Forum on IoT in Smart Sustainable Cities
"The IoT Meets Big Data: A Standards Perspective"
18 January 2016, Singapore

Data in the IoT: considerations on opportunities, challenges and standardization perspectives

Marco CARUGI
Senior Consultant, NEC Corporation
ITU-T Q2/20 Rapporteur

Some trends in the IoT

Fragmented market & vertical applications
One device type per application

Single domain data

Specific and proprietary technologies, In-house IT

Limited service models, device and connectivity focus

Dominance of cellular networks

→

Horizontal solutions
Multi-purpose devices

Data and information driven, Integrated data analytics, Cloud, data marketplace

Application innovation

B2C, B2B, B2B2C services
Hyper-growth of Service Providers, variety of business models, integrated ecosystems

Multi-access networks
Cloud Computing, distributed computing

Standardization has to address this evolution

Opportunities of data in the IoT: using and monetizing data via analytics

(Big) Data analytics are a key revenue opportunity in IoT market

Data analytics can allow organizations to drive revenue by sharing, analyzing and interpreting data, for multiple purposes

- Extraction of tangible business and technology value
- Response and action in real time, improving productivity and business processes, lowering operational and other costs
- Long-range forecasts enabling strategic actions
- Better/New service offer to customers, faster and more efficiently

Surveys indicate increasing Big Data analytics deployments and exponential growth of Big Data analytics revenues

Solution for monitoring and predicting failures for plant facilities [example using NEC's analysis technologies]

Customer's challenge

- Avoid damages by predicting failures, shorten the lead time to identify the cause of failure

