

Distributed and Federated Learning

Federated learning for healthcare data

Yuriy Kochura

Ivan Zhuk

Yuri Gordienko



National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

Us

This material is presented by:

- Theoretical lecture: **Yuri Gordienko**
- Exercise session: **Ivan Zhuk** and **Yuriy Kochura**

Feel free to contact **Yuriy Kochura** at iuriy.kochura@gmail.com or on Telegram – [@y_kochura](https://t.me/y_kochura) for assistance.



Yuriy Kochura



Ivan Zhuk



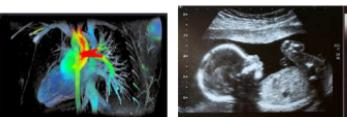
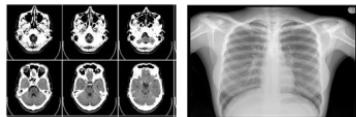
Yuri Gordienko

Agenda

- 🎙 AI in Healthcare
- 🎙 Market Size & Trends
- 🎙 Federated Learning (FL) in Healthcare
- 🎙 Why we need FL?
- 🎙 How does FL work?
- 🎙 Data harmonization challenges for FL
- 🎙 Large-scale standardized dataset (MedMNIST)
- 🎙 Future directions for FL in Healthcare

AI in Healthcare

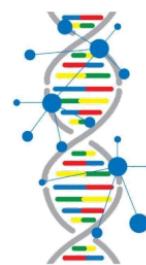
AI in medicine



RADIOLOGY
CT, MR, US, X-RAY



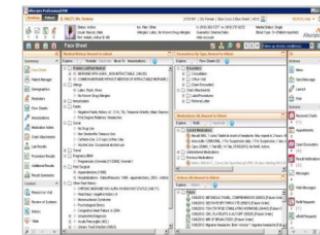
PATHOLOGY
TISSUE & CELL



GENETICS
DERMATOLOGY
OPHTHALMOLOGY



• • •

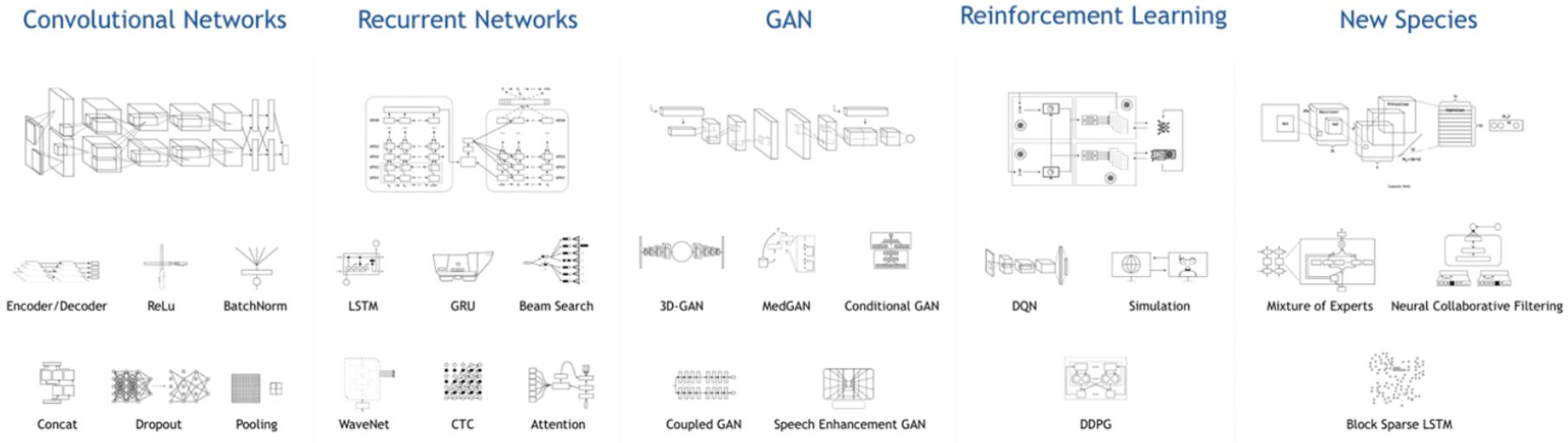


ELECTRONIC HEALTH
RECORDS

Year	FDA-Approved AI Tools	Global Investment (USD)
2021	~128	~\$7B
2024	~169	~\$26.57B

From research ...

Improving state of the art performance in controlled settings



... to applications

Achieving human-level performance on large data and clinical settings

Skin lesion image

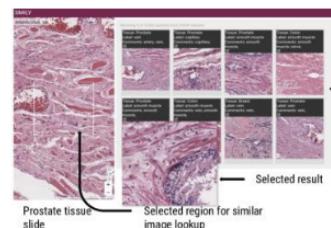
Deep convolutional neural network (Inception v3)

Our classification technique is a deep CNN.
Data flow is from left to right.



- 129450 clinical images
- 2032 diseases
- Skin cancer detection
- comparable to dermatologists

Esteva et al, Dermatologist-level classification of skin cancer with deep neural networks, Nature 2017

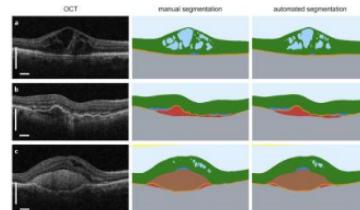


Hegde et al., Similar Image Search for Histopathology: SMILY, Nature digital medicine 2019



- 6 hospitals in China
- 84424 individuals
- 1036496 endoscopy images
- Gastrointestinal cancer detection
- Perf. similar to the expert endoscopist

Luo et al., Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study, Lancet Oncology 2019



- “Only” 14884 OCT 3D scans
- Resolution ~ 5 µm
- Volumetric multi-region segmentation
- Performance comparable to humans

De Fauw et al., Clinically applicable deep learning for diagnosis and referral in retinal disease, Nature Medicine 2018



Camels, Code & Lab Coats: How AI Is Advancing Scienc...

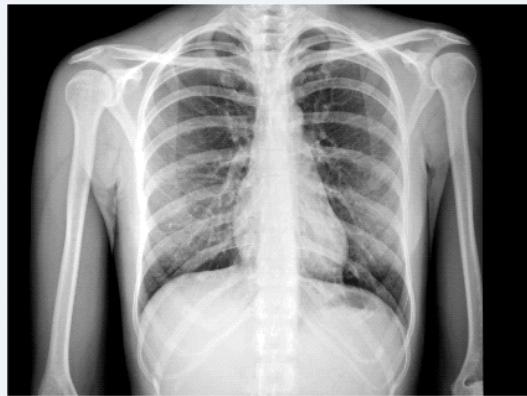


Share



How AI is advancing medicine (Google, 2018)

Can you write me a report analyzing this chest X-ray?



Findings:

- Devices: None.
- Lungs: No pneumothorax. No substantial pleural effusion. Lungs appear clear.
- Cardiomediastinal: Normal heart size. Mediastinal contours within normal limits.
- Other: No acute skeletal abnormality.

Impression:

No active disease seen in chest.

Enter a question here

Med-PaLM 2 (Google) is a large language model tuned for the medical domain. It reaches 85%+ accuracy on the MedQA medical exam benchmark in research.

