



INTERNATIONAL TELECOMMUNICATION UNION

**TELECOMMUNICATION
STANDARDIZATION SECTOR**

STUDY PERIOD 2022-2024

FG-AI4H-P-021-A02

ITU-T Focus Group on AI for Health

Original: English

WG(s): Plenary Helsinki, 20-22 September 2022

DOCUMENT

Source: TG-Symptom Topic Driver

Title: Att.2 – CftGP (TG-Symptom) [same as Meeting N]

Purpose: Engagement

Contact: Topic Driver: Henry Hoffmann Email: henry.hoffmann@ada.com

Abstract: Calling on members of the medical and artificial intelligence communities with a vested interest in "AI-based symptom assessment". We encourage engagement in the group dedicated to establishing a standardized benchmarking platform for "AI-based symptom assessment" within the International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on "Artificial Intelligence for Health" (FG-AI4H).

Call for Topic Group Participation: "AI-based symptom assessment"

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 "AI-based symptom assessment".

About FG-AI4H

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

Thus far, FG-AI4H has established thirteen topic groups. These are concerned with AI and cardiovascular disease risk prediction, child growth monitoring, dermatology, falls among the elderly, histopathology, neuro-cognitive diseases, ophthalmology (retinal imaging diagnostics), psychiatry, radiotherapy, snakebite and snake identification, AI-based symptom assessment (previously named "symptom checkers"), tuberculosis, and volumetric chest computed tomography.

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 health challenge to treat) and be suitable for AI technology to provide a tangible improvement relative to current practice (e.g. better care, results, and/or cost/time effectiveness).

To conduct a rigorous and sound evaluation, undisclosed test data sets must be available (or be collected) for each task. All data must be high quality and compliant with ethical and legal standards. In addition, 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 and properties of the data serving as input to the AI and as the output expected from the AI, as well as the benchmarking metrics, will be 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, each AI will process single samples of the undisclosed test data set and predict output variables, which will be compared with the test set labels representative of "ground truth." The results of the benchmarking will be provided to the AI developers and will appear on a (potentially anonymized) leaderboard.

Topic group: "AI-based symptom assessment"

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 several-fold:

1. Provide a forum for open communication among various stakeholders
2. Agree upon the benchmarking tasks and scoring metrics
3. Facilitate the collection of high quality labeled test data from different sources
4. Clarify the input and output format of the test data
5. Define and set-up the technical benchmarking infrastructure

6. Coordinate the benchmarking process in collaboration with focus group management and working groups

The primary output of the topic group will be a document that describes all aspects relevant to the benchmarking process. The document will be developed cooperatively over several FG-AI4H meetings. Iterative versions of the document will be submitted as input documents for each meeting and the relevant changes will then be discussed and integrated into an official output document that will serve as the framework for the first official benchmarking.

This topic group is dedicated to "AI-based symptom assessment", one of the most promising applications in the field of AI4H. The World Health Organization estimates the shortage of Global Health workers to increase from 7.2 million in 2013 to 12.9 million by 2035¹. The *2017 Global Monitoring Report* by the WHO and the World Bank reported that half of the world's population lacks access to basic essential health services². The growing shortage of health workers is likely to further limit access to proper health care, reduce doctor time, and worsen patient journeys to a correct diagnosis and proper treatment.

At the same time, patient-centered apps for AI-based symptom assessment have become widely available. However, the lack of consistent standardization makes it difficult for organisations like the WHO, governments, and other key players to adopt symptom assessment systems as part of their solutions to address global health challenges.

The implementation of a standardised benchmarking for this class of applications as part of the WHO/ITU's AI for Health Focus Group will therefore be an important step towards addressing this issue. The safe and transparent application of AI technology will help improve access to healthcare for many people. It will enable earlier diagnosis of conditions, more efficient care-navigation through the health system, and ultimately better health as it is currently pursued by WHO's sustainable development goal number 3 (SDG 3).