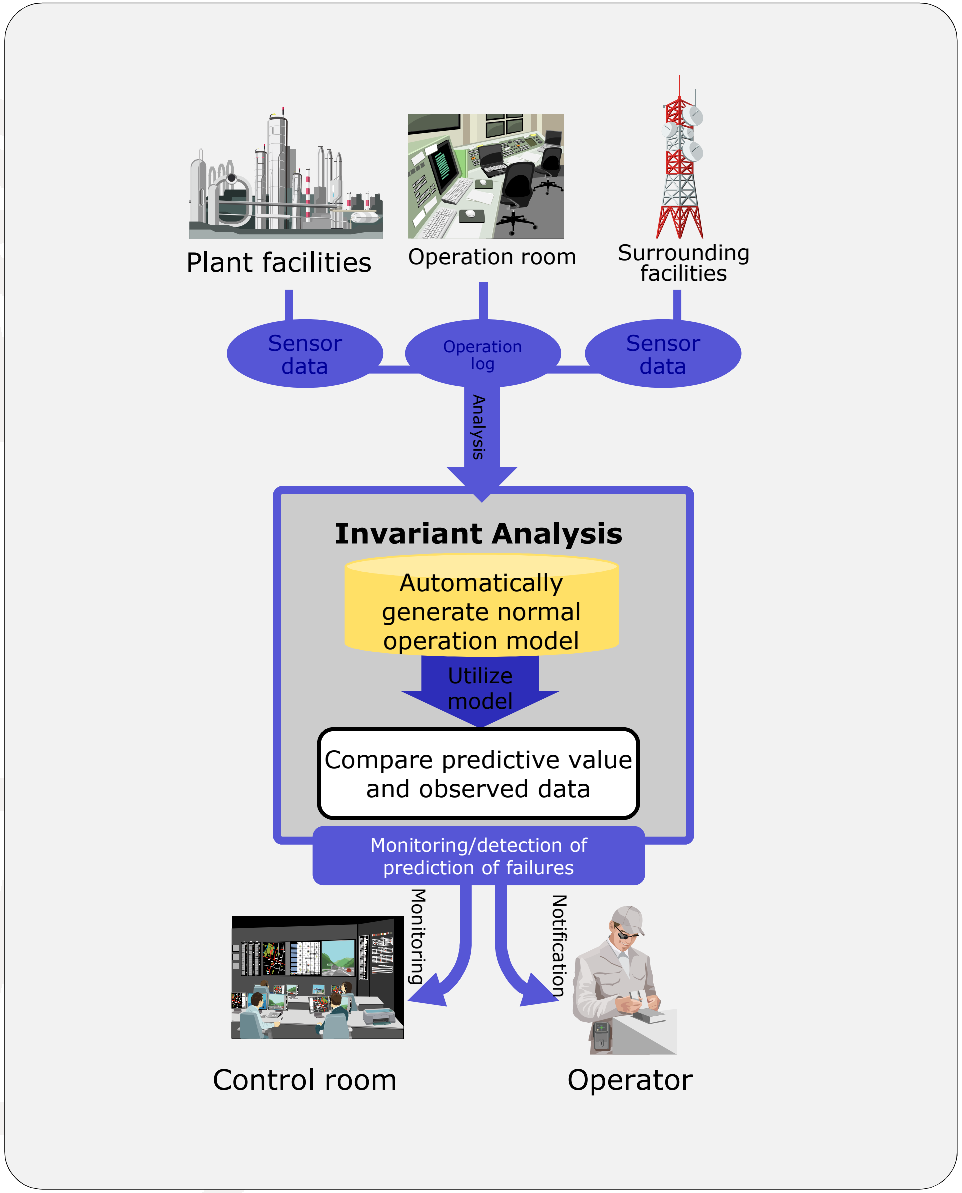
Effect of solution

- Monitor/detect prediction of failures of plant facilities
- Detect abnormalities from large volume sensor data at an early stage, avoid large-scale damage before happens

Point of introduction

- Visualize operational status from the existing data by "Invariant Analysis Technology"
- Utilize massive data in real time and realize high accuracy monitoring/detection of prediction of failures

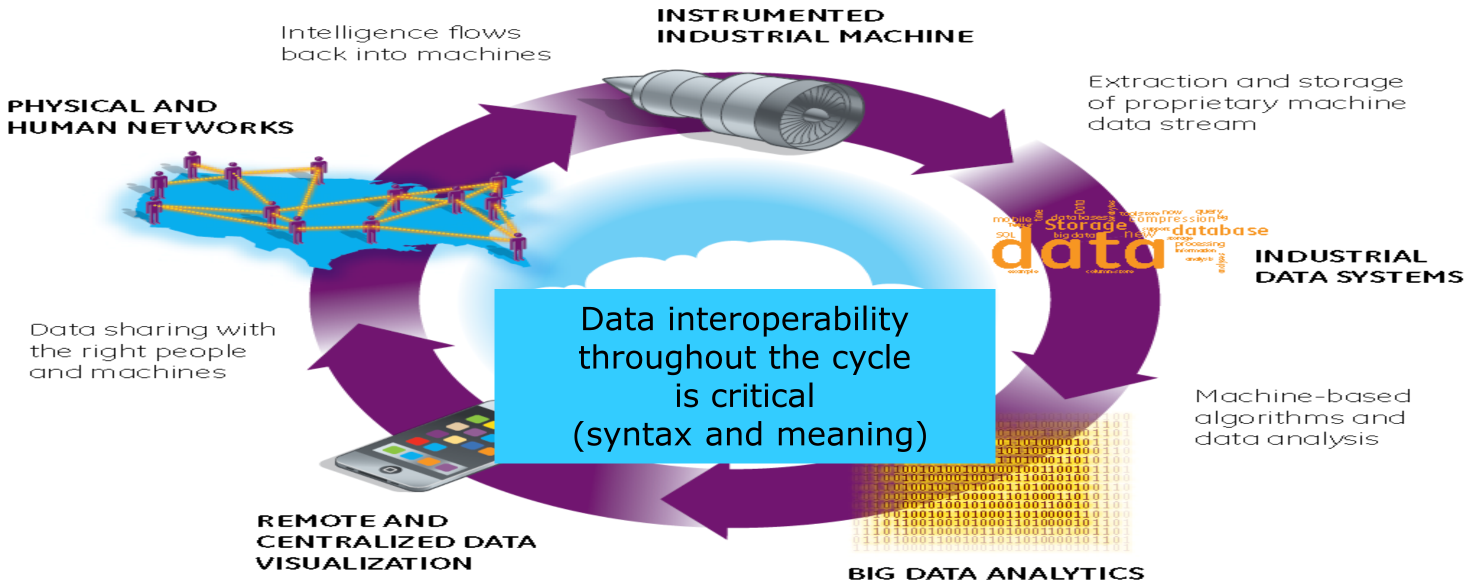
"Invariant Analysis" (NEC's leading analysis technology): analysis of large amounts of metric data collected from multiple sensors to automatically identify relationships and detect anomalies.



