Vision for Health Data Ecosystem

Trustworthy, Inclusive and Responsible

AI/ML powered Data Ecosystem

Erik Rozi, ODSS Civic Fellow, NHLBI

Mentor: Asif Rizwan, PhD, NHLBI

Machine Learning in Healthcare

- Machine Learning (ML) is being developed for
  - disease diagnostics,
  - prognosis,
  - drug discovery,
  - protein structure prediction,
  - imaging,
  - patient interaction,
  - healthcare administration,
  - and many other purposes.
- ML can learn from features from data.
- ML models could help streamline clinical decisions.

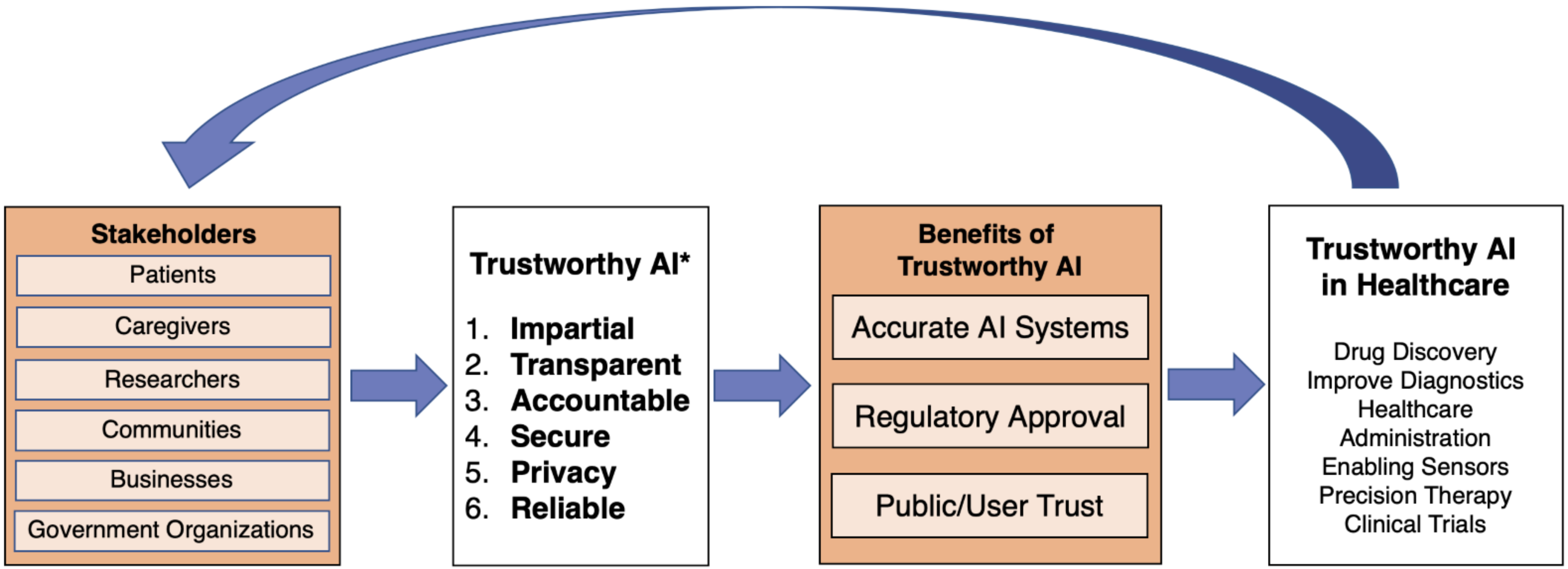
Trustworthy, Inclusive and Responsible ML

- Challenges include*
  - lack of transparency,
  - limited data,
  - lack of explainability and interpretability,
  - lack of inclusiveness,
  - security concerns,
  - and accountability concerns.

*NIH workshops, NIST and FDA workshops, 2019-2022 on *ML in Healthcare*

Receiving Input from Other Programs/Agencies

- Bridge2AI
- AIM-AHEAD
- National Cancer Institute
- National Institute of Biomedical Imaging and Bioengineering
- NIST
- FDA
- Data Science Working Group (DSWG)



Input from all stakeholders leads to more trustworthy AI systems.

**https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf*

HHS Trustworthy AI Principles Applied to R&D

6 Principles

- Robust/Reliable
- Transparent/Explainable
- Fair/Impartial
- Safe/Secure
- Privacy
- Responsible/Accountable

Robust/Reliable

- Many machine learning algorithms are beginning to show great promise.
  - SVM (Support vector machines) and boosting algorithms are used often in cardiovascular medicine.
- There are reliability and accuracy issues that need to be overcome before many algorithms can be deployed.
  - Fundamental and applied research is necessary
- Models must be evaluated and tested to ensure robustness.

