation either in satellite or aerial imagery.

6.22.9.4 Challenges and issues

There is no standard or criteria regulated the process of labelling (manual or automatic) remote sensing (satellite, drone or UAV) images with geographic reference. Development of such a standard is vital to AI algorithms as for guarantees of the quality of training data and for testing and benchmarking.

We consider the following criteria the perfect dataset collection for EO imagery should match:

— 1) Georeference. Simply annotated photos are not enough. Maps for data labeling (e.g. Openstreetmap) require objects' coordinates.

— 2) Time series. To observe places in dynamic and calculate comparative indicators. The main application is "Emergency Mapping" where the detection of changes in residential infrastructure analysis of before and post-event images is required.

— 3) Cartographic styled labeling and classification. Maps make an abstracted interpretation of Earth observation images; we therefore, believe that the previous approach of labeling images with boxes does not satisfied the criteria for accurate image segmentation and won't work. For neural networks it's now necessary to compete with manual mapping and to calculate its accuracy we need at least some Ground Truth that looks like a map.

At the same time there are many other sources beyond the EO imagery that might be useful for mapping, such as POI, collecting field works in order to accumulate addresses. At this moment our goal is to compare ML methods with the information that could be extracted by a cartographer using only optical bands of imagery and some GIS software. For such purposes we proposed the basic classifier that is at the part of training and testing datasets.

— 4) Multispectral. Next, we assume to extend this approach to advanced classification which is comparable to thematic interpretation of satellite imagery with the help of different bands combination. That's why the proposed classifier includes classes which require even more specific training and non-optical bands for better recognition.

Providing Open API and web tools to access and preview datasets. Despite the dataset collection representing structured data, it would be much more capable for further and updated use based on the standards for interoperability of geodata. In our work, we tried to join both mapping and data science approaches in a way we see new tools and services demanded by users. For many users from the data science community, maps and remote sensing are becoming just one of the sources of information that must be structured and classified. And for many mappers that are involved in the process of geodata interpretation and classification, the map itself is the perfect tool to interact with the data; no matter whether implemented in python notebook or loaded in a desktop GIS application.

6.22.9.5 Societal concerns

Global extension of this technology brings society new possibilities of situational awareness and digital instruments for natural and man-made resource management

SDGs to be achieved: Sustainable cities and communities

Annex A
(informative)

Collected use cases

A.1 Explainable Artificial Intelligence for Genomic Medicine

A.1.1 General

ID	1	
Use case name	Explainable artificial intelligence for Genomic Medicine	
Application domain	Healthcare	
Deployment Model	Cloud services	
Status	Prototype	
Scope	To explain reason and basis behind AI-generated findings in genomic medicine	
Objective(s)	To improve the efficiency of investigatory work for experts in genomic medicine.	
Narrative	Short description (not more than 150 words)	This technology was deployed to improve the efficiency of investigatory work for experts in genomic medicine, utilizing training data and a knowledge graph that made use of public databases and medical literature databases in the field of bioinformatics. It was then evaluated to validate that it was possible to find and link the basis supporting findings with regard to phenomena whose interrelationships are only partially understood.
	Complete description	Deep Learning is one of the most representative technologies in recent AI and shows high performance in pattern recognition and analysis. However, as it cannot explain the reasons for its judgment, it is called "black box AI." There is a graph-structured data based machine learning technology called "Deep Tensor" that can directly analyze the relations among numerous pieces of real-world data ranging from intercompany transactions to material structures. Additionally, there is also a technology for building a large-scale knowledge base, which is called a "knowledge graph" and consists of vast knowledge existing around the world such as academic papers, by using our unique technology. This technology identifies the factors (partial graphs) that had a significant influence on an inference and coordinates these with partial graphs from a knowledge graph, building a series of pieces of information in the form of connections in the knowledge graph as the basis for the findings.

		People can combine these two technologies and develop a system that enables AI to explain the reasons and basis (evidence) for its judgment. A use case of applying this explainable AI is genomic medicine (for cancer treatment). The latest genomic medicine helps detect patients' genetic defects that have caused disease (cancer) and uses therapeutic drugs that affect cancer cells produced by such genetic defects. In genomic medicine today, a patient's normal and cancerous cells are analyzed with a next-generation sequencer; then, a medical team uses the obtained genetic data to identify a causal gene and determines the recommended treatment. It takes at least two weeks for the medical team to conduct an examination after completing genetic analysis. Unless the cost and time problems are solved, spreading this advantageous genomic medicine far and wide will be difficult. In this use case, the explainable AI trained Deep Tensor using 180,000 pieces of disease mutation data, successfully embedding more than 10 billion pieces of knowledge from 17 million medical articles and other materials into Knowledge Graph. Inputting genetic mutation data into this system enables Deep Tensor to infer disease-causing factors and enables Knowledge Graph to find medical evidence to justify the obtained results. Medical specialists then simply need to review the flow of obtained inference logic, thereby reducing the period between analysis and report submission significantly— from two weeks to a single day.		
Stakeholders	Doctors of genomic medicine, researchers of genomic medicine, patients			
Stakeholders' assets, values	Reducing the determination periods, maintaining the accuracy of predication as well as manual predication			
System's threats and vulnerabilities	Update knowledge graph lately, huge size of knowledge graph			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy of predication	Proportion of the true positives and true negatives combined in the disease predication by AI	Improve accuracy
	2	Appropriateness of explanation	Proportion of the appropriate flow of obtained inference logic	Improve efficiency
	3	Determination periods	The periods that a medical team uses the obtained genetic data to identify a causal	Improve efficiency

			gene and determines the recommended treatment.	
AI features	Task(s)	Knowledge processing & discovery, Natural Language Processing, Inference, Prediction		
	Method(s)	Knowledge Graph, Deep Learning (Deep Tensor), Natural Language Processing		
	Hardware			
	Topology			
	Terms and concepts used	Knowledge Graph, Deep Learning, Natural Language Processing, Explainable AI		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: To reduce experts' workloads, shortening determination periods in genomic medicine. Issues: The inability to explain the reason behind inferences from the learning algorithm of black-box AI.			
Societal concerns	Description	1, Accountability for using AI in medical examination 2, Incorrect explanation will cause the determination periods increasing.		
	SDGs to be achieved	Good health and well-being for people		

A.1.2 Data

Data characteristics	
Description	Knowledge Graph
Source	Disease mutation data, medical articles and other materials
Type	Graph-structured data in RDF format
Volume (size)	180,000 pieces of disease mutation data, more than 10 billion pieces of knowledge from 17 million medical articles
Velocity (e.g. real time)	Batch
Variety (multiple datasets)	multiple datasets
Variability (rate of change)	Static
Quality	High

A.1.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep tensor) with training data set	Disease mutation data for training is ready	To extract disease mutation data from knowledge graph	

2	Evaluation	Evaluate whether the trained model(deep tensor) can be deployed	Completion of training		Meeting accuracy requirement of predication (e.g. accuracy of predication is 90% or more) is the "success" condition
3	Execution	1, Enables Deep Tensor to infer disease-causing factors 2, Enables Knowledge Graph to find medical evidence to justify the obtained results.	The genetic mutation data is ready	To extract mutation data from knowledge graph	

A.1.4 **Training**

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Disease mutation data for training is ready	Extract training diseases mutation data	Doctors or researchers pf genomic medicine	Extract mutation data from knowledge graph	The software for processing RDF data base has to be provided by the AI solution provider
2	Completion of Step 1	Model training	AI solution provider	Train a model (deep tensor) with the training data set created by Step 1	

Specification of training data	

A.1.5 **Evaluation**

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training	Extract evaluating diseases mutation data	Doctors or researchers pf genomic medicine	Extract diseases mutation data from knowledge graph	The software for processing RDF data base has to be provided by the AI

					solution provider
2	Completion of Step 1	Predication	AI solution provider	Given the mutation data from Step 1, predicate disease / non-disease using deep tensor models that were trained in the scenario of training	
3	Completion of Step 2	Evaluation	Doctors or researchers pf genomic medicine	Compare the result of Step 2 with that of human inspection	

Input of evaluation	
Output of evaluation	

3229 A.1.6 **Execution**

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	The genetic mutation data is ready	Extract genetic mutation data	Doctors or researchers pf genomic medicine	Extract the target of genetic mutation data from knowledge graph	The software for processing RDF data base has to be provided by the AI solution provider
2	Completion of Step 1	Predication	AI solution provider	Given the mutation data from Step 1, predicate disease / non-disease using deep tensor models that were trained in the scenario of training	
3	Completion of Step 2	Inference	AI solution provider	Enables Deep Tensor to infer disease-causing factors	
4	Completion of Step 3	Explanation	AI solution provider and Doctors or researchers	Enables Knowledge Graph to find medical evidence to justify	

			pf genomic medicine	the obtained results	

Input of Execution	
Output of Execution	

A.1.7 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure				Fujitsu	http://journal.jp.fujitsu.com/en/2018/01/23/02/
2	Brochure				Fujitsu	http://www.fujitsu.com/jp/group/labs/en/business/artificial-intelligence/
3	Press Release				Fujitsu	http://www.fujitsu.com/global/about/resources/news/press-releases/2017/0920-02.html
4	Journal				Nature	http://s3-service-broker-live-19ea8b98-4d41-4cb4-be4c-d68f4963b7dd.s3.amazonaws.com/uploads/ckeditor/attachments/8429/04_UK_Fujistu_AI.PDF

A.2 Revolutionizing Clinical Decision-making using Artificial Intelligence

A.2.1 General

ID	2	
Use case name	Revolutionizing clinical decision-making using artificial intelligence	
Application domain	Healthcare	
Deployment Model	On-premise systems	
Status	PoC	
Scope	To improve clinical decision-making and the accurate assessment of risks for individual patients of mental healthcare.	
Objective(s)	Halving the time to pre-screen patient records and giving more time for patient consultations	
Narrative	Short description (not more than 150 words)	The solution has halved the time for the preliminary assessment of patient records, increasing the time available for consultations
	Complete description	Traditional healthcare institutions have extensive paper archives built up over many years, representing a body of data that is often difficult to systematize, locate and interpret. The implementation of the electronic clinical history represents significant progress, facilitating analysis

<table>
<tr><td></td><td></td><td colspan="3">by providing information in an accessible and legible format with centralized access.

However, in a “post-digitization” era, the information generated on a daily basis remains underused. “We have access to a vast quantity of data but it’s hard to extract meaningful information that helps us improve the quality of the care we provide,” explains Dr. Julio Mayol Martínez, Medical Director and Director of Innovation at the San Carlos Clinical Hospital.

The solution has been developed on the back of the company’s in-depth research into applying advanced data analytics for healthcare applications. It has involved working in close collaboration with San Carlos Clinical Hospital’s expert clinicians, applying Fujitsu’s principles of co-creation to deliver tangible value in the field of mental healthcare. It deploys Fujitsu Laboratories’ state of the art anonymization technologies and Fujitsu’s data analytics technologies, tailored to meet the specific needs of the local Spanish healthcare sector. The technology will form the basis of a new Health Application Programming Interface (API), to be deployed in the Fujitsu cloud or delivered locally in a private cluster or cloud.

The field trial took place over a 6-month period, involving senior mental health clinicians from San Carlos Clinical Hospital and a core database of over 36,000 anonymized patient records. Fujitsu leveraged this database to develop its Advanced Clinical Research Information System, based on its advanced artificial intelligence expertise including data analytics and semantic modelling. In the field trial, each of the clinicians looked at issues associated with the main diagnosis, any co-morbidities, potential risks from suicide, substance or alcohol abuse, and the patient history of using the healthcare system. Fujitsu’s system demonstrated a very high degree of risk assessment accuracy, with the system accelerating and systemizing the verification of key clinical data and identification of existing clinical problems. It achieved results of over 85 percent to identify suicide, alcohol and drug abuse risk.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4"></td></tr>
<tr><td>System’s threats and vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">AI features</td><td>Task(s)</td><td colspan="3">Natural language processing</td></tr>
<tr><td>Method(s)</td><td colspan="3">Knowledge Graph</td></tr>
</table>

	Hardware	
	Topology	
	Terms and concepts used	
Standardization opportunities/ requirements		
Challenges and issues	The incorporation of many different types of data is revolutionizing the healthcare sector. The ability to apply semantic and analytic technologies to this heterogeneous mass of data, as well as traditional healthcare data, to discover hidden correlations, identify care patterns and support clinical decision-making is paving the way for a new generation of improved healthcare services	
Societal concerns	Description	Incorrect decision Unexplainable result
	SDGs to be achieved	Good health and well-being for people

A.2.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure				Fujitsu	http://www.fujitsu.com/global/Images/CS_2017Apr_IdISSC_San-Carlos-Hospital_Eng_v.1.pdf
2	Brochure				Fujitsu	http://www.fujitsu.com/global/microsite/vision/customerstories/hospital-clinico-san-carlos/
3	Press Release				Fujitsu	http://www.fujitsu.com/uk/about/resources/news/press-releases/2015/pr-fle20161110.html

A.3 AI Solution to Calculate Amount of Contained Material from Mass Spectrometry Measurement Data

A.3.1 General

ID	3
Use case name	AI solution to calculate amount of contained material from mass spectrometry measurement data
Application domain	Manufacturing
Deployment model	Embedded systems
Status	PoC
Scope	Calculating amount of contained material from mass spectrometry measurement data using chromatography

Objective(s)	To find an accurate and efficient solution to calculating amount of contained material without dependence on individuals	
Narrative	Short description (not more than 150 words)	An AI solution was developed that could automatically pick the peak related to the contained material from measurement data through deep learning. Compared with manual results by an experienced operator, the automated peak picking results using AI had a false detection rate of 7% and an undetected rate of 9%. The peak picking operation time using AI was estimated to be about one fifth.
	Complete description	The technology was developed that utilizes AI (artificial intelligence) to process the vast amounts of data used in analyzing the measurement results, which are essential to analytical processes, acquired from mass spectrometers. Mass spectrometers are used for research and quality control in various areas such as the establishment of early detection techniques for diseases and the measurement of residual pesticides in foods, and because of improvements in sensitivity and speed, the amount of data acquired is enormous. As a result, the data analysis step called "peak picking" has become the bottleneck in the workflow. Complete automation is difficult and to some extent manual adjustments are required. Therefore, there are differences in analysis accuracy depending on each operator and there is a possibility that analytical results might be affected by each operator's practices and data alterations. In recent years, automated data analysis with high accuracy that eliminates this kind of dependence on individuals is now demanded in the fields of healthcare and new drug development. To solve this issue using AI, the three companies investigated the application of deep learning, a neural network technology that imitates brain neurons. Arising to confront this process were two problems: 1) insufficient training data; and 2) learning could not proceed when analytical equipment output data was input, as is, into the deep learning network. The technologies to produce extra data to compensate for the lack of training data and to convert the analysis equipment output features into images were developed. Moreover, the companies developed the feature extraction technology to learn the analytical skills of experienced analysts. By doing this, the deep learning network was able to learn from the over 30,000 items of generated training data. Compared with manual peak picking results by an experienced operator, the automated peak picking results using AI had a false detection rate of 7% and an undetected rate of 9%. These results indicate that an automated peak picking can compare favorably with a peak picking by an experienced operator.
Stakeholders		
Stakeholders’		

assets, values				
System's threats and vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Recall	Proportion of the true positive to positive results by an experienced operator	Improve accuracy
	2	Precision	Proportion of the true positive to positive results by AI	Improve accuracy
	3	Operation time	Ratio of operation time using AI to the conventional one	Improve efficiency
AI features	Task(s)	Recognition		
	Method(s)	Deep Learning		
	Hardware			
	Topology			
	Terms and concepts used	Deep Learning, Data Augmentation		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Achieve the same level as experienced operators for peak picking. Issues: 1) Lack of training data per contained material, 2) how to create good images for deep learning from mass spectrometry measurement data			
Societal concerns	Description			
	SDGs to be achieved			

A.3.2 Data

Data characteristics	
Description	Mass spectrometry measurement data
Source	Mass spectrometry
Type	Numerical data
Volume (size)	
Velocity (e.g. real time)	Batch
Variety (multiple datasets)	Single
Variability (rate of change)	Static
Quality	High

A.3.3 Process scenario

Scenario conditions

No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training samples			
2	Evaluation	Evaluate whether the trained model can be deployed			
3	Execution	Pick peaks using the trained model and calculate the amount of contained material			
4	Retraining	Retrain a model with training samples			

A.3.4 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure				Fujitsu	http://www.fujitsu.com/global/vision/customerstories/shimadzu-corporation/index.html
2	Press Release				Fujitsu	http://www.fujitsu.com/global/about/resources/news/press-releases/2017/1113-01.html

A.4 AI Solution to Quickly Identify Defects during Quality Assurance Process on Wind Turbine Blades

A.4.1 General

ID	4
Use case name	AI solution to quickly identify defects during quality assurance process on wind turbine blades

Application domain	Manufacturing	
Deployment Model	On-premise systems	
Status	In operation	
Scope	Detecting defects in products by inspecting nondestructive testing scanning data	
Objective(s)	To find an accurate and efficient solution to detect defects without compromising the detection of in-material damage and risking a loss in reputation.	
Narrative	Short description (not more than 150 words)	An AI solution was developed that could automatically detect defects through deep learning together with what is called "imagification"; it achieved high coverage of various defects and evaluation of each nondestructive testing scanning was reduced by 80%, which translated into cost savings, reduced production lead times, and increased productivity.
	Complete description	The manufacturer produces over 5,000 wind turbine blades every year for use in on/offshore wind farms. Each blade can be up to 75 meters in length and takes a highly skilled professional quality controller up to 6 hours to evaluate the Ultrasonic Testing (UT) scanning in the quality assurance process. This is because the structure can contain multiple defect types, including how fiberglass can wrinkle during the production process. This has the potential to be catastrophic if this makes the blade crash during operation. The manufacturer must put each wind turbine blade through a stringent quality assurance process. Any defects when a blade is in operation could not only prove catastrophic but also inflict major damage to the company's reputation. Working with the AI solution provider together they co-created an AI solution that could automatically detect defects through deep learning capabilities; it achieved high coverage (more than 95%) of various defects and evaluation of each nondestructive testing scanning reduced by 80%. Another method featured in the AI solution is "imagification," which transforms raw data into image data based on RGB where deep learning-based image recognition can be applied effectively. Quality controllers can focus their efforts on suspicious areas and disregard all clean data; humans only need to examine the blades that are flagged by the AI system. With 5,000 blades produced every year, that adds up to a saving of almost 32,000 man-hours, which translates into significant cost savings, reduced production lead times, and increased productivity. Today, there is a shortage of ultrasonic engineers/inspectors. This solution means the same inspector can do 4 to 5 blades per day instead of 1 previously.
Stakeholders	Manufacturer	
Stakeholders' assets, values	Reputation	
System's threats and vulnerabilities	Changes in defects of in-material damage over time	

Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Coverage	Ratio of defects included/found in the regions of product which are "of interest" for manual inspection. Ideal target is 95%.	Improve accuracy
	2	Split	Proportion of the regions of product which are "of interest" for manual inspection. The less split, the more efficient the total quality assurance process becomes.	Improve efficiency
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware			
	Topology			
	Terms and concepts used	Deep learning, "imagification”, neural network, training, training data set		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Achieve the same level as ultrasonic accredited engineers for detecting critical defects. Issues: 1) Lack of defect data per defect type, 2) how to create good images for deep learning from UT raw data, and 3) back wall detection			
Societal concerns	Description			
	SDGs to be achieved	Affordable and clean energy		

A.4.2 Data

Data characteristics	
Description	UT scanning data
Source	UT scanning instrument
Type	Ultrasonic data from scanner vendor
Volume (size)	
Velocity (e.g. real time)	Batch
Variety (multiple datasets)	Single source
Variability	Static

(rate of change)	
Quality	High (depending on UT equipment)

3248 ### A.4.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training data set	Sample raw data set is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		Meeting KPI requirements (e.g. coverage is 95% or more, split is 20% or less) is the "success" condition
3	Execution	Detect defects (regions including defects) using the trained model	Completion of UT scanning of a blade	The trained model has been evaluated as deployable	
4	Retraining	Retrain a model with training data set	Certain period of time has passed since the last training/re training		

3249 ### A.4.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample raw data set is ready	Imagification	Manufacture r	Transform sample raw data from UT scanning to image data based on RGB	The software for imagification has to be provided by the AI solution provider.
2	Completion of Step 1	Training data set creation	Manufacture r	Create training data set by labelling the output of Step 1 with	

				"defective"/"non -defective"	
3	Completion of Step 2	Model training	AI solution provider	Train a model (deep neural network) with the training data set created by Step 2	

Specification of training data	

A.4.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Imagification	Manufacturer	Transform raw data from UT scanning for blind test to image data based on RGB	
2	Completion of Step 1	Detection	AI solution provider	Given the image data from Step 1, detect defects (regions including defects) using the deep neural network trained in the scenario of training	
3	Completion of Step 2	Evaluation	Manufacturer	Compare the result of Step 2 with that of human inspection	

Input of evaluation	
Output of evaluation	

A.4.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of UT scanning of a blade	Imagification	Manufacture r	Transform raw data from UT scanning to image data based on RGB	
2	Completion of Step 1	Detection	Manufacture r	Given the image data from Step 1,	The trained deep neural

				detect defects (regions including defects) using the trained deep neural network with the output of Step 1 as input	network has to be handed over to the manufacture r.

Input of Execution	
Output of Execution	

A.4.7 Retraining

Scenari o name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Certain period of time has passed since the last training/retraining	Imagification	Manufact urer	Transform sample raw data from UT scanning to image data based on RGB	
2	Completion of Step 1	Training data set creation	Manufact urer	Create training data set by labelling the output of Step 1 with "defective"/"non-defective"	
3	Completion of Step 2	Model training	AI solution provider	Train a model (deep neural network) with the training data set created by Step 2	

Specification of retraining data	Retraining data set has to include recent data

A.4.8 References

References						
No.	Type	Referenc e	Status	Impact on use case	Originator/o rganization	Link
1	Brochure				Fujitsu	http://www.fujitsu.com/global/vision/customerstories/siemens-gamesa/index.html
2	Press release				Fujitsu	http://www.fujitsu.com/fts/about/resources/news/press-releases/2017/emeai-20171107-artificial-intelligence-solution-from.html
3	Press release				Fujitsu	http://www.fujitsu.com/fts/about/resources/news/press-releases/2017/emeai-

						20171002-fujitsu-develops-state-of-the-art-ai.html

A.5 Solution to Detect Signs of Failures in Wind Power Generation System

A.5.1 General

<table>
<tr><td>ID</td><td colspan="4">5</td></tr>
<tr><td>Use case name</td><td colspan="4">Solution to detect signs of failures in wind power generation system</td></tr>
<tr><td>Application domain</td><td colspan="4">Manufacturing</td></tr>
<tr><td>Deployment Model</td><td colspan="4">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="4">PoC</td></tr>
<tr><td>Scope</td><td colspan="4">Detect signs of malfunction (failure) in wind power generators</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Detect signs of failure in wind power generation, earlier than human specialists</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">A system is currently in development that uses machine learning to detect signs of equipment failure that would be difficult to detect from visual inspection. Currently, sensor data is being collected from 43 actual domestic large wind turbines, and large-scale verification testing is being conducted. The goal is for a paradigm shift from responding after the fact to maintenance that prevents problems and maintenance safety</td></tr>
<tr><td>Complete description</td><td colspan="3">"We present a method for detecting anomalies in vibration signals of wind turbine components.
The predominant characteristics of wind turbine vibration signals are extracted by applying a time-frequency feature extraction method based on Fourier local autocorrelation (FLAC) features. For anomaly detection, one-class classification based on an unsupervised clustering approach is applied in consideration of the wind turbine's dynamic operating conditions and environment. To validate the proposed system, we conducted experiments using the vibration data of actual 2 MW wind turbines. The results showed the effectiveness of using the FLAC features, particularly in the case of the low-speed main bearing where the conventional method with traditional features cannot detect the anomalies. "</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats and vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="4">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Time from alert to failure</td><td></td><td></td></tr>
<tr><td>2</td><td>Precision</td><td></td><td></td></tr>
<tr><td>3</td><td>Recall</td><td></td><td></td></tr>
</table>

AI features	Task(s)	Recognition
	Method(s)	Anomaly detection based on machine learning techniques, Accurate feature extraction from vibration signals
	Hardware	
	Topology	
	Terms and concepts used	Fourier Local AutoCorrelation (FLAC) features, Unsupervised learning
Standardization opportunities/ requirements		
Challenges and issues		
Societal concerns	Description	
	SDGs to be achieved	

A.6 Computer-aided Diagnosis in Medical Imaging based on Machine Learning

A.6.1 General

ID	6	
Use case name	Computer-aided diagnosis in medical imaging based on machine learning	
Application domain	Healthcare	
Deployment Model	Hybrid or other (please specify)	
Status	PoC	
Scope	Detecting image anomality	
Objective(s)	Provide AI method to alleviate growing burden of histopathological diagnosis by human	
Narrative	Short description (not more than 150 words)	The advances in image recognition technology enable the machine learning system to support diagnosis in medical imaging. This technology is expected to contribute the great reduction of the burden on doctors and the improvement of diagnostic accuracy when it is used for screening and double checking. Specifically, a support system is currently under development that analyzes histopathological images to automatically detects suspected lesion.
	Complete description	In histopathological diagnosis, a clinical pathologist discriminates between normal tissues and cancerous tissues. However, recently, the shortage of clinical pathologists is posing increasing burdens on meeting the demands for such diagnoses, and this is becoming a serious social problem. Currently, it is necessary to develop new medical technologies to help reduce their burdens. Therefore, as a diagnostic support technology, an extended method of HLAC (Higher-order Local AutoCorrelation) feature extraction for classification of histopathological

		images into normal and anomaly. The proposed method can automatically classify cancerous images as anomaly by using an extended geometric invariant HLAC features with rotation- and reflection-invariant properties from three-level histopathological images, which are segmented into nucleus, cytoplasm and background. In conducted experiments, we demonstrate a reduction in the rate of not only false-negative errors but also of false-positive errors, where a normal image is falsely classified as an image with an anomaly that is suspected as being cancerous.		
Stakeholders				
Stakeholders' assets, values				
System's threats and vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Precision		
	2	Recall		
AI features	Task(s)	Recognition		
	Method(s)	Higher-order Local Auto-Correlation		
	Hardware			
	Topology	Higher-order Local Auto-Correlation		
	Terms and concepts used	Higher-order Local Auto-Correlation		
Standardization opportunities/ requirements				
Challenges and issues				
Societal concerns	Description			
	SDGs to be achieved			

A.7 AI Ideally Matches Children to Daycare Centers

A.7.1 General

ID	7
Use case name	AI Ideally Matches Children to Daycare Centers
Application domain	Public sector
Deployment model	On-premise systems
Status	In operation

<table>
<tr><td>Scope</td><td colspan="4">Assignment pattern that satisfies complex applicants' requirements</td></tr>
<tr><td>Objective(s)</td><td colspan="4">To determine the assignment pattern that will fulfill the preferences of as many applicants as possible automatically.</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">This AI technology automatically determines the assignment pattern while fulfilling as many applicants' preferences as possible by priority ranking by using game theory.</td></tr>
<tr><td>Complete description</td><td colspan="3">The number of children on daycare center waiting lists has become a social issue. Matching children to daycare centers while accommodating each family's preferences is time- and labor-intensive for local governments.
The basic goal of daycare admissions screening is to satisfy the preferences of applicants according to the priority ranking of children in consideration of the number of places in each daycare center. In addition, each local government can incorporate more complex requirements, such as applicants who want their siblings assigned to the same daycare center and who want siblings assigned in the same period, in order to increase the satisfaction of applicants. Saitama city government has eight requirements concerning sibling admissions as well as the timing of the siblings' admissions. The screening rule thus became more complex, and consequently there are cases where multiple assignment patterns can fulfill the rule or no patterns fulfill the rule. This means the city officials are required to take a long time to carefully determine the assignment of applicants to be absolutely sure that the relevant rules have been correctly fulfilled.
This AI technology has made it possible to match children to daycare centers, meeting as many preferences as possible, following a priority ranking. This is done by modelling the dependency relationships of complex requirements, including parents who prioritize siblings going to the same daycare center, or parents who do not mind if their children go to different daycare centers as long as both children get a seat, using a mathematical model based on game theory, which rationally resolves the relationships between people having differing values. When this technology was evaluated using anonymized data from about 8,000 children in the city of Saitama, it successfully calculated an optimal assignment result in just a few seconds.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">City officials, Daycare centers, Applicants</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Maintaining fairness of matching results, Reducing the burden of seat assignment tasks, Leading to return women to the workplace smoothly.</td></tr>
<tr><td>System's threats and vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Accuracy</td><td>The matching rate of assignment</td><td>Automatic assignment</td></tr>
<tr><td>2</td><td>Time</td><td>The computation time to find an</td><td>Time reduction</td></tr>
</table>

			optimal assignment	
AI features	Task(s)	Optimization		
	Method(s)	Game theory		
	Hardware			
	Topology			
	Terms and concepts used	Game theory, Matching theory		
Standardization opportunities/ requirements	Need to consider unique requirements for assignment rules in each local government.			
Challenges and issues	Challenges: Determine an optimal assignment pattern instantly and fairly depending on unique and complex rules in each local government. Issues: Long calculation time is required in the case of a large number of children and siblings			
Societal concerns	Description	Supporting working women Resolving the problem of children waiting for day care		
	SDGs to be achieved	Decent work and economic growth		

A.7.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press Release				Fujitsu	http://www.fujitsu.com/global/about/resources/news/press-releases/2017/0830-01.html
2	Technical Paper	Applying Matching Technology to the MICJET MISALIO Parenting Solution for Young Parents: Field Trial with Saitama City Government			FUJITSU Journal (in Japanese)	http://www.fujitsu.com/jp/documents/about/resources/publications/magazine/backnumber/vol69-4/paper04.pdf
3	Technical Paper	Matching Children to Daycare Centers			Proceedings of the Spring Forum 2018 of the Operations	http://jglobal.jst.go.jp/detail/?JGLOBAL_ID=201802223345266044

					Research Society of Japan (in Japanese)	

A.8 Deep Learning Technology Combined with Topological Data Analysis Successfully Estimates Degree of Internal Damage to Bridge Infrastructure

A.8.1 General

ID	8	
Use case name	Deep Learning Technology Combined with Topological Data Analysis Successfully Estimates Degree of Internal Damage to Bridge Infrastructure	
Application domain	Social infrastructure	
Deployment Model	Cloud services	
Status	PoC	
Scope	Estimate and detect the risk of the catastrophic collapses of old bridges	
Objective(s)	Enables estimation of failure, state of degradation with surface-mounted sensors	
Narrative	Short description (not more than 150 words)	Development of sensor data analysis technology that can aggregate vibration data from sensors attached to the surface of a bridge, and then estimate the degree of the bridge's internal damage
	Complete description	Inspection tasks for bridges are usually performed visually to check the structure for damage. The issue with relying only on information gathered visually, however, is that inspectors can only identify abnormalities or anomalies appearing on the structure's surface, and are consequently unable to grasp information regarding the degree of internal damage. There have been many trials in which sensors were attached to the surface of the bridge deck, using vibration data to evaluate the level of damage. With the methods used until now, accurately understanding the degree of damage within the interior of the deck was an issue. Deep learning AI technology for time-series data can discover anomalies and express in numerical terms degrees of change that demonstrate drastic changes in the status of objects such as structures or machinery, and detect the occurrence of abnormalities or distinctive changes. The technology learns from the geometric characteristics extracted from complex, constantly changing time-series vibration data collected by sensors equipped on IoT devices, thus enabling users to estimate and validate the state of degradation or failure in a variety of social infrastructure or machinery. This technology has now been confirmed through the application of verification test data from RAIMS (Research Association for Infrastructure Monitoring System).

<table>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats and vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Anomaly detection</td><td>The geometric characteristics extracted from the vibration data by this technology would appear as a single cluster when the bridge was intact, but the shape changes when the bridge had developed internal damage.</td><td>Enabling to detect anomalous feature</td></tr>
<tr><td>2</td><td>Change detection</td><td>The degree of abnormality and the degree of change that can be calculated by converting the geometric characteristics to numerical values correspond with the results measured by strain sensors embedded within the bridge deck.</td><td>Precise measure of anomaly</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition</td></tr>
<tr><td>Method(s)</td><td colspan="3">Topological Data Analysis</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Topological Data Analysis, Anomaly Detection, Time Series Classification, Convolutional Neural Network</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Challenges: Detecting the occurrence of internal stress using this technology allows for the estimation of damage in its earliest stages, and can contribute to early countermeasures.
Issues: Conduct trials using vibration data from actual bridges, with the goal of real-world usage.</td></tr>
<tr><td>Societal</td><td>Description</td><td colspan="3"></td></tr>
</table>

concerns		
	SDGs to be achieved	

A.8.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press Release				Fujitsu	http://www.fujitsu.com/global/about/resources/news/press-releases/2017/0828-01.html
2	Press Release				Fujitsu	http://www.fujitsu.com/global/about/resources/news/press-releases/2016/0216-01.html
3	Technical Paper	Time Series Classification via Topological Data Analysis			Transactions of the Japanese Society for Artificial Intelligence	https://www.jstage.jst.go.jp/article/tjsai/32/3/32_D-G72/_article
4	Technical Paper	Topological Data Analysis and its Application to Chronological Data Analysis			FUJITSU Journal (in Japanese)	http://www.fujitsu.com/jp/documents/about/resources/publications/magazine/backnumber/vol69-4/paper15.pdf

A.9 AI Components for Vehicle Platooning on Public Roads

A.9.1 General

ID	
Use case name	AI Components for Vehicle Platooning on Public Roads
Application domain	Transportation
Deployment model	Self-driving vehicles
Status	Prototype
Scope	Trains of vehicles that drive very close to each other at nearly equal speed (platoons) on public roads, in particular platooning trucks on motorways.
Objective(s)	The objectives of truck automation are energy saving and enhanced transportation capacity by platooning, and eventually possible reduction of personnel cost by unmanned operation of following vehicles. In a variant of this concept, platoons of passenger cars follow a truck autonomously.

Narrative	Short description (not more than 150 words)	The overall concept of automated platooning is that the lead vehicle will be driven as normal by a trained (professional) driver, and the following vehicles will be driven fully automatically by the system, allowing the drivers to perform tasks other than driving their vehicles. The EU roadmap for truck platooning (EU project ENSEMBLE) envisions market introduction of multi-brand platooning by 2025 [12]. Several pilot projects have been carried out since about the year 2000 [8,9,10,13,14]. While a few AI components are already used in the pilot projects (e.g. lane keeping), future products are likely to incorporate AI solutions on several functional levels.		
	Complete description	A major development in research on Intelligent Transportation Systems (ITS) is Cooperative Adaptive Cruise Control (CACC). It takes Adaptive Cruise Control (ACC) to the next level by adding direct communication between vehicles. Directly communicating accurate state information allows vehicles to drive much closer to each other without compromising safety. This is the basis of platooning: trains of vehicles that drive very close to each other at nearly equal speed. By CACC, platoons become string stable: changes in the acceleration or deceleration are reduced by the following vehicles instead, of getting amplified. This property is expected to greatly improve the throughput of vehicles on highways, because it is exactly the amplification of acceleration and deceleration that causes many traffic jams. R&D on truck platooning is driven partially by the potential fuel savings and the expectation of an attractive return on investment. Implementations of platooning are complex cyber-physical systems [3]. In freight transportation, for example, a typical system architecture consists of the fleet layer, the cooperation layer, and the vehicle layer. AI components are already used on the vehicle layer (e.g. lane keeping), future products are likely to incorporate AI solutions on several functional levels and all system layers. Lane keeping is an established AI technology in the automotive industry [6]. Some examples for other potential AI components in platooning systems are: • Prediction of behavior of surrounding traffic [4] • Controllers for platooning strategies [1,3] • Road surface recognition [2] • Driver state assessment [7,11] Safe control and safety regions [5]		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives

	1	Efficiency, environmental and economic benefits	• improved on-road safety • greater fuel efficiency and reduced emissions • ease of driving • increased operational efficiency • additional road capacity • reduced labor costs	see above
	2	Societal Acceptance	Safety testing, reporting, benefits analyses, and demonstrations of automated platooning are needed and should be available to the public	see above
	3	Safety	The system must be safe, secure, and reliable	
AI features	Task(s)	Lane keeping, environment perception, prediction, driver monitoring, planning and optimization		
	Method(s)	machine learning, computer vision, logical decision making, pattern recognition, multimodal event detection, multi-agent planning and scheduling, probabilistic predictive modelling, evolutionary algorithm		
	Hardware	commercial road vehicles, positioning sensors, environment sensors (radar, LIDAR, electro-optical cameras, infrared cameras), GPS, V2V communication (UMTS,4/5G, 802.11 networks)		
	Topology			
	Terms and concepts used	autonomous vehicle guidance, environment perception, self perception, planning and scheduling, optimization, human-machine interaction, cyber-physical system		
Standardization opportunities/ requirements				
Challenges and issues	highly unpredictable traffic environment, legislative situation, standardization, stress and comfort of human drivers involved			
Societal Concerns	Description	Stress or boredom for the drivers, Big Brother and constant monitoring, Safety, system security, and reliability, Risk of hacking and hijacking a long-haul freight truck poses great danger, Trust over system reliability when driving next to a computer-controlled platoon.		

	SDGs to be achieved	

A.9.2 References

[1] W. van Willigen, E. Haasdijk, and L. Kester, “A multi-objective approach to evolving platooning strategies in intelligent transportation systems,” in Proceedings of the 15th annual conference on Genetic and evolutionary computation, 2013, pp. 1397–1404.

[2] M. Aki et al., “Road Surface Recognition Using Laser Radar for Automatic Platooning,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 10, pp. 2800–2810, Oct. 2016.

[3] B. Besselink et al., “Cyber–Physical Control of Road Freight Transport,” Proceedings of the IEEE, vol. 104, no. 5, pp. 1128–1141, May 2016.

[4] I. Cara and J.-P. Paardekooper, “THE POTENTIAL OF APPLYING MACHINE LEARNING FOR PREDICTING CUT-IN BEHAVIOUR OF SURROUNDING TRAFFIC FOR TRUCK-PLATOONING SAFETY,” in 25th ESV Conference Proceedings (International Technical Conference on the Enhanced Safety of Vehicles), 2017 [Online]. Available: http://indexsmart.mirasmart.com/25esv/PDFfiles/25ESV-000292.pdf

[5] A. Fermi, M. Mongelli, M. Muselli, and E. Ferrari, “Identification of safety regions in vehicle platooning via machine learning,” in 2018 14th IEEE International Workshop on Factory Communication Systems (WFCS), 2018, pp. 1–4.

[6] J. E. Gayko, “Lane Departure and Lane Keeping,” in Handbook of Intelligent Vehicles, Springer London, 2012, pp. 689–708.

[7] T. Heffelaar, R. Landman, M. Merts, J. M. van Hemert, A. Stuiver, and L. Noldus, “Driver state estimation: from simulation to the real world,” in Proceedings of Measuring Behavior 2014, Wageningen, The Netherlands, 2014 [Online]. Available: https://www.measuringbehavior.org/files/2014/Proceedings/Heffelaar%20T%20-%20MB2014.pdf

[8] R. Janssen, H. Zwijnenberg, I. Blankers, and J. de Kruijff, “TRUCK PLATOONING - DRIVING THE FUTURE OF TRANSPORTATION,” TNO Mobility and Logistics, Delft, The Netherlands, 2015 [Online]. Available: https://www.tno.nl/en/about-tno/news/2015/3/truck-platooning-driving-the-future-of-transportation-tno-whitepaper/

[9] M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds., Autonomous Driving. Springer Berlin Heidelberg, 2016.

[10] T. Robinson and E. Coelingh, “Operating Platoons On Public Motorways: An Introduction To The SARTRE Platooning Programme,” Jul. 2018 [Online]. Available: https://www.researchgate.net/publication/268300380_Operating_Platoons_On_Public_Motorways_An_Introduction_To_The_SARTRE_Platooning_Programme

[11] N. Schoemig, A. Kaussner, H.-P. Krüger, S. Boverie, and F. Flemisch, “THE IMPORTANCE OF DRIVER STATE ASSESSMENT WITHIN HIGHLY AUTOMATED VEHICLES,” Jan. 2009 [Online]. Available: https://www.researchgate.net/publication/255627086_THE_IMPORTANCE_OF_DRIVER_STATE_ASSESSMENT_WITHIN_HIGHLY_AUTOMATED_VEHICLES

[12] K. Sjöberg, “Status of Truck Platooning in Europe,” in 9th ETSI Workshop on Intelligent Transport Systems (ITS), 2018 [Online]. Available: https://docbox.etsi.org/Workshop/2018/20180306_ITS_WORKSHOP/S04_ACCIDENT_FREE_AUTOM_DRIV/TRUCK_PLATOONING_SCANIA_SJOBERG.pdf

[13] P. Slowik and B. Sharpe, “Automation in the long haul: Challenges and opportunities of autonomous heavy duty trucking in the United States,” INTERNATIONAL COUNCIL ON CLEAN TRANSPORTATION, WWW.THEICCT.ORG, 2018 [Online]. Available: https://www.theicct.org/sites/default/files/publications/Automation_long-haul_WorkingPaper-06_20180328.pdf

[14] S. Tsugawa, S. Jeschke, and S. E. Shladover, “A Review of Truck Platooning Projects for Energy Savings,” IEEE Transactions on Intelligent Vehicles, vol. 1, no. 1, pp. 68–77, Mar. 2016.

A.10 Self-Driving Aircraft Towing Vehicle

A.10.1 General

ID	10	
Use case name	Self-Driving Aircraft Towing Vehicle	
Application domain	Transportation	
Deployment model	Self-driving vehicles	
Status	Prototype	
Scope	Self-Driving towing vehicle for aircrafts, operating on an airfield autonomously.	
Objective(s)	A towing vehicle that will, on command, autonomously navigate to an assigned aircraft, attach itself, tow the aircraft to an assigned location (a runway for departures, a gate for arrivals), autonomously detach itself, and navigate to an assigned location, either a staging area or to service another aircraft.	
Narrative	Short description (not more than 150 words)	Self-driving vehicle technology is applied to the problem of towing aircraft at busy airports from gate to runway and runway to gate. Autonomous aircraft towing can be supervised by human ramp controllers, by air traffic controllers (ATC), by pilots, or by ground crew. The controllers provide route information to the tugs, assisted by an automated route planning system. The planning system and tower and ground controllers work in conjunction with the tugs to make tactical decisions during operations to ensure safe and effective taxiing in a highly dynamic environment.
	Complete description	Advances in self-driving automobiles make it technologically feasible to apply this technology for the purpose of taxiing planes to the runway from the terminal gate and vice-versa. Deploying self-driving vehicles for this purpose offers fewer technical challenges than deploying them on roadways and highways. Routes between gates to runways and runways to gates are typically pre-determined, with little or no possibility for alternatives. In addition, to ensure safety, constraints on taxiing operations are rigid and unambiguous. Rules such as separation constraints between taxiing aircraft and those governing right-of-way at intersection points are clearly documented and enforced by ramp and ATC controllers. These rules and procedures reduce the overall uncertainty in the operational environment and

		therefore potentially simplify the models that need to be employed by self-driving vehicles. Nominal autonomous operation of the towing vehicle (tug) is captured as the following sequence (for the case of departures): a tug sits at a tug depot, a designated area of the airport surface where tugs recharge and return when not in service. When the tug receives a message, describing time, route, and gate, it travels to the specified gate following the provided route. As the tug approaches the specified gate, it navigates to a designated ready position. Once the ground marshal attending the gate signals readiness for attachment, the tug assesses the environment to verify the surroundings are obstacle-free before moving to dock with the aircraft. Once a taxi navigation plan is received from the centralized route planner and the aircraft crew and ground marshal both signal ready to push back, the tug pushes the aircraft away from the gate and begins navigation through its assigned route. When reaching a designated location in the takeoff queue near the runway, the tug autonomously detaches from the aircraft, moves to a safe position away from the aircraft, signals to the aircraft's crew through a cockpit display that it is detached, and navigates back to the depot along the route provided by the planner.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Efficiency, environmental and economic benefits	Amount of delay in taxi time and maximizing throughput, reduced fuel emissions, reduced maintenance costs	Advantage of self-driving towing vehicle on busy airports
	2	Complexity of logistics	Complexity of logistics, primarily in the form of workload for flight crew, tower personnel or ground crew	Advantage of self-driving towing vehicle as to reduced workload for personnel
	3	Safety	Safety in the form of things like maintaining separation constraints and	No compromises on safety by the autonomous operation

			avoiding potentially dangerous events such as runway incursions	
AI features	Task(s)	Environment Perception, Path Planning, Obstacle Avoidance, Navigation, Fault Detection, Situational Awareness		
	Method(s)	computer vision , logical decision making, pattern recognition, multimodal event detection, multi-agent planning and scheduling, probabilistic predictive modelling		
	Hardware	host platform: AeroTech Expediter 600; positioning sensors, environment sensors (LIDAR, electro-optical cameras, infrared cameras)		
	Topology	autonomous vehicle guidance, environment perception, self perception, planning and scheduling		
	Terms and concepts used			
Standardization opportunities/ requirements				
Challenges and issues	Safe operations in the airfield environment, minimal changes to the airport infrastructure, minimal impact of their incorporation into normal operations			
Societal Concerns	Description	If labor replacements are involved, then the use of autonomy must provide an equivalent or greater benefit to some portion of the labor pool to offset the potential job loss; furthermore, they must operate in a way that feels common and familiar to humans, and must be perceived as completely safe, simple and non-intimidating.		
	SDGs to be achieved			

A.10.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1					NASA Ames Research Center	www.nasa.gov
2					NASA Johnson Space Center	www.nasa.gov
3					Lockheed Martin Advanced Technology Laboratories	www.lmco.com
4					University of California-Santa Cruz Affiliated Research Center	www.ucsc.edu
5					Carnegie Mellon University	www.cmu.edu

[1] R. Morris *et al.*, "Planning, Scheduling and Monitoring for Airport Surface Operations.," in *Workshop of the Thirtieth AAAI Conference on Artificial Intelligence: Planning for Hybrid Systems*, 2016 [Online]. Available: https://www.aaai.org/ocs/index.php/WS/AAAIW16/paper/download/12611/12430

[2] R. Morris *et al.*, "Self-Driving Aircraft Towing Vehicles: A Preliminary Report.," in *AAAI Workshop: AI for Transportation*, 2015.

[3] C. DiPrima and M. Fong, "TAXIING INTO AUTONOMY, 2018 IATA Competition Submission." 2018 [Online]. Available: https://www.iata.org/events/ighc/Documents/IGHC%20Innovator%202018/taxiing-into-autonomy.pdf

[4] E. V. Cross *et al.*, "SafeTug Semi-Autonomous Aircraft Towing Vehicles," NASA Aeronautics and Research Mission, Ames Research Center, 2015 [Online]. Available: https://www.researchgate.net/publication/311790811_SafeTug_Semi-Autonomous_Aircraft_Towing_Vehicles

A.11 Unmanned Protective Vehicle for Road Works on Motorways

A.11.1 General

ID	11	
Use case name	Unmanned Protective Vehicle for Road Works on Motorways	
Application domain	Transportation	
Deployment model	Self-driving vehicles	
Status	Prototype	
Scope	Unmanned operation of a protective vehicle in order to reduce the risk for road workers in short-time and mobile road works carried out in moving traffic	
Objective(s)	A vehicle that is able to follow mobile road works automatically on the hard shoulder of a German motorway.	
Narrative	Short description (not more than 150 words)	Mobile road works on the hard shoulder of German highways bear an increased accident risk for the crew of the protective vehicle safeguarding road works against moving traffic. The "Automated Unmanned Protective Vehicle for Highway Hard Shoulder Road Works" aims at the unmanned operation of the protective vehicle in order to reduce this risk. The vehicle has first been tested in a real operation on the German autobahn A3 in June 2018 [4]. It is actually the very first unmanned operation of a vehicle on German roads in public traffic. The scientific challenges of the project are strongly related to the general challenges in the field of automated driving.
	Complete description	A typical operational scenario for the automated unmanned protective vehicle looks as follows: In the beginning of the operation, an employee of the road maintenance service manually drives the protective vehicle from the depot to the location of the road works. There the employee stops the protective vehicle and switches to the road maintenance vehicle in front. The employee can activate the automated operation of the protective vehicle via a user interface. The vehicle guidance system then takes over the longitudinal and lateral control of the protective vehicle and follows the

<table>
<tr><td></td><td></td><td colspan="3">road maintenance vehicle in a defined distance at low speeds of about 10 km/h. In unmanned operation the vehicle guidance system operates in one of the three automated modes: Follow Mode, Coupled Mode, and Safe Halt. In Follow Mode, the vehicle guidance system performs the longitudinal and lateral control based on environmental information. The environment perception extracts the lane boundaries, e.g. lane markings, of the highway hard shoulder, the road maintenance vehicle and other obstacles in front of the protective vehicle. If an obstacle is detected, for example an emergency halting car, the system automatically transitions into Safe Halt. The system also performs this transition in case it detects that it is not capable of maintaining unmanned operation. In Coupled Mode, the protective vehicle is controlled by the vehicle guidance system, too. The longitudinal and lateral control is purely based on control commands and state information of the road maintenance vehicle. While lane boundaries are ignored in this mode of operation, obstacles in front of the protective vehicle are still detected. As in Follow Mode, the protective vehicle is able to detect functional system boundaries and to transfer itself to Safe Halt.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">obstacle detection, lane following, scene perception and representation, self perception</td></tr>
<tr><td>Method(s)</td><td colspan="3">computer vision , logical decision making, pattern recognition, multimodal event detection</td></tr>
<tr><td>Hardware</td><td colspan="3">truck vehicle equipped with cameras, radar system, motion and acceleration sensors, rain sensor</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">autonomous vehicle guidance, environment perception, self perception</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Safe operations in public traffic, compliance with ISO 26262</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3"></td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3"></td></tr>
</table>

3343

A.11.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1					MAN Truck & Bus AG	www.mantruckandbus.com
2					ZF Friedrichshafen AG	www.zf.com
3					WABCO Development GmbH	www.wabco-auto.com
4					Hochschule Karlsruhe	www.hs-karlsruhe.de
5					Technische Universität Braunschweig	www.tu-braunschweig.de
6					Hessen Mobil - Road and Traffic Management	mobil.hessen.de
7					BASt - Federal Highway Research Institute	www.bast.de

[1] G. Bagschik, M. Nolte, S. Ernst, and M. Maurer, "A System's Perspective Towards an Architecture Framework for Safe Automated Vehicles," Apr. 2018 [Online]. Available: http://arxiv.org/pdf/1804.07020v3

[2] T. Stolte, A. Reschka, G. Bagschik, and M. Maurer, "Towards Automated Driving: Unmanned Protective Vehicle for Highway Hard Shoulder Road Works," in 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 2015.

[3] G. Bagschik, A. Reschka, T. Stolte, and M. Maurer, "Identification of Potential Hazardous Events for an Unmanned Protective Vehicle," Apr. 2018 [Online]. Available: https://arxiv.org/pdf/1804.08728.pdf

[4] P. Thomas, "Autonome Lastwagen: Folgen Sie mir unauffällig!" 2018 [Online]. Available: http://www.faz.net/aktuell/technik-motor/motor/selbstfahrende-lastwagen-sollen-wanderbaustellen-sichern-15669484.html

A.12 Autonomous Apron Truck

A.12.1 General

ID	12
Use case name	autonomous apron truck
Application domain	Mobility
Deployment model	Embedded systems
Status	PoC
Scope	Automated transportation of luggage (carts) to requested destinations on an airport apron while following local traffic rules and resolve unplanned conflicts.
Objective(s)	Automate transport to increase reliability, precision, efficiency and safety.

	Short description (not more than 150 words)	An AI solution was planned that could operate a luggage truck on an airport apron where it interacts with aircrafts, other machines and humans. It prevents accidents with humans at all times and follows local traffic rules.
Narrative	Complete description	While the number of airplanes visiting German airports steadily increased over the last decades and recently reached a new all-time high the logistics to enable a smooth processing also increased correspondingly in complexity. To further manage even higher number of airplanes a fully automated luggage truck is developed. The truck shall receive tasks from a machine or human coordinator and automatically execute these. For specific tasks as loading and unloading or maintenance further interaction with human workers is needed. Therefore the truck is able to communicate its status and intents to surrounding workers. While operating on the apron the truck shall always obey local traffic rules. The only occasion to violate these rules if an accident is thereby avoided. Human safety is always the truck's first priority. For achieving all these functions an AI system consisting of multiple individual elements which all have to operate collaboratively is designed. The three main modules are a perception module, a behavior generator and an execution module. The truck perceives its environment is by its perception module which consists of multiple submodules, as object detection, recognition, tracking and data fusion blocks for multiple sensor types. The perceived information and their respective uncertainties are further processed to localize, re-project and detect the objects' intend in the trucks coordinate system. The perception unit outputs a context model which the behavior generator receives to decide on what actions to take next. This behavior generator consists of a deep reinforcement learning agent and is supervised by a symbolic rule checker to reassure the agent operates fault free. If a taken action violates a rule either the agent has to determine a new action or, in safety critical situations the rule checker determines safe actions by symbolic reasoning. The execution module executes the behavior determined by the behavior generator. It consists of motion planning, control and communication submodules which execute the intended task while reporting back to the behavior generator to react on unexpected situations. Additionally, the trucks status and intends are constantly reported over communication systems to its surrounding to enable uncomplicated interaction with the truck.
Stakeholders		

Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Safety	Number of accidents weighted by the level of severity.	Reduce accidents
	2	Efficiency	The sum of idle time and covered distance.	Improve efficiency
AI features	Task(s)	Other (please specify) Sense&Plan&Act		
	Method(s)	Symbolic reasoning & sub-symbolic machine learning & Image Processing, Data Fusion		
	Hardware			
	Topology			
	Terms and concepts used	Computer Vision, Symbolic Reasoning, Deep Reinforcement Learning		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Achieve at least the same level as human truck operators. Issues: 1) detect other apron traffic participants (especially aircraft) including intentions 2) Multiplicity of various outside conditions (e.g. signs painted on road but ice and snow covering it), and 3) prediction of human behaviour (e.g. workers in reverse walk)			
Societal Concerns	Description	Changed work environment for workers during loading/unloading with less interactions with co-workers but more non-social interactions (machines).		
	SDGs to be achieved			

A.12.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	IEEE ITSC 2018: @inproceedings{DBLP:conf/itsc/, author = {Martin Buechel, Alois Knoll}, title = {Deep Reinforcement Learning for Predictive Longitudinal Control of Automated Vehicles}, booktitle = {21th {IEEE} International Conference on Intelligent Transportation Systems, {ITSC} 2018, Hawaii, November 4-7, 2018}, pages = {}, year = {2018},		Predictive control of the vehicle	fortiss	

		crossref = {DBLP:conf/itsc/2018}, } (to appear)				
2	Public ation	IEEE ITSC 2018: @inproceedings{DBLP:conf/itsc/, author = {Michael Truong Le, Frederik Diehl, Thomas Brunner, Alois Knoll}, title = {Uncertainty Estimation for Deep Neural Object Detectors in Safety-Critical Applications}, booktitle = {21th {IEEE} International Conference on Intelligent Transportation Systems, {ITSC} 2018, Hawaii, November 4-7, 2018}, pages = {}, year = {2018}, crossref = {DBLP:conf/itsc/2018}, } (to appear)		Estimating the uncertainti es of the vehicles sensor processing	fortiss	
3	Public ation	IEEE ITSC 2018: @inproceedings{DBLP:conf/itsc/, author = {Klemens Esterle, Patrick Christopher Hart, Alois Knoll}, title = {Spatiotemporal Motion Planning with Combinatorial Reasoning for Autonomous Urban Driving}, booktitle = {21th {IEEE} International Conference on Intelligent Transportation Systems, {ITSC} 2018, Hawaii, November 4-7, 2018}, pages = {}, year = {2018}, crossref = {DBLP:conf/itsc/2018}, } (to appear)		The vehicles motion planning with combinato rial reasoning	fortiss	
4	Public ation	IEEE ITSC 2018: @inproceedings{DBLP:conf/itsc/, author = {Tobias Kessler, Pascal Minnerup, Klemens Esterle, Christian Feist, Florian Mickler, Erwin Roth, Alois Knoll}, title = {Roadgraph Generation and Free-Space Estimation in Unknown Structured Environments for Autonomous Vehicle Motion Planning}, booktitle = {21th {IEEE} International Conference on Intelligent Transportation Systems, {ITSC} 2018, Hawaii, November 4-7, 2018}, pages = {}, year = {2018}, crossref = {DBLP:conf/itsc/2018}, }		The vehicles' ability to plan in unknown environme nts	fortiss	

		(to appear)				
5	Public ation	IEEE ITSC 2018: @inproceedings{DBLP:conf/itsc/, author = {Julian Bernhard and Robert Gieselmann and Alois Knoll}, title = {Experience Based Heuristic Search: Robust Motion Planning with Deep Q-Learning}, booktitle = {21th {IEEE} International Conference on Intelligent Transportation Systems, {ITSC} 2018, Hawaii, November 4-7, 2018}, pages = {}, year = {2018}, crossref = {DBLP:conf/itsc/2018}, } (to appear)		Robust motion planning	fortiss	

A.13 AI Solution to Identify Automatically False Positives from a Specific Check for "Untranslated Target Segments" from an Automated Quality Assurance Tool

A.13.1 General

ID	13	
Use case name	AI solution to identify automatically false positives from a specific check for "untranslated target segments" from an automated quality assurance tool	
Application domain	Other (please specify) This will be relevant for content from across any domains	
Deployment model	Cloud services	
Status	PoC	
Scope	The scope of this use case is limited to automated linguistic quality assurance tools, but the outcome of this use case could be applicable to other areas, such as for example: Machine Translation, automated post-editing, Computer Aided Translation Analysis and pre-translation, etc.	
Objective(s)	To reduce the number of false positive issues for check for untranslated target segment for bilingual content with in-house automated quality assurance tool.	
Narrative	Short description (not more than 150 words)	In the future, we aim to build an AI solution that could automatically identify likely false positives issues from the results of the "check for untranslated target segments" following an approach where we could use machine learning based on already identified false positives by our users. The expected outcome would be to increase end user's productivity when reviewing automated quality assurance findings and to change user behaviour to pay more attention to this type of issues by reducing the number of false positives in 80%. In addition, we would like to reduce the amount of time, we spent on a yearly basis on refining this check manually based on users' feedback.

Complete description

Untranslated target segments contain characters, symbols, and words that remain the same in source and target language. These segments can contain, numbers, alphanumeric content, numbers, code, e-mail addresses, prices, proper nouns, etc. or any combination of those. On a yearly basis, this check produces over 1 Million potential issues across over 50 different languages.

Refining this check manually based on annotated false positive data for each specific customer and product and for specific language pairs is very costly, and the coverage is never sufficient, as new content is constantly produced and there are always new opportunities for refining this check via code. In addition, because of the high proportion of false positives over (95.5%) our translators tend to ignore the output from this valuable check and in many cases, we suspect that valid relevant issues for situations when there are real forgotten translations are missed.

There are typically three types of false positives for this type of check:

1) Language specific false positives, for example for situations where source and target segment need to be the same as the words from these segments are "cognates" with the same meaning. For example:

Source Language	Target Language	Source Segment	Target Segment	Checker Message
en-us	es-es	beta	beta	1445 - Untranslated target segment.
en-us	fr-fr	beta	beta	1445 - Untranslated target segment.
en-us	es-es	Monitor	Monitor	1445 - Untranslated target segment.

2) Customer profile specific false positives, for example situations where certain segments are to be left untranslated based on specific guidelines from the customer, for example for segments that jut consist of Company names, Product Names or specific words and segments that have been determined as not to be translated by our customer:

Source Language	Target Language	Source Segment	Target Segment	Checker M
en-us	es-es	Microsoft	Microsoft	1445 - Untr
en-us	es-es	Microsoft Azure	Microsoft Azure	1445 - Untr
en-us	es-es	- Outlook	- Outlook	1445 - Untr

3) Segments that remain the same in source and target, because they act as special type of entities with some special meaning, for example:
alphanumeric segments, for example part numbers, placeholders, code.

Source Language	Target Language	Source Segment	Target Segment	Checker Message
en-us	es-es	public inline virtual const std::string & GetErrorName	public inline virtual const std::string & GetErrorName	1445 - Untranslated target segment.
en-us	es-es	public int64_t Write	public int64_t Write	1445 - Untranslated target segment.
en-us	es-es	SELECT * FROM c	SELECT * FROM c	1445 - Untranslated target segment.

The idea is to create an AI solution that can automatically identify results from the "check for untranslated target segment" that are likely to be a False Positive. With this

<table>
<tr><td></td><td></td><td colspan="3">solution, we expect to reduce the number of potential issues presented by this check to our end users in 80%. This way our end users can focus their efforts on those potential issues that are more likely to be valid corrections because there could have been a forgotten translation. In addition, we will be able to increase the productivity of our end users when reviewing automated quality assurance potential issues from their bilingual content evaluation, and we will be able to save costs internally as we won't have to manually implement code changes in this check based on manual analysis of our data based on user's annotation.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Customers, Translation partners, end users of the translated content.</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Customer's content</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Bias from changes in requirements on the customer's end or inappropriate training data.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Coverage</td><td>Ratio of potential issues which are "of interest" for human evaluation. Ideal target is to reduce the current volume by 80%.</td><td>Improve accuracy</td></tr>
<tr><td>2</td><td>Split</td><td>Proportion of the potential issues which are "more likely to be a valid issue" for our end users.</td><td>Improve efficiency</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine Learning</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Machine Learning</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Challenges: Try to achieve eventually 80% of the accuracy of linguists when identifying false positives for untranslated target segments, preventing as much as possible false negatives.
Issues: segmentation of false positive data by Customer and Product profile could be challenging.</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Not applicable</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3"></td></tr>
</table>

3364 A.13.2 Data

Data characteristics	
Description	Data from end user identification of false positives and valid corrections for the "untranslated target segment check" results of Moravia QA Tools.
Source	RWS Moravia Analytics Portal (https://analytics.moravia.com/Dashboard/459)
Type	Structured content in a table with additional metadata fields (source segment, target segment, source language, target language, valid correction, false positive, customer and product profile, frequency)
Volume (size)	(Data for last 18 months)
Velocity	Every hour
Variety	Data types will be the same but there would be different variables to be considered (source language, target language, customer and product profile)
Variability (rate of change)	No changes
Quality	End-user dependent

3365

3366 A.14 Behavioural and Sentiment Analytics

3367 A.14.1 General

ID	14	
Use case name	Behavioural and sentiment analytics	
Application domain	Security	
Deployment model	On-premise systems	
Status	PoC	
Scope	Derive emotional state and goal of person from their gestures, face, actions	
Objective(s)	Determine if the movements, actions and general behaviour of a person is sign of malevolent intentions. Detect stealing of objects and other criminal behaviours. Prevent undesired behaviour (suicide), adapt narrative to state of person, provide dynamic content according to emotional responses.	
Narrative	Short description (not more than 150 words)	
	Complete description	
Stakeholders	Organizations, end users, community	
Stakeholders' assets, values	Reputation, trustworthiness, fair treatment, privacy	

System's threats & vulnerabilities	Bias, security threats, privacy threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
AI features	Task(s)	Recognition		
	Method(s)	Decision trees, deep learning		
	Hardware	Video camera, microphone, network, cpu, gpu		
	Topology			
	Terms and concepts used	Behavioural analytics, action, visual cues, sentiment, emotion, goal, social media, security, surveillance		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Surveillance cameras often have low resolution, can be in poorly lit environment with bad top-down view angle. A lot of suspicious behaviour can be hidden by passer-by or large crowds. Issues: Unwanted behaviours is MUCH LESS frequent than normal behaviour and can take on various forms			
Societal Concerns	Description	Right to privacy		
	SDGs to be achieved	Peace, justice and strong institutions		

A.15 Generative Design of Mechanical Parts

A.15.1 General

ID	15	
Use case name	Generative design of mechanical parts	
Application domain	Manufacturing	
Deployment model	On-premise systems	
Status	In operation	
Scope	Help mechanical engineers design lighter, strong, better parts	
Objective(s)	Create optimized parts following precise mechanical constraint while permitting cost savings by reducing the amount of material necessary to achieve goals.	
Narrative	Short description (not more than 150 words)	From Wikipedia: Generative design is an iterative design process that involves a program that will generate a certain number of outputs that meet certain constraints, and a designer that will fine tune the feasible region by changing minimal and maximal values of an interval in which a variable of the program meets the set of constraints, in order to reduce or augment the number of outputs to choose from.
	Complete description	https://en.wikipedia.org/wiki/Generative_design https://www.autodesk.com/solutions/generative-design

		http://www.newequipment.com/research-and-development/what-generative-design-and-why-its-future-manufacturing		
Stakeholders	Organizations, Designers, Customers, End users			
Stakeholders' assets, values	Competitiveness, safety, stability			
System's threats & vulnerabilities	Highly dependent on engineer input for constraints and requirements			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Weight reduction	Is the resulting part lighter than original version	Use less material
	2	Mechanical constraints metrics	Various mechanical metrics	Obtain strong, better parts
AI features	Task(s)	Optimization		
	Method(s)	Genetic algorithms, optimisation algorithms, generative adversarial networks		
	Hardware	CPU, GPU		
	Topology			
	Terms and concepts used	Design, generative adversarial network, genetic algorithm, mimicry		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: The engineers using this technology still need to know how to define the constraints, start and end points for the piece. Issues: Pieces generated to satisfy a set of constraint may still have design flaws overlooked because of misunderstanding by the user.			
Societal Concerns	Description			
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.15.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Public	Wikipedia Generative Design webpage			Contributions	https://en.wikipedia.org/wiki/Generative_design
2	Public	Generative design solutions from autodesk			Autodesk	https://www.autodesk.com/solutions/generative-design
3	Public	R&D article on the future of manufacturing			New equipment digest	https://www.newequipment.com/research-and-development/what-

						generative-design-and-why-its-future-manufacturing

A.16 Robotic Prehension of Objects

A.16.1 General

ID	16	
Use case name	Robotic prehension of objects	
Application domain	Other (please specify) Robotics	
Deployment model	Embedded systems	
Status	PoC	
Scope	Outputting end effector velocity & rotation vector in response to view from RGB-D camera located on robot wrist	
Objective(s)	Use reinforcement learning to train the robot to grasp misc. objects in simulation and transfer this learning to real-life robots.	
Narrative	Short description (not more than 150 words)	It may be difficult and time-consuming for clients of assistive robotic arms to control them with the fine degree required for grasping household objects (such as in the context of having a meal). In order to improve their quality of life, we propose a method by which users can select the bounding box around the object they wish grasped, and the robot performs the grasping action. We use methods from reinforcement learning to train first in simulation, in order to reduce total training time and potential robot breakage, and then transfer this learning to real-life.
	Complete description	It can be very difficult and time-consuming for users to perform fine movements with a robot arm, like grasping various household objects. To mitigate this problem, attempts are made to grant users the ability to control the arm at a higher level of abstraction; thus, rather than specifying each translation and rotation of the arm, we would like them to be able to select an object to grasp, and have the arm grasp it automatically. This requires some degree of computer vision, to be able to detect objects in the robot's field of view (a camera will be affixed to its wrist). With that achieved, we will be able to focus on grasping an object selected from the detections. Current literature on robotic grasping One might be tempted to start from a heuristic, geometric approach. That is, to use a set of pre-established rules for picking up objects -- for example, executing pincer grasps from the top along the thinnest dimension of the object that is not too narrow to be grasped. Such approaches work reasonably well in conditions that match the restrictive assumptions on which the rules are built, but fail when encountering even small deviations from those conditions (for example, they do not adapt well to clutter). Attempting to list and plan a proper

		response to all such failure cases heuristically would be an exercise in futility. In contrast, approaches based on machine learning can generalize to unforeseen or novel situations, and, as in the case of object detection, generally perform better than heuristic solutions. Machine learning-based approaches to grasping and object manipulation vary widely. At the simplest level, we can predict the likelihood of grasp success based on an image patch of an object and a given angle of approach. Robot control, in such cases, is beyond the scope of the machine learning model. However, methods can scale up to end-to-end systems which learn to control the robot at the level of its joint actuators in response to a visual stimulus consisting of a bird's eye view of the arm and several objects placed in a bin.		
Stakeholders	Customers, 3rd parties, end users, community			
Stakeholders' assets, values	Trustworthiness, safety, privacy, stability			
System's threats & vulnerabilities	Object or gripper bias, security threats, privacy threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Success rate in simulation	Grasp success rates on both objects seen during training, and new objects, in simulation.	Improve accuracy and generalization.
	2	Success rate in real life	Grasp success rates on both objects seen during training, and new objects, in real life.	Improve accuracy and generalization.
AI features	Task(s)	Planning		
	Method(s)	Reinforcement learning, deep learning		
	Hardware	Depth camera, RGB camera, GPU, actuators, gripper		
	Topology			
	Terms and concepts used	Reinforcement learning, Deep learning, point cloud, depth, scene completion, grasping, transfer learning		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: The camera cannot have a bird's eye view and will instead move with the robot. Sparse rewards may complicate learning. Environment may be cluttered, occlusions of target might occur, objects may move around Issues: For safety reasons, speed and force of robot need to be limited in assistive environment to avoid harm. Human intervention can happen at any time.			
Societal Concerns	Description	Prevent arm to people and animals near robot when it is performing a grasping task		

	SDGs to be achieved	Good health and well-being for people

A.16.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Technical publication	Pinto L, Gupta A. Supersizing Self-supervision: Learning to Grasp from 50K Tries and 700 Robot Hours [Internet]. arXiv [cs.LG]. 2015.				http://arxiv.org/abs/1509.06825
2	Technical publication	Bousmalis K, Irpan A, Wohlhart P, Bai Y, Kelcey M, Kalakrishnan M, et al. Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping [Internet]. arXiv [cs.LG]. 2017				http://arxiv.org/abs/1709.07857
3		Gu S, Holly E, Lillicrap T, Levine S. Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates [Internet]. arXiv [cs.RO]. 2016				http://arxiv.org/abs/1610.00633

A.17 Robotic Vision – Scene Awareness

A.17.1 General

ID	17	
Use case name	Robotic vision – scene awareness	
Application domain	Other (please specify) Robotics	
Deployment model	Embedded systems	
Status	PoC	
Scope	Determining in which environment the robot is and which actions are available to it	
Objective(s)	Robustly identify the scene from video and depth sensors. From the scene and the seen objects, propose the actions to make to human collaborator	
Narrative	Short description (not more than 150 words)	Household robots need to navigate a very diverse set of environments and be able to accomplish different tasks depending on their position and action set. To meet these goals, the robots need to quickly and accurately identify the visual context in which they operate and derive the set of possible actions from this context. They can then propose relevant actions to the end user so that he does not have to define context himself and then sift through a long list of irrelevant actions.
	Complete description	http://places2.csail.mit.edu/challenge.html
Stakeholders	Customers, 3rd parties, end users, community	

<table>
<tr><td></td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Trustworthiness, safety, privacy, stability</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Dynamic environment, security threats, privacy threats</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Classification error</td><td>Min distance between 5 labels and ground truth</td><td>Improve context confidence</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition</td></tr>
<tr><td>Method(s)</td><td colspan="3">Deep learning, decision trees</td></tr>
<tr><td>Hardware</td><td colspan="3">Sensors, processors</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Context awareness, scene recognition, deep learning, action proposal</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Challenges: Environment can be poorly lit leading to difficult context recognition. Issue: Sensors degradation can occur</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Privacy concerns (what data from sensors is kept, reviewed and used to improve models).</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3">Industry, Innovation, and Infrastructure</td></tr>
</table>

A.17.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Public	Places challenge			Bolei Zhou, Aditya Khosla, Antonio Torralba, Aude Oliva	http://places2.csail.mit.edu/challenge.html
2	Peer-Reviewed	B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, "Learning deep features for scene recognition using places database," in In Advances in Neural Information Processing Systems, 2014.			MIT	http://places.csail.mit.edu/places_NIPS14.pdf
3	Peer-Reviewed	L. Herranz, S. Jiang, X. Li, "Scene recognition with CNNs: objects, scales and dataset bias", Proc. International Conference on Computer Vision and Pattern			Key Laboratory of Intelligent Information Processing of Chinese	https://arxiv.org/pdf/1801.06867.pdf

		Recognition (CVPR16), Las Vegas, Nevada			Academy of Sciences	

A.18 AI Solution for Car Damage Classification

A.18.1 General

ID	18	
Use case name	AI solution for Car Damage Classification	
Application domain	Other (Insurance)	
Deployment model	Cloud services	
Status	PoC	
Scope	Car damage classification for common damage types such as bumper dent, door dent, glass shatter, head lamp broken, tail lamp broken, scratch and smash.	
Objective(s)	1. To create an automated system for car damage classification using CNNs. 2. Experiment using transfer and ensemble learning to find which is better for training a CNN for car damage classification.	
Narrative	Short description (not more than 150 words)	Image based vehicle insurance processing is an important area with large scope for automation. We have considered the problem of Car damage classification. We explore deep learning based techniques for this purpose. Initially, we try directly training a CNN. However, due to small set of labeled data, it does not work well. Then, we explore the effect of domain-specific pre-training followed by fine-tuning. Finally, we experiment with transfer learning and ensemble learning. Experimental results show that transfer learning works better than domain specific fine-tuning. We achieve accuracy of 89.5% with combination of transfer and ensemble learning. We hosted the trained model on cloud that can be plugged into applications using API and can be used for automated first level assessment of the damage, in car insurance sector.
	Complete description	Today, in the car insurance industry, a lot of money is wasted due to claims leakage [1] [2]. Claims leakage / Underwriting leakage is defined as the difference between the actual claim payment made and the amount that should have been paid if all industry leading practices were applied. Visual inspection and validation have been used to reduce such effects. However, they introduce delays in the claim processing. There have been efforts by a few start-ups to mitigate claim processing time [3] [4]. An automated system for the car insurance claim processing is a need of the hour. We employ Convolutional Neural Network (CNN) based methods for classification of car damage types. Specifically, we consider common damage types such as bumper dent, door dent, glass shatter, head lamp broken, tail lamp broken, scratch and smash. To the best of our knowledge, there is no publicly available dataset for car damage classification. Therefore, we created our own dataset by collecting images from web and manually

		annotating them. The classification task is challenging due to factors such as large inter-class similarity and barely visible damages. We experimented with many techniques such as directly training a CNN, pre-training a CNN using auto-encoder followed by fine-tuning, using transfer learning from large CNNs trained on ImageNet and building an ensemble classifier on top of the set of pretrained classifiers. We observe that transfer learning combined with ensemble learning works the best. We also devise a method to localize a particular damage type. We achieve accuracy of 89.5% with combination of transfer and ensemble learning. The same technique can be used for localization of damages. Further, only car specific features may not be effective for damage classification. It thus underlines the superiority of feature representation learned from the large training sets. We hosted the trained model on cloud that can be plugged into applications using API and can be used for automated first level assessment of damages, in car insurance sector.		
Stakeholders	Insurance companies, Car owner/user			
Stakeholders' assets, values	competitiveness, reputation, trustworthiness, fair treatment			
System's threats & vulnerabilities	Misclassification of car damage and insurance claims			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy	We performed experiment with transfer learning and ensemble learning. Experimental results show that transfer learning works better than domain specific fine-tuning. We achieve accuracy of 89.5% with combination of transfer and ensemble learning.	Objective 2
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware	c4.2xlarge Amazon AWS EC2 instance which has 8 core Intel Xeon E5-2666 v3 (Haswell) CPUs and 15GB RAM		
	Topology	GPU enabled servers		

	Terms and concepts used	Deep learning, transfer learning, supervised learning, convolutional neural networks
Standardization opportunities/ requirements	ensemble learning, transfer learning, Localization, manual annotation through crowd sourced efforts	
Challenges and issues	1. Small size of the damages 2. Less Quantity of data 3. Ambiguity in damaged and non-damaged images	
Societal Concerns	Description	Insurance agents may need to be re-skilled
	SDGs to be achieved	Decent work and economic growth

A.18.2 Data

Data characteristics	
Description	We created a dataset consisting of images belonging to different types of car damage. We consider seven commonly observed types of damage such as bumper dent, door dent, glass shatter, head lamp broken, tail lamp broken, scratch and smash. In addition, we also collected images which belong to a no damage class
Source	The images were collected from web and were manually annotated
Type	
Volume (size)	
Velocity	
Variety	multiple web sources
Variability (rate of change)	
Quality	Medium

A.18.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference Paper	International Conference on Machine Learning and applications	Published		Tata Consultancy Services Limited	https://ieeexplore.ieee.org/abstract/document/8260613/

A.19 AI to Understand Adulteration in Commonly Used Food Items

A.19.1 General

ID	19

<table>
<tr><td>Use case name</td><td colspan="4">AI to understand adulteration in commonly used food items</td></tr>
<tr><td>Application domain</td><td colspan="4">Agriculture</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">PoC</td></tr>
<tr><td>Scope</td><td colspan="4">Understand the patterns in hyperspectral / NIR or visual imaging specifically for adulteration in milk, banana and mangoes</td></tr>
<tr><td>Objective(s)</td><td colspan="4">To device a simple , cost effective tool to identify the adulteration in food items at point of purchase</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Food adulteration is one of the big evil of modern society. Adulterated milk is hazard for children, many aliments including cancer / kidney failures due to consumption of adulterated food. Hyperspectral technology was evaluated to find out adulteration in food items</td></tr>
<tr><td>Complete description</td><td colspan="3">Food adulteration is becoming menace especially with adulterants that are either carcinogenic or harmful to body parts like kidney. To give few examples, Milk is adulterated with Soda, Urea and detergents. Whereas mangoes and bananas are quickly ripened by calcium carbide and so on. Common man cannot live without these items. There is no frugal way to identify these type of adulterations.
Experiment of controlled adulteration was done and hyperspectral reflectance reading were taken.
AI helped to find the patterns in hyperspectral signature and was able to reliably classify (90% ++) samples that were unadulterated and adulterated.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Consumers, Farmers, Health monitoring agencies</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Health, reputation, trust, fair treatment</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">different sources of bias, incorrect AI system use, improperly trained model, incorrect classification</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Features related to adulterants in radio spectrum</td><td>Intensities around NIR range</td><td>Health</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine learning</td></tr>
<tr><td>Hardware</td><td colspan="3">Hyperspectral camera, GPS servers</td></tr>
<tr><td>Topology</td><td colspan="3">GPU servers</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Deep learning, supervised learning, classification</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Image classification of hyper-spectral camera images</td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Large scale data collection, Miniaturization of frugal NIR / Hyperspectral sensor</td></tr>
</table>

Societal Concerns	Description	If the AI system is rolled out and taken as reliable then it should be able to perform in all cases and scenarios. Incorrect classification can lead to false accusations
	SDGs to be achieved	Good health and well-being for people

A.19.2 Data

Data characteristics	
Description	Hyperspectral signatures (300 nm to 1300 nm @ 30 nm band)
Source	Hyperspectral camera
Type	
Volume (size)	~ 500 samples
Velocity	
Variety	
Variability (rate of change)	
Quality	

A.19.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Published in SPIE Proceedings Vol.9860: Hyperspectral Imaging Sensors: Innovative Applications and Sensor Standards 2016 David P. Bannon, Editor(s)			Tata Consultancy Services Limited	http://spie.org/Publications/Proceedings/Paper/10.1117/12.2223439?origin_id=x4323& start_year=1963

A.20 Detection of Frauds based on Collusions

A.20.1 General

ID	20
Use case name	Detection of frauds based on collusions
Application domain	Fintech
Deployment model	On-premise systems
Status	In operation

<table>
<tr><td>Scope</td><td colspan="4">Validating the predicted collusion set is effort-intensive and needs investigative and legal expertise</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Automatic unsupervised detection of frauds based on collusions</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">A set of unsupervised machine learning algorithms to detect collusion-based frauds, particularly, circular trading and price manipulation in stock market trading</td></tr>
<tr><td>Complete description</td><td colspan="3">Frauds are prevalent across all industries; and they are particularly severe in today’s computerized, web-connected, mobile-accessible, and cloud-enabled business environments. An FBI report states that the insurance industry in the US, which consists of over 7000 companies and collects over $1 trillion in premiums, loses about $40 billion annually in frauds in the non-health insurance sector alone. The aggregate size of the 52 regulated stock exchanges across the world (total market capitalization) was $55 trillion as on Dec. 2012. Given the money involved, it is not surprising that the stock market is a target of frauds.

