

AI in Healthcare: The Example of Imaging

Thursday, December 12, 2024

Curtis P. Langlotz, MD, PhD

Professor of Radiology, Medicine, and Biomedical Data Science

Senior Associate Vice Provost for Research

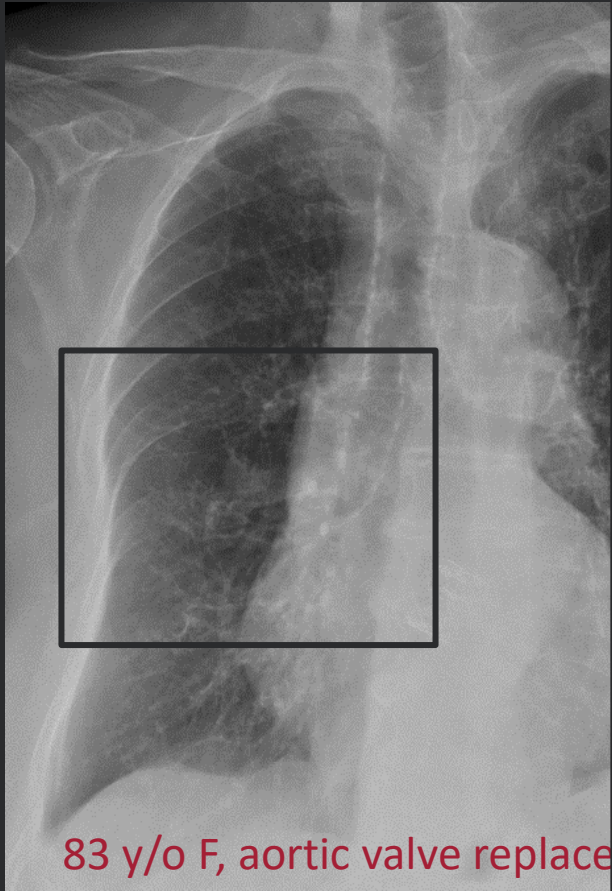
Director, Center for Artificial Intelligence in Medicine & Imaging

