|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | FG-AI4H-I-029-R03 | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | Plenary | E-meeting, 7-8 May 2020 | |
| **DOCUMENT** | | | | |
| **Source:** | | CAICT (China) | | |
| **Title:** | | Proposal to set up an overview document of all deliverables and continuously update it | | |
| **Purpose:** | | Discussion | | |
| **Contact:** | | Shan Xu CAICT China | | Email: [xushan@caict.ac.cn](mailto:edwinjrwu@tencent.com) |

|  |  |
| --- | --- |
| **Abstract:** | This document proposes to set up a continuously updating document on the Deliverables’ structure and a summary of important information of both generalized documents and existing topic groups. This document aims to improve overall understanding of the FG-AI4H activities and could not only be used for external promotion of outcome, but more importantly to strengthen internal collaboration, especially between generalized documents (DEL 1-9) with related scope, and the collaboration between generalized documents (DEL 1-9) and TGs (DEL 10.1-10.19) on a certain specification considerations.  NOTE – This revision provides additional information and some updates as indicated in revision marks, relative to the previous version. |

# Introduction

The ITU/WHO Focus Group on artificial intelligence for health (FG-AI4H) was established by ITU-T Study Group 16 at its meeting in Ljubljana, Slovenia, 9-20 July 2018.  This group is committed to establish a standardized assessment framework for the evaluation of AI-based methods for health, diagnosis, triage or treatment decisions. A list of deliverables for the FG-AI4H was planned and corresponding research groups (or con-calls) was established, with 9 deliverables focus on generalized specifications including ethics, regulatory, requirement, data, training, evaluation, application, etc., and 19 focus on use cases within specific health domains with corresponding AI/ML tasks. As the drafting work progresses, there is an increasingly recognized importance to strengthen the collaboration between generalized specifications (DEL 1-9), and their collaboration between the 19 topic groups (DEL 10.1-10.19). Reasons are as below:

* It is better to have a clear understanding on the scope for editors and contributors generalized specifications (DEL 1-9), to avoid overlap and still maintain specialized.
* Drafting the generalized specifications (DEL 1-9) may need the input from different use cases in specific health scenarios (DEL 10.1-10.19), including data preparation, training and testing methodologies, then generalized into an applicable framework.
* Topic groups (DEL 10.1-10.19) may also wonder novel to deal with heath problem within their domain. AI technologies are general, experiences from other TGs with similar tasks is worth reference.

To improve the possible collaboration above, providing a quick, intuitive, updating and comprehensive information collection on the all deliverables is important. This document proposed to set up an updating document on deliverables structure, abstract summary, structured collection of key messages from all existing topic groups. It can also act as a quick guild for new participants to understand FG-Ai4H activities and choose the work they are most interested in.

# Deliverable structure

A table of planned deliverables was developed. To make the relationship between different deliverables clearer, a figure on deliverables structure was given as below. According to different characteristics, the planned deliverables are divided into two group:

* Generalized specifications (DEL 1-9): focus on generalized specifications including ethics, regulatory, requirement, data, training, evaluation, application, etc. Each part is interconnected and together form a whole. The arrow in the figure is intended to indicate sequential connections from the perspective of software development and implementation.
* Topic groups (DEL 10.1-10.19): focus on use cases in specific health domains with corresponding AI/ML tasks. Each case could be seen as an example of a whole process recommended by generalized specifications (DEL 1-9), and also adapt to some specific application scenarios.

Please kindly note that this figure will be continuously updated according to the scope change of corresponding deliverables and newly established WGs or TGs, suggestions and comments from editors, contributors and experts are all welcome.

