

Dr. Markus Wenzel Fraunhofer HHI, Berlin markus.wenzel@hhi.fraunhofer.de

### About

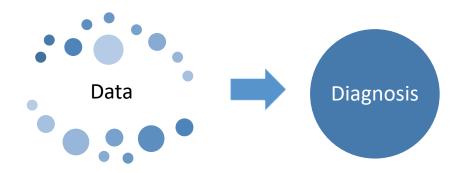


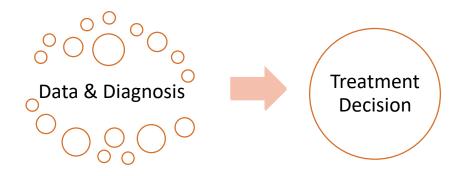
- Al offers substantial potential for improving health
  - Early detection · diagnostics · treatment decision support · self-management · improved outcomes · risk identification
- For worldwide adoption of AI in health, a standardized system of evaluating AI is required
- Focus Group: open platform that will fill this requirement
- Goal: standardized framework for benchmarking/evaluation of AI
- Expected duration 2018-2020

### Diagnostic & decision support with Al



Evaluate mapping: y = f(x)





#### From data...

- Clinical records (SNOMED CT, HPO)
- Text (symptom description)
- Images (microscope, CT, MRI, X-ray)
- Laboratory results
- Sensor time series (ECG, EEG)

#### ...to diagnosis/treatment decision

- International Classification of Diseases (ICD)
- International Classification of Health Interventions (ICHI)



### The Focus Group's process

- 1. Invite proposals for use cases and data
  - Select relevant use cases
  - Collect undisclosed test data (per use case)
  - Description of (format of) test data x and (possible) output y
  - Agree on benchmarking metrics for comparisons
- Invite submissions of Al
- 3. Benchmark AI with undisclosed test data:
  - Al generates output variables y from undisclosed test data x
  - Compare output with "ground truth" (labels/annotations)

[Reports and specifications provide a transparent documentation]



### Current status

25-27 Sep. 2018 · Workshop + Meeting A · WHO-HQ · Geneva

14-16 Nov. 2018 · Workshop + Meeting B · Columbia University · NYC

22-25 January 2019 · Workshop + Meeting C · EPFL · Lausanne

2-5 April 2019 · Workshop + Meeting D · Shanghai

t.b.c.



### Workshop + Meeting A: deliverables

25-27 Sep. 2018 · WHO-HQ · Geneva

- Constructive meeting
- Setup of the basic FG operation and structure
- Talks from medical, AI, public health, and industry experts
- White paper published
- Drafts: thematic classification scheme, data acceptance criteria, and data handling policy
- Call for proposals (CfP) for use cases, benchmarking, and data



# Workshop + Meeting B: deliverables

#### 14-16 Nov. 2018 · Columbia University · NYC

- Talks from medical, AI, regulatory, IT security, public health, and industry experts
- 15 proposals for use cases of "AI for Health" presented (only 7 weeks after CfP!)
- 8 use cases considered for entering the stage of a feasibility study
- Updated CfP for use cases, benchmarking, and data



### Projects adopted for initial feasibility analysis

- 1. Assessing the risk of breast cancer in histopathological images
- 2. Mining medical data to understand Alzheimer's disease
- 3. Guiding anti-venom selection with snake-identification imagery analysis
- 4. Classifying autism through analysis of brain imagery
- 5. Evaluating the accuracy of "symptom checker" applications
- 6. Diagnosing skin lesions through imagery analysis
- 7. Evaluating and reducing the elderly's risk of falling
- 8. Analyzing imagery to detect diabetic retinopathy and prevent vision loss

http://itu.int/go/fgai4h → Login to collaboration site



# (Updated) Call for proposals: Use cases, benchmarking, and data

Title · Overview · Relevance · Impact · Existing work · Feasibility · Data availability · Data quality · Annotation/label quality · Data provenance · Benchmarking · Organizer



# Workshop + Meeting C and D

- C 22-25 January 2019 · EPFL · Lausanne
- **D** 2-5 April 2019 · Shanghai

Work towards step 2 of the process (invite submissions of AI technology)



# Get involved: Submit your data!

- Your data could become the reference data set for a targeted health problem
- Your data (including data collection methods and associated publications) could receive global exposure
- Evaluation could provide insight into your data and assist your future data collection



# Get involved: Submit your Al!

- Your technology will be thoroughly evaluated through our framework
- Governments, regulators, insurance companies, and other users can apply the results of your algorithm evaluation and provide recommendations
- Evaluation could initiate a world-wide adoption of your technology



# Thank you!

- In particular: École Polytechnique Fédérale de Lausanne!
- Meeting C: 23-25 January
- http://itu.int/go/fgai4h
- Meeting D: Shanghai, 2-5 April