AlphaFold: From a sequence of amino acids to a 3D structure

nature

Explore content ▾ About the journal ▾ Publish with us ▾

nature > articles > article

Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

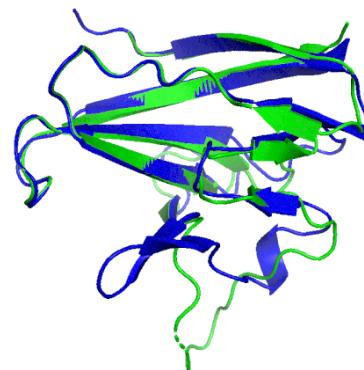
John Jumper , Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis 
+ Show authors

Nature 596, 583–589 (2021) | Cite this article

1.42m Accesses | 12k Citations | 3493 Altmetric | Metrics

Abstract

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort^{1,2,3,4}, the structures of around 100,000 unique proteins have been determined⁵, but this represents a small fraction of the billions of known protein sequences^{6,7}. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid sequence—the structure prediction component of the ‘protein folding problem’⁸—has been an important open research problem for more than 50 years⁹. Despite recent progress^{10,11,12,13,14}, existing methods fall far short of atomic accuracy, especially when no homologous structure is available. Here we provide the



John Jumper

(The 2024 Nobel Prize in
Chemistry)

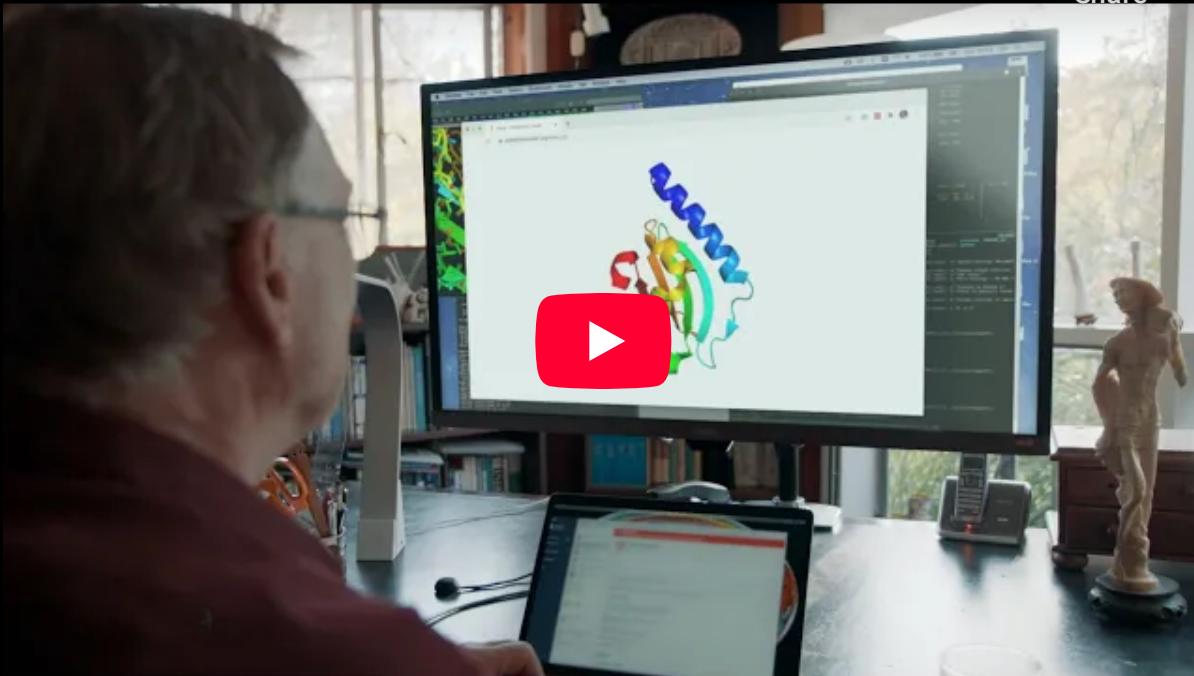




AlphaFold: The making of a scientific breakthrough



Share



Watch on YouTube

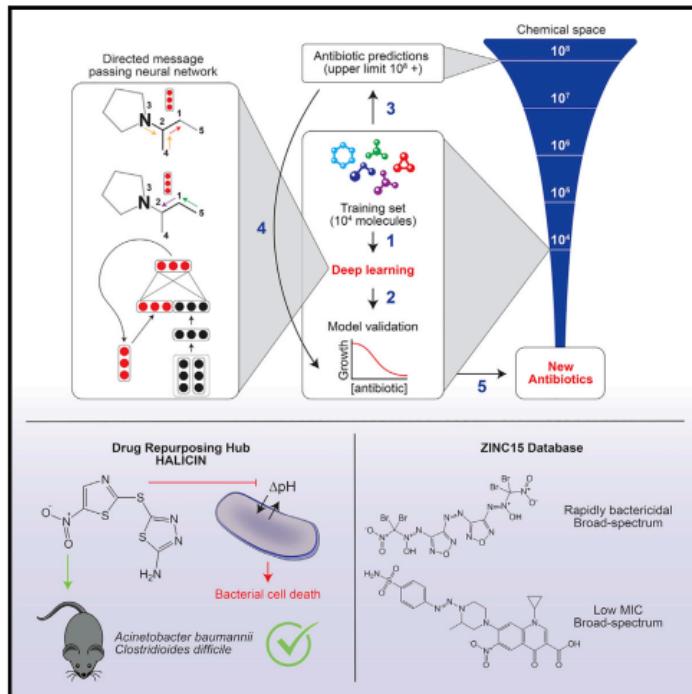
AI for Science (Deepmind, AlphaFold, 2020)

Drug discovery with graph neural networks

Cell

A Deep Learning Approach to Antibiotic Discovery

Graphical Abstract



Authors

Jonathan M. Stokes, Kevin Yang,
Kyle Swanson, ..., Tommi S. Jaakkola,
Regina Barzilay, James J. Collins

Correspondence

regina@csail.mit.edu (R.B.),
jimjc@mit.edu (J.J.C.)

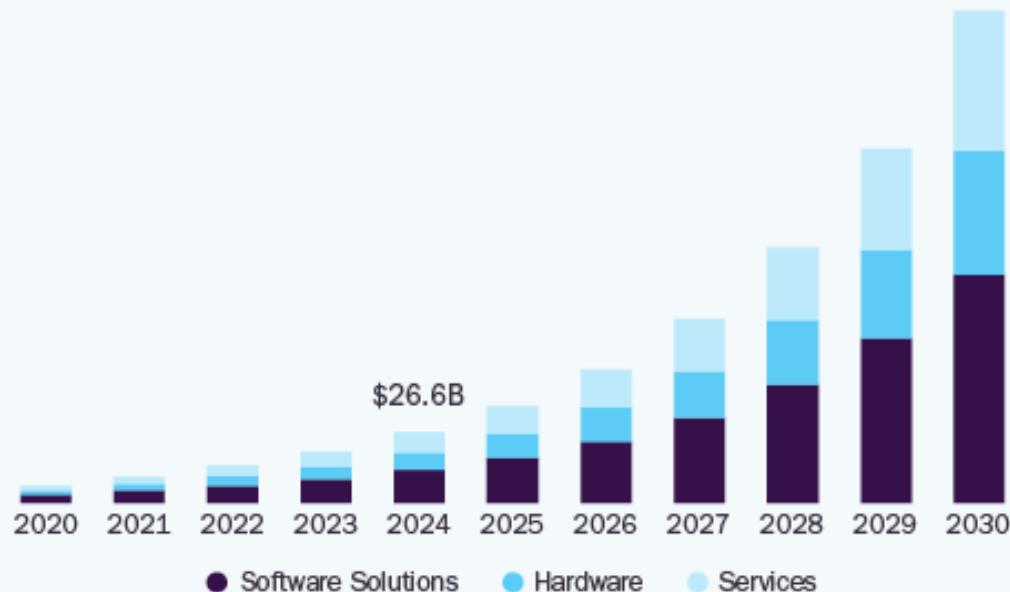
In Brief

A trained deep neural network predicts antibiotic activity in molecules that are structurally different from known antibiotics, among which Halicin exhibits efficacy against broad-spectrum bacterial infections in mice.

Market Size & Trends

AI In Healthcare Market

Size, by Component 2020 - 2030 (USD Billion)



38.6%

Global Market CAGR,
2025 - 2030

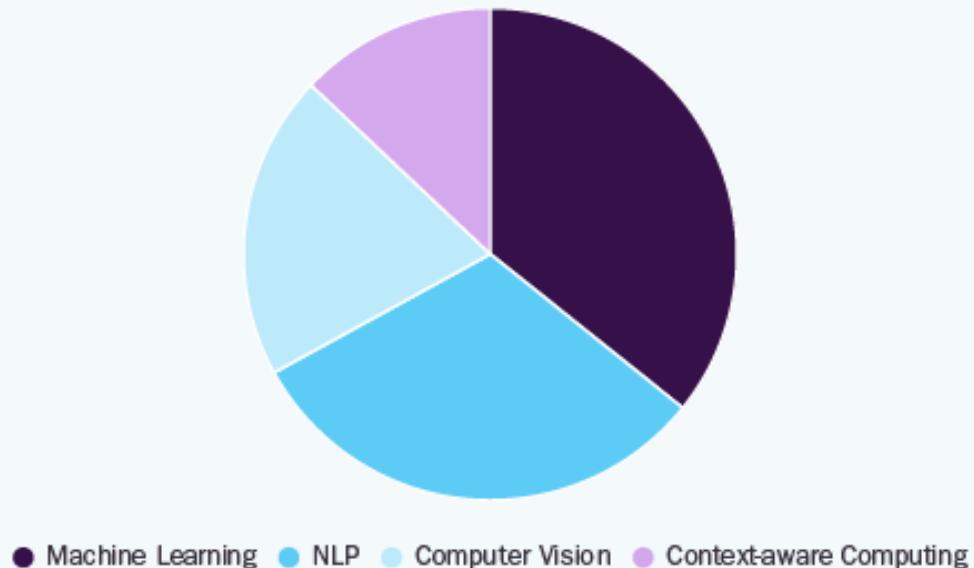
Source:
www.grandviewresearch.com

CAGR (Compound Annual Growth Rate)

A key factor driving market growth is the increasing demand in the healthcare sector for enhanced efficiency, accuracy, and better patient outcomes.