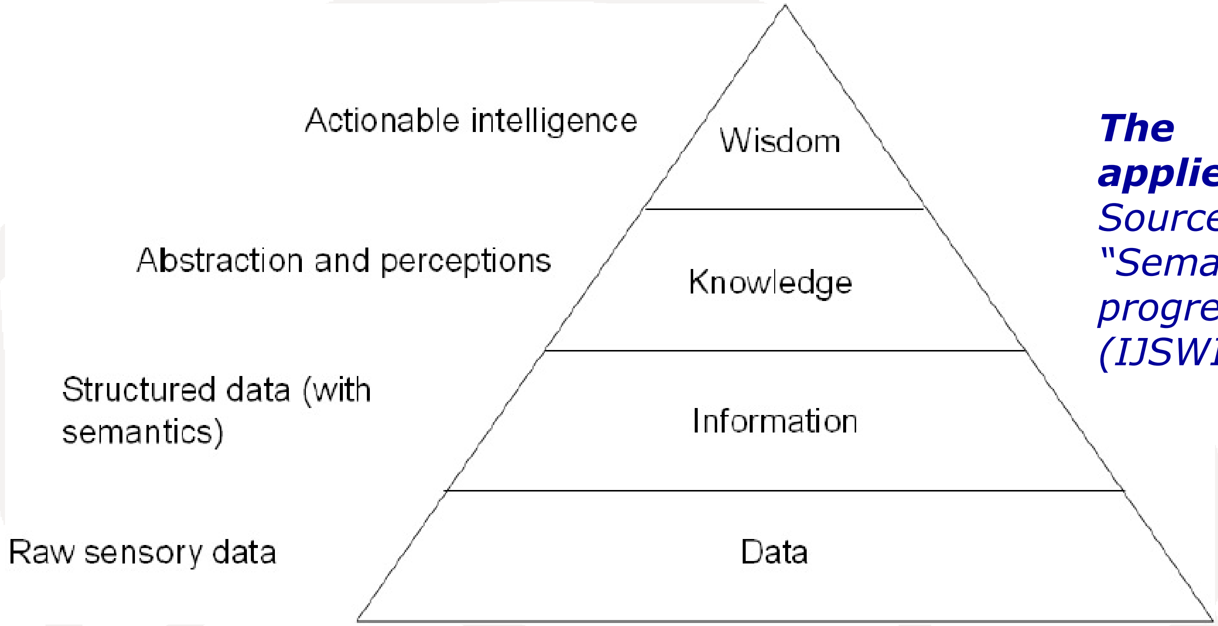
IoT and (Big) Data

- Some analysts indicate that by 2020 40% of data will come from sensors
- Multiple data sources: (real-time) data from things and context, historical and social data (cross-domain data exchange and correlation etc.)
- Data are mainly semi-structured and unstructured
- Data may have different precision and confidence levels
- Various operations to be made on data for the extraction of actionable intelligence (collection, de-noising, aggregation, adaptation, analysis)
- Raw data -> Information -> Knowledge (-> Wisdom)
- Target: the right data, at the right time, at the right location (e.g. cloud versus edge computing)

The Industrial Internet Data loop [source: GE whitepaper]



IoT data transformation



The knowledge hierarchy applied in IoT Data processing
Source: Barnaghi and al., "Semantics for the IoT: early progress and back to the future" (IJSWIS, 2012)

- Things generate raw data
- Additional information enables creation of structured metadata (first step of IoT data enrichment)
- Abstractions and perceptions give detailed insights by reasoning using knowledge (ontologies, rules) of different domains (second step of IoT data enrichment)
- Actionable intelligence allows decision making

Some key technical challenges of data in IoT

- Dealing with the data “V”s: Volume, Variety, Velocity, Variability, Veracity
- Discovery of appropriate device and data sources
- Availability and (open) access to data resources and data, security and privacy of data (incl. doing mining, analytics)
- Data querying
- Integration of heterogeneous devices, networks and data
- Scalability to cope with large numbers of devices, diverse and huge data, computational complexity of data interpretation
- Massive data mining, adaptable learning and efficient computing and processing
- Interpretation: extraction of actionable intelligence from data

