EPFL

Workshop: To trust or not to trust: When should AI be allowed to make decisions?

Hybrid Workshop.

in presence: Lausanne, EPFL, Room BC 410

remote via Zoom (password will be given after registration)

Wednesday, 07.09.2022

Session 1 (in presence)

13:00 - 13:30	Registration
13:30 - 13:45	Welcome (Marcello Ienca, EPFL)
13:45 - 14:15	Felix Gille (UZH): Trust in medical AI: Introducing the Topic
14:15 - 14:45	Marie-Valentine Florin (EPFL): Trust in Medical AI from a Risk-Assessment Perspective
14:45 - 15:30	Alberto Termine (University of Milan) & Alessandro Facchini (USI-SUPSI): TBA
15:30 – 16:00	Break
16:00 - 17:30	Group activity 1: Conceptualizing trust in the context of AI (Moderator: Alberto Termine, Milan)

EPFL

Thursday, 08.09.2022

Session 3 (in presence)

9:00 - 9:20	Janna Hastings (St. Gallen): Meaningful constraints: the role of integrity in trustworthy AI decision-making
9:20 - 9:40	Andrea Ferrario (ETH Zurich): Justified Beliefs about Machine Learning Models: A Reliabilist Account
9:40 - 10:00	Paulina Tomaszewska (Warsaw): Towards responsible Deep Learning solutions in medicine
10:00 - 10:30	Q&A in panel
10:30 - 11:00	Coffee break
11:00 - 12:30	Group activity 2- Trust in Medical AI: What is it and why should we care? (Moderator: Felix Gille, UZH)
12:30 - 13:30	Lunch

Session 4 (in presence)

13:30 - 13:50	Emily Postan (Edinburgh): Who decides who I am? The role of machine learning in healthcare in the generation of human kinds
13:50 - 14:10	Derya Sahin (Cologne): Decision-Support or Manipulator? Making the Case for Human-Centred AI in Medicine
14:10 - 14:30	Giorgia Lorenzini (Basel): AI and doctor-patient-relationship
14:30 - 15:00	Q&A in panel
15:00 - 15:30	Coffee break

EPFL

Session 5 (remote)

15:30 - 15:50   Karina Vold (Toronto): Believing in Black Boxes: Machine Learning for Healthcare Does Not Need Explainability to be Evidence-Based

15:50 - 16:10   Bogdan Kulynych (Harvard): TBA

16:10 - 16:30   Philipp Kellmeyer (Freiburg): Metaphors we design by: On the relation of interpretability, explainability and understandability of AI systems in health care

16:30 - 17:00   Q&A in panel

17:00 – 17:30   Break

17:30 - 18:30   Group activity 3 - Requirements for trustworthy AI in medicine (Moderator: Georg Starke, EPFL)

19:00   Dinner

Friday, 09.09.22

Session 6 (remote)

9:00 - 9:20   Yves Saint James Aquino (Wollongong): Explanation versus outcome: Examining professional perspectives on trustworthy AI in clinical diagnosis

9:20 - 9:40   Elise Racine (Oxford): Trustworthy and Accountable Systems? The Ethics of Using AI for Infectious Disease Surveillance

09:40 - 10:00   Karin Jongsma (Utrecht): On the epistemological and ethical basis of trust in medical AI

10:00 - 10:20   Emilia Niemiec (Lund): When untrustworthy medical AI enters the (healthcare) system - a case from Sweden

10:20 - 10:50   Q&A in panel

10:50 - 11:20   Coffee break

11:20 - 12:30   Group activity 4- Implementing requirements for trustworthy AI in medicine and healthcare (Moderator: Marcello Ienca, EPFL)

12:30 - 13:30   Lunch

EPFL

Session 7 (hybrid)

13:30 - 13:50   Carlo Martini (Milan): Getting the model right and getting the expertise right

13:50 - 14:10   Federico Cabitza, Chiara Natali (Milan): The White-Box Paradox in medicine, its discontents and possible solutions

14:10 - 14:30   Stewart Palmer (Aarhus): Trust Boundaries in AI

14:30 - 14:50   Jamie Webb (Edinburgh): Machine learning, healthcare resource allocation, and consent

14:50 - 15:20   Q&A in panel

15:20 - 16:00   Farewell Coffee; end of conference

EPFL

Workshop: To trust or not to trust: When should AI be allowed to make decisions?

(original program draft is at the end of the document. colours indicate attendance ==in person==, ==remote==, or ==not yet specified==)

Wednesday, 07.09.2022

Session 1:

13:30 - 14:00	Welcome & registration
14:00 - 14:20	==Felix Gille==
14:20 - 14:40	==Alberto Termine & Alessandro Facchini==
14:40 - 15:00	==Georg Starke & Marcello Ienca:== Trust and Distrust in Medical AI
15:00 - 16:30	Q&A in panel
16:30 - 17:00	Break

Session 2

17:00 - 18:30	Group activity 1 (optional): Focus group on trust conceptualisations with ==Stewart Palmer==

Thursday, 08.09.2022

Session 3:

9:00 - 9:20	==Yves Saint James Aquino==: Explanation versus outcome: Examining professional perspectives on trustworthy AI in clinical diagnosis
9:20 - 9:40	==Karin Jongsma==: On the epistemological and ethical basis of trust in medical AI
09:40 - 10:00	==Emilia Niemiec==: When untrustworthy medical AI enters the (healthcare) system - a case from Sweden
10:00 - 10:30	Q&A in panel
10:30 - 11:00	Coffee break
11:00 - 12:30	Group activity 2:
12:30 - 13:30	Lunch