AI In Healthcare Market

Share, by Technology 2024 (%)



\$26.6B

Global Market Size,
2024

Source:
www.grandviewresearch.com

The **context-aware computing** segment is expected to grow at the fastest CAGR between 2025 and 2030. AI algorithms integrate and analyze diverse data sources such as electronic health records (EHRs), real-time vital signs, medical history, environmental factors, and patient activity to understand patient context dynamically.

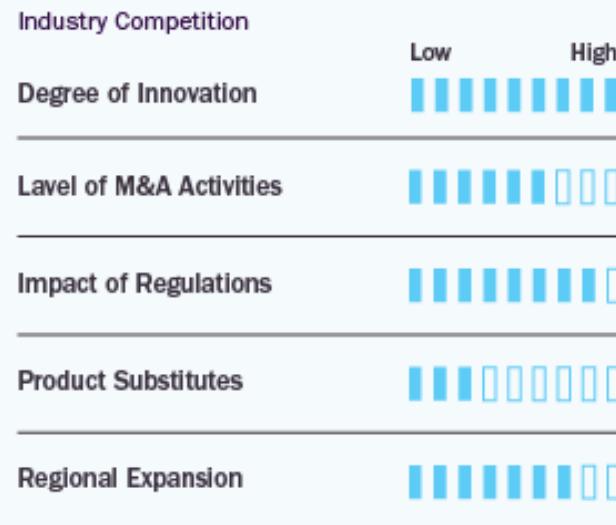
AI In Healthcare Industry Dynamics



Market Concentration



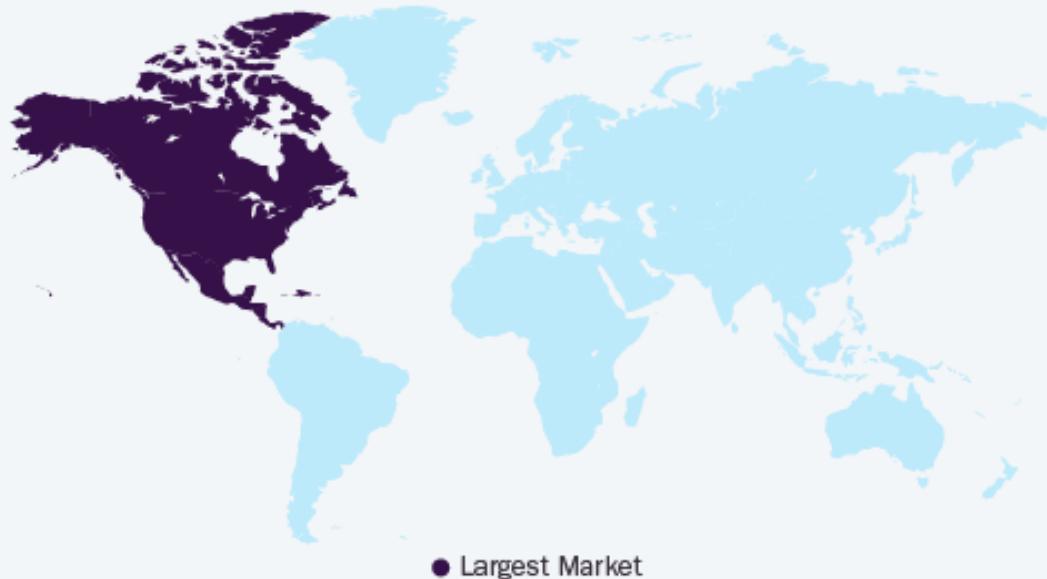
Market Characteristics



The Healthcare AI market exhibits a high degree of innovation, characterized by ongoing advancements in technology. Rapid developments in machine learning, deep learning, natural language processing, and computer vision are driving the evolution of AI-powered healthcare solutions.

AI In Healthcare Market

Trends, by Region, 2025 - 2030



54.0%

North America Market
Revenue Share, 2024

Source:
www.grandviewresearch.com

North America AI in healthcare industry dominated the global market and accounted for the largest revenue share of over 54% in 2024.

AI in healthcare industry in the UK held the largest market share in the European region in 2024.

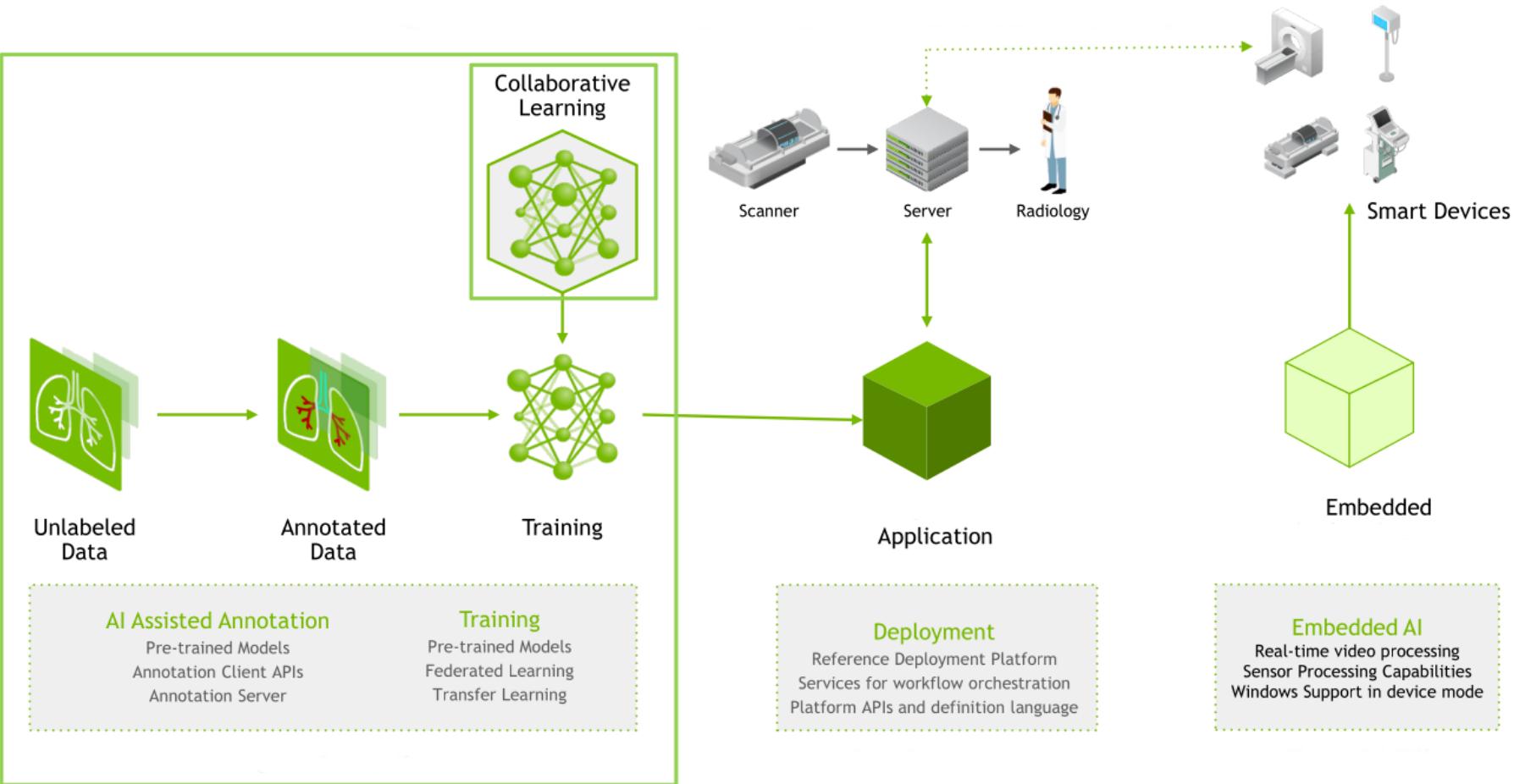
Key Companies

The following are the **leading companies** in the AI in healthcare market. These companies collectively hold the largest market share and dictate industry trends:

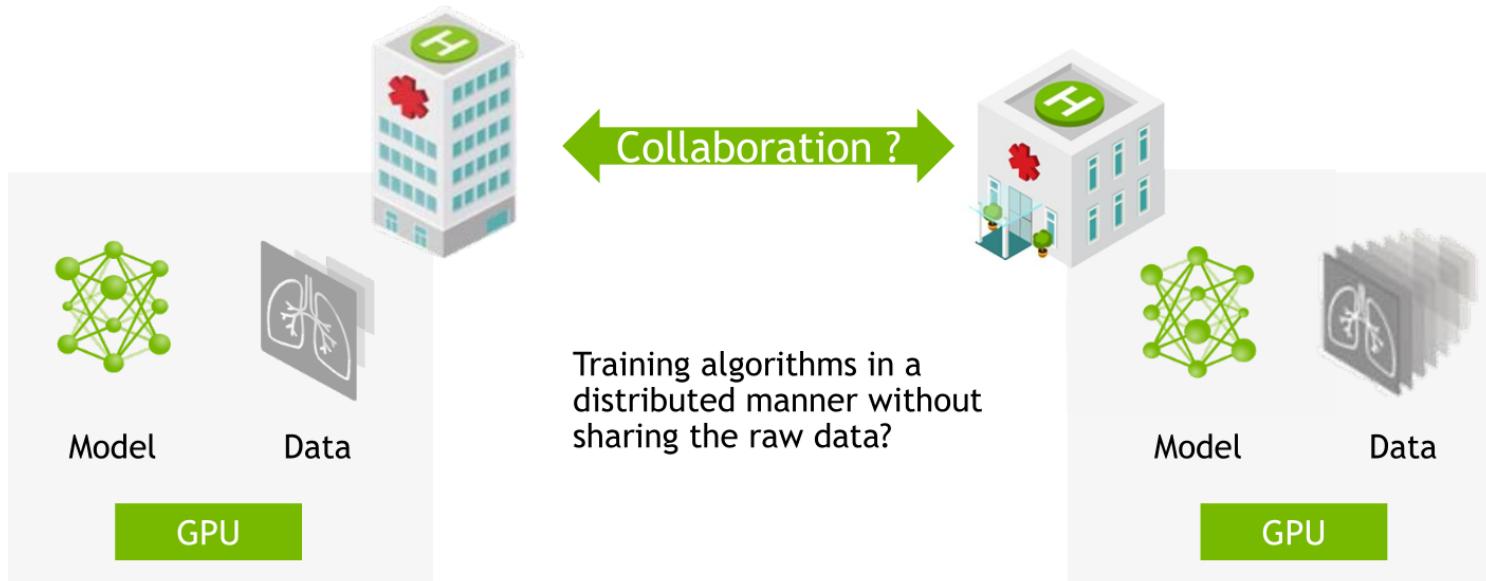
- Microsoft
- IBM
- Google
- NVIDIA Corporation
- Intel Corporation
- Itrex Group
- GE Healthcare
- Medtronic
- Oracle
- Medidata
- Merck
- IQVIA

Federated Learning (FL) in Healthcare

Application framework



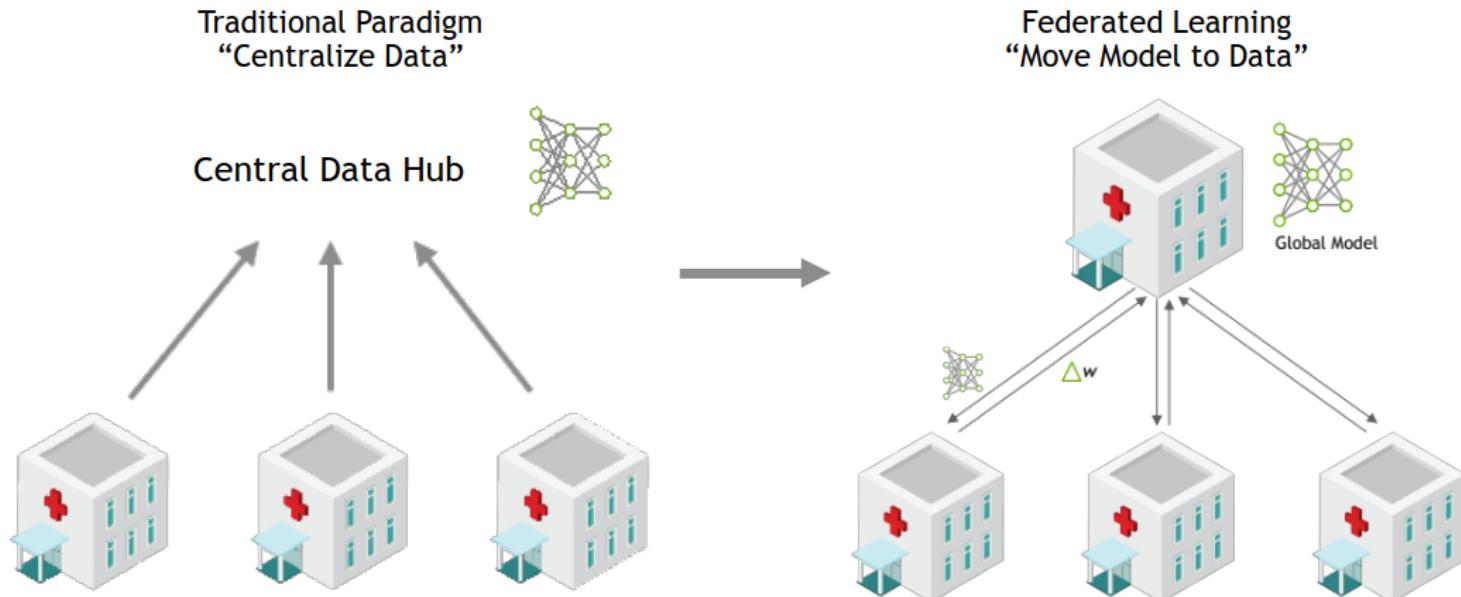
Data-driven medicine requires federated efforts



Possible Solution:

Federated Learning – learning paradigm that allows to integrate knowledge learned from non co-located data that resides within the participating entities into a global machine learning model.

Paradigm Shift



Federated Learning

- Enhanced data privacy and security
- Raw data never leaves the device
- Improved model performance and diversity
- Scalability and efficiency

The future of digital health with federated learning

Positioning Federated Learning (FL) for Healthcare

- Consensus view on FL in healthcare
- Benefits and impact of FL for medical applications
- Key challenges of implementing FL

npj | digital medicine

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [npj digital medicine](#) > [perspectives](#) > [article](#)

Perspective | [Open access](#) | Published: 14 September 2020

The future of digital health with federated learning

[Nicola Rieke](#)  [Jonny Hancox](#), [Wenqi Li](#), [Fausto Milletari](#), [Holger R. Roth](#), [Shadi Albarqouni](#), [Spyridon Bakas](#), [Mathieu N. Galtier](#), [Bennett A. Landman](#), [Klaus Maier-Hein](#), [Sébastien Ourselin](#), [Micah Sheller](#), [Ronald M. Summers](#), [Andrew Trask](#), [Daguang Xu](#), [Maximilian Baust](#) & [M. Jorge Cardoso](#)

[npj Digital Medicine](#) 3, Article number: 119 (2020) | [Cite this article](#)

140k Accesses | 152 Altmetric | [Metrics](#)

Abstract

Data-driven machine learning (ML) has emerged as a promising approach for building accurate and robust statistical models from medical data, which is collected in huge volumes by modern healthcare systems. Existing medical data is not fully exploited by ML primarily because it sits in data silos and privacy concerns restrict access to this data. However, without access to sufficient data, ML will be prevented from reaching its full potential and, ultimately, from making the transition from research to clinical practice. This paper considers key factors contributing to this issue, explores how federated learning (FL) may provide a solution for the future of digital health and highlights the challenges and considerations that need to be addressed.

