Accelerating Medical AI with MONAI FL

Nexus of Open Science 2025

Sai Praneeth Karimireddy

About me
Praneeth

- Assistant Professor @ **CS USC**
- Co-Lead of **Federated Learning and Data Quality** working group at MONAI

- Research Interests
  - Private & Reliable AI
  - Algorithms for AI data (Data-centric AI)
  - AI for health



Medical AI revolution is yet to come

Despite breathless headlines

AI · HEALTH

AI is already touching nearly every corner of the medical field

BY SAGE LAZZARO
CONTRIBUTING WRITER

July 23, 2025 at 10:00 AM EDT

DIGITAL

2 in 3 physicians are using health AI—up 78% from 2023

AMA survey of nearly 1,200 doctors offers four key insights into how physicians use and feel about augmented intelligence (AI) in medicine.

By **Tanya Albert Henry**, Contributing News Writer

Feb 26, 2025 | 5 Min Read

vs.

"Cancer patients face life-threatening delays due to lack of staff, say UK radiologists."

— The Guardian, 8 Jun 2023

Cancer patients face worsening treatment delays due to lack of staff, report finds

— Sky News, 8 June 2023

Faculty Research | September 16, 2025

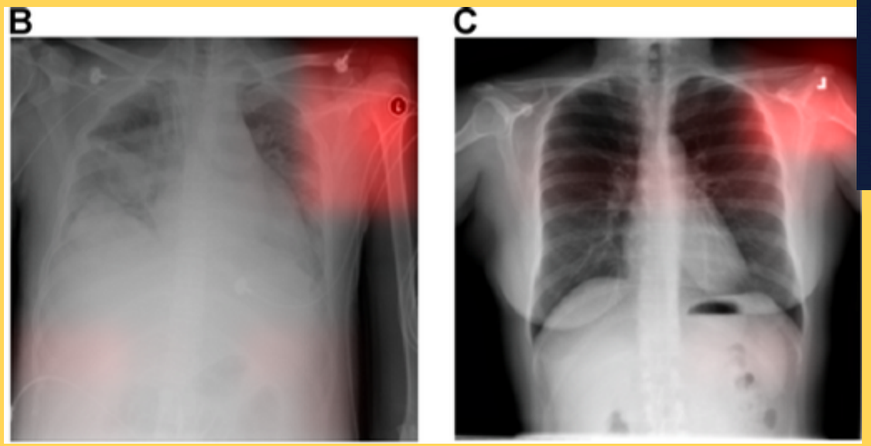
HOME

Doctors who use AI are viewed negatively by their peers, new study shows

- Currently mostly focusing on low-value clerical work.
- Despite real need - acute shortage of radiology+pathology expertise

Medical AI revolution is yet to come

Justified skepticism?



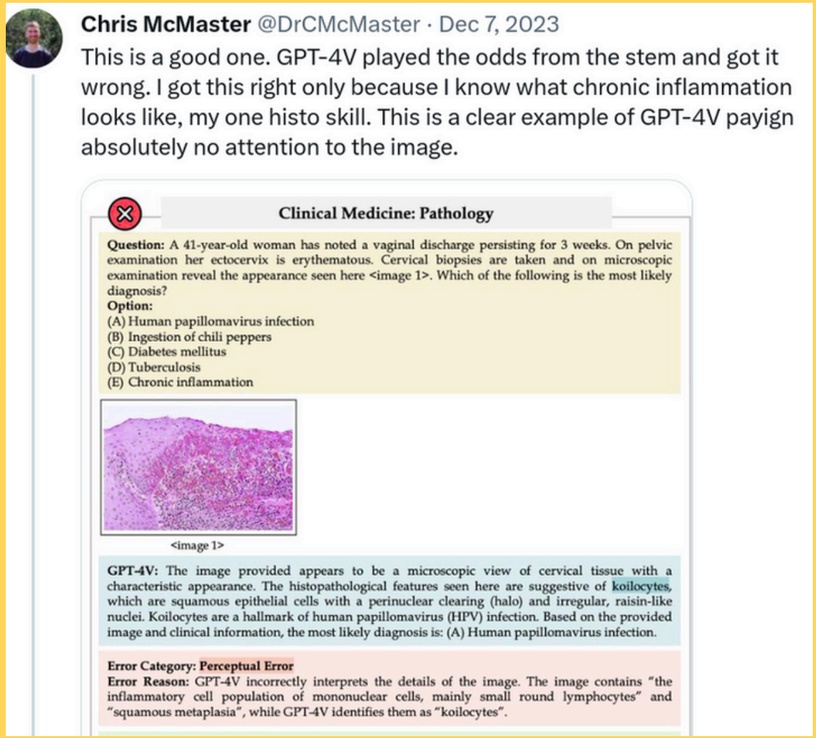
Ringel et al. "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study." PLoS Medicine 2018.

Hundreds of AI tools have been built to catch covid. None of them helped.

Some have been used in hospitals, despite not being properly tested. But the pandemic could help make medical AI better.

By Will Douglas Heaven

July 30, 2021



Chris McMaster @DrCMcMaster · Dec 7, 2023

This is a good one. GPT-4V played the odds from the stem and got it wrong. I got this right only because I know what chronic inflammation looks like, my one histo skill. This is a clear example of GPT-4V payign absolutely no attention to the image.

Clinical Medicine: Pathology

Question: A 41-year-old woman has noted a vaginal discharge persisting for 3 weeks. On pelvic examination her ectocervix is erythematous. Cervical biopsies are taken and on microscopic examination reveal the appearance seen here <image 1>. Which of the following is the most likely diagnosis?
Option:
(A) Human papillomavirus infection
(B) Ingestion of chili peppers
(C) Diabetes mellitus
(D) Tuberculosis
(E) Chronic inflammation

<image 1>

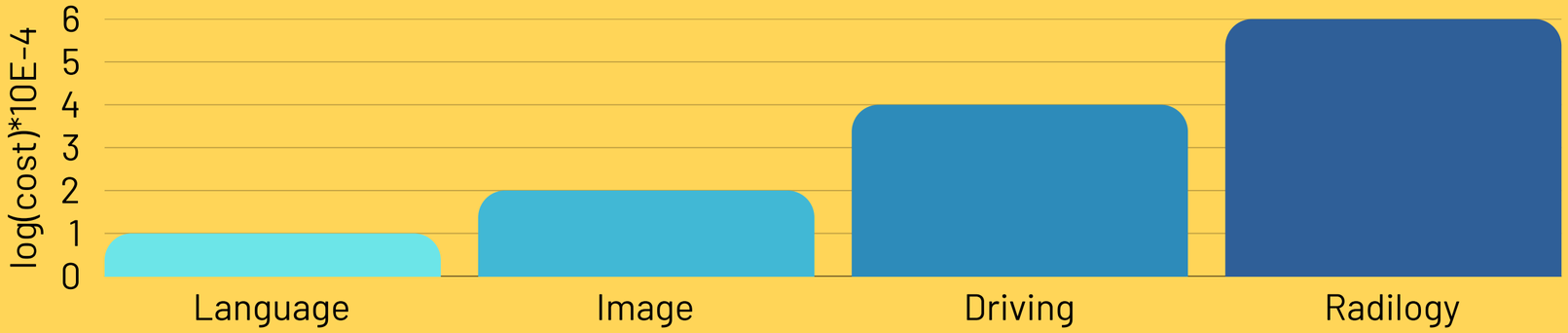
GPT-4V: The image provided appears to be a microscopic view of cervical tissue with a characteristic appearance. The histopathological features seen here are suggestive of koilocytes, which are squamous epithelial cells with a perinuclear clearing (halo) and irregular, raisin-like nuclei. Koilocytes are a hallmark of human papillomavirus (HPV) infection. Based on the provided image and clinical information, the most likely diagnosis is: (A) Human papillomavirus infection.

Error Category: Perceptual Error
Error Reason: GPT-4V incorrectly interprets the details of the image. The image contains "the inflammatory cell population of mononuclear cells, mainly small round lymphocytes" and "squamous metaplasia", while GPT-4V identifies them as "koilocytes".

- SOTA medical AI still seem to rely on "***shortcuts***"
- They have not seen enough ***diverse data***.

Medical AI's data bottleneck

Too costly to collect at scale



Internet scraped text data
$0.0001 per doc
$1k for 1 TB
10 hrs

Annotated internet images
$0.03 per image
$300k for 1M images
1 week
[COCO 2016]

Self-driving cars
$10 per s
$40M+ for 50k miles: $15M in equipment + $25M in driving
3 years
[Ottometric 2024]

Medical imaging
$500 per MRI scan
$1B for 2M scans: $500/hr for MRI scanner + $150/hr for radiologist annotation + $50/hr for volunteers.
20 scans/day.
[Missouri Radiology]

Data Sharing is great!
Decrease costs, improves fairness

Global healthcare fairness: We should be sharing more, not less, data

Kenneth P. Seastedt, Patrick Schwab, Zach O'Brien, Edith Wakida, Karen Herrera, Portia Grace F. Marcelo, Louis Agha-Mir-Salim, Xavier Borrat Frigola, Emily Boardman Ndulue, Alvin Marcelo, Leo Anthony Celi

FDA NEWS RELEASE

FDA Takes Important Steps to Increase Racial and Ethnic Diversity in Clinical Trials

Agency's Focus on Inclusion in Trials for All Medical Products Aligns with Biden Administration's Cancer Moonshot Goal of Addressing Inequities and Beyond

Share | Post | Linkedin | Email | Print

For Immediate Release: April 13, 2022



"Build infrastructure to support research and data-sharing" - 2023

"big area of need that [we are] focusing on is data systems" - 2024

- Data sharing improves diversity of datasets, sharing of collection costs.
- FAIR datasets are the solution, duh!

Why you should NOT share your data

Some fundamental conflicts



Findable



Accessible



Interoperable



Reusable

- **Privacy**: even anonymized medical data can leak private information.
- **Control**: data could be repurposed inappropriately.
- **Costs**: maintaining interoperability is a LOT of costly clerical (boring) work.
- **IP and incentives**: get scooped on research, lose licensing revenues.