Many malpractices in stock market trading, e.g. circular trading and price manipulation—use the modus operandi of collusion. Informally, a set of traders is a candidate collusion set when they have “heavy trading” among themselves, as compared to their trading with others. We formalize the problem of detection of collusion sets, if any, in a given trading database. We show that naïve approaches are inefficient for real-life situations. We adapt and apply two well-known graph clustering algorithms for this problem. We also propose a new graph clustering algorithm, specifically tailored for detecting collusion Sets; further, we establish a combined collusion set. Treating individual experiments as evidence, this approach allows us to quantify the confidence (or belief) in the candidate collusion sets. We have carried out detailed simulation experiments to demonstrate effectiveness of the proposed algorithms. The system is also operational in a government organization. Note that all our collusion detection algorithms are completely unsupervised and do not need any training data.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Stock market regulator, stock traders, stock investors</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Fair price, Prevention of Collusions and frauds</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Incorrect fraud detection may lead to unnecessary alerts</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Prediction accuracy</td><td>How many predicted collusion sets were actually involved in frauds</td><td>Improve accuracy</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>AI features</td><td>Task(s)</td><td colspan="3">Knowledge processing & discovery</td></tr>
</table>

	Method(s)	Machine learning
	Hardware	GPU enabled servers
	Topology	GPU enabled servers
	Terms and concepts used	Deep learning, unsupervised learning, clustering
Standardization opportunities/ requirements	Graph based clustering	
Challenges and issues	Actual examples of collusion-based frauds may not be available easily, even for evaluation and testing	
Societal Concerns	Description	Incorrect detection of Collusions and frauds may cause unnecessary stress in stock traders
	SDGs to be achieved	Decent work and economic growth

A.20.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference				Tata Consultancy Services Limited	D. K. Luna, G. K. Palshikar, M. Apte, A. Bhattacharya, *Finding Shell Company Accounts using Anomaly Detection*, **ACM India Joint International Conference on Data Science and Management (CoDS-COMAD 2018),** Goa, India, Jan 11-13, 2018
2	Journal				Tata Consultancy Services Limited	G. K. Palshikar, M. Apte, *Collusion Set Detection Using Graph Clustering*, vol. 16, no. 2, April 2008, **Data Mining and Knowledge Discovery** journal (Springer-Verlag), pp. 135 – 164
3	Book chapter				Tata Consultancy Services Limited	M. Apte, G.K. Palshikar, S. Baskaran, *Frauds in Online Social Networks: A Review*, accepted as a Book Chapter, in **Social Network and Surveillance for Society**, T. Ozyer and S. Bakshi (ed.s), to be published by Springer in 2018
4	Book chapter				Tata Consultancy Services Limited	G.K. Palshikar, M. Apte, *Financial Security against Money Laundering: A Survey*, Chapter 36 in B. Akhgar, H.R. Arabnia (Ed.s), **Emerging**

						Trends in Information and Communication Technologies Security, pp. 577 – 590, Elsevier (Morgan Kaufman), 2013

A.21 Information Extraction from Hand-marked Industrial Inspection Sheets

A.21.1 General

ID	21	
Use case name	Information Extraction from Hand-marked Industrial Inspection Sheets	
Application domain	Manufacturing	
Deployment model	Cloud services	
Status	PoC	
Scope	Localization and Mapping of machine zones, arrows and text, to extract information from manually tagged inspection sheets.	
Objective(s)	To create a pipeline to build an information extraction system for machine inspection sheets, by mapping the machine zones to the handwritten code using state-of-the-art deep learning and computer vision techniques.	
Narrative	Short description (not more than 150 words)	Inspection Sheets are filled regularly to detect defects and maintain heavy machines. Sheets contains a lot of unstructured information and requires domain experts' intervention to read and digitize. We have proposed a novel pipeline to build an information extraction system for such machine inspection sheets, utilizing state-of-the-art deep learning and computer vision techniques.
	Complete description	In order to effectively detect faults and maintain heavy machines, a standard practice in several organizations is to conduct regular manual inspections. The procedure for conducting such inspections requires marking of the damaged components on a standardized inspection sheet which is then camera scanned. These sheets are marked for different faults in corresponding machine zones using hand-drawn arrows and text. As a result, the reading environment is highly unstructured and requires a domain expert while extracting the manually marked information We have proposed a novel pipeline to build an information extraction system for such machine inspection sheets, utilizing state-of-the-art deep learning and computer vision techniques. The pipeline proceeds in the following stages: (1) localization of different zones of the machine, arrows and text using a combination of template matching, deep learning and connected components, and (2) mapping the machine zone to the corresponding arrow head and the text segment to the arrow tail, followed by pairing them to get the correct damage code for each zone.

		The proposed method yields an accuracy of 83.2% at the end of the pipeline. The organization has 2 million such sheets which are manually processed. This project will enable considerable savings in terms of time and manpower as it takes roughly 5 minutes per sheet for the manual process. The AI system will process a sheet in 20 seconds and can be parallelized for further speed up.		
Stakeholders	Manufacturing companies, Machine Inspectors, Engineers			
Stakeholders' assets, values	Reduced dependence on Expert Engineer time, Possibility of pointing out errors in inspection			
System's threats & vulnerabilities	Trained on one set of inspection sheets can lead to inaccurate classification of another inspector's inspection sheet			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy	Accuracy of system to read the code and map it to the right Machine zone	
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware	GPU enabled desktop / server		
	Topology	GPU enabled servers		
	Terms and concepts used	Deep learning, Feature engineering, Recurrent neural networks (RNN), Convolutional neural network (CNN)		
Standardization opportunities/ requirements	pipeline for information extraction from industrial inspection sheets			
Challenges and issues	Challenges: 1. Quality of Images 2. Structural deformities of individual components(arrows, handwritten code) 3. Quantity of data 4. Cascading effect of error at each stage of the pipeline			
Societal Concerns	Description	Inspection engineers may have to develop other skills		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.21.2 Data

Data characteristics	
Description	a dataset of anonymized inspection sheets provided by a company
Source	a company employing heavy machines in manufacturing
Type	Camera scanned images with resolution of 3210 *2200
Volume (size)	330 scans
Velocity	daily
Variety	Scanned inspection sheets; single source
Variability (rate of change)	Well scanned sheets, poorly scanned sheets, soiled sheets, poorly marked sheets
Quality	Can have missing text, missing arrows etc.

3403

3404 A.21.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Industrial Inspection	Physical inspection of heavy machinery	Scan of machine inspection sheet	Human inspected marked sheets	Digitized information from inspection sheets
2	Training Arrow Detection Model	Train a deep model to recognize arrows in an image	Arrow images	Synthetically generated arrow images	Trained detector with high > 90% accuracy
3	Training Regression model for arrow head and tail	Train a deep model for regressing to head and tail of arrows	Detected Arrow images	Arrow Images	Head and Tail Localization
4	Text Detection	Detect Text via deep model	Detected handwritten text	Handmarked image	Localized handwritten text
5	Reading Handwritten Text	Read text via deep model	Read handwritten text	Isolated handwritten text	Digitized text
6	Mapping of Zones	Zone Mapping	Map each text to a machine zone using arrow	Machine Zone to fault mapping	Final Mapping to database

3405 A.21.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Synthetic Arrow Dataset is ready	Train arrow detector	AI Solution Provider	Train a model to isolate arrows in an image	Needed for mapping text to zones
2	Handwritten text recognition	Train handwritten text recognizer	AI Solution Provider	Train a model to recognize handwritten text	Needed for fault identification
3	Text Detection	Isolate Handwritten text	AI Solution Provider	Train a model to isolate handwritten text	Needed for Text detection

Specification of training data	

3406

A.21.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement

Input of evaluation	Manually annotated sheets, AI System
Output of evaluation	Accuracy

A.21.6 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference Pape	International Conference on Document Analysis and Recognition	Published		Tata Consultancy Services Limited	https://ieeexplore.ieee.org/abstract/document/8270293/

A.22 AI (Swarm Intelligence) Solution for Attack Detection in IoT Environment

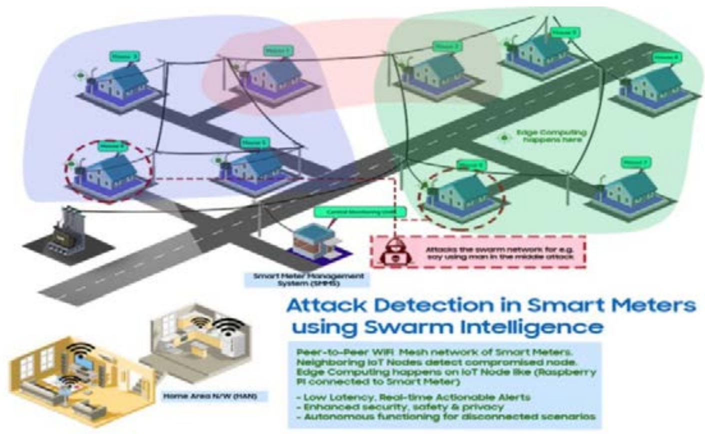
A.22.1 General

ID	22	
Use case name	AI (Swarm Intelligence) solution for Attack Detection in IoT Environment	
Application domain	Security	
Deployment model	Hybrid or other (Agent Based Hub-Spoke)	
Status	Prototype	
Scope	Anomaly Based Attack Detection in IoT environment using Swarm Intelligence	
Objective(s)	Given: AMI (Advanced Metering Infrastructure – Smart Meters in Smart Buildings in Smart Cities. Detect: Detect energy theft / meter tampering by consumer in AMI (Advanced Metering Infrastructure) or hacking attack by an external agent (man in the middle) for edge computing security scenarios with intermitted disconnection, near real-time response without using server or cloud-based analytics.	
Narrative	Short description (not more than 150 words)	This is a unique approach to detect attacks in IoT environment using Anomaly Based Attack Detection using Swarm Intelligence methods. This is a key solution to detect energy theft scenario in Smart Metering. Energy Theft problem varies from 2% in developed countries to 35% in developing countries. This is complimentary to

		traditional AI or other static rule-based analysis which is heavily dependent on analysis of huge amounts of data on centralized cloud infrastructure. This solution is simple, nimble and can be run on low powered edge (IoT Nodes) for near real-time, low latency, low power, small compute, small storage Mist / Edge Computing Scenarios.
	Complete description	**Introduction to Anomaly Based Attack Detection using Swarm Intelligence** Motivation ▪ World-wide statistics shows there will be IoT install based of 12.86 billion units in the consumer segment by 2020. ▪ In Smart city industry, smart security is expected to account for 13.5 percent of global smart city market. There will be more than 1 billion devices installed in smart homes. ▪ India is planning 100 Smart cities to be developed in next 5 years, and security is of paramount importance. Securing Advanced metering Infrastructure (AMI) will be key component for securing smart city infrastructure. ▪ Important aspect of securing AMI is securing the Smart Energy meters and detecting attacks on these smart meters. ▪ While there are many traditional solutions for anomaly and intrusion-based detection based on static preset rules / policies, these solutions are not effective in detecting future attacks that are already not known. A more robust and more secure security solution to detect attacks in edge network is essential. Hence a new innovative approach of using Swarm Intelligence along with Anomaly based Detection has been a technology choice to solve this problem in a unique way. **Problem Statement** Detect energy theft / meter tampering by consumer in AMI (Advanced Metering Infrastructure) or hacking attack by an external agent (man in the middle) for edge computing security scenarios with intermitted disconnection, near real-time response without using server or cloud-based analytics. **Current situation** There are many cloud based centralized solutions available using static rules / policies configured which can detect existing known attack only. Processing in centralized cloud involves transferring data from sensors / actuator to cloud which in itself is a concern in terms of privacy, security, regulations & compliance for some key industry verticals. **Solution Approach**

Swarm Intelligence is a specific branch of AI. A new innovative approach using swarm intelligence (AI) based solution for attack detection. Used collective behavior of decentralized self-organizing swarm of nodes with simple computational rules, interacting locally.

Result: Simple collective algorithms for detection of man in the middle attacks on data / network.
The following Anomaly based attack detection algorithms were used
1. Moving average based
2. Mahalanobis distance based
3. Entropy based



Use-Case: Attack detection of attacks AMI – Smart Metering network.

1. Energy Theft by consumer.
2. Attack launched by external entity (hacker) using say man in-the-middle attack.

Technology: Swarm Intelligence & Anomaly Based attack detection using energy consumption data from Smart Meter to detect attacks using consensus-based anomaly detection algorithms.

Solution Steps:
- Each Smart meter node reads its Energy Consumption data
- Node shares Energy Consumption data with its neighboring nodes
- Node computes anomaly index based on Anomaly Detection algorithm
- Neighboring nodes detect anomalous node(s) based on Anomaly index by consensus
- Neighboring nodes raise alarm indicating attacked / compromised node
- Notify alarm to back end host.

Display monitoring status on host UI.

Stakeholders	

<table>
<tr><td></td><td colspan="4">End users of Smart Metering, Utility Companies</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Competitiveness, trustworthiness, safety, privacy</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Challenges to accountability</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Recommendation</td><td>System can be used to detect even unknown attacks in IoT Environment especially for real-time or near real-time scenarios</td><td>use-case for AMI – Smart Metering with innovative approach</td></tr>
<tr><td>2</td><td>Improve accuracy</td><td>We found the accuracy of the model to be reasonably good</td><td>Improve accuracy</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Inference</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine Learning, Statistics, Heuristics, Anomaly Detection (Distance / Density based).</td></tr>
<tr><td>Hardware</td><td colspan="3">IoT Nodes (like Raspberry PI, Micro-Controllers, Edge Devices, Cloud etc.</td></tr>
<tr><td>Topology</td><td colspan="3">Agent based hub-spoke model.

Anomaly Detection in peer-to-peer mesh network.</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Swarm Intelligence, Anomaly Detection, AMI (Advanced Metering Infrastructure).</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Standardization of use of Swarm Intelligence for specific use case scenarios</td></tr>
<tr><td>Challenges and issues</td><td colspan="4">The problem is challenging because

1. Varied data set for different scenarios - large amount of data needs to be pre-processed to arrive at operation threshold parameters to be used for detection in real-time.
2. IoT (Edge) Nodes Configuration to suite specific environments The Swarm Intelligence System (SIS) involves a swarm of devices. It should be possible to easily configure the entire swarm for different network environments and locations.

Solution: Many reusable modules for Logging, Debugging and configuration through XML has been developed which has enabled binary re-use without having to change any code to suit a new network environment.

3. Flexible to reuse / customize solution for different use-cases / scenarios and scalability</td></tr>
</table>

		The platform needs to be able to provide facilities for different algorithms for anomaly detection to be plugged in with minimum modification, recoding, recompilation. **Solution:** Completely dynamically pluggable Algorithm binaries can be developed that conforms to defined interface Specifications, which gives flexibility to try out new algorithms, without needing to change existing code or re-compile. Use of Swarm Intelligence ensures very less localized communication that is required. Furthermore, the Swarm Intelligence System communication capability also addresses throttling of network traffic because of multi-threading / queuing capability built in.
Societal Concerns	Description	Accuracy of Solution. Fraud (Anomaly Detection) usually incurs a false positive alarm issue.
	SDGs to be achieved	Responsible consumption and production

3413

3414 A.22.2 **Data**

Data characteristics	
Description	Energy consumption data collected from smart meters.
Source	1. 3 years of dataset from smart meters downloaded from publicly available data source. 2. Meter Data Sets received from IIT-Delhi. 3. Sample data collected from Smart Meter setup in the Creative Lab (C-Lab) in Samsung. 4. Analysis & Recommendations on AMI (Advanced metering infrastructure) and Smart Metering scenarios from many research papers. Various online sources on application of Swarm Intelligence as a technology for solving complex problems using simple steps.
Type	Structured Data
Volume (size)	Multi-year Energy Consumption data from smart meters collected at the rate of 2 entries per hour 48 entries in a day; 17520 entries in a year.
Velocity	Batch, near-real time.
Variety	Single source. Similar data from multiple sources of smart meters.
Variability (rate of change)	Static. Datasets vary based on geography, season etc. as energy consumption varies based on these factors.
Quality	Contains some noise. Better quality after pre-processing.

3415

3416 A.22.3 **References**

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Energy Theft	published	High	TSINGHUA SCIENCE AND	https://ieeexplore.ieee.org/docum ent/6787363/

		Detection-AMI			TECHNOLOGY	
2	Paper	Intrusion Detection - AMI	published	High	IEEE University of Illinois	https://ieeexplore.ieee.org/docum ent/5622068/
3	Paper	EPPA	published	High	IEEE University of Waterloo, Waterloo	https://ieeexplore.ieee.org/docum ent/6165271/
4	Report	Quantifying the Extent of Energy Theft	published	Medium	City of Cape Town, SARPA	https://www.smartenergy.com/wpcontent/uploads/Deon%20Louw_ 0.pdf
5	website	About Swarm Intelligence	Available Online	High	TechFerry	http://www.techferry.com/article s/swarm-intelligence.html

A.23 VTrain Recommendation Engine

A.23.1 General

ID	23	
Use case name	VTrain recommendation engine	
Application domain	Education	
Deployment model	On-premise systems	
Status	In operation	
Scope	Based on an employee's career objectives find skill requirements and its training	
Objective(s)	Recommend a personalised list of "best" training courses to an employee, which will help him/her meet his/her career objectives.	
Narrative	Short description (not more than 150 words)	The vTrain system helps employees improve their skills by recommending appropriate training courses from a given list and historical data.
	Complete description	Continuous training is crucial for creating and maintaining the right skill-profile for the industrial organization's workforce. There is a tremendous variety in the available trainings within an organization: technical, project management, quality, leadership, domain-specific, soft-skills etc. Hence it is important to assist the employee in choosing the best trainings, which perfectly suits him/her background, project needs and career goals. In this work, we focus on algorithms for training recommendation in an industrial setting. We formalize the problem of next training recommendation, taking into account the employee's training and work history. We have developed several new unsupervised sequence mining algorithms to mine the past trainings data from the organization for arriving at personalized next training recommendation. Using the real-

		life data about trainings of 118587 employees over 5019 distinct trainings from a large multi-national IT organization, we show that these algorithms outperform several standard recommendation engine algorithms as well as those based on standard sequence mining algorithms.		
Stakeholders	Employees, Job requirements, Training requirements			
Stakeholders' assets, values	Skill profile, Job description requirements			
System's threats & vulnerabilities	Different sources of bias can come based on model training, incorrect AI system use can cause stress in employees			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Prediction accuracy	Number of employees undertaking courses from the recommended list	Improve accuracy
AI features	Task(s)	Recommendation		
	Method(s)	Deep learning		
	Hardware	GPU enabled servers		
	Topology	GPU enabled servers		
	Terms and concepts used	Deep learning, Unsupervised learning, Recommendation		
Standardization opportunities/ requirements	unsupervised sequence mining algorithms to mine the past data			
Challenges and issues	Need large amounts of training data; predicting human behaviour is tricky			
Societal Concerns	Description	Employees may feel challenged or demoralized		
	SDGs to be achieved	Decent work and economic growth		

A.23.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Journal				Tata Consultancy Services Limited	R. Srivastava, G.K. Palshikar, S. Chaurasia, A. Dixit, What's Next? A Recommendation System for Industrial Training, accepted in Data Science and Engineering journal (Springer).
2	Conference				Tata Consultancy Services Limited	R. Srivastava, G.K.Palshikar, S.Chaurasia, What's Next? A Recommendation System for Industrial Training, Proc. of Workshop on Human Capital Management, held as part of International Conference on Data Management (ICDM 2017), New Orleans, USA, 18–21 November, 2017.

3	Conference				Tata Consultancy Services Limited	R. Srivastava, S. Hingmire, G. K. Palshikar, S. Chaurasia, A. Dixit, CSRS: A Context and Sequence Aware Recommendation System, 8th Meeting of the Forum for Information Retrieval Evaluation (FIRE 2016), 7 – 10 December 2016, Kolkata, India.

A.24 AI Solution to Predict Post-Operative Visual Acuity for LASIK Surgeries

A.24.1 General

ID	24	
Use case name	AI solution to predict Post-Operative Visual Acuity for LASIK Surgeries	
Application domain	Healthcare	
Deployment model	Cloud services	
Status	In operation	
Scope	Predicting Post-Operative Visual Acuity for LASIK Surgeries from retrospective LASIK surgery data with patient follow-ups.	
Objective(s)	Given: Pre-operative examination results and demography information about a patient. Predict: Post-operative UCVA after one day, one week and one month of the surgery.	
Narrative	Short description (not more than 150 words)	LASIK (Laser-Assisted in SItu Keratomileusis) surgeries have been quite popular for treatment of myopia, hyperopia and astigmatism over the past two decades. In the past decade, over 10 million LASIK procedures had been performed in the United States alone with an average cost of approximately $2000 USD per surgery. While 99% of such surgeries are successful, the commonest side effect is a residual refractive error and poor uncorrected visual acuity (UCVA). In this work, we aim at predicting the UCVA post LASIK surgery. We model the task as a regression problem and use the patient demography and pre-operative examination details as features. To the best of our knowledge, this is the first work to systematically explore this critical problem using machine learning methods. Further, LASIK surgery settings are often determined by practitioners using manually designed rules. We explore the possibility of determining such settings automatically to optimize for the best post-operative UCVA by including such settings as features in our regression model. Our experiments on a dataset of 791 surgeries provides an RMSE (root mean square error) of 0.102, 0.094 and 0.074 for the predicted post-operative UCVA after one day, one week and one month of the surgery respectively.
	Complete description	**Introduction to LASIK surgeries**

		Refractive surgeries for eye are performed to correct (normalize) the refractive state of the eye, to decrease or eliminate dependency on glasses or contact lenses. This can include various methods of surgical remodelling of the cornea or cataract surgery. LASIK is a refractive eye surgery that uses a laser to correct nearsightedness, farsightedness, and/or astigmatism. In LASIK, a thin flap in the cornea is created using either a microkeratome blade or a femto-second laser. The surgeon folds back the flap, then removes some corneal tissue underneath using a laser. The flap is then laid back in place, covering the area where the corneal tissue was removed. With nearsighted people, the goal of LASIK is to flatten the steep cornea; with farsighted people, a steeper cornea is desired. LASIK can also correct astigmatism by smoothing an irregular cornea into a more normal shape. LASIK surgeries are highly popular; over 10 million LASIK procedures have been performed in the United States alone in the past decade. **Motivation** While overall patient satisfaction rates after primary LASIK surgery have been around 95%, it may not be recommended for everybody for two reasons: (1) high cost with potentially no significant improvement for certain types of patients, and (2) possible eye complications after the surgery. LASIK surgeries cost approximately $2000 USD per surgery. An ability to predict post-operative UCVA can help patients make an informed decision about investing their money in undergoing a LASIK surgery or not. It can also help surgeons recommend the most promising type of laser surgery to the patients. How can we perform this prediction? Further, while performing such surgeries, surgeons need to set multiple parameters like suction time, flap and hinge details, etc. These are often set using manually designed rules. Can we design a data driven automated method to suggest the best settings for a patient undergoing a laser surgery of a certain type? **Problem Definition** In this paper, we address the following problem. <u>Given</u>: Pre-operative examination results and demography information about a patient <u>Predict</u>: Post-operative UCVA after one day, one week and one month of the surgery. <u>Challenges</u> The problem is challenging because (1) large amount of data about such surgeries is not easily available; (2) there are a lot of pre-operative measurements that can be used as signals; and (3) data is sparse, i.e., there are a lot of missing values. **Brief Overview of our Approach**

<table>
<tr><td></td><td></td><td colspan="3">We model the task as a regression problem. We use domain knowledge to preprocess data by transforming a few categorical features into binary features.We also use average values to impute missing values for numeric features. For categorical features, we impute missing values using the most frequent value for the feature. We evaluate multiple regression approaches. Our experiments on a dataset of 791 surgeries provides an RMSE of 0.102, 0.094 and 0.074 for the predicted post-operative UCVA after one day, one week and one month of the surgery respectively.

Summary
– We described a critical problem of predicting post-operative UCVA for patients undergoing LASIK surgeries.
– We modeled the task as a regression problem. We explored the effectiveness of demographic, pre-operative features and surgery settings for the prediction task.
– Using a dataset of 791 LASIK surgeries performed on 404 patients from 2013 and 2014, we tested the effectiveness of the machine learning methods.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Hospitals, Patients undergoing LASIK surgeries.</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4"></td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">different sources of bias; incorrect AI system use</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Recommendation</td><td>The system can be used to automatically recommend the right LASIK surgery to the patient.</td><td>New use-case in healthcare</td></tr>
<tr><td>2</td><td>Improve accuracy</td><td>We found the accuracy of the model to be reasonably good to be practically useful.</td><td>Improve accuracy</td></tr>
<tr><td rowspan="4">AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine Learning, Gradient Boosted Decision Trees Based Regression</td></tr>
<tr><td>Hardware</td><td colspan="3">Machine with 1 CPU and 2 GB RAM. Any Operating system.</td></tr>
<tr><td>Topology</td><td colspan="3">LASIK surgeries, UCVA, Uncorrected visual acuity, Regression</td></tr>
</table>

	Terms and concepts used	
Standardization opportunities/ requirements		
Challenges and issues	The problem is challenging because (1) large amount of data about such surgeries is not easily available; (2) there are a lot of pre-operative measurements that can be used as signals; and (3) data is sparse, i.e., there are a lot of missing values.	
Societal Concerns	Description	
	SDGs to be achieved	Good health and well-being for people

A.24.2 Data

Data characteristics	
Description	The dataset contains information for 404 patients in the age range of 18 to 47 years. 215 of these patients are females, and the rest are males. The 791 LASIK surgeries were done in 2013 and 2014. 397 of the surgeries were performed on the left eye and remaining ones on the right eye. Most of the surgeries are either of the Wavefrontguided-LASIK type or of the Plano-scan-LASIK type. Orbscan is the most popular topography machine used; Oculyzer being the second most popular one. Pre-operative UCVA values vary between 0.15 and 2. Post-operative UCVA values vary between - 0.2 and 1 for day 1, -0.3 and 1 for week 1 and -0.2 and 0.95 for month 1 after the operation. Although usually large datasets improve accuracy of the learned machine learning models, it is difficult to obtain large datasets in this domain.
Source	Measured using various medical machines at the LVPEI Eye Institute, Hyderabad, India.
Type	Structured Data
Volume (size)	791 instances from 404 patients.
Velocity	Batch.
Variety	Single source. Data from multiple centers of the hospital.
Variability (rate of change)	Static.
Quality	Contains some noise. High quality after pre-processing.

A.24.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Pre-processing	Remove unnecessary, noisy, redundant columns. Impute missing	As soon as raw dataset arrives		Pre-processed clean data is ready.

		values. Remove outliers.			
2	Training	Train a model with training samples	Pre-processed clean data is ready.	Pre-processing	Trained regression model
3	Evaluation	Evaluate whether the trained model is of good accuracy	Completion of training/re-training	Training/re-training	Accuracy values
4	Prediction/ Deployment	Test new instances using the trained model	When a new patient visits the hospital for LASIK surgery	Training/re-training	Prediction of post-LASIK surgery outcomes
5	Retraining	Retrain model with more training samples.	Certain period of time has passed since last training/re training and more training samples are available	Pre-processing	Retrained regression model.

A.24.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample Raw data is ready	Pre-processing	AI Cloud Service Provider	Outlier detection, feature selection, missing value imputation	API to perform pre-processsing
2	Completion of step 1	Training sample creation	AI Cloud Service Provider	Create training samples by clearly recognizing relevant features and training label for data from step 1	
3	Completion of step 2	Model training	AI Cloud Service Provider	Train a gradient boosted trees based regression model using	

				training samples from step 2.	

Specification of training data	

A.24.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	New patient visits hospital for LASIK surgery	Pre-processing	AI Cloud Service Provider	Get relevant data from various machines based on patient registration form, and do pre-processing.	
2	Completion of Step 1	Prediction	AI Cloud Service Provider	Given pre-processed instances from step 1 and the trained model, compute predictions for the current patient.	
3	Completion of Step 2	Evaluation	AI Cloud Service Provider	Compare the result of Step 2 with that of the results after surgery.	

Input of evaluation	
Output of evaluation	

A.24.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	New patient comes in	Pre-processing	Hospital	Pre-process input data from patient	
2	Completion of step 1	Prediction	AI Cloud Service Provider	Hospital uses the model hosted on the cloud to predict post-surgery results for the patient based	

				on input from step 1	
3	Completion of step 2	Consultation and surgery recommendation	Hospital	Based on results for various types of LASIK surgeries from step 2, suggest the best suitable surgery to patient.	

Input of Execution	
Output of Execution	

A.24.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Certain period of time has passed since the last training/retraining	Pre-processing	AI Cloud Service Provider	Outlier detection, feature selection, missing value imputation	API/software to perform pre-processing
2	Completion of step 1	Training sample creation	AI Cloud Service Provider	Create training samples by clearly recognizing relevant features and training label for data from step 1	
3	Completion of step 2	Model training	AI Cloud Service Provider	Train a gradient boosted trees based regression model using training samples from step 2.	

Specification of retraining data	

A.24.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Research Paper	LASIK surgery prediction	Published	High	Microsoft, LVPEI	https://link.springer.com/chapter/10.1007/978-3-319-31753-3_39
2	Keynote video snip	LASIK surgery prediction	Available Online	High	Microsoft	https://www.youtube.com/watch?v=mmDz7cwC7CE&t=128s
3	Related Paper	Visual Acuity Prediction	Published	Medium	Visx Inc, Sunnyvale, Calif.	https://www.ncbi.nlm.nih.gov/pubmed/1450116
4	Related Paper	Visual Acuity Prediction for Children	Published	Medium	Department of Ophthalmology, University of Minnesota, Minneapolis, USA.	https://www.ncbi.nlm.nih.gov/pubmed/8965225

A.25 Use of robotic solution for traffic policing and control

A.25.1 General

ID	25	
Use case name	Use of robotic solution for traffic policing and control	
Application domain	Security	
Deployment model	On-premise systems	
Status	PoC	
Scope	Robotics based traffic policing system	
Objective(s)	Efficient traffic control through use of Humanoid robots for traffic control.	
Narrative	Short description (not more than 150 words)	Creation of a humanoid robot which can be deployed for traffic monitoring and control on roads. The solution will use computer vision and will be enabled with IOT for centralized control and data collection. This will relieve the human police from working in polluted environment.
	Complete description	Traffic police needs to stand for long hours in polluted environment which creates stress, other health related issues and may reduce his performance. A humanoid robot equipped with computer vision and IOT can be effectively deployed for effective traffic control. A robotic system can work continuously without any fatigue.

		This system will be centrally controlled and real time data collected can be used to bring efficiency in traffic control.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy of Instructions	The instructions provided by the robot for controlling traffic on various roads.	The controlling instructions should be accurate as per specific traffic conditions.
	2	Response Time	The response required to react to changing traffic condition.	Response time should be minimal (real time) for effective traffic control.
	3	Data collection & control	The robotic system should accurately collect various traffic conditions such as number of vehicles, speed etc. for effective control	The traffic data collected should be accurate for generation of effective control instructions.
AI features	Task(s)	Recommendation		
	Method(s)	Machine Learning, Statistics, Heuristics, Anomaly Detection (Distance / Density based). Artificial Intelligence, Machine Learning, Statistics, Heuristics,Anomaly Detection, Pattern recognition, Computer Vision		
	Hardware	IoT enabled and AI powered Humanoid robots.		
	Topology			
	Terms and concepts used	Automation, Machine Learning, Computer Vision		
Standardization opportunities/ requirements				
Challenges and issues	The problem is challenging because accurate control instructions is crucial for proper traffic control.			
Societal Concerns	Description	Addresses the pressing concern of effective traffic control.		
	SDGs to be achieved	Sustainable cities and communities		

3440

A.25.2 References

[1] J. Zhang, T. Gao, Z. G. Liu, Traffic Video Based Cross Road Violation Detection, In Proc, 2009, International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Vol.3, pages 645-648, April 2009.

[2] D. W. Lim, S. H. Choi, J. S. Jun, Automated detection of all kinds of violations at a street intersection using real time individual vehicle tracking, Image Analysis and Interpretation, 2002. Proceedings. 5th IEEE Southwest Symposium, pages 126-129, 2002.

[3] Y. Chen, C. Yang, Vehicle red-light violation detection base on region, Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference, Vol. 9, pages 700-703, July 2010.

[4] P. KaewTraKulPong and R. Bowden, An improved adaptive background mixture model for real time tracking with shadow detection, In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, Sept 2001.

A.26 Robotic Solution for Replacing Human Labour in Hazardous Condition

A.26.1 General

ID	26	
Use case name	Robotic solution for replacing human labour in Hazardous condition	
Application domain	Security	
Deployment model	On-premise systems	
Status	PoC	
Scope	Building an AI based robotics solution for replacing Human Labour in Hazardous condition	
Objective(s)	Offer AI based robotic solution which can be customized to work in different kind of Hazardous work environment such as Mines, Blast Furnaces, Boilers etc.	
Narrative	Short description (not more than 150 words)	Building an AI based robotic solution enabled with computer vision and equipped with various sensors such as temperature, pressure, smoke detector etc which can effectively replace human labour in risky work environment.
	Complete description	Human labour in Hazardous work environment causes many accidents and loss of life, recent example being NTPC incident that occurred in November 2017 in Unchahar power plant. Working under hazardous conditions also create other serious health related problems including cancer, Asthama etc An AI based robotic system in line with Industry 4.0 fusing technology based automation in manufacturing can replace human labour in hazardous condition and can work efficiently. This also has the potential to reduce incidents caused by human mistakes.
Stakeholders		

Stakeholders' assets, values					
System's threats & vulnerabilities					
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives	
	1	Response Time	Response time required to react to work environment	Response time should be minimal (real time), so that the robotic system can intelligently react in changing work environment.	
	2	Minimum Overshoot	The movement of robotic physical system beyond the intended position should be minimum, ideally zero.	This will enable the robotic system to work accurately in the work environment.	
	3	Reliability	The robotic system should be extremely reliable to avoid any catastrophic failure in the industry. The system should continuously monitor the fitness of its software and hardware component and must have adequate redundancy. It should be able to generate alarm before failure.	Industrial grade robotic solution should be extremely reliable.	
AI features	Task(s)	Automation			
	Method(s)	Artificial Intelligence, Machine Learning, Statistics, Anomaly Detection, Computer Vision			
	Hardware	Robotic Hands, Centralized monitoring and control,			
	Topology				
	Terms and concepts used	Automation, Computer Vision, Reinforced Learning			
Standardization opportunities/ requirements					

Challenges and issues	The problem is challenging because 1. **Solution should be customizable for different work environments**	
Societal Concerns	Description	Addresses the issue of accidents in Hazardous work environment.
	SDGs to be achieved	Decent work and economic growth

A.26.2 References

[1]J. Zhang, T. Gao, Z. G. Liu, Traffic Video Based Cross Road Violation Detection, In Proc, 2009, International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Vol.3, pages 645-648, April 2009.

[2] D. W. Lim, S. H. Choi, J. S. Jun, Automated detection of all kinds of violations at a street intersection using real time individual vehicle tracking, Image Analysis and Interpretation, 2002. Proceedings. 5th IEEE Southwest Symposium, pages 126-129, 2002.

[3] Y. Chen, C. Yang, Vehicle red-light violation detection base on region, Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference, Vol. 9, pages 700-703, July 2010.

[4] P. KaewTraKulPong and R. Bowden, An improved adaptive background mixture model for real time tracking with shadow detection, In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, Sept 2001.

A.27 Credit Scoring using KYC Data

A.27.1 General

ID	27	
Use case name	Credit scoring using KYC data	
Application domain	Banking and Financial Services	
Deployment model	On-premise systems	
Status	PoC	
Scope	Building a risk scorecard for loan applicants using KYC data for better risk management and high population coverage	
Objective(s)	Assigning a risk score to every loan applicant in real time, using just KYC data, which will ensure both new-to-credit and mature customers can be assessed for their creditworthiness, and offered loans on appropriate terms	
Narrative	Short description (not more than 150 words)	It can be often difficult to build a risk scorecard using only KYC data, which often has noisiness and incompleteness issues. However if realized, it can be used to provide a objective score to all loan applicants, even the new-to-credit ones. Non-linear classification algorithms are suitable for this purpose. Several variables are collected from the customer during the KYC process such as Age of customer, Self-reported income, Type of Occupation, Purpose of loan, etc. All these features can be added to a non-linear risk model and their complex interactions allowed to take place.

	Complete description	Financial institutions find it much easier to assess customers with an existing credit history, or those living in urban areas. There are also several credit bureaus who assist them in this endeavor. However, these frameworks don't work as well for new-to-credit customers, especially in rural areas. If only industry wide models or simple heuristics are used to score such customers, many deserving loan applicants will end up not getting a loan or not getting it at deserving terms. Instead, if a good risk scorecard is built using KYC data, which is collected from every loan applicant as a routine and regulated process, it will ensure every applicant receives an objective score. To tackle this problem, non-linear models such as Random Forest and XGBoost are being used which can accommodate many parameters, including categorical ones, and are reasonably resistant to noise in the data.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Delinquency Rate	Percentage of loan defaulters in first X months from loan disbursal vs score bins	Large monotonous decrease in delinquency rate as creditworthiness score increases is desirable, and indicates a good scorecard
	2	Approval rate	Ratio of loan disbursals to loan applicants	Larger approval rate at a predetermined risk level is desirable and indicates a good scorecard
AI features	Task(s)	Credit Scoring		
	Method(s)	Random Forest, XGBoost and Ensemble models		
	Hardware	64 GB RAM, Intel Core i5		
	Topology			
	Terms and concepts used	Classification, Bagging, Boosting, Ensembles		
Standardization				

opportunities/ requirements		
Challenges and issues	1. KYC data obtained from extreme rural areas can be noisy, may have several missing values, and needs appropriate preprocessing and treatment before feeding to the model algorithm 2. Non-linear models like Random Forest and XGBoost need significant computational power during the training phase	
Societal Concerns	Description	We don’t see any societal concerns if it is used
	SDGs to be achieved	

A.27.2 Data

Data characteristics	
Description	Historical KYC data available in internal systems
Source	EDW (Enterprise Data Warehouses)
Type	Structured Data
Volume (size)	10 GB
Velocity	One-time data dump during training phase, real time in production phase
Variety	Mostly Structured
Variability (rate of change)	Moderate
Quality	Moderate

A.27.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	[Breiman 01] Leo Breiman. “Random Forests”. Machine Learning, Volume 45, Issue 1, Pages 5-32. 2001.	Published	High	University of California, Berkeley	https://dl.acm.org/citation.cfm?id=570182
2	Paper	[Chen 16]. Tianqi Chen. “XGBoost: A Scalable Tree Boosting System”. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Pages 785- 794. 2016.	Published	High	University OF Washington, Seattle	https://dl.acm.org/citation.cfm?id=2939785
3	Paper	[Opitz 99]. David Opitz. “Popular ensemble methods: an empirical study”. Journal of	Published	High	University Of Montana, Missoula, MT	https://dl.acm.org/citation.cfm?id=3013549

		Artificial Intelligence Research. Volume 11 Issue 1, Pages 169-198.1999.				

A.28 Recommendation Algorithm for Improving Member Experience and Discoverability of Resorts in the Booking Portal of a Hotel Chain

A.28.1 General

ID	28	
Use case name	Recommendation algorithm for improving member experience and discoverability of resorts in the booking portal of a hotel chain	
Application domain	Leisure and Hospitality	
Deployment model	Cloud services	
Status	In operation	
Scope	Building a personalized recommendation algorithm to help members of the hotel chain to find their desirable hotel for the family holiday	
Objective(s)	Offering personalized recommendations by understanding the member preferences from past holiday patterns and searches in the booking portal. Various member and hotel features were also considered for the model	
Narrative	Short description (not more than 150 words)	Refining existing system and implement a new model that can give personalized recommendations to members and improve bookings at the undiscoverable or not-so-popular hotels. The algorithm would help in reshaping the demand and increase the visibility of the hotels which are at the lower spectrum of demand. We would include member and resort features along with interaction data like members visiting a hotel, and giving a rating to a resort visit etc
	Complete description	The traditional search engine in member portal for booking a hotel is mainly based on the members limited visibility and knowledge of popular holiday destinations. In contrast, a hotel chain might offer a variety of options to members. Each option brings a different holiday experience and possibly include a lot of activities for family members to choose from. In the absence of an intelligent algorithm, many good hotels will be invisible in the large number of hotel lists. This will in turn also increase the burden on some popular hotels which might get disproportionally high bookings, and sometimes run in overcapacity and depriving other hotels of their share of bookings. To solve for this problem, the hybrid recommendation algorithm will help shape the demand and bring up the hotels which are similar to the ones a member has already visited but yet provide a different experience, thus

		encouraging the member to consider an alternative to their usual preferences.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Occupancy %	Percentage of room nights occupied in a hotel	Occupancy in low demand hotels will improve
	2	First time Refusal Rate	Bookings denied because of overdemand in a particular resort	First time refusals will go down
AI features	Task(s)	Recommendation		
	Method(s)	Matrix Factorization and Hybrid Approach		
	Hardware	16 GB RAM, Intel Core i5		
	Topology			
	Terms and concepts used	Matrix Factorization, LightFM, Item and User Features, Latent Features		
Standardization opportunities/ requirements				
Challenges and issues	1. Cold Start Problem: Since the member has only visited certain hotels in the past, the interaction matrix is very sparse 2. The matrix computation at times is computational resource intensive causing system failures			
Societal Concerns	Description	We don't see any societal concerns if it is used		
	SDGs to be achieved			

A.28.2 Data

Data characteristics	
Description	Member Visit Data from booking portals
Source	EDW (Enterprise Data Warehouses)
Type	Structured Data
Volume (size)	1 GB
Velocity	Weekly
Variety	Mostly Structured
Variability (rate of change)	Moderate
Quality	Moderate

A.28.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/ organization	Link
1	Paper	[Kula 15] “Metadata embeddings for user and item cold-start recommendations”. In Proceedings of the 2nd Workshop on New Trends on Content-Based Recommender Systems co- located with 9th ACM Conference on Recommender Systems (RecSys 2015), Vienna, Austria, September 16--20, 2015., pages 14--21, 2015.	Published	High	ACM	https://arxiv.org/abs/1507.08439
2	Paper	[Adomavicius et. al 05]. “Toward the next generation of recommender systems: A survey of the state-of-the- art and possible extensions”. Knowledge and Data Engineering, IEEE Transactions on. 17. 734-749.10.1109/TKDE.2005.99.	Published	Medium	IEEE	https://dl.acm.org/citation.cfm?id=2959 160
3	Paper	Yehuda et. al 09], “Matrix Factorization Techniques for Recommender Systems”, Computer, v.42 n.8, p.30- 37, August 2009 [doi>10.1109/MC.2009.263]	Published	Medium	IEEE	https://dl.acm.org/citation.cfm?id=1608 614

A.29 Enhancing traffic management efficiency and infraction detection accuracy with AI technologies

A.29.1 General

ID	29
Use case name	Enhancing traffic management efficiency and infraction detection accuracy with AI technologies

Application domain	Transportation	
Deployment model	Hybrid or other (please specify) Cloud services and on-premise systems	
Status	In operation	
Scope	Utilizing AI technologies in traffic monitoring and management	
Objective(s)	To increase the accuracy and efficiency of infraction detection, traffic monitoring and flow analysis, while minimizing the human effort and the overall solution cost.	
Narrative	Short description (not more than 150 words)	Big data enabled AI technologies are applied to monitoring and managing the traffic in a large municipality in China. Multi-sourced data (traffic flow, vehicle data, pedestrian movement, etc.) is monitored, from which illegal operation of vehicles, unexpected incidents, surge of traffic etc. are detected and analysed with machine learning (ML) methods. ML tasks (including training and deployment) are carried out on a platform supporting the integration of various ML frameworks, models and algorithms. The platform is based on heterogeneous computing resources. The efficiency and accuracy of infraction detection, and the effectiveness of traffic management are significantly improved, with much reduced human effort and overall solution cost.
	Complete description	With the population and the number of vehicles growing in large cities, managing the heavy traffic in urban areas has become a challenging yet essential task for the municipality. Addressing this issue has become particularly urgent for big cities in China, where millions of people live and commute every day. In this use case, big data based AI technologies are applied to monitoring and managing the heavy traffic in a metropolitan in south China. Previously, significant human resources were involved in the vehicle and road monitoring, and large investment was made to the computing infrastructure specific to certain functionalities. To increase the efficiency of urban transportation, reduce the traffic jam and air pollution, as well as minimize the human effort, machine learning techniques (e.g. deep learning) are applied to image and video analysis, such as traffic flow analysis, infraction detection and incident detection. Example applications include but not limited to 1) detection of traffic rule violation, e.g. over-speeding, wrong driving lanes or parking. AI-enabled detection produces much faster and more accurate result, and helps in enforcing the traffic regulation. 2) traffic light optimization. Based on the modelling and analysis of multi-sourced traffic information (both real-time and historical data), traffic lights are dynamically configured to divert the flow, increase the passing speed of cars and reduce the traffic jam in major junctions. The use of AI has obtained remarkable results: The infraction detection efficiency gets 10X increase, and the

		detection accuracy is greater than 95%. The urban area traffic jam is much alleviated, with vehicles' passing speed through major junctions increases by 9%-25%.		
Stakeholders	Urban citizens (drivers and pedestrians), government, car companies, traffic administrative bureaus, logistics companies, etc.			
Stakeholders' assets, values	Transportation efficiency, controlability and predictability of commute time, pedestrian and vehicle safety, air quality, etc.			
System's threats & vulnerabilities	Low quality pictures, insufficient processing capability			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	accuracy	The accuracy of infraction and incident detection from traffic pictures/videos	To increase the accuracy of traffic monitoring and inspection
	2	split	Proportion of images requiring human inspection. The less the split, the higher the efficiency.	To minimize the human effort in inspection
	3	resource utilization ratio	Achievable resource utilization ratio in the hardware infrastructure (the higher the utilization ratio, the lower amount the required resource)	To reduce the infrastructure investment and overall solution cost
AI features	Task(s)	Recognition		
	Method(s)	Machine learning, Deep learning		
	Hardware	Heterogeneous computing platform (CPU plus heterogeneous accelerators such as GPU, FPGA etc.)		
	Topology			
	Terms and concepts used	Heterogeneous resource pooling, on-demand resource scheduling		
Standardization opportunities/ requirements	• Requirement of computing infrastructure to empower AI applications in the transportation domain, e.g. the integration of acceleration units (GPU, FPGA, etc.), dynamic scheduling and on-demand allocation of heterogeneous resources • Support of mainstream ML frameworks, and the algorithms and models from different vendors, to prevent vendor lock-in			
Challenges and issues	• Constant improvement in hardware architecture to increase the performance and efficiency of running ML/DL tasks • Consistent interfaces between applications, ML engines and heterogeneous resource pools Support of new models and emerging algorithms for growing functionalities			

Societal Concerns	Description	AI's application in urban transportation significantly improves the quality of life for urban citizens, reduces the time wasted in heavy traffic and the air pollution from vehicles.
	SDGs to be achieved	Sustainable cities and communities

3489

3490 A.29.2 **Data**

Data characteristics	
Description	Traffic data (vehicle, road, and pedestrian data)
Source	Traffic camera
Type	Image, video
Volume (size)	~100TB/day
Velocity	Stream and batch
Variety	Traffic flows, vehicle information, pedestrian information, etc.
Variability (rate of change)	Subject to random surge (rush hour, accident, etc.)
Quality	Vary (depending on the weather condition, environment etc.)

3491

3492 A.29.3 **Process scenario**

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (e.g. neural network) with training samples	Sample raw dataset is ready		
2	Evaluation	Evaluate whether the model is properly trained for the detection	Completion of training/re training		Meeting KPI requirements (e.g. accuracy, split) of the particular case
3	Execution	Deploy the model for infraction detection and traffic analysis	Traffic image/vide o data is applied.	The model has been evaluated as properly trained.	
4	Retraining	Retrain a model with training samples	Changes in dataset pattern is expected, or new requiremen		

			t on detection.		

A.29.4 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Journal		Published online		Huawei Technologies Co.,Ltd.	https://www.huaweicloud.com/journal/detail_09.html

A.30 Autonomous Network and Automation Level Definition

A.30.1 General

ID	30	
Use case name	Autonomous network and automation level definition	
Application domain	ICT	
Deployment model	Cyber-physical systems	
Status	PoC	
Scope	Communications network	
Objective(s)	To define autonomous network concept and automation level for the common understanding and consensus	
Narrative	Short description (not more than 150 words)	With the goal of providing common understanding and consensus for autonomous self-driving network, this use case delivers a harmonized classification system and supporting definitions that: • Define the concept of autonomous network • Identify six levels of network automation from “no automation” to “full automation”. • Base definitions and levels on functional aspects of technology. • Describe categorical distinctions for a step-wise progression through the levels. • Educate a wider community by clarifying for each level what role (if any) operators have in performing the dynamic network operations task while a network automation system is engaged.
	Complete description	The telecom CSPs have a dual challenge – to increase agility while reducing network operating cost. 1) The exponential growth of network complexity e.g. 5G will make the traditional network O&M model unsustainable; 2) Digital transformation accelerates service innovation but requires automation capabilities. As CSPs start to evaluate their digital transformation strategies, automation is a central concern. Some operators are already introducing automation to some of their network

		processes, most commonly O&M, planning and optimization. According to Analysys Mason, in 2018, 56% of CSPs globally have little or no automation in their networks. But by 2025, according to their own predictions, almost 80% expect to have automated 40% or more of their network operations, and one-third will have automated over 80%. The introduction of AI/ML (artificial intelligence/machine learning) will be an important part of that process for many CSPs, helping to make the network more intelligent, agile and predictive.

The autonomous self-driving network has two essential elements in common with the autonomous self-driving car:

- There are different levels of automation, relating to different timescales and scenarios
- Intensive use of artificial intelligence (AI) is essential

With the goal of providing common understanding and consensus for autonomous self driving network, this use case delivers a harmonized classification system and supporting definitions that set out six levels of automation for the network.

Level	Name	Definition	Execution (Hands)	Awareness (Eyes)	Decision (Minds)	Experienc (Hearts)
0	**Manual Operation & Maintenance**	Even with auxiliary tools, O&M personnel perform all dynamic tasks.	P	P	P	P
1	**Assisted Operation & Maintenance**	Under the applicable design scope, the system can execute a sub-task repeatedly based on rules.	P/S	P	P	P
2	**Partial Autonomous**	Under the applicable design scope, the system	**S**	P	P	P

			Network	continuously completes the control task of a unit based on the model.					
		3	**Conditional Autonomous Network**	Under the applicable design scope, the system can implement complete closed-loop automation of single-domain scenarios. Users can respond to the requests in a timely manner when the system fails.	S	**S**	P	P	**Domain level**
		4	**Highly Autonomous Network**	Under the applicable design scope, the system can automatically analyze and execute cross-domain and service close-loop automation.	S	S	**P**	P	**Service level**
		5	**Full Autonomous Network**	The system can perform complete dynamic tasks and exception handling in all network environments. O&M personnel do not need	S	S	S	**P/S**	**All Mode**

		to intervene.

P=Personnel (Manual), S=System (Automated)

-Level 0 - manual O&M: The system delivers assisted monitoring capabilities, which means all dynamic tasks have to be executed manually.

-Level 1 - assisted O&M: The system executes a certain sub-task based on existing rules to increase execution efficiency.

-Level 2 - partial autonomous network: The system enables closed-loop O&M for certain units under certain external environments, lowering the bar for personnel experience and skills.

-Level 3 - conditional autonomous network: Building on L2 capabilities, the system can sense real-time environmental changes, and in certain domains, optimize and adjust itself to the external environment to enable intent-based closed-loop management.

-Level 4 - highly autonomous network: Building on L3 capabilities, the system enables, in a more complicated cross-domain environment, predictive or active closed-loop management of service and customer experience-driven networks. This allows operators to resolve network faults prior to customer complaints, reduce service outages and customer complaints, and ultimately, improve customer satisfaction.

-Level 5 - full autonomous network: This level is the ultimate goal for telecom network evolution. The system possesses closed-loop automation capabilities across multiple services, multiple domains, and the entire lifecycle, achieving autonomous driving networks.

The lower levels can be applied now and deliver immediate cost and agility benefits in certain scenarios. An operator can then evolve to the higher levels, gaining additional benefits and addressing a wider range of scenarios.

Network automation is a long run objective with step-to-step process, from providing an alternative to repetitive execution actions, to performing perception and monitoring of network environment and network device status, making decisions based on multiple factors and policies, and providing effective perception of end user experience. The system capability also starts from some service scenarios and covers all service scenarios.

Stakeholders	Communications Service Providers, Suppliers, Industrial and consumer users			
Stakeholders' assets, values	Efficiency; productivity; competitiveness; safety; privacy; availability; experience			
System's threats & vulnerabilities	incorrect AI system use			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Automation level	See the definition in the table	
	2	Accuracy	Predictive & prescriptive decision making&reasoning	
	3	Real-time	The relative response time meets the requirements of operations	
AI features	Task(s)	Other (please specify): All		
	Method(s)	Machine learning, deep learning, Knowledge graph, decision making&reasoning, analytics		
	Hardware	AI training and inference system, and network management system		
	Topology	End-to-end		
	Terms and concepts used	Autonomous network, self-driving network		
Standardization opportunities/ requirements	To standardize autonomous network and automation level			
Challenges and issues	Data usage and sharing, human expertise & competence			
Societal Concerns	Description	None		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.31 Autonomous network scenarios

A.31.1 General

ID	31
Use case name	Autonomous network scenarios
Application domain	ICT

<table>
<tr><td>Deployment model</td><td colspan="2">Cyber-physical systems</td></tr>
<tr><td>Status</td><td colspan="2">PoC</td></tr>
<tr><td>Scope</td><td colspan="2">Communications network</td></tr>
<tr><td>Objective(s)</td><td colspan="2">Clarification and showcases of autonomous network usage</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>Multiple scenarios of autonomous network enabled by AI is addressed for improving operational efficiency, customer experience and service innovation, including wireless network performance improvement, optical network failure prediction, data center energy saving etc.</td></tr>
<tr><td>Complete description</td><td>The leading reason to adopt AI-assisted network automation is to reduce the cost – almost 80% operators placed this in their top three drivers, followed by:

● improvement to customers' network quality of experience
● efficient planning and management of dense networks
● part of an end-to-end automation strategy spanning the network and IT operations
While OPEX reduction is the most important cost-related driver, others include better alignment of network costs to the revenue that is generated; and the ability to defer some capital expenditure (CAPEX) by using existing assets more efficiently.

Obviously, the autonomous self-driving network needs to move from an O&M approach that is focused on network elements, to one based on usage scenarios. This means that process changes relate directly to a particular result, defined by the operator, and with a business value. Progress will be accelerated if a core set of scenarios is defined, which will be of value to all operators. Development of the related autonomous self driving network solutions can then be prioritized accordingly.

The criteria for the selection of scenarios as follows:

● Extent of digitalization: Reflects the technical readiness of the scenarios. Digitalization is the foundation of automation, and the extent to which it is supported determines the extent to which automation can be achieved immediately;
● TCO contribution: Reflects OPEX savings and the improvement to CAPEX efficiency in the given scenario;
● O&M life cycle: Reflects the ability to build differentiation in each phase of the life cycle in order to achieve full autonomous driving across many scenarios. The O&M life cycle spans planning, deployment, maintenance, optimization and provisioning of the network and scenarios have been identified for each one.
Based on those three criteria, we selected six typical key scenarios for the purpose of illustration and clarification.</td></tr>
</table>

		Scenario 1: Base Station Deployment 1) Definition and Description of Scenario The base station deployment scenario refers to the entire process after site survey, including network planning and design, site design, configuration data preparation, site installation, site commissioning and site acceptance. 2) Automation Classification Level 1: The O&M tool helps some elements of the process to be automated, but configuration and site acceptance have to be done manually. Level 2: Some hardware can be detected and configured automatically, and configuration data is simplified based on rules. Level 3: E2E automation: radio parameter self-planning, hardware self-detection and self-configuration, self-acceptance without dialing test. Initial outcomes: Upon the usage of AI, some initial results are achieved as follows: -Site Deployment Time Shortened by 30% -Feature Deployment Time Shortened by 60% -Performance Converging Shortened by 85% **Scenario 2: Network Performance Monitoring** 1) Definition and Description of Scenario The mobile network has entered the stage of very precise planning sites and resources: on the one hand, to identify and forecast high traffic areas, and allocate resources precisely to support business goals; on the other hand, to identify and forecast high-frequency temporary traffic, scheduling resources to meet business objectives. 2) Automation Classification Level 1: Network quality is consistent, and network anomalies can be discovered by tools; Level 2: 3D presentation of network quality and anomalies, and network planning is self-generated; Level 3: E2E closed-loop monitoring and planning: predicting network development according to historical network information, finding value areas and hidden problems, recommending the best network planning and estimating the gain automatically.

		Scenario 3: Fault Analysis and Handling 1) Definition and Description of Scenario The security and reliability is the most important mission of the network, so quick alarm detection and quick fault healing are important. The fault analysis and handling scenario comprises several steps, including alarm monitoring, root cause analysis, and fault remediation. Monitoring: Real-time monitoring of network alarm, performance, configuration, user experience, and other information. Analysis: By analyzing the correlation between alarms and other dimensions data, root cause of fault and fault repairing can be achieved quickly. Healing: Repair fault remotely or by site visiting based on the repairing suggestions. 2) Automation Classification Level 1: Some tools are used to simplify alarm processing, but thresholds and alarm correlation rules are set manually based on expert experience. Level 2: Automatic alarm correlation and root cause analysis. Level 3: Closed-loop of alarms analysis and handling process: Based on the intelligent correlation analysis of multi-dimensional data, accurate location of alarm root cause, precise fault ticket dispatching, and fault self-healing could be reached successfully. Level 4: Proactive troubleshooting: Based on the trend analysis of alarms, performance, and network data, alarms and faults could be predicted and rectified in advance. Initial outcomes: Upon the usage of AI, some initial results are achieved as follows: -Reduction of alarms: 90% **Scenario 4: Network Performance Improvement** 1) Definition and Description of Scenario Wireless networks are geographically very distributed, and activity varies significantly in different places and at different times of day. This makes the network very dynamic and complex. That complexity is further increased by the diversity of services and of terminal performance, and by the mobility of users. If the network cannot achieve the benchmark KPIs or SLAs (service level agreements), or

		enable good user experience, it must be adjusted to meet or exceed those requirements. This is the function of network performance improvement or optimization. The complete process of network performance improvement or optimization includes several stages: ● network monitoring and evaluation ● root cause analysis of performance problems ● optimization analysis and optimization decision-making ● optimization implementation ● post- evaluation and verification 2) Automation Classification Level 2: Drive test evaluation is not required for coverage optimization. Adjustment suggestions are provided automatically. Level 3: Closed-loop of network performance improvement: Automatic identification of network coverage and quality problems, automatic configuration of performance parameters, and automatic evaluation. Level 4: Dynamic adjustment is implemented based on the scenario awareness and prediction to achieve the optimal network performance. Network prediction capability is available: scenario change trends could be perceived, and network configuration could adjusted real-time to achieve optimal performance. Initial outcomes: Upon the usage of AI, some initial results are achieved as follows: -Capacity increase: 30%, -Delivery duration: 2 weeks, non-manual **Scenario 5: Site Power Saving** 1) Definition and Description of Scenario T Site power consumption cost accounts for more than 20% of network OPEX. Although network traffic declines greatly during idle hours, equipment continues to operate, and power consumption does not dynamically adjust to the traffic level, resulting in waste. It is necessary to build the "Zero Bit, Zero Watt" capability. 2) Automation Classification Level 2: Tool aided execution; Level 3: Power-saving closed-loop: Based on the analysis of traffic trends, self-adaptive generation of power-saving strategies, effect and closed-loop KPI feedback;

<table>
<tr><td></td><td></td><td colspan="3">Level 4: Real-time adjustment of power-saving strategies based on traffic prediction. Through integration with third-party space-time platforms, the operator can also add predictive perception of traffic changes, smooth out the user experience, and maximize power-saving.

Initial outcomes: Upon the usage of AI, some initial results are achieved as follows:

-Power saving: 10~15%

Scenario 6: Wireless Broadband Service Provisioning

1) Definition and Description of Scenario
WTTx has become a foundational service for mobile operators because of its convenient installation and low cost of single bit. Rapid launch of WTTx service, accurate evaluation after launch, and network development planning have become important supports for new business development.