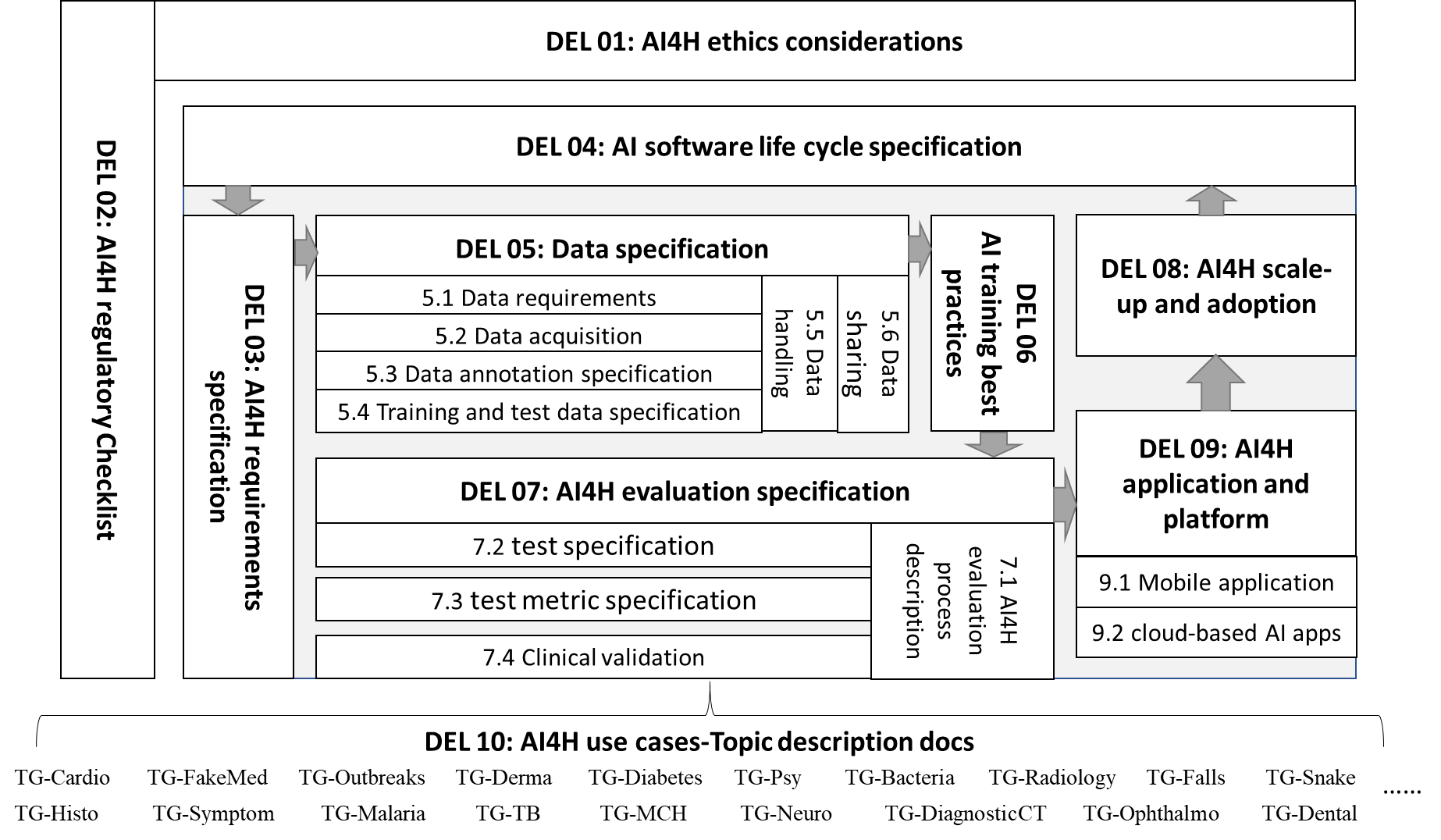


Figure 1 – Deliverables structure and relationship

A table for document links and contacts information is given as below, hopefully makes it easy for collaboration and dialogues between different deliverables.

