|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | FG-AI4H-F-021 | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | Plenary | Zanzibar, 2-5 September 2019 | |
| **DOCUMENT** | | | | |
| **Source:** | | WG-DAISAM Chair | | |
| **Title:** | | WG on Data and AI solution assessment methods (WG-DAISAM) – Approved ToRs | | |
| **Purpose:** | | Admin | | |
| **Contact:** | | Pat Baird Chair, WG-DAISAM Philips, USA | | Tel: +1 262 239-9321 Email: [pat.baird@philips.com](mailto:pat.baird@philips.com) |
| **Contact:** | | Luis Oala Vice-chair, WG-DAISAM Fraunhofer HHI, Germany | | Tel: +49 30 31002-104  Fax: +49 30 31002-190 Email: [luis.oala@hhi.fraunhofer.de](mailto:luis.oala@hhi.fraunhofer.de) |

|  |  |
| --- | --- |
| **Abstract:** | This document contains the Terms of Reference for the FG-AI4H Working Group on data and AI solution assessment methods (WG‑DAISAM), approved by e-mail consultation to the FG AI4H mailing list. |

# Motivation

The Focus Group AI4H set out to establish a standardized assessment framework for the evaluation of AI-based methods for health, diagnosis, triage or treatment decisions. This entails the selection of suitable test data sets and the automated evaluation of AI solutions on these test data sets. Such a selection and evaluation require (1) the identification and definition of assessment methods and (2) their algorithmic implementation to enable automated testing. Point (1) is the main rationale for this WG.

# Scope

This WG is concerned with the identification and definition of methods for data and AI solution assessment, while the implementation of said methods is coordinated with the WG Data and AI Solution Handling.

The modus operandi for this WG is in close cooperation with the TGs, in particular for the definition of task specific assessment methods for data and the AI solutions, as well as the other WGs that work on the technical implementation of the standardized assessment framework.

# Specific issues to be addressed include

Thus, the WG pursues the following tasks:

1. Identify relevant categories of AI solution assessment; e.g. performance, stability, or transparency among others; in close cooperation with the topic drivers from the TGs.
2. Specify concrete methods for AI solution assessment based on (1) task specific considerations and (2) state-of-the art machine learning research. This collection of methods shall include comments on benefits and limits of the included methods.
3. Identify relevant categories of data assessment; e.g. balance, representation or measurement quality among others; in close cooperation with the topic drivers from the TGs.
4. Specify concrete methods for data assessment based on (1) task specific considerations and (2) state-of-the art machine learning research. This collection of methods shall include comments on benefits and limits of the included methods.

# Deliverables

1. Recommendations for categories of data and AI solution assessment
2. Recommendations for data and AI solution assessment methods with comments on benefits and limits of the included methods
3. Implementation specifications of B. to WG Data and AI Solution Handling
4. Progress reports to FG

# Relationships

* Topic drivers (members of FG-AI4H) for the different topic areas
* Chairs of the other working groups of FG-AI4H
* ITU secretariat
* WHO representatives
* Workshop hosts (institutions, academy, industry)

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