2) Automation Classification
Level 1: Blind launch;

Level 2: Automation tools to assist the launch, check the coverage and capacity of the user's location before the business hall, and experience evaluation;

Level 3: Closed-loop for business launch: Integrated with BOSS system to achieve one-step precise launch, remote account launching, CPE installation, fault self-diagnosis and complaint analysis;

Level 4: Auto-balancing of multi-service, automatic value areas identification and network planning recommendation.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Communications Service Providers, Suppliers, Industrial and consumer users</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Efficiency; productivity; competitiveness; safety; privacy; availability; experience</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">incorrect AI system use</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Other (please specify) All</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine learning, deep learning, Knowledge graph, decision making&reasoning, analytics</td></tr>
<tr><td>Hardware</td><td colspan="3">AI training and inference system, and network management system</td></tr>
<tr><td>Topology</td><td colspan="3">End-to-end</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Autonomous network, self-driving network</td></tr>
</table>

Standardization opportunities/ requirements		
Challenges and issues		
Societal Concerns	Description	
	SDGs to be achieved	

A.32 AI Solution to Help Mobile Phone to have Better Picture Effect

A.32.1 General

ID	32	
Use case name	AI solution to help mobile phone to have better picture effect	
Application domain	Mobility	
Deployment model	Hybrid or other (please specify)	
Status	In operation	
Scope	Better understanding the image and improving image effect on smartphone by using DL model which is trained in the cloud or offline.	
Objective(s)	To find an efficient solution to Increase camera image quality on smartphone without Increasing too much operation and power burden for mobile phone.	
Narrative	Short description (not more than 150 words)	An AI solution was developed that could increase smartphone camera image quality. Using deep learning, smartphone can Identify more scenarios and objects than before. Based on the identified scenarios and objects, smartphone can better understand the image and improve image effect.
	Complete description	At present, there are 1.4 billion smart phone shipments in the world every year. Photography is one of the most important functions of smart phones. The industry has been trying to improve the picture quality of mobile phone photography. It hopes to reach even the quality of the professional SLR camera. The traditional image processing algorithm is currently facing the ceiling, many scenes traditional algorithms can not be used, just because the effect is very poor. Deep learning algorithm provides a turning point for solving the above problems. By using the AI solution, smartphones can better "understand" the pictures they take. Based on the deep learning algorithm, the smart phone can analyze the shooting scene in real time and intelligently identify various scenes in the shooting process, such as blue sky, flowers, green plants, night view, snow scene, etc. And the smart phone can also intelligently detect the shooting objects in the scene. Base on scene recognition and object detection ,the smartphone can automatically adjust and set parameters for different pictures, so as to get better photo effects.

		Now the mobile phone can recognize 100 kinds of scenes and can reach hundreds in the future. By using the depth learning algorithm, the mobile phone can now detect the 20 types of subjects, and the future can be detected by hundreds of subjects. Object detection can be used for SmartZoom (auto focus on targets), and portrait segmentation can be used for background blur or light efficiency.		
Stakeholders	mobile phone manufacturer、end users、third party testing and evaluation agency			
Stakeholders' assets, values	Competitiveness			
System's threats & vulnerabilities	new privacy threats (hidden patterns).			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	MIoU(Mean Intersection over Union)	The intersection of prediction area and actual area divided by the union of the predicted area and the actual area. Ideal target is 100%.	Improve accuracy
	2	FAR(false acceptance rate)	Negative samples are identified as positive samples / Total number of negative samples.The low FAR, the more smartphone will get correct scenes and objects	Improve accuracy
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware	NPU、GPU、CPU etc.		
	Topology	No Need		
	Terms and concepts used	Deep learning, "Understand"		
Standardization opportunities/ requirements	The standardized content includes: 1) the format of training picture data; 2) the format of deep learning model generated offline or cloud, which will be transplanted to smart phones; 3) the platform to support the transplanted model in the smart phone; 4) API which can be used by others applications, such as: picture classification, security.			
Challenges and issues	Challenges: Achieve the same level as professional SLR camera for pictures. Issues: 1) Lack of data for certain scene;			

	2) Lack of computing ability on terminal side； 3) Users can feel the improvement of image quality, but may not know that it is brought by AI.	
Societal Concerns	Description	For the wrong object detection, it may lead to racial prejudice or privacy protection problems.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.32.2 Data

Data characteristics	
Description	Annotated pictures
Source	Public picture library /Self collection picture library /Web crawling pictures /Automatic synthesis of pictures
Type	Picture format supported by a training platform and smart phone
Volume (size)	
Velocity	
Variety	Single source
Variability (rate of change)	
Quality	

A.33 Automated Defect Classification on Product Surfaces

A.33.1 General

ID	33	
Use case name	Automated defect classification on product surfaces	
Application domain	Manufacturing processes	
Deployment model	On premise system	
Status	PoC	
Scope	Image Analytics for water taps in sanitary industries.	
Objective(s)	Image analytics using a combination of feature extraction and classification of defects on shining surfaces in sanitary industries.	
Narrative	Short description (not more than 150 words)	A vision system that inspects and identifies the defects on water taps in sanitary industries. The system uses a combination of features for an automatic defect classification on product surfaces. All defects (15 types are identified) are classified into two major categories, real-defects and pseudo-defects. The pseudo-defects cause no quality problem; while the real-defects are critical as they might malfunction the final products. The AI system uses Support Vector Machine (SVM) classifier along with the combined features to identify the defect

		types. With the vision system in place, the quality control process is fully automated without any human intervention.
	Complete description	The proposed vision system has two parts: the hardware part and the software part. The hardware captures the images of product surfaces under a constant illuminating condition. The software is developed to perform image processing tasks and identify defects on product surfaces. The steps of proposed system include image acquisition, preprocessing, segmentation, feature extraction, classification and post-processing. The system presents two software components: Feature Extraction and Classifier Design. These two modules are implemented independently which can be developed in offline platform and can be integrated into vision system and work online. As a first step, the feature extraction is critical and guides the extent to which a classifier can distinguish the defects from one class to another. A combination of features is used like geometry (shape, texture), and statistical features of the segmented images. In the second step, a support vector machine classification model is trained to identify the defect types. The classification results obtained by combining Gabor features, Statistical features, and grayscale features showed comparable performances with human evaluations. Overall, the vision system is modularized with capabilities to self-learn and future extensions.

Stakeholders	Sanitary Industries			
Stakeholders' assets, values	Competitiveness; Quality Check;			
System's threats & vulnerabilities	Incorrect AI System use (AI system affecting quality control); New Security Threats.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Classification Ratio	Real to Pseudo wrong classification	Establishes the quality of identification

AI features	Task(s)	Recognition
	Method(s)	Classification; Feature Extraction
	Hardware	IP Camera and Work Station
	Topology	
	Terms and concepts used	Classification, Feature Extraction, Defect Identification
Standardization opportunities/ requirements	1) Quality acceptance criterion from AI systems: What is the acceptable standard for AI output related to quality? How that can be independently validated? 2) Standards for dealing with AI failures: How/Can standards facilitate dealing with AI failures, w.r.t., quality, productivity criteria?	

Challenges and issues	Real time implementation, accurately identify the nature of defects.	
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.33.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	B. Kuhlenkötter, X. Zhang, C. Krewet, Quality Control in Automated Manufacturing Processes – Combined Features for Image Processing Acta Polytechnica Vol. 46 No. 5/2006.	Published	Use case taken from this reference	Czech Technical University	https://ojs.cvut.cz/ojs/index.php/ap/article/view/868

A.34 Robotic Task Automation: Insertion

A.34.1 General

ID	34	
Use case name	Robotic task automation: Insertion	
Application domain	Manufacturing	
Deployment model	Embedded systems – Cloud service	
Status	PoC	
Scope	Robotic assembly	
Objective(s)	1. Simple programing/instruction and flexibility in usage 2. Automation of tasks lacking analytic description 3. Reliability and efficiency	
Narrative	Short description (not more than 150 words)	Assembly process often includes steps where two parts need to be matched and connected to each other through force exertion. In an ideal case, perfectly formed parts can be matched and be assembled together with predefined amount of force. Due to imperfection of production steps, surface imperfection and other factors such as flexibility of parts, this procedure can become complex and unpredictable. In such cases, human operator can be instructed with simple terms and demonstrations and perform the task easily, while a robotic system will need

		very detailed and extensive program instructions to be able to perform the task including required adaptation to the physical world. The need for such a complex program instruction will make use of automation cumbersome or uneconomical. Control algorithm that are based on machine learning, especially those including reinforcement learning can become alternative solutions increasing and extending the level of automation in manufacturing.
	Complete description	The case described here is a common step in assembly processes in manufacturing industry and includes matching and properly connecting two parts when one needs to be inserted into another. Successful and efficient insertion usually needs action by feeling. It is difficult to describe in terms of mathematical algorithms and therefore is difficult to program. Complexities in programming, or high degree of operational failure make usage of robots, or automation unattractive. Use of machine learning and artificial intelligence is one of promising methods to overcome such difficulties. As will be described below, there are several different phases in the process, where different methodologies can and should be used. To make the methodology usable in a practical case, it should be utilizable by operators without deep technical knowledge with an effort that can be accepted on a production line. Ultimately, such methods must remove the need for programing completely. The assumption here is that the parts to be assembled are properly localized, such that they can be manipulated by a robot in the desired way. The problem concerns the following steps: 1. Identification and picking the first part (A). 2. Moving A to the vicinity of the second part (B). 3. Alignment of the two parts. 4. Exertion of force with simultaneous movement for smooth insertion. 5. Termination of the task when complete insertion is complete. The above task, with all possible challenges, can easily be performed by a human operator. An operator in majority of cases needs very limited amount of information. Using prior knowledge and experiences and the sensory system the task can be completed and all possible exceptions can be handled. With time, a human operator becomes constantly more efficient and performs the task faster and more reliably. The topics to be handled in this use case are how a machine can be instructed, trained, perform and improve to a high level of reliability and efficiency. The process can be divided into following steps:

<table>
<tr><td></td><td></td><td colspan="3">1. Localization of parts: Image processing, object identification, classification and localization.
2. Alignment of parts: Control and optimization with (mainly) vision inputs.
3. Insertion through exertion of forces: Control and optimization with (at least) vision and force sensor feedback
4. Sensing the termination of the process: Pattern recognition in time series.
5. Continuous improvement: Reinforcement learning.

Vision and force sensors are most commonly used sensors in such processes. The objects and environment need to be observed at moderate as well as in very close distances. Force sensors are needed but have the weakness of not being active before a complete contact. Therefore, use of other sensors could be helpful.

The method is used for assembly tasks with the target of reducing the programming effort and increasing flexibility. For that to be achieved, the effort necessary to teach, train and use the system should be minimum and the reliability should come high at short time. This implicitly means that the system should become useful with limited amount of data and at limited amount of time. After an initial relatively stable state is reached, reinforcement can be used to improve the efficiency of the system.

The solution will become more attractive if transfer learning is utilized to further reduce the initial training time.

For benchmarking purpose a specific set of objects to be assembled together should be defined and performance of the methods can be measured by necessary training time, need for computing power and memory as well as time for completion of the task. The objects in the tests can be geometrically relatively simple. Special features such as rough surfaces, tight fitting or flexibility of the objects can be considered for different classes of problems.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Discrete manufacturing industries; Operators</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Competitiveness; Productivity</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Incorrect AI system use; New security threats</td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Ease of use</td><td></td><td>Simplicity and efficiency during initial learning. Teaching process should be easy.</td></tr>
</table>

	2	Training efficiency		Amount of necessary data for training might lead to practical obstacles in application.
	3	Initial success rate		After initial training, the success rate needs to be acceptable such that the system can be put in the production line.
	4	Speed of improvement		Higher convergence speed of the reinforcement algorithm is making the solution more attractive.
	5	Operational efficiency		Cycle time is the primary measure in manufacturing industry.
	6	Success rate		Very high success rate is required for the solution to be accepted.
AI features	Task(s)	Recognition, classification, control, optimization		
	Method(s)	Deep learning, image processing, control, Optimization		
	Hardware	PC equipped with GPU accelerators		
	Topology	NA		
	Terms and concepts used	Reinforcement learning		
Standardization opportunities/ requirements	• Standardization of definition of KPIs; • Standardization of fail-safe options w.r.t. safety and quality; • Standardization towards “Human-Co-working” • Minimum acceptable standards for commercialization; • Standard data set to independently validate the claims;			
Challenges and issues	• Complex and unpredictable assembly process due to imperfection of production steps, surface imperfection and other factors such as flexibility of parts. • Accuracy of sensing • Coworking with humans			
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

3512

3513 A.34.2 **References**

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Fan Dai, Arne Wahrburg, Björn Matthias, Hao Ding: Robot Assembly Skills Based on Compliant Motion Proceedings of 47th International Symposium on Robotics (ISR 2016), At Munich, Germany	Published	Cited to support the detailed description	ABB	https://www.researchgate.net/publication/310951674_Robot_Assembly_Skills_Based_on_Compliant_Motion
2	Conference	Te Tang, Hsien-Chun Lin, Masayoshi Tomizuka, A learning-based framework for robot peg-hole-insertion, Proceedings of the ASME 2015 Dynamic Systems and Control Conference, October 28-30, 2015, Columbus, Ohio, USA	Published	Cited to support the detailed description	University of California	https://www.researchgate.net/publication/314634124_A_Learning-Based_Framework_for_Robot_Peg-Hole-Insertion
3	Publication	Fares J. Abu-Dakka, Bojan Nemec, Aljaž Kramberger, Anders Glent Buch, Norbert Krüger and Aleš Ude, Solving peg-in-hole tasks by human demonstration and exception strategies, Industrial Robot: An International Journal 41/6 (2014) 575–584	Published	Cited to support the detailed description	Jožef Stefan Institute, Dept. of Automatics, Biocybernetics, and Robotics, Slovania Maersk Mc-Kinney Moller Institute, University of Southern Denmark	https://www.researchgate.net/publication/273170116_Solving_peg-in-hole_tasks_by_human_demonstration_and_exception_strategies
4	Publication	Mel Vecerik, Todd Hester, Jonathan Scholz, Fumin Wang, Olivier Pietquin, Bilal Piot, Nicolas Heess, Thomas Rothörl, Thomas Lampe, Martin Riedmiller, Leveraging Demonstrations for Deep Reinforcement, Learning on Robotics Problems with Sparse Rewards, arXiv:1707.08817v2 [cs.AI] 8 Oct 2018	Published	Cited to support the detailed description	Deepmind	https://arxiv.org/pdf/1707.08817.pdf
5	Publi	Mel Vecerik, Oleg Sushkov, David Barker, Thomas Roth¨orl, Todd Hester, Jon	Published	Cited to support the	Deepmind	https://arxiv.o

	catio n	Scholz, A Practical Approach to Insertion with Variable Socket Position Using Deep Reinforcement Learning, arXiv:1810.01531v2 [cs.RO] 8 Oct 2018		detailed description		rg/pdf/1810.0 1531.pdf

A.35 Causality-based Thermal Prediction for Data Center

A.35.1 General

ID	35	
Use case name	Causality-based Thermal Prediction for Data Center	
Application domain	Other (data center) Cooling control is data center. This is mainly intended towards reducing energy requirements towards cooling of data centers.	
Deployment model	On-premise systems	
Status	Prototype	
Scope	Data center cooling control involving use of air cooling to control hot spots in data center.	
Objective(s)	Minimize energy usage in managing data center	
Narrative	Short description (not more than 150 words)	Data centers tend to be overcooled to prevent computing machines from failing due to heat. A reliable fine-grained control that could regulate air control unit (ACU) supply air temperature or flow is needed to avoid overcooling. Methods that are based on correlation-based techniques do not generalize well. Hence, we seek to uncover the causal relationship between ACUs supplying cool air and temperature at the cabinets to prioritize which ACUs should be regulated to control a hot-spot near a cabinet.
	Complete description	First, we perform experiments in 6SigmaRoom for the layout of the data center being studied. We collect time-series data for supply air temperature and flow per ACU, and for inlet temperature at the cabinets. Next, we test the recorded time series for checking if Granger-causality (G-causality) can be established between the supply air temperature from an ACU to a cabinet. G-causality establishes the unidirectional temporal precedence for data center control actions from ACUs that leads to changes in specific cabinet temperatures. A variable X is said to Granger-Cause Y if, including data about past terms from X, leads to a better prediction of the future value of Y (i.e., Yt+1) than predicting Yt+1 based solely on past terms from Y. We show by way of simulation that the ACU flows that Granger-Cause reduction in temperature at a cabinet provide a larger share of influence (based on Zone of Influence/Thermal Correlation Index from the simulation) on the cabinet. This could allow an operator to come up

		with a better control strategy to control hotspots in a data center by regulating ACU supply air temperature/flows.		
Stakeholders	Data center owner; Data center users; Environment			
Stakeholders' assets, values	Competitiveness; Reputation; Stability			
System's threats & vulnerabilities	Incorrect AI system use; Security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Zone of Influence/ Thermal Correlation Index	Extent of influence of ACUs on data center racks.	Helps in improved control.
AI features	Task(s)	Prediction		
	Method(s)	Regression		
	Hardware	64 GB RAM Windows server		
	Topology	NA		
	Terms and concepts used	Granger Causality		
Standardization opportunities/ requirements	• Standardization towards testing robustness • Standardization of input data format and application side information model • Benchmark datasets • Failsafe mode of operation			
Challenges and issues	Data sufficiency			
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.35.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Causality-based Thermal Prediction for Data Center. 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA). Turin, Italy. 4-7 Sept. 2018.	Published	Use case taken from this reference	ABB	https://www.researchgate.net/publication/328995714_Causality-Based_Thermal_Prediction_for_Data_Center

A.36 Powering Remote Drilling Command Centre

A.36.1 General

<table>
<tr><td>ID</td><td colspan="4">36</td></tr>
<tr><td>Use case name</td><td colspan="4">Powering Remote Drilling Command Centre</td></tr>
<tr><td>Application domain</td><td colspan="4">Manufacturing</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Oil and Gas Upstream (Deployed in 150 Oil Rigs and 2.5 Billion+ Data Points each)</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Automatic generation of Daily Performance Report, reduction in overall drilling time, cut down Invisible Loss Time and improve rig asset management</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">It is important for a drilling contractor to have real time monitoring of rig parameters to optimize operations. The customer lacked granular insights during drilling, could not ascertain the root cause of non-productive time, and manual interpretation of signals led to missing of anomalies further degrading performance.</td></tr>
<tr><td>Complete description</td><td colspan="3">Cerebra product extracted and ingested different types of signals from surface and downhole sensors to perform near real-time processing. More than 170 vital signals every second from each oil rig were processed by Cerebra to provide near real time insights into drilling operations. This was achieved by handling Data Format and Data Extraction standards and Cerebra's Visualization Studio provides the flexibility of generating customized asset utilization reports, thus helping the oilfield engineers to understand the root causes of non-productive time and better utilize the assets on field. Rig specific utilization reports, and weekly and monthly utilization reports helped to plan drilling operations improving drilling efficiency.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Oil and Gas Upstream sector; Environment, Humans</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Competitiveness (operational excellence); Safety and Environment</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Challenges to accountability, security threats</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Invisible Loss Time</td><td>Indicates the lost time of the asset in being idle or off or unplanned downtime</td><td>Asset Utilization Reports indicate the effectively utilized time there indicating the lost time and their causes</td></tr>
<tr><td>2</td><td>Overall drilling time</td><td>The time spent on one drilling job inclusive of the all downtimes</td><td>Real Time visibility into operations gives the operations early warnings to</td></tr>
</table>

				take actions immediately.

AI features	Task(s)	Knowledge processing & discovery
	Method(s)	Utilization and Performance Evaluation
	Hardware	Application Server: 64 GB RAM/ 16 Core / 500 GB HDD Data Server: 128 GB RAM/ 16 Core, 3 TB HDD
	Topology	
	Terms and concepts used	**ISO 14224:** • Equipment classification and application • Equipment boundary, taxonomy and time definitions **ISO 13379:** • Condition monitoring set-up and diagnostics requirements • Failure mode symptoms analysis • Elements used for diagnostics • Diagnostic approaches **ISO 13381-1:** • Prognosis Concepts • Failure and deterioration models used for Prognosis • Prognosis Process o ***Existing*** failure mode prognosis process o ***Future*** failure mode prognosis process **ISO 17359:** • Equipment audit o Identification of equipment o Identification of equipment function • Reliability and criticality audit o Reliability block diagram o Equipment criticality o Failure modes, effects and criticality analysis o Alternative maintenance tasks • Monitoring method o Measurement technique o Accuracy of monitored parameters o Feasibility of monitoring o Operating conditions during monitoring o Monitoring interval o Data acquisition rate o Record of monitored parameters o Measurement locations o Initial alert/alarm criteria o Baseline data • Data acquisition and analysis o Measurement and trending o Quality of measurements o Measurement comparison to alert/alarm criteria o Diagnosis and prognosis o Improving diagnosis and/or prognosis confidence

		Determine maintenance action
Standardization opportunities/ requirements	• **Mandate of the key sensors based on the type of equipment** Based on the type of equipment, the makers need to have the basic set on sensors imbibed onto the system. E.g. for a pump – it is important to measure the input flow and output flow rates, vibrations, rotation speed, lube oil temperature and pressure. This will guide the equipment manufactures to provide their customers and their data products to capture the minimum required data and understand the equipment performance. • **Mandate for the organizations to expose the minimum and key parameters** The equipment owners need to enable the basic set of sensors for the equipment health and performance which are required for monitoring the asset from any failures. • **Standards for data formats** Each organization has a different way of capturing data and storing them in different formats. Due to which the solutions are not scalable across organizations though the product behind them is same. It takes customised efforts each time. • **Guidelines for deciding the sampling frequency based on the type of data** We see a need to have a specific set of guidelines to capture data at a minimum required sampling frequency. For e.g. a vibration sensor should capture data at least at 1 ms. • **Guidelines for feature engineering** There must be guidelines as to how the features need to be engineered for AI models. Lack of this would lead to more black box models not explaining how the models behave the way they do. • **Guidelines for standardization of event types and codes** There are multiple events which occur for an asset or in a manufacturing plant. Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical. • **Guidelines for standardization of fault and error codes for an equipment or process** Similar to events, it is also useful to capture fault, failure and error codes in a standard way. • **Process guidelines for event related data (maintenance and work orders)** Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical.	
Challenges and issues	Compliance of organizations	
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

3522

A.36.2 Data

Data characteristics	
Description	Data from an Oil & Gas Rig
Source	Drilling Equipment
Type	Time-Series Sensor Data
Volume (size)	
Velocity	2.5 Billion+ Data Points each day
Variety	Machine Data
Variability (rate of change)	
Quality	

A.36.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Web Page	Upstream Sensor Data + Big Data Analytics = Game Changer in Oil n Gas industry	Published	Use case take from this case study	Flutura Business Solutions Pvt. Ltd.	https://www.flutura.com/blog/Upstream-Sensor-Data--Big-Data-Analytics-=-Game-Changer-in-Oil-n-Gas-industry
2	Web Page	Cerebra creating game changing impact on upstream outcomes	Published	Use case take from this case study	Flutura Business Solutions Pvt. Ltd.	https://flutura.com/case-study-oil-and-gas

A.37 Leveraging AI to Enhance Adhesive Quality

A.37.1 General

ID	37	
Use case name	Leveraging AI to enhance adhesive quality	
Application domain	Manufacturing	
Deployment model	On-premise systems	
Status	In operation	
Scope	Batch/Continuous/Discrete Manufacturing (Deployed in 75+ manufacturing lines in 10+ countries; Specifically identified the contributors to quality; predict potential quality failures).	
Objective(s)	Enhance Adhesive Quality, Performance Benchmarking	
Narrative	Short description (not more than 150 words)	Cerebra IOT signal intelligence platform provides the ability to have a holistic perspective and understanding of the sensitivity of the key parameters affecting output quality and ability to monitor and control the process in real-time.

		This will avoid variations in yields, build-up of inventories and missed customer deadlines.		
	Complete description	Cerebra IOT signal intelligence platform ingested 3+ years of process data and sensor data regarding plant operations from temperature, rpm, torque and pressure sensors which were strapped on to industrial mixers. These are the mandatory sensors for the operations. Cerebra used its episode detection algorithms (deep learning) to filter signal from noise and specifically identify the contributors to quality (anomaly signatures) that can then be used as signals to predict quality. It used its proprietary N-dimensional Euclidian distance-based scoring algorithms to normalize and present a unified score to the business team. This unified health score provided the process team a different lens to benchmark, specifically target and radically improve process efficiencies. Cerebra then leveraged its sophisticated ensemble models to predict potential quality failures allowing the operations team to take real-time actions to control process deviations. The signals identified in the earlier steps provide Model Explainability to the end-user for reasons behind Quality deviation.		
Stakeholders	Manufacturing industries; Suppliers and Buyers; Environment			
Stakeholders' assets, values	Competitiveness (Respond to and exceed customers' and consumers' expectations by providing the best value, quality, service and winning innovations, brands and technologies to create sustainable value).			
System's threats & vulnerabilities	Challenges to accountability, New Security Threats.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Prediction Accuracy	To what extent has the model been able to predict correctly	Provided ability as to % of times the quality complied
AI features	Task(s)	Prediction		
	Method(s)	N-dimensional Euclidian distance-based scoring algorithms		
	Hardware	Application Server: 64 GB RAM/ 16 Core / 500 GB HDD Data Server: 128 GB RAM/ 16 Core, 3 TB HDD		
	Topology			
	Terms and concepts used	ISO 13381-1: ● Prognosis Concepts ● Failure and deterioration models used for Prognosis ● Prognosis Process ➢ ***Existing*** failure mode prognosis process ➢ ***Future*** failure mode prognosis process ISO 17359: ● Monitoring method ➢ Measurement technique ➢ Accuracy of monitored parameters ➢ Feasibility of monitoring ➢ Operating conditions during monitoring		

		➢ Monitoring interval ➢ Data acquisition rate ➢ Record of monitored parameters ➢ Measurement locations ➢ Initial alert/alarm criteria ➢ Baseline data ● Data acquisition and analysis ➢ Measurement and trending ➢ Quality of measurements ➢ Measurement comparison to alert/alarm criteria ➢ Diagnosis and prognosis ➢ Improving diagnosis and/or prognosis confidence ISA 95: ● Identify and work on the boundaries between the enterprise systems and the control systems
Standardization opportunities/ requirements	• **Mandate of the key sensors based on the type of equipment.** Based on the type of equipment, the makers need to have the basic set on sensors imbibed onto the system. e.g. for a pump – it is important to measure the input flow and output flow rates, vibrations, rotation speed, lube oil temperature and pressure. This will guide the equipment manufactures to provide their customers and their data products to capture the minimum required data and understand the equipment performance. • **Mandate for the organizations to expose the minimum and key parameters.** The equipment owners need to enable the basic set of sensors for the equipment health and performance which are required for monitoring the asset from any failures. • **Standards for Data Formats** Each organization has a different way of capturing data and storing them in different formats. Due to this, the solutions are not scalable across organizations though the product behind them is same. It takes customised efforts each time. • **Guidelines for deciding the sampling frequency based on the type of data.** We see a need to have a specific set of guidelines to capture data at a minimum required sampling frequency, e.g. a vibration sensor should capture data at least at 1 ms or less. • **Guidelines for Feature Engineering.** There must be guidelines as to how the features need to be engineered for AI models. Lack of this would lead to more black box models not explaining how the models behave the way they do. • **Guidelines for Standardization of event types and codes.** There are multiple events which occur for an asset or in a manufacturing plant. Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical. • **Guidelines for standardization of Fault and Error Codes for an equipment or process.** Similar to events, it is also useful to capture fault, failure and error codes in a standard way. • **Process Guidelines for event related data (Maintenance and Work Orders):**	

	Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical. • **Guidelines for Training AI models:** A defined set of guidelines for AI models would be useful for the data scientists to follow. It will also aid the consumers of AI models to understand how the outcome has been deduced. • **Guidelines around AI model explainability:** With so many black-box models floating around in the industry, it is difficult for consumers of AI models to understand these models and their output. And with engineers and domain experts coming into the picture, it is very much required to make these models more explainable. • **Process Guidelines and methods for model evaluation (retraining)** Before deployment and post deployment, it is very critical to have standard methods for models. And also post deployment, we must set guidelines for retaining the model on a periodic basis or based on data volatility. This is increasingly becoming important as AI models are being involved in more strategic and operational decision making. • **Guidelines for disaster recovery n autonomous operations:** With the aid of AI models, the operations of an equipment or manufacturing plant are becoming more and more autonomous and self-sufficient. But the human monitoring is also important as any kind of inaccurate prediction can lead to a disaster and it is must to have some standard to recover from this situation and to assess the conditions to go for autonomous operations.	
Challenges and issues	Patented process if any, security restrictions	
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.37.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Web link	Leveraging Cerebra's AI to enhance quality – from Quality Inspection to Quality Assurance	Published as case study	Use case take from this case study	Flutura Business Solutions Pvt. Ltd.	https://flutura.com/case-study-specialty-chemicals

A.38 Machine Learning Driven Approach to Identify the Weak Spots in the Manufacturing of the Circuit Breakers

A.38.1 General

ID	38	
Use case name	Machine learning driven approach to identify the weak spots in the manufacturing of the circuit breakers.	
Application domain	Manufacturing	
Deployment model	Prototype	
Status	On-premise system	
Scope	Detecting the issues in manufacturing process that leads to early failures of the circuit breakers through the data mining of the manufacturing process.	
Objective(s)	To generate actionable intelligence to improve the manufacturing process of circuit breakers through mining of manufacturing related data.	
Narrative	Short description (not more than 150 words)	An approach was developed that can mine the manufacturing data of circuit breakers through multiple machine learning algorithms. The approach could successfully identify the weak spots in the manufacturing where failure rate jumped from 0.2% to 7% (35 fold more probability of failure) and hence candidates for improvement in the manufacturing process.
	Complete description	High voltage circuit breakers are critical component of an electric circuit and it has a normal lifespan of 30-40 years. However, due to various reasons few circuit breakers fail within 0-5 years of operation. As a manufacturer of these circuit breakers, lots of data related to manufacturing aspects are present with the manufacturer. Such data has information about production lot size, material of production, design voltages for sub-components, heater voltages, date of failure etc. In general data related to 49 variables are captured for close to 56000 circuit breakers over a lifespan of several years. The manufacturer is interested to know if there are any weak spots in the manufacturing process which leads to higher failure rates. Circuit breakers can fail not only due to manufacturing defects but also due to wrong operation of the circuit breaker in the field e.g. applying voltages higher than design values. However, operational data of the circuit breakers was not available with the manufacturer. Therefore, the key challenge of this project was knowledge discovery with partial data set using machine learning algorithms. The data scientists applied various machine learning algorithms such as decision tree, random forest, support vector machine, Naïve Bayes classifier, logistic regression and neural network and compared the results of one algorithm verses the other algorithm. Through multiple numerical experimentations on data selection and algorithm hyper parameter tuning, the data scientist team selected the best algorithms and deduced the key weak spots in the manufacturing that are generally associated

		with high failure rates. In conclusion, the work provided a set of 5 actionable rules, where the failure rates jumped drastically from 0.2% to 7% leading to 35-fold higher chance of failure.		
Stakeholders	Manufacturer of HV circuit breakers			
Stakeholders' assets, values	Reliable and safe power supply to customers			
System's threats & vulnerabilities	Incorrect use of AI/ML			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Ratio of ML discovered failure rate to nominal failure rate	What combination of manufacturing processes/decisions leads to higher failure rates compared to nominal failure rate	Actionable intelligence to improve the manufacturing process of HV circuit breakers
AI features	Task(s)	Classification		
	Method(s)	Decision trees, SVM, ANN, Logistic Regression, Random Forest and Naïve Bayes		
	Hardware	64 GB RAM Windows server		
	Topology	NA		
	Terms and concepts used	Classification, Actionable Rules, HV Circuit breakers		
Standardization opportunities/ requirements	Standardization of data representation models comprising of both manufacturing related data and end-use related data.			
Challenges and issues	Discovering actionable insight with partial data set and managing bias in ML models due to limited number of failed cases			
Societal Concerns	Description	Safe and reliable power delivery		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.38.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Kumar, S., K., Jamkhandi, A., G., and Gugaliya, J., K., Achieving Manufacturing Excellence through Data Driven Decisions, IEEE International Conference on Industrial Technology,	Presented in Feb 2019	Use case taken from this reference	ABB	Yet to be published

		Melbourne Australia PP 1267-1273				

A.39 Machine Learning Driven Analysis of Batch Process Operation Data to Identify Causes for Poor Batch Performance

A.39.1 General

<table>
<tr><td>ID</td><td colspan="2">39</td></tr>
<tr><td>Use case name</td><td colspan="2">Machine Learning Driven Analysis of Batch Process Operation Data to Identify Causes for Poor Batch Performance</td></tr>
<tr><td>Application domain</td><td colspan="2">Batch Manufacturing</td></tr>
<tr><td>Deployment model</td><td colspan="2">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="2">Prototype</td></tr>
<tr><td>Scope</td><td colspan="2">Detecting the issues in batch manufacturing process that leads to bad quality products or longer cycle times of batch processing</td></tr>
<tr><td>Objective(s)</td><td colspan="2">Provide insight to the operation team to improve the productivity of batch manufacturing through machine learning on historical operation data</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>An approach was developed that can use machine learning models to identify issues in batch manufacturing.</td></tr>
<tr><td>Complete description</td><td>Batch operation is generally quite complex involving dynamics in the operation and interplay of various process variables. Due to this, sometimes, few batches end up running slower than nominal batch time and few batches also yield bad quality end products resulting in significant production loss. Additionally, often in the industrial context, data size and variety are limited and to develop a robust machine learning model from limited available data sets is a challenging task.

Due to transient nature of batch operation data, the traditional PCA algorithm fails in analyzing the batch data and hence MPCA was applied as logical extension of PCA algorithm. As MPCA naturally considers the dynamics in the data and inter-correlations among the process variables, it provides a valuable insight on the batch data.

The approach was successfully demonstrated on milk pasteurization process data where only 4 batches were provided for modelling. Using such 4 seed batches, the algorithm synthetically creates 50 batches of data and introduction of anomalies in some batches. Concept of design of experiments and stochastic perturbations are used in synthetic generation of the data set.
The work was able to successfully build a robust MPCA model with such data and isolate the bad batches of data from good batches of the data. Additionally, through contribution plots, the algorithm identifies when a certain</td></tr>
</table>

		batch drifted from nominal operation and which variables are the root causes for the bad batch operation.		
Stakeholders	Batch manufacturer such as milk pasteurization, pharmaceutical, paint manufacturing, etc.			
Stakeholders' assets, values	Improve the productivity and avoid the re-work			
System's threats & vulnerabilities	Incorrect use of AI/ML; New Security Threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Closeness to Golden Batch	How close a process is to the best possible batch	Helps in isolation of bad batches from good batches by identifying combination of process variable trajectories that lead to good or bad batch operation.
AI features	Task(s)	Classification		
	Method(s)	Multiway Principal Component Analysis		
	Hardware	64 GB RAM Windows server		
	Topology	NA		
	Terms and concepts used	Classification, MPCA, Anomalies		
Standardization opportunities/ requirements	• Standard data representation models for AI relevant batch data handling • Standard GUI for AI relevant result presentation.			
Challenges and issues	Discovering actionable insight with limited industrial data set, handling dynamics in the process variables			
Societal Concerns	Description	Consistent batch operation lead to enhanced productivity		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.39.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Jeffy, F., J., Gugaliya, J., K., and Kariwala, V. Application of Multi-Way Principal Component Analysis on Batch Data, 2018	Published	Use case taken from this source	ABB	https://www.researchgate.net/publication/328989762_Application_of_Multi-Way_Principal_Co

		UKACC 12th International Conference on Control				mponent_Analysis _on_Batch_Data

A.40 Empowering Autonomous Flow Meter Control- Reducing Time Taken to "Proving of Meters"

A.40.1 General

ID	40			
Use case name	Empowering Autonomous Flow meter control- Reducing time taken to "proving of meters"			
Application domain	Manufacturing			
Deployment model	Cloud services			
Status	In operation			
Scope	Calibration of control devices			
Objective(s)	Reduce the time taken for trial & error methods to set the VFD and FCV setpoints			
Narrative	Short description (not more than 150 words)	The customer had to set VFD and FCV % manually to achieve desired flowrate using trial & error methods, which could take about 3-4 hours. Efficiency for the proving of the meters was very less & improvement was needed to remove any aberration in reading as it was time consuming.		
	Complete description	Cerebra was integrated with the system considering the flow of the fluid. The customer can choose between the available options of high flow rate, low flow rate or multi viscous flow. Then, with the master meter in the loop of testing, the meter from the field was introduced to analyse how much of aberration is there and then proving it more efficiently. Since it took more time for them to get the exact values of VFD & FCV % to achieve the desired flow rate, Cerebra's Prognostics Engine was introduced. Purely based upon machine learning algorithms, the data models for the VFD & FCV % was used to predict the values to be chosen with an accuracy of about 98%. Since there was a presence of a closed-loop system, this predicted value was automatically registered on the valves' monitors which only required small tweaking in the end, thus reduced human efforts.		
Stakeholders	Process Industries; Humans			
Stakeholders' assets, values	Competitiveness; Stability.			
System's threats & vulnerabilities	Challenges to accountability, security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Model Accuracy	Accuracy of the prediction model	The extent to which the setpoints have

				correctly predicted
	2	% Reduction in Calibration Time	The amount of time saved from manually setting the calibration	
AI features	Task(s)	Prediction		
	Method(s)	Random Forest prediction, one hot encoding, cross validation, normalization		
	Hardware	Application Server: 64 GB RAM/ 16 Core / 500 GB HDD; Data Server: 128 GB RAM/ 16 Core, 3 TB HDD		
	Topology			
	Terms and concepts used	**ISO 14224:** • Equipment classification and application • Equipment boundary, taxonomy and time definitions **ISO 13379:** • Condition monitoring set-up and diagnostics requirements • Failure mode symptoms analysis • Elements used for diagnostics • Diagnostic approaches **ISO 13381-1:** • Prognosis Concepts • Failure and deterioration models used for Prognosis • Prognosis Process o ***Existing*** failure mode prognosis process o ***Future*** failure mode prognosis process **ISO 17359:** • Equipment audit o Identification of equipment o Identification of equipment function • Reliability and criticality audit o Reliability block diagram o Equipment criticality o Failure modes, effects and criticality analysis o Alternative maintenance tasks • Monitoring method o Measurement technique o Accuracy of monitored parameters o Feasibility of monitoring o Operating conditions during monitoring o Monitoring interval o Data acquisition rate o Record of monitored parameters o Measurement locations o Initial alert/alarm criteria o Baseline data • Data acquisition and analysis o Measurement and trending o Quality of measurements		

		o Measurement comparison to alert/alarm criteria o Diagnosis and prognosis o Improving diagnosis and/or prognosis confidence • Determine maintenance action **ISA 95:** Identify and work on the boundaries between the enterprise systems and the control systems
Standardization opportunities/ requirements	• **Mandate of the key sensors based on the type of equipment** Based on the type of equipment, the makers need to have the basic set on sensors imbibed onto the system. E.g. for a pump – it is important to measure the input flow and output flow rates, vibrations, rotation speed, lube oil temperature and pressure. This will guide the equipment manufactures to provide their customers and their data products to capture the minimum required data and understand the equipment performance • **Mandate for the organizations to expose the minimum and key parameters** The equipment owners need to enable the basic set of sensors for the equipment health and performance which are required for monitoring the asset from any failures • **Standards for Data Formats** Each organization has a different way of capturing data and storing them in different formats. Due to which the solutions are not scalable across organizations though the product behind them is same. It takes customised efforts each time. • **Guidelines for deciding the sampling frequency based on the type of data** We see a need to have a specific set of guidelines to capture data at a minimum required sampling frequency. For e.g. a vibration sensor should capture data at least at 1 ms or less. • **Guidelines for Feature Engineering** There must be guidelines as to how the features need to be engineered for AI models. Lack of this would lead to more black box models not explaining how the models behave the way they do. • **Guidelines for Standardization of event types and codes** There are multiple events which occur for an asset or in a manufacturing plant. Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical. • **Guidelines for standardization of Fault and Error Codes for an equipment or process** Similar to events, it is also useful to capture fault, failure and error codes in a standard way. • **Process Guidelines for event related data (Maintenance and Work Orders)** Guidelines would help people capture the data in a similar fashion helping the industry to benchmark against one another and at industry level we can understand, which events are the most critical • **Guidelines for Training AI models** A defined set of guidelines for AI models would be useful for the data scientists to follow. It will also aid the consumers of AI models to understand how the outcome has been deduced	

	• **Guidelines around AI model explainability** With so many black box models floating around in the industry, it is difficult for consumers of AI models to understand then and their output. And with engineers and domain experts, coming into the picture, it is very much required to make these models more explainable. • **Process Guidelines and methods for model evaluation (retraining)** Before deployment and post deployment, it is very critical to have standard methods for models. And also post deployment, we must set guidelines for retaining the model on a periodic basis or based on data volatility. This is increasingly becoming important as AI models are being involved in more strategic and operational decision making. • **Guidelines for disaster recovery and autonomous operations** With the aid of AI models, the operations of an equipment or manufacturing plant are becoming more and more autonomous and self- sufficient. But the human monitoring is also important as any kind of inaccurate prediction can lead to a disaster and it is must to have some standard to recover from this situation and to assess the conditions to go for autonomous operations.	
Challenges and issues		
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are all important ways to facilitate sustainable development.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.40.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Web Page	Accelerating shale production through digital technology integration	Published	Use case taken from this source	Flutura Business Solutions Pvt. Ltd. TechnipFMC	https://www.technipfmc.com/en/media/features/accelerating-shale-production-through-digital-technology-integration?type=features
2	Web Page	Fundamentals of meter provers and proving methods	Published	Fundamental definition of Meter Provers	Flow Management Devices	https://asgmt.com/wp-content/uploads/2016/02/011_.pdf

A.41 Improving Productivity for Warehouse Operation

A.41.1 General

ID	41

<table>
<tr><td>Use case name</td><td colspan="4">Improving Productivity for Warehouse Operation</td></tr>
<tr><td>Application domain</td><td colspan="4">Logistics</td></tr>
<tr><td>Deployment model</td><td colspan="4">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="4">PoC</td></tr>
<tr><td>Scope</td><td colspan="4">Big data analysis for enhancing productivity</td></tr>
<tr><td>Objective(s)</td><td colspan="4">To improve productivity of warehouse operation by detecting and changing controllable factors</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">AI-driven operating system that uses big data from work performance information to issue appropriate work instructions has been developed. In PoC, picking operation improvement was conducted in a distribution warehouse. As the result, 8% work reduction was performed.</td></tr>
<tr><td>Complete description</td><td colspan="3">Attempts are being made to increase the efficiency of work improvements through more widespread application of IT to work systems. However, as each new improvement is added or improvements are made with respect to environmental changes, it requires manual changes to the system, leading to increases in work improvement costs. This case has developed an AI system that uses big data such as work performance information, to understand worksite improvements and environmental changes and issue appropriate work instructions. It has conducted a demonstration test, which confirmed the effectiveness of this system for improving distribution warehouse work. In the future, we will continue to work on expanding the AI system to a wide range areas such as manufacturing and distribution.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">warehouse manager</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">reducing cost, reducing labor related problems (e.g. minimizing labors complaint), speed up of operation.</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">possibility of back action</td></tr>
<tr><td rowspan="4">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Number of labors</td><td>reduced % of labors</td><td>improvement of productivity</td></tr>
<tr><td>2</td><td>Number of complaints</td><td>reduced % of labor's complaint</td><td>improvement of productivity</td></tr>
<tr><td>3</td><td>Lead time</td><td>time from order to shipment</td><td>improvement of productivity</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Optimization</td></tr>
<tr><td>Method(s)</td><td colspan="3">modelling of relationship between explaining variables and outcome, and optimization</td></tr>
<tr><td>Hardware</td><td colspan="3">PC, wearable sensor</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Human big data analysis, regression analysis</td></tr>
<tr><td>Standardization</td><td colspan="4">standardization of data format, sensors to be used, and API of IT and mechanical systems</td></tr>
</table>

opportunities/ requirements		
Challenges and issues	understanding of workers' human factors (privacy, additional work etc.)	
Societal Concerns	Description	solving labor shortage problem and improving labor related issues with aiming improving productivity.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.41.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	company's technical journal		Published		Hitachi, Ltd.,	http://www.hitachi.com/rev/archive/2016/r2016_06/106/index.html

[1] F. Kudo T. Akitomi and N. Moriwaki, "An Artificial Intelligence Computer System for Analysis of Social Infrastructure Data," IEEE conf. Business Infomatics (CBI), 2015.

[2] J. Kimura et al., "Framework for Collaborative Creation with Customers to Improve Warehouse Logistics," Hitachi Review, 65, pp. 873-877, 2016.

[3] Hitachi News Release, "Development of Artificial Intelligence issuing work orders based on understanding of on-site Kaizen activity and demand fluctuation," 2015. http://www.hitachi.com/New/cnews/month/2015/09/150904.html

A.42 Emotion-sensitive AI Customer Service

A.42.1 General

ID	42	
Use case name	Emotion-sensitive AI Customer Service	
Application domain	Retail	
Deployment model	On-premise systems	
Status	In operation	
Scope	Extracting sentiment and its intensity from customers' input, and responding with appropriate attitude in order to improve the quality of customers' inquiry.	
Objective(s)	To design an efficient solution for customers' sentiment and intensity detection, especially in the situation of limited training dataset.	
Narrative	Short description (not more than 150 words)	The emotion-sensitive AI customer service of JD.com Int., is supported by AI technology and deep learning method. It is developed for ameliorating accuracy of customer sentiment and intensity. In sentiment classification, it has achieved 74% accuracy and 90% recall score while in intensity detection, it has accomplished 85% accuracy and 85%

<table>
<tr><td rowspan="2"></td><td></td><td colspan="3">recall. During the special sale of “618”, it has increased customer satisfaction by 57%.</td></tr>
<tr><td>Complete description</td><td colspan="3">JD’s customer service representatives need to handle millions of requests on a daily basis. Regular AI customer service systems, 24/7 online, are capable of offering instant assistance, which alleviates the labor resources to a large extent. However, it is quite challenging, if not impossible, for those systems to interpret emotions from customer input and respond as friendly as human.

Under this background, based on huge data set of customer comments and rich experience of Natural Language Processing, our system can automatically detect sentiments like happy, angry, anxious, etc. Moreover, this system can also detect the intensity of customer sentiment. Furthermore, we adapt Convolutional Neural Networks, a widely used techniques in visual computing, to interpret the semantic meaning of customer’s expression. It can improve the system’s performance for sentiment classification and intensity detection. Moreover, with the adoption of transfer learning, the system can also be applied into various types of data. To overcome the difficulty of limited training data, we also use data augmentation method such as reverse translation and data noise to increase the variability of training data.

Up to now, the system has reached 90% recall and 74% accuracy rate for sentiment classification over 7 categories. The overall recall and accuracy for sentiment intensity are also around 85%, it has increased customer satisfaction by 57%.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Customers targeted for the Customer Service system</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Customer experience may be in influnced by the use of AI custemer service</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">The low degree of humanization, and lack of semantic diversity for response;
Reducing the number of human customer service.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Customer Satisfaction</td><td>The ratio of customer satisfaction when using this system for requests. The expectation is 100%</td><td>Increasing its ratio as high as possible</td></tr>
<tr><td>2</td><td>Accuracy</td><td>Among all the predicted customer sentiment classification, the ratio of accurate</td><td>Increasing to 90%</td></tr>
</table>

			prediction, current value is 76.4%	
	3	Recall	Among all the customer sentiment intensity, the ratio of accurate prediction, current overall value is 90%	Increasing to 90%
	4	Accuracy	Among all the predicted customer sentiment intensity, the ratio of accurate prediction, current overall value is 85%	Increasing to 90%
	5	Recall	Among all the customer sentiment intensity, the ratio of accurate prediction, current overall value is 85%	Increasing to 90%
AI features	Task(s)	Natural language processing		
	Method(s)	Deep learning, transfer learning, data augmentation		
	Hardware			
	Topology			
	Terms and concepts used	Deep learning: a class of machine learning algorithms use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Transfer learning: we adopt multi-task learning method in this system. Jointly training different annotated data in same domain, this method improves the model performance for classification problems. Data augmentation: we apply reverse translation to firstly translation Chinese into English and then translate it backward. We also use data noise to improve the data diversity.		
Standardization opportunities/ requirements	The system can be promoted to as many customer cervices companies as possible once provide with enough training data for the specific Application scenario			
Challenges and issues	Challenge: the system's performance should be as good as the human customer server. Issues: 1) limited training data; 2) sentiment classification among seven categories.			

Societal Concerns	Description	Improving the corresponding efficiency of customer service, improving customer service experience； Reducing labor costs, and reducing operating costs.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.42.2 Data

Data characteristics	
Description	For sentiment classification: conversation data from after-sales customer services. It’s annotated by professional annotators into 7 categories of sentiments. For sentiment intensity: Only including sentiment data with “anger” and “anxious”; it’s annotated into 3 degrees of intensity: “low, medium, high”.
Source	Conversation data from JD.com real-time customer services.
Type	Text
Volume (size)	Around 60,000 sentences for sentiment classification and 20,000 for sentiment intensity.
Velocity	Batch Processing
Variety	Real-time data from JD.com, including various categories of products.
Variability (rate of change)	Static
Quality	High

A.42.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Data Augmentation	Using reverse translation and noise processing to increase the size and diversity of data.	Annotated raw data is ready.		Increase the performance of model training.
2	Model Training	Based on the large training data, with deep learning method, to develop model for sentiment classification (7 categories) or sentiment intensity (3 categories).	Augmented data is ready		
3	Evaluation	Evaluate data performance on open dataset and specific data.	Pretrained model is ready		

4	Execution	Apply the trained model on real-time AI customer service.		The trained model has been evaluated as deployable	
5	Retraining	Retraining model with new annotated data and new requirement from industry.			

A.42.4 **Training**

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Complete data augmentation	Design model for training	AI algorithm engineers	Using CNN for sentiment classification and intensity.	
2	Complete model designing	Transfer learning	AI algorithm engineers	Multi-task learning with different data in same domain.	

Specification of training data	

A.42.5 **Evaluation**

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Complete model training	Evaluation on open dataset	AI algorithm engineers	Evaluate different models' performance on open dataset	Their performance shall be as good as state-of-art.
2	Complete model training	Evaluation on own dataset	AI algorithm engineers	Evaluate different models' performance on own dataset	Their performance shall meet certain standard.

Input of evaluation	Independent testing data
Output of evaluation	Accuracy and Recall

A.42.6 **Execution**

Scenario name	Execution

Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Finish model training	Application	AI engineers	Making trained model into application of AI Customer Service system.	
2	Given customer's input	Data processing	AI algorithm engineers	Processing data into required format for model.	
3	Finish data processing	Model prediction	AI algorithm engineers	Predicting sentiment or sentiment intensity.	
4	Completion of Step3	Making response	AI algorithm engineers	Making response according to the preidiction from previous step.	

Input of Execution	
Output of Execution	

A.42.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Certain period of time has passed since the last training/retrainig	Improve architecture of model	AI algorithm engineers	Collecting new requirements for model designing.	
2	Certain period of time has passed since the last training/retrainig	Collecting new data	AI algorithm engineers	Collecting new data based on the further requirements.	
3	Completing Step1&Step2	Model retraining	AI algorithm engineers	Training new model on additional data.	

Specification of retraining data	

A.42.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	IT company	XiaoIce	In operation		Microsoft Asia	

A.43 Deep Learning Based User Intent Recognition

A.43.1 General

ID	43	
Use case name	Deep Learning Based User Intent Recognition	
Application domain	Retail	
Deployment model	On-premise systems	
Status	In operation	
Scope	Recognizing users' intent to solve their problems in e-commerce fields	
Objective(s)	To recognize and understand users' intent by AI and deep learning technologies and apply such technologies to build chat bot systems to further reduce labor cost and to be applied in various fields.	
Narrative	Short description (not more than 150 words)	Intelligent customer service chat bot is mainly used to categorize users' questions, recognize users' intents and answer users' questions intelligently for different business jobs. Currently, this chat bot has been used to handle 90% of online customer service and has enabled JD.com to save over 100 million labor costs every year.
	Complete description	JD.com has been committed to using technology to drive business growth and improve user experience in all customer service fields. Based on the improvement of customer consulting experience and the developing trend of artificial intelligence technology, as early as 2012, JD had decided to develop intelligent chat bots to fulfill the needs of continuous expansion of business, to save customer service costs and increase service capability. Intent recognition is a key and core technology to build such an intelligent customer service chat bot. By applying natural language processing technologies, deep learning technologies, traditional machine learning algorithms, intent recognition accuracy has reached to 95%. Based on accurate intents, and a series of solution finding algorithms, our chat bot can solve the user's problems to a great extent and give the user a high quality consulting experience. Finally, in order to provide diversified and personalized customer services, we are continuously improving the accuracy of intent recognition, personalized solution generation, sentiment recognition, and image recognition. So far, intelligent customer service has revolutionized the traditional customer service consulting business.

Stakeholders	users			
Stakeholders' assets, values	Users' experience			
System's threats & vulnerabilities	high semantic ambiguity, Multiple language expressions in one sentence			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy	The number of correctly recognized users' intent over total number of users. Currently, accuracy reaches 95%.	Improve accuracy of recognizing users' intent
	2	Resolution	The number of answers solved over total number of questions asked	Improve the resolution of questions from users
	3	Satisfaction	The number of users who are satisfied with customer service over total number of users	Improve user experience
AI features	Task(s)	Natural language processing		
	Method(s)	Machine learning and deep learning		
	Hardware	GPU and CPU		
	Topology	TensorFlow		
	Terms and concepts used	Natural language processing, deep learning, CNN, HAN, logistic regression		
Standardization opportunities/ requirements	Process Standardization will Improve Quality and Productivity			
Challenges and issues	Current challenges of deep leaning and intent recognition: 1. high semantic ambiguity, similar sentences can deliver different meanings. 2. Unclear classification rules caused by complicated business logics 3. Hard to answer reasoning questions			
Societal Concerns	Description	1. Solve problems intelligently to increase efficiency 2. Free labors from repetitive work to save large amount of resources for the society		
	SDGs to be achieved	Decent work and economic growth		

A.43.2 Data

Data characteristics	
Description	Question answering data from the JD.com online dialogue log
Source	Customer's dialogue log at JD.com

Type	Text
Volume (size)	Millions
Velocity	Real time
Variety	various scenarios, various business, various categories of products
Variability (rate of change)	Non-linear
Quality	good

A.43.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Based on millions of labeled streaming data, train a model using diversified algorithms, such as a deep learning neural network or a traditional machine learning algorithm	The training sample is ready		
2	Evaluation	Evaluate the performance of the model on online dialogue data	The training procedure has been finished		Each requirement must be satisfied or exceeded to reach the condition of 'success' (e.g. the accuracy should be more than 95%)
3	Execution	Apply the trained model to predict user's intent	Require user's query		
4	Retraining	Take a training sample from online dialogue to retrain the model and	bad cases are feed back to update the training dataset		The requirement is that the new model must be better than the old one

		compare it with the old one by AB test			

A.43.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Raw data stored in the database	Data extraction	Database engineer	Extract related data from the database to generate the raw dataset	
2	Completion of Step 1	Generating training samples	Data labeling team	Label the raw dataset of step one with 300 categories	
3	Completion of Step 2	Pre-process	AI engineer	Segment the sentence into words and convert those words into vectors	
4	Completion of Step 3	Model training	AI engineer	Based on vectors generated on step 3 to train a model using diversified algorithms, such as a deep learning neural network or a traditional machine learning algorithm	

Specification of training data	After manual verifying, the accuracy of labelling should be more than 95%

A.43.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Certain period of time has passed since the last training/retraining	Data Extraction	Database engineer	Randomly take a sample from streaming data to form a test sample	

2	Completion of Step 1	Prediction	AI engineer	Predict the test sample in step 1 by the trained model	
3	Completion of Step 2	Evaluation	Data labeling team	Compare the result of predicted with the result of labeling	

Input of evaluation	the result of labeling and the result of prediction
Output of evaluation	The accuracy and recall rate

A.43.6 **Execution**

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Acquire the user's query	pre-process	AI engineer	Segment the sentence into words and convert those words into vectors	The trained model has been in operation
2	Completion of Step 1	Text classification	AI engineer	Predict the label of user's query	
3	Completion of Step 2	Response	AI trainer	Answer the query based on the result of intent classification	

Input of Execution	
Output of Execution	

A.43.7 **Retraining**

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activit y	Requirement
1	Certain period of time has passed since the last training/retr aining	Data extraction	Database engineer	Randomly take a sample from streaming data to from a training sample	

2	Completion of Step 1	Labeling the sample	Data labeling team	Manually label the sample data	
3	Completion of Step 2	Model training	AI engineer	Combine the new training sample with the old and train a model (deep learning and machine learning)	
4	Completion of Step 3	AB Test	AI engineer	Compare the predicted results of the new model with the results of the old one	The performance of the new model is better than results of the old one
5	Completion of Step 4	Online active of new model	AI engineer	The new model is been active online at JD.com	

Specification of retraining data	After the calibration, the accuracy of labelling should be more than 95%

3591

3592 A.43.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Convolutional Neural Networks for Sentence Classification			New York University	https://arxiv.org/abs/1408.5882
2	Paper	Hierarchical Attention Networks for Document Classification			Carnegie Mellon University, Microsoft Research, Redmond	http://www.aclweb.org/anthology/N16-1174
3	Paper	LIBLINEAR: A library for large linear classification Journal of Machine Learning Research			National Taiwan University	http://www.jmlr.org/papers/volume9/fan08a/fan08a.pdf

3593

3594 A.44 Chromosome Segmentation and Deep Classification

3595 A.44.1 General

ID	44
Use case name	Chromosome Segmentation and Deep Classification

Application domain	Healthcare	
Deployment model	Hybrid or other (please specify)	
Status	PoC	
Scope	Karyotyping of the chromosomes is restricted to healthy patients	
Objective(s)	Automating Karyotyping of the chromosomes in cell spread images. Segmentation of chromosomes in the images using non expert crowd.	
Narrative	Short description (not more than 150 words)	Karyotyping of the chromosomes micro-photographed under metaphase is done by characterizing the individual chromosomes in cell spread images. Currently, considerable effort and time is spent to manually segment out chromosomes from cell images, and classifying the segmented chromosomes. We proposed a method to segment out and classify chromosomes for healthy patients using a combination of crowdsourcing, preprocessing and deep learning, wherein the non-expert crowd from external crowdsourcing platform is utilized to segment out the chromosomes, which are then classified using deep neural network. Results are encouraging and promise to significantly reduce the cognitive burden of segmenting and karyotyping chromosomes.
	Complete description	Metaphase chromosome analysis is one of the primary techniques utilized in cytogenetics. Observations of chromosomal segments or translocations during metaphase can indicate structural changes in the cell genome, and is often used for diagnostic purposes. Karyotyping of the chromosomes micro-photographed under metaphase is done by characterizing the individual chromosomes in cell spread images. Currently, considerable effort and time is spent to manually segment out chromosomes from cell images, and classifying the segmented chromosomes into one of the 24 types, or for diseased cells to one of the known translocated types. Segmenting out the chromosomes in such images can be especially laborious and is often done manually, if there are overlapping chromosomes in the image which are not easily separable by image processing techniques. Many techniques have been proposed to automate the segmentation and classification of chromosomes from spread images with reasonable accuracy, but given the criticality of the domain, a human in the loop is often still required. In this paper, we present a method to segment out and classify chromosomes for healthy patients using a combination of crowdsourcing, preprocessing and deep learning, wherein the non-expert crowd from CrowdFlower is utilized to segment out the chromosomes from the cell image, which are then straightened and fed into a (hierarchical) deep neural network for classification. Experiments are performed on 400 real healthy patient images obtained from a hospital. Results are encouraging and promise to significantly reduce the cognitive burden of segmenting and karyotyping chromosomes.

Stakeholders	Hospitals, Doctors, Cytogeneticists, Patients			
Stakeholders' assets, values	Health, Diagnosis, Privacy			
System's threats & vulnerabilities	Incorrect classification and segmentation, Inadequate training samples for karyotyping of chromosomes, incorrect straightening of bent chromosomes; bias in annotation by crowd-sourcing			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Classifier Accuracy	Without straightening and pre-processing, the average classification accuracy obtained was 68.5%. However, with preprocessing, the classification accuracy improved to 86.7%. These results are very likely to improve with more annotated training data for classification.	
	2	Annotation Completeness	35.9 chromosomes segmented out after crowd annotation, for 50 images having 46 chromosomes	
AI features	Task(s)	Recognition		
	Method(s)	Crowdsourcing and Deep learning		
	Hardware	GPU enabled desktops		
	Topology	Deep models used for training and testing		
	Terms and concepts used	Deep learning, crowd sourcing, non-expert crowd, segmentation, karyotyping		
Standardization opportunities/ requirements	When images are of poor resolution apply super-resolution techniques before feeding the images to any classifier network.			
Challenges and issues	Crowd's job satisfaction Spamming in annotated data			
Societal Concerns	Description	Inaccurate classification of chromosomes can lead to stress in patients in case the classification is not reviewed by expert doctors		

	SDGs to be achieved	Good health and well-being for people

A.44.2 Data

Data characteristics	
Description	The dataset comprised of 400 stained images with varying degrees of overlap between chromosomes, out of which 200 were kept for testing and the remaining for training and validation
Source	Partner hospital
Type	Images
Volume (size)	400
Velocity	
Variety	
Variability (rate of change)	
Quality	

A.44.3 References

[1] Sharma, Monika & Saha, Oindrila & Sriraman, Anand & Hebbalaguppe, Ramya & Vig, Lovekesh & Karande, Shirish. (2017). Crowdsourcing for Chromosome Segmentation and Deep Classification. 786-793. 10.1109/CVPRW.2017.109.

A.45 Anomaly Detection in Sensor Data Using Deep Learning Techniques

A.45.1 General

ID	45	
Use case name	Anomaly Detection in Sensor Data Using Deep Learning techniques	
Application domain	Maintenance & support	
Deployment model	Hybrid or other (Cloud or on premise deployment)	
Status	PoC	
Scope	Temporal Data captured from sensors	
Objective(s)	Identify Anomalies and Events by learning the temporal patterns of sensor data, based on Deep Learning techniques.	
Narrative	Short description (not more than 150 words)	Mechanical devices such as engines, vehicles, aircrafts, etc., are typically instrumented with numerous sensors to capture the behaviour and health of the machine. The sensors temporal data has several complex patterns that are very hard to identify with traditional methods. We have proposed the use of Deep Learning algorithms for analysing such temporal patterns for anomaly/event detection, diagnosis, root cause analysis. Algorithms proposed so far are LSTM-AD, EncDec-AD, online RNN-AD. We used industrial datasets wherever possible and publically available datasets in other scenarios.

		In most of the cases, our algorithms were significantly better than other methods.
	Complete description	Mechanical devices such as engines, vehicles, aircrafts, etc., are typically instrumented with numerous sensors to capture the behaviour and health of the machine. However, there are often external factors or variables which are not captured by sensors leading to time-series which are inherently unpredictable. For instance, manual controls and/or unmonitored environmental conditions or load may lead to inherently unpredictable time-series. Detecting anomalies/events in such scenarios becomes challenging using standard approaches based on mathematical models that rely on stationarity, or prediction models that utilize prediction errors to detect anomalies. LSTM-AD Our Work started with Stacked LSTM network which is trained on non-anomalous data and used as a predictor over a number of time steps. The resulting prediction errors are modeled as a multivariate Gaussian distribution, which is used to assess the likelihood of anomalous behavior. The efficacy of this approach was demonstrated on four datasets: ECG, space shuttle, power demand, and multi-sensor engine dataset. EncDec-AD As an extension to the prior work we proposed a Long Short Term Memory Networks based Encoder-Decoder scheme for Anomaly Detection (EncDec-AD) that learns to reconstruct normal time-series behavior, and thereafter uses reconstruction error to detect anomalies. We experimented with three publicly available quasi predictable time-series datasets: power demand, space shuttle, and ECG, and two real-world engine datasets with both predictive and unpredictable behavior. We had shown that EncDec-AD is robust and can detect anomalies from predictable, unpredictable, periodic, aperiodic, and quasi-periodic time-series. Further, we showed that EncDec-AD is able to detect anomalies from short time-series (length as small as 30) as well as long time-series (length as large as 500). Online-AD The common approach of training one model in an offline manner using historical data is likely to fail under dynamically changing and non-stationary environments where the definition of normal behavior changes over time making the model irrelevant and ineffective. We described a temporal model based on Recurrent Neural Networks (RNNs) for time series anomaly detection to address challenges posed by sudden or regular changes in normal behaviour. The model is trained incrementally as new data becomes available, and is capable of adapting to the changes in the data distribution. RNN is used to make multi-step

		predictions of the time series, and the prediction errors are used to update the RNN model as well as detect anomalies and change points. Large prediction error is used to indicate anomalous behaviour or a change (drift) in normal behaviour. Further, the prediction errors are also used to update the RNN model in such a way that short term anomalies or outliers do not lead to a drastic change in the model parameters whereas high prediction errors over a period of time lead to significant updates in the model parameters such that the model rapidly adapts to the new norm. We demonstrate the efficacy of the proposed approach on a diverse set of synthetic, publicly available and proprietary real-world datasets.		
Stakeholders	Maintenance and support functions, Monitoring, Procurement			
Stakeholders' assets, values	Anomaly/event detection, Diagnosis, Root cause analysis			
System's threats & vulnerabilities	Data biases could result in high number of false negatives and false positives that could result in heavy losses. Accuracy cannot be 100%.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Precision	Correctly Predicted Anomalous scenarios/ Total Anomalous scenarios predicted	
	2	Recall	Correctly Predicted Anomalous scenarios /Total Anomalous Scenarios	
AI features	Task(s)	Prediction		
	Method(s)	Deep Learning		
	Hardware	GPU enabled desktops and servers		
	Topology	Deep models used for training and testing		
	Terms and concepts used	Deep Learning, Recurrent Neural Networks, feature engineering		
Standardization opportunities/ requirements	Sensor data collection			
Challenges and issues	Noisy Data Data with missing temporal features Rarity of Anomalous Data			
Societal Concerns	Description	None		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

3606

A.45.2 Data

Data characteristics	
Description	Multiple datasets(publically available, real industrial) were used
Source	
Type	Temporal data
Volume (size)	
Velocity	
Variety	Space shuttle, ECG, Engine, Power demand
Variability (rate of change)	
Quality	

A.45.3 References

[1] Pankaj Malhotra, Anusha Ramakrishnan, Gaurangi Anand, Lovekesh Vig, Puneet Agarwal, Gautam Shroff, LSTM-based Encoder-Decoder for Multi-sensor Anomaly Detection, https://arxiv.org/abs/1607.00148

[2] Sakti Saurav, Pankaj Malhotra, Vishnu TV, Narendhar Gugulothu, Lovekesh Vig, Puneet Agarwal, Gautam Shroff, Online anomaly detection with concept drift adaptation using recurrent neural networks, CoDS-COMAD '18, Proceedings of the ACM India Joint International Conference on Data Science and Management of Data, Goa, India — January 11 - 13, 2018

A.46 Adaptable Factory

A.46.1 General

ID	46	
Use case name	Adaptable Factory	
Application domain	Manufacturing	
Deployment model	Cyber-physical System, Embedded System	
Status	PoC	
Scope	(Semi-)Automatic change of a production system's capacities and capabilities from a behavioral and physical point of view	
Objective(s)	The objective is to enable flexible production resources which enable fast reconfiguration and adaptation to changing situations, context, and requirements which facilitate optimized resource usage under uncertainty.	
Narrative	Short description (not more than 150 words)	Rapid, and in some cases completely automated, conversion of a manufacturing facility, by changing both production capacities and production capabilities. This use case describes the adaptability of an individual factory by (physical) conversion and/or adaption of a factory's and its machines behavior in order to adjust to changing situations like disruptions, material quality variation, production of new products, etc. A prerequisite is a modular and thereby adaptable design for manufacturing within the factory. The result is a need

		for intelligent and interoperable modules that basically adapted to an altered configuration on their own, and standardized interfaces between these modules.
	Complete description	Use Case description taken from [1,2,3]. Plug & Play – using a home computer and a USB cable, it is easy to connect new devices and use them almost immediately without any additional effort. The flexibility that has been available for quite a while on desktop computers is now gaining importance for industrial production. Demands on adaptability of production infrastructure are already rapidly increasing. Shorter and shorter product and innovation cycles require investment decisions for new production facilities that reflect future demand for production and process changes, where possible. In addition, the growing volatility of orders is hindering the optimal utilization of manufacturing lines with increasing frequency. Flexibility and adaptability will become increasingly important criteria in decisions regarding construction and operation of new production facilities. One example is product labeling. Various printing technologies are available, for example tampon printers (transferring ink from the printing form to the product using an elastic tampon), inkjet printers and/or laser printers. In an adaptable factory this type of operating equipment can be connected directly to the automated production process. Simply put, the material to be printed says: “Print me”, and the tampon printer will ask: “Is the material to be printed greaseless?” The ink jet printer will then ask about the material characteristics, because it uses heat for the drying process, for example. A laser printer will ask about the material receiving the label to ensure sufficient contrast. **Key aspects** The application scenario for adaptable factories describes the rapid, and in some cases completely automated conversion of a manufacturing facility, by changing both production capacities and production capabilities. The key concept for implementation is a modular and thereby adaptable design for manufacturing within the factory. Intelligent and interoperable modules that basically adapted to an altered configuration on their own, and standardized interfaces between these modules allow for quick and simple conversion to adapt to changes in the market and customer demands. Whereas the application scenario Order-Controlled Production emphasizes flexible use of existing manufacturing facilities by means of intelligent connectivity, this scenario describes the adaptability of an individual factory by (physical) conversion. Today, when creating a production line, the focus is usually not only on quality, but also maximization of productivity and profitability of a pre-conceived product range. Individual components are connected statically and are capable of producing the pre-conceived functionalities and

		projected volumes. Frequently, a system integrator takes care of coordinating the individual components and developing a control system for the entire facility. However, if the order level is driven by strong product individuality or high fluctuation in demand, companies can no longer rely on the advantage of particular production lines. In this case, modular, order-oriented and adaptable manufacturing configurations become more attractive: For example, they increase overall utilisation or ability to deliver products. At the same time, however, the demands on individual machines or manufacturing modules increase. Even more important than high variance of specific manufacturing steps will be the ability to combine individual modules with ease and in any situation. In order to achieve this, the modules must contain a self-description regarding their ability to be combined or converted into a machine or plant very rapidly and robustly. The following examples illustrate these requirements: • A new network-enabled field device, for example a drive with a new version of firmware, is hooked up to the production line. The new device must be provided automatically with network connectivity and be made known to all online subsystems. The participating systems must correspondingly be updated. • An unconfigured field device is introduced to production, for example to quickly replace another defective device. The field device now must be individualized and parameterized due to the information located in the software components. • A production facility is converted or modified because a new product variation is planned. The control and software related changes must be detected and automatically transmitted to all participating systems. • After conversion of a plant, it should be possible to move software components for process management around the decentralized control units, while observing certain criteria, such as output or availability. • A (new) function of the Manufacturing Execution System (MES) is inserted or altered, for example the visualization of a situation not previously required. The visualization should be done automatically and access to the necessary information from the field level should also be automatic. This requires the mechanical engineer to design the internal development processes accordingly. Modular machines require “modular” engineering, based on libraries of re-usable modules (“platform development”). Machine architecture must be designed such that combinable mechatronic modules are created, including the Plug & Produce capability of production modules using

<table>
<tr><td></td><td></td><td colspan="3">interoperable interfaces and adaptive automation technology. This requires development of concepts for “services” across manufacturer boundaries, such as archiving, alerting or visualising, as well as a low-cost integration of MES functions.

Effect on value chains
Value added is shifted from the system integrator to the machine provider or its supplier, because the machines or components are enhanced so that they are easier to integrate. The type and quality of system integration change. The present focus on (production) technology shifts to a stronger focus on organization and business processes related to production processes. In extreme cases, the system integrator could become obsolete if intelligent, self-configuring and interoperable manufacturing modules can be created at the level of the machine suppliers.