Table 1 – Doc links and contacts for all deliverables

| No. | Deliverable | Updated initial draft editor | Availability\* |
| --- | --- | --- | --- |
| 1 | AI4H ethics considerations | [Andreas Reis](mailto:reisa@who.int) (WHO) | [G-201](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-201.docx) |
| 2 | AI4H regulatory [best practices | considerations] | [Pradeep Balachandran](mailto:abn.tvm@gmail.com) (India) and [Christian Johner](mailto:christian.johner@johner-institut.de) (Johner Institut, Germany) | [G-202](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-202.docx) |
| 3 | AI4H requirements specification | [Pradeep Balachandran](mailto:pbn.tvm@gmail.com) (India), [Tina Purnat](mailto:purnatt@who.int) (WHO) | [G-203](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-203.docx) |
| 4 | AI software life cycle specification | [Pat Baird](mailto:pat.baird@philips.com) (Philips, USA), [Tina Purnat](mailto:purnatt@who.int) (WHO) | [G-204](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-204.docx) |
| 5 | Data specification | [Marc Lecoultre](mailto:ml@mllab.ai) (ML Lab, Switzerland) | [G-205](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205.docx) |
| 5.1 | Data requirements | [Gupta Saurabh](mailto:drsaurabhmd@gmail.com) (AIIMS, India), [Manjula Singh](mailto:singhmanjula.hq@icmr.gov.in) (ICMR, India) | – |
| 5.2 | Data acquisition | [Rajaraman (Giri) Subramanian](mailto:kinnal@hotmail.com) (Calligo Tech, India), [Vishnu Ram](mailto:vishnu.n@ieee.org) (India) | [G-205-A02](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205-A02.docx) |
| 5.3 | Data annotation specification | [Shan Xu](mailto:xushan@caict.ac.cn) (CAICT, China), [Harpreet Singh](mailto:hsingh@bmi.icmr.org.in) (ICMR, India), [Sebastian Bosse](mailto:sebastian.bosse@hhi.fraunhofer.de) (Fraunhofer HHI, Germany) | [G-205-A03](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205-A03.docx) |
| 5.4 | Training and test data specification | [Luis Oala](mailto:luis.oala@hhi.fraunhofer.de) (Franhofer HHI, Germany), [Pradeep Balachandran](mailto:abn.tvm@gmail.com) (India) | [G-205-A04](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205-A04.docx) |
| 5.5 | Data handling | [Marc Lecoultre](mailto:ml@mllab.ai) (ML Lab, Switzerland) | [G-205-A05](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205-A05.docx) |
| 5.6 | Data sharing practices | [Ferath Kherif](mailto:Ferath.Kherif@chuv.ch) (CHUV, Switzerland), [Banusri Velpandian](mailto:banusrir@gmail.com) (ICMR, India), WHO Data Team | [G-205-A06](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-205-A06.docx) |
| 6 | AI training best practices specification | [Sim Xinming](mailto:xinming@aisingapore.org) and [Stefan Winkler](mailto:stefan@aisingapore.org) (AI Singapore) | [G-206](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-206.docx) |
| 7 | AI4H evaluation specification | [Markus Wenzel](mailto:markus.wenzel@hhi.fraunhofer.de) (Fraunhofer HHI, Germany) | [G-207](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-207.docx) |
| 7.1 | AI4H evaluation process description | [Sheng Wu](mailto:wus@who.int) (WHO) | [G-207-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-207-A01.docx) |
| 7.2 | AI technical test specification | [Auss Abbood](mailto:abbooda@rki.de) (Robert Koch Institute, Germany) | [G-207-A02](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-207-A02.docx) |
| 7.3 | AI technical test metric specification | [Luis Oala](mailto:luis.oala@hhi.fraunhofer.de) (Fraunhofer HHI, Germany) | [G-207-A03](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-207-A03.docx) |
| 7.4 | Clinical validation | [Naomi Lee](mailto:naomi.lee@lancet.com) (Lancet, UK), [Manjula Singh](mailto:singhmanjula.hq@icmr.gov.in) (ICMR, India), [Rupa Sarkar](mailto:rupa.sarkar@lancet.com) (Lancet, UK) | – |
| 8 | AI4H scale-up and adoption | [Sameer Pujari](mailto:pujaris@who.int) (WHO) | – |
| 9 | AI4H applications and platforms | [Manjeet Chalga](mailto:chalgams.hq@icmr.gov.in) (ICMR, India), [Aveek De](mailto:aveek@cms-india.org) (CMS, India) | [G-209](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-209.docx) |
| 9.1 | Mobile applications | [Khondaker Mamun](mailto:mamun@cse.uiu.ac.bd) (UIU, Bangladesh), [Manjeet Chalga](mailto:chalgams.hq@icmr.gov.in) (ICMR, India) | [G-209-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-209-A01.docx) |