EPFL

Session 2

13:30 - 13:50	Paulina Tomaszewska: Towards responsible Deep Learning solutions in medicine
13:50 - 14:10	Andrea Ferrario: Justified Beliefs about Machine Learning Models: A Reliabilist Account
14:10 - 14:30	Stewart Palmer: Trust Boundaries in AI
14:30 - 15:00	Q&A in panel
15:00 - 15:30	Coffee break

Session 3

15:30 - 15:50	Janna Hastings: Meaningful constraints: the role of integrity in trustworthy AI decision-making
15:50 - 16:10	Jamie Webb: Machine learning, healthcare resource allocation, and consent
16:10 - 16:30	Emily Postan: Who decides who I am? The role of machine learning in healthcare in the generation of human kinds
16:30 - 17:00	Q&A in panel
17:00 - 17:30	Break
17:30 - 18:30	Group activity 2
19:00	Dinner

EPFL

Friday, 09.09.22

Session 4

9:00 - 9:20	Carlo Martini: TBD
9:20 - 09:40	Derya Sahin: Decision-Support or Manipulator? Making the Case for Human-Centred AI in Medicine
09:40 - 10:00	Giorgia Lorenzini (and Laura Arbelaez?): AI and doctor-patient-relationship
10:00 - 10:20	Federico Cabitza, Chiara Natali: The White-Box Paradox in medicine, its discontents and possible solutions
10:20 - 10:50	Q&A in panel
10:50 - 11:20	Coffee break
11:20 - 12:30	Group activity 3
12:30 - 13:30	Lunch

Session 5

13:30 - 13:50	Elise Racine: Trustworthy and Accountable Systems? The Ethics of Using AI for Infectious Disease Surveillance
13:50 - 14:10	Karina Vold: TBA.
14:10 - 14:30	Bogdan Kulynych - TBA
14:30 - 14:50	Benjamin Sundholm: Accountability and the black-box problem in medicine
14:50 - 15:20	Q&A in panel
15:20 - 16:00	Farewell Coffee; end of conference

Only as an idea, based on our list of potential speakers: if Tim Miller cannot make it and if, following his suggestion, we actually manage to get Andreas Holzinger, maybe we could dedicate one session to explainability and trustworthiness? We could then also ask Martin Hartmann to give a talk on the relation of trust and transparency in medicine (he just gave a great talk on this at a workshop in Freiburg) and potentially Michele Loi, who co-authored a paper on the relation of explainability and trustworthiness of AI (with Andrea Ferrario). (Georg)

EPFL

Original program draft:

Wednesday, 07.09.22 - Day 1: AI in medicine: capabilities and challenges

    10:00 - 12:00 Informal welcome of arriving participants and registration

Official opening

    12:00 - 12:15 Introduction and presentation of the workshop by the organizers

Session 1 - Chair: Alessandro Facchini - *What are current trends in AI in medicine?*

    12:15 - 12:45   Tim Miller: Current trends in AI for decision-making

    12:45 - 13:15   Group task 1: Interactive development of a AI taxonomy for medicine and health

    13:15 - 14:00   Standing Lunch

Session 2 - Chair: Felix Gille - *What are present capabilities and challenges of AI in medicine?*

    14:00 - 14:30   Ezekiel Emanuel: AI in Medicine: current capabilities and challenges

    14:30 --15:30   Lily Frank: AI and digital health

    15:30 – 15:50   Coffee break

    15:50 - 16:20   Marcello Ienca: What is an AI doctor and should it be trusted?

    16:20 - 18:00   Group task 2: Break-out groups on case study analysis (16:20-17:30) and plenary group presentations (17:30-18:00)

    19:00 - 21:00   Dinner

Thursday, 08.09.2022 - Day ~~2~~1: Requirements for trustworthy AI in medicine

Session 3 - Chair:
Alberto Termine: *What is a suitable conceptual framework for trust in AI in medicine?*

9:00 - 9:30	Shannon Vallor: What does it mean to trust an AI?
9:30 - 10:00	Felix Gille: Trust and public health
10:00 - 10:30	Coffee break
10:30 - 11:00	Group discussion on presentations
11:00 - 12:30	Group task 3: In-depth analysis of three paradigm cases: (A) AI software in diagnostic radiology; (B) AI for medical triage; (C) Neuroadaptive technology in neurology and psychiatry.
12:30 - 13:30	Lunch

Session 4 - Chair:
Felix Gille - *what makes trustworthy AI in medicine?*

13:30 - 14:00	Carissa Veliz: Privacy in AI: a requirement for trust?
14:00 - 14:30	Alberto Termine: AI in decision-making: from probability to reliability
14:30 - 15:30	Group Task 3: Discussing the three case study reports and identifying areas of normative priority.
15:30 - 16:00	Coffee break
16:00 - 16:30	Carina Prunkl: Human Autonomy and Trustworthy AI development
16:30 - 17:30	Group Task 4: Collective sketching of draft guideline documents - Format, structure and identification of priorities
19:00	Dinner

EPFL

Friday, 09.09.22 - Day 3: Closing the Regulatory Gap

Session 5 - Chair:
Marcello Ienca - *How can we build trust in AI in medicine?*

9:00 - 9:30	Mark Coeckelberg:
9:30 - 10:00	Karina Vold
10:00 - 10:30	Coffee Break
11:00 - 13:00	Group task 5: Collective writing of draft guideline document - Content, timeline and implementation
13:00 - 14:00	Lunch
14:00 - 15:30	Group task 6: Parallel sub-sessions on draft guideline development for each case study

Plenary session

15:30 - 16:30	Plenary session on common denominators in guideline development & definition of next steps
16:30	End of conference