Value added for participants
For manufacturing companies, a quick, inexpensive and reliable conversion of manufacturing becomes possible, so that they can react quickly to changes in customer and market demands. Increasing standardization and modularization also expand the possibilities for combining manufacturing entities of various providers and therefore realizing the most economic solution for each individual module.

Machine modularization opens up new areas with scale effects for machinery manufacturers.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Component suppliers (sensors, actuators), Machine builders, system integrators, plant operators (manufacturer)</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4"></td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Automatic reasoning (e.g. [7,8]), AI (task) planning (e.g. [4,6]), distributed coordination and negotiation (e.g. [5])</td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3"></td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Standardization needs for setting up this use case is currently under further investigation. Some initial intentions on standardization needs are the following: a vocabulary with formal semantic for symbolic reasoning about production capabilities across different vendors, standardized negotiation mechanisms,</td></tr>
</table>

	standardized autonomy classes of components, machines, etc. Quality model for trustful learned models and automatic behavior resulting from it.	
Challenges and issues		
Societal Concerns	Description	Enabling flexible and autonomously reconfigurable production systems ease human-machine configuration, facilitate optimized machine use, reduce failures through autonomous compensation, optimized product quality through prediction techniques.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.46.2 References

[1] Working Group on Research and Innovation of the Plattform Industrie 4.0. Aspects of the Research Roadmap in Application Scenarios, Working Paper, German Federal Ministry for Economic Affairs and Energy, url: https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/aspects-of-the-research-roadmap.html , 2016.

[2] Working Group on Research and Innovation of the Plattfom Industrie 4.0 and Alliance Industrie du Futur: Plattform Industrie 4.0 & Alliance Industrie du Futur : Common List of Scenarios. url: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/plattform-i40-und-industrie-du-futur-scenarios.html, 2018

[3] Communication Promoters Group of the Industry-Science Research Alliance and German National Academy of Science and Engineering. Recommendations for implementing the strategic initiative INDUSTRIE 4.0, Final report of the Industrie 4.0 Working Group, url: https://www.acatech.de/Publikation/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group, April 2013

[4] Christoph Legat and Birgit Vogel-Heuser. A configurable partial-order planning approach for field level operation strategies of PLC-based industry 4.0 automated manufacturing systems. Engineering Applications of Artificial Intelligence 66:128-144, DOI: 10.1016/j.engappai.2017.06.014,02017.

[5] Birgit Vogel-Heuser, Jay Lee, and Paolo Leitao. Agents enabling cyber-physical production systems. at – automatisierungstechnik 63(10). DOI: 10.1515/auto-2014-1153, 2015.

[6] Jens Otto and Oliver Niggemann. Automatic Parameterization of Automation. Software for Plug-and-Produce. AAAI Workshop on Algorithm Configuration, 2015

[7] Christoph Legat, Christian Seitz, Steffen Lamparter und Stefan Feldmann. Semantics to the Shop Floor: Towards Ontology Modularization and Reuse in the Automation Domain. IFAC Processings, Vol. 47, Issue 3, pp. 3444 – 3449, Doi:10.3182/20140824-6-ZA-1003.02512, 2014.

[8] Martin Ringsquandl, Steffen Lamparter, Sebastian Brandt, Thomas Hubauer, and Raffaello Lepratti. Semantic-Guided Feature Selection For Industrial Automation Systems. International Semantic Web Conference. Springer, Cham, 2015.

A.47 Order-Controlled Production

A.47.1 General

ID	47	
Use case name	Order-Controlled Production	
Application domain	Manufacturing	
Deployment model	Cloud Services	
Status	Prototype	
Scope	Automatic distribution of production jobs across dynamic supplier networks	
Objective(s)	The objective is to enable automatic supplier contracting for optimized utilization of manufacturing capabilities at suppliers, novel degrees of flexibility in contract manufacturing, and enable (mass) customized customer ordering	
Narrative	Short description (not more than 150 words)	A network of production capabilities and capacities that extend beyond factory and company boundaries allows for a quick order-controlled adaption to changing market and order conditions. The result is a largely fragmented and dynamic value chain network that change as required by the individual order, and thereby make the best use of capabilities and capacities of existing production facilities. The goal is to allow for automated order planning, allocation and execution, thereby considering all production steps and facilities required to facilitate linking external factories into a company's production process, as automated as possible.
	Complete description	Use Case description taken from [1,2,3]. Many contemporary products are changing at an ever-in-creasing rate. Whereas up until just recently, smartphone displays were flat, the first curved displays are already on the market. The array of materials used in the automotive sector is also continually expanding – from aluminum, to high-strength steels and even fiber-reinforced plastics, today many types of materials are used. Innovation and product cycles are getting shorter all the time, and new production technologies are putting pressure on manufacturing companies to react more and more rapidly and make quick investment decisions regarding both consumer goods and investment goods. In order to confront this trend and avoid lengthy investment decisions, companies are starting to increase the network of their production capabilities beyond their own company boundaries. **Key aspects** The Order-Controlled Production application scenario describes a flexible manufacturing configuration. Owing a network of production capabilities and capacities that extend beyond factory and company boundaries, this company can quickly adapt to a changing market and order conditions, and thereby make the best use of capabilities and capacities of existing production facilities. In this way the

		potential provided by a network to other factories out-side of the company's own facilities is used to align the company's own portfolio – and especially its production – to quickly changing customer and market demands. Specifically, manufacturing chains are optimized for various parameters, such as cost and time. At its core, order-controlled production is based on standardization of the individual process steps on the one hand and the self-description of production facility capabilities on the other hand. This standardization allows for auto-mated order planning, allocation and execution, thereby considering all production steps and facilities required. This helps to combine individual process modules much more flexibly and earlier than previously possible, and to make use of their specific capabilities. In this respect, companies offer their available production capacities to other companies and thereby increase the utilization of their own machinery. Other companies may access these capacities as needed, thereby temporarily expanding their own production spectrum. In so doing, available production capacities are utilized better and order fluctuations can be smoothed out. The goal is to facilitate linking external factories into a company's production process, as automated as possible. In particular, the order placement process required for this should be executed automatically. **Effect on value chains** Today's relatively rigid and separately negotiated relation-ships between companies along the value chain will be transformed into a largely fragmented and dynamic value chain network that changes as required by the individual order. This applies both horizontally over the entire manufacturing process as well as vertically, with regard to production depth. Manufacturing companies focus on value-added steps that distinguish them significantly from other competitors. The possibility of creating fast and global client-manufacturer relationships can lead to unexpected competitive situations, because companies may change their role from order to order. Dynamically integrating production capacities will lead to better machine utilization and, as a result, diminishing demand for machinery suppliers. **Value added for participants** On the one hand, manufacturing companies will be able to automatically expand their production capabilities and capacities ad hoc in line with demand, by utilizing external production modules. No investment is required. This enables companies to react very flexibly to changing market and customer demands. On the other hand, companies offering their machines on the market can optimize their utilization rates.

Stakeholders	Customer, Producing companies, Broker			
Stakeholders' assets, values	Customer orders a good via the broker (separate stakeholder), Producing companies operate factories and machine parks.			
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
AI features	Task(s)	Automatic reasoning, AI (task) planning, distributed coordination and negotiation (cf. [5-8] for details and overview)		
	Method(s)			
	Hardware			
	Topology			
	Terms and concepts used			
Standardization opportunities/ requirements	Standardization needs for setting up this use case is currently under further investigation. Some initial intentions on standardization needs are the following: Standardization of data formats and semantic for exchanged data is enabler for this use case where multiple companies and institutions are involved (formal semantics for reasoning about 3d models, task decomposition and planning), standardization of interaction protocols between participants (esp. coordination and negotiation) enables automatic cross-company contracting.			
Challenges and issues				
Societal Concerns	Description	Enabling mass-customized production in global dynamic supply chains, and by that, ease production of small lot sizes for customized products.		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.47.2 References

[1] Working Group on Research and Innovation of the Plattform Industrie 4.0. Aspects of the Research Roadmap in Application Scenarios, Working Paper, German Federal Ministry for Economic Affairs and Energy, url: https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/aspects-of-the-research-roadmap.html , 2016.

[2] Working Group on Research and Innovation of the Plattfom Industrie 4.0 and Alliance Industrie du Futur: Plattform Industrie 4.0 & Alliance Industrie du Futur : Common List of Scenarios. url: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/plattform-i40-und-industrie-du-futur-scenarios.html, 2018

[3] Communication Promoters Group of the Industry-Science Research Alliance and German National Academy of Science and Engineering. Recommendations for implementing the strategic initiative INDUSTRIE 4.0, Final report of the Industrie 4.0 Working Group, url: https://www.acatech.de/Publikation/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group, April 2013

[5] Birgit Vogel-Heuser, Jay Lee, and Paolo Leitao. Agents enabling cyber-physical production systems. at – automatisierungstechnik 63(10). DOI: 10.1515/auto-2014-1153, 2015.

[6] Weyer, Stephan, et al. "Towards Industry 4.0-Standardization as the crucial challenge for highly modular, multi-vendor production systems." Ifac-Papersonline 48.3 (2015): 579-584.

[7] Lasi, Heiner, et al. "Industry 4.0." Business & information systems engineering 6.4 (2014): 239-242.

[8] Monostori, László. "Cyber-physical production systems: Roots, expectations and R&D challenges." Procedia Cirp 17 (2014): 9-13.

A.48 Value-based Service

A.48.1 General

ID	48	
Use case name	Value-based Service	
Application domain	Manufacturing	
Deployment model	Hybrid deployment: Cloud and on-premise deployment in the production field	
Status	PoC	
Scope	Process and status data from production and product use sources are the raw materials for future business models and services.	
Objective(s)	The objective of this use case is the provision of remote services for product and production based on (generic) service platforms. This use case can be seen as a fundament for the deployment of arbitrary AI remote services.	
Narrative	Short description (not more than 150 words)	Service platforms collects data from product use – for example machines or plants – and analyses and processes this data to provide tailor-made individualized services, e.g. optimized maintenance at the proper time, or the timely provision of the correct process parameters for a production task currently being requested. Companies offering these services (service providers) occupy the interface between the product provider and the user.
	Complete description	Use Case description taken from [1,2,3]. In the consumer area, the increased interconnectivity of users which has made it possible to collect user data has made a whole new range of services possible. For example, navigation systems in our cars not only determine the shortest route, but also the quickest, as the traffic situation is assessed in real time based on movement data from other users. Entertainment media is no longer purchased rather made available as needed using streaming services. The services offered extend beyond simply making the products available. The individual customer receives optimized offers, based on user data: the quickest route during rush hour, or music tailored to that customer’s taste. Similar developments are occurring in an increasingly interconnected industrial environment. Services that go significantly beyond simply providing a production unit – a

contemporary example is leasing – are gaining in importance and are changing the classic value-added processes and business models.

Key aspects

At the heart of this application scenario are IT platforms that collect data from product use – for example machines or plants for production purposes – and analyze and process this data to provide tailor-made individualized services. This could include for example optimized maintenance at the proper time, or the timely provision of the correct process parameters for a production task currently being requested. The collected data could be product parameters, for example the machines and plants required for manufacture, the product status information, or data from the production process or the upstream supply process. Even the characteristics of the processed raw materials or the parts of the product could be included. The goal is to use this data as a raw material for optimizing products and production processes and for new services. This can help to not only improve existing value chains but also perhaps create new value-added elements.

Effect on value chains

The industrial environment today is influenced in principle by two actors – the product provider (i.e. manufacturers of production facilities and service providers) and the customer (product users, i.e. production facility operators), who work together with varying degrees of intensity.

With the introduction of Value-Based Services an additional actor enters the scene, operating IT platforms that it uses to provide new services to both classic partners. This platform operator could be a new element of the value chain, that is, an autonomous company. However, this role could be taken on by product providers by increasing their value added compared with the current situation.

Product providers make their product data and parameters available. On the basis of all of this user data, new services can now be developed, such as individual optimized maintenance or specific operating and process parameters that optimize or even expand production capabilities of the existing infrastructure. The companies offering these services (service providers) occupy the interface between the product provider and the user. The result is that the share in the value chain spanning from the product provider to the user can be shifted significantly, compared with the situation today. The user can then distinguish between the products by considering the accompanying services or the possibility of expanding those services even after purchasing the product, and no longer primarily by the (physical) specifications mandated by the product provider. This makes it very attractive for the product provider to use such platforms and to offer new services on them.

<table>
<tr><td></td><td></td><td colspan="3">Value added for participants
In this application scenario the value added for the product provider stems from the availability of a multitude of process data from various application scenarios, which the user can apply to further development of its product port-folio. As an operator of related IT platforms, the product provider can offer new services. In this way, it strengthens customer loyalty and increases its portion of value added.

Value added for the user, on the other hand, can come from better utilization of the product, enhanced product availability from improved maintenance, for example, or optimized product use as a result of optimally adapted product parameters.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Customer (product user), platform provider, service provider, product provider</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Reasoning and autonomous problem solving in the platform, services based on the platform use AI features, e.g. for predictive maintenance, data semantics (cf. [5,6] for an overview)</td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3"></td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Standardization needs for setting up this use case is currently under further investigation. Some initial intentions on standardization needs are the following: For this use case, standardization can be seen as enabler because an agreement on a (small set of) communication protocols would facilitate to connect to the platform and use this protocol also for device2device communication. Since services running on a platform are not aware of an implicit sematic of data sources (machines, sensors, actuators, …), an explicit semantic or a common vocabulary is need describing data and enable reasoning about machine states on premise (on the machine/edge) as well as on the cloud. For cloud2cloud communication and cloud federation, further interoperability standards are required on communication level as well as on data semantics level.</td></tr>
<tr><td>Challenges and issues</td><td colspan="4"></td></tr>
<tr><td>Societal Concerns</td><td>Description</td><td colspan="3">Increasing complexity of modern cyber-physical production systems cannot be managed by humans. AI technologies provide one solution in this context for more reliable, fault-tolerant, safe and secure production systems.</td></tr>
</table>

	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.48.2 References

[1] Working Group on Research and Innovation of the Plattform Industrie 4.0. Aspects of the Research Roadmap in Application Scenarios, Working Paper, German Federal Ministry for Economic Affairs and Energy, url: https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/aspects-of-the-research-roadmap.html , 2016.

[2] Working Group on Research and Innovation of the Plattfom Industrie 4.0 and Alliance Industrie du Futur: Plattform Industrie 4.0 & Alliance Industrie du Futur : Common List of Scenarios. url: https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/plattform-i40-und-industrie-du-futur-scenarios.html, 2018

[3] Communication Promoters Group of the Industry-Science Research Alliance and German National Academy of Science and Engineering. Recommendations for implementing the strategic initiative INDUSTRIE 4.0, Final report of the Industrie 4.0 Working Group, url: https://www.acatech.de/Publikation/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group, April 2013

[4] Bo-hu LI, Bao-cun HOU, Wen-tao YU, Xiao-bing LU, Chun-wei YANG. Applications of artificial intelligence in intelligent manufacturing: a review. Frontiers of Information Technology & Electronic Engineering. 2017

[5] Lee, Jay, Hung-An Kao, and Shanhu Yang. "Service innovation and smart analytics for industry 4.0 and big data environment." Procedia Cirp 16 (2014): 3-8.

A.49 AI Solution for Traffic Signal Optimization based on Multi-source Data Fusion

A.49.1 General

ID	49	
Use case name	AI solution for traffic signal Optimization based on multi-source data fusion	
Application domain	Transportation	
Deployment model	Cloud services	
Status	In operation	
Scope	Generate traffic signal timing plans by analyzing traffic flow status and patterns based on fusing internet data, induction coils data and video data, and control the traffic signal with the generated timing plans in a real-time, self-adaptive and cooperative way	
Objective(s)	To find an effective and efficient solution to improve the road utilization efficiency by increasing traffic flow speed and reducing traffic flow waiting time.	
Narrative	Short description (not more than 150 words)	An AI solution was developed that could recognize real-time traffic flow status and abstract traffic flow patterns by fusing internet data, induction coils data and video data, and could generate optimized traffic signal timing plan by self-adaptively responding to real-time traffic flow

		fluctuation and with regards to traffic flow coordination among multiple intersections within a given region.
	Complete description	By far, traffic administrator produces traffic signal timing plans by observing traffic flow situation on-site at intersections or through videos, and relies on her/his personal experience. Then, the timing plans are input into and executed by the traffic signal control system. The disadvantages of this manual traffic signal timing plan generation approach are as follows: 1. Low computing efficiency, it consumes very long time for traffic administrator to observe and analyze traffic patterns. 2. Low computing precision, traffic administrator only cares about the macro traffic flow tendency at intersections without computing detailed traffic parameters such as speed, queue length in each lane, etc. 3. Slow response to traffic flow fluctuation, it is hard for traffic administrator to produce adaptive timing plan in time with respect to real-time traffic flow fluctuation, due to her/his limited computing ability, not mention to coordinate traffic flows among multiple intersections by controlling the traffic signal in real-time. 4. Experienced traffic administrators are severely in short for cities with the scale of thousands intersections. For solving the above problems, the AI provider applies a multi-source data fusion approach to recognize the traffic flow status and generalize the traffic flow pattern by analyzing the internet data (i.e., vehicle driving trajectory data provided by internet service supplier), detector data collected by induction coils, and structured data recognized from videos. Furthermore, the AI provider develops an optimization method to figure out optimized traffic signal timing plan by self-adaptively responding to real-time traffic flow fluctuation and with regards to traffic flow coordination among multiple intersections. The developed methods have been applied in practice within a given region from a large city. It generates traffic signal timing plans for all the intersections in the region according to their real-time traffic flow fluctuation with an updating frequency of 5 minutes per time. Compared with the manual traffic signal timing plans form the traffic administrators, the plans generated by the new method have increased the average vehicle driving speed by 9%, and reduced the average vehicle waiting time by 15%.

Stakeholders	DOT DOP			
Stakeholders' assets, values	Safety, stability, trustworthiness			
System's threats & vulnerabilities	new privacy threats, new security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives

	1	Average vehicle driving speed	Average vehicle driving speed on all the road sections in a given region	Improve the road utilization efficiency
	2	Average vehicle waiting time	Average vehicle waiting time at all the intersections in a given region	Improve the road utilization efficiency
AI features	Task(s)	Optimization		
	Method(s)	Deep learning, Bayesian network, Time series analysis, Operational research optimization method (i.e., Mixed integer linear programming, etc.)		
	Hardware	ECS		
	Topology	Cloud Service		
	Terms and concepts used	Traffic signal self-adaptive and coordinative control for a large number of intersections.		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Traffic signal self-adaptive and coordinated control for a large number of intersections. Issues: 1. Not all intersections are equipped with detectors such as induction coil or video. 2. The detectors may output abnormal values which need data clean processing.			
Societal Concerns	Description	Relieve urban road congestion		
	SDGs to be achieved	Sustainable cities and communities		

A.49.2 Data

Data characteristics	
Description	Internet data, Induction coil data, Video data
Source	Internet, Detector, Detector
Type	Structured text and number, Structured text and number, Unstructured video
Volume (size)	
Velocity	Internet data updated daily, Induction coil data updated every 5 minutes, Video data updated in real-time
Variety	From multiple domains
Variability (rate of change)	Dynamic
Quality	Exists missing values or abnormal values

A.49.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train multiple models (deep learning,	Dataset is ready		

		Bayesian network, Time series analysis) for recognizing traffic flow volume and abnormal values in the input data			
2	Optimizatio n	Based on the data processed by the trained models, optimize the period length, split, and key phase offsets among multiple intersections for traffic signal timing plans	Completio n of training/re training	Completion of missing values or abnormal values processings	
3	Evaluation	Pre-evaluate the execution effects of the optimized traffic signal timing plans, which include the period lengths, splits, and key phase offsets among multiple intersections	Completio n of optimizati on	Input prediction of traffic flow situation in the next period	The pre-evaluated execution effects of the optimized traffic signal timing plan is superior to the current one
4	Execution	Execute the optimized traffic signal timing plan	Completio n of evaluation	The pre-evaluated execution effects of the optimized traffic signal timing plan is superior to the current one	

3704 A.49.4 **Training**

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Dataset is ready	Transform video data into structured data	AI provider	Transform video data into structured data by deep learning	
2	Completion of Step 1	Data clustering	AI provider	Recognize abnormal value patterns and label them in internet data, induction coil data, and structures video data by data clustering	

3	Completion of Step 2	Processing of missing value and abnormal value	AI provider	Recognize abnormal value and process them, and fill missing values by data clustering, time series analysis and Bayesian network	
4	Completion of Step 3	Data fusion	AI provider	Compute traffic status parameters such as traffic volume, vehicle driving speed, etc. by fusing internet data, induction coil data and structured video data	

Specification of training data	

A.49.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of optimization	Construct the evaluation model of the traffic signal timing plan	AI provider	Construct the evaluation model of the traffic signal timing plan based on traffic engineering theory	
2	Completion of Step 1	Evaluate the effect of the computed traffic signal timing plan	Traffic administrator	Pre-evaluate the effect of the computed traffic signal timing plan with the evaluation model	

Input of evaluation	
Output of evaluation	

A.49.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of evaluation	Execute the computed traffic signal timing plan	Traffic administrator	Input the computed traffic signal timing plan into the traffic signal control system and execute it	The pre-evaluated execution effects of the optimized traffic signal timing plan is superior to the current one

Input of Execution	
Output of Execution	

A.49.7 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	patent	ZHANG MAOLEI;WEI LIXIA;CHEN XIAOMING;LI JIN., ”Crossing traffic jam judging and control method and system based on sensing detectors ”.CN201310395431.2013			QINGDAO HISENSE TRANS TECH CO	http://www.pss-system.gov.cn/sipopublicsearch/patentsearch/showViewList-jumpToView.shtml
2	patent	ZHANG MAOLEI;WEI LIXIA;CHEN XIAOMING;LIU XIN;LIU HONGMEI;LI JIN.“Multi-strategy and multi-object self-adaptation traffic control method”. CN201310548921.2013			QINGDAO HISENSE TRANS TECH CO	http://www.pss-system.gov.cn/sipopublicsearch/patentsearch/showViewList-jumpToView.shtml
3	patent	WANGMENGJIA;MINWANLI.” Road traffic optimization method and device and electronic equipment			ALIBABA GROUP HOLDING LTD;	http://www.pss-system.gov.cn/sipopublicsearch/patentse

		". CN201710081075.2017				arch/showViewList-jumpToView.shtml
4	patent	HUA XIANSHENG" Assessment method and device of traffic condition ". CN201610645412.2016			ALIBABA GROUP HOLDING LTD;	http://www.pss-system.gov.cn/sipopublicsearch/patentsearch/showViewList-jumpToView.shtml
5	patent	HUA XIANSHENG,REN PEIRAN,SHEN CHEN,CHU WENQING,LIU YAO." Road intersection traffic flow control method and device". CN201610644132.2016			ALIBABA GROUP HOLDING LTD;	http://www.pss-system.gov.cn/sipopublicsearch/patentsearch/showViewList-jumpToView.shtml
6	paper	Liang Yu,Jingqiang Yu,Maolei Zhang, Xin Zhang, Yuehu Liu."Large Scale Traffic Signal Network Optimization-a Paradigm Shift Driven by Big Data". ICDE2019			Alibaba Cloud Computing Hangzhou, China	
7		M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y.Wang, "Review of road traffic control strategies," Proceedings of the IEEE, vol. 91, no. 12, pp. 2043–2067, 2003.				
8	paper	P. Lowrie, "Scats, sydney co-ordinated adaptive traffic system: A traffic responsive method of controlling urban traffic," 1990.				
9	paper	F. Corman, A. D'Ariano, D. Pacciarelli, and M. Pranzo, "Evaluation of green wave policy in real-time railway traffic management," Transportation Research Part C: Emerging Technologies, vol. 17, no. 6, pp. 607–616, 2009.				
10	paper	L. Singh, S. Tripathi, and H. Arora, "Time optimization for traffic signal control using genetic algorithm," International Journal of Recent Trends in Engineering, vol. 2, no. 2, p. 4, 2009.				

3712

A.50 AI Solution to Quality Control of Electronic Medical Record(EMR) in Real Time

A.50.1 General

ID	50	
Use case name	AI solution to quality control of Electronic Medical Record(EMR) in real time	
Application domain	Healthcare	
Deployment model	Cloud services	
Status	In operation	
Scope	Detecting defects in EMR by inspecting unstructured data based on Natural Language Processing(NLP) ability	
Objective(s)	To insure the completeness, consistency, punctuality and medical-compliance of EMR written by physicians	
Narrative	Short description (not more than 150 words)	This AI solution in ET Medical Brain Medical service support system was developed that could simultaneously detect mistakes while physicians wrote EMR（Electronic Medical Record）. Using NLP（Natural Language Processing） ability, it can process a large amount of unstructured text and judge the accuracy according to recognized medical reference. It achieved 80% coverage of all the EMR quality control requirements issued by Chinese government, and human labour of EMR QC （Quality Control ） was reduced 60%, which translated into cost savings, and enhanced physician education.
	Complete description	Medical records are the records of the occurrence, development and prognosis of patients' diseases, as well as the medical activities such as examination, diagnosis and treatment. A high-quality medical record has great value at medical and legal level. When medical records are converted from handwritten to electronic input, delayed, uncompleted writing and copying are endangering the quality of medical records. Once the medical record data does not meet the requirements, it will greatly affect the health of patients, the development of medicine and the judgment of responsibility in medical accidents. Nowadays, hospital has a Medical Records Department to control medical records quality manually. However, as the number of medical records increases, the inspection requirements become more complex, and the medical professional knowledge requirements are improved, so the medical records quality inspection becomes harder. The intelligent electronic medical record quality control

		system is based on NLP. When a doctor writes medical records, it can analyze unstructured medical record text, and control the quality based on government requirements, ensure the integrity, consistency, timeliness and compliance of medical records. ET（Evolutionary Technology） Medical Brain Medical service support system has learning ability to learn more medical knowledge including clinical pathway, drug compatibility taboo etc. it can learn the habits and rules of doctor's manual review to inspects records profoundly. The current system has covered 189 medical records quality inspection requirements, saved 60% review time for medical record department, which greatly saved the cost of the hospital, reduced the inspection time and repeated work, and will help doctors put more energy into the education and training.		
Stakeholders	Doctor, Hospital, Patient			
Stakeholders' assets, values	Safety, privacy, fair treatment, trustworthiness			
System's threats & vulnerabilities	New privacy threats, new security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Coverage	Ratio of EMR QC requirements done in the solution/all issued EMR QC requirements in China. Ideal target is 100%.	Improve accuracy
AI features	Task(s)	Natural language processing		
	Method(s)	SimHash		
	Hardware	ECS		
	Topology	Cloud Service		
	Terms and concepts used	Jaccard index		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Achieve all EMR QC requirements in different disease areas Issues: 1) Lack of medical reference data 2) Lack of medical knowledge graph			
Societal Concerns	Description	Achieved 80% coverage of all the EMR quality control requirements issued by Chinese government, and human labour of EMR QC （Quality Control） was reduced 60%, which translated into cost savings, and enhanced physician education.		

	SDGs to be achieved	Good health and well-being for people

A.50.2 Data

Data characteristics	
Description	EMR text data
Source	EMR system
Type	Text data from EMR system vendor
Volume (size)	
Velocity	Real time
Variety	Multiple datasets
Variability (rate of change)	Static
Quality	High (depending on EMR system)

A.50.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training samples	Sample raw dataset is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		
3	Execution	Detect defects (regions including defects) using the trained model	Completion of deploymen t in EMR system	The trained model has been evaluated as deployable	
4	Retraining	Retrain a model with training samples	Certain period of time has passed since the last training/re trainig		

A.50.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Raw data preparation	Raw data to cloud	AI solution provider	Transform sample raw data from EMR system to server on cloud	The software for data transform has to be provided by the AI solution provider.
2	Completion of Step 1	Training sample creation	AI solution provider	Create training samples by labelling the output of Step 1 with "defective"/"non -defective"	
3	Completion of Step 2	Model training	AI solution provider	Train a model (deep neural network) with the training samples created by Step 2	

Specification of training data	

A.50.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Preparation	AI solution provider	Transform sample raw data from EMR system to server on cloud	
2	Completion of Step 1	Detection	AI solution provider	Given the image data from Step 1, detect defects (regions including defects) using the deep neural network trained in the scenario of training	

3	Completion of Step 2	Evaluation	Manufacturer	Compare the result of Step 2 with that of human inspection	

Input of evaluation	
Output of evaluation	

A.50.6 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1						https://et.aliyun.com/brain/healthcare?spm=a2c17.92424.1146454.87.254f1a43dCNCpb
2	Patent	A medical symptom knowledge base classification system construction algorithm and device based on lexical cluster similarity	In application			Patent number: 100424310
3	Patent	Electronic medical record named entity recognition method and device combining Section feature information	In application			Patent number: 100557465
4	Patent	Algorithm and device for recognizing nested medical named entities based on two-layer recurrent neural network	In application			Patent number: 100609063
5	Patent	Algorithm and device for unsupervised keyword-based medical image report key information extraction	In application			Patent number: 100619640
6	Patent	Medical record text structure analysis algorithm and device based on pseudo corpus generation	In application			Patent number: 100558223
7	Patent	Algorithm and device for improving accuracy of medical record quality	In application			Patent number: 100558228

		assurance system by using doctor behavior log				
8	Patent	Medical record text structure analysis algorithm and device based on context-free grammar parsing technology	In application			Patent number: 100549098
9	Patent	Algorithm and device for structural analysis of medical records combined with visual features	In application			Patent number: 100605377
10	Patent	Method and device for Chinese medical record named entity recognition by using Iterated Dilated CNN with condition random field model based on Chinese character structure	In application			Patent number: 100554136
11	Patent	Method and device for Chinese medical field relationship extraction by using residual convolution attention network model	In application			Patent number: 100558469
12	Patent	Method and device to detect similar electronic medical records	In application			

A.51 Machine Learning Tools in Support of Transformer Diagnostics

A.51.1 General

ID	51	
Use case name	Machine Learning Tools in Support of Transformer Diagnostics	
Application domain	Performance evaluation and diagnostics	
Deployment model	Prototype	
Status	Under development	
Scope	Power Transformers operation and maintenance	
Objective(s)	Use of Machine Learning (ML) algorithms as supporting tools for the automatic classification of power transformers operating condition	
Narrative	Short description (not more than 150 words)	The successful use of ML tools may find multiple applications in the industry such as providing fast ways of analysing new data streaming from online sensors, evaluating the importance of individual variables in the context of transformer condition assessment and also the need or adequacy of data imputation in the so widely common problem of missing data
	Complete description	The work consists of training 12 ML algorithms with real data from 1,000 (one thousand) transformers that were individually analyzed by human experts.

		Each transformer in the database is scored with a ‘green’, ‘yellow’ or ‘red’ card depending on the data, the interpretation of human experts, or even after some calculations carried out by the company’s internal algorithms frequently utilized by the experts to identify units with technical operational issues. The ML algorithms, however, do not utilize or are given any of the engineering tools employed by the human experts. The algorithms only employed the raw data in a supervised learning process in which a column named ‘Class’ was added to the transformer information with the classification red, yellow or green provided by the human expert.		
Stakeholders	Transformers end users			
Stakeholders’ assets, values	Enhanced diagnostic of transformer fleet with consequent improvement on predictive maintenance and therefore electrical grid reliability			
System’s threats & vulnerabilities	Lack of enough data to perform the analysis			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Algorithm accuracy	Output when compared to the human expert analysis of the same data	See reference
AI features	Task(s)	Statistical learning		
	Method(s)	12 ML methods used for the comparison exercise: Linear Algorithms 1. General linear regression (logistic regression) - GLM 2. Linear discriminant analysis - LDA Non-linear Algorithms 1. Classification and regression trees (CART and C5.0) 2. Naïve Bayes algorithm (NB) 3. K-Nearest Neighbor (KNN) 4. Support Vector Machine (SVM) Ensemble Algorithms 1. Random Forest (stochastic assembly of a large number of CART algorithms) 2. Tree Bagging (Tree Bagging) 3. Extreme Gradient Boosting Machine (xGBM1 and xGBM2) 4. Artificial Neural Networks (ANN)		
	Hardware	Standard laptop		
	Topology	NA		
	Terms and concepts used	Machine Learning Algorithms, Transformer Diagnostics, Condition Assessment, Automated Tool		
Standardization opportunities/ requirements	Standardization of asset performance data format and analysis			

Challenges and issues	Data availability, missing data, imbalanced classes	
Societal Concerns	Description	Safe and reliable power delivery
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.51.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Conference	Cheim, Luiz V. Machine Learning Tools in Support of Transformer Diagnostics Cigre General Session Paris 2018, paper reference A2-206	Presented in Aug 2018	Use case taken from this reference	ABB	Cigre web page

A.52 Automated Travel Pattern Recognition using Mobile Network Data for Applications to Mobility as a Service

A.52.1 General

ID	52	
Use case name	Automated Travel Pattern Recognition using Mobile Network Data for Applications to Mobility as a Service	
Application domain	Other (please specify) Transport	
Deployment model	Activity- based Modelling for New mobility Services	
Status	PoC	
Scope	Detect automatically travel pattern recognition from anonymized and aggregated Mobile phone Network Data	
Objective(s)	Phase 1: Attribute trip purpose and mode of transport to multimodal door-to-door journeys from Mobile phone Network Dataset using AI and machine learning techniques (Activity based model) Phase 2: Generate daily activities for static agents in the Agent Based Model Phase 3: Optimisation of New Mobility services in integration with mass transit	
Narrative	Short description (not more than 150 words)	Activity- based modelling has the capability to exploit big data source generated by smart cities to create a digital twin of urban environments to test Mobility as a Service schemes. MND data have been used to create activities for an Agent Based Model. AI is used to automatically detect purpose and mode of transport in multimodal round trips, obtained by anonymized and aggregated MND trip-chains dataset. Data fusion techniques and SQL queries were also used to consider land

		use and facilities in the urban area of interest.		
	Complete description	Activity- based modelling has the capability to exploit big data source generated by smart cities to create a digital twin of urban environments to test Mobility as a Service schemes. Given the rise of location- based data and Mobile phone Network Data (MND) for transport modelling purpose, Agent based modelling has become a viable tool to explore a sustainable introduction of mobility services, exploring the integration with mass transit. AI is used in detecting purpose and mode of transport in multimodal round trips and assign purpose and mode of transport to trip- chains dataset coming from MND. The methodology has been developed for the Innovate UK funded Mobility on Demand Laboratory Environment (MODLE) project and will undergo a validation process during the Demand Modelling and Assessment through a Network Demonstrator (DeMAND) project for the Department for Transport (UK)		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Generation of Activities (land use information and time of travel)	Purpose of activities is assigned based on land use information and time of travel. Cnesus data and national/ local travel surveys will provide validation for the process	Phase 1
	2	Generation of agents (travel times, speed on links)	Agents generated will build up in the network creating realistic conditions of congestion. Speed on links	Phase 2
		Operation of service (number of users for the service)	Optimisation of route and operation time in the day. Validation provided using data collected by Mobility service operators during the operation of service	Phase 3
AI features	Task(s)	Assign purpose of each trip in the chain, assign model of transport for each trip in the chain, generate daily activity plans, generate static agents (users), generate dynamic agents (service)		
	Method(s)	Agent Based Models with Activity based approach		
	Hardware	NA		

	Topology	
	Terms and concepts used	Data fusion, machine learning techniques
Standardization opportunities/ requirements		
Challenges and issues	The use of Mobile Phone Network data is still not precise for shorter trips and internal trips which might be not detected. However, with the introduction of 5G, MND will be even more reliable and available to use in transport modelling.	
Societal Concerns	Description	The use of anonymization techniques minimise the risk of disclosing personal information when analyzing location based data and Mobile phone Network Data
	SDGs to be achieved	

A.52.2 References

[1] Franco P, Johnston R, McCormick (2019) Demand Responsive transport: generation of activity patterns from mobile phone network data to support the operation of flexible mobility services. - Special issue of Transportation Research Part A (TRA) on developments in Mobility as a Service (MaaS) and intelligent mobility (forthcoming) https://www.sciencedirect.com/journal/transportation-research-part-a-policy-and-practice/vol/121/suppl/C

[2] Franco P, Johnston R, McCormick E (2018) Role of Intelligent Transport Systems applications in the uptake of mobility on demand services, United Nation “Transport and Communications Bulletin for Asia and the Pacific, 2018, No. 88 - Intelligent Transport Systems”, https://www.unescap.org/sites/default/files/Ch02-Role%20of%20Intelligent%20Transport%20Systems%20%28ITS%29%20applications%20in%20the%20uptake%20of%20mobility%20on%20demand%20services_0.pdf

[3] Franco P, McCormick E, Johnston R (2018) Multimodal activity Modelling for supporting mobility service operations, ITS World Congress Copenhagen, 17-21 September 2018

[4] Franco P, McCormick E, Van Leeuwen K, Ryan Johnston, Gregor Engelmann (2017) Multi-Modal Activity-Based Models to support Flexible Demand Mobility Services. ITS World Congress 2017, Montreal 29 October- 2 November 2017. Awarded Best Paper

[5] Franco P, McCormick E, Van Leeuwen K (2017) Framework for modelling MaaS using ABM and real-time data from ride-sharing services. 12ve ITS Europe Congress 2017, Strasbourg, 19-22 June 2017. Proceedings

A.53 Improving conversion rates and RoI (Return on Investment) with AI technologies

A.53.1 General

ID	53
Use case name	Improving conversion rates and RoI (Return on Investment) with AI technologies
Application domain	Digital marketing
Deployment	On-premise systems

<table>
<tr><td>model</td><td colspan="4"></td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Utilizing AI technologies in digital marketing</td></tr>
<tr><td>Objective(s)</td><td colspan="4">1) help the operation team identify new business scenarios and seize more market opportunities,
2) increase conversion rate and marketing effectiveness,
3) improve user experience by providing individually customized services</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Personalized digital marketing has become increasingly important in response to the needs of providing different services to different consumers. The combination of big data and AI algorithms is the core of personalized digital marketing. By modelling user preferences, we can predict the services that users may be interested in, improve marketing effectiveness and enhance user experience.</td></tr>
<tr><td>Complete description</td><td colspan="3">With the economic development, consumers are more emphatic about self-personality. Digital Marketing has also begun to focus more on the consumer's personality instead of the commonality. Personalized digital marketing has become increasingly important in response to the needs of providing different services to different consumers.

The combination of big data and AI algorithms is the core of personalized digital marketing. By modelling user preferences, we can predict the services that users may be interested in, improve marketing effectiveness and enhance user experience. There are three main parts of personalized marketing technology: 1) Audience Targeting: Forecasting people who may be interested in the marketing activities, focusing on high-conversion probability populations to increase conversion rates; 2) Smart subsidy: Different marketing subsidies for different users to achieve higher conversion rates at lower cost ; 3. Personalized Recommendation : Predict user preferences for services or items, and recommend to users what they are most likely to be interested in, to increase conversion rates.

Through the application of AI technology, personalized digital marketing has achieved very significant results: the predicted population’s conversion rates has achieved more than 30% improvement; in subsidy scenario it has achieved a cost reduction of more than 10% while the 2% increase in conversion rate; in the coupon recommendation scenario, the conversion rate has been improved by more than 70%.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Third-party payment companies, end users, merchants</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">User experience, digital marketing RoI, conversion rate, marketing cost</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Abuse of personal information, Falsified or dirty data</td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td></td><td>Conversion rate</td><td>the percentage of users who accept the marketing</td><td>To increase the conversion rate</td></tr>
</table>

			(e.g., clicks) out of the total number of visitors	
		RoI	RoI=conversion_rate*(1-k*cost) k is the cost impact factor and it can be adjusted to get higher conversion rate or lower cost	To increase the marketing effectiveness
AI features	Task(s)	Audience Targeting, Smart Pricing, Personalized Recommendation		
	Method(s)	Machine learning, Deep learning		
	Hardware			
	Topology			
	Terms and concepts used	Attribution Analysis, Fatigue control, Smart Pricing, Off-line Batch Computation, OLAP Analysis		
Standardization opportunities/ requirements	• Technical framework of AI-enabled digital marketing system • Guidelines for collecting, storing and handling of digital marketing data • Guidelines for applying AI technology to digital marketing			
Challenges and issues	• How to collect, utilize and protect user information within the scope of what is permitted by relevant national and regional legislation and regulations • How to let the system evolve and improve continuously with applying new AI models and algorithms ……			
Societal Concerns	Description	For Users: enjoy better service at a lower cost For Merchants: Increase profits and decrease costs For Cities and communities: Promote economic prosperity and develop green economy		
	SDGs to be achieved	Sustainable cities and communities		

A.53.2 Data

Data characteristics	
Description	sample and feature data of marketing campaign
Source	Customers
Type	Log Text
Volume (size)	~500GB/day
Velocity	Stream and batch
Variety	Device information, location information, conversion information (clicks, transactions), active level
Variability (rate of change)	Subject to digital marketing effort (Festival, on sale)
Quality	Vary (depending on position of data collection and data reflow mechanism)

A.53.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training				
2	Evaluation				
3	Execution				
4	Retraining				

A.53.4 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Journal		Published online	implementation	Ant Financial Services Group	https://martech.alipay.com

A.54 bioBotGuard

A.54.1 General

ID	54	
Use case name	bioBotGuard	
Application domain	Agriculture	
Deployment model	Cloud services	
Status	PoC	
Scope	Use visual recognition to identify and help fight parasites attacking organic farms.	
Objective(s)	The use case shows how AI contributing to modernize Agriculture industry.	
Narrative	Short description (not more than 150 words)	BioBotGuard defines itself as an initiative of Precision Farming as a Service. From an IT perspective it uses drones with GPS and high-resolution cameras to monitor the crops; the images are then processed by computer vision API in order to spot diseases and harmful insect attacks, building a georeferenced risk map of the crop. This can be used to send operational drones to put the treatment (or antagonist insects) only when and where it is needed.
	Complete description	BioBotGuard main goals are to cut the use of Phyto-sanitary treatments to contain the environmental health risk by estimating the probability of incubation and development of plant diseases or harmful insects attacks and anticipate treatments. BioBotGuard monitors microclimatic conditions with high accuracy measurement and prediction models to optimize irrigations. From the technology point of view, it employs: AgroDrones to patrol and map the culture filed that are equipped with 20Mx high-resolutions cameras to capture in real-time

		images. On the backend the drone send data to computer vision API for image classifications and pattern detections. Among others, the system is able to detect harmful insects and build a georeferenced risk map of the crop. As a result, bioBotGuard can help AgriFood producers to change the cost structure of the industry, by requiring less water and less treatment, as well as a significant reduction in labor costs.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Optimize Phyto-sanitary treatments	The objective is to contain the environmental health risk by estimating the probability of incubation and development of plant diseases or harmful insects attacks and anticipate treatments.	Improve healthy
	2	Reduced field mapping time	The objective is to reduce the time as well as achieve a more frequent monitoring time of the crop and the field microclimate.	Reduce Time
	3	Reduced Labor Costs	Reduction of the labor costs due to autonomous monitoring systems	Reduce Costs
AI features	Task(s)	Deep Learning, Pattern Recognition		
	Method(s)	Drones		
	Hardware			
	Topology	Drones, Agriculture, Image Recognition, Computer Vision		
	Terms and concepts used	Deep Learning, Pattern Recognition		
Standardization opportunities/ requirements				

Challenges and issues	Acquire filed as well as crop images at different distances and normalize image recognition and pattern detection.	
Societal Concerns	Description	
	SDGs to be achieved	

A.54.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1		bioBotGuard project Web site and presentation				https://www.blueit.it/biobotguard/ https://vimeo.com/238174241

A.55 RAVE

A.55.1 General

ID	55	
Use case name	RAVE	
Application domain	Learning	
Deployment model	Hybrid Cloud or other	
Status	PoC	
Scope	Use of advanced an multimodal sensing ability to facilitate a complex task	
Objective(s)	Avatar and social robot interact with deaf babies for facilitating language learning.	
Narrative	Short description (not more than 150 words)	RAVE system is an integrated multi-agent system involving a robot and virtual human designed to augment language exposure for 6-12 month old infants. The system is an engineered robot and avatar to provide visual language to effect socially contingent human conversational exchange. The team demonstrated the successful engagement of our technology through case studies of deaf and hearing infants.
	Complete description	The RAVE system is designed as a dual-agent that uses a physical robot and a virtual human to engage 6-12month old deaf infants in linguistic interactions. The system was bolstered by a perception system capable of estimating infant attention and engagement through thermal imaging and eye tracking. RAVE has been designed and experienced for a unique population (deaf infants) during a three period of observation and developing three case studies. This system has been successful at soliciting infant attention, directing attention to the linguistic content, and keeping the infant engaged for developmentally appropriate lengths of time. It has been also observed instances of infants copying robot behavior, of infants producing signs displayed by the avatar, and of infants producing signs to the non-signing robot agent that they

		had observed the virtual human perform. These initial experiences give the hope that longer-term exposure to a system based on this work may be able to impact long-term learning in this unique population.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Soliciting infant attention	The objective is to have a system able to capture the infant attention status and decode his "ready to learn" moment to provide content	Improve learner attention
	2	Keeping Infant engaged	The objective is to keep the learning engaged during the learning process	Improve learner engagement
AI features	Task(s)	Virtual Humans and 3D model reconstruction, Robot, Biometric status by using thermal cameras, eye tracking, Motion Capture		
	Method(s)	Deep Learning, Pattern Recognition		
	Hardware	Robot, Thermal Camera, Screen		
	Topology			
	Terms and concepts used	Learning, thermal camera, eye tracking, Image Recognition, Computer Vision		
Standardization opportunities/ requirements				
Challenges and issues	Ability to decode a learner cognitive status and his attention level.			
Societal Concerns	Description			
	SDGs to be achieved			

A.55.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1		Nex2U - RAVE Application with Thermal Camera				http://www.next2u-

						solutions.com/featured-projects/

[2] Brian Scassellati, Jake Brawer, Katherine Tsui, Setareh Nasihati Gilani, Melissa Malzkuhn, Barbara Manini, Adam Stone,Geo Kartheiser, Arcangelo Merla, Ari Shapiro, David Traum, Laura-Ann Petitto3. Teaching Language to Deaf Infants with a Robot and a Virtual Human, http://petitto.net/wp-content/uploads/2014/04/Petitto_CHI18.pdf

A.56 Logo and Trademark Detection

A.56.1 General

ID	56	
Use case name	Logo and Trademark Detection	
Application domain	Digital marketing Retail and Other (e.g. Fashion)	
Deployment model	Cloud services or on-premises systems	
Status	PoC	
Scope	Identification of logos / trademarks in pictures, optionally performing sentiment analysis associated to the product	
Objective(s)	Understand usage of retail or fashion products and optionally sentiment associated to it, according to pictures posted on the internet or social networks by customers	
Narrative	Short description (not more than 150 words)	The case is about being able to identify logos and trademarks in pictures provided to the AI systems, and optionally derive a positive or negative sentiment for the product based on the written context that was provided with the picture.
	Complete description	In order to provide business and marketing with a better understanding of how/in what context products are used, AI can be leveraged to help determine customer segments, anticipate changes in brand perception and customer preferences and help generate ideas for designers. The use case involves several steps: Confirm scope (including countries, targets, logos/trademarks) and business metrics Select and gather a suitable data set for training and testing the visual recognition algorithm. Optionally determine the rules that identify a proper context to be analysed with NLP techniques, to understand the sentiment associated to the logo/trademark contained in the picture when posted online.Pictures can be crawled from social networks, forums, and other websites, from which textual context (comments, etc) is obtained as well. Deploy to production and manage the lifecycle of AI, while providing business with the outcomes of the AI analysis.
Stakeholders		
Stakeholders' assets, values		
System's threats & vulnerabilities		

Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Number of logos/trademarks identified correctly	This is a technical precision/recall/accuracy measurement of how the visual recognition classifier is performing	Refers to the main objective
	2	Sentiment of Logo / trademark	This is a business measurement, that allows to understand the sentiment associated to a certain logo/trademark. The KPI is usually segmented by picture source, or other variables from the context	Refers to the main objective
AI features	Task(s)	Object detection and localization in pictures, Classification, Sentiment and Tone Analysis		
	Method(s)	Convolutional Neural Networks, Natural Language Processing		
	Hardware	None		
	Topology			
	Terms and concepts used	Visual Recognition, Sentiment Analysis, Tone Analysis		
Standardization opportunities/ requirements				
Challenges and issues	The primary challenge is to be able to correctly identify trademarks in all situations (with bad lighting, image distortions, dirt, etc.) and interpret the sentiment and tone in different countries and languages, as people might use slang and irony.			
Societal Concerns	Description	Automated analysis of public posts on social networks might be seen unethical in certain cultures.		
	SDGs to be achieved			

A.57 Virtual Bank Assistant

A.57.1 General

ID	57
Use case name	Virtual Bank Assistant

<table>
<tr><td>Application domain</td><td colspan="4">Banking</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Use of advanced chatbots and dialogue systems to automatize part of the call center activities</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Provide better quality help desk support to employees</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">The Virtual Assistant of the Bank is the first point of contact for branch operators, who receive immediate answers at any time - it allows to optimize the time of the "human operators" of the Service Desk, which they are dedicated to activities of greater value.</td></tr>
<tr><td>Complete description</td><td colspan="3">A bank in Italy has created a virtual consultant to support internal staff in their operations and interaction with customers.
The solution enabled a significant change in the service model of the bank, allowing to achieve important results in terms of greater contact volumes, extension of service hours and reduction of low-value human-centric activities.
The Virtual Assistant has been conceived as the first (and only) access point for assistance, it is easy to use and responds with a high level of reliability to the questions of branch colleagues. The virtual assistant has been not designed as a simple "chatbot" trained on a specific topic, but the virtual "colleague" to turn to for any question, completely integrated into the bank knowledge chain. To date, Virtual Bank Assistant manages all fourteen knowledge domains of the bank receiving thousands of answers.
From the beginning of its use (January 2018), the Virtual Assistant manages 100% of the requests, partly independently and partly in collaboration with the human operators of Service Desk.
The effectiveness of the solution is evidenced by the very high level of satisfaction, with positive feedback from users exceeding 90% and the reduction in the time spent by Service Desk operators in providing support to the branches, which today can be quantified in a reduction of 25 %.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Greater contact volumes with the bank</td><td>The objective is to expand the quantity of internal support activities provided by the</td><td>Improve productivity of service desk operators (already measured an</td></tr>
</table>

			bank its employees.	improvement of 25%)
	2	Extension of service hours	Expand the internal support activities 24/7	Always on
	3	Reduction of low-value human-centric activities	Reduction of the low level labor activities and let employees concentrate on more added value activities.	Improve the quality of work
AI features	Task(s)	Natural Language Dialogue systems		
	Method(s)	NLP		
	Hardware	Web based solution		
	Topology			
	Terms and concepts used	Natural Language Processing, Chat Bot, Dialogues Systems		
Standardization opportunities/ requirements				
Challenges and issues	Provide a natural and consistent interaction with users from different levels of experience (and thus terminology) and background			
Societal Concerns	Description			
	SDGs to be achieved			

A.58 Video on Demand Publishing Intelligence Platform

A.58.1 General

ID	58	
Use case name	Video on Demand Publishing Intelligence Platform	
Application domain	TMT Industry, Technology Department	
Deployment model	On-premises	
Status	Delivered Project	
Scope	Predictive maintenance platform on a Video on Demand Content Preparation Process	
Objective(s)	The goals of the project are: 1. Process fault comprehension 2. Fault prediction 3. Fault recovery through a recommendation engine 4. Productive interaction between the fault prediction and recovery recommendation engines for a proactive process maintenance	
Narrative	Short description (not more than 150 words)	An E2E platform was developed in order to achieve accurate fault prediction with Machine Learning and useful recovery action recommendation using Reinforcement Learning

<table>
<tr><td></td><td>Complete description</td><td colspan="3">The Fault Prediction engine allows to simulate the outcome of a process instance. The Machine Learning engine predicts the outcome using:
· The current state of the target applications
· The current state of the target IT systems
· The recent state of target applications (20 minutes)
· The recent state of the target IT systems (20 minutes)
The ML models give insights on the most important variables in predicting the outcome. These variables might point directly to the error cause, or be related to it.
The recovery recommendation engine is able to:
· Use the ML models to find a data-driven optimal action
· Incorporate user feedback to add custom actions
· Incorporate user feedback in order to further improve its recommendation strategy
Model and user defined actions challenge each other in order to provide the current best action. User feedback is incorporated in a reinforcement learning fashion.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="5">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Error frequency</td><td>Error frequency to be reduced</td><td>Productive interaction between the fault prediction and recovery recommendation engines for a proactive process maintenance</td></tr>
<tr><td>2</td><td>Lateness</td><td>Number of time consumed tasks to be reduced</td><td>Productive interaction between the fault prediction and recovery recommendation engines for a proactive process maintenance</td></tr>
<tr><td>3</td><td>Model AUC</td><td>KPI to monitor the classification quality of the models</td><td>Fault prediction</td></tr>
<tr><td>4</td><td>User feedback</td><td>User feedback is used to tune the recommendation engine</td><td>Fault recovery through a recommendation engine</td></tr>
<tr><td>AI features</td><td>Task(s)</td><td colspan="3">The fault prediction engine and the fault recovery recommendation engine work in synergy: the first yields a</td></tr>
</table>

		fault probability based on the current and recent state of applications and IT systems, providing the latter with a recommended recovery action. This action is challenged by other user-defined actions in the recommendation engine. The platform suggests the winning action to the user. The user can then give a feedback, allowing the recommendation engine to improve in a reinforcement learning fashion.
	Method(s)	Random Forest, Variable Importance evaluation, Reinforcement Learning
	Hardware	Virtual Machines
	Topology	
	Terms and concepts used	Machine Learning, Reinforcement Learning, Recommendation Engine, Environmental logs, Application log, Next Best Action, Process Mining
Standardization opportunities/ requirements		
Challenges and issues	The Machine Learning Engine processing time had to be very short	
Societal Concerns	Description	
	SDGs to be achieved	

A.59 Predictive Testing

A.59.1 General

ID	59	
Use case name	Predictive Testing	
Application domain	TMT Industry – Application development	
Deployment model	On-premises	
Status	PoC	
Scope	Automatic detection of inaccurate test outcomes in an application development process	
Objective(s)	The goal of the project is the improvement of the automation level in the application testing process. This is achieved by the automatic identification of inaccurate test outcomes, reducing the number of failure alerts	
Narrative	Short description (not more than 150 words)	The solution adopts machine learning to analyze event logs of test results in order to reduce the number of wrongly failed tests
	Complete description	The testing phase represents a critical point for many companies with a strong technological impact. The test execution is often not completely automated, thus requiring a significant effort in terms of people and time. The event log analysis of tests can prevent the presence of false positives (failed tests not related to failures in the target application), can help in the identification of the stage

		in which the error occurred and can help identifying the actual outcome of the test. The solution consists in adopting Machine Learning methodologies to analyze the available data (coming from different applications and sources involved in the tests), identify correlations and patterns in order to identify false positives, automate testing phases and recommend mitigation actions		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	False positive	Reduce false positives	
	2	Test efficiency	Shorten testing phase	
AI features	Task(s)	Data analysis, Anomaly Detection, Complex event correlation		
	Method(s)	Process Mining, Markov Chains, Machine Learning		
	Hardware			
	Topology			
	Terms and concepts used	Data integration, compress and denoise, probability distribution of events, complex patterns		
Standardization opportunities/ requirements				
Challenges and issues	Being able to manage and handle different types of data (including contextual information), integrating the solution in the processes and procedures of the company			
Societal Concerns	Description			
	SDGs to be achieved			

A.60 Predictive Data Quality

A.60.1 General

ID	60
Use case name	Predictive Data Quality
Application domain	Other (please specify) Data Management
Deployment model	On premise / cloud
Status	PoC
Scope	A solution for assessing Data Quality in data collection systems

<table>
<tr><td>Objective(s)</td><td colspan="4">Using machine learning techniques for identifying complex or unknown correlation among data in order to score its quality and enhance the confidence for data consumer in using data for the decision making processes</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">The solution adopt machine learning methods to analyze data collected in order to identify complex correlation on data (unknown at priori) and predict data quality issues.</td></tr>
<tr><td>Complete description</td><td colspan="3">The solution relies on four elements:
· Sources: the data sources represent the subject of the assessment. This sources can be heterogeneous (structured and semi-structured)
· Model: the representation of the ontology used as a reference for identifying the non-conformity on data
· Processes: the set of processes that produce and consume data, whose execution could be affected by the quality of data
· Organization and governance: the set of policies, procedures for governing data and handling the advanced data quality techniques.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Conformity Indicator</td><td>An indicator of the intrinsic data quality</td><td></td></tr>
<tr><td>2</td><td>Robustness Indicator</td><td>An indicator of the completeness of the set of data quality controls</td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Data analysis, Anomaly Detection, Complex event correlations</td></tr>
<tr><td>Method(s)</td><td colspan="3">Bayesian network, Support Vector Machine, CNN</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Data integration, data linkage, correlation analysis</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Being able to manage and handle different type of data, link data to reference knowledge model, change management in the organization</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3"></td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3"></td></tr>
</table>

A.61 Robot consciousness

A.61.1 General

ID	61	
Use case name	Robot consciousness	
Application domain	Other (please specify) Robotics	
Deployment model	Embedded systems	
Status	PoC	
Scope	A robot for museum tours equipped with the main capabilities of functional consciousness, accepted and transparent to untrained users.	
Objective(s)	The robot “CiceRobot” offering guided tours in indoor and outdoor museum and equipped with capabilities of functional consciousness, with no concern on the robot qualitative experience. The objective of case study is the acceptance and transparency of the autonomous behavior of the robot in an environment populated with untrained users as the museum visitors.	
Narrative	Short description (not more than 150 words)	The “CiceRobot” is a robot with capabilities associated with functional aspects of consciousness. CiceRobot offered indoors guided tours and outdoors guided tours. The outcome of the project is the acceptance and transparency of the autonomous behavior of the robot towards untrained visitors.
	Complete description	The “CiceRobot” is a robot with the capabilities associated with the functional aspects of consciousness. The architecture was instantiated on a wheeled robot for indoor use, on a wheeled robot for outdoor use and currently is instantiated on a humanoid robot. The robot has capabilities associated with the functional aspects of consciousness: · to build and to maintain an internal model of the environment and itself; · to pay attention to the relevant entities in the environment; · to integrate information from different sources and different parts of the same source; · to generate expectations about the possible events in the environment; · to self-monitor; · to simulate emotional states; · to process information by making it globally available to the robot. The primary outcome of the case study is the acceptancy and transparency of the autonomous behavior of the robot in an environment populated by untrained users as museum tourists.
Stakeholders		
Stakeholders’ assets, values		
System’s threats & vulnerabilities		

Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Robot transparency	The capability of the robot to act in a transparent way to tourists. The transparency of robot behavior is measured by questionnaires, M.O.S. on scale 1 – 5. The transparency of personal data handling and mitigation of cyberattack is pursued by local use of data (what happens to the robot remains on the robot and it is part of its personal history) and measured by questionnaires, M.O.S. on scale 1 – 5.	
	2	Robot acceptance	The capability of the robot to be accepted by tourists as a museum guide is measured by user satisfaction questionnaires, M.O.S. on scale 1-5.	
AI features	Task(s)	Internal model generation, attention, self-modelling, global workspace, expectation generation, information integration		
	Method(s)	Neural networks, symbolic representation systems, hybrid symbolic-subsymbolic systems, global representations.		
	Hardware	Wheeled indoor robot; wheeled outdoor robot; humanoid robot.		
	Topology			
	Terms and concepts used	Consciousness, attention, information integration, self-monitoring, expectation generation, internal modelling, global workspace.		
Standardization opportunities/ requirements				

Challenges and issues	The primary challenge of robot consciousness is the transparency and acceptance of robot operations, important in environments populated by untrained people as tourists in an archaeological museum.	
Societal Concerns	Description	The main concern may be the capability of the robot to act in a way which may is considered unethical to humans.
	SDGs to be achieved	

A.61.2 References

[1] Reggia, J.A. The rise of machine consciousness: Studying consciousness with computational models (2013) Neural Networks, 44, pp. 112-131. https://doi.org/10.1016/j.neunet.2013.03.011

[2] Chella, A., Gaglio, S. Synthetic phenomenology and high-dimensional buffer hypothesis (2012) International Journal of Machine Consciousness, 4 (2), pp. 353-365. https://doi.org/10.1142/S1793843012400203

[3] Chella, A., Manzotti, R. Machine consciousness: A manifesto for robotics (2009) International Journal of Machine Consciousness, 1 (1), pp. 33-51. https://doi.org/10.1142/S1793843009000062

[4] Chella, A., Macaluso, I. The perception loop in CiceRobot, a museum guide robot (2009) Neurocomputing, 72 (4-6), pp. 760-766. https://doi.org/10.1016/j.neucom.2008.07.011

[5] Chella, A., Frixione, M., Gaglio, S. A cognitive architecture for robot self-consciousness (2008) Artificial Intelligence in Medicine, 44 (2), pp. 147-154. https://doi.org/10.1016/j.artmed.2008.07.003

[6] Gamez, D. Progress in machine consciousness (2008) Consciousness and Cognition, 17 (3), pp. 887-910. https://doi.org/10.1016/j.concog.2007.04.005

[7] Chella, A., Barone, R.E. Panormo: An emo-dramatic tour guide (2008) AAAI Spring Symposium - Technical Report, SS-08-04, pp. 10-16. https://aaai.org/Papers/Symposia/Spring/2008/SS-08-04/SS08-04-003.pdf

[8] Chella, A., Barone, R.E., Pilato, G., Sorbello, R. An emotional storyteller robot (2008) AAAI Spring Symposium - Technical Report, SS-08-04, pp. 17-22. https://aaai.org/Papers/Symposia/Spring/2008/SS-08-04/SS08-04-004.pdf

[9] Barone, R., Macaluso, I., Riano, L., Chella, A. A brain inspired architecture for an outdoor robot guide (2008) AAAI Fall Symposium - Technical Report, FS-08-04, pp. 27-34. https://aaai.org/Papers/Symposia/Fall/2008/FS-08-04/FS08-04-005.pdf

[10] Macaluso, I., Chella, A. Machine consciousness in CiceRobot, a museum guide robot (2007) AAAI Fall Symposium - Technical Report, FS-07-01, pp. 90-95. https://www.aaai.org/Papers/Symposia/Fall/2007/FS-07-01/FS07-01-016.pdf

[11] Chella, A., Liotta, M., Macaluso, I. CiceRobot: A cognitive robot for interactive museum tours (2007) Industrial Robot, 34 (6), pp. 503-511. https://doi.org/10.1108/01439910710832101

[12] Macaluso, I., Ardizzone, E., Chella, A., Cossentino, M., Gentile, A., Gradino, R., Infantino, I., Liotta, M., Rizzo, R., Scardino, G. Experiences with cicerobot, a museum guide cognitive robot (2005) Lecture Notes in Artificial Intelligence, 3673 LNAI, pp. 474-482. https://link.springer.com/chapter/10.1007%2F11558590_48

A.62 AI Sign Language Interpretation System for the Hearing-Impaired

A.62.1 General

<table>
<tr><td>ID</td><td colspan="4">62</td></tr>
<tr><td>Use case name</td><td colspan="4">AI Sign Language Interpretation System for the Hearing-Impaired</td></tr>
<tr><td>Application domain</td><td colspan="4">Public sector</td></tr>
<tr><td>Deployment model</td><td colspan="4">Embedded systems</td></tr>
<tr><td>Status</td><td colspan="4">Prototype</td></tr>
<tr><td>Scope</td><td colspan="4">Increase the convenience of public services to hearing-impaired people by providing a service to translate sign language image information into natural language</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Supporting communication between hearing-impaired and non-disabled people</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">In this use case scenario, hearing impaired and non-disabled people are able to communicate each other through the AI sign language-natural language interpretation service.</td></tr>
<tr><td>Complete description</td><td colspan="3">This service supports seamless conversation with a non-disabled person by converting the sign language image sequences of a hearing-impaired person into voice or natural language text.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Government or public institutions</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Welfare fund or budget for the disabled people</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">It is difficult to understand the dialectical expressions and other domain vocabularies that are not used as training data in sign languages and natural languages.</td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Sentence unit translation accuracy</td><td>A performance measurement that calculates the ratio of sentences translated into the correct natural language among the sentences of evaluation data (in %)</td><td>Accurate communication between disabled and non-disabled people</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition, Generation</td></tr>
<tr><td>Method(s)</td><td colspan="3">Computer vision, translation modelling, speech synthesis, video synthesis</td></tr>
<tr><td>Hardware</td><td colspan="3">Camera, speaker, monitor, microphone</td></tr>
<tr><td>Topology</td><td colspan="3">Deep learning-based sequence to sequence model</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Sign language recognition, automatic translation, sign language generation</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Multi-modal data input/output format, the interface definition of structures, and interface specifications between modules</td></tr>
</table>

Challenges and issues	Multimodal interactions Translation from visual information to textual information Translation from textual information to visual information	
Societal Concerns	Description	Promoting welfare and supporting social activities for the disabled
	SDGs to be achieved	Good health and well-being for people

A.62.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Academic paper	Neural Sign Language Translation based on Human Keypoint Estimation	Accepted	Developing prototype system	Sang-Ki Ko, Chang Jo Kim, Hyedong Jung, Choongsang Cho / KETI	https://www.mdpi.com/journal/applsci (Journal of Applied Sciences)

A.63 Dialogue-based social care services for people with mental illness, dementia and the elderly living alone

A.63.1 General

ID	63	
Use case name	Dialogue-based social care services for people with mental illness, dementia and the elderly living alone	
Application domain	Medical sector	
Deployment model	Client and server systems	
Status	Prototype	
Scope	Daily life support AI services that provide an interaction with humans using natural language	
Objective(s)	Dialogue-based interaction between people and machines utilizing artificial intelligence technology helps people with accessibility issues to IT devices	
Narrative	Short description (not more than 150 words)	Daily life support services based on artificial intelligence conversation technology that can perform information processing tasks through natural language conversation with users
	Complete description	This use case is related to the spread of digital and unmanned services. A variety of reasons, including unfriendly user interfaces, mental or physical limitations, make some people uncomfortable with the latest IT device-based services. This causes inequality in the benefits of the latest technology. Artificial intelligence conversation technology, which can interact with users through natural

		language, can help reduce this inequality. This technology supports the interaction of people and technologies in a digital environment.		
Stakeholders	Government or public institutions			
Stakeholders' assets, values	Welfare fund or budget for the elderly people or the mental illness people			
System's threats & vulnerabilities	Since the service is closely related to an individual's daily life, if the system is hacked during the service process, the hacked information can be exploited for personal information leaks or various personal crimes.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Task completion rate	The performance is calculated by dividing the number of cases that have been completed successfully by the total number of assigned tasks. The success or failure of a task is set according to the criteria of each system.	Accurate task completion using the AI system
	2	Mean Opinion Score	A measure used in the domain of Quality of Experience and telecommunications engineering, representing overall quality of a stimulus or system	Providing a human-friendly system interface through natural language-based conversations
AI features	Task(s)	Daily chitchat, Question and Answering, condition checking, e-commerce		
	Method(s)	Dialogue system, Generation, agent, knowledge bases, information retrieval, speech recognition, speech synthesis		
	Hardware	Camera, speaker, monitor, microphone		
	Topology	Deep learning-based sequence to sequence model, Tacotron, Wavenet, and so on.		
	Terms and concepts used	Dialogue system, generation, agent, knowledge bases, information retrieval, speech recognition, speech synthesis, sequence to sequence model, Tacotron, Wavenet		
Standardization opportunities/ requirements	Multimodal information input/output formats and the technical process guideline Knowledge base format Knowledge base query format			
Challenges and issues	Multimodal data handling based multimodal interaction Multimodal data analysis			

	Multimodal data-based inferences	
Societal Concerns	Description	Promoting welfare and supporting social activities for the inconvenient
	SDGs to be achieved	Good health and well-being for people

A.63.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Blog article	Improve patient engagement and efficiency with AI powered chatbots	published	Justification of service	David Houlding / Microsoft Azure	https://azure.microsoft.com/ko-kr/blog/improve-patient-engagement-and-efficiency-with-ai-powered-chatbots/

A.64 AI Situation Explanation Service for the Visually Impaired

A.64.1 General

ID	64	
Use case name	AI Situation Explanation Service for the Visually Impaired	
Application domain	Public sector	
Deployment model	Client and server systems	
Status	Prototype	
Scope	A real-time situation explanation service through voice for the visually impaired	
Objective(s)	Recognizing Texts around the visually impaired Recognizing Faces around the visually impaired Recognizing Objects around the visually impaired Assisting the mobility of the visually impaired Describe scenes and photos for the visually impaired	
Narrative	Short description (not more than 150 words)	A daily life support service, based on artificial intelligence technologies, that can explain the situation around visually impaired people while moving
	Complete description	The use case supports the daily life of visually impaired people through AI vision technologies. This service helps to recognize or avoid dangerous objects on the move, identify people, text, and objects, and acquaintances by taking into account various surrounding situations. This also supports captioning service to understand the current situation or photos.
Stakeholders	Personal services	
Stakeholders' assets, values	Welfare fund or budget for the impaired people	

System's threats & vulnerabilities	Since the services is closely related to the individual's life, if the system is hacked in the service process, the hacked information can be exploited for personal information leakage and various personal crimes.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	CIDEr	Consensus-based Image Description Evaluation. The evaluation metric measures the similarity of a generated sentence against a set of ground truth sentences written by humans. CIDEr metric shows high agreement with consensus as assessed by humans.	Accurate task completion using the AI system
AI features	Task(s)	Generation of the most proper natural language sentence from an image input		
	Method(s)	Image captioning Object detection Face detection		
	Hardware	Camera, speaker, monitor, microphone		
	Topology	Variational Auto Encoder (VAE), Generative Adversarial Nets (GAN)		
	Terms and concepts used	Image captioning, Object detection, Face detection, Natural language generation		
Standardization opportunities/ requirements	Image data input/text output interface structures and specifications Minimum image quality and communication environment guidelines for reliable performance Guidelines for building training data for commercial services and the minimum size of learning data construction and structure			
Challenges and issues	Vision			
Societal Concerns	Description	Promoting welfare and supporting social activities for the blind		
	SDGs to be achieved	Good health and well-being for people		

A.64.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

1	News articles	Horus Technology Launches Early Access Program for AI-Powered Wearable for the Blind; Rebrands Company as Eyra	published	Related application service	Saverio Murgia / Eyra	https://www.prnewswire.com/news-releases/horus-technology-launches-early-access-program-for-ai-powered-wearable-for-the-blind-rebrands-company-as-eyra-300351430.html

A.65 Social humanoid technology capable of multi-modal context recognition and expression

A.65.1 General

ID	65	
Use case name	Social humanoid technology capable of multi-modal context recognition and expression	
Application domain	Service robot, HCI	
Deployment model	Embedded systems	
Status	Prototype	
Scope	Human-AI sympathetic technology expressing dynamic immersive dialogue with humans through a combination of various artificial intelligence technologies	
Objective(s)	Sympathetic dialogue technology in order to understand socio-cultural consensus and emotions Creation of para-verbal expressions to induce sympathy with a speaker Representing non-verbal expressions reflecting the emphasis and intention of each utterance Deep dialogue management and combination of multimodal expressions for in-depth sympathy while conversations	
Narrative	Short description (not more than 150 words)	A highly immersive sympathetic conversation technology based on artificial intelligence that includes integrated understanding and expression skills of verbal, nonverbal, and para-verbal information to derive complete communion with humans
	Complete description	Immersive sympathetic dialogue technique is a technology that allows AI's interactions to share ideas and emotions with people through in-depth understanding of complex information beyond simple information exchange. Sympathetic dialogue technology means cognition, understanding, reasoning, management, and generation techniques for mutual context and information sharing and creation using dialogue with a human. These sympathetic dialogue techniques include the understanding and representation of verbal, para-verbal, and non-verbal information to understand in-depth intents for more human-like communications. The verbal interaction means the interaction of language understanding and

		representations. The para-verbal interaction is vocal size, height, tremor, speed, clarity, turn taking, etc. Non-verbal interactions can be defined by the expression, gaze, or action of an emotion through understanding of the surrounding situation.		
Stakeholders	Dialogue service services using display devices Service robot			
Stakeholders' assets, values	R&D fund from a governmental research project or major companies			
System's threats & vulnerabilities	Definition of system ethics of communication and decision-making process			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Mean Opinion Score	A measure used in the domain of Quality of Experience and telecommunications engineering, representing overall quality of a stimulus or system	Providing a human-friendly system interface through natural language-based conversations
AI features	Task(s)	Emotionally sympathetic dialogue understanding and generation Non-verbal / para-verbal dialogue understanding and representations Dynamically simultaneous synthesis of verbal / non-verbal / para-verbal conversations		
	Method(s)	Image captioning Object detection Face detection		
	Hardware	Camera, speaker, monitor, microphone		
	Topology	Generative Adversarial Nets (GAN), Deep learning-based sequence to sequence model		
	Terms and concepts used	Natural Language Understanding, Natural language generation, Machine Reading and Comprehension, Spoken Dialogue System		
Standardization opportunities/ requirements	Multimodal information input/output formats and typical process guideline Transformation from Multimodal input to Knowledge base query format Knowledgebase interfacing format Reference functional module structure and their typical interface structure and formats Information synchronization and sharing issues in Knowledgebase			
Challenges and issues	Multimodal data understanding / inference / representation			
Societal Concerns	Description	The increase in the elderly population and the decrease in the total population are increasing the inequality of social welfare and benefits according to generation, class and region.		