| 9.2 | Cloud-based AI applications | [Khondaker Mamun](mailto:mamun@cse.uiu.ac.bd) (UIU, Bangladesh), [Manjeet Chalga](mailto:chalgams.hq@icmr.gov.in) (ICMR, India) | [G-209-A02](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-209-A02.docx) |
| 10 | AI4H use cases: Topic description docs. | [Eva Weicken](mailto:eva.weicken@hhi.fraunhofer.de) (Fraunhofer HHI, Germany) | [G-210](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-G-210.docx) |
| 10.1 | Cardiovascular disease risk prediction (TG-Cardio) | [Benjamin Muthambi](mailto:brm5@caa.columbia.edu) (Watif Health, South Africa) | [H-006-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-006-A01.docx) |
| 10.2 | Dermatology (TG-Derma) | [Maria Vasconcelos](mailto:maria.vasconcelos@fraunhofer.pt) (Fraunhofer Portugal) | [H-007-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-007-A01.docx) |
| 10.3 | Diagnosis of bacterial infection and anti-microbial resistance (TG-Bacteria) | [Nada Malou](mailto:nada.malou@paris.msf.org) (MSF, France) | – |
| 10.4 | Falls among the elderly (TG-Falls) | [Inês Sousa](mailto:ines.sousa@fraunhofer.pt) (Fraunhofer Portugal) | [H-012-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-012-A01.docx) |
| 10.5 | Histopathology (TG-Histo) | [Frederick Klauschen](mailto:frederick.klauschen@charite.de) (Charité Berlin, Germany) | [H-013-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-013-A01.docx) |
| 10.6 | Malaria detection (TG-Malaria) | [Rose Nakasi](mailto:g.nakasi.rose@gmail.com) (Makerere University, Uganda) | [H-014-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-014-A01.docx) |
| 10.7 | Maternal and child health (TG-MCH) | [Raghu Dharmaraju](mailto:rdharmaraju@gmail.com) (Wadhwani AI, India) and [Alexandre Chiavegatto Filho](mailto:alexdiasporto@usp.br) (University of São Paulo, Brazil) | [H-015-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-015-A01.docx) |
| 10.8 | Neurological disorders (TG-Neuro) | [Marc Lecoultre](mailto:ml@mllab.ai) (ML Labs, Switzerland) | [H-016-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-016-A01.docx) |
| 10.9 | Ophthalmology (TG-Ophthalmo) | [Arun Shroff](mailto:arunshroff@gmail.com) (MedIndia) | [H-017-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-017-A01.docx) |
| 10.10 | Outbreak detection (TG-Outbreaks) | [Stéphane Ghozzi](mailto:ghozzis@rki.de) (Robert Koch Institute, Germany) | [H-018-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-018-A01.docx) |
| 10.11 | Psychiatry (TG-Psy) | [Nicolas Langer](mailto:n.langer@psychologie.uzh.ch) (ETH Zurich, Switzerland) | [H-019-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-019-A01.docx) |
| 10.12 | AI for radiology (TG-Radiology) | [Darlington Ahiale Akogo](mailto:darlington@gudra-studio.com) (minoHealth AI Labs, Ghana) | – |
| 10.13 | Snakebite and snake identification (TG-Snake) | [Rafael Ruiz de Castaneda](mailto:Rafael.RuizDeCastaneda@unige.ch) (UniGE, Switzerland) | [H-020-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-020-A01.docx) |
| 10.14 | Symptom assessment (TG-Symptom) | [Henry Hoffmann](mailto:henry.hoffmann@ada.com) (Ada Health, Germany) | [H-021-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-021-A01.docx) |
| 10.15 | Tuberculosis (TG-TB) | [Manjula Singh](mailto:singhmanjula.hq@icmr.gov.in) (ICMR, India) | [H-022-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-022-A01.docx) |
| 10.16 | Volumetric chest CT (TG-DiagnosticCT) | [Kuan Chen](mailto:ckuan@infervision.com) (Infervision, China) | [H-009-A01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-009-A01.docx) |
| 10.17 | Dental diagnostics and digital dentistry (TG-Dental) | [Falk Schwendicke](mailto:falk.schwendicke@charite.de) and [Joachim Krois](mailto:Joachim.krois@charite.de) (Charité Berlin, Germany) | [H-010-A1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-H-010-A01.docx) |
| 10.18 | Falsified Medicine (TG-FakeMed) | [Franck Verzefé](mailto:fverzefe@gmail.com) (TrueSpec-Africa, DRC) | – |
| 10.19 | Primary and secondary diabetes prediction (TG-Diabetes) | [Andrés Valdivieso](mailto:avaldivieso@anastasia.ai) (Anastasia.ai & Tecnigen, Chile) | – |