Senior Fellow, Institute for Human-Centered Artificial Intelligence

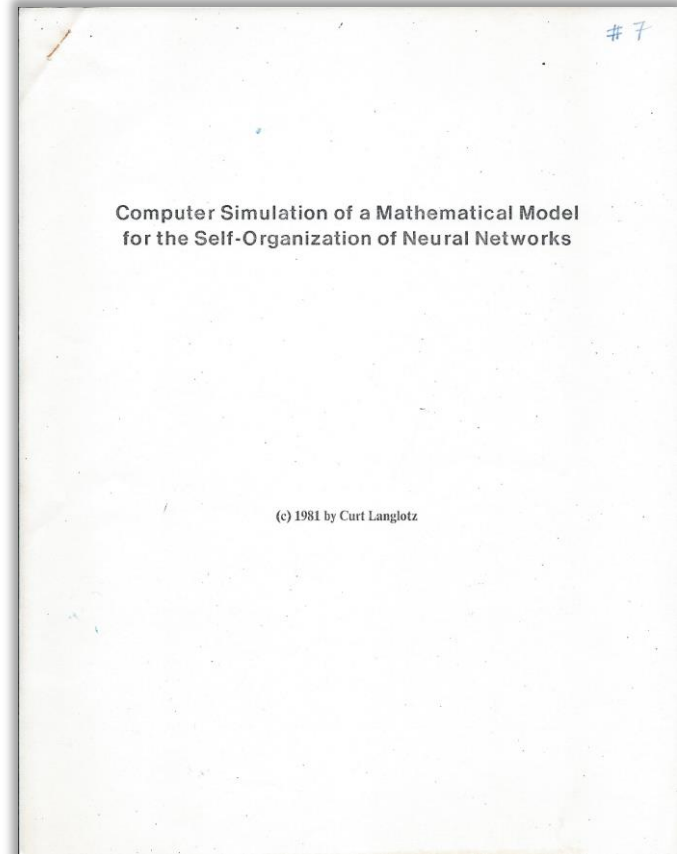
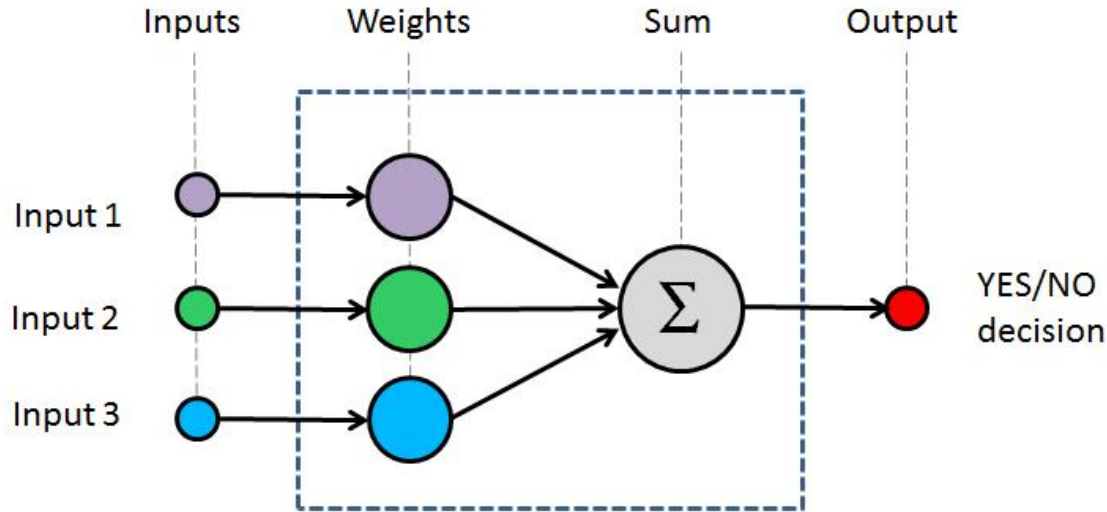
@curtlanglotz



Stanford | MEDICINE



Perceptron, 1960-70s



Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

Garden spider, *Aranea diademata*

A spider common in European gardens

1372 pictures

63.48% Popularity Percentile

Wordnet IDs

Numbers in brackets: (the number of synsets in the subtree)

- ImageNet 2011 Fall Release (32326)
 - plant, flora, plant life (4486)
 - geological formation, formation (175)
 - natural object (1112)
 - sport, athletics (176)
 - artifact, artefact (10504)
 - fungus (308)
 - person, individual, someone, somebody, mortal, soul (6978)
 - animal, animate being, beast, brute, creature, fauna (3998)
 - invertebrate (766)
 - arthropod (579)
 - trilobite (0)
 - arachnid, arachnoid (41)
 - harvestman, daddy longlegs, Phalangium opilio (0)
 - scorpion (0)
 - false scorpion, pseudoscorpion (1)
 - book scorpion, Chelifer cancroides (0)
 - whip-scorpion, whip scorpion (1)
 - vinegarroon, Mastigoproctus giganteus (0)
 - spider (10)
 - orb-weaving spider (0)
 - black and gold garden spider, Argiope aurantia (0)
 - barn spider, Araneus cavaticus (0)
 - garden spider, *Aranea diademata* (0)
 - comb-footed spider, theridiid (0)
 - black widow, Latrodectus mactans (0)