	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.66 Expansion of AI training dataset and contents using artificial intelligence techniques

A.66.1 General

ID	66			
Use case name	Expansion of AI training dataset and contents using artificial intelligence techniques			
Application domain	IT, AI, Future services			
Deployment model	Server system			
Status	Research			
Scope	Data self-propagation and validation service for deep learning and contents services			
Objective(s)	Self-propagation of data to enhance the performance of application systems and to support the expansion of data for deep learning Self-propagated data evaluation for qualitative verification			
Narrative	Short description (not more than 150 words)	The service expands the data used for deep learning for rapid commercialization of artificial intelligence technologies. The service includes quantitative extensions of the amount of learning data for high-quality in-depth learning and qualitative verification of extended data applied to machine learning or commercial content services.		
	Complete description	As the artificial intelligence technology develops, the services to which the technologies are applied are increasing. The development of artificial intelligence using machine learning or deep learning requires vast amounts of data for learning. However, because such a sufficient amount of data is rare, technological polarization in the artificial intelligence area is getting serious. In order to alleviate these problems and to support artificial intelligence research and various commercialization, training data should be available at relatively low cost. The service utilizes artificial intelligence technology to multiply training data of artificial intelligence systems itself and perform qualitative verification of these automatically generated data.		
Stakeholders	AI research AI technology-based services Contents providers			
Stakeholders' assets, values	R&D fund from a governmental research project or companies			
System's threats & vulnerabilities	In some cases, the distribution of data in the real world may not be reflected in the automatic data expansion process.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives

	1	Performance improvement rate	Performance enhancement rate due to the additional utilization of propagated training data	Improving the performance of machine learning engines
AI features	Task(s)	Self-propagation of training data Self-propagation of service data or contents Validation of self-generated data for qualitative verification		
	Method(s)	Machine learning Algorithms for a generative model		
	Hardware			
	Topology	Generative Adversarial Nets (GAN)		
	Terms and concepts used	Machine learning, Generative models		
Standardization opportunities/ requirements	Generated data quality guidelines for use as learning data or services Qualified evaluation guideline for generated data validation in various data types			
Challenges and issues	The optimized self-propagation techniques for various types of data			
Societal Concerns	Description	The technology polarization in artificial intelligence technical area becomes serious more and more.		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.67 Pre-screening of cavity and oral diseases based on 2D digital images

A.67.1 General

ID	67	
Use case name	Pre-screening of cavity and oral diseases based on 2D digital images	
Application domain	Medical services	
Deployment model	Client and server systems	
Status	Prototype	
Scope	Artificial intelligence-based oral examination platform	
Objective(s)	AI based oral disease self-examination solution Cavity, periodontal disease, oral disease, tooth care and oral care self-care prevention management	
Narrative	Short description (not more than 150 words)	This service utilizes artificial intelligence technology to analyze the oral condition by sending oral images to the diagnostic server without visiting the dentist.
	Complete description	The oral condition self-diagnosis service is easy to use. Artificial intelligence technology analyses oral health status such as periodontal disease, gingivitis, periodontitis, and cavities and provides oral status reports. This service provides sufficient guidelines for preliminary diagnosis through artificial intelligence techniques before dental

		visits. This system comprehensively manages the oral health of individuals and families.		
Stakeholders	Dentist Public			
Stakeholders' assets, values	R&D fund from a governmental research project or dentists			
System's threats & vulnerabilities	Personal information utilization issue			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Diagnostic accuracy	A performance measure that calculates the percentage of correct into the correct diagnosis among evaluation data in %	Accurate diagnosis before visiting dentists
AI features	Task(s)	Vision Oral image analysis Lesion segmentation		
	Method(s)	Machine learning Algorithms for the classification model		
	Hardware	Smartphone (including camera)		
	Topology	CNN, ResNet		
	Terms and concepts used	Machine learning, Medical AI, Data eco system		
Standardization opportunities/ requirements	Guidelines for capturing oral image and the minimum quality of the images for diagnosis Guidelines for a provision of the diagnostic results			
Challenges and issues	Dental image processing using artificial intelligence			
Societal Concerns	Description	Elimination of inequalities in regional health care services		
	SDGs to be achieved	Good health and well-being for people		

A.67.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Google play	Apo-AI	Beta service released	Related application service	Pusan national university dental hospital, Korea	https://play.google.com/store/apps/details?id=com.qtt.ea4&hl=en_US 5

A.68 Real-time patient support and medical information service applying spoken dialogue system

A.68.1 General

<table>
<tr><td>ID</td><td colspan="4">68</td></tr>
<tr><td>Use case name</td><td colspan="4">Real-time patient support and medical information service applying spoken dialogue system</td></tr>
<tr><td>Application domain</td><td colspan="4">Medical services</td></tr>
<tr><td>Deployment model</td><td colspan="4">Client and server systems</td></tr>
<tr><td>Status</td><td colspan="4">Prototype</td></tr>
<tr><td>Scope</td><td colspan="4">Medical business support system using artificial intelligence based human computer interface technology</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Acquisition, retrieval and provision of patients and related data needed by medical staffs in real time through a voice dialogue interface during medical treatment</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">The service is a medical system that provides patient information and related data for treatment in real time based on a voice dialogue interface to help medical hands-on medical activities, such as dental, first aid, and surgery.</td></tr>
<tr><td>Complete description</td><td colspan="3">Dental care and medical procedures that directly treat patients require a variety of identification and integration of patient data and related health information. It is difficult for medical practitioners to search, analyse and organize data during direct treatment. The voice dialogue interface based medical information provision and management system combined with various artificial intelligence technologies is beneficial to both the medical staffs and the patient by increasing convenience and efficiency of medical treatment.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Dentist
Hospital</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">R&D fund from a governmental research project or dentists</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Utilizing personal medical information in artificial intelligence research and systems</td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Mean Opinion Score</td><td>A measure used in the domain of Quality of Experience and telecommunications engineering, representing overall quality of a stimulus or system</td><td>Providing a human-friendly system interface through speech-based conversations</td></tr>
<tr><td>AI features</td><td>Task(s)</td><td colspan="3">Speech recognition
Natural language processing
Knowledge based question and answering
Speech synthesis</td></tr>
</table>

	Method(s)	Speech dialogue system Question and answering Information retrieval Human computer interface
	Hardware	Camera, speaker, monitor, microphone
	Topology	Deep learning-based sequence to sequence model, Tacotron, Wavenet, and so on.
	Terms and concepts used	Machine learning, Medical AI, Data eco system
Standardization opportunities/ requirements	Guidelines for collecting patient data for dental care	
Challenges and issues	Dialogue service in medical data and knowledge Question and answering in a medical expert system Multi-task handling in a dialogue-based interfacing environment Remote speech recognition	
Societal Concerns	Description	Improving medical service efficiency and patient satisfaction
	SDGs to be achieved	Good health and well-being for people

A.68.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Company homepage	DEXvoice - The SMART Solution for Your Dental Workflow	Service released	Related application service	Kavo	https://www.kavo.com/en-us/dexvoice-smart-solution-your-dental-workflow

A.69 Integrated recommendation solution for prosthodontic treatments

A.69.1 General

ID	69
Use case name	Integrated recommendation solution for prosthodontic treatments
Application domain	Medical services
Deployment model	Client and server systems
Status	Prototype
Scope	In order to support complicated prosthetic treatments according to the patient's condition, the artificial intelligence technology provides a comprehensive analysis of the given information and situations to recommend various prosthetic treatment methods and visualize them to support doctors and patients.

<table>
<tr><td>Objective(s)</td><td colspan="4">Various knowledge in dentistry and related patient data for prosthodontic treatment are collected in advance
Suggesting recommended cases and possible solutions for the prosthesis</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">This service includes sufficient dental knowledge and patient data for prosthodontic treatment, and uses a variety of artificial intelligence techniques to provide recommended practices and possible solutions for prosthodontics.</td></tr>
<tr><td>Complete description</td><td colspan="3">The prosthodontic treatment depends on the experience and ability of the medical staff, and the patient satisfaction varies accordingly. This technology has sufficient knowledge of dental and patient data for prosthetic treatment in advance to improve health care efficiency and patient satisfaction. During the diagnosis process, this service proposes recommended practices and possible solutions by applying a variety of artificial intelligence techniques to help medical staffs for accurate diagnoses.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Dentist
Hospital</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">R&D fund from a governmental research project or dentists</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Utilizing personal medical information in artificial intelligence research and systems
Diagnosis of AI
Depends on the performance of the AI system in the diagnostic process</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Recommendation accuracy</td><td>A performance measure that calculates the percentage of correct into the correct solution that medical staff selected among evaluation data
in %</td><td>Accurate discovery of a solution before diagnosis for medical experts</td></tr>
<tr><td>2</td><td>Mean Opinion Score</td><td>A measure used in the domain of Quality of Experience and telecommunications engineering, representing overall quality of a stimulus or system</td><td>Suggesting optimized solutions using this service</td></tr>
<tr><td rowspan="2">AI features</td><td>Task(s)</td><td colspan="3">Natural language processing
Knowledge based question and answering
Data mining
Searching similar cases
Recommendation of optimal solutions</td></tr>
<tr><td>Method(s)</td><td colspan="3">Information retrieval</td></tr>
</table>

		Recommendation
	Hardware	
	Topology	K-means, Graph clustering, Ranking, Dynamic time Warping, Genetic algorithms
	Terms and concepts used	Recommendation engine, Discovery engine, Medical AI, Data eco system
Standardization opportunities/ requirements	Guidelines for collecting patient data for dental care Medical knowledgebase representation format Medical knowledgebase search format	
Challenges and issues	Discovery satisfied solutions based on medical knowledge and clinical data Reasoning novel cases by combining expert knowledge and case studies	
Societal Concerns	Description	Improving medical service efficiency and patient satisfaction
	SDGs to be achieved	Good health and well-being for people

A.70 A judging support system for gymnastics using 3D sensing

A.70.1 General

ID	70	
Use case name	A judging support system for gymnastics using 3D sensing	
Application domain	ICT	
Deployment model	On-premise systems	
Status	PoC	
Scope	Skeleton recognition for gymnastics	
Objective(s)	To support judgement of difficult element by high-level and high-speed.	
Narrative	Short description (not more than 150 words)	We have been developing a judging support system for artistic gymnastics to enhance accuracy and fairness in judging. We developed a skeleton recognition technique using the learned model that we trained using a large amount of depth images of gymnastics created from CG in advance. With this technology, it is possible to recognize a human 3D skeleton from depth image.
	Complete description	In gymnastics, wrong scoring is a problem, when it is difficult to judge by high-level and high-speed. Therefore, 3D sensing technology is required to reduce burden of referee by recognizing skeleton of gymnast. We developed a technique to recognize heatmaps of body parts using the learned model that we trained using a large amount of depth images of gymnastics created from CG in advance. We calculate 3D skeleton position using heatmaps of body parts. With this technology, it is possible to recognize a human 3D skeleton from depth image.
Stakeholders	Federation International Gymnastics(FIG)	
Stakeholders' assets, values		

System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware			
	Topology	CNN		
	Terms and concepts used	Deep learning, Convolution neural network, training, training data set		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Recognize skeleton of all gymnastics element. Issues: Recognize 3D skeleton in gymnastics that are complex movements from depth image.			
Societal Concerns	Description	Positive: Fairness of scoring, reducing burden of referee, and technical improvement of gymnast. Negative:		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.70.2 Data

Data characteristics	
Description	Depth images, 2D data of skeleton
Source	Motion capture
Type	Images
Volume (size)	
Velocity	Non-real time
Variety	Single dataset
Variability (rate of change)	Static
Quality	High

A.70.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model with training data set.	Training data set is ready		Evaluation
2	Evaluation	Evaluate whether	Completion of	Training/Retraining	Execution

		the trained model can be deployed cg data	training/re training		
3	Execution	Recognize real data gained 3D laser sensor	Get real data by 3D laser sensor	Evaluation	Retaining
4	Retraining	Retrain a model with added training data set.	Recognitio n accuracy of real data is low	Execution	

A.70.4 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/o rganization	Link
1	Pressr elease				Fujitsu	http://pr.fujitsu.com/jp/ne ws/2018/11/20.html

A.71 Active Antenna Array Satellite

A.71.1 General

ID	71	
Use case name	Active Antenna Array Satellite	
Application domain	ICT	
Deployment model	Cyber-physical systems	
Status	Prototype	
Scope	Determine optimal spot beam patterns for communication satellites in order to react to changing geographic distribution and bandwidth requirements of terminals	
Objective(s)	Optimise service quality and bandwidth allocation for users of satellite system	
Narrative	Short description (not more than 150 words)	Future high throughput satellites (HTS) will be equipped with an active antenna array instead of a fixed multiple spot beam pattern. This allows generating multiple spot beams with different number, size and shape. Moreover, the parameters, i.e. number, size and shape, can be adapted in a flexible way.
	Complete description	The problem tackled in this use-case is to find the optimum setup of the spot beams with respect to non-uniform distributed users on the service area. For training purposes, the ML algorithm would be fed with different, e.g. randomly generated, terminal distributions, and a set of spot beam parameters. The performance of the

		solution is assessed by analyses of the possible network-wide throughput. This takes into account: · The contour losses of the terminals at their position. · The interference from spot beams transmitting on the same frequency band (4-coloring scheme assumed). · The HTS multi-spot beam antenna pattern.		
Stakeholders	Operators of satellite communication systems Users of satellite communication systems Regulation authorities Space agencies			
Stakeholders' assets, values	Reliability of the service, coverage of the service, bandwidth optimisation			
System's threats & vulnerabilities	Potential for attack via terminal data to disturb system performance			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	System throughput	Overall throughput for a particular terminal configuration	
	2	Update time	Time required to determine a new antenna configuration	
AI features	Task(s)	Optimization		
	Method(s)	semi-supervised clustering [2,3], generative networks, self organizing maps [1]		
	Hardware	Server at ground control station		
	Topology	GANs		
	Terms and concepts used	machine learning, semi-supervised learning		
Standardization opportunities/ requirements	Robustness requirements and metrics			
Challenges and issues				
Societal Concerns	Description	Potential to provide demand-adapted service coverage in sparsely populated areas that might not be well served in a fixed configuration scenario		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.71.2 Data

Data characteristics	
Description	Terminal positions and bandwidth requirements
Source	Simulations
Type	Time series of position updates
Volume (size)	~10^4 terminals

Velocity	updates of configurations within seconds
Variety	-
Variability (rate of change)	terminals will appear/disappear all the time
Quality	position updates may be incomplete

3875

3876 A.71.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Li, Jiaxin, Ben M. Chen, and Gim Hee Lee. "So-net: Self-organizing network for point cloud analysis." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.	Published			
2	Paper	Eick, Christoph F., Nidal Zeidat, and Zhenghong Zhao. "Supervised clustering-algorithms and benefits." 16th IEEE International Conference on Tools with Artificial Intelligence. IEEE, 2004.	Published			
3	Paper	Basu, Sugato, Mikhail Bilenko, and Raymond J. Mooney. "A probabilistic framework for semi-supervised clustering." Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004.	Published			

3877

3878 A.72 Carrier interference detection and removal for satellite communication

3879 A.72.1 General

ID	72
Use case name	Carrier interference detection and removal for satellite communication
Application domain	ICT

Deployment model	On-premise systems	
Status	PoC	
Scope	Machine-learning-based detection, classification and removal of interference signal for satellite communication systems	
Objective(s)	Detection (and possibly classification) of interfering signals in satellite communication systems (e.g., DVB-S2 or DVB-S2x), and removal of the interfering signal using the gained knowledge about the interfere characteristics, with the aim of reducing the error rate at the receiver.	
Narrative	Short description (not more than 150 words)	In satellite communication systems, unintended or intended interferences are quite common. For instance, interferences might originate from a mis-pointed terminal antenna, a radar signal or from another terrestrial radio source. In this use-case, the intention is to detect the presence of an interferer in addition to a desired carrier and potentially classify it. The setting for this use-case is as follows: · The terminal receives a desired carrier. · The details of the desired carrier are known, e.g. a DVB-S2x carrier with known symbol rate and modulation scheme. · There might be an interferer present with unknown frequency, bandwidth and structure. · The objective is to detect the presence of such an interferer and to classify the interferer, e.g. in terms of power, bandwidth and type. · Additionally, it may be desired to remove the influence of the interferer from the signal.
	Complete description	The ML-algorithm operates on the received samples of the signal consisting of the desired carrier and the interferer. The ML-algorithm searches for repetitive patterns in the signal, which are not expected from the known carrier signal. For instance, the interfering signal could be another DVB-S2 or DVB-S2x carrier from an adjacent satellite, a radar signal, or a terrestrial radio relay systems. Each of these interfering signals contains a repetitive pattern for instance in form of pilot symbols or unique words. Regarding the type of the ML method, both supervised and unsupervised learning could be feasible. However, the supervised learning scenario requires to train on a number of previously known interferers. This would limit the detection to a class of selected interfering signals. The use case can be broken down into different sub-problems: · A: Interference detection: This problem can be treated as anomaly detection, and learning a model for the undistorted signal from clean data. · B: Interference classification: Given sufficient training data for different types of inference signals, the problem can be treated as a classification problem into undistorted signals and signals overlapping with a particular type of distortion. This approach provides the type of distortion as a result, but may produce unreliable results under presence

		of distortions not trained for. A case that may need to be handled specifically is that over interfering signals of the same type, e.g., a DVB signal overlapping with another DVB signal, as the statistics of the two signals will be similar, but just the time offset of synchronisation symbols will enable the identification of the signals. · C: Signal separation: If interference has been identified, signal separation could be desired for further processing. Parts of the carrier are known (pilot sequence) or it is possible to transmit known data signal over the carrier, such that the desired carrier can be reconstructed at the receiver. The ML-algorithm is trained by comparison of the received (and interfered) signal with the (known) transmitted signal from the carrier, and determines a model how the interfering samples add to the carrier. Then the interference is reduced symbol by symbol from the carrier based on the trained states of the ML-algorithm.		
Stakeholders	Operators of satellite communication systems Operators of other communication systems (satellite or non-satellite) that are potential sources of interference Users of satellite communication systems Regulation authorities Space agencies			
Stakeholders' assets, values	Reliability of the service, costs to provide a certain service level			
System's threats & vulnerabilities	Potential for malicious attacks on classification of interference signal type			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Detection ratio	Ratio of correct detection of presence of interferer	
	2	Classification accuracy	Accuracy of correct classification of type and properties of interferer	
	3	SNR improvement	Improvement of signal to noise ratio by removing interferer	
AI features	Task(s)	Optimization		
	Method(s)	anomaly detection, time series classification, source separation		
	Hardware	FPGA		
	Topology	autoencoders, RNNs		
	Terms and concepts used	machine learning, supervised learning		
Standardization	Performance measurements and robustness requirements			

opportunities/ requirements		
Challenges and issues	performance and robustness needs probably be defined w.r.t. a certain class of signals (e.g. DVB-S but not generally)	
Societal Concerns	Description	
	SDGs to be achieved	

A.72.2 Data

Data characteristics	
Description	Carrier and interferer data
Source	Simulations
Type	Time series from different types of signals
Volume (size)	~100 Mbits/s
Velocity	training can be done offline, inference must be done in real-time
Variety	broad range of possible interferer signals
Variability (rate of change)	low change
Quality	interferer signals may be weak

A.72.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Tarem Ahmed, Boris Oreshkin and Mark Coates, Machine learning approaches to network anomaly detection, Proceedings of the 2nd USENIX workshop on Tackling computer systems problems with machine learning techniques, page 1-6, 2017.	Published			
2	Article	Kiran, B., Dilip Thomas, and Ranjith Parakkal. "An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos." Journal of Imaging 4.2 (2018): 36.	Published			
3	Paper	Weninger, Felix, et al. "Discriminatively	Published			

		trained recurrent neural networks for single-channel speech separation." 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2014.				
4	Paper	Hershey, John R., et al. "Deep clustering: Discriminative embeddings for segmentation and separation." 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016.	Published			

A.73 Jet Engine Predictive Maintenance Service

A.73.1 General

ID	73	
Use case name	Jet Engine Predictive Maintenance Service	
Application domain	Civilian Aviation Maintenance	
Deployment model	Cloud services	
Status	Prototype	
Scope	Use of jet engine telemetry data to train predictive maintenance algorithms	
Objective(s)		
Narrative	Short description (not more than 150 words)	ML-based jet-engine predictive maintenance technology predicts the next maintenance tasks proactively using machine learning model trained by jet engine telemetry data and maintenance history
	Complete description	By collecting large quantities of telemetry data from jet engines installed on commercial airliners as well as their maintenance history, machine learning algorithms can be trained to predict how those engines could fail in the future. Having made such predictions, maintenance can be performed proactively on the airliner engines before the problems actually occur, improving safety and lower cost by having more reliable and predictable equipment, making airline flights less prone to disruption. To allow collection of large quantities of jet engine telemetry and maintenance logs (Big Data) for use in ML model training, both airlines operating the planes as well as jet engine manufacturers are required to participate. But jet

<table>
<tr><td></td><td></td><td colspan="3">engine telemetry data or maintenance logs could contain proprietary and confidential corporate data under exclusive control of the jet engine manufacturers.

Therefore, the use of the proprietary data in model training by the company that develops the maintenance service needs to be explained and be transparent so the airlines and engine manufacturers would know how they data is used, and to ensure that their proprietary data is not shared with their competition.

The process of training models and how data is used needs to be explainable and transparent, and use of de-identifications techniques applied to parts of data that contain proprietary information need to be described to ensure trustworthiness. Such level of transparency and explainability can then be used in contracts necessary to enable data sharing across the industry. Without such transparency and explainability of data use in ML model training, data sharing will not proliferate and adoption of ML technologies will be hindered.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Airline industry, Jet Engine industry, Airline maintenance industry, cloud-based AI providers, airline insurance industry</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Leak of corporate technical intellectual property data and trade secrets to competition</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Confidentiality</td><td>Jet engine manufacturers and airline companies are confident enough about how their data is used to model training, and are satisfied that their trade secrets are not leaked to competition by the data sharing needed to allow models to be trained</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">AI features</td><td>Task(s)</td><td colspan="3">- Recognition of patterns and making predictions
- Explainability and transparency about how data was used in the model training phase</td></tr>
<tr><td>Method(s)</td><td colspan="3">Deep NN</td></tr>
</table>

	Hardware	1) High performing CPU nodes or GPUs in Cloud Computing Data Centers to train the DNN model 2) Cloud based VMs to run the trained DNN model
	Topology	
	Terms and concepts used	Deep Neural Networks – customized for infant facial recognition
Standardization opportunities/ requirements	- Need for transparency about the properties and sources of large quantities of jet engine telemetry data and engine maintenance history used to train DNN model – generalized to be applicable to all organizational IP and trade secrets containing data from data principals who are not human, rather may be IoT devices, for example. - Need for transparency of aspects of training data such as portions of data principal's data that need to be de-identified in order for corporate IP and trade secrets to be protected when shared with competition or partners. - Need for transparency and explainability of model training processes and the stages involved; and how data is used in each stage, and what de-identification techniques can be used to ensure corporate trade secrets are protected when data is shared with the outside. Such fundamental transparency and explainability can be used in contracts and agreements for data sharing	
Challenges and issues	- Explainability and transparency regarding the training data used, from the perspective of corporate confidentiality concerns, - Need a structured, common and standardized way to describe the stages of the machine learning model training process, and the types and aspects of the data used in the various stages of the process so the stakeholders (policy makers, partners and customers) can build confidence and trust in such ML-based product or service, ensuring that their corporate trade secrets are not leaked when they contribute to shared pools of data used for model training. The various aspects of data are described in ISO/IEC 19944 and the new version of it.	
Societal Concerns	Description	Ability for industry players to share their data with their partners to develop ML-based algorithms while protecting their IP and interest would allow for flourishing of commercial AI/ML applications and solutions.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.73.2 Data

Data characteristics	
Description	Jet engine telemetry and maintenance logs
Source	Airlines and jet engine manufacturers
Type	Numeric values representing telemetry of various components in the engine
Volume (size)	Very large, terra-bytes
Velocity	High. A Jet engine can produce extremely large quantities of telemetry during regular operation

Variety	Telemetry and maintenance logs
Variability (rate of change)	
Quality	

A.73.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Data collection	Large telemetry data set containing data from selected jet engine obtained via satellite link during the flight, or from the engine itself after the flight has landed			
2	Data preparation	Process and normalize the training data obtained from the first step to prepare them for use in DNN model training for data pattern recognitions			
3	Model Training	Large training data set, with deep learning method, to develop model for predictive maintenanc			

		e of jet engines			

A.74 Infant SID

A.74.1 General

<table>
<tr><td>ID</td><td colspan="4">74</td></tr>
<tr><td>Use case name</td><td colspan="4">Infant SID</td></tr>
<tr><td>Application domain</td><td colspan="4">Healthcare</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">Prototype</td></tr>
<tr><td>Scope</td><td colspan="4">Use of facial recognition in healthcare</td></tr>
<tr><td>Objective(s)</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">ML-based facial recognition technology detects when infant is lying on her back or face down, alerting care taker to intervene when infant in on her stomach, hence lowering the statistical chance of infant death syndrome (SID)</td></tr>
<tr><td>Complete description</td><td colspan="3">Statistical analysis has shown that the chance of infants dying from Infant Death Syndrome (SID) is lower when the infants lie on their back, as opposed to faced down.

A cost-effective solution could be built for infant monitoring and alert system using a Webcam connected over the Internet to a customized facial recognition technology implemented as a cloud service.

The cloud service analyzes the periodic snapshots taken from the infant and uploaded to the cloud service. Once the ML-based facial recognition software in the cloud analyzes the snapshot from the infant and determines that the infant is no longer lying on her back, the service alerts the parents or care takers to attend to the infant by sending them a SMS text message or making an automated phone call.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Healthcare industry, cloud-based AI providers, healthcare public policy makers, parents of young children, insurance industry</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Privacy concerns, data subjects</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Accuracy</td><td>Accurately and reliably recognize the infant position</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>

AI features	Task(s)	Recognition
	Method(s)	Deep NN
	Hardware	1) High performing CPU nodes or GPUs in Cloud Computing Data Centers to train the DNN model 2) Cloud based VMs to run the trained DNN model
	Topology	
	Terms and concepts used	Deep Neural Networks – customized for infant facial recognition
Standardization opportunities/ requirements	- Need for transparency about the properties and sources of large quantities of infant facial data used to train DNN model - Need for transparency of aspects of training data such as PII, and potentially racial or ethnic bias in the data due to the size, source and content of the training data used may affect the effectiveness of the trained algorithm when used to recognize infant from different race or ethnicity	
Challenges and issues	- Explainability and transparency regarding the training data used, from the perspective of privacy concerns, and racial and ethnics biases which may be unintentionally built into the trained model. - Need a structured, common and standardized way to describe the stages of the machine learning model training process, and the types and aspects of the data used in the various stages of the process so the stakeholders (policy makers, privacy advocates and customers) can build confidence and trust in such ML-based product or service. The various aspects of data are described in ISO/IEC 19944 and the new version of it.	
Societal Concerns	Description	- Cost and availability of the ML-based service for low income populations who may not have access to high speed internet access or may not afford the ML-based cloud service - Any unintentional bias built into the training data used which may hinder effectiveness of the algorithm when used with infants from other races or ethnic backgrounds
	SDGs to be achieved	Good health and well-being for people

A.74.2 **Data**

Data characteristics	
Description	Infant photos
Source	Public or private collections of infant photos
Type	Unstructured photo images
Volume (size)	Very large, terra-bytes
Velocity	
Variety	photos
Variability	Quality and resolution of photos in training set could vary

(rate of change)	
Quality	Quality of training data (infant photos) could vary

A.74.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Data collection	Large data set containing diverse types of infant photos from different parts of the world			
2	Data preparation	Process and normalize the training data obtained from the first step to prepare them for use in DNN facial recognition model training			
3	Model Training	Large training data set, with deep learning method, to develop model for facial recognition of infants			

A.75 CRWB Recommendation benchmark

A.75.1 General

<table>
<tr><td>ID</td><td colspan="4">75</td></tr>
<tr><td>Use case name</td><td colspan="4">CRWB Recommendation benchmark</td></tr>
<tr><td>Application domain</td><td colspan="4">Other (please specify) Cooking recipe, nutrition, health</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">Prototype</td></tr>
<tr><td>Scope</td><td colspan="4">Cooking recipe execution plan decision support and nutrition recommendation</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Machine Data understandable</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Recommendation benchmark based on a cooking recipe dataset of cooking recipe execution plans</td></tr>
<tr><td>Complete description</td><td colspan="3">Recommendation benchmark is based on a cooking recipe data expressed in a Machine understandable language including Explicit knowledge on the way to proceed the cooking recipe actions.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Cookware and kitchenware industry</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Healthiness trust</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Nutrition rules</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Satisfaction</td><td>User satisfaction</td><td></td></tr>
<tr><td>2</td><td>Optimal experience</td><td>perception</td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recommendation</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine learning-based multi-dish generation and optimisation</td></tr>
<tr><td>Hardware</td><td colspan="3">cloud</td></tr>
<tr><td>Topology</td><td colspan="3">distributed</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Natural language processing, robotic process automation</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Health recommendation</td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Personal expectation related to flavor, taste and texture</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Local Production for Local Consumption</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3">Responsible consumption and production</td></tr>
</table>

A.75.2 Data

<table>
<tr><td colspan="2">Data characteristics</td></tr>
<tr><td>Description</td><td>CRWB data set (cooking recipes without border)</td></tr>
</table>

Source	Private cooking recipe collection
Type	Unsupervised structured multimedia/multisemedia
Volume (size)	gigabytes scales
Velocity	Daily
Variety	Cooking recipes
Variability (rate of change)	Depending on the community members and activity rate
Quality	Quality assessment during the data ingestion

A.75.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Carbon footprint estimation	Evaluation of carboon footprint impact of cooking recipes	New Recipe ingestion or citizen request	No estimation and any recipe ingredient or action update	carbon footprint debit or credit
2	Nutrition estimation and Recommended Dietary Allowance	Evaluation of the nutrition estimation of cooking recipes	New Recipe ingestion or citizen request	New nutritional constraints	Nutritional qualification of the estimated cooking recipes
3	Allergen elimination	Allergy Elimination and ingredient replacement	No free allergen recipe	Existing allergen ingredients	Free-allergen recipe

A.75.4 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	article	10.1109/ICDEW.2018.00032	Publisher: IEEE	Start point	Frederic Andres	shorturl.at/epVZ8
2	event		@ICDE2018	Community increase	DECOR workshop 2018	shorturl.at/AIQS7
3	event		@ICDE2019	Community increase	DECOR workshop 2019	shorturl.at/kxBE2
4	event		@ICDE2020	Community increase	DECOR workshop 2020	To be added

A.76 Flavorlens

A.76.1 General

<table>
<tr><td>ID</td><td colspan="4">76</td></tr>
<tr><td>Use case name</td><td colspan="4">Flavorlens</td></tr>
<tr><td>Application domain</td><td colspan="4">Other (please specify) Tasting sharing experience</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">Prototype</td></tr>
<tr><td>Scope</td><td colspan="4">Multi-sensing Dish tasting experience sharing in a social media ecosystem</td></tr>
<tr><td>Objective(s)</td><td colspan="4">users share their experiences and dish recommendation</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Social network to enable dish tasting experiences</td></tr>
<tr><td>Complete description</td><td colspan="3">Flavorlens, a mobile AI-based application for sharing dish tasting experiences. Each dish tasting experience is an observation which consists of one or more photographs, a title, a location tag, a description, a rating, a sensoring experience reporting about flavors, textures, and odors of a particular dish.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Cookware and kitchenware industry</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Healthiness trust</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Nutrition rules</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Satisfaction</td><td>User satisfaction</td><td></td></tr>
<tr><td>2</td><td>Optimal experience</td><td>perception</td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recommendation</td></tr>
<tr><td>Method(s)</td><td colspan="3">Approach using vector machines, artificial neural network and natural language processing</td></tr>
<tr><td>Hardware</td><td colspan="3">cloud</td></tr>
<tr><td>Topology</td><td colspan="3">distributed</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Multimedia processing, robotic process automation</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Food preference recommendation</td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Personal expectation related to flavor, taste and texture</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Local healthy dish for user satisfaction and preference</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3">Good health and well-being for people</td></tr>
</table>

A.76.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	event		@ICDE2018	Community increase	DECOR workshop 2018	shorturl.at/AIQS7
2	event		@ICDE2019	Community increase	DECOR workshop 2019	shorturl.at/kxBE2
3	event		@ICDE2020	Community increase	DECOR workshop 2020	To be added

[1] Alexandra Fritzen, Frederic Andres, and Maria Leite. 2018. Introducing Flavorlens: A Social Media Platform for Sharing Dish Observations. In Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction (MHFI'18). ACM, New York, NY, USA, Article 7, 7 pages. DOI: https://doi.org/10.1145/3279954.3279961

A.77 Water Crystal Mapping

A.77.1 General

ID	77			
Use case name	Water Crystal Mapping			
Application domain	Other (please specify) Water quality monitoring			
Deployment model	Cloud services			
Status	Prototype			
Scope	Increase citizen awareness on the quality of water			
Objective(s)	Map of the similarity of water crystals			
Narrative	Short description (not more than 150 words)	Deep learning-based approach to automatically classify water crystals.		
	Complete description	The deep learning approach identifies several kinds of symmetry for each water crystal in the EPP dataset. It will enable to extract similarities of three-dimensional structural data.		
Stakeholders	citizens, municipality, county, regions, UN,			
Stakeholders' assets, values	Sustainable Development Goal 6 - UN Sustainable Development (water)			
System's threats & vulnerabilities	Nutrition rules			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Water quality	Water crystal ranking	
	2	Crystal similarity	Crystal classification in the water crystal map	
AI features	Task(s)	Water Crystal similarity ranking		

	Method(s)	Deep learning approach and crystal structure embeddings
	Hardware	cloud
	Topology	distributed
	Terms and concepts used	Water crystal structure
Standardization opportunities/ requirements	water crystal knowledge standardisation	
Challenges and issues	Water quality, ice memory	
Societal Concerns	Description	Sustainable Development Goal 6 - UN Sustainable Development (water)
	SDGs to be achieved	Clean water and sanitation

A.77.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	book	1		Scientific foundation	Prof. Pollack	
2	event		2018	Research Community increase	Water conference on the Physics, Chemistry, and Biology of Water	https://archives.waterconf.org
3	Article	2		New challenge in the field	Fritz-Haber-Institut der Max-Planck-Gesellschaft	https://www.nature.com/articles/s41467-018-05169-6
4	event		2019	Research Community increase	Water conference on the Physics, Chemistry, and Biology of Water	https://waterconf.org/
5	event		2020	Community increase	Water conference on the Physics, Chemistry, and Biology of Water	To be added

[1] The Fourth Phase of Water: Beyond Solid, Liquid, and Vapor. By Gerald H. Pollack, Ebner & Sons Publishers, 2013; 357 Pages. ISBN 978-0-9626895-4-3

[2] Angelo Ziletti, Devinder Kumar, Matthias Scheffler & Luca M. Ghiringhelli. Insightful classification of crystal structures using deep learning, Nature Communications, volume 9, Article number: 2775 (2018)

A.78 Ontologies for Smart Buildings

A.78.1 General

ID	78
Use case name	Ontologies for Smart Buildings

Application domain	Smart Buildings	
Deployment model	Hybrid (Cloud but also locally in the buildings)	
Status	Prototype	
Scope	Renovation of buildings, improve the life's quality of residents - limited to data issues in a building, - Audience: citizen, public and private actors, companies involved in the ICT System managing the building. Building Management System (BMS) is not the limited scope, we would like to open it to data produced by residents, coupled with data coming from BMS.	
Objective(s)		
Narrative	Short description (not more than 150 words)	The general question is How to build and to standardize ontologies for data produced, in a broad sense, in a building. Data are coming both from the System managing the building but also from residents.
	Complete description	Seminal and technical papers introducing the vocabulary, definitions, concepts of smart buildings are [1,2,3,4,5]. The common view and shared definition of the community is that a smart building is a construction with an appropriate design and technological support to maximize its functionalities and comfort for their occupants with the compromise to reduce their operational costs, and extend the life of the physical structure [1]. In [2] authors presented an initial guide to understand the layers, taxonomy of services and best practices for the development of smart buildings. Open standards are claimed in order to increase interoperability between layers and services. In [3] authors explained variations between different notions. The findings of the paper allow to clarify and to define the border between the intelligent and the (more advanced) Smart Building. The upper bound of the Smart Building is defined by (the future development of) the predictive building. To simplify a little, from a System point of view, we may think an Intelligent Building as a building reacting to some events whereas Smart Buildings "are buildings which integrate and account for intelligence, enterprise,control, and materials and construction as an entire building system, with adaptability, not reactivity, at the core, in order to meet the drivers for building progression: energy and efficiency, longevity, and comfort and satisfaction." The INTEL online document [4] is oriented towards Internet of Things and Building Management System (BMS). Analogous to a supervisory control and data acquisition (SCADA) system used in manufacturing, a building management system (BMS) monitors and controls various building systems, such as heating, ventilation, air conditioning (HVAC), and lighting with additional and often separate systems to control elevators, fire, safety, security, and access controls. We will explain later on that our work,

<table>
<tr><td></td><td></td><td>at the System level, is not about BMS that we consider to be not able to learn on the data it is managing.

The technical document [5] gives more details about BMS, Direct Digital Control (DDC), Building Automation System (BAS), Facility Master System Integrator (FMSI) all of them are defined according to a System point of view. The system we propose is more like an operating system for the building or like an orchestrator of machine learning tasks or computing tasks and it does not looks like any of these systems.

At last the Residential Buildings System project, from the Berkeley Lab (https://homes.lbl.gov/publications) is also a good source of papers, from 1978 until today, related to Smart Buildings with a special focus on the movement of air and associated penalties involving distribution of pollutants, energy and fresh air.

The ISO process or technology regulations related to Smart Buildings are ISO 16484-2:2004 (Building automation and control systems hardware), ISO 16484-6:2009 (Building automation and control systems data communication conformance testing), ISO 16484-5:2012 (Building automation and control systems data communication protocol), ISO 16484-3:2005 (Building automation and control systems functions). They are not related to AI nor to data produced by residents.

The objective of the use case is to study existing (open) data, and to build new tools to collect data produced in a building in order to classify them in ontologies. To be short, an ontology is a knowledge as a set of concepts. The idea behind the standardisation, here, is to “put” some order in the brute data and to extract general knowledge. There is a lack, in the Smart Building field, to structure the data, all types of data in order to infer and based decisions or reactions on general knowledge instead of scattered facts. We are also guessing here that a ‘collective’ intelligence/knowledge helps a lot for taking ‘good’ decisions for people living in buildings.</td></tr>
<tr><td>Stakeholders</td><td colspan="2">Those that can affect the AI system: since it is under the supervision of a university, the data exchange with the building is controlled by the Networking team of the university and the person in charge of the Security. A university network is not so open! It is not like with Internet for individuals. A group of persons in charge of the GDPR (General Data Protection Regulation) will also be deployed during the use case.</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="2">Residents/users of a building (the initial use case is related to a university building). We need other buildings.</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="2">Physical intrusions on sensors located in the building are possible if sensors are not protected (physical) as well as servers. Injections into the database is also possible if not managed. Intrusion Detection Systems are already deployed. Another threats will be ‘data stolen’ and (re)identification of persons. This</td></tr>
</table>

	implies that the database should be designed with respect to GDPR (as promoted in Europe)			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
AI features	Task(s)			
	Method(s)	A call of volunteers (equipped with Smartphones) and sensors will send data in a server.		
	Hardware	Sensors with communicating according to the IoT standards+server to collect the data + a time series database such as InfluxDB and Timescale.		
	Topology	1 Server + N sensors connected to the server + volunteers		
	Terms and concepts used			
Standardization opportunities/ requirements				
Challenges and issues				
Societal Concerns	Description			
	SDGs to be achieved			

3927

3928 A.78.2 Data

Data characteristics	
Description	A project of a French team working at the university of Paris 13, related to smart buildings, has been selected this year with Reves de Scenes Urbaines (RSU(http://www.urbanisme-puca.gouv.fr/plaine-commune-93-divd-reve-de-scenes-urbaines-a822.html)), the industrial demonstrator of the sustainable city, located in St Denis (department 93 in the north of Paris). This non-funded project in partnership with Qarnot Computing is part of the building renovation of the Institute of Technology (IUT) of St Denis. In terms of "demonstration", the project aims to deploy a sensor infrastructure in the IUT and collect the data using the OASIS Qarnot tool (https://www.qarnot.com/oasis_os_building/). Real data coming from a real building should be available for our purpose.
Source	
Type	
Volume (size)	
Velocity	
Variety	
Variability (rate of change)	
Quality	

3929

A.78.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1				Terminology	Johnson Controls	see below
2				Description of ISO standards on Smart Buildings	IEEE-GDL CCD	see below
3				Terminology	Position paper	see below
4				Position paper of INTEL regarding Smart Buildings (example)	INTEL	see below
5				Technical paper at the System Level - Definition and terminology	Lonmark	see below

[1] T. Hoffmann, “Smart Buildings,” Johnson Controls, Inc., pp. 1-8, October, 2009. Available at https://www.scribd.com/document/259029136/Smart-Buildings

[2] V.M. Larios, J.G. Robledo, L. Gómez,and R. Rincon, “IEEE-GDL CCD Smart Buildings Introduction”, white paper of the working group of physical infrastructure available online at https://smartcities.ieee.org/images/files/pdf/whitepaper_phi_smartbuildingsv6.pdf

[3] A.H. Buckman M. Mayfield Stephen B.M. Beck, "What is a Smart Building?", Smart and Sustainable Built Environment, Vol. 3 Iss 2 pp. 92 - 109, (2014). Permanent link to this document:http://dx.doi.org/10.1108/SASBE-01-2014-0003

[4] Intel, “Designing More Affordable Smart Buildings Solutions”, white paper available at https://www.intel.com/content/dam/www/public/us/en/documents/solution-briefs/iot-smart-building-solutions-brief.pdf

[5] Ron Bernstein, “Building Automation Training and LonMark Certification Institute Programs” Available at https://www.lonmark.org/connection/presentations/2017/AHR/Session%202/Session%202%20-%20Ron%20Bernstien%20Smart%20Buildings%20Course%20101%20-%20Key%20Concepts,%20Definitions%20and%20Elements.pdf

A.79 Discharge Summary Classifier

A.79.1 General

ID	79

<table>
<tr><td>Use case name</td><td colspan="4">Discharge Summary Classifier</td></tr>
<tr><td>Application domain</td><td colspan="4">Healthcare</td></tr>
<tr><td>Deployment model</td><td colspan="4">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Decision Tree, Random Forest, SVM, BNN, Deep Learning</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Classification of Discharge Summaries</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">This system proposes a method for construction of classifiers for discharge summaries.</td></tr>
<tr><td>Complete description</td><td colspan="3">This system proposes a method for construction of classifiers for discharge summaries. First, morphological analysis is applied to a set of summaries and a term matrix is generated. Second, correspond analysis is applied to the classification labels and the term matrix and generates two dimensional coordinates. By measuring the distance between categories and the assigned points, ranking of key words will be generated. Then, keywords are selected as attributes according to the rank, and training example for classifiers will be generated. Finally learning methods are applied to the training examples. Experimental validation shows that random forest achieved the best performance and the second best was the deep learner with a small difference, but decision tree methods with many keywords performed only a little worse than neural network or deep learning methods.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Medical Staff</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Quality of Medical Care</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Bias in Hospital Texts</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Accuracy</td><td>Classification Accuracy</td><td>Check of Decision Summaries</td></tr>
<tr><td>2</td><td>Length of Stay</td><td>Length of Stay in Inpatient Ward</td><td>Management of Ward</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Knowledge processing & discovery</td></tr>
<tr><td>Method(s)</td><td colspan="3">Text Mining, Decision Tree, Random Forest, SVM, BNN, Deep Learning</td></tr>
<tr><td>Hardware</td><td colspan="3">Servers for Analytics (PREMERGY, Z8), Data Servers (Primergy)</td></tr>
<tr><td>Topology</td><td colspan="3">Network of Data and Analytics Servers</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Text Mining, Decision Tree, Random Forest, SVM, BNN, Deep Learning</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4">Big Data Analytics</td></tr>
<tr><td>Challenges and issues</td><td colspan="4">Computational Complexity</td></tr>
</table>

Societal Concerns	Description	Refinement of Medical Texts Medical Hospital Management
	SDGs to be achieved	Good health and well-being for people

A.79.2 Data

Data characteristics	
Description	
Source	Hospital Information System
Type	Text, Numerical: Time-series
Volume (size)	Text: 1GB
Velocity	Real time
Variety	Text, Numerical, (Time series)
Variability (rate of change)	Every hours
Quality	Records: Dependent on Medical Staff, Numerical: Automatic

A.79.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Shusaku Tsumoto, Tomohiro Kimura, Haruko Iwata, Shoji Hirano:Construction of Discharge Summaries Classifier. ICHI 2017: 74-82	On Demand Usage	Original	Medical Informatics, Shimane University Hospital	https://doi.org/10.1109/ICHI.2017.92

A.80 Generation of Clinical Pathways

A.80.1 General

ID	80	
Use case name	Generation of Clinical Pathways	
Application domain	Healthcare	
Deployment model	On-premise systems	
Status	In operation	
Scope	Decision Tree, Clustering	
Objective(s)	Nursing clinical pathway	
Narrative	Short description (not more than 150 words)	This system proposes a temporal data mining method to construct and maintain a clinical pathway used for schedule management of clinical care.
	Complete description	This system proposes a temporal data mining method to construct and maintain a clinical pathway used for schedule

		management of clinical care. Since the log data of clinical actions and plans are stored in hospital information system, these histories give temporal and procedural information about treatment. The method consists of the following four steps: First, histories of nursing orders are extracted from hospital information system. Second, orders are classified into several groups by using clustering and multidimensional scaling method. Third, by using the information on groups, feature selection is applied to the data and important features for classification are extracted. Finally, original temporal data are split into several groups and the first step will be repeated. After the grouping results are stable, a new pathway is constructed .based on the induced results.		
Stakeholders	Nursing Staff			
Stakeholders' assets, values	Quality of Medical Care			
System's threats & vulnerabilities	Bias in Hospital Data			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Pathway Complexity	Complexity of Nursing Orders	Management of Nursing Orders
	2	Length of Stay	Length of Stay in Inpatient Ward	Management of Ward
AI features	Task(s)	Knowledge processing & discovery		
	Method(s)	Decision Tree, Clustering		
	Hardware	Servers for Analytics (PREMERGY, Z8), Data Servers (Primergy)		
	Topology	Network of Data and Analytics Servers		
	Terms and concepts used	Decision Tree, Clustering, OLAP		
Standardization opportunities/ requirements	Big Data Analytics			
Challenges and issues	Computational Complexity			
Societal Concerns	Description	Good Practice of Medical Services		
	SDGs to be achieved	Good health and well-being for people		

A.80.2 Data

Data characteristics	
Description	
Source	Hospital Information System
Type	Text, Numerical: Time-series
Volume (size)	Text: 1GB
Velocity	Real time
Variety	Text, Numerical, Image (Time series)

Variability (rate of change)	Every minutes/hours
Quality	Records: Dependent on Medical Staff, Numerical/Image: Automatic

A.80.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Haruko Iwata, Shoji Hirano, Shusaku Tsumoto:Maintenance and Discovery of Domain Knowledge for Nursing Care using Data in Hospital Information System. Fundam. Inform. 137(2): 237-252 (2015)	On Demand Usage	Original	Medical Informatics, Shimane University Hospital	https://doi.org/10.3233/FI-2015-1177
2	Paper	Shusaku Tsumoto, Shoji Hirano, Haruko Iwata:Data decomposition and dual clustering for clinical care management. BigData 2015: 1475-1584	On Demand	Original	Medical Informatics, Shimane University Hospital	https://doi.org/10.1109/BigData.2015.7363923

A.81 Hospital Management Tools

A.81.1 General

ID	81	
Use case name	Hospital Management Tools	
Application domain	Healthcare	
Deployment model	On-premise systems	
Status	In operation	
Scope	Temporal Data Mining, Visualization	
Objective(s)	Hospital Management	
Narrative	Short description (not more than 150 words)	Temporal Data Mining Methods (Multi-scale comparison with clustering and Temporal Frequent Item Sets) is applied to Hospital Data.
	Complete description	A scheme for innovation of hospital services based on data mining. Then, based on this scheme, data mining techniques are applied to data extracted from hospital information systems. The results included several interesting findings, which suggests that the reuse of stored data will provide a powerful tool to improve the quality of hospital services.

Stakeholders	Hospital Administrator			
Stakeholders' assets, values	Visualization of Medical Staff Behavior in Hospital			
System's threats & vulnerabilities	Bias in Hospital Data			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Waiting Time	Waiting Time of Outpatient Clinic	Management of Outpatient Clinic
	2	Length of Stay	Length of Stay in Inpatient Ward	Management of Ward
AI features	Task(s)	Knowledge processing & discovery		
	Method(s)	Temporal Data Mining, Clustering		
	Hardware	Servers for Analytics (PREMERGY, Z8), Data Servers (Primergy)		
	Topology	Network of Data and Analytics Servers		
	Terms and concepts used	Trajectories Mining, Clustering, OLAP		
Standardization opportunities/ requirements	Big Data Analytics			
Challenges and issues	Computational Complexity			
Societal Concerns	Description	Good Practice of Medical Services		
	SDGs to be achieved	Good health and well-being for people		

A.81.2 Data

Data characteristics	
Description	
Source	Hospital Information System
Type	Text, Numerical, Images: Time-series
Volume (size)	Text: 1GB, Images: 4TB
Velocity	Real time
Variety	Text, Numerical, Image (Time series)
Variability (rate of change)	Every second/hours
Quality	Records: Dependent on Medical Staff, Numerical/Image: Automatic

A.81.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Shusaku Tsumoto, Haruko Iwata, Shoji Hirano, Yuko	On Demand Usage	Original	Medical Informatics, Shimane	https://doi.org/10.1016/j.future.2013.10.014

		Tsumoto:Similarity-based behavior and process mining of medical practices. Future Generation Comp. Syst. 33: 21-31 (2014)			University Hospital	
2	Paper	Toshihiko Kawamura, Tomohiro Kimura, Shusaku Tsumoto:Estimation of Service Quality of a Hospital Information System Using a Service Log. The Review of Socionetwork Strategies8(2): 53-68 (2014)	On Demand	Original	Medical Informatics, Shimane University Hospital	https://doi.org/10.1007/s12626-014-0044-x

A.82 Surgeries Improvement of productivity of semiconductor manufacturing

A.82.1 General

ID	82	
Use case name	Improvement of productivity of semiconductor manufacturing	
Application domain	Manufacturing	
Deployment model	On-premise systems	
Status	In operation	
Scope	Analysis of data taken from production equipment and improvement of productivity based on the analysis	
Objective(s)	Cost reduction of semiconductor manufacturing	
Narrative	Short description (not more than 150 words)	In modern semiconductor manufacturing, huge amount of data are gathered and used to improve yields. However, it is difficult even for skilled engineers to promptly achieve the improvements by means of manual analysis because of the complexity of the production process and the scale of the data. In Yokkaichi operation, where more than 5,000 pieces of equipment are working and two billion records of data are daily created, it is difficult to secure enough engineers to resolve problems arise in the production. Toshiba Memory Corporation tackled the issue with AI technology including machine learning. The endeavor resulted in improvement of the productivity through the stable quality based on semi-automated data analysis.
	Complete description	This use case consists of the following three themes. 1. Support of analysis of cause of failure based on wafer map patterns

		At the final stage of semiconductor manufacturing, each chip on a wafer is tested and a pattern how the failure chips are distributed on the wafer is produced (Fig.1). Analysis of the cause of the failure is carried out based on the pattern and the history of usage of manufacturing devices. The analysis is supported by the following four technologies. 1.1 Clustering of wafer map patterns Clustering of the wafer map patterns are carried out in order to grasp the overview of the occurrence of the failure. Because there are 200 thousands of wafers per month, a fast clustering algorithm is required to promptly provide information to engineers. Making use of Scalable k-means++, the clustering process is 72.5 times faster than the previous method. 1.2 Cause estimation based on pattern mining If a manufacturing device frequently occurs in the history of a wafer belongs to a wafer map cluster and the device seldom occurs in the history of other wafers then the device is likely to be the cause of the failure. The candidates of the cause of the failure and their likelihoods are calculated based on the number of occurrences of the combinations of the devices promptly counted by a pattern mining algorithm FPGrowth and ranking through chi-square test. 1.3 Wafer map classification based on CNN A wafer map is classified into registered typical wafer maps in order to monitor the recurrence of the failure. The classification accuracy (F1 score) with SVM was 0.898. Making use of CNN, the accuracy is improved to 0.95. 1.4 Web portal for yield analysis The information provided by the above technologies are shown in a web portal (Fig.2). The portal has improved the average analysis time from six hours to two hours. 2. Automatic classification of SEM images of defects Tests of wafers are carried out not only at the final stage of the production but also between processes, where the result of the previous processes is checked. One of the tests is classification of images of microscopic aspects of the defects observed by scanning electron microscope (SEM) (Fig.3). Thirty thousands of the images are daily taken. It is an important test because the class of a defect may provide valuable insight for cause estimation. Previously the classification was carried out semi-automatically by an engineer with a tool with classification function. However, human work load was relatively high because the tool's ability was quite limited. Making use of CNN, the number of

		defect categories that are automatically classifiable has dramatically increased. Now the automation ratio is 83%, improved from 49%.

3. Analysis of cause of variation of quality characteristic value

In Yokkaichi operation, the cause of the variation of a quality characteristic value is identified and the yield is kept by countermeasure against it. For quick identification, various data including process parameters and sensor measurements from a manufacturing device are stored in DB, therefore the number of attributes becomes huge at the completion of the production. It is not uncommon for the number of the attribute to be much greater than the number of products to be analyzed, sometimes by several orders. Making use of Lasso regression for data with 23,600 attributes and 303 products, a regression model predicting a quality characteristic value has been built, with automatic feature selection. Engineers' cause identification tasks are also supported by a network diagram visualizes causal structure of the selected features. As a result, the average analysis time is improved to one day from seven days.

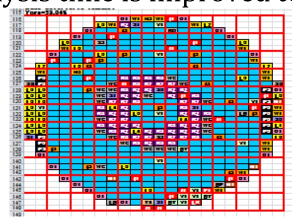


Fig.1 Wafer map sample

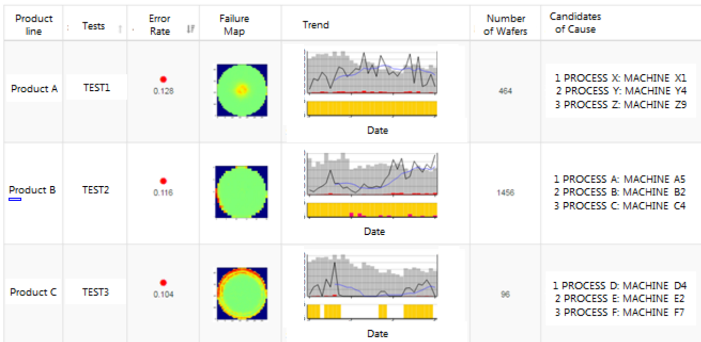


Fig.2 Sample screen shot of the web portal

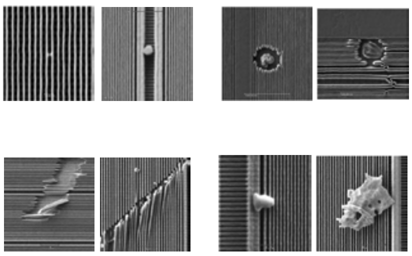


Fig.3 SEM images

		This proposal is based on the use case collection initiative promoted by Japanese Society of Artificial Intelligence (JSAI).		
Stakeholders	Executives of semiconductor manufacturing companies			
Stakeholders' assets, values	Competitive edge based on manufacturing cost reduction Business continuity based on the fewer number of required data scientists			
System's threats & vulnerabilities	Delay of the analysis tasks caused by inaccurate AI outputs Delay of countermeasure deployment caused by a fact that the physical model of a failure is unknown			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy of wafer map classification	Classification accuracy in the theme 1.3	How accurately to detect the recurrence of a failure
	2	Time to identify the cause of failure	Time to complete the task corresponds to the theme 1	How quickly to identify the cause of a failure
	3	Accuracy of defect classification	Classification accuracy in the theme 2	How accurately to classify the defect SEM images
	4	Accuracy of feature selection	Accuracy of feature selection in the theme 3	How accurately to select important features to quality characteristic values
AI features	Task(s)	Other (please specify) Recognition, Prediction, Optimization, Interactivity, Recommendation		
	Method(s)	Clustering, Pattern Mining, CNN, Web Portal, Lasso Regression		
	Hardware	PC cluster with GPU		
	Topology			
	Terms and concepts used	Yield analysis, Wafer map pattern, Defect SEM images, Quality prediction, Web portal		
Standardization opportunities/ requirements	Standardization of kinds and formats of data taken from manufacturing devices Standardization of kinds and formats of outputs from AI			
Challenges and issues	Guarantee of correctness of analysis by AI Automatic physical model building for a failure			
Societal Concerns	Description	Hollowing out of analytic know-how		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.82.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

[1] Nakata, K., Orihara, R., Mizuoka, Y. and Takagi, K. A Comprehensive Big-Data-Based Monitoring System for Yield Enhancement in Semiconductor Manufacturing. IEEE Transactions on Semiconductor Manufacturing, November 2017, vol. 30, no. 4, pp.339-344.

[2] Imoto, K., Nakai T., Ike T., Haruki K. and Sato, Y. A CNN-based Transfer Learning Method for Defect Classification in Semiconductor Manufacturing, Proc. ISSM 2018, 2018.

[3] Takada, M., Saiki, S., Sueyoshi, S., Eguchi, H., and Nishikawa T. Intelligent Causal Analysis System for Wafer Quality Control using Sparse Modelling. Proceedings of AEC/APC Symposium Asia, 2017.

A.83 IFLYTEK Intelligent marking system

A.83.1 General

ID	83	
Use case name	IFLYTEK Intelligent marking system	
Application domain	Education	
Deployment model	On-premise systems	
Status	In operation	
Scope	It can realize intelligent detection and grading of all subjective questions	
Objective(s)	To reduce a lot of labor and organizational costs	
Narrative	Short description (not more than 150 words)	Iflytek intelligent marking system is based on the core technology design research, including iflytek independent intellectual property rights handwritten recognition, natural language understanding, intelligent evaluation and other artificial intelligence and so on. It can realize the detection of blank questions for all types of questions except multiple choice questions, and the computer intelligent evaluation of Chinese, English composition, English translation, Literature synthesis category short answer questions and English blank questions. At the same time, for Chinese composition and English composition, it can also effectively detect the abnormal answer papers which are highly similar to the dry content of the test paper or the content of the external model text.
	Complete description	The intelligent marking system can provide a new generation of intelligent scanning network evaluation solution for large-scale paper and pen examination combined with the mature scanning network evaluation technology. In the process of scanning, the detection and screening of similar volume, blank volume and the intelligent evaluation of subjective questions are carried out in real time. Taking the data outputted from the scanning link as the objective third party quality evaluation standard, the online or offline quality monitoring of the marking paper is carried out to improve the quality of the marking paper. At the same time, the computer intelligent evaluation of subjective questions can assist manual marking to a certain extent, and effectively reduce the workload of manual marking of subjective questions. Intelligent marking system has many advantages. First, it has a scientific and unified scoring standard, which can

		avoid the difference of scoring scale and subjective interference among different reviewers, and ensure the fairness of the marking results. Second, it only needs to invest a small number of technical personnel and servers, which can reduce the organizational cost of existing manual marking by about 50%. Third, it can detect the abnormal situation of the answer, such as blank questions, similar volumes. At the same time, through the real-time comparison with the manual marking data to achieve the quality monitoring.		
Stakeholders	Marking teacher and technicist			
Stakeholders' assets, values	Efficiency			
System's threats & vulnerabilities	Accuracy			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Cost	Reduce the cost of existing manual marking	
	2	Efficiency	Improve the efficiency of existing manual marking	
	3	Accuracy	Improve the accuracy of existing manual marking	
AI features	Task(s)	Natural language processing		
	Method(s)	Deep learning, semantic recognition		
	Hardware			
	Topology			
	Terms and concepts used	Deep learning: a class of machine learning algorithms use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Semantic recognition: analytical techniques for the meaning and emotion of discourse		
Standardization opportunities/ requirements	After repeated training, the system can achieve at least 96% accuracy.			
Challenges and issues	The accuracy of marking paper needs to be further improved.			
Societal Concerns	Description	There is a scientific and unified scoring standard, which can ensure the fairness of the marking results. Reduced a lot of labor and organizational costs		
	SDGs to be achieved	Quality education		

3982

A.83.2 Data

Data characteristics	
Description	Scanning student papers
Source	The data from scanning student papers
Type	Text, Picture
Volume (size)	
Velocity	Batch Processing
Variety	Single source
Variability (rate of change)	Static
Quality	High

A.83.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training data set	Sample raw data set is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		Meeting KPI requirements
3	Execution	Intelligent marking using training Model	Complete the scoring of the scanned test paper	The trained model has been evaluated as deployable	
4	Retraining	Retrain a model with training data set	Certain period of time has passed since the last training/re training		

A.83.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement

1	Sample raw data set is ready	Get the test paper that passes the system scan	Intelligent marking system		
2	Completion of Step 1	Training data set creation	Intelligent marking system		
3	Comparison	Comparing the results of manual reading and intelligent system marking	Teachers		
4	Completion of Step 2 and 3	Model training	Intelligent marking system		

Specification of training data	

A.83.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Get the test paper that passes the system scan	Intelligent marking system		
2	Completion of Step 1	Detection and grading	Intelligent marking system		
3	Completion of Step 2	Evaluation	Intelligent marking system		

Input of evaluation	
Output of evaluation	

A.83.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of the scanning of a student paper				
2	Completion of Step 1	Complete the scoring of the scanned test paper	Intelligent marking system		

Input of Execution	
Output of Execution	

A.83.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Certain period of time has passed since the last training/retraining	Get the test paper that passes the system scan	Intelligent marking system		
2	Completion of Step 1	Training data set creation	Intelligent marking system		
3	Comparison	Comparing the results of manual reading and intelligent system marking	Teachers		
4	Completion of Step 2 and 3	Model training	Intelligent marking system		

Specification of retraining data	Retraining data set has to include recent data

A.83.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press release				IFlytek	https://www.iflytek.com/
2	Press release				IFlytek	https://mp.weixin.qq.com/s?__biz=MzA5NjYyMTA0OA%3D%3D&idx=1&mid=501756147&sn=c8f94e3f905fd5cf07a3cfae4b72ee43

A.84 Intelligent educational robot

A.84.1 General

ID	84
Use case name	Intelligent educational robot
Application domain	Education

Deployment model	On-premise systems			
Status	In operation			
Scope	It's the best partner of a child, and make the child learn in play			
Objective(s)	To improve the pleasure of learning.			
Narrative	Short description (not more than 150 words)	Educational robot is a new teaching tool to cultivate students' comprehensive ability. It mainly uses artificial intelligence technology, speech recognition technology and bionic technology to cultivate students' various abilities. Educational robots have hearing, vision, oral skills, recognition, emotional detection and the ability to interact for a long time.		
	Complete description	Educational robot is a new teaching tool to cultivate students' comprehensive ability. It mainly uses artificial intelligence technology, speech recognition technology and bionic technology to cultivate students' various abilities. Educational robots have hearing, vision, oral skills, recognition, emotional detection and the ability to interact for a long time. Recently, a popular educational robot called Little handsome Robot belongs to the educational robot of children's entertainment education. Its appearance is very cute, especially easy to be favored by children, it is suitable for children, primary school, junior high school students for study or entertainment. It has teaching materials and lectures for famous teachers in all grades and disciplines, and students can accept high-quality teaching without leaving home, and it can also present the knowledge forgotten by students and solve students' learning problems in time. Moreover, Correct students' dependence, hating to get out of bed, playfulness and other bad habits by giving instructions, intelligent reminders, so as to cultivate students' good learning behavior and living habits. For learning English, the handsome robot can train students' oral English ability by practicing dialogue with students, and can also make students' pronunciation more standard and improve students' communication ability.		
Stakeholders	Students, Parents, Teachers.			
Stakeholders' assets, values	Students' grades and learning interest			
System's threats & vulnerabilities	Teaching effect of intelligent robot			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Interest	Improve students' interest in learning	
	2	Grades	Improve students' academic performance.	
AI features	Task(s)	Recognition		

	Method(s)	Deep learning, Automatic Speech Recognition, Bionics techniques
	Hardware	
	Topology	
	Terms and concepts used	Deep learning: a class of machine learning algorithms use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Automatic Speech Recognition: converts the lexical content of human speech into computer-readable input Bionics techniques: bionic technology studies the functional principles and mechanism of various biological systems as biological models, and finally realizes the design of new technologies and makes better new instruments and machines.
Standardization opportunities/ requirements	After repeated training, the intelligent educational robots can accompany students to study like teachers	
Challenges and issues	Be able to sense students' emotions like teachers. Accurately capture students' gestures, postures, face information, etc.	
Societal Concerns	Description	To give students emotional support Stimulate students' interest in learning
	SDGs to be achieved	Quality education

3998

3999 ### A.84.2 Data

Data characteristics	
Description	Learner input, including pronunciation, visual information, keystrokes, etc.
Source	The data from learner
Type	Voices, Visual information, Keystrokes, etc.
Volume (size)	
Velocity	Batch Processing
Variety	Multiple source
Variability (rate of change)	Static
Quality	High

4000

4001 ### A.84.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with	Sample raw data set is ready		

		training data set			
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		Meeting KPI requirements
3	Execution	Intelligent educational robot using training Model	Complete the scoring of the scanned test paper	The trained model has been evaluated as deployable	
4	Retraining	Retrain a model with training data set	Certain period of time has passed since the last training/re training		

A.84.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample raw data set is ready	Get the data from the input of learner	Intelligent educational robot		
2	Completion of Step 1	Training data set creation	Intelligent educational robot		
3	Completion of Step 2	Model training	Intelligent educational robot		

Specification of training data	

A.84.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Get the data from the input of learner	Intelligent educational robot		
2	Completion of Step 1	Feedback	Intelligent educational robot		

3	Completion of Step 2	Evaluation	Intelligent educational robot		

Input of evaluation	
Output of evaluation	

A.84.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Get the data from the input of learner		Intelligent educational robot		
2	Completion of Step 1	Feedback	Intelligent educational robot		

Input of Execution	
Output of Execution	

A.84.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activit y	Requirement
1	Sample raw data set is ready	Get the data from the input of learner	Intelligen t educatio nal robot		
2	Completion of Step 1	Training data set creation	Intelligen t educatio nal robot		
3	Completion of Step 2	Model training	Intelligen t educatio nal robot		

Specification of retraining data	

A.84.8 References

References

No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Journal	[1]王兴月.人工智能在教育领域中的应用案例分析及发展前景[J].中小学电教,2019(Z1):30-34.	Published online			http://www.cnki.com.cn/Article/CJFDTotal-ZXDJ2019Z1012.htm

A.85 AI solution to intelligence campus

A.85.1 General

ID	85	
Use case name	AI solution to intelligence campus	
Application domain	Education	
Deployment model	Cloud services	
Status	In operation	
Scope	It is a full range of products and integrated solutions for teaching, examination, evaluation, management, learning	
Objective(s)	This scheme provides a comprehensive intelligent sensing environment and comprehensive information service platform for teachers and students, so as to realize the integration of human and business information.	
Narrative	Short description (not more than 150 words)	Based on big data and artificial intelligence technology, the scheme brings teaching, examination, learning and management into the integrated system of mutual cooperation, based on accompanying data acquisition and dynamic big data analysis, combined with process evaluation, to help teachers and students to realize teaching according to their aptitude and individualized learning, to help managers to supervise and assist decision-making, and to greatly promote the transformation of education, learning and management to intelligence.
	Complete description	In teaching, iFLYTEK built an intelligent and efficient classroom based on the cloud network end. Through docking the resource cloud platform and school-based resource library, it can realize synchronous push of high-quality resources and help teachers prepare class efficiently. In the examination, relying on iFLYTEK's leading voice and artificial intelligence core technology, iFLYTEK oral evaluation technology has only passed the certification of the National language Commission, widely used in the national Chinese Mandarin online test, and used in classroom teaching. IFLYTEK applies the industry exclusive artificial intelligence core technology to the examination and the automatic approval of traditional offline homework, which greatly reduces the burden of teachers' work and data the daily examination process. Big data analysis technology can be used to promote personalized teaching and learning.