- Non-technical challenges are essential though (e.g. data ownership and governance)

Semantics for the IoT (and data in the IoT)

Semantics is the study of meaning

- A definition: "The rules and conventions governing the interpretation and assignment of meaning to constructions in a language" [ITU-T Z.341]
- Semantics for the IoT: Shared vocabularies for IoT entities and their relationships

Requirements for interoperability, scalability, consistency, discovery, reusability, composability, automatic operations, analysis and processing of data and services are becoming more and more essential in the IoT

- Given the growing number of interconnected things and connections, variety of devices and connectivity, volume and types of generated data, number and type of services

Semantics based approaches reveal outstanding features to support these dynamic and cross-domain requirements of the IoT

Value proposition of semantics based approaches

Outstanding features for data and services, including an answer to some key technical challenges of data in the IoT

- Interoperability: the interoperability level of data and services of the IoT within one application domain and/or among different application domains can be improved.
- Scalability: data and services can be managed locally to IoT components. This increases the IoT technical component independence (loosely coupled components) and decentralizes the management, leading to enhanced scalability. Service reachability can be more easily expanded to reach more users, functional evolution of services can be rationalized.
- Consistency: data and services can refer to same meaning across time, location, IoT components.
- Re-usability: data and services can be reused and composed to construct new data, services.
- Analytics and actionable knowledge: merging, correlation and analysis of diverse data generated by the IoT, together with data from external sources such as social media, events and news, can be facilitated in order to produce actionable knowledge.
- Human-machine interaction: on one hand, since semantic technologies are based on natural human concepts, data and services become easier for humans to understand. On the other hand, since semantic technologies are formal ways to express concepts, data and services can also be understood by machines. This can improve the interaction between humans and the IoT.

But still issues concerning the full applicability of semantic technologies in the IoT

The benefits of semantic technologies for the IoT will be realized in incremental way: current issues include

- Lack of elaborated use cases as drivers to validate the value proposition (value for the different stakeholders of the value chain)
- Insufficient link with IoT architectures
- Immaturity of tools, essential to establish semantics based bridging (domain specific ontologies formalizing the meaning of domain data and information models; semantic merging, matching and alignment strategies; mediators to enable integration of disparate data resources; bridges etc.)
- Semantic discovery of services, devices, things and their capabilities
- Semantic metadata framework and IoT base/core ontology (horizontal cross-domain integration)
- Participation of domain specific communities and creation of social and business incentives for sharing
- Education (entrepreneurs/domain experts – developers interaction)

The initial results are promising, but further research, development, validation and standardization is required.

ITU-T studies on semantics for IoT

Initial study: ITU-T Y.2076 "*Semantics based requirements and framework for the IoT* "*[currently in AAP]*

- Value proposition of the usage of semantics in IoT systems, semantics based requirements at different layers and cross-layer capabilities of the IoT Reference Model [Y.2060], semantics based capability framework of the IoT
- Liaised to W3C and oneM2M to promote collaboration
- Some content contributed as ITU-T liaison into the Alliance for IoT Innovation (AIOTI) – Y.2076 is referred in the Oct 2015 Technical Report on "Semantic Interoperability" from AIOTI WG03 (the IoT standardization WG)