Treemap Visualization

Images of the Synset

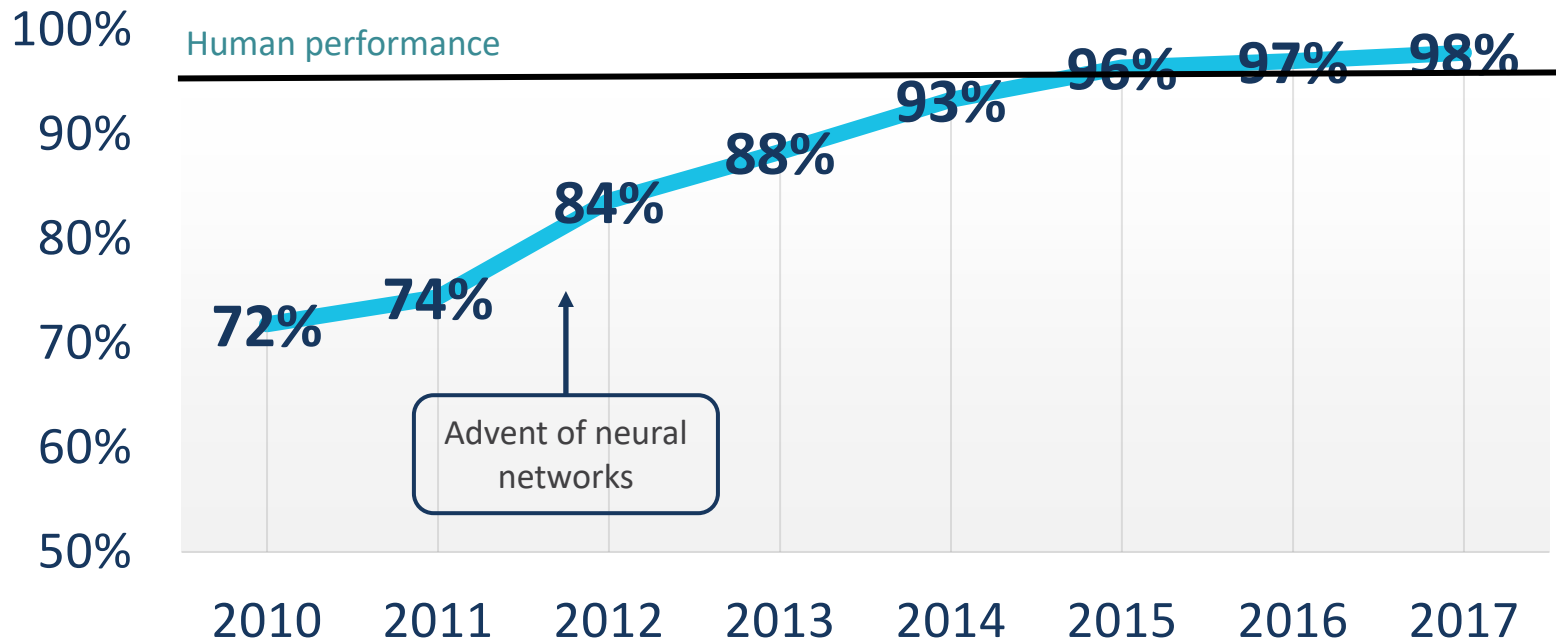
Downloads

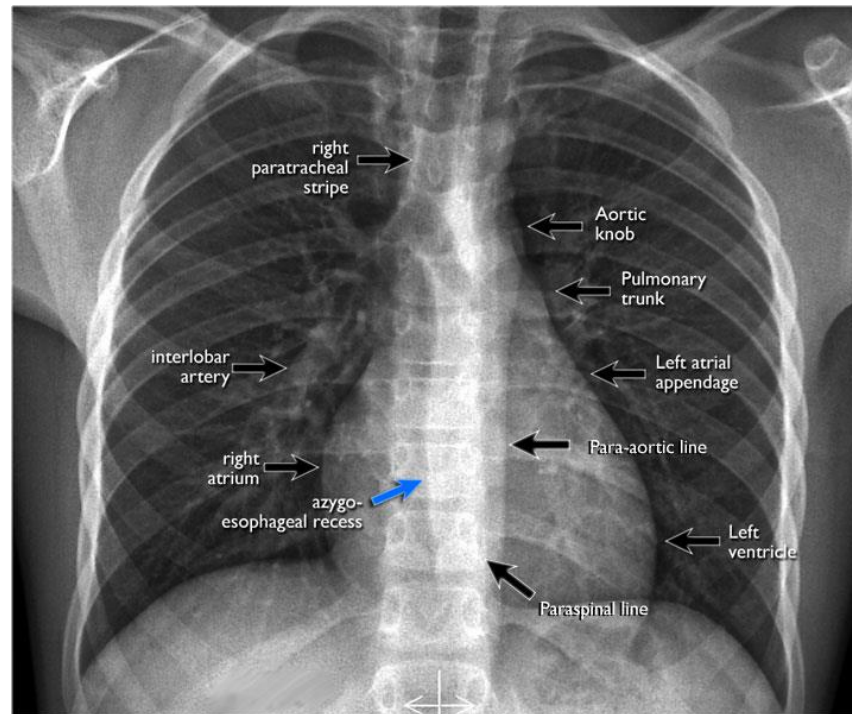
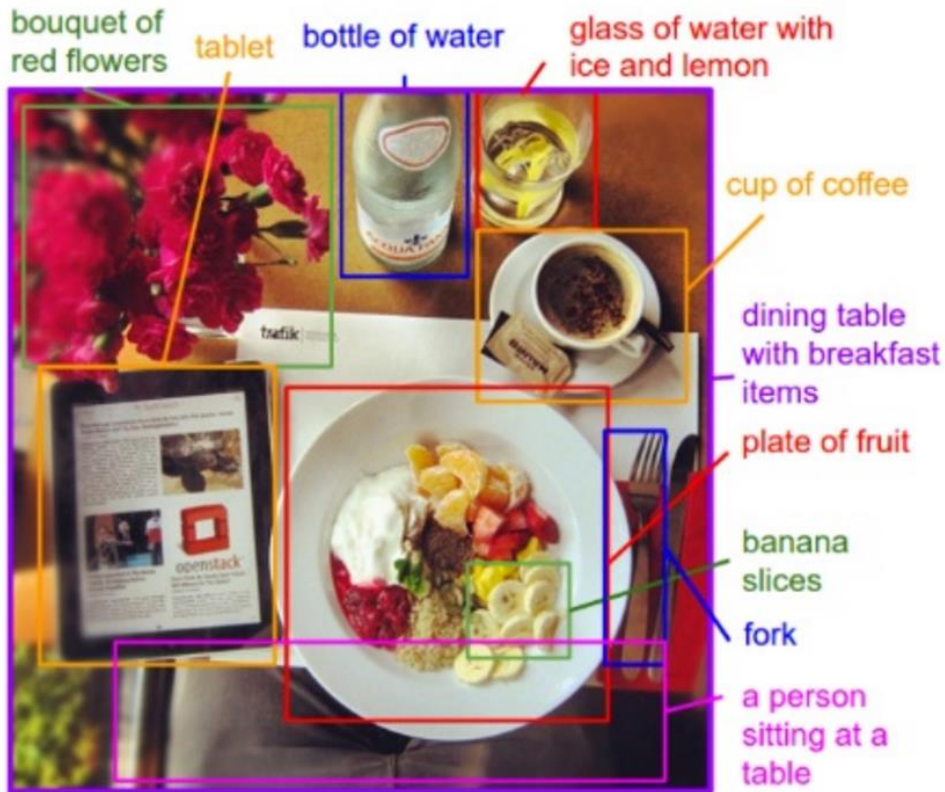


- 14 million images
- 21,841 distinct labels:
 - 856 types of bird
 - 993 types of tree
 - 157 musical instruments

- Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis. 2015;115(3):211-252.
- <https://www.economist.com/news/special-report/21700756-artificial-intelligence-boom-based-old-idea-modern-twist-not>
- <http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/>

