AI and Facial Recognition Technologies in Healthcare: Do We Need Them to be Explainable?

BY **DR. RITA MATULIONYTE,** MACQUARIE UNIVERSITY
RAILS WEBINAR, 11 APRIL 2022

MACQUARIE
University

In this presentation

1. AI and its challenges in healthcare: A brief review
2. Explainability in Healthcare AI: What is that? Do we need it?
3. FRT in healthcare – to which extent AI explainability is relevant?

1. AI AND ITS CHALLENGES IN HEALTHCARE: A BRIEF REVIEW

AI in healthcare

AI can be used in:

- diagnosis;
- treatment and prediction of diseases;
- improving clinical workflow;
- high quality direct-to-consumer services eg

wearable monitoring devices;
- genome interpretation;
- biomarker discovery;
- Automated robotic surgery.

Healthcare AI CHALLENGES

Will it serve or harm humans?

Quality, accuracy

Trust

Opacity, black-box nature

Patient privacy

Fairness, bias, discrimination

etc

A call for ETHICAL AI

WHO Ethical AI principles:
- Human autonomy;
- Human safety and well-being;
- **Transparency, explainability and ineligibility**;
- Responsibility and accountability;
- Inclusiveness and equity; and
- Responsiveness and sustainability

2. EXPLAINABILTY IN HEALTHCARE AI: WHAT IS IT? DO WE NEED IT?

What is explainable AI?

Different concepts:

- AI explainability as an ethical principle – very broad concept, often overlapping with transparency, interpretability
- A right of explanation (GDPR, etc) – a right of affected person to require explanation of AI-generated decisions (narrow)
- XAI research stream – technical explainability for different stakeholders
- Our definition: **explainability principle requires an explanation for different stakeholders on how a particular AI module functions, ie how a particular output, inference or result was generated**
- Distinguishable from AI transparency principle that refers to information about how the module was developed, its goals and general parameters, validation (quality assurance) info, etc

Example

CheXNet: Heat map produced by a post-hoc explanation method for a deep learning model designed to detect pneumonia in chest x-rays

Do we need Healthcare AI to be explainable?

TRUST/CLINICIAN-PATIENT
RELATIONSHIP

IMPROVED CLINICAL
DECISION MAKING/QUALITY
ASSURANCE

ALLOCATION OF
ACCOUNTABILITY AND
LIABILITY

BUT

Will explainability lead to trust? Can it be achieved using other approaches?

Can explainability improve quality of clinical decision making?

Will explainability lead to more clear and fair allocation of liability?

3.FRT in healthcare:
to which extent explainability is relevant?

FRT in healthcare and ethics

FRT use in healthcare – in early stages, e.g.
- Patient check in and check out
- Patient monitoring and diagnosis
- Caretaking robots

FRT in healthcare should comply with ethical AI principles: safety, inclusiveness, **explainability and transparency**, responsibility and accountability, etc

Case study: FRT used to confirm the identity

- Should FRT outcomes be explained to
  - patients?
  - clinicians/medical staff?
  - healthcare organization?
  - lawyers?
- Perhaps not..

Case study: FRT that assists in diagnosis of a disease

Would explainability of this outcome help:

- Patient/their parents to gain trust and make an informed decision?
- Clinician to trust the AI, identify possible errors and override them?
- Court to determine liability for a wrongful decision?

Interim conclusions

- (Technical) explainability in healthcare context is of limited use, e.g. quality assurance during AI development process
- Transparency is what most stakeholders need
- Different levels/kinds of transparency for different stakeholders (future research needed):
  - Patients – general information about the use of AI, advantages and risks involved
  - Clinicians – general understanding of technology; info about training, clinical validation, error rates; level of transparency will depend on level of expertise
  - Courts – transparency around the AI development process, whether AI developer/clinician did everything (reasonably) possible to prevent harm (industry/sector standards would play a role)
  - !! AI in higher-risk scenarios would need higher levels of transparency
  - !! Specialists would need more transparency than GPs or patients

THANK YOU

Any questions: rita.matulionyte@mq.edu.au

Further reading:

R Matulionyte, A Haneef, 'Towards explainable AI in Law Enforcement', 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW) Proceedings (2022)

R Matulionyte, 'Reconciling AI explainability and trade secrets', (2022) 44 (1) EIPR 36

MACQUARIE
University