To create a standardized benchmarking, the topic group will consider all relevant aspects. In addition to the aforementioned general objectives, this will include more topic specific questions like:

- Which output features should be benchmarked (triage, next-steps advice, differential diagnosis, explanations, question-flow, etc.)?
- Which scores and metrics should be used to assess the different dimensions of "AI-based symptom assessment" performance within the different fields of application?
- Which terminologies and ontologies (SNOMED, ICD, etc.) should be used for representing labeled test cases as well as for the communication with the AI-systems?
- How to set up benchmarking in a way that protects the intellectual property of the AI contributors, the rights of the data providers, and is still transparent?
- What should the first minimal viable benchmarking look like and how can this be refined over time to eventually cover all relevant aspects of symptom-based decision support?

All aspects will be discussed during regular focus group meetings and incorporated in the corresponding *topic description document* for this topic group.

More details about the activities of the topic group can be found in the documents:

- [A-020](#): Towards a potential AI4H use case "diagnostic self-assessment apps"
- [B-021](#): Proposal: Standardized benchmarking of diagnostic self-assessment apps

¹WHO; "Global health workforce shortage to reach 12.9 million in coming decades"; WHO 2013; <http://www.who.int/mediacentre/news/releases/2013/health-workforce-shortage/en/>

²Tracking Universal Health Coverage: 2017 Global Monitoring Report; WHO, The World Bank <http://pubdocs.worldbank.org/en/193371513169798347/2017-global-monitoring-report.pdf>

- [C-019](#): Status report on the “Evaluating the accuracy of ‘symptom checker’ applications” use case
- [C-025](#): Clinical evaluation of AI triage and risk awareness in primary care setting

These can be accessed with a free ITU account (cf. “Get involved”).

Current members of the topic group on "AI-based symptom assessment" include:

1Doc3

1DOC3 is a digital health startup based in Colombia and Mexico, was founded in 2014 and provide the first layer of access to affordable healthcare for Spanish speaking people on their phone. 1DOC3 has developed a Medical Knowledge graph in Spanish and a proprietary AI assisted technology to improve user experience by effectively symptom checking, triaging and pre diagnosing, optimizing doctors’ time allowing 1DOC3 to serve 350K consultations a month.

Ada Health GmbH

Ada Health GmbH is a digital health company based in Berlin, Germany, developing decision support systems since 2011. In 2016, Ada launched the Ada app, a diagnostic self-assessment application for smartphone users, which has been used by more than 5 million users for over 9 million health assessments. The app is currently available worldwide in 5 languages. Ada joined the focus group at the very first meeting and submitted the initial proposals that lead to the creation of this topic group. Ada is currently the topic driver for this topic group.

Babylon Health

Babylon Health is a London-based digital health company which was founded in 2013. Leveraging the increasing penetration of mobile phones, Babylon has developed a comprehensive, high-quality, digital-first health service. Users are able to access Babylon health services via three main routes: i) Artificial Intelligence (AI) services, via our chatbot, ii) “Virtual” telemedicine services and iii) physical consultations with Babylon’s doctors (only available in the UK as part of our partnership with the NHS). Babylon currently operates in the U.K., Rwanda and Canada, serving approximately 4 million registered users. Babylon’s AI services will be expanding to Asia and opportunities in various LMICs are currently being explored to bring accessible healthcare to where it is needed the most.

Baidu

Baidu is an international company with leading AI technology and platforms. After years of commercial exploration, Baidu has formed a comprehensive AI ecosystem and is now at the forefront of the AI industry in terms of fundamental technological capability, speed of productization and commercialization, and “open” strategy. Baidu Intelligent Healthcare—an AI health-specialized division established in 2018—is seeking to harness Baidu's core technology assets to use evidence-based AI to empower primary health care. The division’s technology development strategy was developed in collaboration with the Chinese government and industry thought leaders. It's building capacity in China’s public health-care facilities at a grassroots level through the development of its Clinical Decision Support System (CDSS), an AI software tool for primary health-care providers built upon medical natural language understanding and knowledge graph technology. By providing explainable suggestions, CDSS guides physicians through the clinical decision-making process like diagnosis, treatment plans, and risk alert. In the future, Baidu

will continue to enhance user experience and accelerate the development of AI applications through the strategy of “strengthening the mobile foundation and leading in AI”.