ImageNet Visual Recognition Accuracy



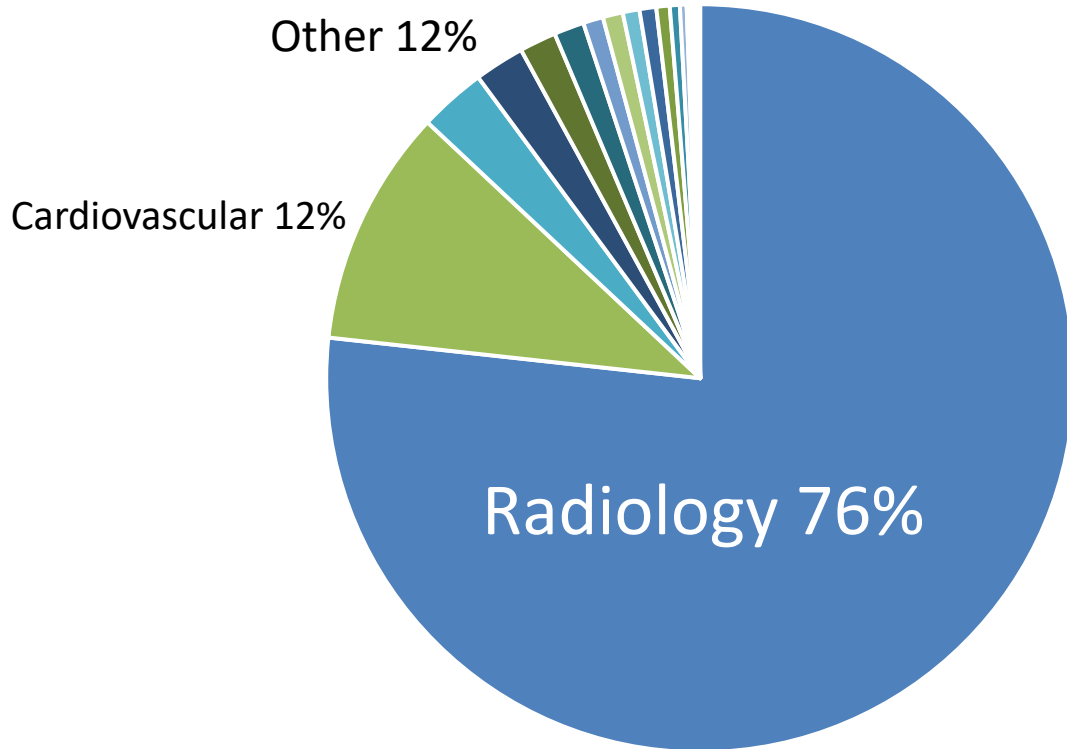


Karpathy, Andrej & Li, Fei Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions, CVPR, 2015

<http://www.radiologyassistant.nl/>

The Health AI Industry

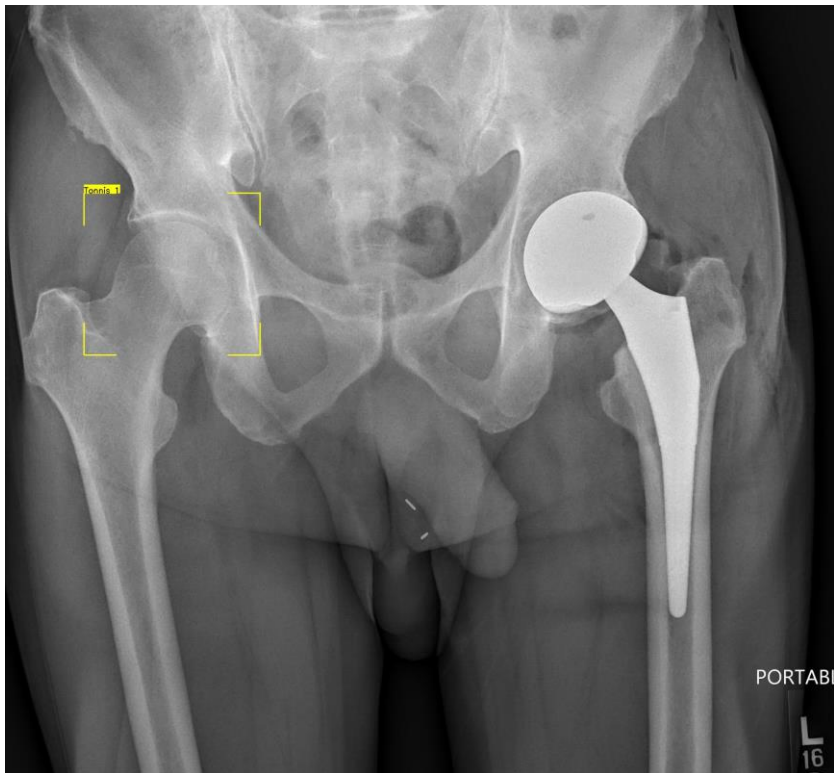
FDA Cleared Algorithms (N=950)