# Roadmap of collaboration

<figure- to be added>

* **Horizontal perspective-** Generalized specifications (DEL 1-9): focus on generalized specifications including ethics, regulatory, requirement, data, training, evaluation, application, etc. It is a perspective of standards or specification development.
* **Vertical** **perspective** -Topic groups (DEL 10.1-10.19): focus on use cases in specific health domains with corresponding AI/ML tasks. It is better to have it as a guideline in specific direction

# Summary of generalized documents (DEL 1-9)

A summary of generalise documents from DL 1-9 will help to provide a clear understanding on each to avoid overlap and still maintain specialized. Some statements are extracted from documents updated in Brasilia meeting (22-24 January 2020) as examples, but it is better to have corresponding editors to review and provide an accurate version.

Table 2 – Summary of generalized documents (DEL 1-9)

| **Deliverable** | | **Scope/ overview/ objectives** |
| --- | --- | --- |
| **1- AI4H ethics considerations** | | This initial draft of the abstract describes the topics to be addressed in the forthcoming deliverable “AI for Health Ethics Considerations” to help seed future content. Digital technologies, machine learning and Artificial Intelligence (AI) are revolutionizing the fields of medicine, research and public health in an unprecedented manner. While holding great promise, this rapidly developing field raises a number of ethical, legal and social concerns, e.g. regarding equitable access, privacy, appropriate uses and users, liability and bias and inclusiveness. These issues are trans-national in nature, as capturing, sharing and using data generated and/or used by these technologies goes beyond national boundaries. The tools, methods and technologies used in “Big Data” and AI are being applied to improve health services and systems. However, many questions remain unanswered concerning the ethical development and use of these technologies, including how low- and middle-income countries will benefit from AI developments. A number of government agencies, academic institutions, NGOs and National Ethics Committees have started to address the ethical issues and challenges posed by digital technologies in general, but there remains no international guidance on the specific case of health. There is an urgent need to develop harmonised ethics guidance for the design and implementation of AI in global health. Moreover, to secure AI benefits at the global scale, a new collaborative research agenda should be established. |
| **2- AI4H regulatory [best practices |considerations]** | | High-level perspective and guidelines on the best practices for regulation of AI for health towards the safe application of AI solutions. It will include common elements of AI regulation across different verticals as well as some regional and country-specific regulations. |
| 2.1 | Mapping of IMDRF essential principles to AI for health software | AI for health (AI4H) software provides a number of new aspects that have not been considered when developing the regulatory framework for software as a medical device (SaMD) as described by the IMDRF Essential Principles (EPs) in  “Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices”, IMDRF Good Regulatory Review Practices Group, IMDRF GRRP WG/N47 FINAL, 31 October 2018.  This document provides a suggested mapping of the EPs to related aspects of AI4H software. Its purpose is to cover all aspects considered in the regulation of SaMDs and whether and if yes, how they are applicable to AI4H. |
| 2.2 | Regulatory checklist | This document defines a set of guidelines intended to serve the regulators and the device manufacturers with a common understanding on the best practices and processes that support a comprehensive requirements analysis to achieve regulatory compliance for AI for Medical Devices (AI-MD)  The **regulatory scope** of AI-MD, include (a) regulated and non-regulated medical devices, (b) medical devices with or without enforcement of regulations  The **regulatory requirements scope,** in this context**,** of AI-MD pertain only to technical aspects of AI/ML products; not to commercial or business aspects, such as strategic positioning, market assessment, profitability, etc.  The **product scope** of AI-MD,   * DOES include c) Software-as-a-Medical Device (SaMD) , (d) Software-in-a-Medical Device (SiMD) and (e) healthcare applications intended to improve medical outcomes or efficiency of healthcare system * DOES NOT include software applications for (a)healthcare facility administrative support, (b) for maintaining or encouraging healthy lifestyle, behaviour and wellness   This set of guidelines are not intended 1) to be comprehensive and/or 2) to replace any regulation, directive, standard, or similar legally-binding regulatory framework or guidance document of any geographic jurisdiction. |