Accelerating Medical AI with MONAI FL

Talk outline

01

MONAI

Open source Medical AI

02

Federated Learning

Privacy & Control

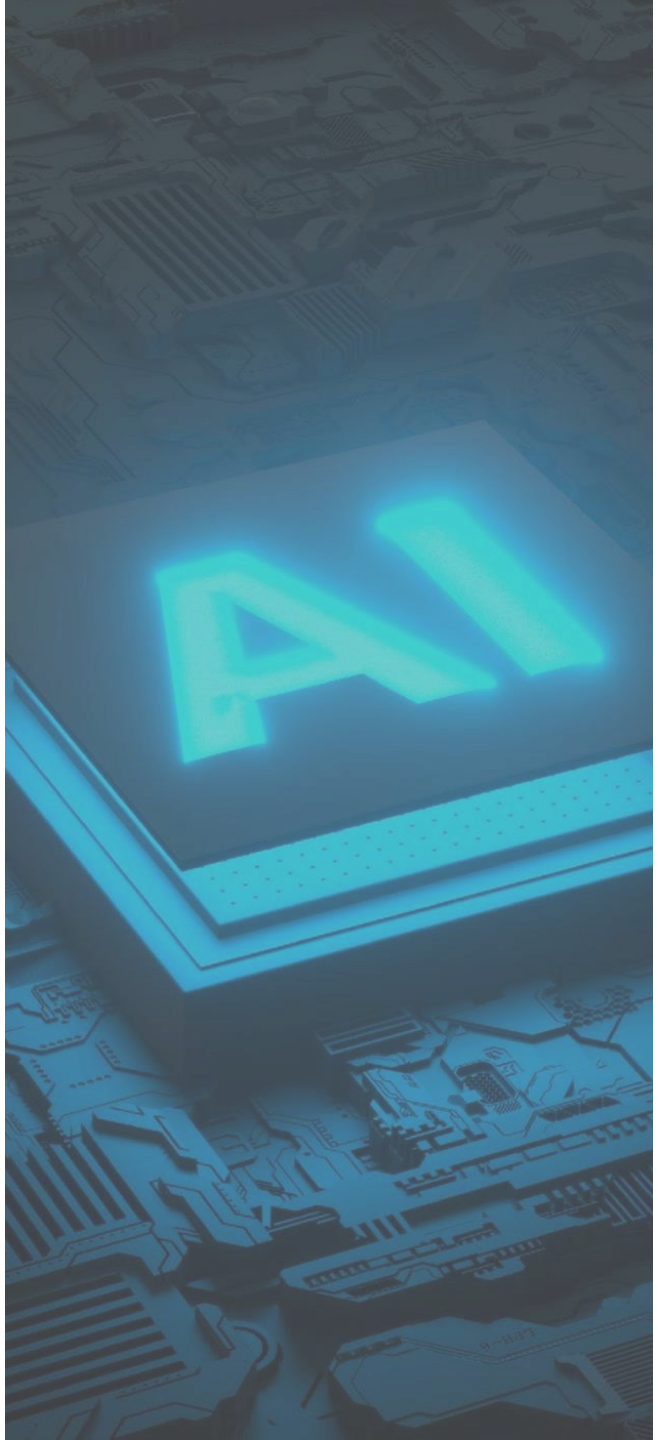
03

Automating Interoperability

Reduce costs



Medical Open Network for AI



What is MONAI?

Project MONAI is a collaborative open-source initiative built by academic and industry leaders for deep learning in healthcare imaging.

MONAI Working Groups.



Imaging I/O

Focus: define how data is read into and written out from memory in MONAI.

Data

Focus: Defining support for bioinformatics, biomarkers, and metadata that are in scope for MONAI.

Transformations

Focus: Topics related to data preprocessing and augmentation modules in MONAI.

Federated Learning

Focus: Unify the disparate methods of Federated Learning in a common MONAI framework.

Evaluation, Reproducibility, and Benchmarking

Focus: Provide the infrastructure and tools for quality-controlled validation and benchmarking of medical image analytics methods.

Research

Focus: Establish MONAI as a catalyst for scientific progress and real-life impact.

Community Development

Focus: Establish MONAI as a common software foundation that the medical imaging research and development community can build upon.

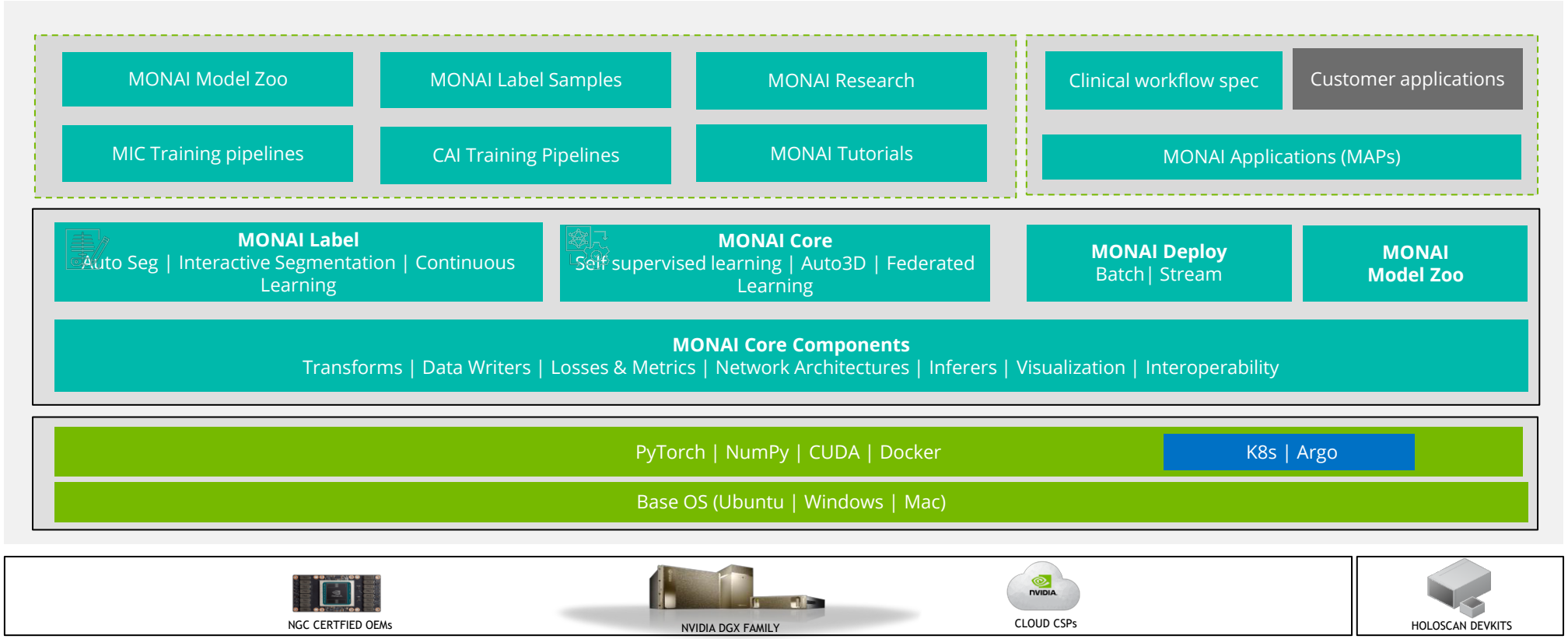
Deploy

Focus: Close the existing gap from research and development to clinical production environments by bringing AI models into the medical workflow.

Digital Pathology

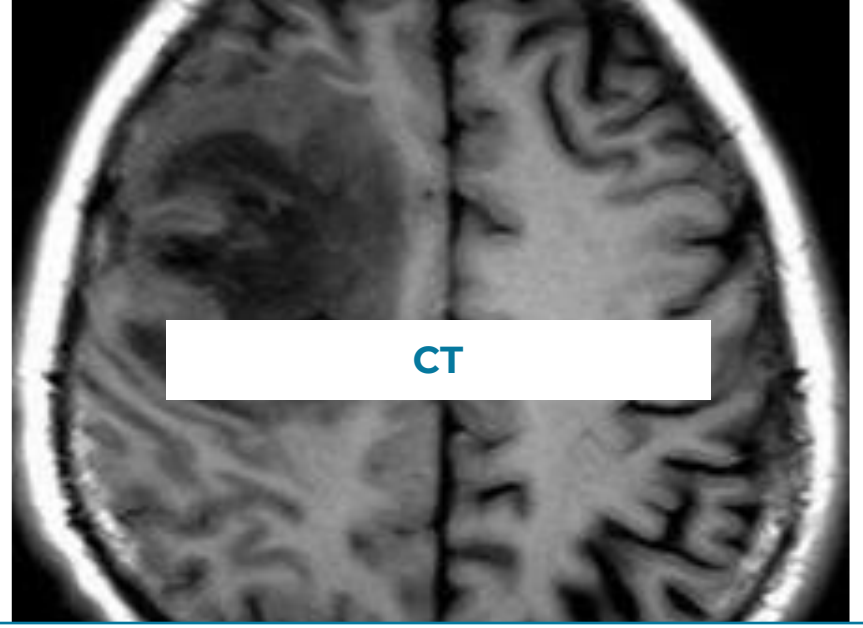
Focus: Creating a standard pipeline for preprocessing, analysis, and visualization of pathology images.

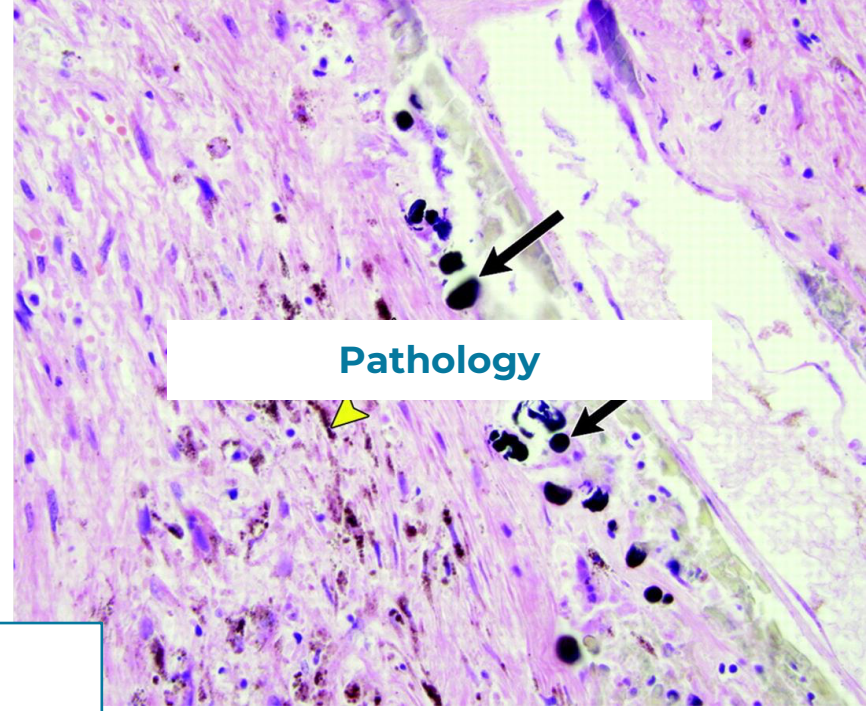
MONAI Stack.