		In learning, iFLYTEK realizes students' online adaptive learning by building question bank system, evaluation system and online learning system. Through the analysis of the students' examination results, we can evaluate the mastery of the students' knowledge points and the stability of their grades, and then combine the key points of the teaching materials with the high frequency test points. Through intelligent analysis, the optimal learning path recommendation can be given. In management, iFLYTEK Smart Campus solution covers more than 10 departments such as academic Affairs Office, Student Office, School Office and so on. The system provides more than 60 applications to meet the needs of normal campus management. It is worth mentioning that in order to cope with the challenges of educational administration brought by the new curriculum reform and the new college entrance examination reform, iFLYTEK, based on the classification algorithm of deep neural network, puts forward the intelligent course arrangement system, effectively avoids the conflict of course selection, and realizes the optimal voluntary satisfaction rate under the premise of the same teachers and classroom resources, so that every student can attend classes according to his own volunteers.		
Stakeholders	Student,Teacher,School, Government			
Stakeholders' assets, values	Privacy			
System's threats & vulnerabilities	Disclosure of privacy data for teachers and students			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Efficiency	Improve student's learning effect and teacher's office efficiency	Improve efficiency
AI features	Task(s)	Knowledge processing & discovery		
	Method(s)			
	Hardware			
	Topology			
	Terms and concepts used			
Standardization opportunities/ requirements				
Challenges and issues	The implementation of intelligent campus makes the data of students and teachers be collected and processed in large quantities, which is likely to lead to the disclosure of private data. Therefore, the establishment of data privacy protection mechanism should be strengthened in intelligent platform.			
Societal Concerns	Description	Intelligent campus solution leads artificial intelligence technology into the campus, into the classroom, promotes		

		students' learning and teachers' teaching, and facilitates teaching management.
	SDGs to be achieved	Quality education

4013

4014 A.85.2 Data

Data characteristics	
Description	The data comes from students and teachers as well as from their learning and office processes.
Source	Intelligent education products or platforms
Type	Structured/Unstructured data
Volume (size)	
Velocity	In real time
Variety	students information, teachers information, information generated during the course of teaching, learning and management.
Variability (rate of change)	In real time
Quality	

4015

4016 A.85.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training samples			
2	Evaluation	Evaluate whether the model is properly trained for the detection			Meeting KPI requirements (e.g. efficiency) of the particular case
3	Execution	Pick peaks using the trained model			
4	Retraining	Retrain a model with training samples			

4017 A.85.4 References

References

No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press Release				iFlytek	https://mp.weixin.qq.com/s?__biz=MzA4NjM4ODQzNQ%3D%3D&idx=1&mid=2651544421&sn=87bf38741ed5901fe6f6fd83ab98aa40
2	website				iFlytek	https://max.book118.com/html/2018/1202/8124132074001135.shtm

A.86 Product failure prediction for critical IT infrastructure

A.86.1 General

<table>
<tr><td>ID</td><td colspan="4">86</td></tr>
<tr><td>Use case name</td><td colspan="4">Product failure prediction for critical IT infrastructure</td></tr>
<tr><td>Application domain</td><td colspan="4">ICT</td></tr>
<tr><td>Deployment model</td><td colspan="4">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Building an AI solution to augment QA engineers</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Reduce the likelihood of releasing defective batches of hardware</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">A deep learning model to learn from a visual representation of the number of items that failed in a specific batch of hardware as well as the type of defect.</td></tr>
<tr><td>Complete description</td><td colspan="3">The hardware manufacturing company was using a few QA engineers to make subjective calls on whether or not a specific batch is good enough to be released into the market. The graphical representation of the shortfalls and defects was also done manually. This led to inconsistent labeling and many unsatisfied customers. To augment the QA engineers, a deep learning AI model was developed to do a more accurate and consistent labeling of which batches could be most defective and the major type of defects.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">QA engineers, Manufacturing line technicians, Technical sales</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Customer satisfaction index, cost of returned merchandise, time spent on QA</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">If the retraining model is compromised due to significant changes in the input data, the prediction model could generate incorrect outcomes and cost the hardware manufacturer serious loss.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Prediction Accuracy</td><td>Consistency of prediction compared to actual defect rates</td><td>Prediction accuracy should be 80% or more to ensure only the true-negative batches are inspected.</td></tr>
<tr><td>2</td><td>Time saved</td><td>Time for QA engineers to</td><td>The prediction model highlights</td></tr>
</table>

			inspect every batch	the most obvious defective batches and allows the QA engineers to spend time only on high-discretion tasks
	3	Customer Satisfaction	The number of returns from the manufacturer's customers	The satisfaction goes up when the number of defects is reduced upfront before the sales process.
AI features	Task(s)	Prediction		
	Method(s)	Deep Learning		
	Hardware	Private on premise severs		
	Topology	Bus and Hybrid		
	Terms and concepts used	Deep Learning, Dockers, Microservices		
Standardization opportunities/ requirements	Failure prediction models can improve global standards in manufacturing by reducing the waste of materials used and energy & water consumed.			
Challenges and issues	Challenges in identifying which deep learning model gives the best performance output, and challenges in indexing raw flat files into visualization images.			
Societal Concerns	Description	Address issues of sustainable manufacturing and high-value technical jobs		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.86.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/org anization	Link
1	Techni cal Paper	Support vector regression for warranty claim forecasting (Wu and Akbarov, 2011)				
2	Techni cal paper	Analysis of warranty claim data (Karim and Suzuki, 2005)				

A.87 Predicting relapse of a dialysis patient during treatment

A.87.1 General

ID	87
Use case name	Predicting relapse of a dialysis patient during treatment
Application domain	Healthcare
Deployment model	Cloud services

Status	In operation			
Scope	Build an AI solution to augment dialysis nurses			
Objective(s)	Use AI to predict if a patient may relapse during dialysis to reduce patient trauma			
Narrative	Short description (not more than 150 words)	A deep learning model to learn from historical and real-time parameters about a patient to identify the probability he or she may relapse during dialysis		
	Complete description	The private dialysis clinic was relying solely on the discretion of trained nurses to make a call whether or not a patient can get started for a dialysis session or should be taken to a hospital ahead of the treatment due to possible relapse. This created inconsistencies in the patient's experience and 10% of the patients would relapse and suffer trauma in the middle of their sessions. The deep learning model was able to provide a more consistent call about the likelihood of relapse, upon which the trained nurses could decide proactively for or against starting the dialysis session.		
Stakeholders	Dialysis nurses, Dialysis patients, Partner Hospitals			
Stakeholders' assets, values	Percentage of relapses as a total of all sessions, cost of incomplete sessions			
System's threats & vulnerabilities	If the equipment to identify the on-premise vital stats of the patient is incorrect or inaccurate, these would feed incorrect data into the model and the prediction output would also be inaccurate, leading to misguided decisions.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Prediction Accuracy	Consistency of prediction compared to actual relapse rates	Prediction accuracy should be 90% or more to ensure only the true-positive relapses are proactively sent to hospitals.
	2	Ease of Use	Ease of interpreting the inference of the models	The output of the model should be easily understandable for the nurses.
	3	Money Saved	The loss incurred for incomplete sessions	The proactive decisions to not commence high-relapse-chance patients' sessions to reduce the cost of incomplete sessions.
AI features	Task(s)	Prediction		
	Method(s)	Deep Learning		
	Hardware	Clinic computers and laptops		
	Topology	Hybrid		

	Terms and concepts used	Deep Learning, API
Standardization opportunities/ requirements	Prediction models can improve global quality of care for patients of kidney diseases or failure, and can allow the services to be more federated and standardized.	
Challenges and issues	Challenges in feature engineering the scores of datasets into a logical format that allows the prediction model to retrain without need for high compute.	
Societal Concerns	Description	Lack of reliable and accessible healthcare facilities
	SDGs to be achieved	Good health and well-being for people

4024 ### A.87.2 Data

Data characteristics	
Description	Dialysis appointment history data
Source	Dialysis company database
Type	Structured Data with Boolean, Numerical and Alphanumberical data
Volume (size)	
Velocity	Batch
Variety	Single
Variability (rate of change)	Dynamic, Weekly updated
Quality	High

4025 ### A.87.3 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Feature data into a form more suitable for prediction using Deep Learning	Featuring			
2	Training a deep learning model with training data	Training			
3	Test and reconcile outcomes of the model with actual results on the historical patient data.	Testing			
4	Correction and retraining of the model to improve prediction results.	Execution			

Specification of training data	

4026

4027 ## A.88 Improving the quality of online interaction

4028 ### A.88.1 General

ID	88
Use case name	Improving the quality of online interaction

<table>
<tr><td>Application domain</td><td colspan="4">Work & life</td></tr>
<tr><td>Deployment model</td><td colspan="4">Cloud services</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Build an AI solution to recommend relevant ideas to users in a chat interface</td></tr>
<tr><td>Objective(s)</td><td colspan="4">To improve the quality of conversations and translating online chat to meet ups</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">A recommendation engine operating live in a chat interface to help both users decide on the next steps they can take of high interest to both.</td></tr>
<tr><td>Complete description</td><td colspan="3">The dating platform prides itself on focusing on quality over quantity of matches made. Their online platform is assisted by downstream in person and sociological interventions to help newly met couples move towards a more meaningful relationship. The recommendation engine was to bring the sociological intervention more upstream in the engagement by infusing relevant recommendation of ideas of mutual interest in the chat interface.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Dating platform, Singles in Singapore</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Conversions of online to offline meet ups, Customer acquisition cost, customer life-cycle value.</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">If the recommendations made by the AI model are superficial, generic or inaccurate, the AI element could lead to a complete opposite of the desired outcome of bringing engagements online to in-person.</td></tr>
<tr><td rowspan="4">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Recommendation Accuracy</td><td>The number of recommendations accepted by both users.</td><td>If the recommendations are accepted by both users in the chat, then the engine performs better in future iterations.</td></tr>
<tr><td>2</td><td>Latency</td><td>Speed of recommendations appearing to the users</td><td>The latency should be low to bring the AI-enabled sociological interventions at the right time in the engagement.</td></tr>
<tr><td>3</td><td>Customer satisfaction</td><td>The number of users who manage to build a positive rapport after meeting online</td><td>The recommendation engine should improve the quality of online conversation.</td></tr>
<tr><td rowspan="3">AI features</td><td>Task(s)</td><td colspan="3">Recommendation</td></tr>
<tr><td>Method(s)</td><td colspan="3">Recommendation Engine, Natural Language Understanding</td></tr>
<tr><td>Hardware</td><td colspan="3">Users' individual internet connected devices</td></tr>
</table>

	Topology	Tree and Hybrid
	Terms and concepts used	Natural Language Understanding, Recommendation, API
Standardization opportunities/ requirements	Building a global corpus of language lexicon that if shared, can be used by AI systems to better identify online bullying, racism or other errant behavior.	
Challenges and issues	Translating sociological theories, customized to Singapore's context, and translating then into data labeling for the first step of NLU.	
Societal Concerns	Description	Improper use of online engagements that compromise on the culture of mutual respect and dignity.
	SDGs to be achieved	Good health and well-being for people

A.88.2 Data

Data characteristics	
Description	Google Search results and Quora Forum text
Source	Google, Quora
Type	Unstructured Text
Volume (size)	
Velocity	Real Time
Variety	Multiple
Variability (rate of change)	Static
Quality	Medium

A.88.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Research paper	Ultra-low fertility in Pacific Asia: Trends, causes and policy issues (Paulin Straughan, Angelique Chan, Gavin Jones, 2008)				http://books.google.com/books?hl=en&lr=&id=L_Z8AgAAQBAJ&oi=fnd&pg=PP1&dq=info:TfLQwqiHnWkJ:scholar.google.com&ots=AGo0gnIZME&sig=cvx7ZnE8tuYry0eCS1x5aLeOaKc

A.89 Instant triaging of wounds

A.89.1 General

ID	89
Use case name	Instant triaging of wounds
Application domain	Healthcare
Deployment model	Cloud services

<table>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Build an AI solution to augment triaging decisions of wound nurses</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Use AI to identify and classify the intensity of wounds</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">A computer vision model able to use RGB and IR wavelengths to measure the size, depth and intensity of a wound.</td></tr>
<tr><td>Complete description</td><td colspan="3">A wound nurse is the first line of medical attention when a patient comes to the hospital suffering from serious external wound injuries. The problem is more chronic in diabetic patients. The wound nurse has to spend time to view and decide how to triage the seriousness of the wound before sending the patient to the doctor. A CV model was built that can use a 2 megapixel mobile camera and off-the-shelf IR camera attachments to visualize wounds within seconds, to help the wound nurse make faster & more consistent triaging decisions.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Wound nurses, diabetes patients, hospitals</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Time and accuracy of triaging wounds</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Externalities like poor lighting or damages in the phone camera can ingest incorrect data into the CV model and output inaccurate visualisations.</td></tr>
<tr><td rowspan="4">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Visualisation accuracy</td><td>The visual representation of the wound is close to the actual condition</td><td>Unburden the nurse from the stress of accurately identifying the severity of wounds.</td></tr>
<tr><td>2</td><td>Ease of Use</td><td>Ease of interpreting the visual models of the wound</td><td>The visualisation of the wound should be easily understandable for the wound nurses.</td></tr>
<tr><td>3</td><td>Time saved</td><td>The time taken to view, assess and triage each patient.</td><td>The CV model would create a visualization of the wound within seconds which may otherwise take a wound nurse 10-30 minutes</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Knowledge processing & discovery</td></tr>
<tr><td>Method(s)</td><td colspan="3">Computer Vision</td></tr>
<tr><td>Hardware</td><td colspan="3">Mobile phones, hospital computers</td></tr>
<tr><td>Topology</td><td colspan="3">Bus</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Machine Learning, CNN, API</td></tr>
</table>

Standardization opportunities/ requirements	Using computer vision can make medical attention more globally accessible, in particular for poor and remote areas without compromising on the quality of care.	
Challenges and issues	Challenges in integrating RGB models and IR models into a single, interpretable visualization for the nurses.	
Societal Concerns	Description	Shortfalls in access to trained nurses and medical imaging technology.
	SDGs to be achieved	Good health and well-being for people

A.89.2 Data

Data characteristics	
Description	Images of wounds in RGB and IR spectrum
Source	
Type	Image data
Volume (size)	250GB
Velocity	Batch
Variety	Single
Variability (rate of change)	Static
Quality	High

A.90 Detection of fraudulent medical claims

A.90.1 General

ID	90	
Use case name	Detection of fraudulent medical claims	
Application domain	Maintenance & support	
Deployment model	On-premise systems	
Status	In operation	
Scope	Build a ML model to classify if a particular claim could be fraudulent	
Objective(s)	Upgrade from a only-human-interpretation to an ML-assisted fraud detection	
Narrative	Short description (not more than 150 words)	A machine learning model to identify true anomalies and trends of fraudulent claims customized to the source of fraud.
	Complete description	The Third Party Administrator (TPA) company has a very good visualization dashboard to eyeball trends by patient, by doctor and by condition of the medical claims submitted to the insurance companies the TPA serves. However, the identification of anomalies from the visual representation was still done on a subjective judgement basis. The ML model was developed to identify anomalies in claims that could have fraudulent activites by the patient, by the doctor or by both in collusion.
Stakeholders	TPA, Medical Insurance companies	
Stakeholders' assets, values	Percentage of true-positive fraudulent claims detected out of total set of claims.	

System's threats & vulnerabilities	If the features of the model are not updated every few years, the model may not be able to detect modes of fraud that have never ever been seen before.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Inference accuracy	Number of true-positives detectives vs false-positives	The better accuracy of the model, the more surgical would be the TPA's intervention in identifying & controlling fraud.
	2	Time of Inference	The latency in the model to retrain and generate new inferences	The latency in the model should be reasonable to allow the TPA to make faster action against fraudulent activities.
	3	Insurance Company Client's Satisfaction	The reduction in number of fraudulent claims that the insurance company client has to disburse money to.	The loss to the clients of the TPA ie insurance companies would reduce if more fraudulent claims are detected.
AI features	Task(s)	Inference		
	Method(s)	Machine Learning		
	Hardware	TPA's own devices and servers		
	Topology	Ring and Hybrid		
	Terms and concepts used	Machine Learning, Batch Retraining		
Standardization opportunities/ requirements	Machine learning models to detect frauds can be used globally to protect the integrity of public or private funds that are meant for essential services like medical care, housing, education or sanitation.			
Challenges and issues	The challenge was in building separate models for the each major sources of fraudulent claims.			
Societal Concerns	Description	Unintended or unlawful use of funds that are meant for essential services to people.		
	SDGs to be achieved	Sustainable cities and communities		

4038 ### A.90.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Research Paper	Big Data and Analytics in Healthcare: Introduction to the Special Issue			Information Systems	

		(Kankanhalli, A., Hahn, J., Tan, S. and Gao, G. 2016)			Frontier s	
2	Book	Actionable Intelligence: A Guide to Delivering Business Results with Big Data Fast! (2014)			Keith Carter	

A.91 Forecasting prices of commodities

A.91.1 General

ID	91			
Use case name	Forecasting prices of commodities			
Application domain	Fintech			
Deployment model	On-premise systems			
Status	In operation			
Scope	Build a neural network to forecast the price of base metal commodities			
Objective(s)	Use forecasted prices to interpret trading trends			
Narrative	Short description (not more than 150 words)	A trading company needed to improve the forecast accuracy of price points for specific commodities.		
	Complete description	The trading company has access to very good data to develop regression models. However, the model was insufficient to different impact of long term versus short term externalities. As such, a neural network was developed to ingest both structured market data as well as unstructured aggregate social media data to improve the inference and retraining ability to forecast prices.		
Stakeholders	Trading company, Manufacturers, Suppliers,			
Stakeholders' assets, values	Loss in spread of trades, Market research for clients			
System's threats & vulnerabilities	Possible tightening of aggregate data access policies of social media platforms which may require the neural network to be remodeled.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Forecast accuracy	Difference between forecasted and actual price	The use case depends on higher and timely accuracy of the price for necessary trades.
	2	Model latency	The latency for the model to retrain and output inferences	As trading sector gets more automated, it was important for the model to reduce latency.

	3	Money Saved	The loss incurred in poor or negative spreads	The trading company can use better forecasts to save their clients' money and reduce stress on cash flow.
AI features	Task(s)	Prediction		
	Method(s)	Neural Networks		
	Hardware	On cloud accessible by secure API		
	Topology	Star		
	Terms and concepts used	Neural Networks, NLP, API		
Standardization opportunities/ requirements	AI to predict the price and flow of goods can be used to hedge against unpredictable externalities such as civil unrest or territorial disputes when the accuracy of prices and amounts is critical to the mission.			
Challenges and issues	Challenge in modelling a neural network model that ingest large and wide array of data, while calibrating for variables that have short term versus long term impact.			
Societal Concerns	Description	Unpredictable flow of materials and commodities due to price shocks.		
	SDGs to be achieved	Reducing inequalities		

A.92 AI based dynamic routing SaaS

A.92.1 General

ID	92	
Use case name	AI based dynamic routing SaaS	
Application domain	Logistics	
Deployment model	Cloud services	
Status	In operation	
Scope	Build an ML model that dynamically corrects routes	
Objective(s)	Incorporate last minute human-driven factors into optimising delivery routes	
Narrative	Short description (not more than 150 words)	A machine learning model that dynamically corrects the delivery route and time to delivery.
	Complete description	The SaaS company used to provide routing service to delivery and e-commerce operators. However, the routes using heuristic models did not leave any scope of real-time changes to traffic, weather and driver behavior. As such, an ML model was required for the route to self-correct in real time to improve the satisfaction of the operators' clients.
Stakeholders	Delivery & logistics operators, Delivery personnel	
Stakeholders' assets, values	Speed of delivery, inaccurate routing, inaccurate estimate time of delivery	

System's threats & vulnerabilities	The real time correction of the model heavily relies on connectivity with the source of the real time data e.g. delivery personnel's internet device.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Fastest route	Ability of the model to correct to find the fastest route	Finding the fastest route with both static and dynamic parameters
	2	Latency	The time taken for the model to retrain and output new inference	The latency affects the delivery operators' ability to take action on dynamically advised routes
	3	Customer satisfaction	Satisfaction of the customers of the operators of delivery services	If the SaaS company's model performs well, the operators' delivery service improves and in turn customers receiving deliveries are more satisfied.
AI features	Task(s)	Planning		
	Method(s)	Machine Learning		
	Hardware	Personal internet connected devices		
	Topology	Bus and Mesh		
	Terms and concepts used	Machine Learning, API		
Standardization opportunities/ requirements	Dynamic routing using AI can improve the amount of fossil fuels consumed in supporting the growing e-commerce sector in urban areas around the world.			
Challenges and issues	Challenges in feature engineering static and dynamic variables, and over reliance on internet connectivity of the dynamic routing device.			
Societal Concerns	Description	Over utilization of resources and emittance of greenhouse gases to fulfil the trend of e-commerce.		
	SDGs to be achieved	Climate action		

A.92.2 Data

Data characteristics	
Description	Location Data and Delivery Reports
Source	Industry Partner
Type	Numerical
Volume (size)	
Velocity	Real Time
Variety	Multiple
Variability (rate of change)	Static

Quality	High

A.92.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1		Evaluating existing Vehicle Routing Problem solution effectiveness			
2		Evaluating use of 3rd party open source VRP software			
3		Evaluation of VRP performance in the face of data changes during execution of the solver.			
4		Use of constraint propagation in constraint solver with fine-grained algorithms.			

A.93 Non-intrusive detection of malware

A.93.1 General

ID	93	
Use case name	Non-intrusive detection of malware	
Application domain	Security	
Deployment model	Cloud services	
Status	In operation	
Scope	Build an AI solution that detects malware activities	
Objective(s)	User ML to flag out activities induced by malware without access to personal data on local devices	
Narrative	Short description (not more than 150 words)	A machine learning model that interprets phone activities like use of battery, data, location services or microphone to flag out possible malware in a local mobile device.
	Complete description	Off the shelf malware detection softwares have two limitations. One, that they are limited by the security features created by mobile device manufacturers, in particular for Android devices. And two, that they require the software to have access to device owners' personal data like name, phone number, IMEI or authentication credentials. As such, a non-intrusive ML based detection is a more sustainable and better accepted malware detection service which can flag out anomalies in the local device's generic activities like battery usage, data use, location services, camera or microphone.
Stakeholders	General users of internet connected devices	
Stakeholders'	Privacy of information, Cybersecurity	

assets, values				
System's threats & vulnerabilities	The model may require updates and tweaks if and when new applications get popular which impose new patterns of use of local device battery, location services, cameras and so on.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Detection accuracy	The number of true-positives detected versus false-positives	The usability of the malware detection depends on how accurately it detects true-positive malware.
	2	Ease of deployment	Ease of a local mobile device accessing this software.	The AI model of the detection software is on cloud accessed by API, making deployment easy in terms of compute capacity.
	3	Customer satisfaction	The number of customers safeguarded against malware	The more actual malware detection by this ML model, the more satisfied and reassured the users of the software would be.
AI features	Task(s)	Prediction		
	Method(s)	Machine Learning		
	Hardware	Personal mobile devices		
	Topology	Bus		
	Terms and concepts used	Machine Learning, API		
Standardization opportunities/ requirements	As one of the major victims of unsophisticated cyber-attacks is general public, using non-intrusive ML-based malware detection software has more wide ranging and affordable applications around the world.			
Challenges and issues	The model has limitations of the malware attacks are highly sophisticated and not easily detectable.			
Societal Concerns	Description	Disparate non-institutional sources of cyber attacks		
	SDGs to be achieved	Sustainable cities and communities		

A.93.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

1	Research Paper	Measuring the Declared SDK Versions and Their Consistency with API Calls in Android Apps (Daoyuan Wu, Ximing Liu, Jiayun Xu, David Lo, and Debin Gao, 2017)			Proceedings of the 12th International Conference on Wireless Algorithms, Systems, and Applications	
2	Research Paper	Classification of software behaviors for failure detection: a discriminative pattern mining approach (D Lo, H Cheng, J Han, SC Khoo, C Sun, 2009)			Proceedings of the 15th ACM SIGKDD international conference	
3	Research Paper	A generic framework for three-factor authentication: Preserving security and privacy in distributed systems (X Huang, Y Xiang, A Chonka, J Zhou, RH Deng, 2010)			IEEE Transactions on Parallel and Distributed Systems 22	

A.94 Predictive maintenance of public housing lifts

A.94.1 General

ID	94	
Use case name	Predictive maintenance of public housing lifts	
Application domain	Public sector	
Deployment model	Embedded systems	
Status	PoC	
Scope	Build an AI solution that can predict malfunction in a lift	
Objective(s)	Use RNN to predict possibility and type of malfunction in a lift	
Narrative	Short description (not more than 150 words)	An AI model that helps the facilities management company of public housing to move from a reactive to predictive maintenance of lifts.
	Complete description	The facilities management (FM) company currently adopts a reactive lift maintenance approach, where repair technician is notified to rectify specific faults, based on fault codes obtained from the lift monitoring device (LMD), after they happen. With additional telemetry data collected from ongoing installations of sensor box development and data acquisition systems, the FM company has begun exploring predictive approaches, using more than a year's accelerometer data and historical data on fault codes, for 600 lifts.
Stakeholders	FM company, residents in public housing	
Stakeholders' assets, values	Public housing lifts, repair technicians	

System's threats & vulnerabilities	The accuracy of the model is limited by quite a few hardware and sensors which may change with new sensors, new lifts or random spikes in activities by the residents.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Prediction accuracy	The number of true-positive failures predicted	The usability of the predictive maintenance system depends on how accurately it can predict true-positive failures.
	2	Time to failure	How soon can the model predict failure	The sooner the model can predict failure, the more proactive FM can be.
	3	Customer satisfaction	Reduction in number of lift breakdowns	The customer satisfaction would go up if the lift predicted to breakdown are proactively repaired.
AI features	Task(s)	Prediction		
	Method(s)	RNN		
	Hardware	Lifts, Servers		
	Topology	Hybrid		
	Terms and concepts used	Neural networks		
Standardization opportunities/ requirements	Predictive maintenance models are very useful when the stakes of the "down time" are very high for public infrastructure such as public transport which reduces the impact on climate.			
Challenges and issues	The model may at times predict false-positives which may lead to unnecessary deployment of repair & maintenance manpower			
Societal Concerns	Description	Disruptions to public due to breakdown of shared infrastructure		
	SDGs to be achieved	Climate action		

A.94.2 Data

Data characteristics	
Description	Lift maintenance log records
Source	Public sector
Type	Alphanumerical
Volume (size)	8 years of data across 10,000 lifts
Velocity	Real Time
Variety	Single
Variability	Batch

(rate of change)	
Quality	High

A.95 Tax Rules Updates and Classification

A.95.1 General

ID	95			
Use case name	Tax Rules Updates and Classification			
Application domain	Legal			
Deployment model	On-premise systems			
Status	PoC			
Scope	Build an AI solution that identify updates on tax laws and classify them			
Objective(s)	Use NLP to identify new tax laws from different countries and classify them			
Narrative	Short description (not more than 150 words)	An NLP model that helps a investment firm identify tax laws and trends that have an impact on their current and future portfolio		
	Complete description	The investment firm currently relies on tax analysts and external consultants to advice then on local, international and cross-border tax laws which have an impact on the income statements of their portfolio companies. This is a time consuming exercise and slow the firm's ability to react to changes. As such, the AI model shortens to process of identifying and classifying these changes such that the investment firm analysts can spend time instead on assessing the impact to their portfolio.		
Stakeholders	Investment firm			
Stakeholders' assets, values	Investment Risk, Return on Investment			
System's threats & vulnerabilities	The classification end of the model would need to be periodically updated if and when major nations make major tax or monetary policy changes.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accurate parsing	Identifying the correct terms and keywords in the new tax policy	The ability of the model to accurately parse unstructured text is essential to the next step of classification
	2	Accurate classification	The accuracy in classifying the type of tax rule or policies	Classifications in the user case include tax reclaims & refunds, indirect transfers and 7 other such policies types.

	3	Time to react	Reduction in time from change in tax rules to action by the investment firm	The NLP model for this use case helps to reduce the time to react to relevant tax rules or policies
AI features	Task(s)	Natural language processing		
	Method(s)	NLP		
	Hardware	Personal devices		
	Topology	Mesh and Hybrid		
	Terms and concepts used	NLP		
Standardization opportunities/ requirements	Providing fact-based transparency for tax laws applicable globally.			
Challenges and issues	The classes are pre-determined, and if these are changed, it will affect the ability of the model to re-classify.			
Societal Concerns	Description	Erratic changes in local and cross-border tax rules which have repercussions on economic growth.		
	SDGs to be achieved	Decent work and economic growth		

A.96 Ecosystems management from causal relation inference from observational data

A.96.1 General

ID	96	
Use case name	Ecosystems management from causal relation inference from observational data	
Application domain	Agriculture, Knowledge management, ICT	
Deployment model	Cloud Services, On-premise systems, Embedded systems, Hybrid	
Status	PoC	
Scope	Infer important latent variables to control whole ecosystem from database including human observation and sensor data.	
Objective(s)	To provide some suggestions for managing ecosystems and repeatedly improve it with the introduction of possibly latent variables and new data.	
Narrative	Short description (not more than 150 words)	We can find diverse relations between climate, animals and plants that infer ecologically consistent structure. To determine the factors that support a species niche is necessary to diversify the polyculture in ecological optimum, which is a complex entanglement that depends on environmental condition, associated biodiversity, farming option, etc. In our Synecoculture project, polyculture with ecological optimum requires a huge amount of information on biodiversity, interactions, and vegetation succession parameters, generally sparse possibly biased, open-ended, etc., because it relies on human observation. Still, it can

		bring useful information and intriguing insight on the management if powerful algorithmic analysis is combined with appropriate human evaluation.		
	Complete description	Please refer to a publication [1]: Foundation of CS-DC e-Laboratory: Open Systems Exploration for Ecosystems Leveraging (Masatoshi Funabashi, Peter Hanappe, Takashi Isozaki, AnneMarie Maes, Takahiro Sasaki, Luc Steels, and Kaoru Yoshida) — Section 2 Synecoculture — Section 3 Open Systems Data Analytics		
Stakeholders	Environment, Ecosystem			
Stakeholders' assets, values				
System's threats & vulnerabilities	Huge Database. Security threats.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Observed species	Maximize biodiversity	Management ecosystems
	2	yield	Maximize yield.	Management ecosystems
AI features	Task(s)	Infer important latent variables from database		
	Method(s)	Statistical causal discovery. [2]		
	Hardware	Sensors, processors		
	Topology			
	Terms and concepts used	Bayesian networks, causal discovery, conditional independence tests, structural equation modes		
Standardization opportunities/ requirements				
Challenges and issues				
Societal Concerns	Description			
	SDGs to be achieved	No poverty Zero Hanger Good health and well-being Clean water and Sanitation Decent work and economic growth Industry, innovation and infrastructure Reduce inequalities Responsible consumption and production Climate action Life on land Partnerships for the goals		

4062

A.96.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Article	Foundation of CS-DC e-Laboratory: Open Systems Exploration for Ecosystems Leveraging	Published	High	Masatoshi Funabashi, et al	https://hal.archives-ouvertes.fr/hal-01291104/document
2	Paper	A Robust Causal Discovery Algorithm against Faithfulness Violation	Published	High	Takashi Isozaki	https://www.jstage.jst.go.jp/article/imt/9/1/9_121/_pdf/-char/en
3	Article	Open Systems Exploration – An Example with Ecosystems Management	Published	High	Masatoshi Funabashi	https://hal.archives-ouvertes.fr/hal-01291125/document
4	Article (Web-site)	Creating abundant ecosystems through new agricultural methods Synecoculture	Published	Low	Sony CSL	https://www.sony.net/SonyInfo/sony_ai/synecoculture.html
5	Article (Web-site)	Synecoculture	Published	Low	Sony CSL	https://www.sonycsl.co.jp/tokyo/407/

A.97 System for Real-Time Earthquake Simulation with Data Assimilation

A.97.1 General

ID	97	
Use case name	System for Real-Time Earthquake Simulation with Data Assimilation	
Application domain	Social infrastructure	
Deployment model	On-premise systems	
Status	Prototype	
Scope	This system provides accurate information for evacuation in earthquake disaster.	
Objective(s)	The system conducts large-scale simulation of 3D Seismic Wave Propagation, and results are improved based on real-time data assimilation using observation and machine-learning.	
Narrative	Short description (not more than 150 words)	This system provides accurate information for evacuation in earthquake disaster. The system integrates Simulation, Data Analytics and Learning (S+D+L) on the BDEC System with h3-Open-BDEC which will be introduced at the

<table>
<tr>
<td></td>
<td></td>
<td>University of Tokyo in April 2021. It conducts large-scale simulation of 3D Seismic Wave Propagation, and results are improved based on real-time data assimilation using observation and machine-learning. Observations of seismic activities at more than 2,000 points in Japan are obtained by JDXnet developed by ERI/U.Tokyo through SINET in real-time manner. Construction of the detailed and accurate underground model is crucial for accurate simulations. Optimized underground model is also constructed by integration of (S+D+L). The BDEC system is 40+PF heterogeneous supercomputer system which includes Simulation Nodes for S, Data/Learning Nodes for D and L, and Integration Nodes. h3-Open-BDEC is a software infrastructure for application development towards integration of (S+D+L) supported by the Japanese Government (JSPS KAKENHI Kiban-S).</td>
</tr>
<tr>
<td></td>
<td>Complete description</td>
<td>

1 New Directions in Supercomputing

Majority of SCD/ITC/U.Tokyo's (Supercomputing Research Division, The University of Tokyo) supercomputer system users belong to the fields of CSE (Computational Science & Engineering), including engineering simulations (fluid dynamics, structural dynamics, and electromagnetics), earth sciences (atmosphere, ocean, solid earth, and earthquakes), and material sciences, as shown in the A pie chart of Fig.1, which shows usage rate of each research area on Oakleaf/Oakbridge-FX system (commercial version of the K computer) based on CPU hours in FY.2017. Recently, the number of users related to data science, machine learning, and artificial intelligence (AI) has been increasing, as shown in the B pie chart of Fig.1, which shows usage rate on Reedbush-H system with GPU's in FY.2018. Examples of new research topics are weather prediction by data assimilation, medical image recognition, and human genome analyses. Towards Society 5.0, a new type of method for solving scientific problems which integrates "Simulation (S)", "Data (D)" and "Learning (L)" (S+D+L) is emerging.

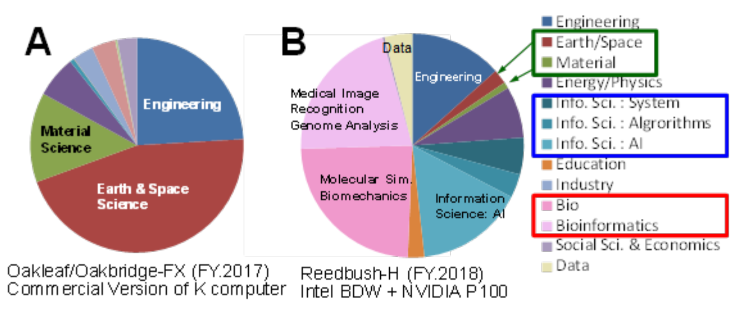


Fig.1 Research Area based on CPU Hours

2 BDEC: Big Data & Extreme Computing

The BDEC system (Big Data & Extreme Computing), which is scheduled to be introduced to SCD/ITC in April 2021, is a Hierarchical, Hybrid, Heterogeneous (h3) system. The BDEC is the platform for integration of "Simulation, Data and Learning (S+D+L)", and consists of computing nodes for

</td>
</tr>
</table>

computational science, those for data science/machine learning, and those for integration. The aggregated peak performance of the BDEC system is expected to be 40+ PFLOPS with aggregated memory bandwidth: 5.00+PB/sec, and it will comprise three types
of compute nodes, “Simulation Nodes (SIM, 90% of total resources)” for traditional supercomputing applications, “Data/Learning Nodes (DL, 5%)” for data and learning, and “Integration Nodes (INT, 5%)”. Architecture of SIM and INT must be same, while that of DL could be different. Some of the DL nodes will be connected to external resources (e.g. data storage, servers, sensor networks, and etc.) directly through an external network (e.g., SINET, Japan). DL and INT will share a fast file system (Fast File System, FFS, Capacity: 4+PB, Bandwidth: 2+TB/sec), while all nodes will share large-scale file system (Shared File System, SFS, 60+PB, 500+GB/sec).

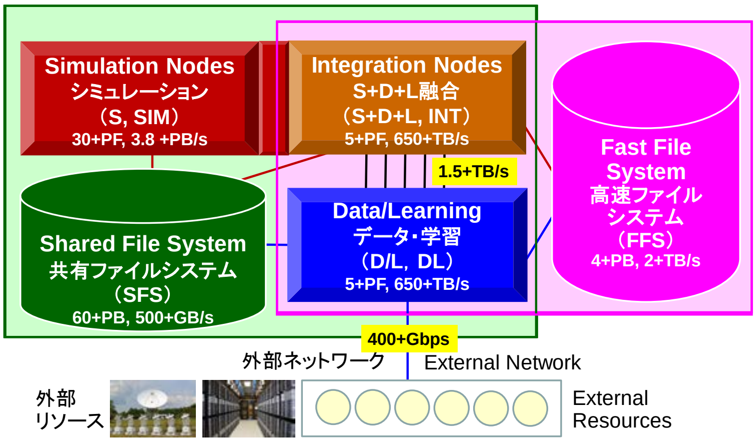


Fig.2 Overview of the BDEC System

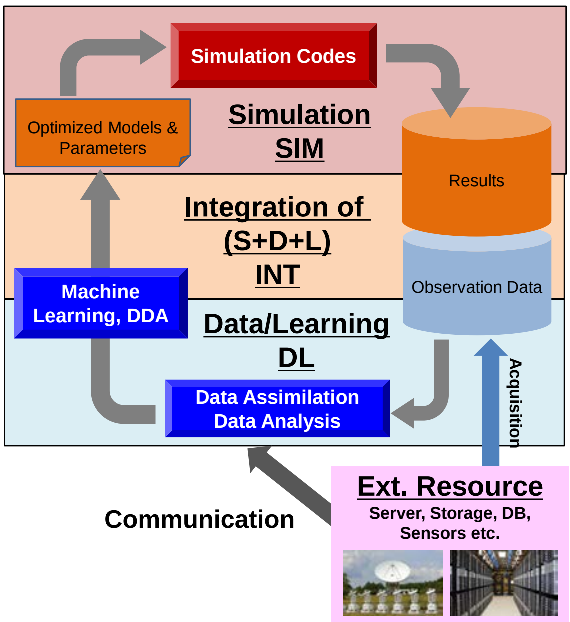


Fig.3 Integration of (S+D+L) by BDEC

3 h3-Open-BDEC: Innovative Software Platform for Integration of (S+D+L)

We develop an innovative software platform “h3-Open-BDEC” for integration of (S+D+L) and evaluate the effects of integration of (S+D+L) on the BDEC (Fig.4). The h3-Open-BDEC is designed for extracting the maximum performance of the supercomputers with minimum energy consumption focusing on (1) Innovative method for numerical analysis wit high-performance/high-reliability/power-saving based on the new principle of computing by adaptive precision, accuracy verification and automatic tuning, and (2) Hierarchical Data Driven Approach (hDDA) based on machine learning. This work will be supported by Japanese Government from FY.2019 to FY.2023 (JSPS Grant-in-Aid for Scientific Research (S), P.I.: Kengo Nakajima (ITC/U.Tokyo)).

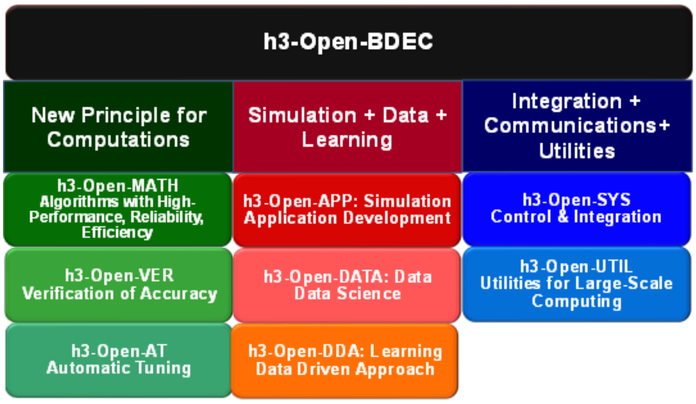


Fig.4 Overview of h3-Open-BDEC

In Data Driven Approach (DDA), technique of machine learning is introduced for predicting the results of simulations with different parameters. DDA generally requires a lot of simulations for generation of teaching data. We propose the hDDA, where simplified models for generating teaching data are constructed automatically by machine learning with Feature Detection, MOR (Model Order Reduction), UQ (Uncertainty Quantification), Sparse Modeling and AMR (Adaptive Mesh Refinement) (Fig.5).
The h3-Open-BDEC is the first innovative software platform to realize integration of (S+D+L) on supercomputers in the Exascale Era, where computational scientists can achieve such integration without supports by other experts. Source codes and documents are open to public for various kinds of computational environments. This integration by h3-Open-BDEC enables significant reduction of computations and power consumptions, compared to those by conventional simulations.
Idea of h3-Open-BDEC is extension of that of “ppOpen-HPC (https://github.com/Post-Peta-Crest/ppOpenHPC)” “ppOpen-HPC” is part of a (five+three)-year project (FY.2011–2015, FY.2016-2018) supported by JST-CREST and DFG-SPPEXA in Germany.

<table>
<tr>
<td></td>
<td></td>
<td>

Possible applications on the BDEC system with h3-Open-BDEC are combined simulations/data assimilations for climate/weather simulations and earthquake simulations, and real-time disaster simulations, such as flood, earthquake and tsunami (Fig.6).

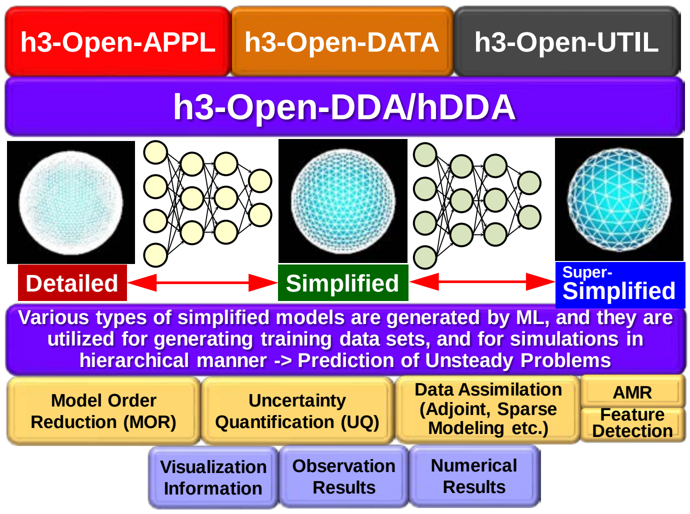


Fig.5 Hierarchical Data Driven Approach (hDDA)

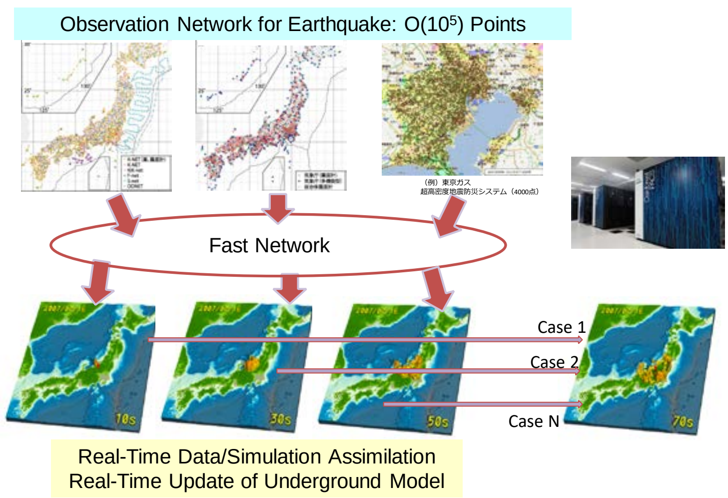


Fig.6 Real-Time Earthquake Simulation with Data Assimilation: Integration of (Simulation+Data+Learning) using h3-Open-BDEC and the BDEC System (c/o Prof. T. Furumura (ERI/U.Tokyo)

</td>
</tr>
</table>

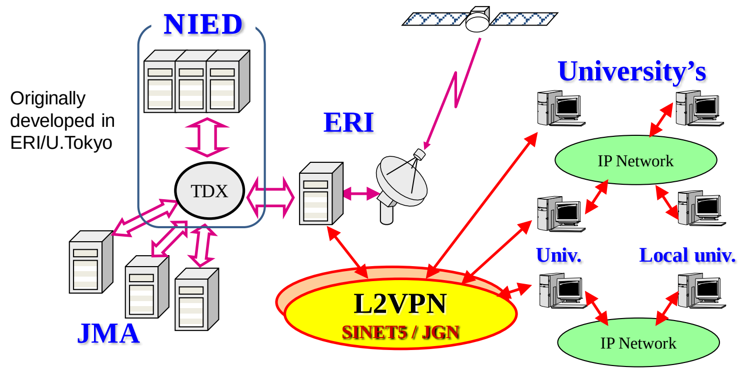
<table>
<tr><td></td><td></td><td colspan="3">

Fig.7 Real-Time Sharing of Seismic Observation is possible in Japan by JDXnet with SINET
4 System for Real-Time Earthquake Simulation with Data Assimilation
The system conducts large-scale simulation of 3D Seismic Wave Propagation, and results are improved based on real-time data assimilation using observation and machine-learning. Observations of seismic activities at more than 2,000 points in Japan are obtained by JDXnet developed by ERI/U.Tokyo through SINET (Fig.7) in real-time manner. Construction of the detailed and accurate underground model is crucial for accurate simulations. Optimized underground model is also constructed by integration of (S+D+L). </td></tr>
<tr><td>Stakeholders</td><td colspan="4">Information Technology Center (ITC), The University of Tokyo
Earthquake Research Institute (ERI), The University of Tokyo
National Institute of Informatics (NII)
National Research Institute for Earth Science and Disaster Resilience (NIED)
Japan Meteorological Agency (JMA), Meteorological Research Institute (MRI)
Local Governments in Japan
Transportation Companies (Railway, Highway)</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Disaster Prevention by Earthquakes</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Shutdown of Electricity by Earthquakes
Shutdown of Network by Earthquakes</td></tr>
<tr><td rowspan="4">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Time to Solution</td><td>Total computation time to achieve accurate solution after several simulations, data assimilations, and inferences</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
</table>

	Method(s)	Data Driven Approach, Hierarchical Data Driven Approach, Uncertainty Quantification (UQ), Model Order Reduction (MOR)
	Hardware	BDEC System in the University of Tokyo
	Topology	Results will be delivered through
	Terms and concepts used	
Standardization opportunities/ requirements		
Challenges and issues	Construction of reasonable and realistic underground model for simulation Real-time earthquake simulation with data assimilation	
Societal Concerns	Description	Earthquake Disasters
	SDGs to be achieved	Sustainable cities and communities

A.97.2 Data

Data characteristics	
Description	Seismic Observation
Source	2,000+ observation points in Japan operated by ERI, NIED and JMA
Type	Numbers
Volume (size)	O(10^2) GB/day
Velocity	100Hz
Variety	Deformation in 3-directions
Variability (rate of change)	Large deformation in earthquake events
Quality	Noise could be included, filtering methods have been already developed

A.97.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Overview of the Project	Innovative Methods for Scientific Computing in the Exascale Era by Integrations of (Simulation+Data+Learning)	On-Going	Software Infrastructure	JSPS	https://kaken.nii.ac.jp/en/grant/KAKENHI-PROJECT-19H05662/

A.98 Data compression with AI techniques

A.98.1 General

ID	98
Use case name	Data compression with AI techniques
Application domain	ICT
Deployment model	On-premise systems

Status	Prototype			
Scope	Data center/Supercomputing center			
Objective(s)	Fast data transfer via WAN			
Narrative	Short description (not more than 150 words)	Improving Data Compression with Deep Predictive Neural Network for Time Evolutional Data		
	Complete description	Scientific applications/simulations periodically generate huge intermediate data. Storing or transferring such a large scale of data is critical. Fast I/O is important for making this process faster. One of the approaches to achieve fast I/O is data compression. Our goal is to achieve a delta technique that can improve the performance of existing data compression algorithms for time evolutional intermediate data. In our approach, we compute the delta values from original data and data predicted by the deep predictive neural network. We pass these delta values through three phases which are preprocessing phase, partitioned entropy coding phase, and density-based spatial delta encoding phase. In our poster, we present how our predictive delta technique can leverage the time evolutional data to produce highly concentrated small values. We show the improvement in compression ratio when our technique, combined with existing compression algorithms, are applied on the intermediate data for different datasets.		
Stakeholders	High performance computing (HPC) communities			
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Fast data transfer	10x faster data transfer	(Under development)
AI features	Task(s)			
	Method(s)	Deep recurrent neural network/TensorFlow		
	Hardware	Tesla V100		
	Topology	A single node		
	Terms and concepts used	Neural network		
Standardization opportunities/ requirements	The software will be open-sourced			
Challenges and issues	More accurate prediction to data to be compressed			
Societal Concerns	Description			
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

4073 ### A.98.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Peer-reviewed publication	Rupak Roy, Kento Sato, Jian Guo, Jens Domke, Weikuan Yu, Takaki Hatsui and Yasumasa Joti, “Improving Data Compression with Deep Predictive Neural Network for Time Evolutional Data”, In Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis 2019 (SC19), Regular Poster, Denver, USA, Nov, 2019.	Published	Under development	RIKEN/ FSU	https://sc19.supercomputing.org/proceedings/tech_poster/tech_poster_pages/rpost181.html

4074

4075 ## A.99 Optimization of software configurations with AI techniques

4076 ### A.99.1 General

ID	99	
Use case name	Optimization of software configurations with AI techniques	
Application domain	ICT	
Deployment model	On-premise systems	
Status	Prototype	
Scope	Data center/Supercomputing center	
Objective(s)	Optimization of software configurations	
Narrative	Short description (not more than 150 words)	Optimizing Asynchronous Multi-level Checkpoint/Restart Configurations with Machine Learning
	Complete description	With the emergence of fast local storage, multi-level checkpointing (MLC) has become a common approach for efficient checkpointing. To utilize MLC efficiently, it is important to determine the optimal configuration for the checkpoint/restart (CR). There are mainly two approaches for determining the optimal configuration for CR, namely modeling and simulation approach. However, with MLC, CR becomes more complicated making the modeling approach inaccurate and the simulation approach though accurate, very slow. In this poster, we focus on optimizing the performance of CR by predicting the optimized checkpoint count and interval. This was achieved by combining the simulation approach with machine learning and neural network to leverage its accuracy without spending time on simulating different CR parameters. We demonstrate that our models can predict the optimized parameter values

		with minimal error when compared to the simulation approach.		
Stakeholders	High performance computing (HPC) communities			
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Optimization	20% improvement	(under development)
AI features	Task(s)			
	Method(s)	Three-layer neural network/TensorFlow		
	Hardware	Tesla V100		
	Topology	No interconnect (a single node)		
	Terms and concepts used	Neural network		
Standardization opportunities/ requirements	The software will be open-sourced			
Challenges and issues	More accurate prediction for the optimization			
Societal Concerns	Description			
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.99.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Peer-reviewed publication	Tonmoy Dey (Florida State University), Kento Sato (RIKEN Center for Computational Science (R-CCS)), Jian Guo (RIKEN Center for Computational Science (R-CCS)), Bogdan Nicolae (Argonne National Laboratory), Jens Domke (RIKEN Center for Computational Science (R-CCS)), Weikuan Yu (Florida State University), Franck Cappello (Argonne National Laboratory), Kathryn Mohror (Lawrence Livermore National Laboratory) "Optimizing Asynchronous Multi-Level Checkpoint/Restart Configurations with Machine	Published	Under development	RIKEN/FSU/ANL/LLNL	https://sc19.supercomputing.org/proceedings/tech_poster/tech_poster_pages/rpost180.html

		Learning", In Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis 2019 (SC19), Regular Poster, Denver, USA, Nov, 2019.				

A.100 Better human-computer interaction with advanced language models

A.100.1 General

ID	100			
Use case name	Better human-computer interaction with advanced language models			
Application domain	ICT			
Deployment model	Hybrid or other (please specify)			
Status	Prototype			
Scope	Human-computer interaction			
Objective(s)	Improve quality of human-computer interaction			
Narrative	Short description (not more than 150 words)	Better language models are crucial for improving the quality of human-computer interaction, for example tasks like question answering, summarization etc. We use large-scale compute systems to develop better language models by exploiting neural architecture search, large datasets and holistic evaluation framework.		
	Complete description	Natural language processing (NLP) technologies are crucial for interaction of social systems and artificial intelligence algorithms. AI models used in NLP are typically trained on large amounts of text (corpora) in order to make them "learn" language in general (language models), then these models are fine-tuned to particular down-stream tasks like question answering, paraphrasing, fake news detection etc. High quality models require large amount of data to train and thus large compute systems are needed. Additionally we employ neural architecture search and automated hyper-parameter optimization techniques to derive better models. Finally, we built a rigorous evaluation framework to explore how model architectures and hyper-parameters, including source corpora and preprocessing methods affect performance of the models on each of the down-stream tasks.		
Stakeholders	end users			
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives

	1	Linguistic benchmarks	Measured performance on a range of downstream tasks	
AI features	Task(s)	Natural language processing		
	Method(s)	Self-supervised pre-training, transfer learning, neural architecture search		
	Hardware	Large clusters of AI-capable devices for training		
	Topology	any		
	Terms and concepts used	Transfer learning, neural architecture search		
Standardization opportunities/ requirements	The software will be open-sourced			
Challenges and issues	High computational costs			
Societal Concerns	Description			
	SDGs to be achieved	Partnerships for the goals		

A.100.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Peer-reviewed paper	Subcharacter information in japanese embeddings: when is it worth it?	published	Prototyping some of the methods	Marzena Karpinska, Bofang Li, Anna Rogers, and Aleksandr Drozd.	https://www.aclweb.org/anthology/W18-2905/
2	Peer-reviewed paper	Subword-level composition functions for learning word embeddings.	Published	Prototyping some of the methods	Bofang Li, Aleksandr Drozd, Tao Liu, and Xiaoyong Du.	https://www.aclweb.org/anthology/W18-1205/

A.101 Accelerated acquisition of magnetic resonance images

A.101.1 General

ID	101	
Use case name	Accelerated acquisition of magnetic resonance images	
Application domain	Healthcare	
Deployment model	Hybrid or other (please specify)	
Status	Prototype	
Scope	Innovations in MRI image formation	
Objective(s)	Developing new approaches to MRI image formation aimed at reducing image acquisition time while maintaining the diagnostic image quality.	
Narrative	Short description (not more than 150 words)	Magnetic resonance imaging (MRI) is an essential instrument in precision diagnostics of neurological, oncological, musculoskeletal and other diseases. However, long acquisition times combined with the requirement for patient stillness pose a challenge for both patient and the radiology department, leading to high exam costs. Recent advances in sparse raw signal acquisition and specific image reconstruction show that it is possible to significantly reduce the acquisition time.
	Complete description	The excellent soft tissue contrast and flexibility of magnetic resonance imaging (MRI) makes it a very powerful diagnostic tool for a wide range of disorders, including neurological, musculoskeletal, and oncological diseases. However, the long acquisition time in the MRI machine, which can easily exceed 30 minutes, leads to low patient throughput, problems with patient comfort and compliance, artifacts from patient motion, and high exam costs. Increasing imaging speed has been a major ongoing research goal since the advent of the MRI. By combining both hardware developments (such as improved magnetic field gradients) and software advances (such as new pulse sequences), it has been possible to significantly reduce the image acquisition times. One noteworthy development in this context is parallel imaging, introduced in the 1990s, which allows multiple data points to be sampled simultaneously, rather than in a traditional sequential order [1, 2]. Compressed sensing [3, 4] techniques speed up the MR acquisition by acquiring less measurement data than was previously required to reconstruct diagnostic quality images. Artifacts that are introduced by the violation of the Nyquist-Shannon sampling theorem can be eliminated in the course of image reconstruction. This can be achieved by incorporating additional a priori knowledge during the image reconstruction process. The last two years have seen the rapid development of machine learning approaches for MR image reconstruction, which hold great promise for further acceleration of MR image acquisition [5, 6, 7]. To speed up the algorithm development, public datasets are being provided to the

<table>
<tr><td></td><td></td><td colspan="3">research community. For example, the fastMRI challenge [8] introduced standardized evaluation criteria and freely-accessible datasets to help the community make rapid advances in state-of-the-art MR image reconstruction.
In machine learning based approaches, the reconstruction function is learned from the dataset of the input-output pairs of samples drawn from a population. Such techniques also leverage previous exam data to learn the spatial structure of anatomy and typical image artifacts caused by under-sampling. These attributes allow CNN-based methods to reconstruct highly under-sampled data at higher fidelity than CS schemes in certain cases [9].
The developed reconstruction algorithms may be deployed either directly into the scanner console, or on the dedicated reconstruction workstation or even on the cloud, depending on the computational requirements.
The main challenge in clinical application of such deep learning based image formation algorithms is to guarantee safety. For any device it is necessary to guarantee that the AI system is not leading to diagnostic errors by removing or introducing pathologies or other image features. It is also necessary to guarantee image quality for all possible combinations of MRI sequence parameters, anatomical areas, patient cohorts, or to be very conservative in defining the limits of applicability.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Radiology departments, MRI vendors</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Safety/robustness, patient throughput, acquisition cost, scanner utilization</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Increased image reconstruction time, image quality reduction/artifacts</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Image quality</td><td>The quality of the image obtained with accelerated technology should be enough for making a diagnosis</td><td>Image quality</td></tr>
<tr><td>2</td><td>Reconstruction time</td><td>Time for reconstructing the image from the raw signal</td><td>Acquisition time</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Inference</td></tr>
<tr><td>Method(s)</td><td colspan="3">Neural networks for image generation</td></tr>
<tr><td>Hardware</td><td colspan="3">GPU</td></tr>
<tr><td>Topology</td><td colspan="3">Depends on the deployment. May be either edge (on the scanner), or on the dedicated HW, or on the cloud.</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3"></td></tr>
</table>

Standardization opportunities/ requirements	1) Quality acceptance criteria: it is necessary to guarantee that AI system is not leading to diagnostic errors by removing or introducing pathologies or other image features	
Challenges and issues	1) Image quality measurements shall correlate with the diagnostic value – extensive clinical validation and A/B testing is needed, but it is expensive 2) It is necessary to guarantee quality for all possible combinations of MRI sequence parameters, anatomical areas, patient cohorts, or to be very conservative in defining the limits of applicability	
Societal Concerns	Description	(If safety/quality is guaranteed), MRI imaging will be used more often, more images will be generated which will increase radiologists' workloads. Development of AI-assisted image interpretation tools will be very much demanded.
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.101.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	Daniel K Sodickson and Warren J Manning. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. Magnetic resonance in medicine, 38(4), 1997.	published			
2	Publication	Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. SENSE: sensitivity encoding for fast MRI. Magnetic resonance in medicine, 42(5), 1999.	published			
3	Publication	Emmanuel J Cand`es, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory, 52(2), 2006.	published			
4	Publication	Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging. Magnetic Resonance in Medicine, 58(6), 2007.	published			
5	Publication	Kerstin Hammernik, Florian Knoll, Daniel K Sodickson, and Thomas Pock. Learning a Variational Model for Compressed Sensing MRI Reconstruction. In Magnetic Resonance in Medicine (ISMRM), 2016.	published			
6	Publication	Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. In	published			

		IEEE International Symposium on Biomedical Imaging (ISBI), 2016.				
7	Publication	Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P. Recht, Daniel K. Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine, 2018.	published			
8	Publication	Zbontar, Jure, et al. "fastmri: An open dataset and benchmarks for accelerated mri." arXiv preprint arXiv:1811.08839 (2018).		Use case taken from this publication		
9	Preprint	Sandino, Christopher M., et al. "Deep convolutional neural networks for accelerated dynamic magnetic resonance imaging." preprint (2017).				

A.102 AI Adaptive Learning Platform for Personalized Learning

A.102.1 General

ID	102	
Use case name	AI Adaptive Learning Platform for Personalized Learning	
Application domain	Education	
Deployment model	Cloud services	
Status	In operation	
Scope	2,5 million users (09.19) [1]	
Objective(s)	Open access, Interactive tasks, Personalization, User-generated content, Learning graph. Summarizing - equal access to high-quality education [2]	
Narrative	Short description (not more than 150 words)	Adaptive learning platform (AiEd platform) [3] is an elearning platform and course-builder which uses AI for forming adaptive learning paths [4]
	Complete description	Adaptive learning platform is a cloud-based platform designed to create and distribute interactive educational content, enhanced by various types of automatically graded assignments with a real-time feedback. The platform is suitable for any kind of e-learning activity, from private on-campus classes to MOOCs (massive open online courses). The platform is designed keeping the needs of computer science education in mind. The platform aims to apply data mining techniques to make education more efficient and to improve the way people learn and teach. Adaptive [5] and personalized learning are one of the key priorities of our platform. [6]
Stakeholders	Students, teachers (content providers), third-party services (via xAPI), academic researchers (sets of eduDATA)	
Stakeholders' assets, values	Personal data concerning interests and preferences, safety, privacy (learners); reputation, trustworthiness, high quality content (teachers, content providers); safety (third-party actors)	

System's threats & vulnerabilities	Verification of new content			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Performance	Increasing educational results by personalized learning process	Personalization
	2	Variability	Educational content makes up learning graph. AI Adaptive learning engine based on learning graph allows automatically create huge number of education programs, acceptable for everyone.	Learning graph
AI features	Task(s)	Optimization		
	Method(s)	Recommendation-based approach which uses: Item Response theory, the ELO rating system [8]		
	Hardware	none, cloud-based solution is used		
	Topology			
	Terms and concepts used	Adaptive learning is made in the form of a recommendation system that advises the user which lesson they should learn next, depending on prior actions. [8]		
Standardization opportunities/ requirements	After repeated training, adaptive learning system will be highly efficient			
Challenges and issues	Edstories (micro-learning video stories) should be included to satisfy the pedagogical model of movement-based learning			
Societal Concerns	Description	The system should be integrated into secondary and tertiary school-systems that still face legal boundaries and limitations for scaling		
	SDGs to be achieved	Quality education		

A.102.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

1	Service tutorial		Valid	definition, basic information	Stepik	https://support.stepik.org/hc/ru/articles/360000172234
2	Service tutorial		Valid	definition, basic information	Stepik	https://support.stepik.org/hc/en-us/articles/360000172234-What-is-Stepik
3	Conference paper		Valid	definition, basic information	Springer	https://link.springer.com/book/10.1007/978-3-319-59044-8
4	Stepik business presentation		Valid	Stepik overview	Stepik	https://te-st.ru/wp-content/uploads/2016/10/Stepik.pdf
5	Conference paper		Valid	definition, basic information	Springer	https://link.springer.com/chapter/10.1007/978-3-030-23207-8_33
6	Service tutorial		Valid	Stepik functionality tutorial	Stepik	https://support.stepik.org/hc/en-us/articles/360000172234-What-is-Stepik
7	Service tutorial		Valid	Stepik functionality tutorial	Stepik	https://support.stepik.org/hc/en-us/articles/360000173074-Points-and-certificates
8	Stepik official blog on Habr		Valid	Stepik adaptive learning concept	Stepik	https://habr.com/ru/company/stepic/blog/325206/

A.103 AI based text to speech services with personal voices for speech impaired people

A.103.1 General

ID	103	
Use case name	AI based text to speech services with personal voices for speech impaired people	
Application domain	Healthcare	
Deployment model	On-premise systems	
Status	Prototype	
Scope	All people who has some sort of speech impairments including but not limited to three basic types: articulation disorders, fluency disorders, and voice disorders.	
Objective(s)	People with speech impairments will be fully integrated into social processes without communication restrictions.	
Narrative	Short description (not more than 150 words)	Communication with other people can be difficult for those who have speech disorders. This seriously complicates communication with the surrounding domestic processes and the involvement of a person in society. A personal wearable device is capable of online synthesizing voice over text or correcting distorted speech. The voice can be fully

		synthesized with individually selected tone, timbre and pronunciation style settings.		
	Complete description	Communication with other people can be difficult for those who have speech disorders. This seriously complicates communication with the surrounding domestic processes and the involvement of a person in society. A personal wearable device is capable of online synthesizing voice over text or correcting distorted speech. The voice can be fully synthesized with individually selected tone, timbre and pronunciation style settings. Moreover, the voice can be a copy of the voice of the owner, which he/she retained. The device itself can be implemented as a bracelet or a special medical device. Implementation as software for a smartphone, laptop, etc. is also possible.		
Stakeholders	People with speech impairments			
Stakeholders' assets, values	Social integration processes of people with speech impairments			
System's threats & vulnerabilities	Quality of voices			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	mos	Mean opinion score	Score from stakeholders whom use new services\hardware	Lowering communication barriers
	use_scale	Scale of use	Percentage of stakeholders using the service\hardware to the total number of stakeholders.	Service\hardware distribution scale
AI features	Task(s)	Text to speech		
	Method(s)	Deep learning		
	Hardware	Cloud hardware, wearable devices		
	Topology	Tacotron2, LPCNet		
	Terms and concepts used	Text to speech, deep learning, Tacotron2, LPCNet		
Standardization opportunities/ requirements	Minimum hardware requirements for wearable devices. Voices package/format standardization.			
Challenges and issues	1. Minimization of source records to create a synthesized voice from tens of hours to several tens of minutes 2. Hardware requirements for voices based on neural networks should be reduced to the level available on wearable devices. 3. The ability to control intonations, speech style should be expanded for use in a natural dialogue between people.			
Societal	Description	We don't see any societal concerns if it is used		

Concerns	SDGs to be achieved	

A.104 AI Decryption of Magnetograms

A.104.1 General

ID	104			
Use case name	AI Decryption of Magnetograms			
Application domain	Manufacturing, Gas & Oil			
Deployment model	Client and server systems			
Status	In operation			
Scope	Oil and gas transportation. AI solution to quickly identify defects during the quality assurance process on field pipeline			
Objective(s)	Detection of internal defects (pits, ulcers, etc.) Detection of structural elements (welds, bends, etc.)			
Narrative	Short description (not more than 150 words)	A solution has been developed that allows for the detection of internal defects and structural elements		
	Complete description	In the territory of the Russian Federation, there are tens of thousands of kilometers of small diameter production pipelines under varying degrees of condition facing varying numbers of internal defects (pits, ulcers, etc.) and structural elements (welds, bends, etc.) There are in-tube flaw detectors that allow the signal from the magnetometer sensors to be read. These robots are not widely used due to the speed of data interpretation. Automation of the recognition of structural elements and defects will reduce the pipeline diagnostics process by at least 160 times		
Stakeholders	Manufacturer			
Stakeholders' assets, values	Decision speed			
System's threats & vulnerabilities	Condition of the flaw detector			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Coverage for welds detection	Detection accuracy's ideal target is 95%	Improved accuracy
	2	Coverage for defects detection	Detection accuracy's ideal target is 100% with 50% defect depth. Target is 90% with 30% defect depth.	Improved accuracy
AI features	Task(s)	Recognition		

	Method(s)	Machine learning, classic computer vision
	Hardware	Flaw detector
	Topology	Trees, Random forest
	Terms and concepts used	Machine learning, computer vision, training, training data set
Standardization opportunities/ requirements		
Challenges and issues	To achieve high level accuracy recognizing defects and welds; To reduce the processing time of magnetograms.	
Societal Concerns	Description	Minimizing the risk of environmental disasters associated with oil spills
	SDGs to be achieved	Zero pipeline breakthroughs per year

A.104.2 Data

Data characteristics	
Description	Data from 64 robot sensors
Source	Flaw detector
Type	Raw data, transformed into .csv
Volume (size)	60 Gb
Velocity	Batch
Variety	Different source
Variability (rate of change)	Static
Quality	Low

A.104.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model with training data set	Sample raw data set is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		Meeting KPI requirements (e.g. accuracy of detection welds is 0.95) is the "success" condition
3	Execution	Detection	Completion detection	The trained model has been evaluated as deployable	

4	Retraining	Retrain a model with training data set	New examples of magnetogr ams		

A.104.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample data set is ready	Data preparation	Manufacture r	Transform raw data into .csv.	The software for data preparation has to be provided by the AI solution provider.
2	Completion of Step 1	Training data set creation	Manufacture r	Create training data set by manual marking of magnetograms for further analysis the output of Step 1 with different classes and balancing	
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	

Specification of training data	

A.104.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Model evaluation	Manufacturer	Compare the result of model work with that of human inspection	

Input of evaluation	
Output of evaluation	

A.104.6 Execution

Scenario name	Execution

Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Acquisition of raw magnetogram	Detection	Manufacturer	Detection of defects and welds	A trained model should convey the results of the work to the manufacturer.