| **3- AI4H requirements specification** | | The purpose of this document is to define the System Requirements Specifications (SyRS) that explains the informational, functional, behavioural and operational aspects a generic AI for health (AI4H) system.   * SyRS scope includes a requirements model that defines the informational, functional, behavioural and operational aspects of the AI4H system under consideration. Specific objectives include the following:   + Best practices for defining the AI software requirements and the task that the AI should solve without any ambiguity. This includes a clear description of the intended use   + Procedure to classify AI4H software vis-a-vis existing health interventions. Important considerations include, among others: Does the AI4H software replace components in existing health intervention workflows? Does it represent a new type of intervention?   + Risk management guidelines * This SyRS is generic in nature and shall be applicable across all domain specialties/ topic groups of AI4H FG. It may be modified, customized or extended appropriately to include the specific requirements and needs of the particular topic group under consideration * Requirement specifications may be defined in terms of use cases, graphical methods, mathematical models, documentation, etc. or combination of these |
| **4-AI software life cycle specification** | | 1. Identification of all standards and best practices that are relevant for the AI for health software life cycle. Similar to other software life cycle processes, the AI software life cycle process needs to be specified. 2. Summary and critical review of the identified documents including a discussion of their limits/gaps and need for action. 3. Identification of life cycle steps that are specific/characteristic for AI for health software, such as training and test procedures based on data that potentially need to be annotated. 4. Specification of the AI for health software life cycle and definition of best practices for the different life cycle steps in one document (under consideration of a, b, and c). Overview and examples of best practices |
| **5-Data specification** | | This deliverable combines a set of six separate deliverables as umbrella, which address six important aspects related to data specification when used for artificial intelligence (AI) and machine learning (ML) models/methods for health purposes. Each editor will propose an initial outline (=Table of Contents), define the objectives of the future deliverable, and collect a bibliography of existing literature and material relevant for the development of the respective document. A short call for participation, the expertise profile of potential contributors, a time plan, and a brief characterisation of the target audience serve as preface |
| 5.1 | Data requirements | <TBC> |
| 5.2 | Data acquisition | This document contains the proposed initial structure for the FG-AI4H Deliverable 5B, “Data Acquisition”. It presents a framework for public healthcare data acquisition and management model based on standard protocol for its easy adoption by any country or international health organizations. This paper assumes basic digitization of electronic health record (EHR) at basic health facilities. There is a gap in developing an integrated and comprehensive framework that addresses the use of EHR in a standardized way for public health, privacy issue by anonymizing patient specific information, fusing multiple records with slight changes in the same information, augmenting a broad spectrum of contextual data, and so on. |
| 5.3 | Data annotation specification | This document describes the topics to be addressed in the forthcoming deliverable “DEL05-A03: Data Annotation Specification”. Data annotation would be one of the most dependable factors on model performance, it serves as one important aspect of data quality control on Artificial Intelligence for health. This document is committed to give a general guideline of data annotation specification, including definition, background and goals, framework, standard operating procedure, scenario classifications and corresponding criteria, as well as recommended metadata, etc. A questionnaire is attached to seek input and collaboration with topic groups in FG-AI4H regarding data annotation |
| 5.4 | Training and test data specification | This document is intended to guide the target audience with a systematic way of preparing technical requirements specification for datasets used in training and testing of machine ML models  This document explains the best practices of data quality assurance aimed at minimizing the data error risks during the training and test data preparation phase of machine learning process lifecycle  The training and test data requirement specifications follow the data integrity, data security and data safety norms of the AI data governance lifecycle process |
