|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | **FG-AI4H-C-030** | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | N/A | Lausanne, 22-25 January 2019 | |
| **DOCUMENT** | | | | |
| **Source:** | | Ada Health GmbH | | |
| **Title:** | | Suggestion for a generic topic description document (TDD) outline | | |
| **Purpose:** | | Discussion | | |
| **Contact:** | | Henry Hoffmann Ada Health GmbH Germany | | Tel: +49 177 6612889 Email: [henry.hoffmann@ada.com](mailto:henry.hoffmann@ada.com) |
| **Contact:** | | Andreas Kühn Ada Health GmbH Germany | | Tel: +49 30 60031987 Email: [andreas.kuehn@ada.com](mailto:andreas.kuehn@ada.com) |
| **Contact:** | | Johannes Schröder Ada Health GmbH Germany | | Tel: +49 30 60031987 Email: [johannes.schroeder@ada.com](mailto:johannes.schroeder@ada.com) |

|  |  |
| --- | --- |
| **Abstract:** | This document contains a draft outline of topic description document (TDD), prepared on request after the discussion on topic groups held on Thu 24 Jan. |

### ITU Document Header

...

# Introduction

## Document Structure

* overview of the whole document

## Topic Description

* description ot topic
* categorization of the topic according to categorization guideline (currently C-0xx)
* relevance of the health topic
* gold standard of current health topic handling
* possible impact of AI in this topic
* expected impact of the benchmarking

## Ethical Considerations

* ethical considerations on usage of AI
* ethical consideration of and benchmarking including its data acquisition

## Existing AI Solutions

* current systems available with their inputs, output, focus/bias
* existing benchmarking including self stated performance

## Existing work on benchmarking

* papers on existing attempts to benchmark solutions on the topic
* clinical evaluation attempts, RCT, etc.
* including existing numbers

# Method

* Overview of the benchmarking

## AI Input Data Structure

* possible inputs for benchmarking
* ontologies, terminologies
* data format

## AI Output Data Structure

* outputs to benchmark
* ontologies, terminologies
* data format

## Test Data Labels

* label types
* ontologies, terminologies
* data format

## Scores & Metrics

* which metrics & scores to use for benchmarking
* considering relation to parameters stakeholders need for decision making
* considering scores that providers use
* considering the scope providers designed their solutions for
* considering the state of the art in RCT, statistics, AI benchmarking etc.
* considering bias transparency

## Undisclosed Test Data Set Collection

* raw data acquisition / acceptance
* test data source(s): availability, reliability,
* labeling process / acceptance
* bias documentation process
* quality control mechanisms
* discussion of the necessary size of the test data set for relevant benchmarking results
* specific data governance derived by general data governance document (currently C-004)

# Benchmarking Methodology and Architecture

* technical architecture
* hosting (IIC, etc.)
* possibility of an online benchmarking on a public test dataset
* protocol for performing the benchmarking (who does what when etc.)
* AI submission procedure including contracts, rights, IP etc. considerations

# Reporting Methodology

* Report publication in papers or as part of ITU documents
* Online reporting
* public Leaderboards vs. private
* Credit-Check like on approved sharing with selected stakeholders
* Report structure including an example
* Frequency of benchmarking

## AI4H Topic group

* Topic group structure
* Topic group participation
* Tools/process of TG cooperation
* TG interaction with WG, FG
* Current topic group and topic status
* Contributors so far
* Next meetings
* Next steps for the work on this document

### Results

* insert here the reports of the different benchmarking runs

### Discussion

* Discussion of the insights from executing the benchmarking on
  + external feedback on the whole topic and its benchmarking
  + technical architecture
  + data acquisition
  + benchmarking process
  + benchmarking results

### Declaration of Conflict of Interest

* by each contributor to this document

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