|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-C-107 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | Lausanne, 23-25 January 2019 |
| **DOCUMENT** |
| **Source:** | Focus Group on “Artificial Intelligence for Health” (FG-AI4H) |
| **Title:** | Call for Topic Group Participation: Standardized benchmarking of <topic> |
| **Purpose:** | Engagement |
| **Contact:** | FG-AI4H secretariat | Tel: +41-22-730-6805Fax: +41-22-730-5853E-mail: tsbfgai4h@itu.int |

|  |  |
| --- | --- |
| **Abstract:** | Calling on members of the medical and artificial intelligence communities with a vested interest in <topic[[1]](#footnote-1)>! Become engaged in the group dedicated to establishing a standardized benchmarking platform for <topic1> within the International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG-AI4H). |

**Call for Topic Group Participation: <topic1>**

The International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG-AI4H; <https://www.itu.int/go/fgai4h>) seeks engagement from members of the medical and artificial intelligence (AI) communities (including clinicians, technologists, entrepreneurs, potential benchmarking data providers, machine learning experts, software developers, researchers, regulators, policy-makers, companies/institutions, and field experts) with a vested interest in shaping the benchmarking process of <topic1>.

# About FG-AI4H

Over the past decade, considerable resources have been allocated to exploring the use of AI for health, which has revealed an immense potential. Yet, due to the complexity of AI models, it is difficult to understand their strengths, weaknesses, and limitations. If the technology is poorly designed or the underlying training data are biased or incomplete, errors or problematic results can occur. AI technology can only be used with complete confidence if it has been quality controlled through a rigorous evaluation in a standardized way. Towards developing this standard assessment framework of AI for health, the ITU has established FG-AI4H in partnership with the WHO.

Thus far, FG-AI4H has established ten topic groups. The topic groups are: AI and cardiovascular disease risk prediction, dermatology, falls among the elderly, histopathology, neuro-cognitive diseases, ophthalmology (retinal imaging diagnostics), psychiatry, snakebite and snake identification, symptom checkers, and tuberculosis.

Each topic group agrees upon representative benchmarking tasks in a pragmatic, best-practice approach, which can later be scaled and expanded to similar tasks. Every benchmarking task should address a health problem of relevance (e.g. impacting a large and diverse part of the global population or challenging to treat) and for which AI technology would provide a tangible improvement relative to the current practice (e.g. better care, results, and/or cost/time effectiveness).

For a rigorous and sound evaluation, undisclosed test data sets must be available (or have to be collected) for each task. All data must be of high quality and compliant with ethical and legal standards. In addition, the data must originate from a variety of sources so that it can be determined whether an AI algorithm can generalize across different conditions, locations, or settings (e.g. across different people, hospitals, and/or measurement devices). The format/properties of the data serving as input to the AI and of the output expected from the AI, as well as the benchmarking metrics are agreed upon and specified by the topic group.

Finally, the AI-to-be-evaluated will be benchmarked with the undisclosed test data on FG-AI4H computing infrastructure. Here, the AI will process single samples of the undisclosed test data set and predict output variables, which will be compared with the "ground truth." The results of the benchmarking will be provided to the AI developers and will appear on a (potentially anonymized) leaderboard.

# Topic group: <topic1>

A topic group is a community of stakeholders from the medical and AI communities with a shared interest in a topic. The objectives of the topic groups are manifold:

1. to provide a forum for open communication among various stakeholders,
2. to agree upon the benchmarking tasks of this topic and scoring metrics,
3. to facilitate the collection of high quality labeled test data from different sources,
4. to clarify the input and output format of the test data,
5. to define and set-up the technical benchmarking infrastructure, and
6. to coordinate the benchmarking process in collaboration with the Focus Group management and working groups.

The primary output of a topic group is one document that describes all aspects of how to perform the benchmarking for this topic. (The document will be developed in a cooperative way by suggesting changes as input documents for the next FG-AI4H meeting that will then be discussed and integrated into an official output document of this meeting. The process will continue over several meetings until the topic description document is ready for performing the first benchmarking.)

This topic group is dedicated to <topic1>. <Here, please briefly introduce the topic, explain the relevance of the topic, clarify how AI could assist with this topic, describe the intended benchmarking task, and provide details about available (public data and) undisclosed test data.>

More details about the activities of the topic group can be found in the documents <insert links to all relevant documents on the ITU collaboration site>. These can be accessed with a free ITU account (cf. “Get involved”).

Current members of the topic group on <topic> include <please provide information about the existing members and their background/qualifications>. The topic group would benefit from further expertise of the medical and AI communities and from additional data. <Please provide any specific requirements for this topic group.>

# Get involved

To participate in the topic group on <topic1>, please complete the document template found at : [https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FG-AI4H-Doc-template.docx.](https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FG-AI4H-Doc-template.docx) First, update the document header as follows. WG(s): *N/A - place, date.* Source: *Your name.* Title: *Participation in the topic group on <topic1>.* Purpose: *Discussion.* Contact: *Insert contact name, contact organization, country, telephone, e-mail.* Abstract: *Short summary of main text.*

In the main text, introduce yourself/your organization, explain your interest in the topic group on <topic1>, describe your relevant experience and expertise, and indicate the potential availability of undisclosed test data. This document will be published on the FG-AI4H collaboration site. Meeting attendance is expected. Please email the completed document template to tsbfgai4h@itu.int.

Participation in FG-AI4H is free of charge and open to all. To attend the workshops and meetings (upcoming events are 2–5 April 2019 in Shanghai, China; ​29-31 May 2019 in Geneva, Switzerland; September 2019, t.b.d.; November 2019 in New Delhi, India), please create a free ITU account by registering on the Focus Group website (<https://www.itu.int/go/fgai4h>). On this website, you can also find the whitepaper and key output documents, and sign up to the mailing list.

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

1. AI for cardiovascular disease risk prediction, AI for dermatology, AI to prevent falls among the elderly, AI for histopathology, AI against neuro-cognitive diseases, AI for ophthalmology (retinal imaging diagnostics), AI for psychiatry, AI for snakebite and snake identification, AI-based symptom checkers, AI against tuberculosis. [↑](#footnote-ref-1)