MONAI Apps
MONAI Model Zoo
MONAI Label Samples
MONAI Research
MIC Training pipelines
CAI Training Pipelines
MONAI Tutorials
Clinical workflow spec
Customer applications
MONAI Applications (MAPs)
MONAI Tools
MONAI Label
Auto Seg | Interactive Segmentation | Continuous Learning
MONAI Core
Self supervised learning | Auto3D | Federated Learning
MONAI Deploy
Batch | Stream
MONAI Model Zoo
MONAI Core Components
Transforms | Data Writers | Losses & Metrics | Network Architectures | Inferers | Visualization | Interoperability
System Software
PyTorch | NumPy | CUDA | Docker
K8s | Argo
Base OS (Ubuntu | Windows | Mac)
Systems
NGC CERTFIED OEMs
NVIDIA DGX FAMILY
CLOUD CSPs
HOLOSCAN DEVKITS
SOTA Research to Clinical Deployments

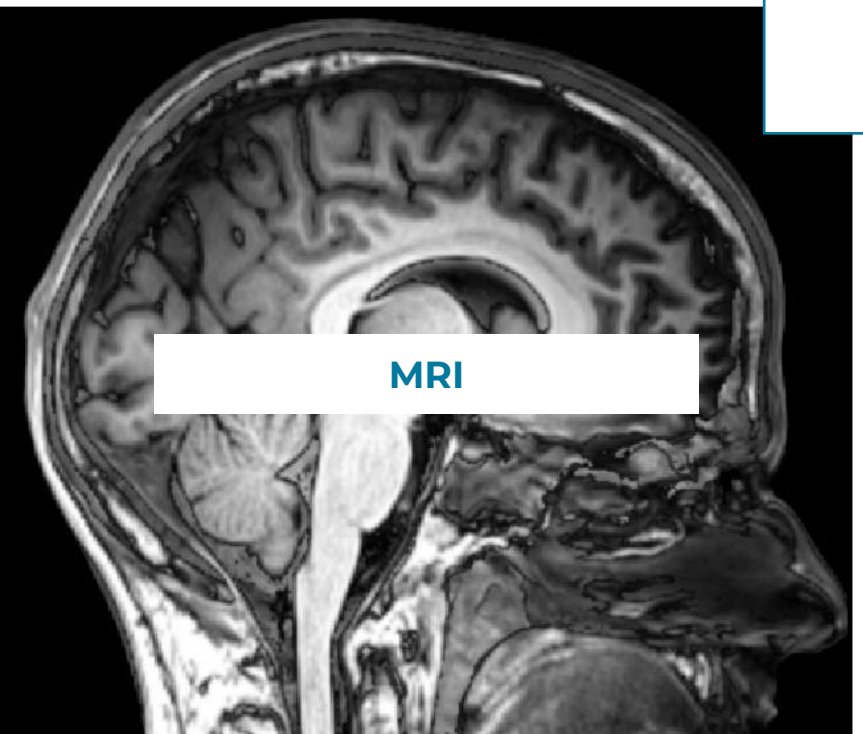
MONAI Model Zoo.

Use pre-trained models to jumpstart AI development for all organs and disease types

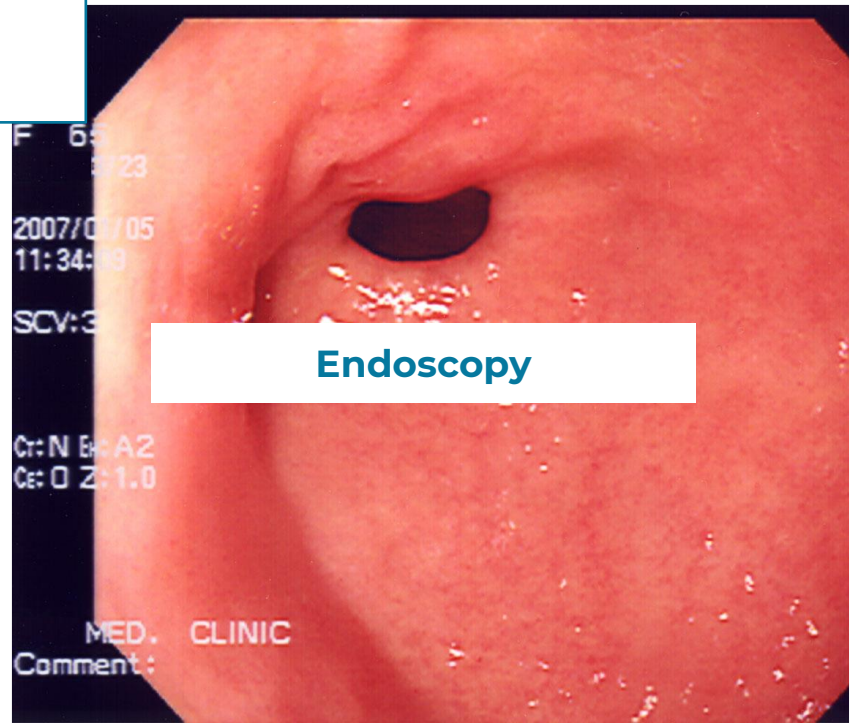




Modalities







Model Zoo Contributors.