Fair/Impartial

- A 2022 study investigated four AI models with >70% success rate in identifying liver diseases through blood tests
- The models missed **44%** of liver disease cases for **women**
- Meanwhile, they only missed **23%** of cases for **men**

- Impartiality must be considered throughout the entire ML development process
  - Data collection process
  - Mitigating biases due to model design
  - Testing continuously for biases that were not considered
  - Considering Social Determinants of Health as factors

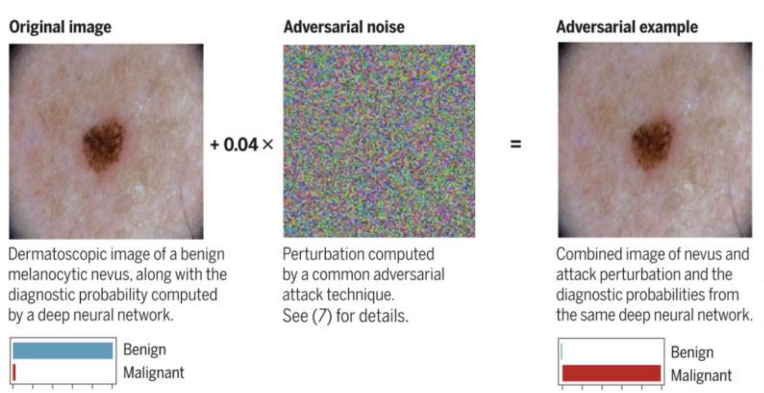
Fair/Impartial Continued

- A study published on July 11, 2022 found that pulse oximeters perform worse on Asian, Black, and Hispanic patients when compared to White patients
- This exacerbates medical biases against racial minorities
- Furthermore, machine learning algorithms have been created that use pulse oximeter data as input
- It is essential to establish that data sources and algorithms are fair

https://www.wbez.org/stories/racial-bias-in-pulse-oximeters-led-to-delayed-covid-19-treatment-for-people-of-color/fc5ad790-0d68-47d2-91a7-df07ca37efad

Safe/Secure

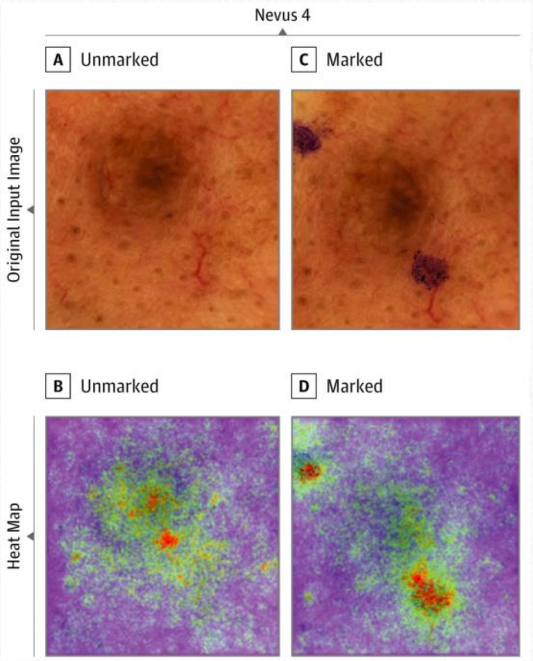
- Datasets are often created by scraping online data, or are distributed online for download
- One study found that a dataset injected with even 0.1% of harmful data will make machine learning models trained on that data useless
- Noise can also be applied on any image or data



https://www.science.org/doi/10.1126/science.aaw4399

Transparent/Explainable

- Both images are of benign nevi/moles.
- The model classified marked (surgical skin marker) images as malignant and unmarked as benign.
- Heat maps show that the ML model isn't paying attention to the right features.



https://jamanetwork.com/journals/jamadermatology/fullartide/2740808

Transparent/Explainable

- Potential applications of explainable methods:
  - Sickle cell research, blood clot detection, etc.
  - Understanding biases in ML models
  - Identifying adversarial examples

Privacy

- When researchers are collecting data for datasets, is this being done ethically?
  - NHLBI's Biomedical Catalyst
- Are patients providing data adequately informed and giving consent?
- Is data being properly de-identified?
  - Studies have shown that there are ways that algorithms can re-identify data that has been de-identified

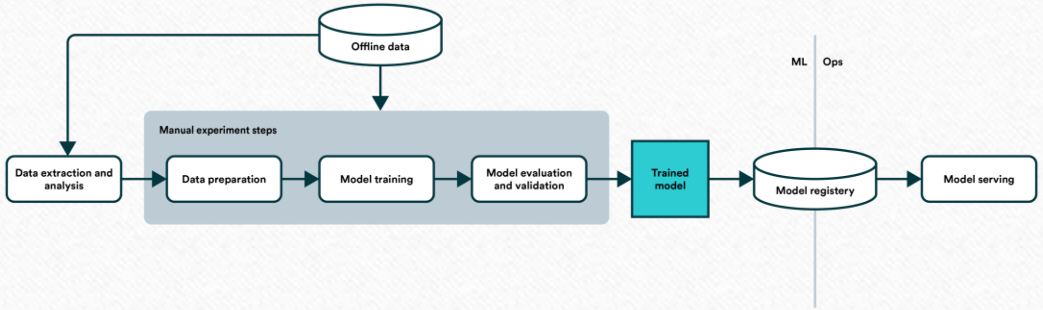
Privacy Research

- Research focusing on privacy preservation in ML algorithms has begun:
  - Privacy-preserving generative models
  - An NHLBI funded project: Distributing algorithms rather than data to other hospitals to maintain privacy