Input of Execution	
Output of Execution	

A.104.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Getting new data	Data preparation	Manufacturer	Transform raw data into .csv.	1
2	Completion of Step 1	Training data set creation	Manufacturer	Create training data set by manual marking of magnetograms for further analysis the output of Step 1 with different classes and balancing	2
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	3

Specification of retraining data	Retraining data set has to include data from different robot types

A.104.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure		In operation		Gazprom neft	

A.105 AI Platform for Chest CT-Scan Analysis (early stage lung cancer detection)

A.105.1 General

ID	105	
Use case name	AI Platform for Chest CT-Scan Analysis (early stage lung cancer detection)	
Application domain	Healthcare	
Deployment model	Cloud services	
Status	In operation	
Scope	Detecting malignant neoplasms (lungs) on chest CT-scans	
Objective(s)	To facilitate early stage oncology chest CT-scans through the application of the Botkin.AI platform based on artificial intelligence	
Narrative	Short description (not more than 150 words)	"Botkin.AI" is a software platform for the diagnosis and assessment of pathology risks using artificial intelligence technologies. The product supports radiologists and oncologists, facilitating the analysis and recognition of diagnostic images of CT-scans, digital X-rays and mammography. The project aims to reduce costs and improve diagnostic accuracy, while detecting pathologies at early stages.
	Complete description	Botkin.AI implements its own-patented technology to create a digital model of the patient. This allows for state-of-the-art results derived from the company's algorithms, confirmed by scientific publications. The Botkin.AI platform core goals are improved oncology detection at early stages and prioritization of patient flow. The company provides its own developed DICOM viewer. The platform may be integrated into any type of PACS/central archive of medical images such as SaaS solutions, or as part of a medical institution's closed infrastructure. The company is ready to provide customizable integration options to fit the needs of varying customers. "Hybrid Intelligence" technology allows for the combination of the AI platform's sensitivity with the specificity of a skilled radiologist. "Botkin WorkFlow" technology may also be used to manage the flow of different radiological studies. Botkin.AI plans to demonstrate a platform which increases the efficiency and effectiveness of radiological analysis. This product addresses two main medical issues: an undersupply of radiologists in the workforce; and missed malignant neoplasms on chest CT-scans. With the introduction of this technology, thousands of lives could be saved via improved early stage oncology.
Stakeholders	Healthcare authorities	
Stakeholders' assets, values	Reputation, saved lives, cost savings	
System's threats & vulnerabilities	Loss of trust	

Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Accuracy	93% detection rate of malignant neoplasm on chest CT-scans (AUC) for Botkin.AI	Improved accuracy
	2	Speed	From 4-10 min (depending on Internet speed)	Improved speed
AI features	Task(s)	Recognition		
	Method(s)	Deep learning		
	Hardware			
	Topology			
	Terms and concepts used	Deep learning, imagification, neural network, training, training data set		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: Achieving a higher confirmed level than accredited radiologists in the detection of lung cancer			
Societal Concerns	Description			
	SDGs to be achieved	Good health and well-being for people		

A.105.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure				Intellogic LLC	https://botkin.ai/wp-content/uploads/2019/08/Botkin_AI_Brochure_ENG_PRINT_curves.pdf
2	Scientific article				R&D Director at Intellogic LLC	https://openreview.net/forum?id=rkexLAH0FE

A.106 AI Virtual Assistant for Customer Support and Service

A.106.1 General

ID	106
Use case name	AI Virtual Assistant for Customer Support and Service
Application domain	Maintenance & support
Deployment	Embedded systems

<table>
<tr><td>model</td><td colspan="4"></td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Customer support service, product and service consulting
Limitations - support for dialogs exclusively within MTS products
Target audience - b2b, b2c clients of MTS Russia</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Optimization of company resources for support and customer service by automating the customer service process. As a result of the implementation of the system, the company was able to cover a greater volume of customer requests without needing to increase its staff of operators. This allowed the prevention of an increase in the company's operating expenses.</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">The system automatically answers customer questions in the application and on the company website. At peak, service automation reaches 85%.</td></tr>
<tr><td>Complete description</td><td colspan="3">Chatbot assists the client in the selection of tariffs and services, and advises on the financial condition of the account. Chatbot promotes new products without the need for an operator. The client can ask a question in free form; the system will understand the request. If necessary, the system may ask additional questions before delivering its answer to the client.
Chatbot is integrated with internal billing systems, CRM, with a product catalog and many other key services of the company. This allows each client to be provided with an individualized service. If chatbot is unable to help the client, or if the service procedure requires an operator, the dialogue is transferred to the operator. Currently, chatbot serves more than 1 million requests per month, working 24/7 to serve customers in all regions of Russia.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Customer Service Department</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Customer service department - maintaining / increasing customer loyalty, saving resources</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Information security, communication secrecy and the safety of personal data</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Automation</td><td>Solving a customer issue with a chatbot without operator intervention</td><td>Optimization of customer service costs</td></tr>
<tr><td>2</td><td>Quality</td><td>Customer satisfaction rating</td><td>Ensuring high customer loyalty to the MTS brand</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Optimization, natural language understanding, dialogue management</td></tr>
<tr><td>Method(s)</td><td colspan="3">Deep learning, NLP</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Natural language processing, chatbot, dialogue systems</td></tr>
<tr><td>Standardization</td><td colspan="4"></td></tr>
</table>

opportunities/ requirements		
Challenges and issues	- The readiness of external systems' API for integration with the bot platform - Biased customer attitudes towards chatbots	
Societal Concerns	Description	
	SDGs to be achieved	Affordable and clean energy

A.106.2 Data

Data characteristics	
Description	
Source	Customer profiles and a history of questions
Type	Text, voice
Volume (size)	Millions of hits (historical data)
Velocity	In real time
Variety	Collected datasets
Variability (rate of change)	The system is updated daily with new scenarios.
Quality	High

A.106.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Initiate change	Studying the platform for new topics	The appearance of a new company product and the need to train the chatbot on it	The customer provides a logic diagram of the new scenario; additionally, other examples of customer requests for training	
2	Training	Platform training	Receiving data for training	The model is trained on labeled data	
3	Integration	Receiving data from external systems	If necessary, integration into a specific scenario; the system has an API	Functions are implemented based on API methods	
4	Testing	Testing the operation of the API and scripts	Readiness of the previous stages.	Problem solving; and refining the script process without error	

4125 ### A.106.4 Initiation

Scenario name	Initiation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Request from customer	Elaboration of logic	Stakeholders	The process of creating script logic	
2	Step 1	Data markup	Developers	Markup data for model training	

4126 ### A.106.5 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Initiation process completed	Training data preparation	Developer	Preparation of tagged data for training models	
2	Step 1	Model training	AI	The model is trained on the tagged data, taking into account the logical construction of the script	
3	Step 2	Verify model performance	Developer	The model is checked for accuracy	

Specification of training data	

4127

4128 ### A.106.6 Integration

Scenario name	Integration				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of model checks for correct operation	Preparation of a list of systems for integration with the bot	Developers / Managers	Creating a list of systems with which the bot will interact	
2	Step 1	Exploring API documentation	Developers	Studying the documentation; testing methods of system interaction	
3	Step 2	Integration with external and internal systems	Integrators	The bot is connected to all necessary external and internal	

				systems for interaction	

Input of evaluation	
Output of evaluation	

A.106.7 Testing

Scenario name	Testing				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of the integration process	Formation of test cases for system verification	Stakeholders	A list of test cases for the bot is formed	
2	Step 1	Test Case Run	Developers	Test cases are run for all scenarios, taking into account connected integrations	

A.106.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press release				MTS	https://moskva.mts.ru/about/media-centr/soobshheniya-kompanii/novosti-mts-v-rossii-i-mire/2019-06-03/mts-nachala-prodavat-robotov
2	Advertising				MTS	https://www.youtube.com/watch?v=flMkRsV8Gvo

A.107 AI-based design of pharmacologically relevant targets with target properties

A.107.1 General

ID	107
Use case name	AI-based design of pharmacologically relevant targets with target properties
Application domain	Healthcare
Deployment model	On-premise systems
Status	Prototype
Scope	AI-based engineering of G protein-coupled receptors with enhanced stability

Objective(s)	Given: protein template in a form of a protein sequence or structure; target properties Predict: protein sequence that satisfies target properties and has minimal differences from the given template	
Narrative	Short description (not more than 150 words)	Molecular design is one of the most important and rapidly developing fields in biotechnology, where the protein engineering plays a significant role in major topics. With an accumulation of biophysical data, AI-based approaches become beneficial in protein design for biotechnology. A particular case is to design stable forms of pharmacological targets, such as G protein-coupled receptors (GPCRs). Malfunctions of these receptors typically lead to various diseases: neurodegenerative, oncological and cardiovascular diseases, asthma, depression, obesity, drug dependence, etc. GPCR receptors are one of the main targets for pharmacological companies, and about 1/3 of all drugs produced in the world are oriented on GPCRs. Obtaining the spatial structure of a single receptor is an extremely difficult and resource-intensive task. We developed an innovative AI-based digital platform for GPCR design, which allowed for a technological breakthrough in obtaining spatial structures of GPCR [1, 2] for the rational development of a new generation drugs.
	Complete description	Molecular design is one of the most important and rapidly development field in biotechnology. Optogenetics tools in neurobiology, fluorescent proteins in cellular biology, sequencing nanopores in molecular biology, drug discovery in medicinal chemistry and many other examples in modern biotechnology are based on protein engineering. With an accumulation of biophysical data, AI-based approaches become beneficial in protein design for biotechnology. Typically, protein design starts with a template—protein from a human or any other living organism—and with a target property, for example, protein stability or spectral shift. Then, the goal is to modify the template to obtain engineered protein with the target property. A particular case is to design stable forms of pharmacological targets, such as G protein-coupled receptors (GPCRs). Malfunctions of these receptors typically lead to various diseases: neurodegenerative, oncological and cardiovascular diseases, asthma, depression, obesity, drug dependence, etc. GPCR receptors are one of the main targets for pharmacological companies, and about 1/3 of all drugs produced in the world are oriented on GPCRs. For the development of more efficient and safer drugs, as well as personalized drugs that take into account the characteristics of the human genome (mutation), it is necessary to understand how GPCRs work on structural level. Obtaining the spatial structure of a single receptor is an extremely difficult and resource-intensive task. We developed an innovative AI-based digital platform for GPCR design, which allowed for a technological breakthrough in obtaining spatial structures of GPCR [1, 2]. Thanks to the developed technology for the last few years, spatial

		structures of ~10 GPCR receptors were determined, i.e. >10% of all receptors with a known spatial structure to date. These include relevant pharmacological targets, such as the human cannabinoid receptor [3], the human serotonin receptor [1,6], the human prostaglandin receptor [4], the «frizzled» human receptor [5], the human adenosine receptor [7], the human cysteine receptors of types one [8] and type two [9]. Structural analysis of each new receptor has opened up opportunities for the rational development of a new generation drugs.		
Stakeholders	Pharmacy companies, biomedical researchers			
Stakeholders' assets, values	Competitiveness, reputation, trustworthiness, safety			
System's threats & vulnerabilities	Different sources of bias, incorrect AI system use			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	CompoMug – structure-based learning	Four modules: knowledge-based, sequence-based, structure-based, and machine-learning-based	Composition of a list of the candidate point mutations, which can improve the stability of a GPCR
	2	Machine learning	Alanine scanning mutagenesis data for GPCRs	Training benchmark: Definition of (non-)stabilizing point mutations
AI features	Task(s)	Inference		
	Method(s)			
	Hardware	HPC		
	Topology			
	Terms and concepts used	GRCR, structure-based recognition, classifier, mutagenesis, mutations, drug development		
Standardization opportunities/ requirements				
Challenges and issues	Biophysical data is typically very noisy, and the results critically depend on the used experimental assay and initial conditions. Therefore, the training data must be carefully processed with expert knowledge. Consequently, the derived prediction models must rigorously analyzed for robustness, domain applicability, and generalizing power			
Societal Concerns	Description	Discovery of more efficient, safer and personalized drugs		
	SDGs to be achieved	Good health and well-being for people		

A.107.2 References

References

No.	Type	Reference	Status	Impact on use case	Originator /organization	Link
1	Paper	P Popov et al., “Computational design of thermostabilizing point mutations for G protein-coupled receptors”, eLife, 2018	Published	High		https://elifesciences.org/articles/34729
2	Paper	P Popov et al., “Computational design for thermostabilization of GPCRs”, Current Opinion in Structural Biology, 2019	Published			
3	Paper	X Li et al., “Crystal Structure of the Human Cannabinoid Receptor CB2”, Cell, 2019	Published			
4	Paper	M Audet et al., “Crystal structure of misoprostol bound to the labor inducer prostaglandin E 2 receptor”, Nature Chemical Biology, 2019	Published			
5	Paper	S Yang et al., “Crystal structure of the Frizzled 4 receptor in a ligand-free state”, Nature, 2018	Published	High		
6	Paper	Y Peng et al., “5-HT2C Receptor Structures Reveal the Structural Basis of GPCR Polypharmacology”, Cell, 2018	Published			
7	Paper	A Batyuk et al., “Native phasing of x-ray free-electron laser data for a G protein–coupled receptor”, Science Advances, 2016	Published			
8	Paper	A Luginina et al., “Structure-Based Mechanism of Cysteinyl Leukotriene Receptor Inhibition by Antiasthmatic Drugs”, Science Advances, 2019	Published			
9	Paper	A Gusach et al., “Structural Basis of Ligand Selectivity and Disease Mutations in Cysteinyl Leukotriene Receptors”, Nature Communications, 2019	Published			
10	Grant	Skoltech STRIP. Digital Platform for GPCR-specific drug discovery	Realized	High		https://sip.skoltech.ru/en/digital-platform-for-gpcr-specific-drug-discovery/

A.108 AI-based mapping of optical to multi-electrode catheter recordings for Atrial Fibrillation Treatment

A.108.1 General

<table>
<tr><td>ID</td><td colspan="2">108</td></tr>
<tr><td>Use case name</td><td colspan="2">AI-based mapping of optical to multi-electrode catheter recordings for Atrial Fibrillation Treatment</td></tr>
<tr><td>Application domain</td><td colspan="2">Healthcare</td></tr>
<tr><td>Deployment model</td><td colspan="2">Embedded systems</td></tr>
<tr><td>Status</td><td colspan="2">PoC</td></tr>
<tr><td>Scope</td><td colspan="2">Predicting possible targets for Atrial Fibrillation Ablation based on explanted human heart data of two modalities (multi-electrode mapping and near-infrared optical imaging)</td></tr>
<tr><td>Objective(s)</td><td colspan="2">Given: Recordings from multi-electrode catheter grid, with ground-truth labels from near-infrared optical mapping, obtained from explanted hearts. Output: possibility of recordings to be from source (driver) region of atrial fibrillation.</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>Atrial fibrillation (AF) is the leading cause of stroke with low treatment rate maintained by micro-anatomic intramural re-entry called drivers. Unfortunately, the current clinical method to look for drivers (multi-electrode mapping, MEM) suffers from many limitations, including poor resolution and only-surface tissue mapping. On the other hand, near-infrared optical mapping (NIOM) has 1000 times higher resolution and records electrical activity from the depth of atrial tissue (up to 5 mm), but needs specific voltage-sensitive dye. For our research, we used simultaneous recordings of AF episodes from Ohio State University. We predicted the possibility of AF drivers to be visible in the MEM recording as trained by the Optical ex-vivo data. We created the machine learning classifier with ground-truth labels based on NIOM maps. As features, we used characteristics from the Fourier spectra of MEM recordings.</td></tr>
<tr><td>Complete description</td><td>Atrial fibrillation (AF) is the most common cardiac arrhythmia and the leading cause of stroke. The success rate of current AF treatment is low, 50-70%. Several experimental and clinical studies suggest that AF may be caused and maintained by micro-anatomic intramural re-entry called drivers. Physical destruction of the driver, or driver ablation, leads to the termination of AF. Unfortunately, the current clinical method to look for drivers (multi-electrode mapping, MEM) suffers from many limitations, including poor resolution and only-surface tissue mapping. On the other hand, near-infrared optical mapping (NIOM) has 1000 times higher resolution and records electrical activity from the depth of atrial tissue (up to 5 mm), but needs specific voltage-sensitive dye to color the tissue and therefore can be used only for explanted specimens. For our research, we used unique data of the experiments with explanted human atria from Ohio State University – simultaneous recordings of AF episodes by</td></tr>
</table>

<table>
<tr><td></td><td></td><td colspan="3">MEM and NIOM. In this work, we predicted the possibility of AFib drivers to be visible in the MEM recording as trained by the Optical ex-vivo data. We created the machine learning classifier with ground-truth labels based on NIOM maps. As features, we used characteristics from the Fourier spectra of MEM recordings. Our experiments on a dataset of more than 20000 spectra provided an accuracy and f1-score of 97.3% and 0.89, respectively.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Hospitals, cardiologists</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Trustworthiness, competitiveness, fair treatment, stability</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">New privacy threats, new security threats</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Classifier</td><td>ML classifier with ground-truth labels based on NIOM maps</td><td>ML features generated for each Fourier spectrum calculated for NIOM and MEM data</td></tr>
<tr><td>2</td><td>f1-score</td><td>Effect of adding NIOM features to MEM features on f1-score</td><td>Metric</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3">HPC</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Machine learning, deep learning, classification, fourier spectrum, features</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">1) There is only one laboratory in the world that provide the needed explanted human atria; 2) the number of experiments is limited (approximately 20 atria per year), and collecting the data is difficult; 3) only a few experiments consist of two modalities recordings and are therefore suitable for this research</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Better life quality for Atrial Fibrillation patients, diminishment of stroke accidents caused by Atrial Fibrillation genesis; as a result, decreased mortality of such patients</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3">Good health and well-being for people</td></tr>
</table>

A.108.2 Data

Data characteristics	
Description	Simultaneous recordings of AF episodes by MEM and NIOM

Source	Experiments with explanted human atria
Type	Electrical activity spectra
Volume (size)	> 20000 spectra
Velocity	Prediction of the possibility of AFib drivers to be visible in the MEM recording
Variety	ML features were generated for each Fourier spectrum calculated for NIOM and MEM data; classification algorithms for MEM data; adding NIOM features to MEM features on f1-score was also tested
Variability (rate of change)	Moderate
Quality	Accuracy: 97.3%, f1-score: 0.89

4145

4146 A.108.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organi zation	Link
1	Abstra ct	Alexander Zolotarev, Ekaterina Ivanova, Brian J. Hansen, Katelynn M. Helfrich, Dmitry Dylov, Vadim V. Fedorov. Machine Learning Trained with Optical Mapping Improves Detection of Atrial Fibrillation Drivers for Clinical Multi-Electrode Mapping	publis hed	High	Dorothy M. Davis Heart and Lung Research Institute, Wexner Medical Center, Ohio State University, Columbus, OH	https://heartlun g.osu.edu › Lists › Attachments
2	Abstra ct	Dmitry Dylov. Towards autonomous surgical suturing: augmented stitching of coronal incision. World Congress on Medical Physics and Biomedical Engineering 2018 (WC2018)	publis hed	Moderate	Czech Society for Biomedical Engineering and Medical Informatics, Prague	https://guarant. topinfo.cz/iupes m2018/en/prog ramme-in-details
3	Master Thesis	Ivanova EA. Multi-modal Machine Learning Toolset for Spatio-Temporal Characterization of Atrial Fibrillation Drivers in Human Heart	Defen ded	Moderate	HSE	https://www.hs e.ru/edu/vkr/2 96286627

4147

A.109 AI-dispatcher (operator) of large-scale distributed energy system infrastructure

A.109.1 General

<table>
<tr><td>ID</td><td colspan="2">109</td></tr>
<tr><td>Use case name</td><td colspan="2">AI-dispatcher (operator) of large-scale distributed energy system infrastructure</td></tr>
<tr><td>Application domain</td><td colspan="2">Energy</td></tr>
<tr><td>Deployment model</td><td colspan="2">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="2">PoC</td></tr>
<tr><td>Scope</td><td colspan="2">Monitoring, optimization and control of large scale distributed energy systems using Deep Reinforcement Learning (gas, oil, power, heat, water transmission and distribution infrastructure systems)</td></tr>
<tr><td>Objective(s)</td><td colspan="2">To develop an effective industrial AI solution which is able to recommend the optimal control of energy infrastructure systems in real-time in order to:
— satisfy the energy demand of consumers
— minimize possible negative impacts on the environment
— reduce operational costs through systems' real-time continuous optimization in self-adaptive manner</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>An AI solution is currently in development that uses hybrid models (based on both traditional physics models and artificial neural networks), “digital twins,” and deep reinforcement learning to support decision making and control of energy infrastructure systems in real-time.</td></tr>
<tr><td>Complete description</td><td>Motivation
The existing technologies do not provide an effective solution to the problem of optimization of distributed energy systems in real time. At the same time, the effects of optimization in the energy sector are substantial.

Objects (systems) under consideration
Real large-scale distributed energy systems (gas, oil, power, heat, water transmission and distribution infrastructure systems). The main features of systems under consideration:
— Territorial distribution and a large number of interconnected units of equipment with individual characteristics
— The complex physics of technological processes
— Huge amounts of real-time information from various sensors

Problem statement
The central goal of the AI solution that is being developed is formulated as follows: to ensure the supply of energy of a certain quality at the right time to all consumers of a</td></tr>
</table>

<table>
<tr><td></td><td></td><td colspan="3">distributed energy system, taking into account all technological limitations and minimizing the operational costs through systems’ real-time continuous optimization. Solving this problem requires solving a number of subtasks.

Solution Approach
The AI solution uses an approach of industrial system modeling based on hybrid models which combine the benefits of traditional physics-based modeling and machine learning capabilities. We use the reliable “digital twins” of energy systems and virtual simulators to simulate the systems’ physics (dynamics) and we train deep reinforcement learning models of these systems.

Current results
PoC of the AI system has been developed, which consists of:
— “digital twins” of real gas infrastructure systems

— reliable physics-based models and virtual simulators of these systems, actively used in industry

— model-free deep reinforcement learning algorithms, connected with the above-mentioned virtual simulators

— services for training models, visualizing and analyzing the results

Computational experiments proved that the initial objective can be achieved with the help of modern AI technologies. The results show the effectiveness of using AI based technologies to optimize and control of distributed energy systems and that these solutions can outperform both human capabilities and traditional optimization algorithms that were proposed earlier.

Technologies
Physics-based modeling, deep reinforcement learning technologies, deep learning frameworks, big data technologies, streaming platforms, cloud-native architecture of AI-system.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Energy companies focused on AI solutions to drive the energy production, transition and distribution in large territories</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Safety and environment, competitiveness, stability</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Sources of bias, security threats</td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Accuracy of forecasts</td><td>Convolved ratio of actual system’s parameters over predicted parameters</td><td>Improve accuracy</td></tr>
</table>

	2	Optimization rate	The ability to optimize real energy systems (expert assessment)	Improve efficiency
	3	Response time	The response time required to react to new conditions (changes in the environment parameters)	Improve reliability
AI features	Task(s)	Prediction, optimization, recommendation		
	Method(s)	Time series analysis; artificial neural networks; deep reinforcement learning, decision making and control; physics-based simulation of infrastructure systems; Monte Carlo tree search		
	Hardware	High performance CPU and GPU		
	Topology	Agent based topology which resembles the topology of real systems' dispatching control.		
	Terms and concepts used	Deep learning, neural networks, training, reinforcement learning, automation.		
Standardization opportunities/ requirements	**Standards (Guidelines) for Virtual Simulators APIs** Currently virtual simulator developers usually use their own data formats and often do not provide an API to access simulator services. Standardization is due to the need to embed these simulators in AI systems, which will use them to train ML-models and AI-agents. **Standards (Guidelines) for Reference Architectures of "digital twins" of industrial objects** Different companies offer the development of "digital twins" for different industrial objects. To avoid the "patchwork digitalization" and to ensure the compatibility of the "digital twins zoo" within common AI-solutions, it is necessary to standardize their typical Reference Architectures. **Standards (Guidelines) for Reference Architectures of AI-systems that are used in energy sector** We need special standards for AI-systems of the energy sector due to their importance in our every-day life.			
Challenges and issues	— To achieve a high level of efficiency of complex energy system's optimization and dispatching control — To learn from human-beings, including machine teaching techniques — To employ meta-learning techniques in real industrial environments, which can help AI-agents to adopt efficiently to different systems (for example, from small scale to large scale industrial systems, from gas to oil transmission system, from power to heat infrastructure systems, and vice versa) — To deal effectively with partially observed systems — To develop an AI-solution which reacts reliably to rare events			

Societal Concerns	Description	Safety, security and reliability of AI solutions that are used in energy infrastructure management.
	SDGs to be achieved	Affordable and clean energy

A.109.2 Data

Data characteristics	
Description	We use reliable virtual simulators to generate synthetic data of technological regimes of energy systems based on physics-based modeling
Source	Virtual simulations
Type	Structured data
Volume (size)	~10 GB of synthetic data
Velocity	Real time emulation
Variety	Mostly structured
Variability (rate of change)	Moderate
Quality	Moderate

A.109.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Book	Reinforcement Learning: An Introduction. Richard S. Sutton and Andrew G. Barto. Second Edition. MIT Press, Cambridge, MA, 2018	Published	High	MIT Press, Cambridge	
2	Monograph	Pipeline energy Systems: mathematical and computer modeling. Novosibirsk: “Science”. 2014 (in Russian).	Published	High	Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences	
3	Ph.D thesis	A.V. Belinsky Elaborating methods, algorithms and software for development and reconstruction of territorial gas supply systems. Moscow, Gubkin Russian state University of oil and gas. 2009 (in Russian).	Published	High	Gubkin Russian State University of Oil and Gas	
4	Book	S.A. Sardanashvili Calculation methods and algorithms (pipeline gas transportation). Moscow: “Oil and gas”. Gubkin Russian state University of oil and gas. 2005 (in Russian).	Published	High	Gubkin Russian State University of Oil and Gas	

A.110 Analyzing and Predicting Acid Treatment Effectiveness of Bottom Hole Zone

A.110.1 General

<table>
<tr><td>ID</td><td colspan="4">110</td></tr>
<tr><td>Use case name</td><td colspan="4">Analyzing and Predicting Acid Treatment Effectiveness of Bottom Hole Zone</td></tr>
<tr><td>Application domain</td><td colspan="4">Manufacturing, gas & oil</td></tr>
<tr><td>Deployment model</td><td colspan="4">Client and server systems</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Mining of oil and gas; digital assistant for analyzing and predicting the effectiveness of acid treatments of the bottom hole zone</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Predict the effectiveness of acid treatments of the bottom hole zone</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Predicting the technological and economic efficiency of acid treatments of the bottom-hole zone of the well</td></tr>
<tr><td>Complete description</td><td colspan="3">Currently, a long and subjective selection of candidate wells for acid treatments is being carried out.
An application with mathematical models for automating statistical analyses and predicting the technological and economic efficiency of acid treatments of the bottom hole zone of the well in the form of additional oil and well production.
The ranking of wells according to the degree of effectiveness of acid treatment of the bottom hole zone.
Determining the significance of various factors on the regression model for the field.
The goal is a convergence of the obtained forecast of the mathematical model with historical data of at least 80%.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Manufacturer</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Decision quality</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Coefficient of determination</td><td>Prediction accuracy's ideal target is 0.9.</td><td>Prediction of future outcomes</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine learning</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
<tr><td>Topology</td><td colspan="3">Trees, random forest, boosting</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Machine learning, training, training data set</td></tr>
<tr><td>Standardization</td><td colspan="4"></td></tr>
</table>

opportunities/ requirements		
Challenges and issues	Challenges: To achieve high level accuracy of prediction efficiency of acid treatments	
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, are important for facilitating sustainable development
	SDGs to be achieved	Industry, innovation, and infrastructure

A.110.2 Data

Data characteristics	
Description	Data from different well sensors
Source	
Type	Structured data, .csv
Volume (size)	100 Mb
Velocity	Real time
Variety	Different source
Variability (rate of change)	Static
Quality	Position updates may be incomplete

A.110.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model with training data set	Sample raw data set is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training/re training		Meeting KPI requirements (e.g. Coefficient of determination is 0.9) is the "success" condition
3	Execution	Prediction	Completion prediction	The trained model has been evaluated as deployable	
4	Retraining	Retrain a model with training data set	New well data		

A.110.4 Training

Scenario name	Training

Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample data set is ready	Data preparation	Manufacture r	Transform raw data into .csv.	The software for data preparation has to be provided by the AI solution provider.
2	Completion of Step 1	Training data set creation	Manufacture r	Create training data set for further analysis the output of Step 1	
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	

Specification of training data	

A.110.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Model evaluation	Manufacturer	Compare the result of model work with that of human inspection	

Input of evaluation	
Output of evaluation	

A.110.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Acquisition of data	Prediction	Manufacture r	Prediction efficiency of acid treatments	A trained model should convey the results of the work to the manufacture r.

Input of Execution	

Output of Execution	

A.110.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Getting new data	Data preparation	Manufacturer	Transform data into .csv.	
2	Completion of Step 1	Training data set creation	Manufacturer	Create training data set for further analysis the output of Step 1 with different classes and balancing	
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	

Specification of retraining data	Collecting data for new wells

A.110.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure		In operation		Gazprom neft	

A.111 Application of Strong Artificial Intelligence

A.111.1 General

ID	111
Use case name	Application of Strong Artificial Intelligence
Application domain	Other (please specify): Hi-Tech Labor Market
Deployment model	Hybrid or other (please specify): Artificial Agents
Status	In operation
Scope	Economic sectors and social services
Objective(s)	Find accurate and universal application of strong artificial intelligence

Narrative	Short description (not more than 150 words)	Strong artificial intelligence is a digital twin of human intelligence, capable of learning, retraining, self-realization and development by improving functional activities through the mastery of creative and innovative high-tech professional and behavioral skills and competences according to a criteria of preferences with qualitative choices.		
	Complete description	Technology that studies the development of digital twins capable of acquiring, processing and applying human knowledge and skills, purchased through training, to solve problems and adapt to changing circumstances with or without human or external control in physical work, as well as in mental or cognitive work. The technology builds models by analyzing quantitative and qualitative data from different perspectives and measurements, classifying them and summarizing potential relationships and impacts. The technology uses natural language processing and machine learning to interact more naturally and expand human experience and knowledge on a permanent basis during operation. The technology has robust mechanisms by which to ensure security in ways that humans understand. The technology shows smart behavior comparable to a human across a range of cognitive abilities. The technology models the spectrum of human abilities by retraining. It relies on an infrastructure of interconnected actors, people, systems and information resources from high-tech industries and social spheres, as well as on services that process and respond to information from the physical and virtual world of social and cognitively smart robots, such as guides, sellers, teachers, nurses, volunteers, guards, and administrators.		
Stakeholders	Highly technological producers			
Stakeholders' assets, values	Reputation			
System's threats & vulnerabilities	Legal and ethical aspects of interaction with society.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	AI management of professional cooperation process	The technology of creative process control can itself predict optimal terms of execution of certain stages on the basis of accumulated information about labour intensity, selection of the route of staff loads and competences of employees. It optimizes processes during execution, automatically delegating tasks and taking into account the load of employees and their competences.	Improve accuracy
	2	Productivity and quality AI	Strong artificial intelligence works with fewer mistakes and is safer. Strong artificial intelligence improves human quality of life and addresses society's daily concerns, as well as benefits productivity in	Improve accuracy

			high-tech industries and production.	
AI features	Task(s)	Other (please specify): Creative activity		
	Method(s)	Deep learning		
	Hardware	Supercomputer with Strong Artificial Intelligence		
	Topology	Distributed Modular Interconnect Topology		
	Terms and concepts used	Deep learning, "imagification", neural networks, training, training data set		
Standardization opportunities/ requirements	Strong artificial intelligence requires process standardization, as does every human activity.			
Challenges and issues	Qualitatively new type of thinking not available to humans			
Societal Concerns	Description	Security and ethical and legal aspects		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.111.2 Data

Data characteristics	
Description	Strong Artificial Intelligence Data
Source	Model and technology of Strong Artificial Intelligence
Type	Strong
Volume (size)	Hi-tech labor market
Velocity	Supercomputing velocity
Variety	Streams of multiple datasets
Variability (rate of change)	Retraining
Quality	High

A.111.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model (deep neural network) with training data set	Technological process raw data set is ready	Formatting of data	Management of safety

2	Evaluation	Expansion of the trained model	Development of technological thinking and behaviour	Cognitive thinking patterns and psychological behaviors	Meeting KPI requirements is condition of development
3	Execution	Model and Technology Tooling	Interaction	Activization of Model	Completion of interaction
4	Retraining	Retrain a model with training data set	A certain period of time has passed since the last training/retraining	Additional data and knowledge	Combining data and knowledge

A.111.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample raw data set is ready	Specification and classification	Manufacturer	Transform sample of raw data	Strong AI Software
2	Completion of Step 1	Creating Set of Experimental Data	Manufacturer	Development of a set of experimental data through job modelling	Software of modelling
3	Completion of Step 2	Model training	AI solution provider	Train a model (deep neural network) with experimental data set created by Step 2	Big Data

Specification of training data	Big Smart Data

A.111.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Research	Manufacturer	Train model (deep neural network) with experimental data set created	Big Data
2	Completion of Step 1	Identification	AI solution provider	Based on data, detect	Smart Data

				execution using a deep neural network trained in a learning scenario	
3	Completion of Step 2	Evaluation	Manufacturer	Comparison of phase 2 results with human performance	Efficiency and quality

Input of evaluation	Productivity
Output of evaluation	Efficiency and quality

4181

4182 ### A.111.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of comparison of modeling results with human performanc e	Research	Manufacture r	Development of a set of experimental data through job modelling	Quality
2	Completion of Step 1	Identification	Manufacture r	Based on modified data train model (deep neural network) with experimental data set created	Compatibilit y

Input of Execution	Modification
Output of Execution	Compatibility

4183

4184 ### A.111.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activit y	Requirement
1	Certain period of time has passed since the last training/retraining	Research	Manufact urer	Additional data and knowledge	Completeness
2	Completion of Step 1	Experimental data set creation	Manufact urer	Combining Data and Knowledge Based on	Compatibility

				modified data train model (deep neural network) with experimental data set created	
3	Completion of Step 2	Model training	AI solution provider	Comparison of phase 2 results with human performance	Efficiency and quality

Specification of retraining data	Retraining data set has to include recent data

4185

4186 A.111.8 **References**

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	Evgeniy Bryndin. Cognitive Robots with Imitative Thinking for Digital Libraries, Banks, Universities and Smart Factories. International Journal of Management and Fuzzy Systems. V.3, N.5, 2017, pp 57- 66.	Published	Strong AI of Robots Technological Artificial Intelligence in the social sphere and industry	Research Center “NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=353&doi=10.11648/j.ijmfs.20170305.11
2	Paper	Evgeniy Bryndin. Program Hierarchical Realization of Adaptation Behavior of the Cognitive Mobile Robot with Imitative Thinking. International Journal of Engineering Management. Volume 1, Issue 4. 2017, pp. 74-79.	Published	Strong AI of Robots Program realization of Technological Artificial Intelligence	Research Center “NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=522&doi=10.11648/j.ijem.20170104.11
3	Paper	Evgeniy Bryndin. Technological Thinking, Communication and Behavior of Androids. Communications. Vol. 6, No. 1, 2018. Pages: 13-19.	Published	Strong AI of Robots Technological Artificial Intelligence	Research Center “NATURE INFORMATIC Russia, Novosibirsk	http://article.sciencepublishinggroup.com/pdf/10.11648.j.com.20180601.13.pdf
4	Paper	Evgeniy Bryndin. Communicative Associative Logic of Cognitive Professional Robot with Imitative Thinking. Journal Engineering Mathematics, Volume 2,	Published	Strong AI of Robots Technological Artificial thinking	Research Center “NATURE INFORMATIC Russia, Novosibirsk	http://article.sciencepublishinggroup.com/pdf/10.11648.j.engmath.20180202.14.pdf

		Issue 2. 2018. Pages: 79-85.				
5	Paper	Evgeniy Bryndin. Social Cognitive Smart Robots: Guide, Seller, Lecturer, Vacuum Cleaner, Nurse, Volunteer, Security Guard, Administrator. Communications. Volume 7, Issue 1. 2019. Pages: 6-12.	Published	Strong AI of Robots Technological Artificial Intelligence in the social sphere	Research Center "NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=139&doi=10.11648/j.com.20190701.12
6	Paper	Evgeniy Bryndin. System retraining to professional competences of cognitive robots on basis of communicative associative logic of technological thinking. International Robotics Automation Journal. 2019; 5(3.):112–119	Published	Strong AI of Robots Artificial Intelligence in technological training	Center "NATURE INFORMATIC Russia, Novosibirsk	https://medcraveonline.com/IRATJ/
7	Paper	Evgeniy Bryndin. Human Digital Doubles with Technological Cognitive Thinking and Adaptive Behaviour. Software Engineering, Volume 7, Issue 1, 2019. P. 1-9.	Published	Strong AI Technological Artificial Intelligence	Center "NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/j/se
8	Paper	Evgeniy Bryndin. Robots for Communication in Public in High-Tech Industry Life and Space. Frontiers Journal of Current Engineering Research. Volume 1, Issue 1, 2019. P. 1-10.	Published	Strong AI of Robots Technological Artificial Intelligence in the social sphere	Center "NATURE INFORMATIC Russia, Novosibirsk	https://fmpublishers.org/admin/uploads/journals/pdfs/1567063131.pdf
9	Paper	Evgeniy Bryndin. Mainstreamig technological development of industrial production based on artificial intelligence. COJ Technical & Scientific Research, 2(3). 2019. Pages: 1-5.	Published	Strong AI: Paradigms, Architectures, and Methods Technological development on artificial intelligence	Center "NATURE INFORMATIC Russia, Novosibirsk	https://crimsonpublishers.com/cojts/pdf/COJTS.000539.pdf
10	Paper	Evgeniy Bryndin. Robots with Artificial Intelligence and Spectroscopic Sight in Hi-Tech Labor Market. International Journal of Systems Science and Applied Mathematic, V.	Published	Strong AI of Robots Hi-Tech technological artificial intelligence	Center "NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=245&doi=10.11648/j.ijssam.20190403.11

		4, No 3, 2019. Pages: 31-37				
11	Paper	Evgeniy Bryndin. Collaboration Robots as Digital Doubles of Person for Communication in Public Life and Space. American Journal of Mechanical and Industrial Engineering, Volume 4, Issue 2, 2019. Pages: 35-39.	Published	Strong AI of Robots Technological Artificial Intelligence in the social sphere and Space	Center "NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=248&doi=10.11648/j.ajmie.20190402.12
12	Paper	Evgeniy Bryndin. Collaboration Robots with Artificial Intelligence as Digital Doubles of Person for Communication in Public Life and Space. Budapest International Research in Exact Sciences (BirEx-Journal), Volume 1, No. 4, 2019. P: 1-11.	Published	Strong AI of Robots Technological Artificial Intelligence in the social sphere and Space	Center "NATURE INFORMATIC Russia, Novosibirsk	https://bircu-journal.com/index.php/birex/article/view/473/pdf
13	Paper	Evgeniy Bryndin Formation Smart Data Science for Automated Analytics of Modeling of Scientific Experiments. American Journal of Software Engineering and Applications. Volume 8, I. 2, 2019. Pages: 36-43.	Published use case: "Application of Strong Artificial Intelligence"	Strong AI	Center "NATURE INFORMATIC Russia, Novosibirsk	http://www.sciencepublishinggroup.com/journal/archive?journalid=137&issueid=-1
14	Paper	Evgeniy Bryndin. Supercomputer BEG with Artificial Intelligence of Optimal Resource Use and Management by Continuous Processing of Large Programs. International Journal of Research in Engineering, Vol. 1, Issue 2, 2019. Pages: 9-14.	Published	Super computer BEG with AI	Center "NATURE INFORMATIC Russia, Novosibirsk http://www.engineeringpaper.net/article/view/9/1-2-13	
15	Paper	Evgeniy Bryndin. Practical Formation of Creative Life-Saving Strong Artificial	In the press	Strong AI	Center "NATURE INFORMATI	

		Intelligence. International Journal Artificial Intelligence Research, Vol.8, No.2, 2019.			C Russia, Novosibirsk	

A.112 Automatic Classification Tool for Full Size Core

A.112.1 General

ID	112			
Use case name	Automatic Classification Tool for Full Size Core			
Application domain	Manufacturing, Gas & Oil			
Deployment model	Client and server systems			
Status	In operation			
Scope	Oil and Gas exploration, classification of rock types, oil saturation, carbonate and fracture according to core images			
Objective(s)	Classification of rock types Classification of oil saturation Classification of carbonate and Classification of fracture according of core			
Narrative	Short description (not more than 150 words)	A solution has been developed that allows for the classification of rock types into four classes. This resulted in an 80% reduction in core image analysis.		
	Complete description	To describe the core of an exploratory well, 3 specialists are allocated for a period of up to 2 weeks with travel expenses. The results of the description are subjective and may contain conflicting positions of experts. Automation of the process of classifying rock types, saturation, carbonate and rock layer degradation by daylight and ultraviolet images using machine learning mechanisms can reduce the lithotype typing time to 3 days.		
Stakeholders	Manufacturer, geologist			
Stakeholders' assets, values	Decision speed			
System's threats & vulnerabilities	Quality of images received from special equipment			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Coverage	Rock type accuracy's ideal target is 80%	Improved accuracy
	2	Splits	Detection of splits in the rock with an accuracy of 2 centimeters	Improved efficiency
AI features	Task(s)	Recognition		
	Method(s)	Machine learning, classic computer vision		

	Hardware	Camera
	Topology	Trees, Random forest
	Terms and concepts used	Machine learning, computer vision, training, training data set
Standardization opportunities/ requirements		
Challenges and issues	To achieve the same level of accuracy of recognition of rock types as expert lithologists; To minimize the set of laboratory tests due to visual recognition of rock types and their parameters from core images	
Societal Concerns	Description	Promoting sustainable industries, and investing in scientific research and innovation, is important for facilitating sustainable development.
	SDGs to be achieved	Industry, innovation, and infrastructure

A.112.2 Data

Data characteristics	
Description	DL and UV core photos
Source	UT scanning instrument
Type	Photo
Volume (size)	50 Gb
Velocity	Batch
Variety	Single source
Variability (rate of change)	Static
Quality	Middle

A.112.3 Process scenario

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Training	Train a model with training data set	Sample raw data set is ready		
2	Evaluation	Evaluate whether the trained model can be deployed	Completion of training / retraining		Meeting KPI requirements (e.g. accuracy of classification is 0.8 on multiclass classification) is the "success" condition
3	Execution	Classificati on of rock types, saturation, carbonate	Completion classificatio n	The trained model has been evaluated	

				as deployable	
4	Retraining	Retrain a model with training data set	New examples of tagged images from other fields are obtained		

A.112.4 Training

Scenario name	Training				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Sample data set is ready	Data preparation	Manufacture r	Transform sample photos to segments for further analysis	The software for image preparation has to be provided by the AI solution provider.
2	Completion of Step 1	Training data set creation	Manufacture r	Create training data set by labelling the output of Step 1 with different classes and balancing	
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	

Specification of training data	

A.112.5 Evaluation

Scenario name	Evaluation				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Completion of training/retraining	Model evaluation	Manufacturer	Compare the result of model work with that of human inspection	

Input of evaluation	
Output of evaluation	

A.112.6 Execution

Scenario name	Execution				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Acquisition of core images for analysis	Classification	Manufacturer	Classification of rock types, oil saturation, carbonate and fracture according of core	A trained model should convey the results of the work to the manufacturer.

Input of Execution	
Output of Execution	

A.112.7 Retraining

Scenario name	Retraining				
Step No.	Event	Name of process/Activity	Primary actor	Description of process/activity	Requirement
1	Getting new data	Data preparation	Manufacturer	Transform sample photos to segments for further analysis	
2	Completion of Step 1	Training data set creation	Manufacturer	Create training data set by labelling the output of Step 1 with different classes and balancing	
3	Completion of Step 2	Model training	AI solution provider	Train a model with the training data set created by Step 2	

Specification of retraining data	

A.112.8 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure		In operation		Gazprom neft	

4204 A.113 Autonomous Trains (Unattended Train Operation (UTO))

4205 A.113.1 General

ID	113	
Use case name	Autonomous Trains (Unattended Train Operation (UTO))	
Application domain	Transportation	
Deployment model	Self-driving vehicles	
Status	prototype	
Scope	Freight and passenger trains operate autonomously, excluding any crew presence on board, but with remote operator attention involved (GoA 4).	
Objective(s)	The critical objective of automation in trains is to provide extra reliability, safety and to prevent accidents on railways, which tend to be caused by human error. Moreover, the provided innovation leads to energy consumption optimization, transport capacity increases, and, eventually, possible reduction of personnel costs due to the autonomous operation.	
Narrative	Short description (not more than 150 words)	Regarding passenger transportation, UTO enables unattended operation of trains according to schedule. The system is responsible for the train's acceleration, braking, speed control, station departure, doors opening and closing, obstacle detection, management of hazardous conditions, and emergency situations. Autonomous trains obtain data from sensors (internal - GPS, various types of cameras, LIDARs, RADARs) and traffic control systems (train schedule, movement authority), in order to interact with passengers, other vehicles, and obstacles based on information about the environment.
	Complete description	There is a lot of information about self-driving automobiles. Developing computer vision technology, reliable navigation, and radio communication makes creating self-driving trains technologically feasible. Compared to cars, trains have a long braking distance. This means that autonomous trains have to have a unique obstacle detection system, which can spot obstacles up to 1000 meters away and more. Both conventional and autonomous railway systems consist of fleet and infrastructure. Current interaction between locomotive and dispatcher is realized by voice communication. For autonomous trains use, digital communication with formal commands for train control is necessary. Key AI development realized into the obstacle detection module can be fulfilled with both computer vision methods by processing data received from sensors (LIDARs, RADARs, infrared and electro-optical cameras) and by positioning and localization based on prior electronic map information and obtained data from GPS information. This system can work under differences in light, weather, and timing conditions. The data collected from sensors with a varied

<table>
<tr><td></td><td></td><td colspan="3">range of actions and purposes is processed by classical image analysis and deep learning approaches; it is then fused. Methods such as semantic segmentation, object detection, LIDAR points clustering, tracking, localization, and mapping are used. All in all, this leads to clear scene perception and safety system responses. The machine can trigger the alarm, halt, apply the brakes, or accelerate based on information about the environment.

However, a remote driver is still needed to resolve complicated cases, which the on-board system is not able to process correctly. Considering that the system's priority is safety, such examples most commonly include false-positive object occurrences. It is important to stress that one remote driver operator can control the performance of several automated trains at the same time.

Three autonomous shunting locomotives are already in operation at Luzhskaya Marshalling Yard in Moscow, Russia; and parallel deployment for passenger trains is current under test on the Moscow Central Ring.</td></tr>
<tr><td>Stakeholders</td><td colspan="4"></td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4"></td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Efficiency and economic benefits</td><td>Reduced fuel consumptions, reduced operation cost, capacity increase</td><td>Advantage of autonomous train on railways</td></tr>
<tr><td>2</td><td>Safety</td><td>Safety due to lack of fatigue and applying current sensors able to detect obstacles in bad weather conditions</td><td></td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition, Planning, Prediction, (from here – not from the given list) Environment perception</td></tr>
<tr><td>Method(s)</td><td colspan="3">Computer vision, image processing, obstacle detection, semantic segmentation, tracking, decision making, localization & positioning, data fusion</td></tr>
<tr><td>Hardware</td><td colspan="3">Obstacle detection unit: GPS sensors, LIDAR, RADAR, electro-optical and infrared cameras, computer module with GPU</td></tr>
<tr><td>Topology</td><td colspan="3">CNN and others</td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">autonomous vehicle guidance, environment perception, self-perception, computer vision, deep learning, convolutional neural networks</td></tr>
</table>

Standardization opportunities/ requirements		
Challenges and issues		
Societal Concerns	Description	Safety, reliability, security, (potential) job loss
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.114 Finance Advising and Asset Management with AI

A.114.1 General

ID	114	
Use case name	Finance Advising and Asset Management with AI	
Application domain	Fintech	
Deployment model	Cloud service	
Status	In operation	
Scope	Financial advising and portfolio management for financial institutions and consumers	
Objective(s)	Designed to manage exchange-traded securities portfolios of conservative investors in real time, using asset price data and macroeconomic data, to make the most accurate decisions at a given yield and moderate risk. Prediction of significant depreciation of exchange-traded asset prices as a result of a sharp monetary contraction called financial crises.	
Narrative	Short description (not more than 150 words)	The core of the system carries out a structured collection from open sources and multi-threaded parallel analysis of information; it regulates the application of basic algorithms and rules for changing these algorithms that change the purpose of the task. (Intermediate goal setting is one of the elements of "Strong AI"). One of the tasks is to assess market trends, as well as market and interest rate risk. Changes in the algorithm of actions depend on the macroeconomic information received from the outside. It translates notoriously weakly formalized parameters into specific decisions on the formation of investment portfolios and issues orders to brokers to purchase, rebalance, or sell assets in stock exchanges. The macroeconomics unit is an autonomous system that generates indicators of time periods and geographical areas with different weights of investment potential.
	Complete description	For the purposes of efficiency, which cannot be achieved by competitors, the project uses more complex technologies than offer standard solutions for building neural systems. All algorithms of the basic core of the project are developed by the creators themselves. The idea that neural systems are absolute, impenetrable "black boxes" is mythologized. Therefore, by understanding exactly what technologies are used to achieve analysis

		goals, overloaded "boxed" solutions can be optimized. This was done in the project. The algorithm of simple regression analysis of prices (model William Sharpe/Harry Markowitz - Nobel laureates) does not lead to the required efficiency. Therefore, the project uses the "complex" model when weighting factors and the algorithms of simple regression analysis of prices change depending on the "field," formed by the regression assessment of other economic parameters. The William Sharpe/Harry Markowitz model is unacceptably simplified precisely because it is very resource-intensive. This is particularly true when it comes to the hundreds of asset names around the world for the diversification needed in this model. If we consider applying a straight-line approach to the assessment of dozens or even hundreds of additional macroeconomic parameters of each of the dozens of different countries (and today it is clear that the world economy is interrelated), we are talking either about supercomputers and very expensive neural models, or about building a fundamentally new economic model for the AI core. In this project, the regression evaluation of higher-order macroeconomic indicators "guides" all subsequent lower-order models. Resource issue resolved.		
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Portfolio yield	The percentage return of the portfolio compared to the benchmark	Long-term, from 10 to 20 years, the retention of positive returns is significantly higher than the base
	2	Sharpe ratio	Risk assessment strategies	A higher Sharpe ratio is an indication of a higher level of control reliability (1 to 2 or more)
AI features	Task(s)	Prediction, advising and management		
	Method(s)	Ensemble models		
	Hardware	64 GB RAM, 2 x Intel Core i7		
	Topology			
	Terms and concepts used			
Standardization				

opportunities/ requirements		
Challenges and issues	1. Data can be noisy, may have several missing values, and needs appropriate pre-processing and treatment before feeding to the model algorithm 2. Working with financial assets requires high reliability of computing systems and replication systems	
Societal Concerns	Description	
	SDGs to be achieved	No poverty

A.114.2 Data

Data characteristics	
Description	1. Historical and real-time securities price data 2. Historical and real-time macroeconomic data
Source	1. Securities prices from exchanges 2. Open source, websites of Central banks and the IMF
Type	Structured Data
Volume (size)	4 TB
Velocity	Real-time data replenishment 100 mbps
Variety	Mostly Structured
Variability (rate of change)	high
Quality	high

A.114.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	“Finance advising and asset management with AI”	High quality company whitepapers and presentations	High	AI Sys Financial	http://aisfin.ru/wp-content/uploads/2019/10/Sk_AISFin_101019.pdf
2	Paper	Botvinnik M.M. Chess method for solving iterative problems. - Moscow, Soviet Sport, 1989	Published	Low	Botvinnik M.M.	
3.	Paper	"Capital Asset Prices – A Theory of Market Equilibrium Under Conditions of Risk". Journal of Finance. XIX (3): 425–442	Published	Low	William Forsyth Sharpe	https://onlinelibrary.wiley.com/doi/full/10.1111/j.1540-6261.1964.tb02865.x

A.115 Generation of Computer Tomography scans from Magnetic Resonance Images

A.115.1 General

<table>
<tr><td>ID</td><td colspan="2">115</td></tr>
<tr><td>Use case name</td><td colspan="2">Generation of Computer Tomography scans from Magnetic Resonance Images</td></tr>
<tr><td>Application domain</td><td colspan="2">Healthcare</td></tr>
<tr><td>Deployment model</td><td colspan="2">Embedded systems</td></tr>
<tr><td>Status</td><td colspan="2">PoC</td></tr>
<tr><td>Scope</td><td colspan="2">Restoration of naturally distorted microscopy images for following visualization and analysis of meaningful patterns of protein formation inside living cells.</td></tr>
<tr><td>Objective(s)</td><td colspan="2">Create a method for automatic analysis and clustering of cell microscopy images, including microscopy of multilayer 3D objects, and implement the developed method for processing of 2D/3D images of cultured human cell models and clustering based on protein modification patterns</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>Patterns of protein modification inside cells play an important role in the regulation of gene expression. Here, we aim to develop a method allowing for a detailed analysis of the meaningful protein formation inside living cells with visualization and the processing of microscopy cell images. However, the observed microscopy images suffer from visible artifacts related to blurriness and noise. In this work, we aim to implement AI methods throughout the pipeline of microscopy cell image restoration and analysis. Thereafter, we plan to implement AI approaches for the extraction of meaningful patterns of protein modifications inside cells and use this information for effective cell clustering. Our experiments are on 2D images as well as multilayer 3D objects. To the best of the author’s knowledge, this is the first work to apply AI for living cells featuring extraction and clustering.</td></tr>
<tr><td>Complete description</td><td>Patterns of protein modification inside cells play an important role in the regulation of gene expression. In this work we aim to develop a method allowing for a detailed analysis of the meaningful protein formation inside living cells with visualization and the processing of microscopy cell images. However, the observed microscopy images suffer from visible artifacts related to blurriness and noise. One of the main modern approaches to the processing of microscopic images of cell cultures is computer vision using deep learning methods and artificial intelligence (AI). In this work, we aim to implement AI methods throughout the pipeline of microscopy cell images restoration and analysis. The proposed scheme involves the implementation of deep learning methods for image restoration, segmentation, and time and space localization of cells. Thereafter, we plan to implement AI approaches for the extraction of meaningful patterns of protein modifications inside cells and use this</td></tr>
</table>

		information for effective cell clustering. Our experiments are on 2D images as well as multilayer 3D objects. To the best of the author's knowledge, this is the first work to apply AI for living cells featuring extraction and clustering.		
Stakeholders	Biochemical, metabolomics and imaging branches of biomedicine			
Stakeholders' assets, values	Trustworthiness, competitiveness, fair treatment, stability, reputation			
System's threats & vulnerabilities	New privacy threats, new security threats, different source of bias			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Image processing	Restoration, segmentation, and time and space localization of cells	2-D to 3-D processing
	2	Protein modification patterns extraction	Extraction of meaningful patterns	Cell clustering
AI features	Task(s)	Recognition		
	Method(s)			
	Hardware	HPC		
	Topology			
	Terms and concepts used	Machine Learning, Deep Learning, Radiology, Computed Tomography, Magnetic Resonance Imaging		
Standardization opportunities/ requirements				
Challenges and issues	(1) An effective localization of living cells without losing meaningful information must be done; (2) multilayer 3D objects require more computational time and resources, as well as slightly different restoration approaches, due to the 3D object formation model, compared to 2D images			
Societal Concerns	Description	The developed method of analysis of protein modifications inside living cells is applicable to a wide range of biological and biomedical tasks, far beyond the scope of this project.		
	SDGs to be achieved	Good health and well-being for people		

A.115.2 Data

Data characteristics	
Description	EPO-lnternal, PAJ, WPI data, BIOSIS, INSPEC
Source	Human cell data
Type	Images
Volume (size)	
Velocity	Batch
Variety	Different cell cultures
Variability (rate of change)	Static

Quality	MAE: 60.83 HU, PSNR 17.21 dB, SSIM 0.8

A.115.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Patent	WO 2014/070082 Al	Published	Use case is based on this patent	World Intellectual Property Organization, International Bureau	https://patentimages.storage.googleapis.com/24/28/3a/63e3ebeb1f94c3/WO2014070082A1.pdf
2	Article	Keshavamurthy, K. N., Dylov, D. V., Yazdanfar, S., Patel, D., Silk, T., Silk, M., … Durack, J. C. (2019). Spectroscopy and Machine Learning Based Rapid Point-of-Care Assessment of Core Needle Cancer Biopsies	Published			https://doi.org/10.1101/745158

A.116 Generation of Computer Tomography Scans from Magnetic Resonance Images

A.116.1 General

ID	116	
Use case name	Generation of Computer Tomography Scans from Magnetic Resonance Images	
Application domain	Healthcare	
Deployment model	Embedded systems	
Status	PoC	
Scope	Train a model that generates CT images from MRI scans. Synthetic CT image may be used for radiation dose calculation in radiation therapy	
Objective(s)	Generation a CT image from a given MRI image	
Narrative	Short description (not more than 150 words)	Generating radiological scans has grown in popularity in recent years. Here, we generate synthetic Computed Tomography (CT) images from real Magnetic Resonance Imaging (MRI) data. Our architectures were trained on unpaired MRI-CT data and then evaluated on a paired brain dataset. The MRI-CT translation approach holds the potential to eliminate the need for the patients to undergo both examinations and to be clinically accepted as a new tool for radiotherapy planning.

	Complete description	In this project, we investigate approaches to generating synthetic Computed Tomography (CT) images from the real Magnetic Resonance Imaging (MRI) data. Generating radiological scans has grown in popularity in recent years due to its promise to enable single-modality radiotherapy planning in clinical oncology, where the co-registration of the radiological modalities is cumbersome. We rely on Generative Adversarial Network (GAN) models with cycle consistency, which permit unpaired image-to-image translation between the modalities. We also introduce the perceptual loss function term and the coordinate convolutional layer to further enhance the quality of translated images. The Unsharp masking and the Super-Resolution GAN (SRGAN) were considered to improve the quality of synthetic images. The proposed architectures were trained on unpaired MRI-CT data and then evaluated on paired brain dataset. The resulting CT scans were generated with a mean absolute error (MAE), a peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) scores of 60.83 HU, 17.21 dB, and 0.8, respectively. DualGAN, with perceptual loss function term and coordinated convolutional layer, proved to perform best. The MRI-CT translation approach holds the potential to eliminate the need for the patients to undergo both examinations and to be clinically accepted as a new tool for radiotherapy planning.		
Stakeholders	Oncology hospitals, oncologists			
Stakeholders' assets, values	Trustworthiness, competitiveness, fair treatment, stability			
System's threats & vulnerabilities	New privacy threats, new security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	DualGAN architecture	Cycle of two image generators and two discriminators	Unpaired image-to-image translation between modalities
	2	The perceptual loss function	Feature matching, where high-level representations of two images are compared by mean squared error	Features are extracted in an identical way for both compared images
	3	The coordinate convolutional layer (CC)	Concentration of two additional x and y coordinates slices with the tensor	Distinguishing the black pixels of MRI image, which could represent either bone or air
AI features	Task(s)	Prediction		
	Method(s)			

	Hardware	HPC
	Topology	
	Terms and concepts used	Machine Learning, deep learning, radiology, computed tomography, magnetic resonance imaging
Standardization opportunities/ requirements		
Challenges and issues	(1) Large amounts of paired MRI-CT data is not easily available; and (2) doctors are reluctant to accept synthetic CT scans	
Societal Concerns	Description	Savings for oncologic patients. Reduced radiation dosage.
	SDGs to be achieved	Good health and well-being for people

A.116.2 Data

Data characteristics	
Description	Computed Tomography (CT) scans of cancer patients
Source	Patients Magnetic Resonance Imaging (MRI) data
Type	Images
Volume (size)	Three medical datasets; each set was divided into a train and tested in a 7:3 ratio. The first included the MRI T1-weighted images of 7 patients; each 3D volume of a patient contains 22 – 24 slices in the axial anatomical plane. The second cancer dataset consisted of CT scans of 61 patients and 3D volumes include 61 - 94 slices. The third consisted of images of 10 patients; the volumes were 66–137 slices.
Velocity	MRI-to-CT image translation: unpaired image-to-image translation between the modalities, perceptual loss function term, coordinate convolutional layer, training on unpaired MRI-CT data, evaluation on paired brain dataset
Variety	DualGAN architecture: constant comparison of the reconstructed image and original to evaluate the quality of generators without the need of paired data
Variability (rate of change)	Moderate
Quality	MAE: 60.83 HU, PSNR 17.21 dB, SSIM 0.8

A.116.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Brochure	D. Prokopenko, J.V. Stadelmann, H. Schulz, S. Renisch & D.V. Dylov. “Synthetic CT Generation from MRI Using Improved DualGAN”. Medical Imaging with Deep Learning 2019. Accepted 06 May 2019.	Published	High	Medical Imaging with Deep Learning 2019 conference, London	https://openreview.net/pdf?id=S1em7ZOkFN

2	Paper	Prokopenko, D., Stadelmann, J. V., Schulz, H., Renisch, S., & Dylov, D. V. (2019). Unpaired Synthetic Image Generation in Radiology Using GANs. In Artificial Intelligence in Radiation Therapy (pp. 94–101).	published	high	Medical Image-to-Image Translation in Radiology, Philips Innovation Labs RUS	https://doi.org/10.1007/978-3-030-32486-5_12
3	Abstract	Denis Prokopenko, Jo¨el Valentin Stadelmann, Heinrich Schulz, Steffen Renisch, Dmitry V. Dylov. Unpaired Synthetic Image Generation in Radiology Using GANs	published	High	MICCAI, Shenzhen, China	https://www.miccai2019.org/wp-content/uploads/2019/10/MICCAI-Programme-Book-for-web-1.pdf
4	Relater paper	National Cancer Institute Clinical Proteomic Tumor Analysis Consortium CPTAC. Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme [CPTAC-GBM] collection [Data set]. The Cancer Imaging Archive.	Published	High	National Cancer Institute	https://doi.org/10.7937/k9/tcia.2018.3rje41q1, 2018.
5	Related paper	Rosanne Liu, Joel Lehman, Piero Molino, Felipe Petroski Such, Eric Frank, Alex Sergeev, and Jason Yosinski. An intriguing failing of convolutional neural networks and the coordconv solution.	arXiv preprint	high		arXiv:1807.03247, 2018

A.117 Improving the knowledge base of prescriptions for drug and non-drug therapy and its use as a tool in support of medical professionals

A.117.1 General

ID	117
Use case name	Improving the knowledge base of prescriptions for drug and non-drug therapy and its use as a tool in support of medical professionals
Application domain	Healthcare
Deployment model	Cloud services
Status	Prototype
Scope	Providing the medical professional with methods and means that will allow, within the time allotted for the appointment of a patient with a known nosology,

	to make a high-quality choice of drugs and to formulate a prescription corresponding to "good medical practices"	
Objective(s)	Helping a medical professional consider the influence of a selected drug therapy, as well as monitor the patient's vital characteristics to reduce the risk of wrong prescriptions and to prevent negative consequences from the prescribed drugs	
Narrative	Short description (not more than 150 words)	Services are developed designed to improve the efficiency and quality of medical care in third-level medical organizations, which have in their structure units providing high-tech medical care. A knowledge base of prescribed drug and non-drug therapy was formed based on the RLS® database. For its improvement and scaling throughout the industry, it is advisable to use AI methods.
	Complete description	The complexity of choosing an optimal drug therapy can be illustrated by the example of a great number of possible combinations that arise when considering a nosology such as "arterial hypertension" (hypertension and Hypertensive diseases, ICD-10 version 2016: I10-I15) ... The basic factors initially influencing the choice of therapy for hypertension = 6 (gender male and female, as well as 3 gradations of age). The next step is to establish a correspondence between the vital characteristics (VC) of the patient and the specific features of the use of a drug. An informational "portrait" of a patient can be compiled using trivial and composite VC (currently, more than 500 already exist). Considering the individual characteristics of the patient (comorbidity, data from laboratory and instrumental methods of research, genetic factors, eating habits, etc.), the number of VCs can be increased by orders of magnitude. Associated hypertension of nosologies and conditions that have a specific section in the existing CR = 17. Clinical recommendations (CR) in the framework of concomitant nosologies – more than 15 (it is impossible to say for sure, because the lack of specificity by sections of the CR makes it impossible to determine the total number of CR). Pharm group (FG) of drugs = 25 (8 groups of antihypertensive drugs + 17 groups of other drugs, for example, used in the treatment of concomitant nosologies that increase blood pressure. Active substances (AS) = 72 (36 antihypertensive + 15 other used in the treatment of concomitant nosologies, for example, antidiabetic or AS, which increase the blood pressure + 21 antihypertensive and others, whose names are not in the CR, but are included in the FG mentioned in the CR). Fixed combinations considering different dosages = 45. And the number of instructions for medical usage of drugs (IMU), information from which must be considered = 218. In total, for every one considered nosology there are thousands of pages of text and tens of thousands of parameters to one degree or another, directly or indirectly connected, and sometimes even in contradiction. A single mistake poses a negative outcome.
Stakeholders	Doctors and Patients	
Stakeholders' assets, values	Doctor's reputation, patient health	
System's threats & vulnerabilities	Incorrect AI system use (AI system affecting quality control)	

Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	C_CR	Conformity_CR	Comply with CR	Improve accuracy
	C_IMU	Conformity_IMU	Comply with IMU	Improve accuracy
AI features	Task(s)	Recommendation		
	Method(s)	Classification and Categorization		
	Hardware	Cloud infrastructure (e.g. Microsoft Azure)		
	Topology			
	Terms and concepts used	Classification, Categorization		
Standardization opportunities/ requirements	It is necessary to consider the difference in regulations governing the use of CR and IMU			
Challenges and issues	1. The existence in parallel of several CR used by doctors. 2. The difference in the information of CR and IMU. 3. The need for complementing the information of CR and IMU. 4. The discrepancy between the information of CR and the real situation in the pharmaceutical market.			
Societal Concerns	Description	The widespread use of the solution will allow the doctor: — develop competencies in the field of drug selection, considering VC and drug interactions when prescribing; — reduce the risks of erroneous prescriptions; — improve the quality of medical care In the end, this will allow: — preserve the health of the patient, and of their loved ones; — extend the quality of a full life		
	SDGs to be achieved	Good health and well-being for people		

A.117.2 Data

Data characteristics	
Description	Rules for prescribing drug and non-drug therapy
Source	All information used in the services' databases contains only digitalized information from the texts of IMU approved by the Ministry of Health and approved for use and CR prepared by professional communities.
Type	Structured / unstructured text
Volume (size)	
Velocity	
Variety	
Variability (rate of change)	

Quality	Completeness and accuracy of the data with respect to semantic content as well as syntax of the data (such as presence of missing fields or wrong meanings)

A.117.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Patent	Method for effecting computer implemented decision-support in the selection of the drug therapy of patients having a viral disease (US7010431B2 dated March 7, 2006)	Published	Use case taken from this reference		https://patents.google.com/patent/US7010431B2
2	Patent	Optimization and individualization of medication selection and dosing (WO2007064675A2 dated June 7, 2007)	Published	Use case taken from this reference		https://patents.google.com/patent/WO2007064675A2
3	Patent	Medical risk assessment method and program product (US7306562B1 dated December 11, 2007)	Published	Use case taken from this reference		https://patents.google.com/patent/WO2007064675A2
4	Certificate on state registration	Database «Basic terminological dictionaries of vital characteristics v 1.0» (2019621394 dated July 30, 2019)	Published	High	Federal Service for Intellectual Property of the Russian Federation (RosPatent)	http://new.fips.ru/registers-doc-view/fips_servlet?DB=DB&DocNumber=2019621394
5	Certificate on state registration	Certificate on state registration of the database «Conditions for the applicability of drugs in terms of vital characteristics v 1.0» (20169620990 dated June 5, 2019)	Published	High	Federal Service for Intellectual Property of the Russian Federation (RosPatent)	http://new.fips.ru/registers-doc-view/fips_servlet?DB=DB&DocNumber=2019620990

A.118 Intelligent Technology to Control Manual Operations on Video — “Norma”

A.118.1 General

<table>
<tr><td>ID</td><td colspan="2">118</td></tr>
<tr><td>Use case name</td><td colspan="2">Intelligent Technology to Control Manual Operations on Video — “Norma”</td></tr>
<tr><td>Application domain</td><td colspan="2">Manufacturing</td></tr>
<tr><td>Deployment model</td><td colspan="2">On-premise systems</td></tr>
<tr><td>Status</td><td colspan="2">Prototype</td></tr>
<tr><td>Scope</td><td colspan="2">Tooltip visualization technology (augmented reality) based on technological process and manual operations control in the assembly, maintenance, and repair of engineering products.</td></tr>
<tr><td>Objective(s)</td><td colspan="2">“Norma” technology will reduce the number of errors made by technical personnel during manual assembly of products to the lowest possible minimum. It visualizes the correct sequence of actions to the user-assembler on top of the parts through augmented reality glasses. Norma controls the correctness of manual operations and the tool used. It fixes the detected deviations in the electronic passport of the product. Additionally, Norma promptly reports identified violations of the process to the quality control department. Norma will provide a dramatic improvement in the quality of production and technological operations without the widespread use of industrial robotics, which will avoid the negative social consequences caused by automation of production.</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td>The Norma technology is designed to control manual operations during assembly, maintenance, and repair of engineering products using video data.</td></tr>
<tr><td>Complete description</td><td>The quality of assembly, maintenance, and repair of engineering products substantially depends on the number of errors made during manual operations. According to data from various sources (railway, nuclear, aviation, and other industries), the percentage of failures caused by a violation of maintenance and production technologies by technical personnel reaches 54%.

The mistakes made during assembly, maintenance, and repair of engineering products are based on the following problems:
- lack of constant (objective) control of manual operations at all stages of assembly, maintenance, and repair of engineering products;
- lack of reliable data on the operations performed and the mistakes made throughout the life cycle of the product, which makes it difficult to exclude errors from the design of the product.

“Norma” technology is designed to control manual operations during assembly, maintenance, and repair of engineering products. “Norma” technology will monitor compliance with the process through the analysis of video data and record every step in the electronic passport of the product.</td></tr>
</table>

<table>
<tr><td></td><td></td><td colspan="3">Engineers performing maintenance and repair or operator-assemblers will wear augmented reality (AR) glasses which will show all information about the technological process and step-by-step instructions. In these glasses, all parts of engineering product will be recognized and marked up with bounding boxes showing which parts are required at each step of technological process.

The AR glasses user will be notified about errors made during the technological operation so they can fix the problem before proceeding to the next step. The AR glasses will record and store video of technological process performed and create electronic passport of engineering product.

The engineer will train machine learning models included in the "Norma" technology in an automated manner using 3D CAD models of engineering products and descriptions of technological process.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Industrial enterprises, repair enterprises, repair shops, operators of engineering products.</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Improving the quality and reducing the number of errors due to the fault of technical personnel in the assembly, maintenance, and repair of engineering products.</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">There is a risk of video leaking during the process.
Norma technology is installed in the internal network of the enterprise and the safety of all materials is ensured by the enterprise's security policy.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Detection of parts, assemblies and products.</td><td>To assess the quality of detection, the mAP@0.5 metric is used</td><td>Monitoring the assembly process and maintenance of engineering products</td></tr>
<tr><td>2</td><td>Classification of manual operations</td><td>To assess the quality of classification of manual operations, the accuracy metric is used</td><td>Monitoring the assembly process and maintenance of engineering products</td></tr>
<tr><td rowspan="3">AI features</td><td>Task(s)</td><td colspan="3">Other (please specify):
training on datasets synthesized on the basis of 3D CAD models of product parts;
detection of parts, assemblies and the product as a whole;
hand tool detection;
classification of manual operations;
automatic step detection of a technological process;
automatic control of the correct assembly of the product.</td></tr>
<tr><td>Method(s)</td><td colspan="3">Deep learning, Convolutional Neural Networks, Domain Randomization, Action Recognition, Object Detection</td></tr>
<tr><td>Hardware</td><td colspan="3">GPU server, AR-glasses</td></tr>
</table>

	Topology	GPU server, AR-glasses
	Terms and concepts used	Machine Learning, Computer vision, Human-machine teaming, AI system, Convolutional / deep convolutional neural networks, Domain Randomization
Standardization opportunities/ requirements	Desirable to standardize input formats of CAD models and technological process descriptions.	
Challenges and issues	1. Small (or none) number of real photos for training — neural networks shall be trained on a synthetic data 2. Synthetic data shall be generated to cover all possible light conditions in which system can be used 3. System shall operate in real time	
Societal Concerns	Description	Norm technology will provide quality improvement in production without the use of robotic systems, which will not lead to a reduction in jobs and will therefore avoid negative social consequences
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.119 Loan in 7 minutes

A.119.1 General

ID	119	
Use case name	Loan in 7 minutes	
Application domain	Banking and Financial Services	
Deployment model	On-premise systems	
Status	In operation	
Scope	A completely automated solution which analyzes customer behavior and makes loan offers best for the customer	
Objective(s)	Create lending product for clients of medium and large businesses (LMB) with the shortest delivery time possible taking into account the extremely detailed customer profile	
Narrative	Short description (not more than 150 words)	Loan in 7 minutes is the first solution in the world where the credit decision is made by artificial intelligence without human participation in just a few minutes. A complex machine learning settlement system was implemented on one of the largest Hadoop-cluster in Eastern Europe (tens of petabytes of data) and integrated into the business process of corporate lending of the Bank. The new project has significantly improved customer experience: — eliminated the need for the client to contact the Bank in person for a loan;

<table>
<tr><td></td><td></td><td colspan="3">— requires no additional documents from the client to get a decision;
— Bank's automated systems were improved in terms of automatic transaction creation;
— substantially simplified the process of issuing a loan.</td></tr>
<tr><td></td><td>Complete description</td><td colspan="3">If the client requires a loan he fills out a short form in the Bank online system to reflect the recent changes of the business. As soon as the client provides necessary information the solution kicks in.