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN

MAYO CLINIC

VANDERBILT UNIVERSITY

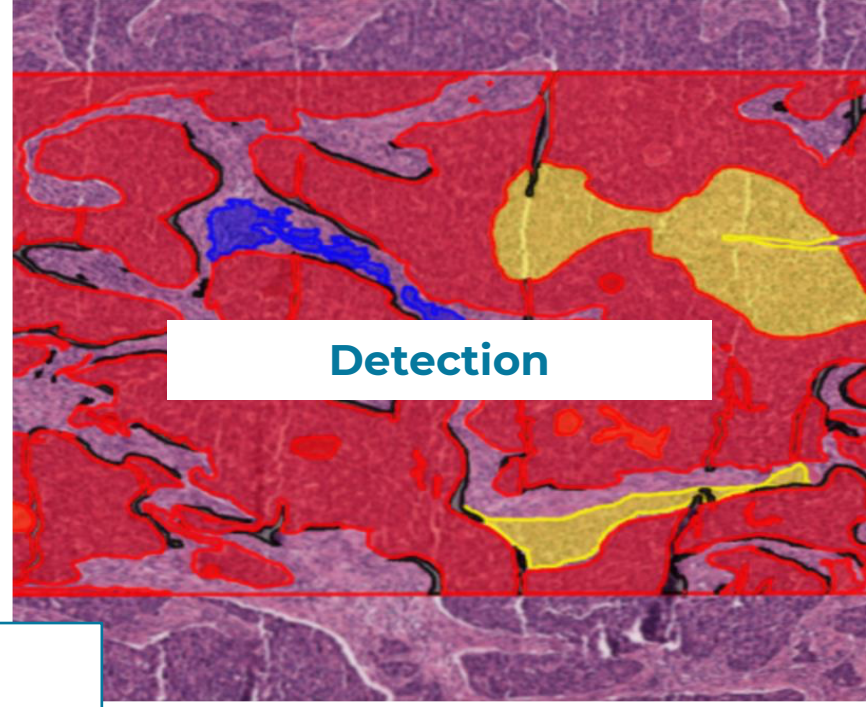
KING'S College LONDON

NVIDIA

WARWICK THE UNIVERSITY OF WARWICK

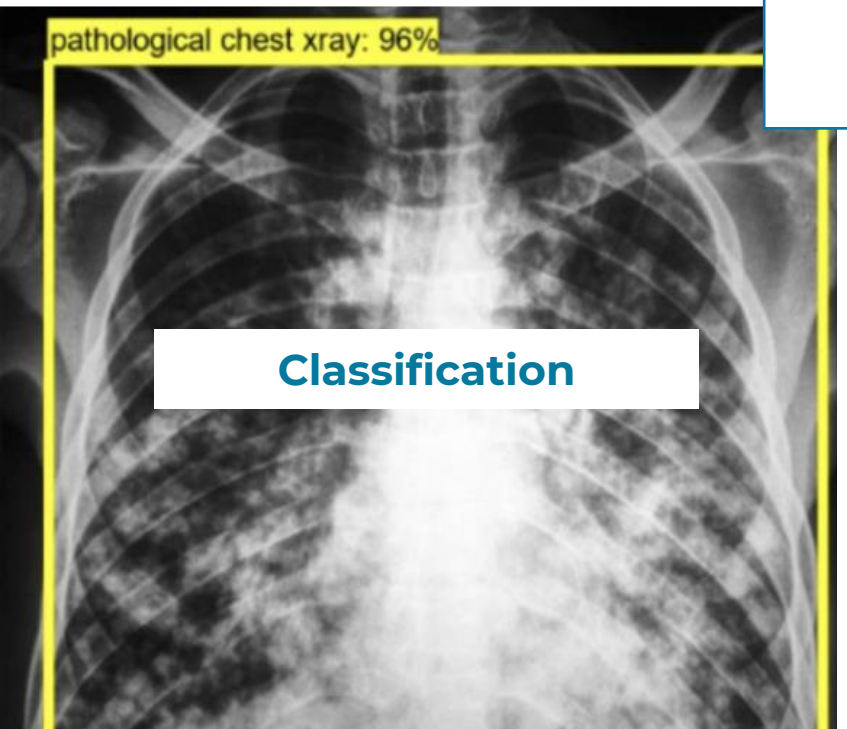


Segmentation

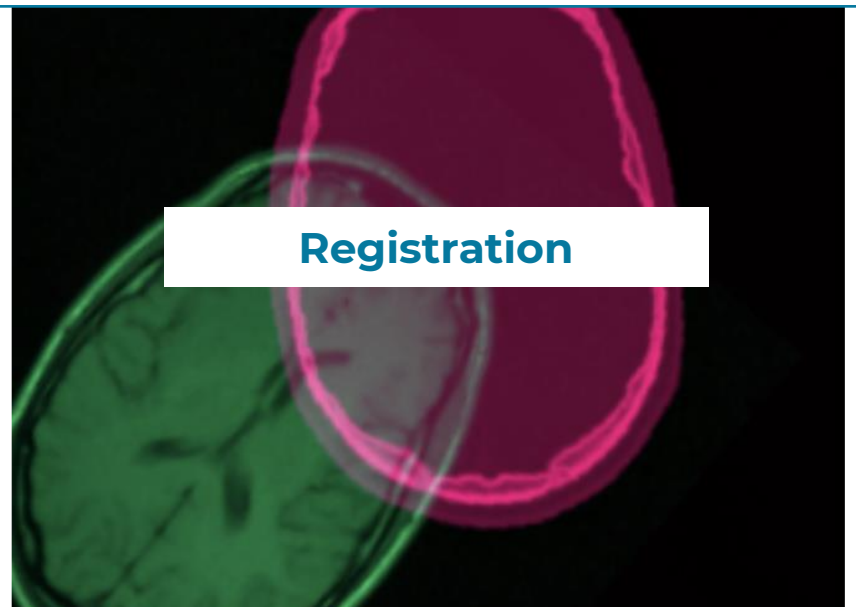


Detection

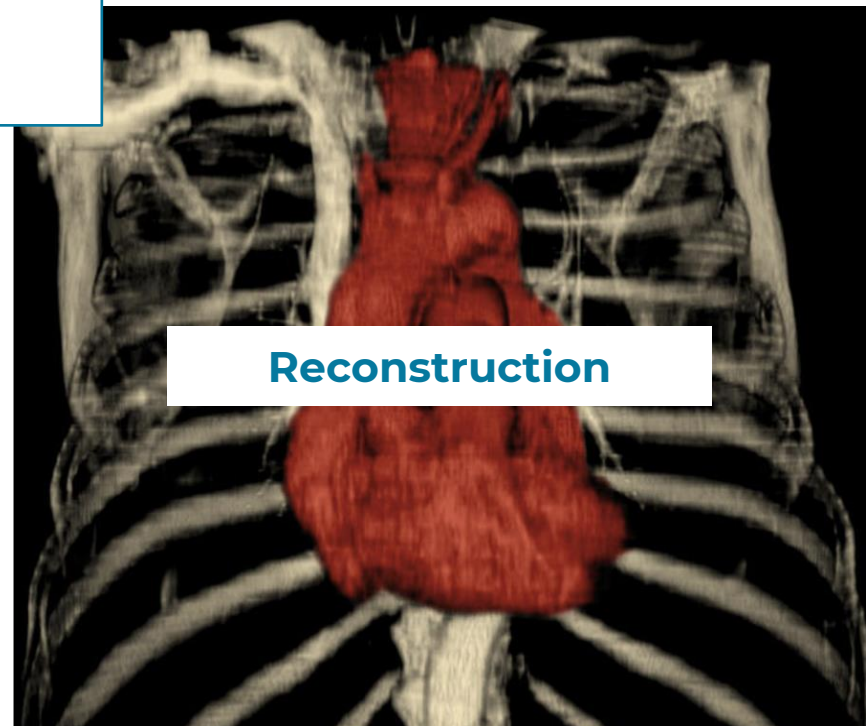
Applications



Classification



Registration



Reconstruction

Available Models

39 models available (34 Bundle, 5 HF)

All Models | Bundle Models | HF Models

Spleen CT segmentation

Bundle Model

MONAI team

A pre-trained model for volumetric (3D) segmentation of the spleen from CT image

v0.6.0

View Details > Download

Pancreas CT DiNTS segmentation

Bundle Model

MONAI team

Searched architectures for volumetric (3D) segmentation of the pancreas from CT image

v0.5.1

View Details > Download

BraTS MRI segmentation

Bundle Model

MONAI team

A pre-trained model for volumetric (3D) segmentation of brain tumor subregions from multimodal MRIs based on BraTS 2018 data

v0.5.3

View Details > Download

Spleen DeepEdit annotation

Bundle Model

MONAI team

This is a pre-trained model for 3D segmentation of the spleen organ from CT images using DeepEdit.

v0.5.7

View Details > Download

Swin UNETR BTCV segmentation

Bundle Model

MONAI team

A pre-trained model for volumetric (3D) multi-organ segmentation from CT image

v0.5.7

View Details > Download

Ventricular short axis 3 label segmentation

Bundle Model

Eric Kerfoot

This network segments full cycle short axis images of the ventricles, labelling LV pool separate from myocardium and RV pool

v0.3.4

View Details > Download

MedNIST GAN

Bundle Model

MONAI Team

This example of a GAN generator produces hand xray images like those in the MedNIST dataset

v0.4.3

View Details > Download

Pathology tumor detection

Bundle Model

MONAI team

A pre-trained model for metastasis detection on Camelyon 16 dataset.

v0.6.3

View Details > Download

Renal structures UNEST segmentation

Bundle Model

Vanderbilt University + MONAI team

A transformer-based model for renal segmentation from CT image

v0.2.6

View Details > Download

MONAI Downloads and Contributors.

780k+

MONAI Core

20k+

MONAI Label

10k+

MONAI Deploy

150+

Individual Contributors

30+

Institutions

MONAI Resources.

- MONAI Website: https://monai.io/
- MONAI Slack: https://forms.gle/QTxJq3hFictp31UM9
- MONAI Docs:
  - MONAI Core: https://docs.monai.io/en/stable/
  - MONAI Label: https://docs.monai.io/projects/label/en/latest/index.html
  - MONAI Deploy App SDK: https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/
- MONAI Github: https://github.com/Project-MONAI
  - MONAI Core: https://github.com/Project-MONAI/MONAI
  - MONAI Label: https://github.com/Project-MONAI/MONAILabel
  - MONAI Deploy: https://github.com/Project-MONAI/monai-deploy
- MONAI YouTube: https://www.youtube.com/c/Project-MONAI
  - Overview Videos, Deep Dive Series, Bootcamp and other event recordings
- MONAI Twitter: https://twitter.com/ProjectMONAI
  - Follow for the latest announcements
- MONAI Medium: https://monai.medium.com/
  - Read about our latest releases and our upcoming research interview series

Accelerating Medical AI with MONAI FL

Talk outline

01

MONAI

Open source Medical AI

02

Federated Learning

Privacy & Control

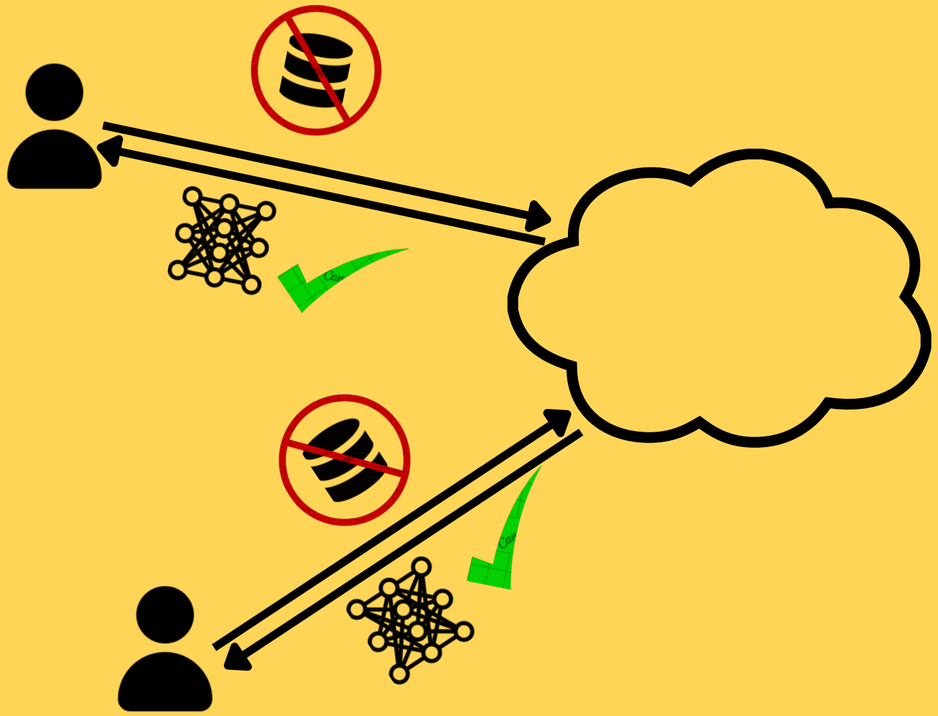
03

Automating Interoperability

Reduce costs

Federated Learning

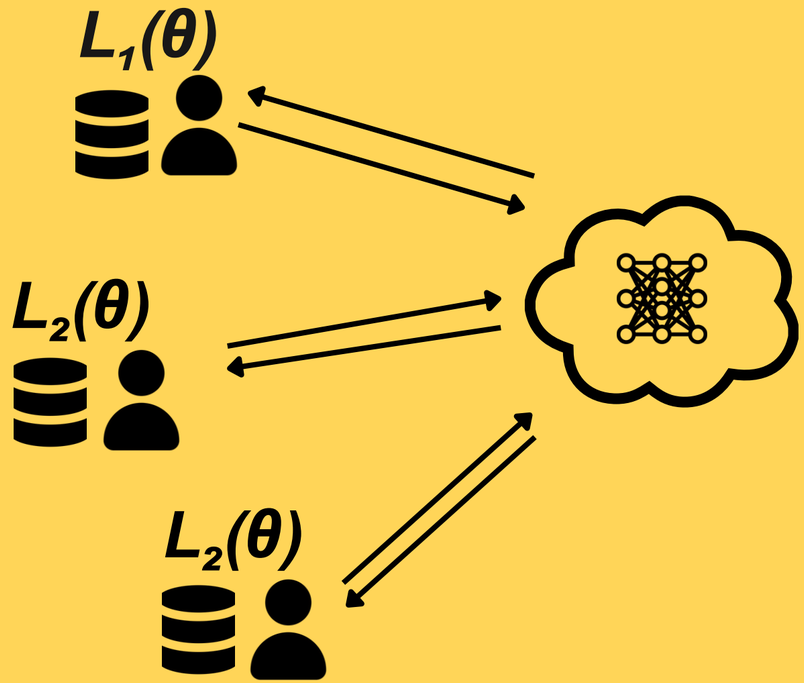
What is it?



- Multiple centers (hospitals) collaborate
- Train AI models on their ***combined data***
- Without ***sharing raw data***
  - only aggregates (models, ...)