Why we need FL?

Data is the Key to Generalizable, Robust AI Models

Decentralized AI = federated compute + decentralized data, model



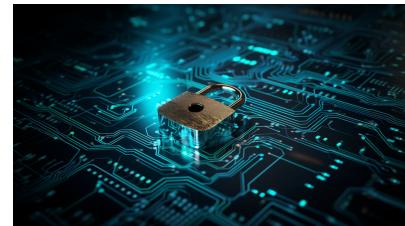
Regulations

Health data is highly sensitive, subject to regulations and cannot easily be shared.



Data Availability

More data in private domain than public. Distributed data cross-country, rare data siloed and sparsely distributed.



Preserve Privacy

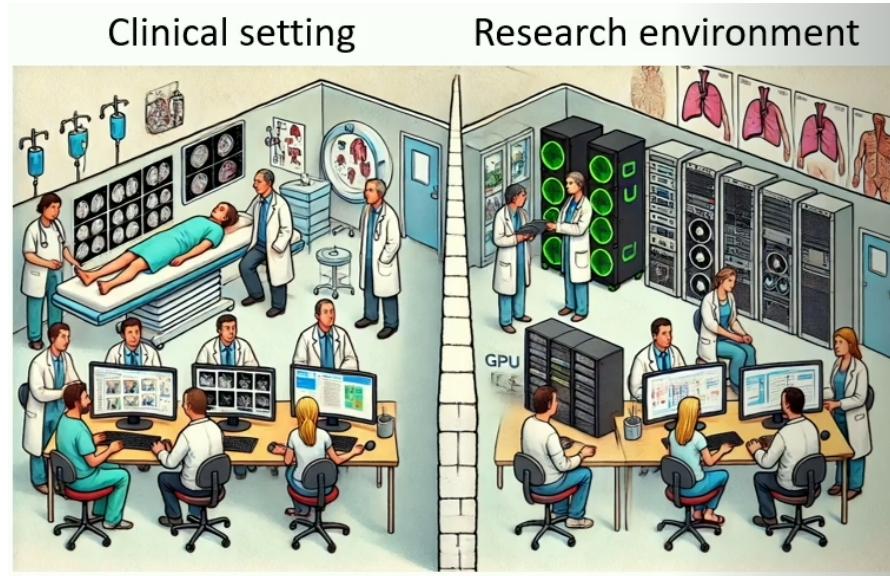
Leveraging private data while preserving data privacy is the mission of **federated learning**.

Real-world medical AI development needs **external validation, multiple institutions, prospective data**. Among 500+ published medical AI studies, only 6% of have external validation. Few included multiple institutions.

Design Characteristic	All Articles (n = 516)	Articles Published in Medical Journals (n = 437)	Articles Published in Non-Medical Journals (n = 79)	P*
External validation				1.000
Used	31 (6.0)	27 (6.2)	4 (5.1)	
Not used				
	485 (94.0)	410 (93.8)	75 (94.9)	
In studies that used external validation				
Diagnostic cohort design	5 (1.0)	5 (1.1)	0 (0)	1.000
Data from multiple institutions	15 (2.9)	12 (2.7)	3 (3.8)	0.713
Prospective data collection	4 (0.8)	4 (0.9)	0 (0)	1.000
Fulfillment of all of above three criteria	0 (0)	0 (0)	0 (0)	1.000
Fulfillment of at least two criteria	3 (0.6)	3 (0.7)	0 (0)	1.000
Fulfillment of at least one criterion	21 (4.1)	18 (4.1)	3 (3.8)	1.000

Challenges: data sharing, regulation, legal, privacy, technical ...

Separation of clinical data and compute

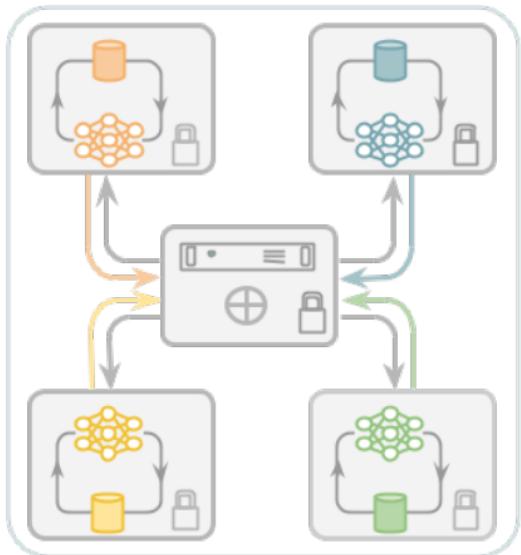


- Compute clusters typically operate outside Protected Health Information (PHI) compliant environments, limiting direct access to medical data.
- Secure, compliant environments often have limited compute resources.
- Strict firewalls and security measures complicate connecting to federated learning servers.

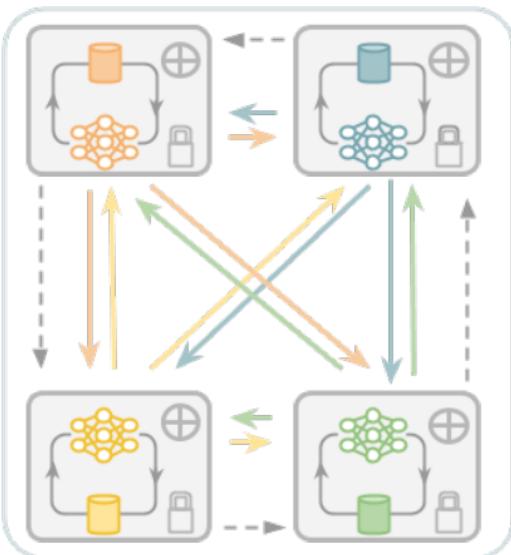
How does FL work?