| 5.5 | Data handling | This document outlines how data will be handled, once they are accepted. Health data are one of the most valuable and sensitive types of data. Handling this kind of data is often associated with a strict and factual framework defined by data protection laws. It is important to set a strict data policy which will ensure confidence in FG-AI4H not only among contributors, but across all stakeholders. There are two major issues that the data handling policy should address: (a) compliance with regulations dealing with the use of personal health data; and (b) non-disclosure of the undisclosed test data held by FG-AI4H for the purpose of model evaluation. |
| 5.6 | Data sharing practices | This deliverable aims to provide an overview of the existing best practices for data sharing of health-related data. The scope of this document includes a description of all the necessary steps and requirement to enable secure data sharing. The document specifies the role of the data providers, data processors and the data receivers. The document outlines established data sharing methods and novel methods based on distributed and federated environments for privacy preserving AI/ML models. |
| **6-AI training best practices specification** | | This document comprises two parts. The first summarises challenges encountered during the training of AI models and provides some suggested best practices, while the second recommends a framework which could be used for the transparent reporting of trained AI models. |
| **7-AI4H evaluation specification** | | In this introductory document, characteristics of health AI evaluation that are novel or otherwise essential are identified, best practices for the health AI model assessment are collected from selected sources, and requirements for a benchmarking platform are considered. This sets the scene for the four related documents DEL07.1-4 that dive into the details of health AI evaluation including a process description, technical tests, metrics and clinical validation |
| 7.1 | AI4H evaluation process description | The AI4H evaluation process description serves as overview of the state of the art of AI evaluation principles and methods and a forward-looking initiator for the evaluation process of AI4H. This process description includes a review of existing evaluation principles and methods, evaluation need and solutions specific for AI4H. It will also look into ethics and risks aspects of AI4H evaluation. Furthermore, based on the fundamentals of AI, the description will gain insights on the direction of how the current evaluation methods evolve towards the concept of REAL AI. |
| 7.2 | AI technical test specification | This document specifies how an AI can and should be tested in silico. Among other aspects, best practices for test procedures known from (but not exclusively) AI challenges will be reviewed in this document. Important testing paradigms that are not exclusively related to AI applications should be mentioned too. |
| 7.3 | AI technical test metric specification | Review of test metrics and criteria that should be considered (performance measures, robustness, transparency/explainable machine learning, uncertainty quantification etc.). Etc. |
| 7.4 | Clinical validation | <TBC> |
| **8-AI4H scale-up and adoption** | | <TBC> |
| **9-AI4H applications and platforms** | | This document contains a draft set of rules for development of AI tool for Health using Mobile Applications & Cloud-based AI applications, their testing and benchmarking. This document also invites Medical & AI researchers to collaborate in development of Cloud-based / Mobile Application based AI tools for Health within the International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG-AI4H). |
| 9.1 | Mobile applications | This document contains a draft set of rules for development of AI tool for Health using Mobile Applications, their testing and benchmarking. The objectives of the topic groups are:   1. to prepare the rules for development of AI tool for Health using Mobile Applications 2. to discuss the regulatory/ethical rules for Mobile Apps with AI for Healthcare 3. to provide a forum for open communication among various stakeholders, 4. to develop benchmarking for the Apps, 5. to coordinate the complete process in collaboration with the Focus Group management and working groups. |
| 9.2 | Cloud-based AI applications | This document contains a draft set of rules for development of Cloud-based AI applications, their testing and benchmarking. The objectives of the topic groups are as follows:   1. to discuss on technology, security and legal issues related to cloud-based AI tools 2. to provide a forum for open communication among various stakeholders, 3. to coordinate the benchmarking process in collaboration with the Focus Group management and working groups. |