Federated Learning
Overview

1. Each center sets up local training using their local data

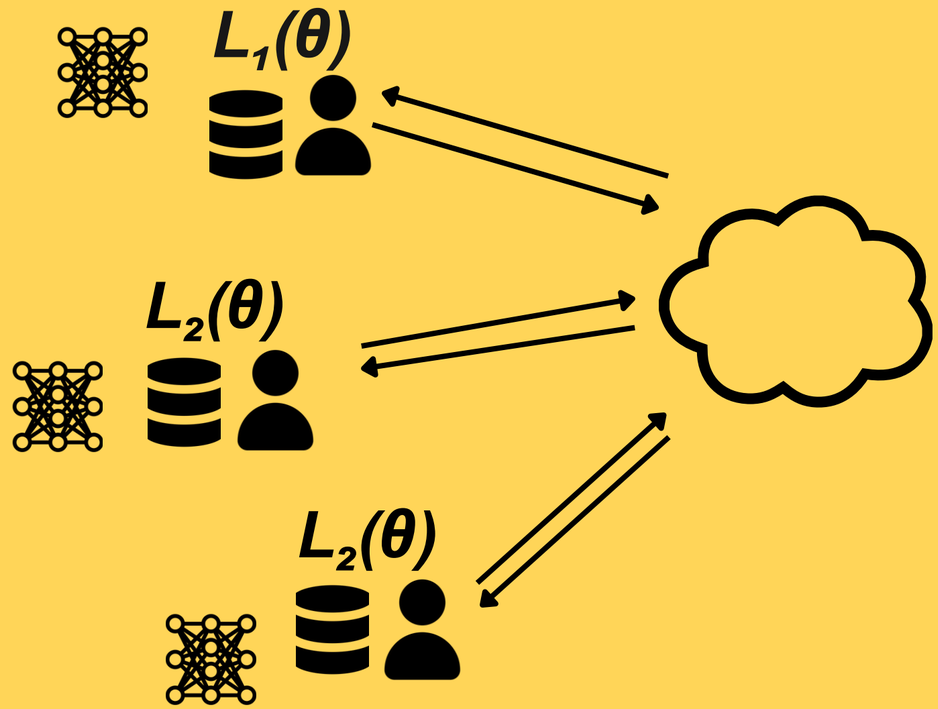


Federated Learning
[McMahan et al. 2016]

Federated Learning

Overview

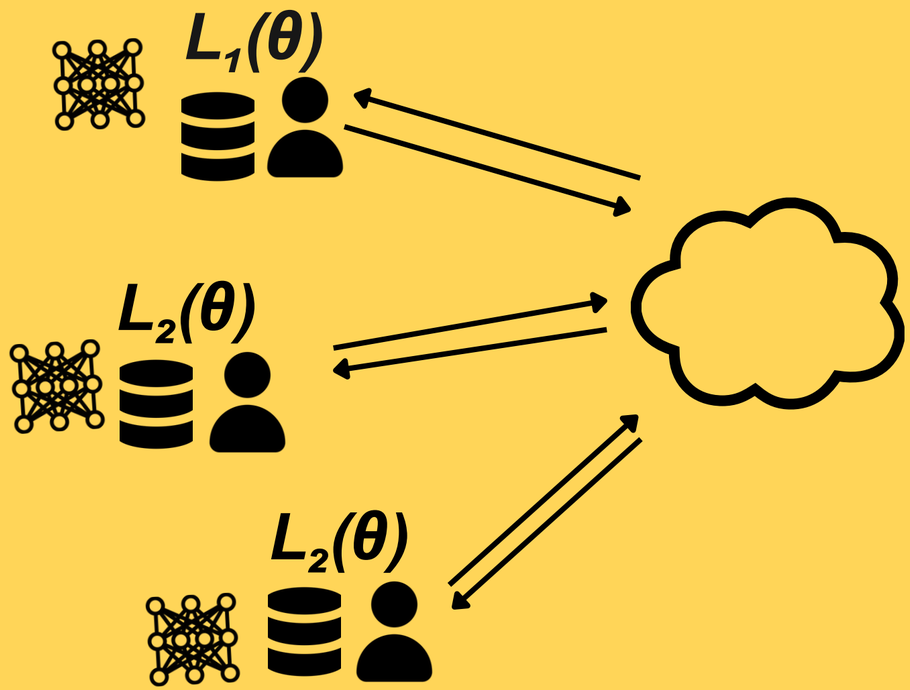
1. Each center sets up local training using their local data
2. Iteratively on every user,
   - Copy global model θ



Federated Learning
[McMahan et al. 2016]

Federated Learning
Overview

1. Each center sets up local training using their local data
2. Iteratively on every user,
    - Copy global model θ
    - Train model on local data to obtain local model θ_i



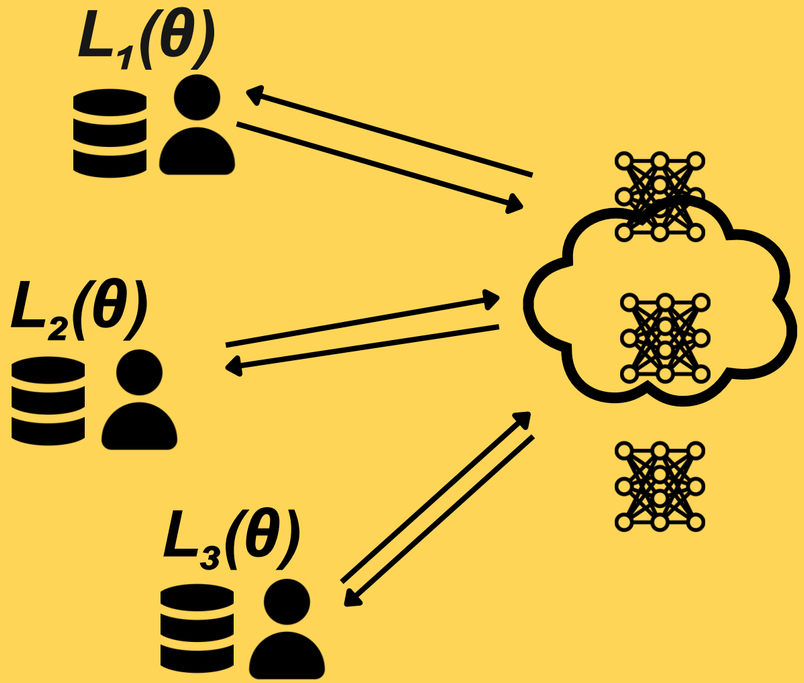
Federated Learning
[McMahan et al. 2016]

Federated Learning

Overview

1. Each center sets up local training using their local data
2. Iteratively on every user,
   - Copy global model θ
   - Train model on local data to obtain local model θ_i
   - Combine local models

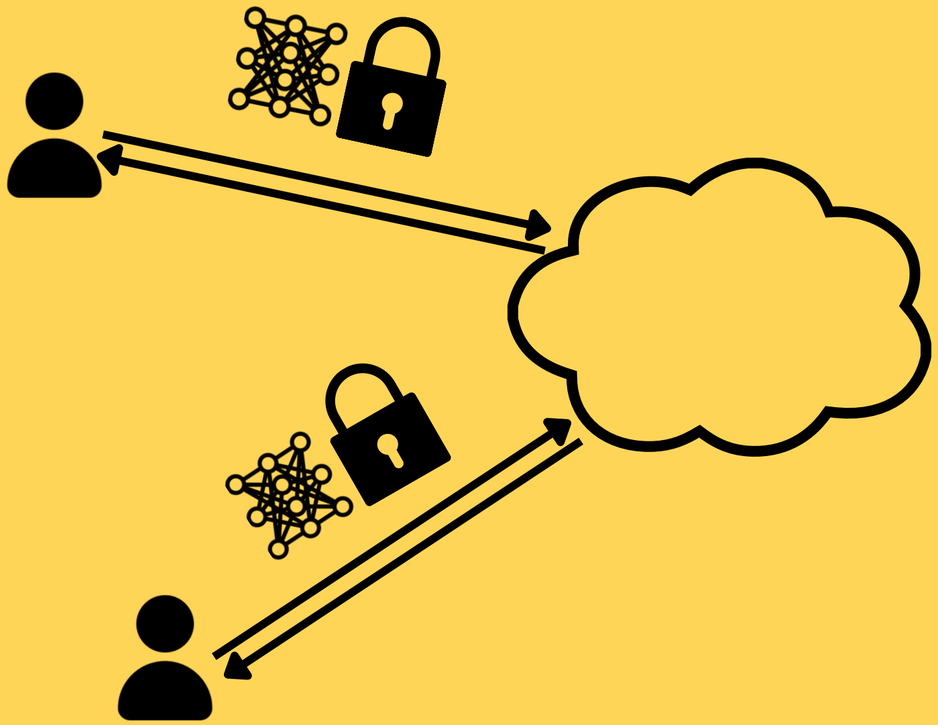
$$\theta = \frac{1}{m} \sum_i \theta_i$$



Federated Learning
[McMahan et al. 2016]

Federated Learning

Privacy and Control

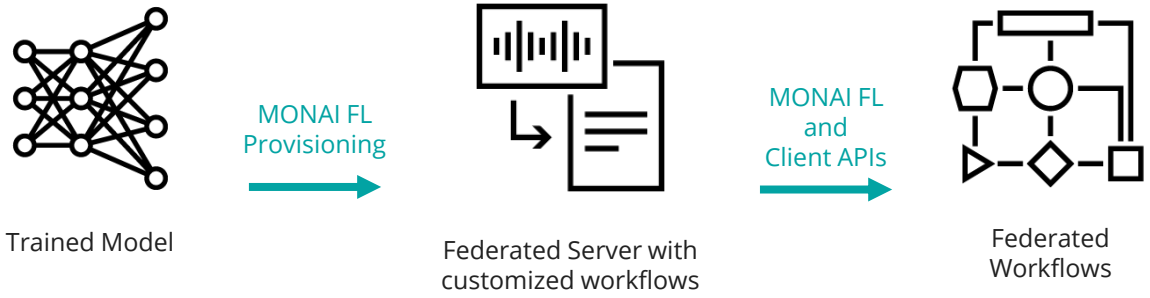


- Fine-grain control over data usage
- Only aggregates shared => more private
- For formal privacy and security,
  - **Differential Privacy**: share noisy aggregates
  - Trusted Execution Environments (**TEEs**) or homomorphic encryption

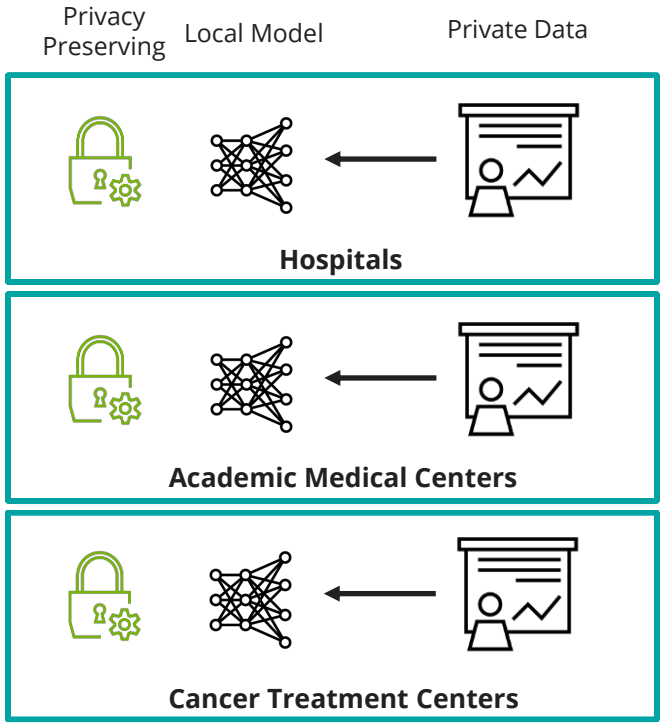
Federated Learning with MONAI.