Communication Architectures

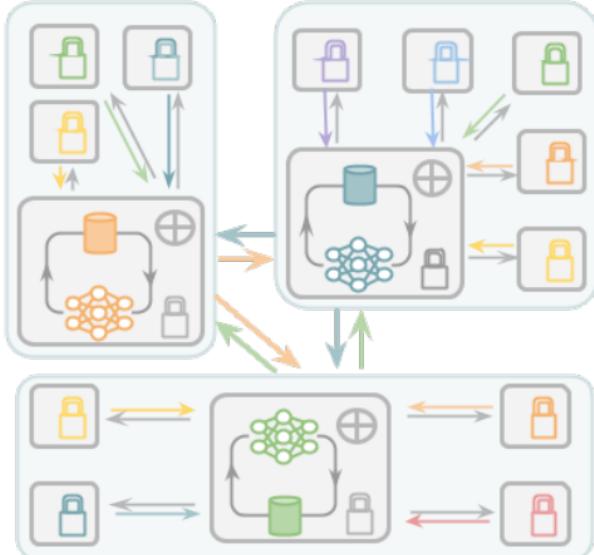
a) Client/Server Architecture



b) Peer to Peer Architecture



c) Hybrid Architecture



Federation
of Nodes

Medical
Database

Consensus Model
Redistribution

Model
Aggregation

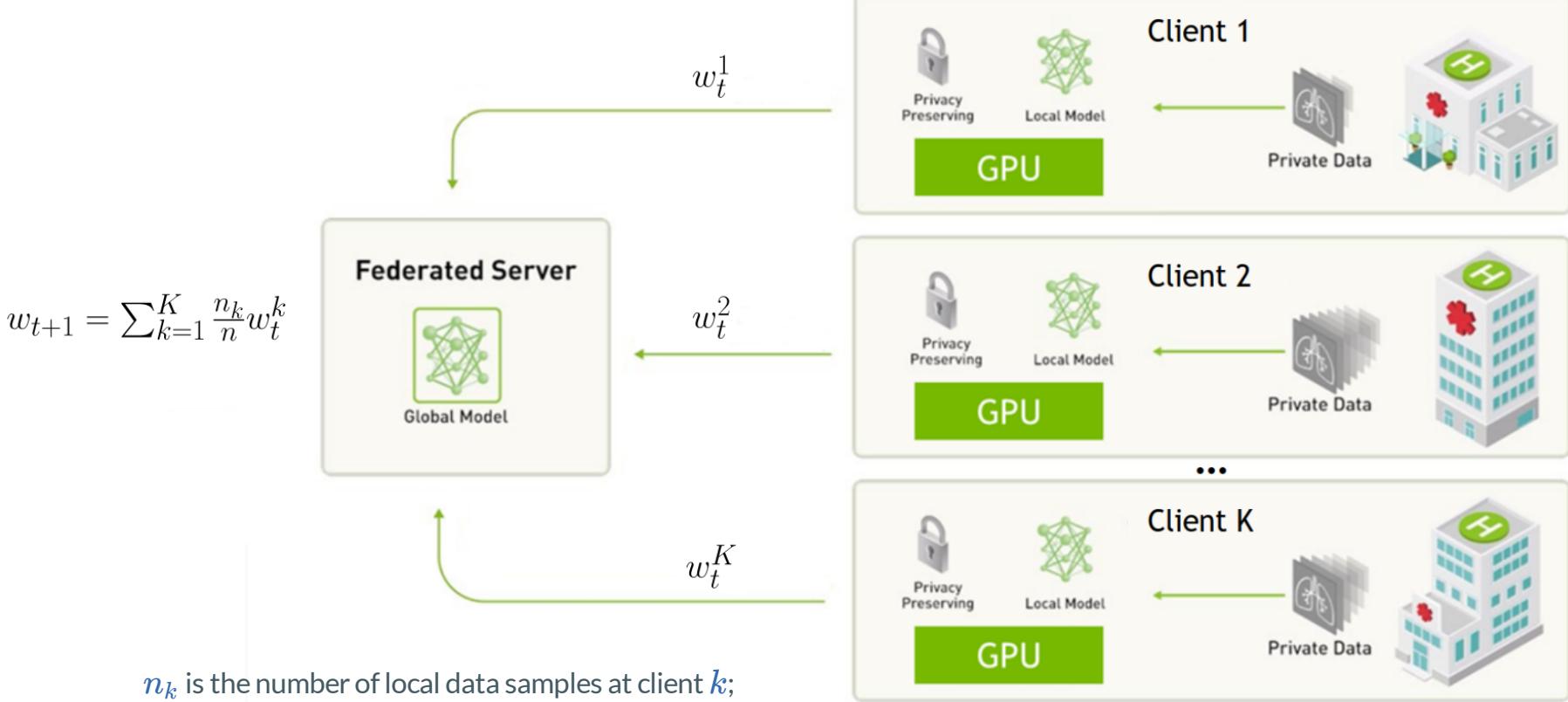
Locally Trained
Model

Secure Compute
Node

Model
Forwarding

Cyclic Learning

Server-Client Federated Learning: Averaging



n_k is the number of local data samples at client k ;

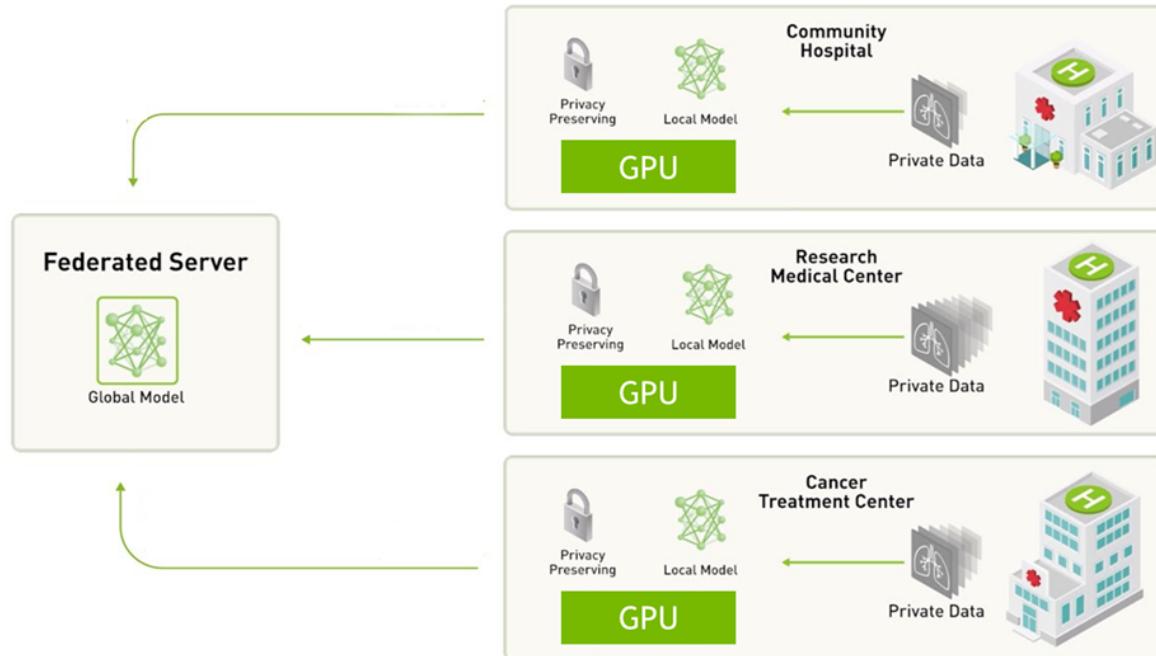
n is the total number of samples across all participating clients;

w_t^k – model weights trained locally by client k ;

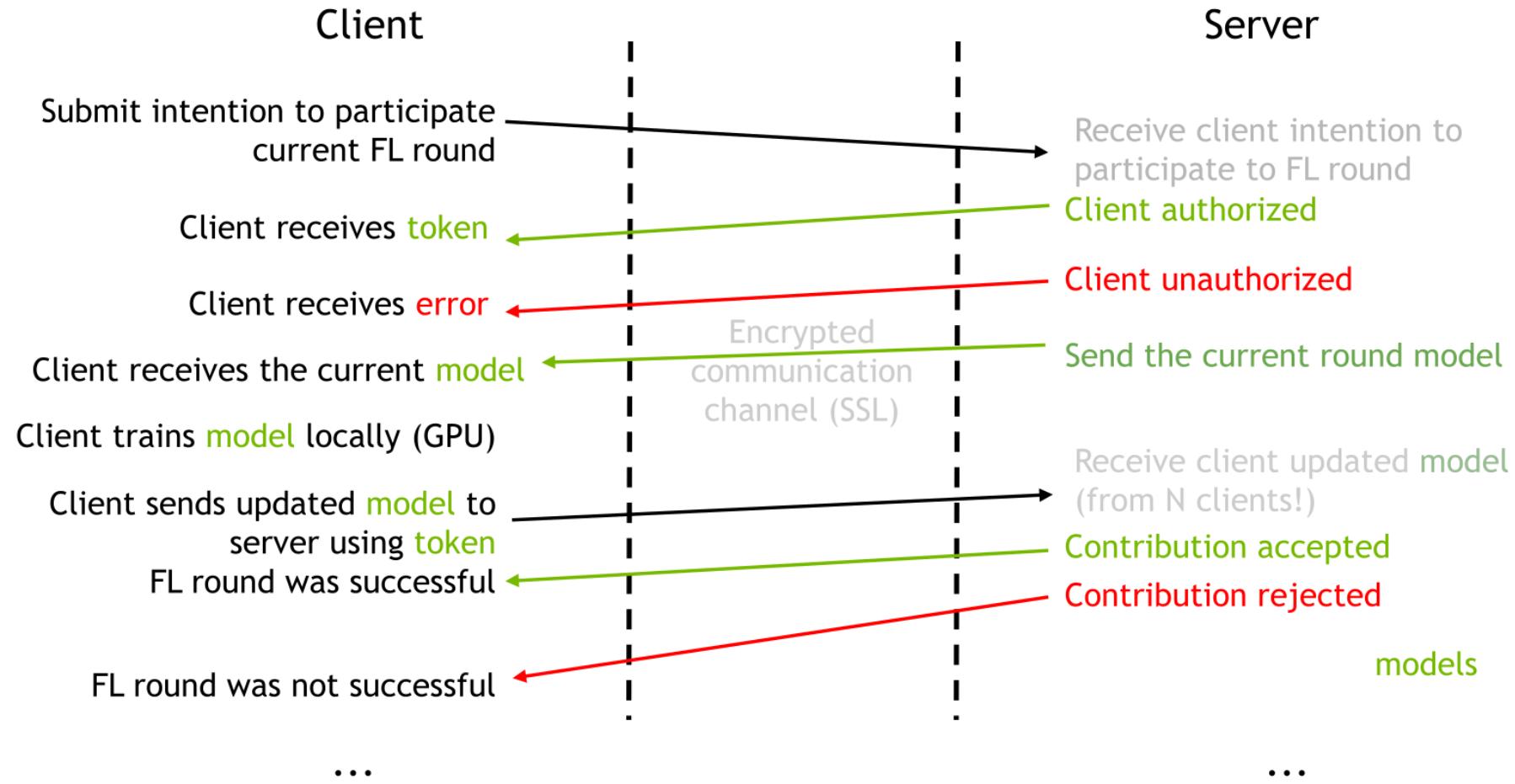
t is a communication round.