- The NIH/NHLBI can support this research

Responsible/Accountable

- If a machine learning algorithm leads to a bad outcome, who's at fault? Is anyone at fault?
  - The researcher who created the technique?
  - The business implementing the research?
  - The doctor watching over the algorithm?
  - The hospital using the algorithm?
- More importantly, how do we resolve these issues?
- Can we trace issues back to their root cause?



https://valohai.com/machine-learning-pipeline/

Programs Supporting Trustworthy AI

Generate diverse community interest in AI for healthcare to preserve equity → Develop ethical and trustworthy datasets → Research trustworthy AI methods → Train trustworthy AI models → Test robustness, reliability, usefulness → Ensure widespread adoption of trustworthy AI/ML methods

AIM AHEAD

DS-I Africa
Data Science for Health Discovery and Innovation in Africa

- AIM-AHEAD: Brings diverse researchers and communities with to advance AI methods while addressing health disparities
- DSI-Africa: Develop expertise among African scientists and establish network of investigators

BioData CATALYST

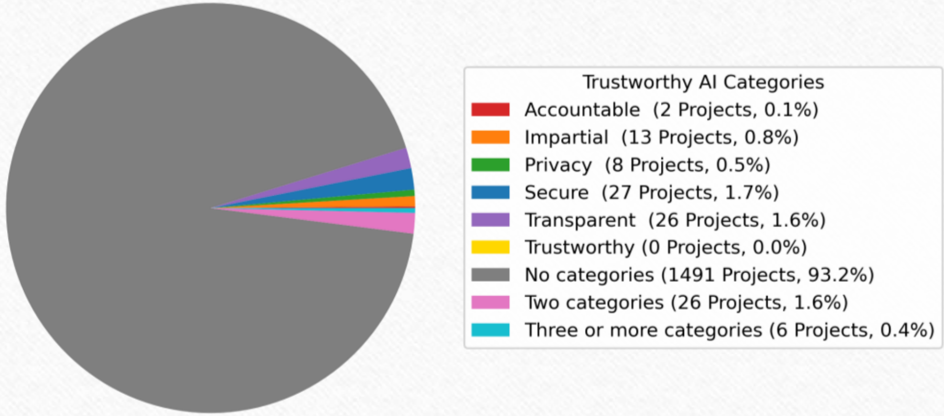
- Bridge2AI: Creates flagship datasets that are ethically sourced, trustworthy, well-defined, and accessible.
- Midrc: Data sharing in response to COVID-19
- Biodata Catalyst: Provides FAIR datasets for NHLBI

No current overarching Trustworthy AI Program

Trustworthy AI

- Impartial (AIM-AHEAD, DSI-Africa, Bridge2AI)
- Transparent (No Program)
- Accountable (No Program)
- Secure (No Program)
- Privacy (No Program)
- Reliable (No Program)

Number of Active Machine Learning Projects Containing Trustworthy AI/ML Keyword



Category	Keywords
Impartial	Impartial, Ethics, Ethic, Fair
Transparent	Explainable, Explanations, Interpretable, Transparent, XAI
Accountable	Accountable, Responsible
Privacy	Privacy, Private
Secure	Secure, Safe

Comparison of Trustworthy AI Grant Appearances in NIH

Category	ML/AI Projects in 2018-2019	ML/AI Projects in 2020-2021
Accountable	1 Project (0.1%)	6 Projects (0.1%)
Impartial	4 Projects (0.3%)	35 Projects (0.8%)
Privacy	5 Projects (0.4%)	23 Projects (0.5%)
Secure	27 Projects (2.2%)	86 Projects (2.0%)
Transparent	**11 Projects (0.9%)**	**68 Projects (1.6%)**
Trustworthy	0 Projects (0.0%)	0 Projects (0.0%)
No Categories	**1155 Projects (94.0%)**	**3906 Projects (93.0%)**
Two Categories	24 Projects (2.0%)	63 Projects (1.5%)
Three or More Categories	2 Projects (0.2%)	14 Projects (0.3%)
Total Number of Projects	**1229 Projects**	**4201 Projects**

Trustworthy AI/ML Projects between 2018-2019 and 2020-2021

- The total amount of ML projects increased from 1229 projects to 4201 projects

- The number of projects containing a Trustworthy AI/ML keyword increased

Trustworthy AI/ML Projects between 2018-2019 and 2020-2021

- The share of projects containing Trustworthy AI keywords increased only slightly from 6.0% to 7.0%
- 0.7% of this increase can be attributed to Transparent/Explainable AI projects
- We can support further research in other categories of Trustworthy AI

Next Steps

- Host **workshop** with multiple community stakeholders and researchers to further Trustworthy ML/AI adoption.
- Engage with researchers directly to emphasize fundamental Trustworthy ML/AI research.
- Support Trustworthy AI Funding Initiatives

Thank You