# Key message of TGs (DEL 10.1-10.19)

Maintaining an index of key message of different topics group may provide a quick, structured and targeted input from different use cases in specific health scenarios (DEL 10.1-10.19) to develop generalized documents (DEL 1-9). Meanwhile, AI technologies are general, experiences from other TGs with similar tasks is worth reference on dealing with specific health problems in topic groups (DEL 10.1-10.19).

Table 3 – Key message of Topic Groups

| Topic Groups (Examples) | Domain (Cardiovascular/ Dermatology/ Histopathology/‌etc) | Task (Classification/ detection/ segmentation/ prediction/‌etc) | Gold Standard (state-of-the-art task intervention method) | Input data type (Text/ Image/ video/ audio/ numerical/‌etc) | Testing/ Training dataset (Public dataset/ Collected by myself/‌etc) | Data annotation (Procedure/ annotator number/ tool/‌etc) | Algorithm (specific model used in this TG) | Evaluation (Metrics used in this TG) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| TG-Cardio |  |  |  |  |  |  |  |  |
| TG-Derma |  |  |  |  |  |  |  |  |
| TG-Bacteria |  |  |  |  |  |  |  |  |
| TG-Falls |  |  |  |  |  |  |  |  |
| TG-Histo |  |  |  |  |  |  |  |  |
| TG-Malaria |  |  |  |  |  |  |  |  |
| TG-MCH |  |  |  |  |  |  |  |  |
| TG-Neuro |  |  |  |  |  |  |  |  |
| TG-Ophthalmo |  |  |  |  |  |  |  |  |
| TG-Outbreaks |  |  |  |  |  |  |  |  |
| TG-Psy |  |  |  |  |  |  |  |  |
| TG-Radiology |  |  |  |  |  |  |  |  |
| TG-Snake |  |  |  |  |  |  |  |  |
| TG-Symptom |  |  |  |  |  |  |  |  |
| TG-TB |  |  |  |  |  |  |  |  |
| TG-DiagnosticCT |  |  |  |  |  |  |  |  |
| TG-Dental |  |  |  |  |  |  |  |  |
| TG-FakeMed |  |  |  |  |  |  |  |  |
| TG-Diabetes |  |  |  |  |  |  |  |  |

# Ongoing collaboration activities

Some ongoing collaboration activities can be recorded here as information sharing.

* DEL 02 regulatory checklist sent to Regulatory considerations on AI for health (WG-RC) for review?
* DEL 5.3 Data annotation specification designed a questionnaire to seek input from different Topic Groups: <https://forms.gle/3fYrm3SZSrNQu3eeA>

# Update mechanism

* Confirm with editors of different deliverables if the initial content in Table 2 is accurate. If not, kindly ask them to provide an abstract paragraph as scope.
* Review all TGs, and fill in the blanks in Table 3, and invite topics drives to have a review
* After a first mature version, periodically review it every meeting of FG-Ai4H, and provide an update after the meeting.
* Revise it if there is any scope change of corresponding deliverables or newly established WGs or TGs

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_