Barkibu

[Barkibu](#) is a pet health care and insurance company based in Coruña, Spain and founded in 2015. Through the Barkibu app, pet parents can get assistance on how to take care of their pets, check their symptoms and get immediate triage, talk to a live vet or find the best suited clinic for their pet’s problem. We do this through a combination of an AI powered vet assistant that runs our proprietary algorithms fed with real case data, a chat & video telehealth platform and a comprehensive insurance coverage policy.

Buoy

Deepcare

Deepcare is a Vietnam based medtech company. Founded in 2018 by three co-founders. Actually, we provide a Teleconsultation system for Vietnamese market. AI-based symptom checker is our core product. It actually is available only in Vietnamese language.

EQL

EQL is a digital health-tech organisation based in London, UK, which focuses on MSK conditions and physiotherapy. EQL’s product, Phio Access, provides a conversational AI-enabled digital solution to support triage for MSK conditions. Phio Access is currently available to 9.5 million people in the UK and in active use by several major healthcare providers, including Circle, BMI, Connect Health, Healthshare. EQL is currently working on its next-generation products, with the extended application of AI and ML technology for MSK medicine and physiotherapy.

Flo Health

Flo Health is an international company with offices in London, Vilnius, Minsk, Cyprus and USA. We are focused purely on women’s health and our aim is to help women and girls prioritise their health by giving access to expert information, knowledge and support. We encourage our users to better understand how physiology affects their wellbeing. We are mainly a B2C company, available in 22 languages, we offer products including a menstrual cycle tracker, symptoms tracking and predictions, and a dialog service that provides potential differentials for symptoms they are experiencing. We also have a very large content library for users to use and educate themselves on conditions and symptoms.

Infermedica

Infermedica, Inc. is a US and Polish based health IT company which was founded in 2012. The company provides customizable white-label tools for patient triage and preliminary medical diagnosis to B2B clients, mainly health insurance companies and health systems. Infermedica is available in 15 language versions and offered products include Symptom Checker, Call Center Triage and Infermedica API. To date the company’s solutions provided over 3.5 million health assessments worldwide.

Inspired Ideas

Inspired Ideas is a technology company in Tanzania that believes in using technology to solve the biggest challenges across the African continent. Their intelligent Health Assistant, Dr. Elsa, is powered by data and artificial intelligence and supports healthcare workers in rural areas through symptom assessment, diagnostic decision support, next step recommendations, and predicting disease outbreaks. The Health Assistant augments the capacity and expertise of healthcare providers, empowering them to make more accurate decisions about their patients' health, as well as analyzes existing health data to predict infectious disease outbreaks six months in advance. Inspired Ideas envisions building a complete end-to-end intelligent health system by putting digital tools in the hands of clinicians all over the African continent to connect providers, improve health outcomes, and support decision making within the health infrastructure that already exists.

Isabel Healthcare

Isabel Healthcare is a social enterprise based in the UK. Founded in 2000 after the near fatal misdiagnosis of the co-founder's daughter, the company develops and markets machine learning based diagnosis decision support systems to clinicians, patients and medical students. The Isabel DDx Generator has been used by healthcare institutions since 2001. Its main user base is in the USA with over 160 leading institutions but also has institutional users around the world, including emerging economies such as Bangladesh, Guatemala and Somalia . The DDx Generator is also available in Spanish and Chinese. The Isabel Symptom Checker and Triage system has been available since 2012. This system is freely available to patients and currently receives traffic from 142 countries. The company makes its APIs available so EMR vendors, health information and telehealth companies can integrate Isabel into their own systems. The Isabel system has been robustly validated since 2002 with several articles in peer reviewed publications.

Kahun

Kahun is an Israeli based med-tech venture, founded in 2018, developed an AI virtual clinical intake technology. Kahun has built an evidence-based medical knowledge graph (20M+ relations) and an AI engine that utilizes the graph to generate real-time insights. It enables Kahun to perform a patient interview, and supply a patient decision support dashboard to the provider.