Server-Client Federated Learning

- **Server:** manages job lifecycle, assigns computation tasks to clients & aggregate.
- **Clients:** perform tasks assigned by the server.
- **Workflow:** a distributed workflow with specific optimization algorithms and communication strategies.



High-level implementation



Data harmonization challenges for FL

- **Protocol and equipment variability**

- Differences in imaging methods and machines between institutions.

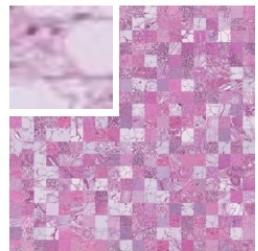
- **Protocol and equipment variability**
 - Differences in imaging methods and machines between institutions.
- **Data interoperability**
 - Using common formats and processing methods for easier analysis.
 - In medical imaging DICOM standard is widespread, though clinical data such as labs is less standardized.

- **Protocol and equipment variability**
 - Differences in imaging methods and machines between institutions.
- **Data interoperability**
 - Using common formats and processing methods for easier analysis.
 - In medical imaging DICOM standard is widespread, though clinical data such as labs is less standardized.
- **Labeling differences**
 - Inconsistent terminology and labels are used across different organizations.

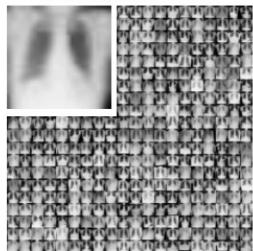
Large-scale standardized dataset (MedMNIST)

The **MedMNIST** dataset consists of **12 datasets for 2D** and **6 datasets for 3D**. Covers key medical imaging modalities (X-Ray, OCT, Ultrasound, CT, Electron Microscopy). Supports tasks (**binary/multi-class classification, ordinal regression, multi-label**), and scales from **10^2** to **10^5** samples. Multiple size options: **28 (MNIST-Like)**, **64**, **128**, and **224**.

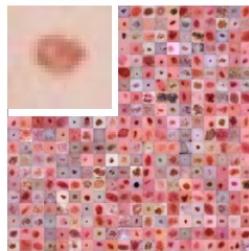
PathMNIST



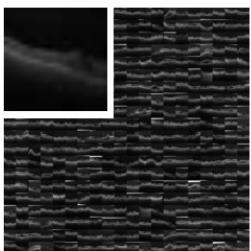
ChestMNIST



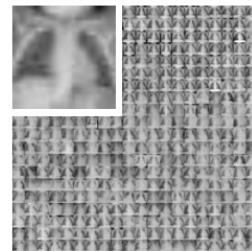
DermaMNIST



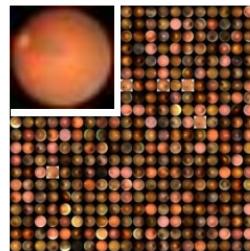
OCTMNIST



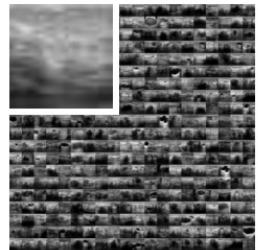
PneumoniaMNIST



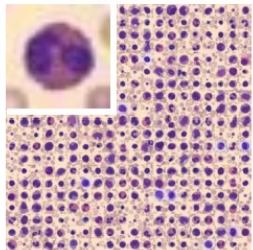
RetinaMNIST



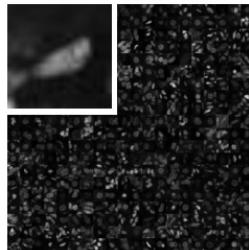
BreastMNIST



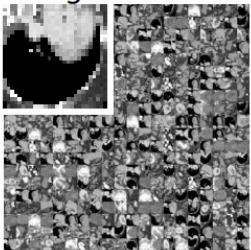
BloodMNIST



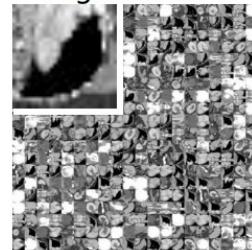
TissueMNIST



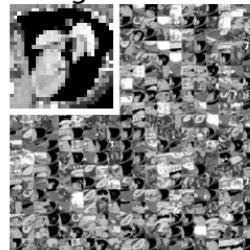
OrganAMNIST



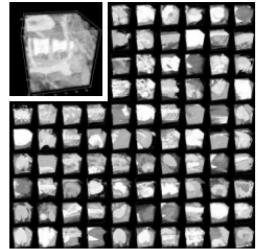
OrganCMNIST



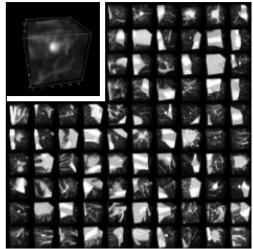
OrganSMNIST



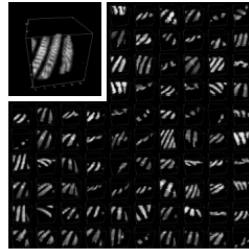
OrganMNIST3D



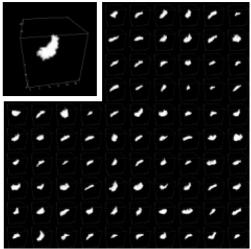
NoduleMNIST3D



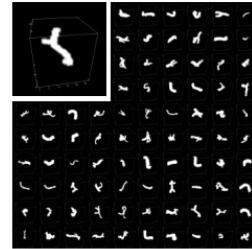
FractureMNIST3D



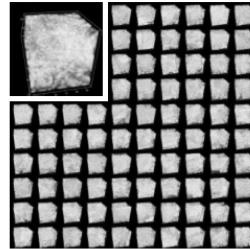
AdrenalMNIST3D



VesselMNIST3D



SynapseMNIST3D



MedMNIST2D

MedMNIST2D	Data Modality	Tasks (# Classes/Labels)	# Samples	# Training / Validation / Test
PathMNIST	Colon Pathology	Multi-Class (9)	107,180	89,996 / 10,004 / 7,180
ChestMNIST	Chest X-Ray	Multi-Label (14) Binary-Class (2)	112,120	78,468 / 11,219 / 22,433
DermaMNIST	Dermatoscope	Multi-Class (7)	10,015	7,007 / 1,003 / 2,005
OCTMNIST	Retinal OCT	Multi-Class (4)	109,309	97,477 / 10,832 / 1,000
PneumoniaMNIST	Chest X-Ray	Binary-Class (2)	5,856	4,708 / 524 / 624
RetinaMNIST	Fundus Camera	Ordinal Regression (5)	1,600	1,080 / 120 / 400
BreastMNIST	Breast Ultrasound	Binary-Class (2)	780	546 / 78 / 156
BloodMNIST	Blood Cell Microscope	Multi-Class (8)	17,092	11,959 / 1,712 / 3,421
TissueMNIST	Kidney Cortex Microscope	Multi-Class (8)	236,386	165,466 / 23,640 / 47,280
OrganAMNIST	Abdominal CT	Multi-Class (11)	58,830	34,561 / 6,491 / 17,778
OrganCMNIST	Abdominal CT	Multi-Class (11)	23,583	12,975 / 2,392 / 8,216
OrganSMNIST	Abdominal CT	Multi-Class (11)	25,211	13,932 / 2,452 / 8,827