Build collaborative AI with privacy-preserving NVIDIA FLARE integration

AI Research and Development



Federation Participants



Interoperability

MONAI Federated learning (FL) client algorithm APIs can run AI applications on any FL platform

Distributed Multi-Party Collaboration

Mitigate risk of compromising data security and privacy with built-in algorithms and workflows

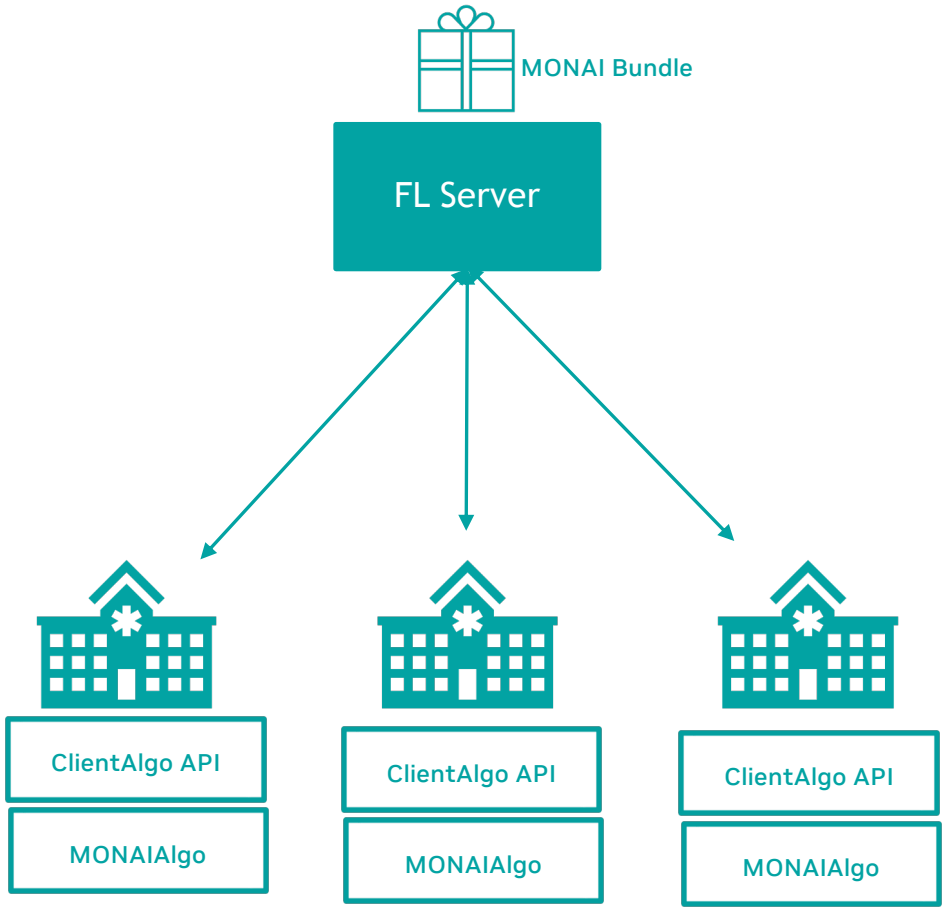
Generalizable Models

Adapt existing AI workflows to an FL paradigm and build generalizable models for any data set

MONAI FEDERATED LEARNING

NEW in v1.0

Low Code Framework for high quality 3D Segmentation Models



Features	Benefits
Federated Learning Client Algo APIs	Ability to defining a MONAI client app that can run on any FL platform. Enabler for FL Toolkits interoperability.
NVIDIA FLARE Integration	Nvidia FLARE, the federated platform developed by Nvidia already integrated with MONAI FL Client APIs
MONAI Bundle compatibility	Allow for seamlessly extending any bundle from MONAI zoo to a federated paradigm

FEDERATED 2.2 NEW FEATURES

From research simulation to real world deployment

FL SIMULATOR

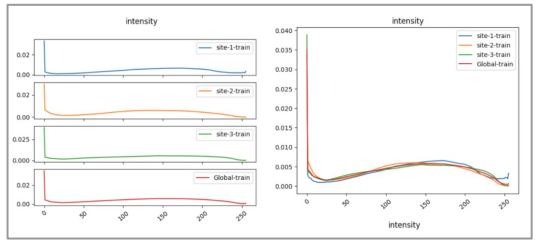
Rapid Development and Debugging

```
def run_simulator(simulator_args):
    simulator = SimulatorRunner(
        job_folder=simulator_args.job_folder,
        workspace=simulator_args.workspace,
        clients=simulator_args.clients,
        n_clients=simulator_args.n_clients,
        threads=simulator_args.threads,
        gpu=simulator_args.gpu,
        max_clients=simulator_args.max_clients,
    )
    run_status = simulator.run()

    return run_status
```

FEDERATED STATS

Analyze data distributions



FRAMEWORKS INTEGRATION

MONAI & XGBOOST

dmlc XGBoost

MONAI+

FLARE DASHBOARD

Streamlined Operation & Deployment



UNIFIED CLI

Simplified Operations



CLIENT CONTROLLED PP

Real-world Site Security



SECURITY & PRIVACY

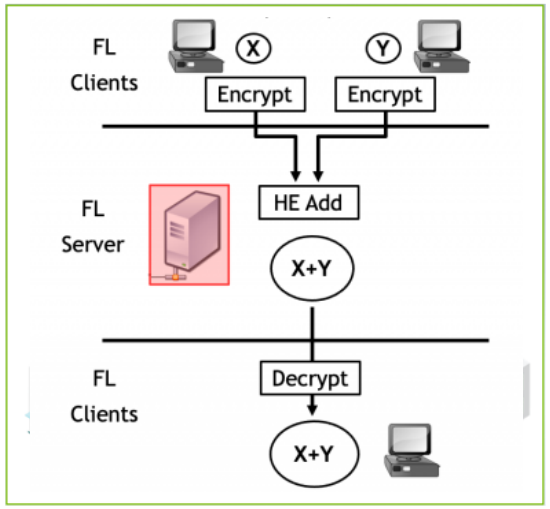
Homomorphic encryption & differential privacy

Federated Learning with Homomorphic Encryption

What if I don't trust the server?

Homomorphic encryption (HE)

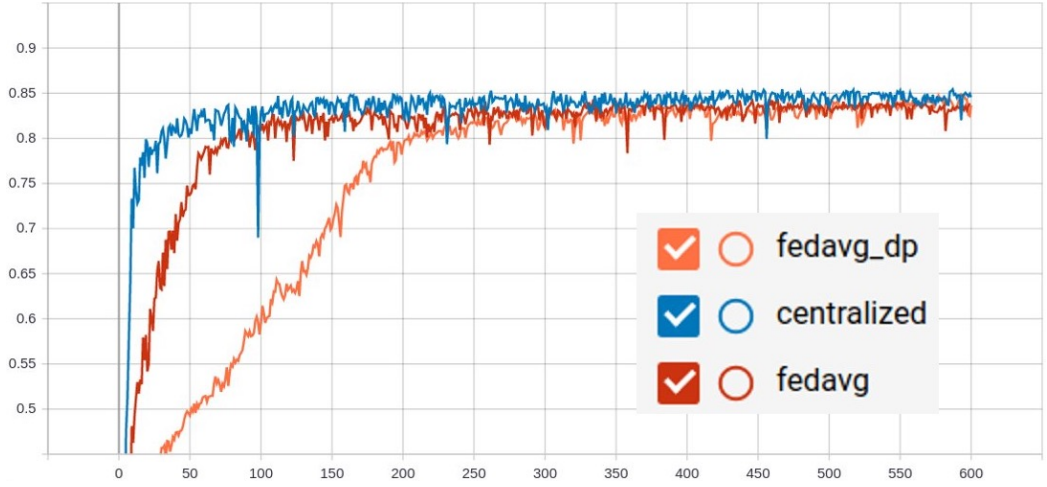
A form of encryption that permits users to perform computations on encrypted data



Blog: https://developer.nvidia.com/blog/federated-learning-with-homomorphic-encryption/
Example: https://github.com/NVIDIA/NVFlare/tree/main/examples/cifar10

Differential Privacy for BraTS18 Segmentation

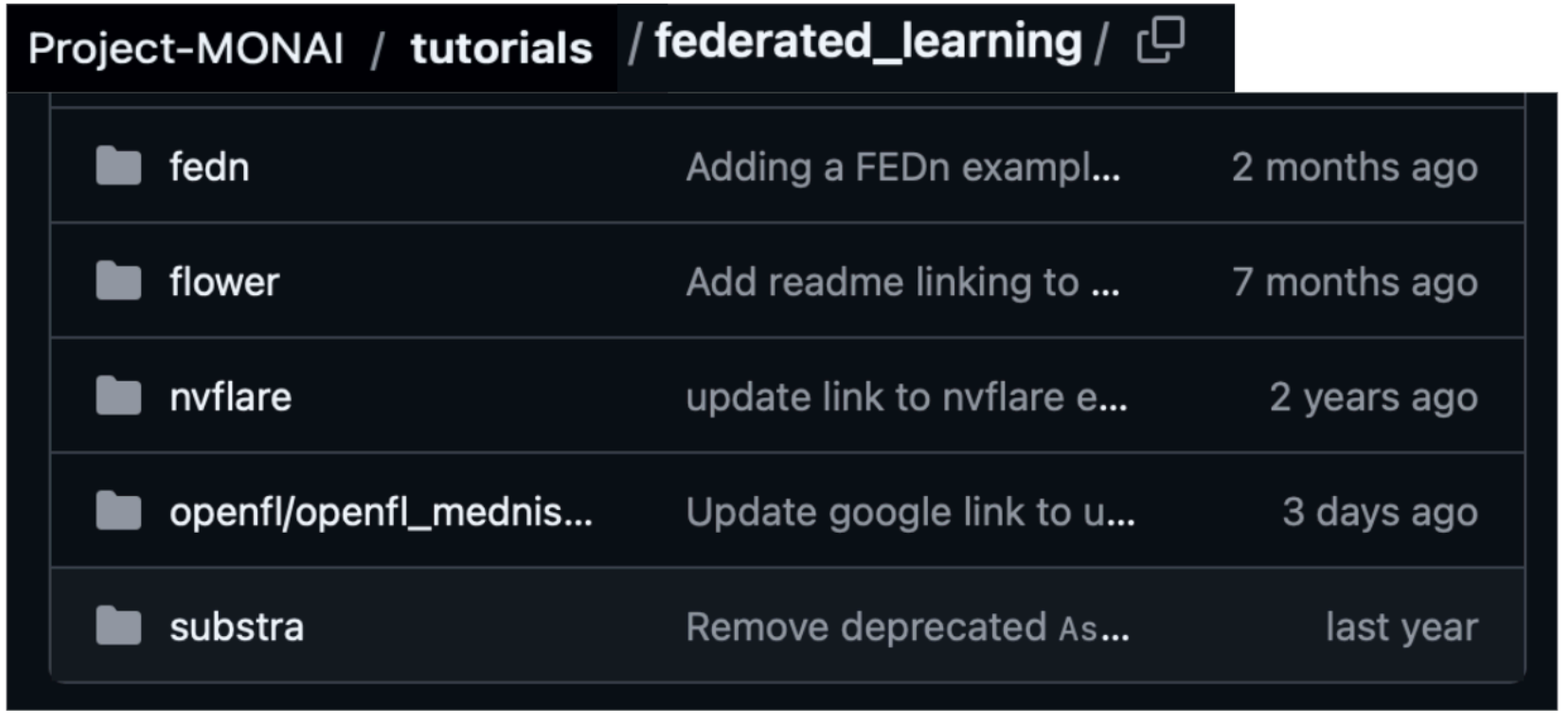
validation Dice scores of the global model for 600 training epochs:



Example: https://github.com/NVIDIA/NVFlare/tree/main/examples/brats18

Federated Learning Hub

- Hub for Federated Medical Imaging: integration with existing federated learning frameworks



Accelerating Medical AI with MONAI FL

Talk outline

01

MONAI

Open source Medical AI

02

Federated Learning

Privacy & Control

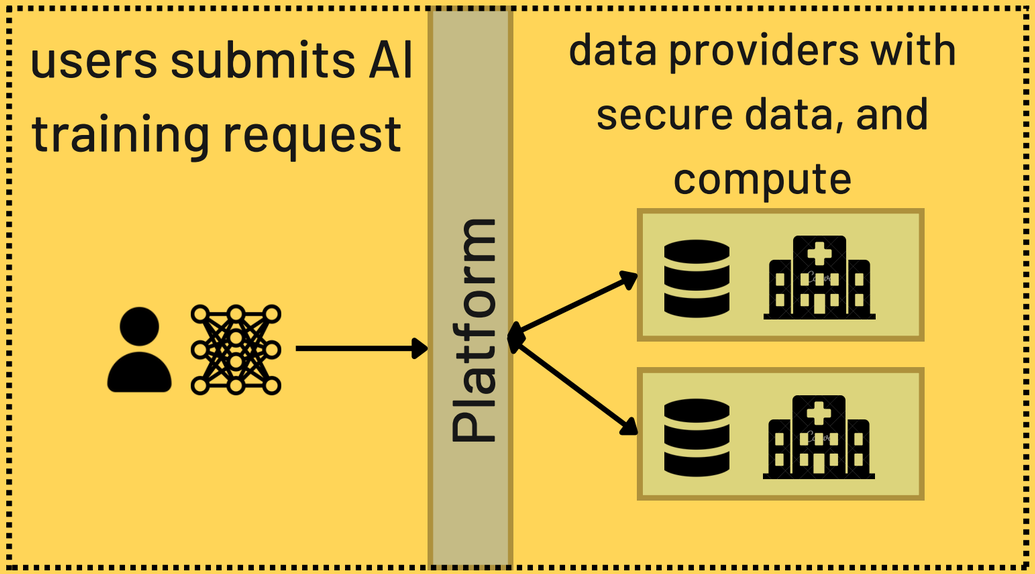
03

Automating Interoperability

Reduce costs

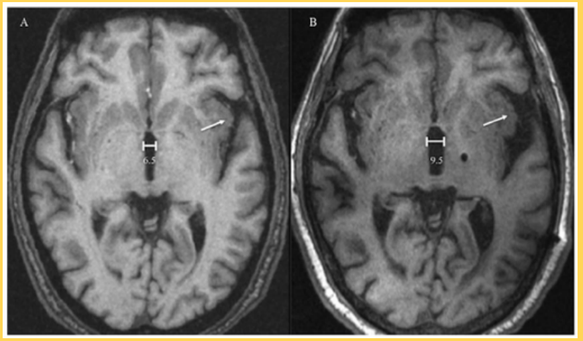
Federated Data Discovery
How to find relevant AI datasets?

- Suppose all datasets are registered on a common platform.
- How do I “search” for datasets relevant for my AI training tasks? **Nontrivial.**
  -



Federated Data Discovery
How to find relevant AI datasets?

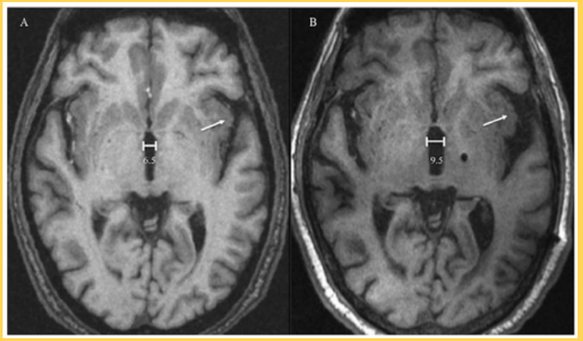
- Suppose all datasets are registered on a common platform.
- How do I “search” for datasets relevant for my AI training tasks? **Nontrivial**
- Can I simply **compare the x’s?**
    -



[Werner, ..., Karimireddy. TMLR 2023]

Federated Data Discovery
How to find relevant AI datasets?

- Can I simply **compare the x's? No**
  - Need **y**. Relevant data for segmentation and classification are different.



[Werner, ..., Karimireddy. TMLR 2023]

Federated Data Discovery
What distinshes "rare" from "bad" data?

- Can I simply **compare the x's? No**
- Can I simply **look at data?**



3 or 5? samples labeled 5.

[Werner, ..., Karimireddy. TMLR 2023]

Federated Data Discovery

What distinshes “rare” from “bad” data?

- Can I simply **compare the x’s? No**
- Can I simply **look at data? No**
  - small model capacity can only use easy samples larger models benefit from rare/weird samples
  - **criterion is also model dependent**.
  - If model can fit both tasks, it is rare and included. Otherwise, it is bad and excluded.



3 or 5? samples labeled 5.

[Werner, ..., Karimireddy. TMLR 2023]

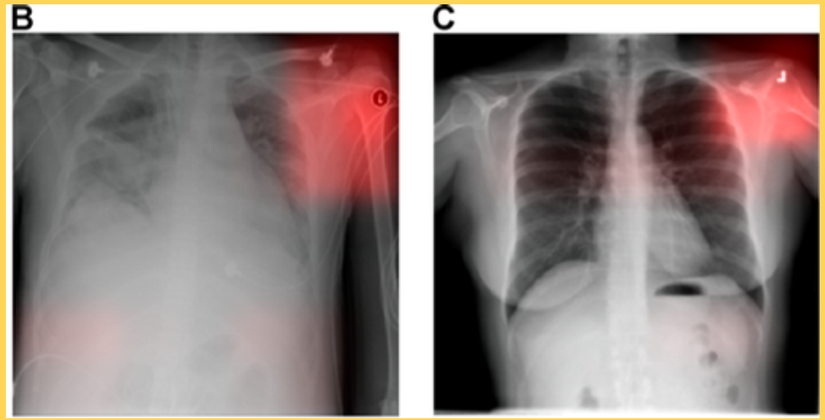
Quality issues in AI Datasets

Hundreds of AI tools have been built to catch covid. None of them helped.

Some have been used in hospitals, despite not being properly tested. But the pandemic could help make medical AI better.

By Will Douglas Heaven

July 30, 2021

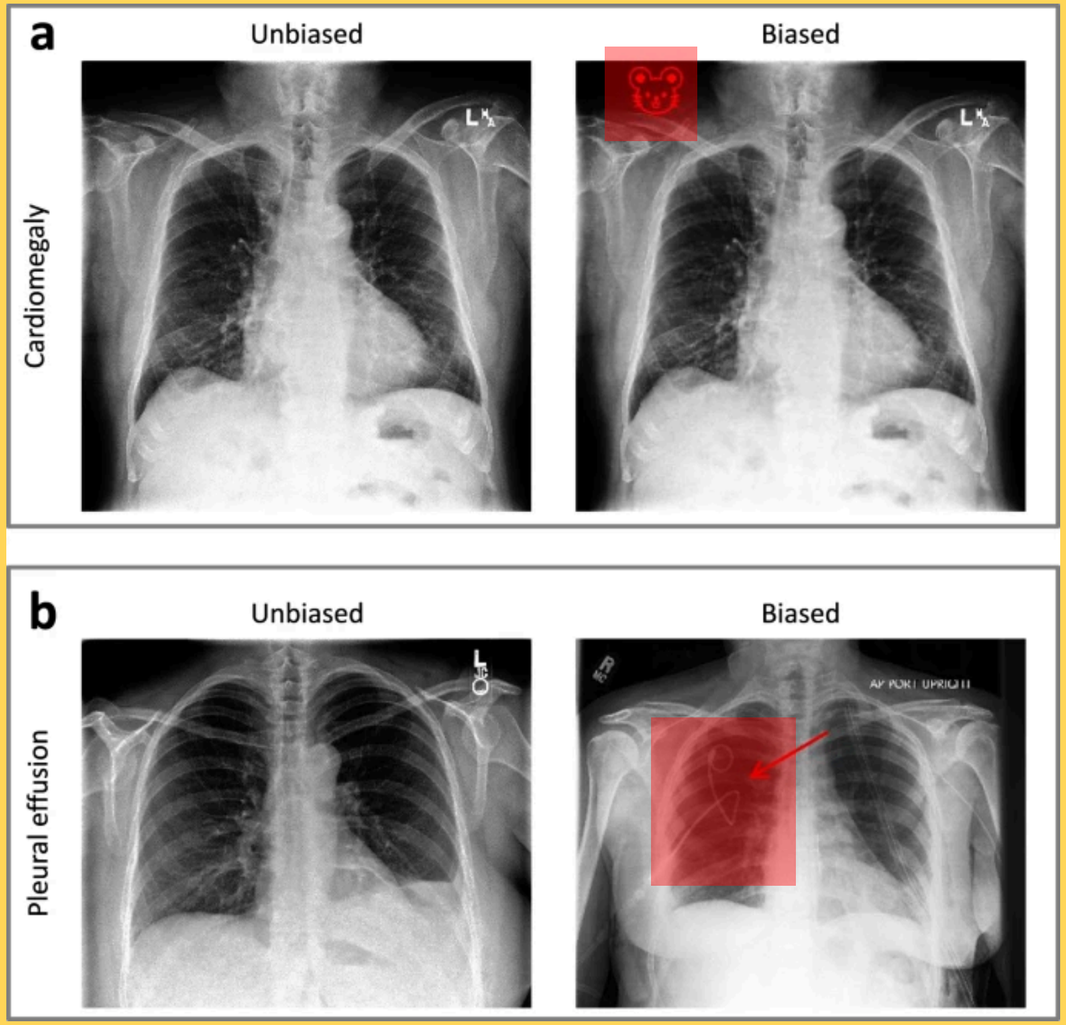


Ringel et al. "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study." PLoS Medicine 2018.