Mfine

MyDoctor

Nivi

PnP

Symptify

Visiba Care

Visiba Care supplies and develops a software solution that enables healthcare providers to run own-brand digital practices. The company offers a scalable and flexible platform with facilities such as

video meetings, secure messaging, drop-ins and booking appointments. Visiba Care enables larger healthcare organisations to implement digital healthcare on a large scale, and include multiple practices with unique patient offers in parallel. The solution can be integrated with existing tools and healthcare information systems. Facilities and flows can be added and customised as needed.

Visiba Care was founded in 2014 to make healthcare more accessible, efficient and equal. In a short time, Visiba Care has been established as a market-leading provider of technology and services in Sweden, enabling existing healthcare to digitalise their care flows. Through its innovative product offering and the value it creates for both healthcare providers and patients, Visiba Care has been a driving force in the digitalisation of existing healthcare. Through our platform, thousands of patients today can choose to meet their healthcare provider digitally. As of today, Visiba Care is active in 4 markets (Sweden, Finland, Norway and UK) with more than 70 customers and has helped facilitate more than 130.000 consultations. Most customers are present in Sweden today, and our largest client is the Västra Götaland region with 1.6 million patients.

We have been working specifically with AI-based symptom assessment and automated triage for 2 years now, and this becomes a natural step to expand our solution and improve patient onboarding within the digi-physical careflow.

Xund.ai

Your.MD

Your.MD is a Norwegian company based in London. We have four years' experience in the field, a team of 40 people and currently deliver next steps health advice based on symptoms and personal factors to 570,000 people a month. Your.MD is currently working with Leeds University's eHealth Department and NHS England to scope a benchmarking approach that can be adopted by organisations like the National Institute of Clinical Excellence to assess AI self-assessment tools. We are keen to link all these initiatives together to create a globally recognised benchmarking standard.

Tom Neumark (Independent Contributor)

I am a postdoctoral research fellow, trained in social anthropology, employed by the University of Oslo. My qualitative and ethnographic research concerns the role of digital technologies and data in improving healthcare outcomes in East Africa. This research is part of a European Research Council funded project, based at the University of Oslo, titled 'Universal Health Coverage and the Public Good in Africa'. It has ethical approval from the NSD (Norway) and NIMR (Tanzania); in accordance with this, the following applies: Personal information (names and identifiers) will be anonymized unless the participant explicitly wishes to be named. No unauthorized persons will have access to the research data. Measures will be taken to ensure confidentiality and anonymity. More information available on request.

Muhammad Murhaba (Independent Contributor)

Alejandro Osornio (Independent Contributor)

Salman Razzaki (Independent Contributor)

Pritesh Mistry (Independent Contributor)

Reza Jarral (Independent Contributor)

Yura Perov (Independent Contributor)

The topic group would benefit from further expertise from the following fields:

Subject Matter Experts

Benchmarking of “AI-based symptom assessment” is a complex task with many details to consider, ranging from the required benchmarking inputs and outputs, the proper metrics and scores, and the infrastructure design that protects the integrity of the benchmarking data and the AI systems. The topic group would therefore benefit strongly from experts with relevant experience in benchmarking health assessments.

Data Providers

To benchmark “AI-based symptom assessment”, high quality labeled case datasets are needed that have not been previously accessed by any of the participating AI providers. Including partners in the topic group that could provide non-disclosed datasets or have experience with creating such datasets is a key success factor.

Developers of “AI-based symptom assessment” applications

The creation of a benchmarking framework requires the involvement of a diverse set of companies and institutions developing “AI-based symptom assessment”. While participation in the later benchmarking is open to any “AI-based symptom assessment” provider, it is important to collaborate with interested providers in order to define a viable and accepted benchmarking framework that covers all relevant aspects and satisfies the needs of all stakeholders.

Get involved

To join this topic group, please send an e-mail to the focus group secretariat (tsbfgai4h@itu.int) and the topic driver (henry.hoffmann@ada.com). Please use a descriptive e-mail subject (e.g. "Participation topic group AI-based symptom assessment), briefly introduce yourself and your organization, concisely describe your relevant experience and expertise, and explain your interest in the topic group.

Participation in FG-AI4H is free of charge and open to all. To attend the workshops and meetings, please visit the Focus Group website (<https://www.itu.int/go/fgai4h>), where you can also find the whitepaper, get access to the documentation, and sign up to the mailing list.