MedMNIST3D

MedMNIST3D	Data Modality	Tasks (# Classes/Labels)	# Samples	# Training / Validation / Test
OrganMNIST3D	Abdominal CT	Multi-Class (11)	1,742	971 / 161 / 610
NoduleMNIST3D	Chest CT	Binary-Class (2)	1,633	1,158 / 165 / 310
AdrenalMNIST3D	Shape from Abdominal CT	Binary-Class (2)	1,584	1,188 / 98 / 298
FractureMNIST3D	Chest CT	Multi-Class (3)	1,370	1,027 / 103 / 240
VesselMNIST3D	Shape from Brain MRA	Binary-Class (2)	1,908	1,335 / 191 / 382
SynapseMNIST3D	Electron Microscope	Binary-Class (2)	1,759	1,230 / 177 / 352

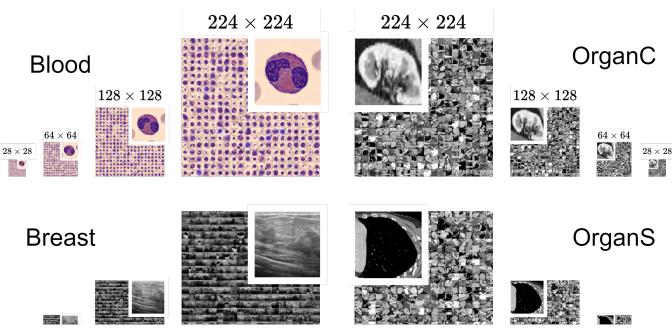
Benchmarking

	PathMNIST		ChestMNIST		DermaMNIST		OCTMNIST		PneumoniaMNIST		RetinaMNIST	
Methods	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC
ResNet-18 (28)	0.983	0.907	0.768	0.947	0.917	0.735	0.943	0.743	0.944	0.854	0.717	0.524
ResNet-18 (224)	0.989	0.909	0.773	0.947	0.920	0.754	0.958	0.763	0.956	0.864	0.710	0.493
ResNet-50 (28)	0.990	0.911	0.769	0.947	0.913	0.735	0.952	0.762	0.948	0.854	0.726	0.528
ResNet-50 (224)	0.989	0.892	0.773	0.948	0.912	0.731	0.958	0.776	0.962	0.884	0.716	0.511
auto-sklearn	0.934	0.716	0.649	0.779	0.902	0.719	0.887	0.601	0.942	0.855	0.690	0.515
AutoKeras	0.959	0.834	0.742	0.937	0.915	0.749	0.955	0.763	0.947	0.878	0.719	0.503
Google AutoML Vision	0.944	0.728	0.778	0.948	0.914	0.768	0.963	0.771	0.991	0.946	0.750	0.531
	BreastMNIST		BloodMNIST		TissueMNIST		OrganAMNIST		OrganCMNIST		OrganSMNIST	
Methods	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC
ResNet-18 (28)	0.901	0.863	0.998	0.958	0.930	0.676	0.997	0.935	0.992	0.900	0.972	0.782
ResNet-18 (224)	0.891	0.833	0.998	0.963	0.933	0.681	0.998	0.951	0.994	0.920	0.974	0.778
ResNet-50 (28)	0.857	0.812	0.997	0.956	0.931	0.680	0.997	0.935	0.992	0.905	0.972	0.770
ResNet-50 (224)	0.866	0.842	0.997	0.950	0.932	0.680	0.998	0.947	0.993	0.911	0.975	0.785
auto-sklearn	0.836	0.803	0.984	0.878	0.828	0.532	0.963	0.762	0.976	0.829	0.945	0.672
AutoKeras	0.871	0.831	0.998	0.961	0.941	0.703	0.994	0.905	0.990	0.879	0.974	0.813
Google AutoML Vision	0.919	0.861	0.998	0.966	0.924	0.673	0.990	0.886	0.988	0.877	0.964	0.749

Benchmarking Performance on MedMNIST2D.

Methods	OrganMNIST3D		NoduleMNIST3D		FractureMNIST3D		AdrenalMNIST3D		VesselMNIST3D		SynapseMNIST3D	
	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC	AUC	ACC
ResNet-18 + 2.5D	0.977	0.788	0.838	0.835	0.587	0.451	0.718	0.772	0.748	0.846	0.634	0.696
ResNet-18 + 3D	0.996	0.907	0.863	0.844	0.712	0.508	0.827	0.721	0.874	0.877	0.820	0.745
ResNet-18 + ACS	0.994	0.900	0.873	0.847	0.714	0.497	0.839	0.754	0.930	0.928	0.705	0.722
ResNet-50 + 2.5D	0.974	0.769	0.835	0.848	0.552	0.397	0.732	0.763	0.751	0.877	0.669	0.735
ResNet-50 + 3D	0.994	0.883	0.875	0.847	0.725	0.494	0.828	0.745	0.907	0.918	0.851	0.795
ResNet-50 + ACS	0.994	0.889	0.886	0.841	0.750	0.517	0.828	0.758	0.912	0.858	0.719	0.709
auto-sklearn	0.977	0.814	0.914	0.874	0.628	0.453	0.828	0.802	0.910	0.915	0.631	0.730
AutoKeras	0.979	0.804	0.844	0.834	0.642	0.458	0.804	0.705	0.773	0.894	0.538	0.724

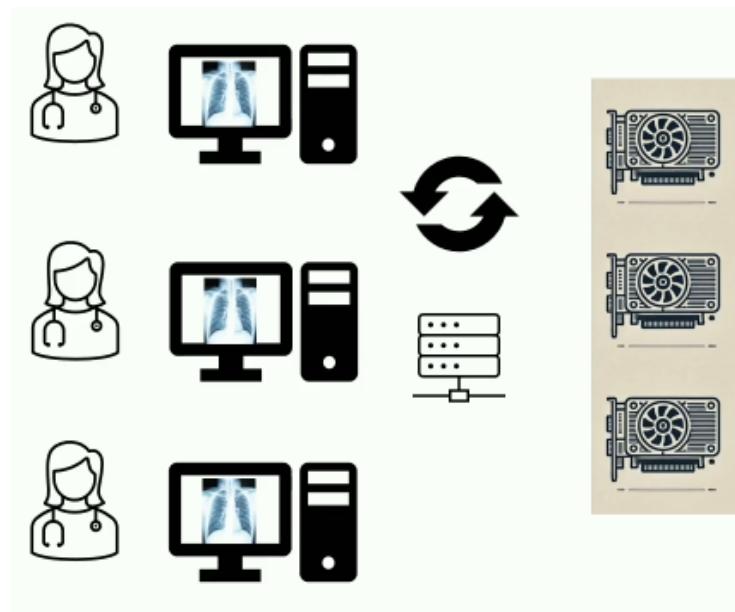
Benchmarking Performance on MedMNIST3D.



Methods	Accuracy (ACC)				Area Under the ROC Curve (AUC)			
	28 × 28	64 × 64	128 × 128	224 × 224	28 × 28	64 × 64	128 × 128	224 × 224
END-TO-END								
VGG16	82.34±0.88	85.33±1.16	86.64±0.82	86.70±0.93	92.66±0.41	94.24±0.28	95.16±0.27	95.30±0.22
AlexNet	78.92±0.81	82.94±0.78	85.04±0.74	85.74±0.64	91.14±0.43	92.72±0.33	94.29±0.30	94.90±0.23
ResNet-18	79.66±0.74	83.42±0.65	85.73±0.66	86.22±0.58	90.92±0.28	92.49±0.50	93.91±0.27	94.51±0.24
DenseNet-121	80.32±0.93	84.62±0.80	87.13±0.56	87.11±0.64	91.75±0.55	93.59±0.23	94.57±0.21	95.03±0.23
EfficientNet-B4	73.18±1.61	79.37±1.10	82.52±0.79	82.44±1.11	87.04±0.82	90.07±0.65	91.89±0.39	91.64±0.73
ViT-B/16	78.23±0.88	83.17±0.92	84.94±0.93	86.06±0.92	90.54±0.47	92.53±0.69	93.25±0.35	94.08±0.38
CLIP ViT-B/16	76.73±0.80	80.39±0.99	82.33±1.02	82.75±1.01	89.22±1.11	90.91±0.51	91.51±0.31	91.83±0.58

Future directions for FL in Healthcare

Real time updates to clinical models



Today basically all imaging AI models deployed and used clinically are static.

Though there's a need to work with the FDA, federated learning may soon facilitate real time refinement and updates of deployed models.