It interacts with the internal (e.g. transactional data) and external (e.g. credit bureaus) systems, collects all detailed information about the client, applies algorithms based on artificial intelligence and machine learning methods, automatically performs risks estimation and calculates appropriate offers for clients.

The client chooses appropriate lending terms. The solution calculates interest rate, generates electronic version of credit documentation and sends it to the client via web interface.

Along with the terms of the loan the list of legal documents which should be requested from the customer for the deal to succeed is formed. The function of a legal officer is performed automatically by the Robot Lawyer which does the same set checks on client documents as a human lawyer in a standard credit process would do.

The client signs the documentation using his electronic certificate. The signature applied has full legal force and may be verified automatically in a certificate authority.

The loan conditions chosen by the client are reflected in the Bank's internal accounting system.

The speed of the decision making on a loan application in the solution is unprecedented and it is important step in the development of corporate lending in Russia for LMB segment.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Customers</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Fair treatment</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Non-performing loans ratio</td><td>Ratio of a sum of borrowed money upon which the</td><td>Improve efficiency</td></tr>
</table>

			debtor has not made the scheduled payments for a specified period to the total loans	
	2	Time to decision	Minutes for the generating appropriate loan offers	Shorten delivery time
AI features	Task(s)	Natural language processing, Decision Making, Graph		
	Method(s)	— NLP: Neural Networks CNN + bi-LSTM, BERT + Attention & Few-shot Learning (Proto-NER) — Decision Making for loan approval: NN, XGBoost, LogReg + L1/L2 regularization — Graph: investigation of companies influence on each other to consider it in decision making		
	Hardware			
	Topology			
	Terms and concepts used	Segmentation, Embedding, Boosting, Ensembles		
Standardization opportunities/ requirements	Standardization needs for setting up this use case is currently under further investigation.			
Challenges and issues	Non-linear models based on big data need significant computational power during the training phase			
Societal Concerns	Description	Investment in technological innovation and infrastructure are crucial drivers of higher levels of productivity and economic growth		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

A.119.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press release		Published	High	Sberbank	https://www.sberbank.ru/en/press_center/all/article?newsID=3d7cd460-ae60-48a9-a4f0-4f78578a6988&blockID=1539&regionID=77&lang=en
2	Press release		Published	High	Sberbank	https://www.sberbank.com/news-and-media/press-releases/article?newsID=421abe14-3082-4969-8f93-

						90a81e656885&blockID=7&regionID=77&lang=en

A.120 AI Contract Management

A.120.1 General

ID	120	
Use case name	AI Contract Management	
Application domain	Legal	
Deployment model	On-premise systems	
Status	In operation	
Scope	Building an AI Contract Management solution for the business process of documents automation: data classification, automatic data extraction and contract monitoring.	
Objective(s)	Creating a solution that is able to standardize contract management process, improve quality of work on problematic contracts and claims and optimize lawyers' working process and relieve them from routine tasks.	
Narrative	Short description (not more than 150 words)	MTS AI Contract Management solution is built on our AI legal core, which includes technology that enables to convert different types of documents into digital format, replicate the natural human-like text recognition and extract data to automate business tasks.
	Complete description	It's a platform for automatic reading and analysis of legal documents, extraction of data with astonishing high level of accuracy. Based on the extracted data automatic contract monitoring and execution can be performed. The following features of the AI Contract Management can be highlighted: — Structured digital documents archive, — Hierarchical chain and connections of all documents in relation to the primary document, whether it is a contract, order or anything else, — Monitoring and control of key contract terms, — Creation of all necessary documents: notifications, claims, etc. — Autofilling the required ERP systems with relevant data.
Stakeholders	Procurement department, legal department	
Stakeholders' assets, values	Acceleration and rising quality of legal operations and processes	
System's threats & vulnerabilities	Security threats, privacy threats. Usually contracts contain trade secrets, disclosure of which can lead to serious financial losses. For this reason, the solution operates in a closed client protected form.	

	Bias due to changes in requirements on the customer's end or inappropriate training data.			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Recall	Also known as sensitivity is the fraction of the total amount of relevant instances that were actually retrieved	
	2	Precision	Also called positive predictive value is the fraction of relevant instances among the retrieved instances	
	3	Customer Satisfaction	The ratio of customer satisfaction when using this system for requests. The expectation is 100%	Increasing its ratio as high as possible
	4	Algorithm accuracy	Output when compared to the human expert analysis of the same data	
	5	Task completion rate	The performance is calculated by dividing the number of cases that have been completed successfully by the total number of assigned tasks. The success or failure of a task is set according to the criteria of each system.	Accurate task completion using the AI system
	6	Cost	Minimize the financial costs and reduce the risk of penalties under the contracts	
	7	Efficiency	Improve the efficiency of existing manual	

			document processing	

AI features	Task(s)	Contract Management
	Method(s)	OCR, NLP and Knowledge representation, NLU, Neural networks, Machine Learning, CV
	Hardware	40 CPU, 80 Gb RAM, SSD ~3.9 Tb
	Topology	
	Terms and concepts used	Data Classification, Information Extraction, Computer Vision, Natural Language Processing, Image Segmentation
Standardization opportunities/ requirements	The opportunity to bring the working process on a contract to a single standardized format (meta-document process) with ability to extract a key data set.	
Challenges and issues	1. Noisy data (different scans quality) 2. Working with private data (information security) 3. Non-linear models need significant computational power during the training phase	
Societal Concerns	Description	We create the helpful industrial solution that can optimize the current contract management process and assist to make easier the legal departments job
	SDGs to be achieved	Industry, Innovation, and Infrastructure

A.120.2 Data

Data characteristics	
Description	Different type of documents: contracts, additional agreements, NDA, etc
Source	DW (Data Warehouses)
Type	Structured/unstructured text, images
Volume (size)	
Velocity	Real time in production phase
Variety	Different types of source with mostly structured data
Variability (rate of change)	Moderate
Quality	Moderate

A.120.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1.		The project (JureCloud) was included in the PwC LegalTech Research				https://www.pwc.ru/ru/services/legal-services/news-archive/legal-tech-russia.html https://www.pwc.ru/ru/services/pwc-legal-tech-map-ru.pdf

						https://www.kommersant.ru/doc/3744362
2.		Inhouse MTS departments (Procurement, Legal)	In operation	Cost saving by monitoring key terms and data	Inhouse MTS departments	

A.121 Neural Network Formation of 3D-models orthopedic insoles

A.121.1 General

ID	121	
Use case name	Neural Network Formation of 3D-models orthopedic insoles	
Application domain	Healthcare	
Deployment model	Client and server systems	
Status	In operation	
Scope	Artificial intelligence methods using to construction of individual medical products to reduce the risk of developing diseases of the musculoskeletal system	
Objective(s)	Development of comfortable, individualized, anatomically correct orthopedic 3D insoles for the treatment of flat feet	
Narrative	Short description (not more than 150 words)	Using artificial intelligence methods, the system converts a pre-scanned foot print into an innovative, medically-based 3D-insole. The AI-system will independently make a medical decision based on the collected medical history, and anthropometric data. Initial training of the AI-system will take place together with the doctor. In the future, the system will begin by independently choosing the most suitable location options for a patient vaults and indentations and plan an anatomically correct and secure 3D-insole.
	Complete description	The system consists of two parts, hardware and software. The hardware scans 3D / 2D foot images patients and receives a production file format ready for loading into a specialized machine or a 3D printer. In the software, a local orthopedic 3D model of the insole is formed according to a unique author's technique using a local software package based on artificial intelligence. The received data is stored on a cloud platform. 3D-method allows more accurately to orthose complex pathologies and atypical deformations due to the used sophisticated equipment and accurate removal of anatomical physiological parameters of the foot up to

		10,000 points per 1 sq. sm. The patient's foot is scanned in the sitting position; it is not exposed to loads; the 3D laser scanning is 6 cm high, which allows for the obtaining of full-color 3D models of the patient's legs with an accuracy of half a millimeter. Further automatic milling is highly accurate for orthopedic shoes. The process of creating insoles is completely autonomous, personalized, and does not require the intervention of an orthopedic doctor. Overall, the system is modularized with capabilities to self-learn and for future extensions.		
Stakeholders	Medicine, public sector			
Stakeholders' assets, values	Improving the quality of life			
System's threats & vulnerabilities	Incorrect AI system use			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Individualized, anatomically correct orthopedic 3D insoles	Local orthopedic 3D model of the insole is formed	Reducing the risk of developing diseases of the musculoskeletal system
AI features	Task(s)	Construction		
	Method(s)	Neural Networks		
	Hardware	3D printer, scanner, cloud platform		
	Topology			
	Terms and concepts used	Classification, feature extraction, anatomically correct orthopedic 3D insoles		
Standardization opportunities/ requirements	Tolerance criteria for predicted product characteristics			
Challenges and issues	None identified			
Societal Concerns	Description	None identified		
	SDGs to be achieved	Good health and well-being for people		

A.121.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	The 3D-printing advantage for foot orthotics	Published	analogues	Dr. Bruce Williams, DPM	https://www.fitstation.com/http://h20195.www2.hp.com/v2/GetDocument.aspx?docname=4AA7-5747ENW

2		Footscan - plantar pressure measurement product	Web site	analogues	RSscan International NV	https://rsscan.com/footscan/
3	Publication	HISTORICAL BACKGROUND OF THE DEVELOPMENT OF BIPEDIC MOVEMENT (WALKING)	Published	Research	Polukarov N.V., Achkasov E.E.	https://rucont.ru/efd/375087
4	Publication	INFLUENCE OF THE INDIVIDUAL APPROACH OF CONSERVATIVE THERAPY OF PLANOSCOPY ON THE REDUCTION OF PAIN SYNDROME AND IMPROVEMENT OF THE QUALITY OF LIFE OF PATIENTS	Published	Research	Zhukova E.V., Achkasov E.E., Polukarov N.V.	http://vvmr.ru/about/svezhiy-nomer/
5	Publication	INFLUENCE OF WALKING BIOMECHANICS ON THE FORMATION OF STOP PATHOLOGY	Published	Research	Zhukova EV, Achkasov EE, Polukarov NV, Gridin LA, Osadchuk MA, Puzin S.N.	http: //www.phdynasty.ru/katalog/zhurnaly/voprosy-prakticheskoy-pediatrii/2018/tom-13-nomer-4/34305

A.122 Open spatial dataset for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data

A.122.1 General

ID	122
Use case name	Open spatial dataset for developing AI algorithms based on remote sensing (satellite, drone, aerial imagery) data
Application domain	Other (please specify) earth science, digital cartography
Deployment model	On-premise systems
Status	In operation
Scope	Analytical services for automatic detection of changes of the state of ground surface objects for administrative, government, and social purposes in different use-cases, such as: — Urban monitoring: cadastral data, land management, estimation of the living population etc. — Emergency mapping: estimation of disaster damages

<table>
<tr><td></td><td colspan="4">— Security and risk management monitoring of protected zones (powerlines, railroads, pipelines): detection of vegetation growth, control of the safety etc.</td></tr>
<tr><td>Objective(s)</td><td colspan="4">The growth of the Russian market of geo-analytical cloud-services based on remote sensing data and AI technologies; open benchmark datasets for the R&D community; and bringing the power of AI and the global coverage of remote sensing imagery closer to the people.</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Despite the increasing number of datasets and competitions in remote sensing data science (e.g. Spacenet) there is still a lack of geographical diversity, of training classes, and of interoperability of datasets.
The proposed approach is to be extended to different types of remote sensing data and application domains based on classification of the natural and man-made objects that have a clear interpretation either in satellite or aerial imagery.</td></tr>
<tr><td>Complete description</td><td colspan="3">Despite the increasing number of datasets and competitions in remote sensing data science (e.g. Spacenet) there is still a lack of geographical diversity, of training classes, and of interoperability of datasets.
The proposed approach is to be extended to different types of remote sensing data and application domains based on classification of the natural and man-made objects that have a clear interpretation either in satellite or aerial imagery.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Community</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="4">Trustworthiness, safety, competitiveness</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">New privacy and security threats, challenges to accountability</td></tr>
<tr><td rowspan="5">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Georeference</td><td>Maps (e.g. Openstreetmap) for data labeling require objects' coordinates</td><td>Simply annotated photos are not enough</td></tr>
<tr><td>2</td><td>Time series</td><td>Emergency Mapping requires the detection of changes in residential infrastructure analysis before and post-event images</td><td>To observe places in dynamic and calculate comparative indicators</td></tr>
<tr><td>3</td><td>Cartographic styled labeling and classification</td><td>Competition of network with manual mapping</td><td>Maps make an abstracted interpretation of Earth observation images</td></tr>
<tr><td>4</td><td>Advanced classification</td><td>The help of different bands combination</td><td>Thematic interpretation of satellite imagery</td></tr>
</table>

<table>
<tr><td></td><td>5</td><td>Open API and web tools</td><td>Integrate both mapping and data science approaches in ways demanded by users</td><td>To access and preview</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Recognition</td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3">HPC</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">Machine learning, mapping, open spatial dataset, recognition, remote sensing</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">There is no standard or criteria regulated the process of labelling (manual or automatic) remote sensing (satellite, drone or UAV) images with geographic reference. Development of such a standard is vital to AI algorithms as for guarantees of the quality of training data and for testing and benchmarking.
We consider the following criteria the perfect dataset collection for EO imagery should match:
— 1) Georeference. Simply annotated photos are not enough. Maps for data labeling (e.g. Openstreetmap) require objects’ coordinates.

— 2) Time series. To observe places in dynamic and calculate comparative indicators. The main application is “Emergency Mapping” where the detection of changes in residential infrastructure analysis of before and post-event images is required.

— 3) Cartographic styled labeling and classification. Maps make an abstracted interpretation of Earth observation images; we therefore, believe that the previous approach of labeling images with boxes does not satisfied the criteria for accurate image segmentation and won’t work. For neural networks it’s now necessary to compete with manual mapping and to calculate its accuracy we need at least some Ground Truth that looks like a map.

At the same time there are many other sources beyond the EO imagery that might be useful for mapping, such as POI*, collecting field works in order to accumulate addresses. At this moment our goal is to compare ML methods with the information that could be extracted by a cartographer using only optical bands of imagery and some GIS* software. For such purposes we proposed the basic classifier that is at the part of training and testing datasets.

— 4) Multispectral. Next, we assume to extend this approach to advanced classification which is comparable to thematic interpretation of satellite imagery with the help of different bands combination. That’s why the proposed classifier includes classes which require even more specific training and non-optical bands for better recognition.

Providing Open API and web tools to access and preview datasets. Despite the dataset collection representing structured data, it would be much more capable</td></tr>
</table>

	for further and updated use based on the standards for interoperability of geodata. In our work, we tried to join both mapping and data science approaches in a way we see new tools and services demanded by users. For many users from the data science community, maps and remote sensing are becoming just one of the sources of information that must be structured and classified. And for many mappers that are involved in the process of geodata interpretation and classification, the map itself is the perfect tool to interact with the data; no matter whether implemented in python notebook or loaded in a desktop GIS application.	
Societal Concerns	Description	Global extension of this technology brings society new possibilities of situational awareness and digital instruments for natural and man-made resource management
	SDGs to be achieved	Sustainable cities and communities

A.122.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	V. Ignatiev, A. Trekin, V. Lobachev, G. Potapov, and E. Burnaev. Targeted change detection in remote sensing images	Published		Proc. SPIE 11041, 11th International Conference on Machine Vision (ICMV 2018), 110412H (15 March 2019)	doi: 10.1117/12.2523141
2	Paper	Novikov G., Trekin A., Potapov G., Ignatiev V., Burnaev E. (2018) Satellite Imagery Analysis for Operational Damage Assessment in Emergency Situations. In: Abramowicz W., Paschke A. (eds) Business Information Systems	Published	High	BIS 2018. Lecture Notes in Business Information Processing, vol 320, pp. 347-358. Springer, Cham.	https://doi.org/10.1007/978-3-319-93931-5_25
3	Software	"Program for Protected Areas Monitoring "		High	Registration Certificate No. 2019662525,	https://aeronetlab.space/
4	Data repository	“Open spatial dataset”		High		https://github.com/aeronetlab/open-datasets

5	Press release	“Buildings height estimation”			medium.com	https://medium.com/geoalert-platform-urban-monitoring/buildings-height-estimation-7babe6420893
6	Press release				Medium.com	https://medium.com/geoalert-platform-urban-monitoring/buildings-damaged-in-florida-ef1f2089c8c7
7	Press release				Medium.com	https://medium.com/geoalert-platform-urban-monitoring/moscow-surface-parking-how-large-is-the-free-parking-space-and-whats-the-occupancy-616ac46c9a8f
8	Press release				Medium.com	https://medium.com/geoalert-platform-urban-monitoring/%D0%BA%D0%B0%D1%80%D1%82%D1%8B-%D0%B8-%D0%BD%D0%B0%D0%B2%D0%BE%D0%B4%D0%BD%D0%B5%D0%BD%D0%B8%D1%8F-9c30a98a6351

A.123 Optimization of ferroalloy consumption for a steel production company

A.123.1 General

ID	123	
Use case name	Optimization of ferroalloy consumption for a steel production company	
Application domain	Manufacturing	
Deployment model	Embedded systems	
Status	PoC	
Scope	Recommendation for the optimal consumption of ferroalloys at ladle furnace treatment during secondary steelmaking	
Objective(s)	Reducing the usage of ferroalloys in metallurgical plants while maintaining alloy quality standards for steel. Improving production efficiency	
Narrative	Short description (not more than 150 words)	Digital advisor in steel ladle treatment. Recommends the optimal consumption of ferroalloys at ladle furnace treatment during secondary steelmaking. The solution is based on physico-chemical technological models and machine learning models. Datana Smart uses historical data, different factors and correlations, with high accuracy based on real dependencies on the physical process.
	Complete description	Datana Smart’s application area concerns manufacturing process optimization. The solution increases equipment

		productivity, largely removes the human factor, and reduces energy and material resource consumption. Joint usage of physico-chemical technological models and machine learning models cancels mutual disadvantages and strengthens the advantages of the models. Datana Smart uses historical data, including: — Steel grades specifications — Results of chemical analyses — Chemical composition requirements and standards for ferroalloy use		
Stakeholders	Steelmaking, steel industry			
Stakeholders' assets, values	Competitiveness, quality check			
System's threats & vulnerabilities	Different sources of bias, incorrect AI system use			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Economical efficiency	Ratio of a unit cost of recommended ferroalloys to the unit cost of ferroalloys used without recommendation. Satisfying result is 0.97 or less	Reduce the usage of ferroalloys.Improve production efficiency
AI features	Task(s)	Optimization		
	Method(s)	Machine learning models, physico-chemical technological models		
	Hardware			
	Topology			
	Terms and concepts used	Machine learning, big data		
Standardization opportunities/ requirements	Quality acceptance criterion from AI systems: prediction of a chemical composition of steel in the case of implementation of the recommendations should be equal to 95% or more			
Challenges and issues	There is no data available for creating mathematical models. Incorrect/insufficient data; outliers, gaps, accumulated errors, and inaccurate measurements.			
Societal Concerns	Description	Promoting sustainable industries, and investing in innovation, are important for facilitating sustainable development		
	SDGs to be achieved	Industry, Innovation, and Infrastructure		

4266

A.123.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	"Industry Components 4.0: Artificial Intelligence". Rational Enterprise Management. p. 20-36 No.1-2/2019	Published	Average	Magazine Rational Enterprise Management	http://www.remmag.ru/upload_data/files/2019-0102/RT.pdf?fbclid=IwAR1Qd8s5fXcvGitBgZzB5NLdUCfl2_r4CMxfc84O_gz6Rws7mdcxZMIfZjA
2	Presentation	Company whitepapers and presentations	Published	Average	Datana	https://yadi.sk/i/bTTwgc9ZUGwopg
3	Press release	Press release 03.04.2019 Vedomosti	Published	Average	Vedomosti	https://www.vedomosti.ru/press_releases/2019/04/03/kompaniya-datana-pomozhet-promishlennikam-sekonomit

A.124 AI Adaptive Learning Mobile App

A.124.1 General

ID	124	
Use case name	AI Adaptive Learning Mobile App	
Application domain	Education	
Deployment model	Hybrid or other (mobile app)	
Status	In operation	
Scope		
Objective(s)	Providing easy, convenient and adaptive learning of English with the help of a virtual teacher based on artificial intelligence	
Narrative	Short description (not more than 150 words)	A mobile application for learning English, which is based on a program that adapts content to the student and learns with them. During registration, the program analyzes the user's account on a social network and draws up an individual training plan based on the student’s interests
	Complete description	The application analyzes successes and develops a curriculum adapted for each user (2). The user is required to first indicate their level of knowledge of the language and follow the instructions of the virtual teacher.

		The program pays more attention to developing vocabulary and learning grammar rules. Notably, the program collects various information while the student interacts with it, including the user's training rate, the percentage of correct and erroneous answers, how well the user knows and understands various grammar rules, etc. By collecting this information, the application can tailor different activities to meet goals that have already been achieved and those toward which the student still wants to strive. The virtual teacher suggests choosing a level of difficulty, and then monitors the execution of tests and tasks, analyzing errors. If the student cannot cope, it offers to repeat the material. Solves without errors - skips on (1)		
Stakeholders	All age groups with a goal to learn a foreign language			
Stakeholders' assets, values	Learning interest, effectiveness of acquiring new knowledge, the involvement in the educational process through gamification			
System's threats & vulnerabilities	Teaching effect of virtual teacher			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Efficiency	Improve student's learning effect through an adaptive learning format	Improve efficiency
	2	Interest	Improve students' interest in learning	Improve involvement
AI features	Task(s)	Optimization		
	Method(s)	In-depth study of user actions and user information		
	Hardware			
	Topology			
	Terms and concepts used	The application is based on a program that adapts to the student and learns with them on the basis of the database of data		
Standardization opportunities/ requirements	Use of a virtual teacher in the educational process enables the analysis of student actions and builds individualized learning tracks based on the data received			
Challenges and issues	The development of a personalized approach to learning			
Societal Concerns	Description	This case of the use of artificial intelligence in the educational process can complement teachers as knowledge transmitters and make education accessible to everyone. At the same time, artificial intelligence, performing the functions of analytics, packaging and personalization of educational content, is much more effective than a person in the role of an assistant to a teacher and shifts the role of a classical teacher towards mentoring.		
	SDGs to be achieved	Quality education		

4271

A.124.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	The article of RBC (Russian Business Consulting) about the mobile application with using artificial intelligence		Valid	Parla overview	Parla	https://www.rbc.ru/own_business/20/09/2017/59c25e659a7947f26210ac80

A.125 Predictive analytics for the behavior and psycho-emotional conditions of eSports players using heterogeneous data and artificial intelligence

A.125.1 General

ID	125	
Use case name	Predictive analytics for the behavior and psycho-emotional conditions of eSports players using heterogeneous data and artificial intelligence	
Application domain	Other (please specify) eSports	
Deployment model	Cyber-physical systems	
Status	Prototype	
Scope	Prediction of psycho-emotional conditions of eSports players. To form predictions, we collect the physiological data from wearables/video cameras/eye tracker, game telemetry data from keyboard/mouse/demo files, and environmental conditions followed by the application of machine learning methods for the analysis of the collected data.	
Objective(s)	Predict psycho-emotional conditions of eSports players in particular game scenarios based on collected heterogeneous data	
Narrative	Short description (not more than 150 words)	eSports is organized video gaming, where single players or teams compete against each other with the aim of achieving a specific goal by the end of the game. The eSports industry has progressed considerably within the last decade: a huge number of professional and amateur teams take part in numerous competitions where the prize pools amount to tens of millions of dollars USD. Its global audience has already reached 380 million in 2018 and is expected to reach more than 550 million in 2021. However, a lack of tools exists to help assess the physiological and psycho-emotional conditions of eSports players. In this project, we collect three classes of data (physiological, game telemetry, and environmental conditions) followed by a data analysis using artificial intelligence based on machine learning algorithms. For example, we apply machine learning and recurrent neural networks with attention to assessing player performance dynamics.

	Complete description	eSports is organized video gaming where single players or teams compete against each other with the aim of achieving a specific goal by the end of the game. The eSports industry has progressed considerably within the last decade: a huge number of professional and amateur teams take part in numerous competitions where the prize pools achieve tens of millions of dollars USD. Its global audience has already reached 380 million in 2018 and is expected to reach more than 550 million in 2021. However, a lack of tools exists to help assess the physiological and psycho-emotional conditions of eSports players. In this project, we collect three classes of data (physiological, game telemetry, and environmental conditions) followed by a data analysis using artificial intelligence based on machine learning algorithms. For example, we apply machine learning and recurrent neural networks with attention to assessing player performance dynamics.		
Stakeholders	End users			
Stakeholders' assets, values	Trustworthiness, reputation, privacy, stability			
System's threats & vulnerabilities	New privacy threats, incorrect AI system use			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Create a model and develop a prototype data acquisition system	physiological, contextual and game telemetry	real-time measurements
	2	Experimental methodology	Recognition and noise reduction algorithms	collecting physiological data from professional cyber-sportsmen
	3	Identify the characteristic multidimensional sequences of movements and physiological parameters	Processing of the interaction between a person and Internet	Determine the psycho-emotional state
	4	Development of an algorithm for detecting abnormal psycho-emotional states	Consideration of multidimensional data of time series of measured physiological indicators	
	5	Development of algorithms and methods of	To process a wide range of various modalities of	To solve the problem of identifying a reliable psycho-

		predictive analytics	psychophysiological and other data	emotional state of a person and predictive determination of his behavior
	6	Obtaining a labeled set of multi-modal data	(a) processing of the video stream to obtain the primary markup of data, (b) assessment of game telemetry (game statistics)	Obtaining high-precision markup
	7	Determination of the dependence of the emotion experienced by the athlete on the measured data	Constructing display from the space of signs into the space of emotions	Emotional patterns
AI features	Task(s)	Recommendation		
	Method(s)			
	Hardware	Wired and wireless sensors, HPC, online data		
	Topology			
	Terms and concepts used	Predictive analytics, sensor networks, machine learning, artificial intelligence, neural networks, psycho-emotional state, neurophysiological state, e-sports, online games, people's behavior on the Internet		
Standardization opportunities/ requirements				
Challenges and issues	The challenges are associated with data collection and data analysis. To create a reasonably large dataset, a high number of Pro eSports athletes is required. Moreover, it is not a trivial task to collect the data during competitions; the sensors must ensure unobtrusive sensing. At the same time, the collected data is truly heterogeneous, e.g. video/time-series/tests, requiring new methods of data storage and data analysis.			
Societal Concerns	Description	Although eSports has evolved from amateur video gaming to a developing and innovative industry, there is a skeptical attitude to eSports in our society. A common understanding in particular communities is that eSports could be dangerous and cannot serve as a profession of the future.		
	SDGs to be achieved	Good health and well-being for people		

4277

4278 A.125.2 **References**

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link

1	Paper	V. Lebedev, E. Laukhina, V. Laukhin, A. Somov, A. M. Baranov, C. Rovira, J. Veciana. Investigation of sensing capabilities of organic bi-layer thermistor in wearable e-textile and wireless sensing devices. Organic Electronics. 42: 146-152, 2017. Impact Factor: 3.680				https://www.sciencedirect.com/science/article/pii/S1566119916305742
2	Paper	A. Somov, E. F. Karpov, E. Karpova, A. Suchkov, S. Mironov, A. Karelin, A. Baranov, D. Spirjakin. Compact low power wireless gas sensor node with thermo compensation for ubiquitous deployment. IEEE Transactions on Industrial Informatics 11(6): 1660-1670, 2015. Impact Factor: 5.43				https://ieeexplore.ieee.org/document/7088611
3	Paper	A. Somov, A. Baranov, D. Spirjakin, A. Spirjakin, V. Sleptsov, R. Passerone. Deployment and evaluation of a wireless sensor network for methane leak detection. Sensors and Actuators, A: Physical 202(11): 217-225, 2013. Impact Factor: 2.311				https://www.sciencedirect.com/science/article/pii/S0924424712007297
4	Paper	B. B. Velichkovsky. Consciousness and working memory: Current trends and research perspectives. Consciousness and Cognition, 55: 35-45, 2017. Impact Factor: 2.272				https://www.sciencedirect.com/science/article/pii/S1053810017301654
5	Paper	B. B. Velichkovsky, A. N. Gusev, A. E. Kremlev, S. S. Grigorovich. Cognitive control influences the sense of presence in virtual environments with different immersion levels. Lecture Notes in Computer Science, 10324 LNCS, pp. 3-16, 2017.				https://www.scopus.com/inward/record.uri?eid=2-s2.0-85021234651&doi=10.1007%2f978-3-319-60922-5_1&partnerID=40&md5=4e0d7b445de841e937da3dfb7b293d39
6	Paper	B. B. Velichkovsky. The relationship between interference control and sense of presence in virtual environments. Psychology in Russia: State of the Art, 10(3): 165-176, 2017. Impact Factor: 0.213				https://www.scopus.com/inward/record.uri?eid=2-s2.0-85031997117&doi=10.11621%2fpir.2017.0311&partnerID=40&md5=2983ec1a015dc1076533a32dbe06e189
7	Paper	F. Cong, A.-H. Phan, P. Astikainen, Q. Zhao, Q. Wu, J. K. Hietanen, T. Ristaniemi, A. Cichocki: "Multi-domain feature extraction for small event related potentials through nonnegative multi-way array decomposition from low dense array EEG". International Journal of Neural				https://www.ncbi.nlm.nih.gov/pubmed/23578056

		Systems, 23(2), 2013. Impact Factor: 4.58				
8	Paper	A. Cichocki, D. Mandic, A.-H. Phan, C. Caiafa, G. Zhou, Q. Zhao and L. De Lathauwer, “Tensor decompositions for signal processing applications from two-way to multiway component analysis”, IEEE Signal Processing Magazine, peer review, 32 (2): pp. 145–163, 2015. Impact Factor: 7.451				http://ieeexplore.ieee.org/abstract/document/7038247
9	Paper	I.V. Strelnikova, G.V. Strelnikova. "The developing potential of computer games." Computer Sports (eSports): Problems and Prospects: Materials of the 3rd All-Russian Scientific and Practical Conference (in the format of an online conference). -Moscow: Russian State University Of Physical Education,Sport, Youth And Tourism (SCOLIPE), 2014. -pp. 95-97				https://elibrary.ru/item.asp?id=24090561
10	Paper	E.V. Burnaev., G.K. Golubev. On one problem in Multichannel Signal Detection. Problems of Information Transmission, October 2017, Volume 53, Issue 4, pp 368–380. Impact Factor: 0.359.				https://link.springer.com/article/10.1134/S0032946017040056
11	Paper	A. Artemov, E. Burnaev. Optimal estimation of a signal perturbed by a fractional Brownian noise. Theory of Probability and Its Applications, 2016, vol. 60, № 1, pp. 126-134. Impact Factor: 0.41.				https://epubs.siam.org/doi/10.1137/S0040585X97T98752
12	Paper	E. Burnaev, A. Zaytsev. Large Scale Variable Fidelity Surrogate Modeling. Ann Math Artif Intell (2017), pp. 1-20. doi:10.1007/s10472-017-9545-y Impact Factor: 0.899				https://link.springer.com/article/10.1007/s10472-017-9545-y
13	Paper	M. Belyaev, E. Burnaev, E. Kapushev, M. Panov, P. Prikhodko, D. Vetrov, D. Yarotsky. GTApprox: Surrogate modeling for industrial design. Advances in Engineering Software 102 (2016) 29–39 Impact Factor: 3.198.				https://www.sciencedirect.com/science/article/pii/S0965997816303696

A.126 Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI

A.126.1 General

ID	126	
Use case name	Real-time segmentation and prediction of plant growth dynamics using low-power embedded systems equipped with AI	
Application domain	Agriculture	
Deployment model	Embedded system	
Status	In operation	
Scope	The project is devoted to the development of a low-power embedded system and AI algorithm for real-time plant segmentation and prediction of its growth. The proposed distributed system is aimed for use in greenhouses and remote areas, where edge-computing autonomous systems are in demand. A branch of this project also aims to develop the payload for drones for the segmentation of harmful plants in real-time.	
Objective(s)	Prediction of harvest, biomass/leaf area dynamics, leaf index, parameters describing the quality of produced food, consumption of resources from sequences of images of plant growth (including multispectral), data from sensors that describe environmental conditions and artificial growing system parameters representing the state of the growing system.	
Narrative	Short description (not more than 150 words)	Research efforts towards low-power sensing devices with fully-functional AI on board are still fragmented. In our project, we present an embedded system enriched with AI that ensures the continuous analysis and in-situ prediction of the plant leaf growth dynamics and other important growth parameters. The embedded solutions grounded on a low-power embedded sensing system with a Graphics Processing Unit (GPU) are able to run the neural networks-based AI on board. Advantages of the proposed system include portability and ease of deployment. The proposed approach guarantees the system autonomous operation for 180 days using a standard Li-ion battery. We rely on state-of-the-art mobile graphic chips for smart analysis and control of autonomous devices. The data was used for training and testing the Recurrent Neural Network, Convolutional Neural Network algorithms, and the segmentation algorithms. All this allows for high performance in-situ optimization of plant growth dynamics and resource consumption.
	Complete description	Research efforts towards low-power sensing devices with fully-functional AI on board are still fragmented. In our project, we present an embedded system enriched with the AI that ensures the continuous analysis and in-situ prediction of plant leaf growth dynamics and other important growth parameters. The embedded solutions, grounded on a low-power embedded sensing system with a Graphics Processing Unit (GPU), are able to run the neural networks-based AI on board. Advantages of the proposed system include portability and ease of deployment. We use a sequence of Convolutional Neural Network (CNN) and a

<table>
<tr><td></td><td></td><td colspan="3">Recurrent Neural Network (RNN) called the Long-Short Term Memory network (LSTM) as the core of the AI in our system. The proposed approach guarantees the system autonomous operation for 180 days using a standard Li-ion battery. We rely on state-of-the-art mobile graphic chips for smart analysis and control of autonomous devices. We used 5514 images as a source for automated leaf area calculation and follow the training of AI algorithms. Over 1000 records from sensors provide additional information about environmental conditions. All this data was used for training and testing the Recurrent Neural Network, Convolutional Neural Network algorithms, and the segmentation algorithms. Our solution provides a Root Mean Squared Error (RMSE) close to 4 sq.cm in a 3-hour prediction horizon. All this allows for high performance in-situ optimization of plant growth dynamics and resource consumption.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Agriculture, ecology management, sanitary services</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Stability, reputation, trustworthiness, competitiveness</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Hidden patterns, incorrect AI system use</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>CNN</td><td></td><td></td></tr>
<tr><td>2</td><td>RNN</td><td>Long-Short Term Memory network</td><td>core of AI</td></tr>
<tr><td rowspan="5">AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
<tr><td>Method(s)</td><td colspan="3"></td></tr>
<tr><td>Hardware</td><td colspan="3">GPU</td></tr>
<tr><td>Topology</td><td colspan="3"></td></tr>
<tr><td>Terms and concepts used</td><td colspan="3">CNN, RNN, autonomous device, segmentation, resource</td></tr>
<tr><td>Standardization opportunities/ requirements</td><td colspan="4"></td></tr>
<tr><td>Challenges and issues</td><td colspan="4">(1) The plant growth data significantly depends on multiple factors, including used solutions, illumination characteristics (for greenhouses), weather and seasonal conditions (for outdoors); (2) The architecture of the neural network should have both high accuracy, high framerate, but low amount of layers and trained parameters for further inference on low-power embedded systems. These controversial factors should be met since embedded systems have limited processing capabilities; and (3) high diversity of data types and no standardization of data obtained by farmers.</td></tr>
<tr><td rowspan="2">Societal Concerns</td><td>Description</td><td colspan="3">Good health and well-being for people; elimination of hunger; availability of cheap and healthy food for everyone; colonization of harsh environments on Earth and in space exploration.</td></tr>
<tr><td>SDGs to be achieved</td><td colspan="3">Good health and well-being for people; zero hunger</td></tr>
</table>

4283

A.126.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator /organization	Link
1	Patent	A. Menshchikov. “Airflow 2.0” RU #2018618762, 2018. Topic: “2D Computational Fluid Dynamics Simulator and Optimizer of 2D Airfoils”.	Published	High		#2018618762
2	Grant	#9189ГУ/2015 in UMNIK program (2015-2018). Topic: “Design and Development of Adaptive Wing for Unmanned Aerial Vehicle with Electric Power Source”	Realized	High		#9189ГУ
3	Paper	Menshchikov, A. M., and Somov, A. S., “Morphing wing with compliant aileron and slat for unmanned aerial vehicles”, Physics of Fluids Journal, Vol. 31, No. 3, March 2019.				
4	Paper	A. Menshchikov. “Development of Adaptive Wing with Double Hinge Aileron for Unmanned Aerial Vehicles”, Austrian Journal of Natural and Technical Science, pp. 150-159, Jun. 2018				
5	Abstract	A. Menshchikov, I. Dranitsky, D. Ermilov, L. Kupchenko, M. Panov, A. Somov and M. Fedorov. “Data-Driven Body-Machine Interface for Drone Intuitive Control through Voice and Gestures”, IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society (IES)	Published			
6	Paper	A. Menshchikov, D. Shadrin, S. Sosnin, E. Tsykunov, V. Prutyanov, D. Lopatkin, E. Iakovlev, A. Somov “Fighting Against Hogweed in Real-time: Airborne Platform Empowered by Deep Learning”, Computers and Electronics in Agriculture	Submitted			6
7	Abstract	D. Shadrin, A. Menshchikov*, A. Somov and M. Fedorov “Enabling Precision Agriculture through Embedded Sensing with Artificial Intelligence”, IEEE Transactions on Instrumentations and Measurements, pp. 1-10.				7
8	Related Paper	D. Shadrin, A. Menshchikov*, D. Ermilov, and A. Somov, “Designing Future Precision Agriculture: Detection of Seeds Germination Using Artificial Intelligence on a Low-Power Embedded System”, IEEE Sensors Journal, pp. 1-10	Published		doi: 10.1109/JSEN.2019.2935812.	8

A.127 Search of undiagnosed patients

A.127.1 General

<table>
<tr><td>ID</td><td colspan="4">127</td></tr>
<tr><td>Use case name</td><td colspan="4">Search of undiagnosed patients</td></tr>
<tr><td>Application domain</td><td colspan="4">Healthcare</td></tr>
<tr><td>Deployment model</td><td colspan="4">Social networks</td></tr>
<tr><td>Status</td><td colspan="4">In operation</td></tr>
<tr><td>Scope</td><td colspan="4">Search of undiagnosed patients with orphan diseases, define patients’ journey</td></tr>
<tr><td>Objective(s)</td><td colspan="4">Deep semantic analysis of unstructured texts (based on meaning, rather than keywords, i.e. using natural language processing technology)</td></tr>
<tr><td rowspan="2">Narrative</td><td>Short description (not more than 150 words)</td><td colspan="3">Knowledge extraction from the massif of user posts in patient forums, and physicians’ professional networks, health-related portals, etc.</td></tr>
<tr><td>Complete description</td><td colspan="3">Full-scale crawling of Google and Yandex environment.
Semantic and statistical analysis of found posts related to description of particular symptoms, description of clinical analyses, diagnostic procedures, etc.
Identification of insights and presentation of results.
Semantic artificial intelligence (AI) tools that can read and interpret electronic free text at scale. Real patient journey, patient subgroups, etc. are to be evaluated.
A unified medical and social image of the user (patient) can be created.</td></tr>
<tr><td>Stakeholders</td><td colspan="4">Patients, government affairs, physicians, pharma companies.</td></tr>
<tr><td>Stakeholders’ assets, values</td><td colspan="4">Personal data of the subjects planned to be identified, especially patients’, i.e. special health information could potentially be in risk area.</td></tr>
<tr><td>System’s threats & vulnerabilities</td><td colspan="4">Difficulties with ordering and finding patients.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Patient journey</td><td>Real patient journey is to be clarified based on obtained data. Disease guidelines are to be changed accordingly</td><td>Inflamm Bowel Dis _ Volume 23, Number 7, July 2017</td></tr>
<tr><td>2</td><td>Effectiveness</td><td>% of totally identified patients should be close to number pf patients predicted by prevalence data</td><td>National disease and patient registries,</td></tr>
<tr><td rowspan="3">AI features</td><td>Task(s)</td><td colspan="3">Natural language processing</td></tr>
<tr><td>Method(s)</td><td colspan="3">Crawling, natural language processing</td></tr>
<tr><td>Hardware</td><td colspan="3"></td></tr>
</table>

	Topology	
	Terms and concepts used	AI and deep linguistic processing, Patient Journey verification,
Standardization opportunities/ requirements		
Challenges and issues	Personal data of the subjects planned to be identified, especially patients', i.e. special health information could potentially be in risk area.	
Societal Concerns	Description	
	SDGs to be achieved	Good health and well-being for people

4288

4289 A.127.2 Data

Data characteristics	
Description	
Source	
Type	
Volume (size)	
Velocity	
Variety	Real time
Variability (rate of change)	Multiple
Quality	

4290

4291 A.127.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Manuscript	Inflamm Bowel Dis. 2017 Jul;23(7):1057-1064. Patient Understanding of the Risks and Benefits of Biologic Therapies in Inflammatory Bowel Disease: Insights from a Large-scale Analysis of Social Media Platforms. Martinez B1, Dailey F, Almario CV, et al	published	Use case taken from this reference	*Cedars-Sinai Center for Outcomes Research and Education (CS-CORE)	doi: 10.1097/MIB.0000000000001110

4292

4293 A.128 Semantic Analysis of Legal Documents

4294 A.128.1 General

ID	128
Use case name	Semantic Analysis of Legal Documents
Application domain	Legal

Deployment model	On-premise systems	
Status	Prototype	
Scope	Semantic analysis of legal documents in the course of its development, verification and improvement	
Objective(s)	Machine understanding of the meaning of legal documents. The obtaining of semantic visual images of documents; the detection of contradictions and inaccuracies in legal documents describing similar objects of law for the task of classifying documents; quick document comprehension; and analyzing the consistency of the legal base.	
Narrative	Short description (not more than 150 words)	The software tool is oriented on the analysis and representation content of normative documents in the form of formal ontology (OWL ontology) and the construction of their visual images for the subsequent detection of inaccuracies and contradictions using logical inference and visual analysis methods.
	Complete description	The most important condition for ensuring the integrity of the legal base is the identification and elimination of contradictions, which are often found when using existing or developing new legal acts and documents relating to various aspects of the same objects of law. To solve this problem, a software tool has been developed to control the integrity of the legal base in the development and use of legal documents. The software tool accepts an initial set of legal documents as input, performs its syntactic and semantic analysis. For parsing and determining the morphological characteristics of words, a grammar dictionary of the Russian language, the WordNet thesaurus, and the SyntaxNet library for determining syntactic relationships, are used. The result of the analysis of the document is a weighted semantic image of the document, which is a semantic network of concepts and relations between them. A fragment of this network related to documents and concepts of interest to an expert is described as a set of OWL expressions - an ontology of documents. As the base of this ontology - ontology for legal knowledge representation - LKIF-Core is used. On the resulting ontology logical inference by using the JFACT++ reasoner is performed to identify possible inconsistencies and notify the expert. At the same time, the generated semantic image of the document is visualized in various ways, in which vertex incidence (degree of detail of the description of a concept in the document), the weight of the edges (the importance of the relationship between the concepts), paths (structure of the definition of concept in document) are presented using visual effects. The expert can therewith quickly comprehend a document, identify documents that are similar in meaning, and identify possible problem places in the definition of legal concepts and relations between them.

Stakeholders	legislative institutions, management institutions			
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Relative time of document analysis	The relative time of the analysis of the document is defined as the ratio of the number of words in the document to the time of its analysis by an expert	Identification of contradictions and inaccuracies in regulatory documents
AI features	Task(s)	Knowledge processing & discovery		
	Method(s)	Ontology learning		
	Hardware			
	Topology			
	Terms and concepts used	visual analysis, ontology, ontology web language, inference engine, reasoner, ontology learning		
Standardization opportunities/ requirements				
Challenges and issues	Different levels of abstraction of concepts in documents.			
Societal Concerns	Description			
	SDGs to be achieved			

A.128.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	scientific paper	Lomov P. A., Development Of The Ontology Based Technology For Legal Documents Consistency Checking And Coordination Support / Lomov P.A., Oleynik A. G // Proceedings of Institute of System Analysis: Mathematical models of socio-economic processes. Decision Making Methods. Numerical methods	published			http://www.isa.ru/proceedings/index.php?option=com_content&view=article&id=782

		of solution. Economic and sociocultural problems of the information society. Risk and security management. Vol.63. Book.2 2013. – p. 62-69. (ISBN 978-5-396-00530-3)				
2	scientific paper	Vicentiy A.V., Dikovitsky V.V., Shishaev M.G. The Semantic Models of Arctic Zone Legal Acts Visualization for Express Content Analysis // Advances in Intelligent Systems and Computing. 2019. Vol. 763, pp. 216-228.	published			https://link.springer.com/chapter/10.1007/978-3-319-91186-1_23

A.129 Support system for optimization and personification of drug therapy

A.129.1 General

ID	129	
Use case name	Support system for optimization and personification of drug therapy	
Application domain	Healthcare	
Deployment model	On premise system	
Status	PoC	
Scope	It is a full-range of integrated solutions for the selection of the optimal type of drug, its dose, and its combination with other drugs	
Objective(s)	Support system for optimization of the medical therapy of the patient taking into account their individual physiological features, type, and disease severity	
Narrative	Short description (not more than 150 words)	Data from the laboratory and clinical examinations of a particular patient are displayed in a single integrative medical record. There is currently a significant amount of patient data available electronically. Based on the pool of data of patients receiving a known drug, training is conducted in the recommendation system using AI, taking into account their individual physiological characteristics, type, and severity of the disease, as well as the particular drug's combined administration with other drugs. When requesting recommendations for a patient, after entering information of their current condition, the system will give individualized recommendations for optimizing drug therapy. Furthermore, the system in the course of treatment, receiving fresh data, makes recommendations for the correction of therapy.
	Complete description	For the doctor at the present time it may be a problem to choose a specific drug and the selection of its optimal dosage in the treatment of a disease. There are, however, a number of more experienced therapists, in whose practice may have repeatedly occurred cases of atypical courses of

		disease, characteristics of patients, and the combined administration of several drugs. A thorough analysis of documented cases will provide recommendations and generalizations for these patient groups. However, clusters of case histories for each patient history group must first be created. It is expected that the number of cases will be very unevenly distributed among groups. Although for the most typical cases, recommendations are also typical and can be given, including inexperienced (novice) doctors, for cases of diseases falling into clusters with a small amount of data, in the presence of individual physiological characteristics of the patient and the presence of other drugs, the accumulation of data and training of the AI system based on the recommendations of doctors is of particular importance. The main body of the analyzed data is text data, namely transcripts of the results of the analysis of patients and doctors' appointments. However, input data can also contain images (snapshots), which implies a more complex data analysis based on deep and full entanglement of neural networks.		
Stakeholders	public and private healthcare system, pharmaceutical companies			
Stakeholders' assets, values	Safety, privacy, fair treatment, trustworthiness			
System's threats & vulnerabilities	New privacy threats, new security threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Appropriateness of treatment	Proportion of the appropriate flow of obtained inference logic	Improve efficiency
AI features	Task(s)	Knowledge processing & discovery, Natural Language Processing, Image recognition, Inference, Prediction,		
	Method(s)	Classification, Feature Extraction, Knowledge Graph, Deep Learning, Natural Language Processing		
	Hardware			
	Topology			
	Terms and concepts used	Knowledge Graph, Deep Learning, Natural Language Processing, Explainable AI		
Standardization opportunities/ requirements				
Challenges and issues	In addition to the classic data analysis with new technologies to find hidden patterns in relation to health care, the possibility of using methods and technologies to analyze a heterogeneous mass of data with a significant percentage of emissions and uneven distribution of data by classes and			

	categories is a challenge. Of challenge is well is identifying hidden correlations and thereby improving the quality of medical services.	
Societal Concerns	Description	Incorrect decision Unexplainable result
	SDGs to be achieved	Improving the effectiveness of drug therapy

A.130 Syntelly - computer aided organic synthesis

A.130.1 General

ID	130	
Use case name	Syntelly - computer aided organic synthesis	
Application domain	Other (please specify) - Drug design, digital pharma	
Deployment model	Hybrid or other (please specify) - System for the prediction of the properties of pharmaceutically relevant molecules	
Status	Prototype	
Scope	Recent progress in deep learning has made a revolution in many areas of science and technology. However, the potential of this method in drug discovery has not yet been fully elaborated. The Syntelly project intends to close this gap. We are developing a web-based platform that helps chemists navigate through chemical space by predicting synthetic availability and ways of synthesis for new drug candidates that have not yet been studied; it also estimates the potential efficiency and safety of specific molecules. We hope that the successful implementation of our project will reduce drug discovery costs and related risks, which will stimulate pharmaceutical companies to search for unexplored molecules as a base for a new generation of drugs.	
Objective(s)	Exploration of chemical space is a very complicated task due to a large number of predicted chemical molecules. The number of described molecules is only several million compounds, but the estimated number of potentially synthetically accessible molecules is enormous: around 10^60 [4], and neither man nor machine can directly process such a volume of data. The only hope is the development of methods and tools, based on deep learning, which will trigger a chemist-machine alliance to analyze chemical Big Data.	
Narrative	Short description (not more than 150 words)	The Syntelly project is directed to help organic chemists in chemical space exploration. Due to high risks and cost of new molecule trials, pharmaceutical companies do not prefer to open new chemical space areas in an experimental way. Using deep learning based on the chemical reaction databases, we predict the best retrosynthesis pathway to achieve the easiest way to a molecule synthesis. The next task is the prediction of the toxicity and bioconcentration of the molecule.
	Complete description	It requires approximately $1,000,000,000 to bring a new drug to the global market. Moreover, 30% of drugs fail the first stage of clinical trials due to unexpected side effects [5]. Chemical space is close to being dried out. Pharmaceuticals companies are trying to find new molecules that are similar to existing ones because the exploration of the new scaffolds is risky; a company's losses may be very high if the

		drug candidate fails. However, there is a strong demand for new kinds of drugs, especially for antibacterial and antiviral therapy due to emerging resistance. Long-term consumption of medicines often leads to a lack of sensitivity, and this fact also motivates us to develop methods for the exploration of chemical space to investigate unexplored scaffolds. Multitask deep learning allows heterogeneous chemical data to be processed. With our platform Syntelly, we implemented multitask models for acute toxicity of organic compounds for different toxicological endpoints (an endpoint is a combination of animal type, type of administration, and type of toxicity). We trained our deep neural networks on a broad dataset of more than 87,000 compounds. Our best models achieved high performance ($R^2 > 0.7$) while having a broad and diverse applicability domain. This result is better than previous state-of-the-art approaches without multitask learning [2]. Recent progress in deep generative models raised not only the extensive grow of intellectual assistants (chat bots) but also inspired a new paradigm in the de-novo generation of molecules with desired properties. We implemented a generator of promising drug candidates satisfying the criteria of high affinity to a target receptor, low toxicity, and good synthetic accessibility. As mentioned, the final decisions are for humans, which is why there is a need for tools to represent chemical space in a convenient way. Deep learning can also support humans in perceiving large chemical data. We implemented a parametric t-SNE based mapper of chemical compounds to the 2D surface, such that similar compounds group together [3]. On the base of this method, we created a tool that helps chemists work with large chemical databases.		
Stakeholders	Organic chemists, medical chemists, synthesists			
Stakeholders' assets, values	Trustworthiness, robustness, reputation			
System's threats & vulnerabilities	New security threats, new privacy threats			
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	3-D descriptors	The modelling of molecular descriptors	Physicochemical properties
	2	CNN	Multitask deep learning	Neglecting data diversity
AI features	Task(s)	Prediction		
	Method(s)	Data-driven modeling, CNN		
	Hardware	CPU and GPU		
	Topology			
	Terms and concepts used	CNN, chemical space, machine learning, multitask deep learning, chemical reactions		

Standardization opportunities/ requirements		
Challenges and issues	— a) The large size of chemical space implies the development of machine learning algorithms in two directions: to generate molecules and estimate their parameters, and for chemical space customization for new synthetic pathways — b) Characteristics of organic compounds are extremely diverse. They are collected from different sources and may be represented in many ways (i.e. toxicity can be measured on different animals). — c) There are only two major players on the market of chemical and reaction data, and the possibilities to obtain the whole datasets required for deep learning are heavily restricted. — d) Synthetic and medical chemists prefer to ignore computer-based approaches.	
Societal Concerns	Description	Our primary goal is to make the drug discovery process easier and cheaper. It will stimulate pharmaceutical companies and academic researchers to study new compounds and new scaffolds. Finally, society will obtain new effective drugs against the most dangerous bacterial and viral diseases. Reducing risks will generate interest in developing drugs for orphan diseases, which is now one of the biggest problems for society.
	SDGs to be achieved	Good health and well-being for people; responsible consumption and production

A.130.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Paper	A Survey of Multi - Task Learning Methods in Chemoinformatics. Mol. Inf..Sosnin, S. , Vashurina, M., Withnall, M., Karpov, P., Fedorov, M. and Tetko, I. V. (2018),	Published	High		doi:10.1002/minf.201800108
2		Comparative study of multitask toxicity modeling on a broad chemical space / S. Sosnin, D. Karlov, I. V. Tetko, M. Fedorov. Journal of Chemical Information and Modeling. 2019, 59, 3	Published	High		
3	Paper	3D matters! 3D-RISM and 3D convolutional neural network for accurate bioaccumulation prediction, 2018 J. Phys.: Condens. Matter 30 32LT03	Published	High		

4	Related Paper	Kirkpatrick, P., & Ellis, C. (2004). Chemical space. Nature, 432(7019), 823–823	Published	High		
5	Related paper	Wong, C. H., Siah, K. W., & Lo, A. W. (2018). Estimation of clinical trial success rates and related parameters. Biostatistics				
6	Paper	Karlov, D. S., Sosnin, S., Tetko, I. V., & Fedorov, M. V. (2019). Chemical space exploration guided by deep neural networks. RSC Advances, 9(9), 5151–5157	Published	High		
7	Grant	STRIP program in Skoltech: Syntelly – Computer aided organic synthesis	Realized	High		https://sip.skoltech.ru/en/supported-projects/program-2018-2019/syntelly-computer-added-organic-synthesis/

A.131 WebioMed Clinical Decision Support System

A.131.1 General

ID	131	
Use case name	WebioMed Clinical Decision Support System	
Application domain	Healthcare	
Deployment model	Cloud services	
Status	In operation	
Scope	Screening for cardiovascular disease risk prediction with machine and deep learning methods	
Objective(s)	Advances in precision medicine will require an increasingly individualized prognostic evaluation of patients in order to provide the patient with appropriate therapy	
Narrative	Short description (not more than 150 words)	Cardiovascular disease (CVD) continues to be the most relevant health problem of most countries in the world, including the Russian Federation. According to the World Health Organization, more than 17 million people die each year from CVD worldwide, including more than 7 million from coronary heart disease (CHD). The machine learning models outperformed traditional approaches for CVD risk prediction (such as SCORE,

		PROCAM, and Framingham equations). This approach was used to create a clinical decision support system (CDSS). It uses both traditional risk scales and models based on neural networks. Of notable importance is the fact that the system can calculate the risk of cardiovascular disease automatically and recalculate immediately after adding new information to the EHR. The results are delivered to the user's personal account.
	Complete description	The **CDSS WebioMed** is a ready-made, trained solution to identify high-risk patients and prevent morbidity and mortality. — Automatic risk stratification of patients — A more efficient organization of preventive work aimed at a personal group of patients with a high risk of complications and death — The ability to route patients depending on the assessment obtained — Reduced morbidity and mortality — Reliable digital assistance, trained on the results of evidence-based medicine and modern clinical guidelines. — Automatic Identification of risk factors — Automatic determination of the likelihood of developing a disease — Compliance with clinical practice guidelines — Reduced time of the patient risk assessment — Powerful artificial intelligence to evaluate medical data and identify risk factors without development costs. — The addition of medical decision support functions — Ready service for evaluating EHR and to identify the risk factors — Reducing the costs of development of medical information system
Stakeholders	— End-users (physician, nurse, laboratory technologist, pharmacist, patient) — Sales and marketing team — CDSS product development and maintenance team (system administrator, system developer, system architect, project manager, and system maintenance)	
Stakeholders'	Competitiveness, cost savings	

<table>
<tr><td>assets, values</td><td colspan="4"></td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="4">Injuries and error. The most obvious risk is that AI systems will sometimes be wrong, and that patient injury or other health-care problems may result.
Data availability. Training AI systems requires large amounts of data from sources such as electronic health records, pharmacy records, insurance claims records, or consumer-generated information like fitness trackers or purchasing history. But health data are often problematic. Data are typically fragmented across many different systems.
Privacy concerns. Another set of risks arise around privacy. The requirement of large datasets creates incentives for developers to collect data from many patients. Some patients may be concerned that this collection may violate their privacy, and lawsuits have been filed based on data-sharing between large health systems and AI developers.
Bias and inequality. There are risks involving bias and inequality in health-care AI. AI systems learn from the data on which they are trained, and they can incorporate biases from those data. For instance, if the data available for AI are principally gathered in academic medical centers, the resulting AI systems will know less about—and therefore will treat less effectively—patients from populations that do not typically frequent academic medical centers.</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>AUC ROC</td><td>Area under curve receiver operating characteristic</td><td>AUC provides an aggregate measure of performance across all possible classification thresholds.</td><td>To determine the quality and correctness of classification models</td></tr>
<tr><td>TP,FP,TN,FN</td><td>Confusion matrix</td><td>Metrics that can be used to measure the performance of a classifier or predictor</td><td>Some of these people have the disease, and our test correctly says they are positive. They are called true positives (TP). Some have the disease, but the test incorrectly claims they don't. They are called false negatives (FN). Some don't have the disease, and the test says they don't – true negatives (TN). Finally, there might be healthy people who have a positive test result – false positives (FP). These can be arranged into a</td></tr>
</table>

				2×2 contingency table (confusion matrix), conventionally with the test result on the vertical axis and the actual condition on the horizontal axis.
	Accuracy, precision and recall	Metrics	Evaluation metrics for machine learning	To evaluate the performance of a model in ML
AI features	Task(s)	Natural language processing		
	Method(s)	SpaCy, NLTK, StanfordNLP,Tensorflow,Keras		
	Hardware	CPU, TPU,		
	Topology	Colaboratory Google, web-services,		
	Terms and concepts used	Classification, features extraction, NLP, logit regression, data driven application		
Standardization opportunities/ requirements				
Challenges and issues	Challenges: to provide physician tools to easily calculate cardiovascular risk anywhere in a world			
Societal Concerns	Description	One of the major concerns about AI-assisted CDSS is how the machines reach decisions, and whose decision should prevail when there is disagreement between the CDSS and the medical professional. This lack of transparency is referred to as the 'black box' of AI. In addition to the lack of transparency, the necessary use of large training data sets coupled with mathematical and statistical algorithms and sometimes neural networks, whether with or without full understanding of the internal workings, presents a challenge in educating doctors to use these tools in a clinically relevant way.		
	SDGs to be achieved	Good health and well-being for people		

A.131.2 Data

Data characteristics	
Description	FHS-Cohort
Source	Biologic Specimen and Data Repository Information Coordinating Center
Type	Structured/unstructured text: time-series
Volume (size)	84 Mb
Velocity	
Variety	
Variability (rate of change)	never

Quality	presence of missing fields or incorrect values

A.131.3 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Publication	I Korsakov, A Gusev,T Kuznetsova,D Gavrilov,R Novitskiy Deep and machine learning models to improve risk prediction of cardiovascular disease using data extraction from electronic health records. European Heart Journal, Volume 40, Issue Supplement_1, October 2019, ehz748.0670, Published: 21 October 2019	Published: 21 October 2019	Yes	K-Lab, Ltd./K-SkAI	https://doi.org/10.1093/eurheartj/ehz748.0670

A.132 Device Control Using both cloud AI and embedded AI

A.132.1 General

<table>
<tr><td>ID</td><td colspan="2">132</td></tr>
<tr><td>Use case name</td><td colspan="2">Device Control Using both cloud AI and embedded AI</td></tr>
<tr><td>Application domain</td><td colspan="2">Manufacturing</td></tr>
<tr><td>Deployment model</td><td colspan="2">Hybrid or other (please specify) (Learning in both Cloud and Embedded)</td></tr>
<tr><td>Status</td><td colspan="2">In operation</td></tr>
<tr><td>Scope</td><td colspan="2">Learn the user's preferred temperature for each situation for the control of home appliances (air conditioning equipment)</td></tr>
<tr><td>Objective(s)</td><td colspan="2">Keep comfortable room status by driving home appliances (air conditioning equipment) at the user's preferred temperature according to the situation</td></tr>
<tr><td>Narrative</td><td>Short description (not more than 150 words)</td><td>Because temperature that the user feels comfortable depending on the situation, such as the time of day and the day of the week, the user changes set temperature every time the user feels uncomfortable.
By Learning the user's preferred temperature for each situation, home appliances (air conditioning equipment) can keep room comfortable state automatically.
For the learning of the operation with long-term cycle, such as a fixed operation for each day of the week, it is effective learning from the accumulated operation history. So, A model is learning on the cloud.
For sudden operation pattern changes, e.g., when the temperature of the day rises suddenly and user react to it,</td></tr>
</table>

		high frequency online machine learning inside the equipment can adjust the model immediately. The consistency between the model learned on the cloud and one adjusted inside the equipment should be kept.
	Complete description	**Motivation:** — The temperature at which the user feels comfortable varies depending on the outside conditions of the air conditioner, such as outside temperature, sunshine, time of day, day of the week, etc. — Always maintain a comfortable state by eliminating the need for this setting change **Problem statement:** — Though temperature that the user feels comfortable depending on the situation, such as the time of day and the day of the week changes, it is impossible to preset these settings at the time of product shipment. — Even if designer of the product provides a method to let user set such setting, the user himself/herself does not know he/she should set what degree on what time. — Long-term data cannot be stored in the device, but forced to learn in the cloud, only the learning of the batch in the cloud is longer time to be able to cope with the variation of the sudden driving pattern of the user. **Current situation:** — The temperature is set using the controller every time the user feels uncomfortable **Solution Approach and Solution Steps:** In addition to learning the model using long-term historical data in the cloud, the model is also adjusted by learning frequently in embedded devices. — When the user changes the temperature setting using the controller, in addition to the setting contents, its data is stored with the accompanying data, such as the setting time, in the air conditioner. — The operating status data, such as temperature sensor values installed for the control of the air conditioner, in the air conditioner. — Upload data stored in the air conditioner to the cloud instance held by the manufacturer periodically. — The latest weather forecast information, etc. is kept on the cloud at all times. — Create a model to represent what set temperature should be in accordance with the external situation of the air conditioner (including the forecast) by learning

<table>
<tr><td></td><td></td><td>for each air conditioner on the cloud periodically. The model is delivered to the corresponding air conditioner.

— Online machine learning is performed based on the data stored inside the air conditioner, and the internal parameters of the model are adjusted. This embedded learning is performed frequently, e.g., once an hour, and it is possible to reflect sudden changes in the user's usage pattern to the model.

— The online machine learning algorithm inside air conditioner and batch machine learning algorithm in the cloud is tuned as close as possible to prohibit radical model change from adjusted model by online machine learning when the model is delivered from the cloud and overwritten the adjusted model.

— Air conditioner predicts the preferred temperature with the model, and the result is used as the set temperature of the air conditioner.

— Air conditioner, as in normal operation, performs control so that the temperature of the room keeps set temperature.

Results and Effects:
— Since the prediction is done by the air conditioner (embedded), it works in case of a network failure or a cloud failure. The only impact of a failure is the inability to upload data and the inability to update the model by learned by the cloud.

— The learning of the operation with long-term cycle, such as a fixed operation for each day of the week, is effective if the model is learned from the accumulated operation history. A model with this effect is created mainly by learning on the cloud.

— In case of sudden operation pattern changes, e.g., when the temperature of the day rises suddenly and the user react to it, high frequency of online machine learning inside the air conditioner can adjust the model immediately.</td></tr>
<tr><td>Stakeholders</td><td colspan="2">Equipment users, manufacturers, distributors</td></tr>
<tr><td>Stakeholders' assets, values</td><td colspan="2">Equipment users: comfort, unintended (unpleasant) behavior, riskless behavior, privacy
Manufacturer: Competitiveness, Reputation, Reliability, Safety
Distributor: No claims for unintended (unpleasant) behavior</td></tr>
<tr><td>System's threats & vulnerabilities</td><td colspan="2">— Creating an incorrect model by machine learning using the child's mischievous operation history</td></tr>
</table>

<table>
<tr><td></td><td colspan="4">— Create an incorrect model by machine learning using the history of operations based on user misunderstandings, for example, operations that set the temperature extremely low when the user want to cool immediately,
— When resold, the use pattern of the original user leaks to the resale destination by using the air conditioner.
— Threats to the cloud in general</td></tr>
<tr><td rowspan="3">Key performance indicators (KPIs)</td><td>ID</td><td>Name</td><td>Description</td><td>Reference to mentioned use case objectives</td></tr>
<tr><td>1</td><td>Number of cancel operations</td><td>Air conditioner changes temperature setting based on the prediction. When a user notices unintended(unpleasant) setting, the user operates controller to cancel setting based on the prediction.</td><td></td></tr>
<tr><td>2</td><td>Distance between models</td><td>The difference between the model learned in the cloud and the one learned in the embedding equipment(air-conditioner)</td><td></td></tr>
<tr><td rowspan="4">AI features</td><td>Task(s)</td><td colspan="3">Prediction</td></tr>
<tr><td>Method(s)</td><td colspan="3">Machine Learning, Online Machine Learning</td></tr>
<tr><td>Hardware</td><td colspan="3">PC (pre-validation), cloud, cloud-to-device communication (Internet), embedded equipment</td></tr>
<tr><td>Topology</td><td colspan="3">— All air conditioners are connected to one cloud.
— On the cloud, keep a history of past operations and operating conditions for all air conditioners.
— Learning for each air conditioner on the cloud, and delivering the created model to the air conditioner.
— The air conditioner retains the operation history and operation status history data for a certain period of time, and also maintains the delivered model. The model is adjusted regularly by executing online machine learning in the air-conditioning.</td></tr>
</table>

		— Change the set temperature based on the prediction based on the model in the air conditioner.
	Terms and concepts used	Cloud AI, Embedded AI,
Standardization opportunities/ requirements	Standardization of architecture in which multiple AIs (Online algorithms and batch algorithms) in multiple place (embedded and cloud) work together for the same purpose	
Challenges and issues	— During actual use, there is a possibility of significant difference between the model learned by cloud and the model adjusted in air-conditioner. It leads significant change of temperature setting when the model in the air conditioner is overridden by the model learned by the cloud. — How and when to detect whether there has been a significant difference. — How does air-conditioner explain a significant difference when it is detected. Criteria for determining whether or not to explain	
Societal Concerns	Description	By automatically adjusting the temperature so that the user feels comfortable, it can suppress unnecessary power due to overtemperature or overcool.
	SDGs to be achieved	Affordable and clean energy

A.132.2 References

References						
No.	Type	Reference	Status	Impact on use case	Originator/organization	Link
1	Press release		Published		Fujitsu	https://www.fujitsu-general.com/jp/news/2019/09/19-N04-19/index.html (In Japanese)
2	Press release		Published		Fujitsu	https://www.fujitsu-general.com/shared/jp/pdf-fcjp-news-19-n04-19-02.pdf (In Japanese)
3	patent	JP2019/033811			Fujitsu	
4	patent	JP2019-05309			Fujitsu	

Annex B
(informative)

Impact Analysis Items

Table 8 —List of Impact Analysis Items

No.	Impact analysis items
1	Context or Application Area
2	Technologies
3	Title
4	Scope
5	Identified Benefits
6	Identified Challenges
7	Identified Societal Concerns
8	Data
9	Are all key stakeholders identified?
10	Are there any vulnerable stakeholders (e.g., children, mothers with young children, racial minorities, cultural minorities, ethnic minorities, displaced persons, incarcerated persons, refugees, etc.)?
11	If there are vulnerable stakeholders, do they have an identified voice in the process or technology?
12	If they don't have a voice, how will their interests be protected?
13	Is the application, technology, system or process well-defined?
14	Is the application, technology, system or process transparent to the developers and engineers?
15	Is the application, technology, system or process transparent to the users of the system?
16	Is the application, technology, system or process transparent to other stakeholders?
17	Are there environmental or sustainability issues involved? (e.g., water management and access, pollution, energy, etc.)

18	Are there health and wellness issues?
19	Are there gender equality issues?
20	Are there workforce or economic equality issues?
21	Are there data or privacy issues that could adversely affect or unduly benefit specific individuals or stakeholders?
22	Are there Intellectual Property Rights that need to be considered and protected?
23	Does this technology, system, or process manipulate, bias, or alter (or seek to manipulate, bias, or alter) an individual's behavior, attitudes, ideas, or actions?
24	Are there any aspects of the technology, system, or process that would deny essential services to some stakeholders?
25	Are there cultural, economic, political, social, or technical biases in the evaluation process?
26	Are there biases in the benefits of this technology or application?
27	Can there be any unanticipated feedback or interactions because of the complexity of the system, technology, or process?

Bibliography

[1] B. Du Boulay. "Artificial Intelligence as an Effective Classroom Assistant". IEEE Intelligent Systems, V 31, p.76–81. 2016.

[2] S. Hong. "Artificial intelligence audio apparatus and operation method thereof". N US 9,948,764, Available at: https://patents.google.com/patent/US20150120618A1/en. 2018.

[3] M.R. Sumner, B.J. Newendorp and R.M. Orr. "Structured dictation using intelligent automated assistants". N US 9,865,280, 2018.

[4] J.HENDLER, S. ELLIS, K. MCGUIRE, N. NEGEDLY, A. WEINSTOCK, M. KLAWONN and D.Burns. "WATSON@RPI, Technical Project Review". Available at:https://www.slideshare.net/jahendler/watson-summer-review82013final. 2013.

[5] AI white paper editing committee of Information Technology Promotion Agency, Japan. "Artificial Intelligence White Paper" (in Japanese). Kadokawa Ascii Research Laboratories, Inc, 2017.

[6] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: http://ai100.stanford.edu/2016-report.