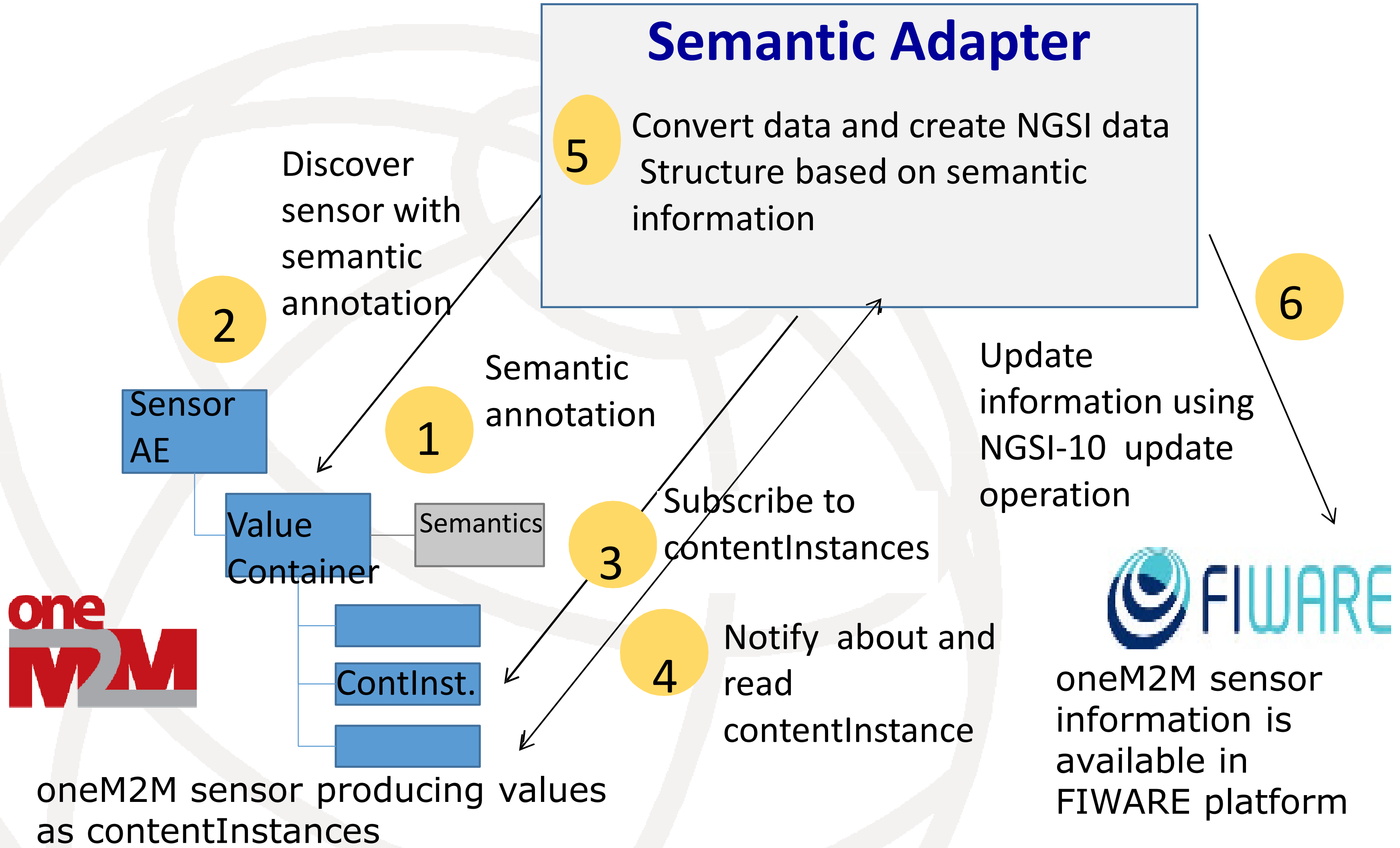
Further work expected in SG20 in short-medium term

- Other studies of common interest and application-domain specific studies – contributions in this Singapore meeting !

Two other SDOs with relevant studies on semantics

- W3C – has pioneered with the Semantic Web, other studies ongoing
- oneM2M – ongoing study on base ontology (ETSI TS 103 264 "Smart Appliances Reference Ontology (SAREF) " will be fully mapped to it)

Semantic interoperability between oneM2M and FIWARE



Demo realized between Easy Global Market, KETI, NEC, Sejong University and Telefonica at the Dec 2015 ETSI M2M workshop

Big Data technologies: value proposition and current critical aspects

Big Data technologies address several data challenges in the IoT

- Scalability, data integration, massive data mining, data accessibility

and create value (Big Data Analytics)

But some aspects are still critical from a deployment perspective

- Integration and interoperability with legacy environments and applications
- System performances and reliability
- Wide spectrum of technologies and products, organizational impact, skilled personnel
- And analytics is becoming a multi-dimensional challenge (data at rest versus data in motion, and the related data cycle operations)
- Data privacy, access, security

Big Data technical standardization

Potential areas of standardization related to Big Data interoperability [Y.Suppl.BigData-RoadMap, *AFNOR's 2015 Big Data white paper*]

- Requirements and use cases
- Architecture, data model and APIs
- Network-driven data analytics
- Security and data protection, anonymization and de-identification of personal data
- Framework for data quality and veracity
- Standards and guidelines to address issues related to legal implications of big data in telecommunications (e.g. data ownership)
- Framework and standards for telecom big data exchange
- *Benchmarks for system performance evaluation (e.g. Hadoop)*
- *Standardized visualization methods*
- *NoSQL query languages*

