

# TG Symptom call 2019-05-08

## 1. Welcome - approval of agenda

- Jason Maude from Isabel introduces himself
  - Isabel comes from a focus on healthcare professionals
  - <https://symptomchecker.isabelhealthcare.com/>
- Baidu has nominated Dr Chao Lu as their new representative for the topic group
  - Not on the call

## 2. Introduction / Organisational Topics

### Outreach

- There are some key players in the symptom-checker field that should be involved or at least made aware of the work
- The general call can be used to reach out to various key players to improve

### Are there constraints on the size of a functional TG?

- Can split into non-autonomous subgroups / can split the TG along thematic lines
- Though the TG is regarded as one organisational unit (no sub-topic organisations)

### Collaboration tools - Topic Description Document

- For now it's a google doc, Baidu can access, though not ideal
- Eventually migrate to TSBcloud (NextCloud with Collabora)

### Future online TG meetings

- Biweekly or on a needs basis?
- Work needs to be progressed between meetings rather than at meetings
- We possibly start bi-weekly after the next FG meeting

## 3. Topic Description Document (TDD) work planning

(The last published version of the TDD was [D-016](#) (Free ITU account required))

The goal of the Topic Description Document is to provide a complete specification of all relevant aspects of the benchmarking of AI based symptom assessment systems. Following the structure of [C-105](#), the general approach is to identify all relevant systems in this field, analyze their inputs, their benchmarkable outputs and their focus/specializations to then derive from this the scores, metrics and methodology for a first minimal viable benchmarking. This will also include exchange with field experts and all the stakeholders we aim to support with the benchmarking results in their decision making.

The scores and metrics section has not been written yet

- Proposal to progress this section to something that can be presented at the FG meeting
- Companies to look at metrics they use, and identify common denominators

## TDD structure

- From the FG output doc: [C-105](#)
- Possible to make some changes to improve legibility?
  - Chapter 1 and section 1.1 can be more accessible
  - Make more accessible where possible, without compromising on accuracy
    - Your.MD and Isabel agreed to have a look on how to improve this section
  - Changes welcome in remarks / suggestion mode
- Can contribute proposed TDD structure changes back upstream to FG for new output doc
- ITU deals with large documents 600 pages
  - Can be useful to split if structure permits, especially have more static and more changing documents
- The most important thing to change for the FG meeting is to introduce the sub-topic separation
  - The TD-driver suggested to not use the [FGAI4H-D-022](#) but to prefer a separate “method” chapter for each sub-topic.
- The list of other things to change for the next FG meeting are:
  - Adding Baidu and Isabel to the author-header, the system lists, and the declaration of interest.

## Benchmarking hierarchies and versions

- There are some questions about having parallel benchmarks, as well as updating benchmarks (keeping every deprecated version for reference).

## 4. AI crowd evaluation

Input document [FGAI4H-D-01](#) suggested to use AI Crowd as benchmarking platform. The FG created an Adhoc-Group for evaluating it at meeting D. The individual topic groups are asked to have a look at AI Crowd and assess if it can be used for their topic.

- The TG driver will have a look at AI crowd
- TG members might have a look at it (Isabel and Your.MD confirmed this)
- Aspects to look for are e.g.
  - Is it possible to benchmark our AI solutions in a technology independent way?
  - Can we implement all the specific scores, metrics and categories we need for our benchmarking?
  - Is the report generation flexible enough? For instance, if a stakeholder wants to know the best offline preclinical triage solution for pregnant 16-25 year old women in South Africa?
  - Can it potentially handle the complexities of simulating a dialog in later benchmarking iterations?
- We add a detailed section on AI Crowd to our TDD reflecting our joint assessment
- We given the Adhoc Group access to our insights
- Relevant readings/links include:
  - [https://github.com/Alcrowd/Alcrowd/blob/master/doc/grading\\_overview.md](https://github.com/Alcrowd/Alcrowd/blob/master/doc/grading_overview.md)
  - <https://github.com/Alcrowd/Alcrowd>
  - [www.aicrowd.com](http://www.aicrowd.com) (previously [www.crowdai.org](http://www.crowdai.org) , old challenges available here)

## 5. Meeting E, AI4Good Summit and others

- The deadline for TDD change submission is: **22nd May 23:59 CET**
- Please add contribution to the TDD until 20th May.
- Can the group sit down together?

- ITU will try to provide a room for a breakout session for the TG
- TG will be presented at AI for good health breakthrough session

## Annex A: Agenda

[1. Welcome - approval of agenda](#)

[2. Introduction / Organisational Topics](#)

[2.a FGAI4H-D-022](#)

[2.b FGAI4H-D-016](#)

[3. Topic Description Document \(TDD\) work planning](#)

[4. AI crowd evaluation](#)

[4.a FGAI4H-D-011](#)

[5. Meeting E. AI4Good Summit and others](#)

[Annex A: Agenda](#)

[Annex B: Attendees](#)

(from: <https://docs.google.com/document/d/1khfiAQldVCP21XTGxXyXmQ-4vJRpecNvVdgLo2INHa0/>)

## Annex B: Attendees

1. Henry Hoffmann ([Ada](#)) - Topic Driver
2. Andreas Kuhn ([Ada](#))
3. Shubs Upadhyay ([Ada](#))
4. Jonathon Carr-Brown ([Your.MD](#))
5. Jason Maude ([Isabel](#))
6. Pradeep Balachandran (FG-AI4H)
7. Vishnu Ram OV (FG-ML5G)
8. +393402378412
9. Simao Campos (ITU)
10. Bastiaan Quast (ITU)