- Ad-hoc data merging = bad idea!
- “Frankenstein” datasets: often have concealed systematic biases.
  - most non-covid examples were children. So model learnt to **recognize kids**.
  - different hospitals treat different severities. Models learnt to **recognize hospitals by the font** on the scans.
- But **can’t look** at data (privacy) :(

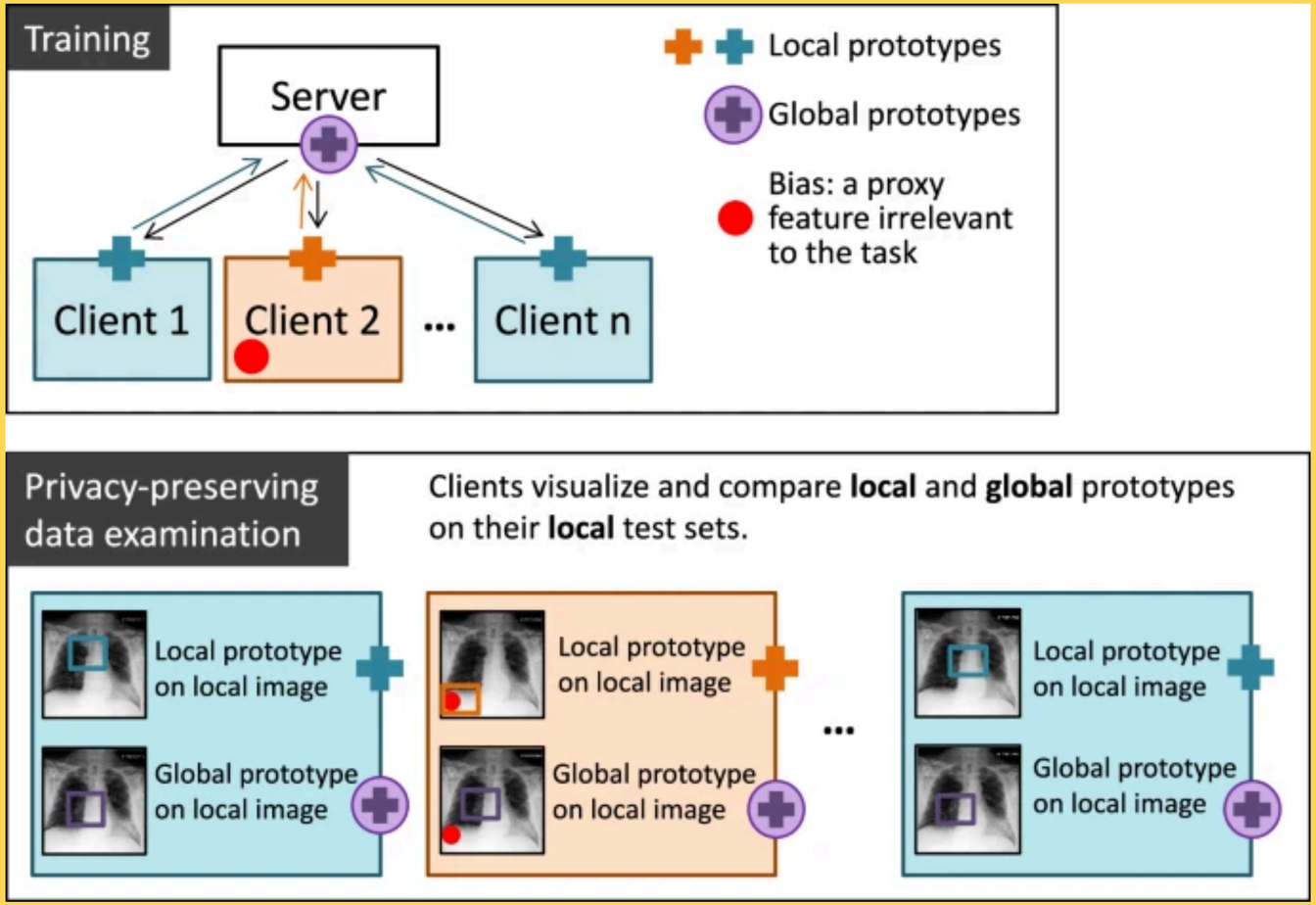
Is my dataset biased?

- Suppose one user has
  - a small red emoji to a positive class in cardiomegaly classification
  - chest drains to a positive class in pleural effusion classification
- How can I tell if my data is biased without looking at other's data?



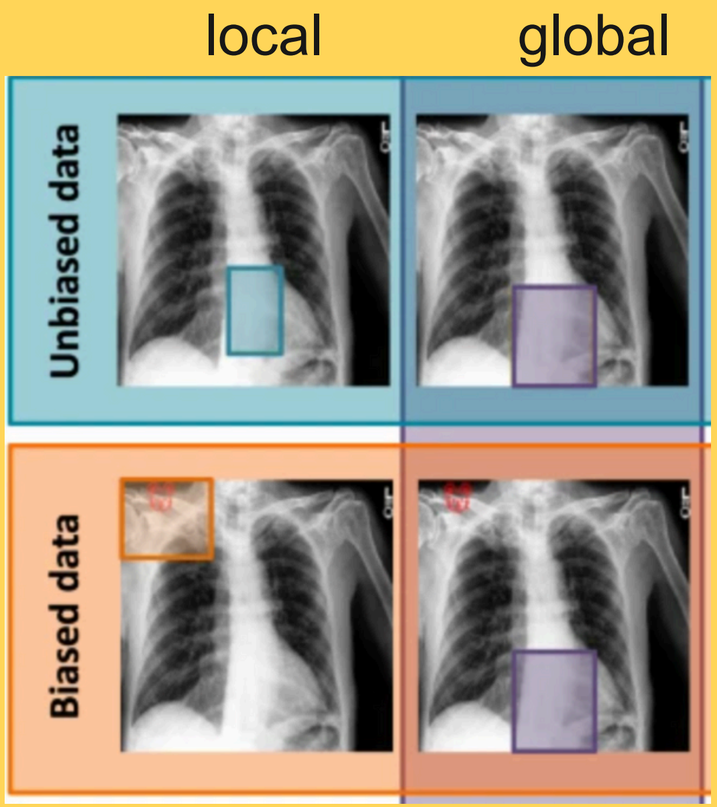
MyTH: MyThisYourThat

- Idea: train a **robust explainable model** on the **global** dataset, and on **local** data.
- If there is bias, local model would have picked it up.
- Compare the **explanations** given by the two models.

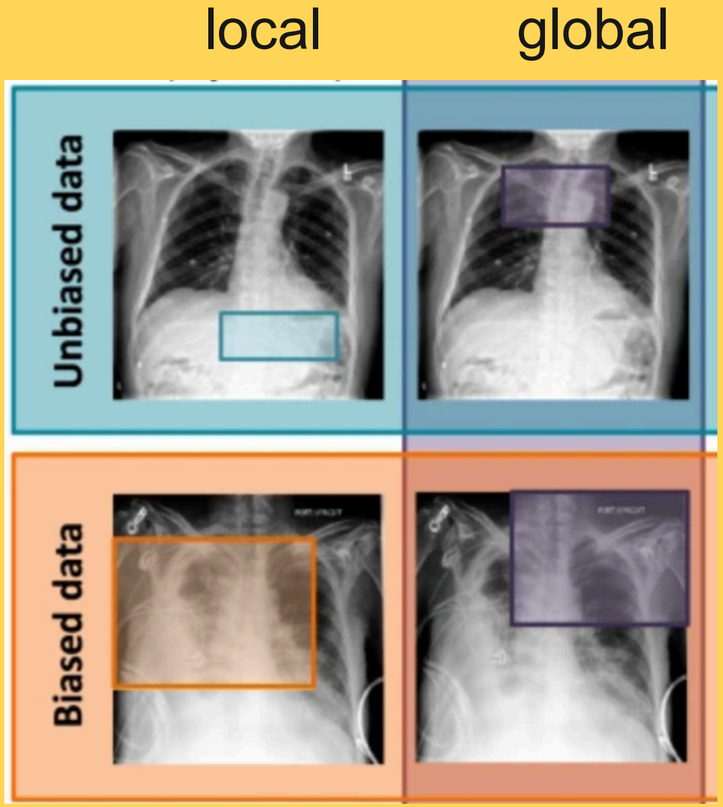


[Naumova, Devos, Karimireddy et al. npj Digital Medicine 2024]

MyTH: MyThisYourThat
Comparing Protoyptes



Cardiomegaly classification

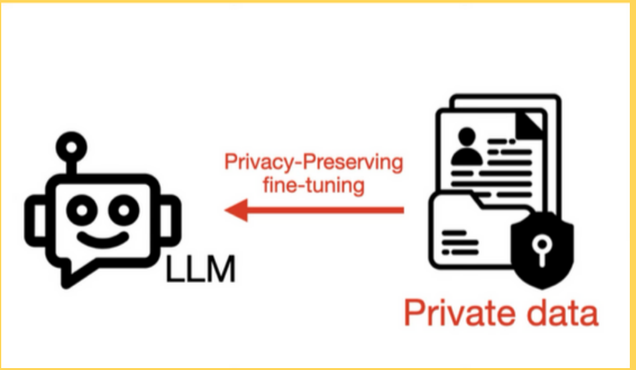


Pleural Effusion classification

[Naumova, Devos, Karimireddy et al. npj Digital Medicine 2024]

Interoperability leveraging GenAI

- Use GenAI to generate **private synthetic data**
  - private => low quality
  - but sufficient for debugging



- Use Agentic AI to **automatically tranform datasets**
  - Reason about compatibiltiy using meta-data
  - Given xyz format, transform into DICOM
  - Automatically infer missing fields



Works in progress

Current State

- Extensive outreach to folks working on real world FL to understand their data compatibility issues

- Incorporated Data Quality Assessment Tools into MONAI FL - the AI readiness tools in APPFL.

  Completeness, Outliers, Duplicates, Privacy, Fairness, FAIR Compliance, Feature Correlations, Feature Relevancy, Class Imbalance

- Working on outlier detection and data transformation techniques to improve cross-center data compatibility for medical image segmentation using GenAI.



Table 1: Data Quality Assessment Tools

Tool	Completeness	Outliers	Duplicates	Privacy	Fairness	FAIR Compliance	Feature Correlations	Feature Relevancy	Class Imbalance
Informatica [30]	✓								
DQLearn [47]	✓	✓	✓						
Gupta et al [25]	✓	✓	✓		✓		✓	✓	✓
Data Readiness Report [5]	✓	✓				✓	✓		
AI360 [8]					✓				
FAIR Cookbook [43]						✓			
FAIRassist [19]						✓			
ESS-DIVE FAIR [18]						✓			
AIDRIN	✓	✓	✓	✓	✓	✓	✓	✓	✓

Conclusion: FAIR for Medical AI Data

Lots of open research questions



- **Privacy and Control**: using **federated learning**
- Mitigate **Costs**: by **automating data interoperability** and discovery **using AI.**
- **IP and incentives**: data market mechanisms. ask me later!