Main ITU-T developments related to Big Data

Completed studies

- **Y.2066:** *Common requirements of the IoT*
  - It includes requirements on data aspects of the IoT
- **Y.2068:** *Functional framework and capabilities of the IoT*
  - It includes IoT capabilities for data management and for integration of Big Data technologies and Cloud Computing technologies with the IoT
- **Y.3600**: *Big data – cloud computing based reqts and capabilities*

Ongoing studies

- **Y.IoT-BigData-reqts**: *Specific reqts and capabilities of IoT for Big Data*
- **Y.Suppl.BigData-RoadMap:** *Big Data Standardization Roadmap*
  - standards landscape in ICT sector and al., feasibility studies on Big Data standardization, identification of technical areas (e.g. IoT), roadmap
- **Y.BigDataEX-reqts:** *Big Data Exchange Framework and Reqts*
- **Y.BDaaS-arch:** *Functional architecture of Big Data as a Service*

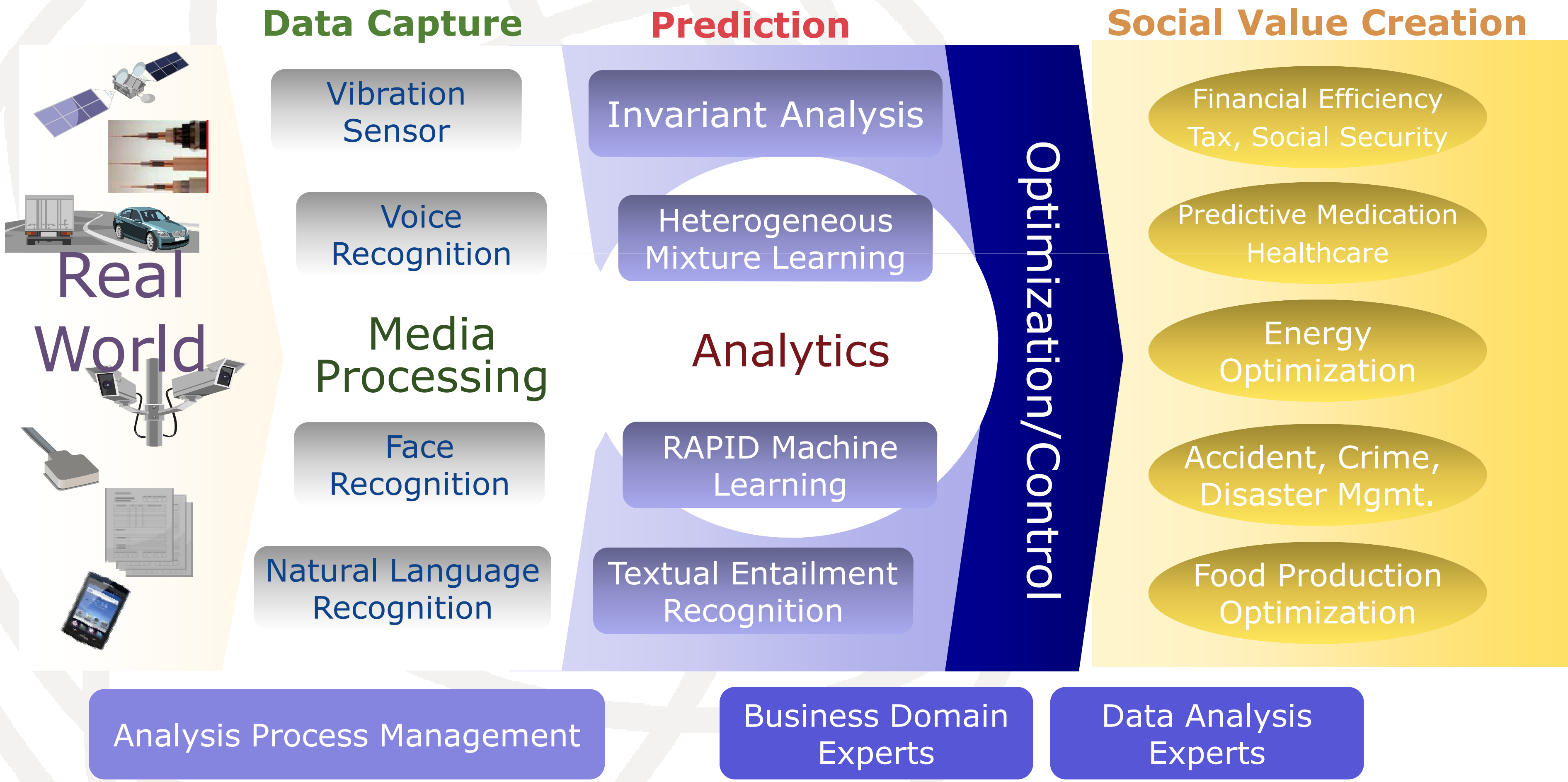
A number of other standards initiatives beyond ITU-T

- ISO/IEC JTC1 WG9, W3C Data Activity group, others

Thank you very much for your attention

NEC's technologies to transform Big Data into Values

Social value creation based on 1) edge devices and data collection platform, 2) advanced analytics technologies, and 3) advanced analysis expertise and process management



World leading analysis technologies from NEC

World leading analysis technologies for knowledge acquisition from big data

World 1st
Analyzes large amounts of metric data collected from multiple sensors to automatically identify relationships and detect anomalies.
Invariant Analysis

World 1st
Automatically detects massive patterns hidden in big data. Achieves high-precision predictions and anomaly detections that are difficult by handcrafted data analytics.
Heterogeneous Mixture Learning(HML)

World #1※1
Higher speed and lighter processing deep learning based matching engine
RAPID Machine Learning

World #1※2
Delivers best-in-class performance in NIST testing. Recognizes semantic content between texts.
Recognizing Textual Entailment

World 1st
Collecting personal face features from surveillance camera, and checking with particular person in real time
Face Recognition

※1 PTP: Penn Treebank Project 2011
※2 NIST : National Institute of Standard and Technology

Marco Carugi

Marco Carugi works as consultant on advanced telecommunication technologies and associated standardization, and is currently contractor for NEC representing the company in standards development activities.
During his professional career, he has worked as Telecommunication Engineer in the Solvay group, Research Engineer in Orange Labs, Senior Advisor in the Nortel Networks CTO division and Senior Expert in the Technology Strategy department of ZTE R&D.
Marco is active in standardization since long time, and has held numerous leadership positions, including ITU-T SG13 Vice-Chair, ITU-T Rapporteur in last three study periods, ITU-T FG Cloud Computing WG Chair, OIF Board member, IETF Provider Provisioned VPN WG co-Chair.
He is currently Rapporteur for Question 2 - "Requirements and use cases for IoT" - in ITU-T SG20 (Internet of Things and its applications including smart cities and communities), still maintaining the Rapporteurship for Question 2 - "Requirements for NGN evolution and its capabilities including support of IoT and SDN" - in ITU-T SG13 (Future networks) where he also acts as SG13 Mentor.
Marco has led the development of technical specifications on requirements, capabilities and services for IoT/M2M since the creation of the ITU-T IoT Global Standards Initiative, acting also as convenor of the IoT work plan sessions within IoT-GSI. He is currently the ITU-T JCA-IoT/SG20 Liaison Officer to ISO/IEC JTC1/WG10 and to the Alliance for IoT Innovation (AIOTI) (WG3 on IoT standardization), participating regularly in the activities of the European Commission's Internet of Things Research Cluster [co-author of IERC books edition 2014 and 2013, and AIOTI WG3 Rel.1/2 deliverables on IoT standards landscape, high level architecture and semantic interoperability]. He has also acted as vice-chair of the past ITU-T Focus Group on M2M Service Layer.
NGN evolution and Future Networks, SDN, Cloud Computing and Big Data are other technical areas in which he is involved at present.
Marco has led the development of numerous standards specifications and published in technical journals and books.
He holds an Electronic Engineering degree in Telecommunications from University of Pisa (Pisa, Italy), a M.S. in Engineering and Management of Telecommunication Networks from National Institute of Telecommunications (Evry, France) and a Master in International Business Development from ESSEC Business School (Paris, France). He is currently completing a Certificate on Big Data at Ecole Centrale (Paris, France).