Medical Specialty	FDA devices
Radiology	723
Cardiovascular	98
Neurology	34
Hematology	18
GI/Urology	14
Ophthalmology	10
Anesthesiology	9
Chemistry	8
Pathology	8
Surgery	6
Microbiology	5
Orthopedics	5
Hospital	4
Dental	3
Ear/Nose/Throat	2
OB/Gyn	1
Immunology	1
Physical Med	1
Total	950



Automatic Radiology Report Generation



Comparison: XR 08/08/2022, XR 12/02/2021, MR 11/19/2019

Findings:

1 view of the pelvis was acquired. No acute fracture or dislocation.

RIGHT hip: Tonnis category 1 - Mild osteoarthrosis.

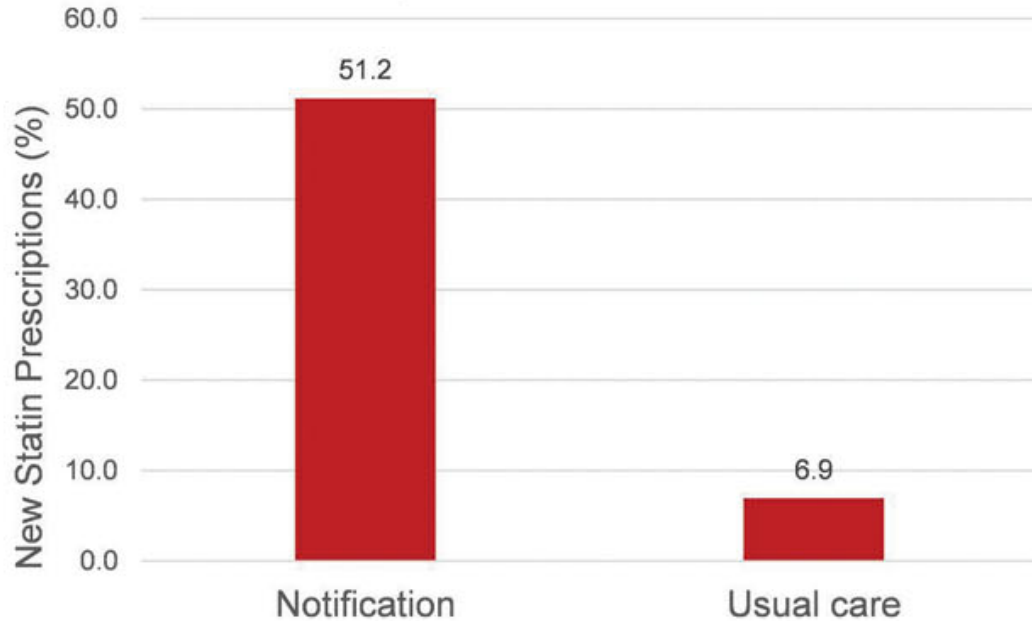
LEFT hip: Left hip arthroplasty is in position. Hardware is intact. Alignment is anatomic. There is no peri-hardware lucency or fracture. Postsurgical appearance of the regional soft tissues.

Impression:

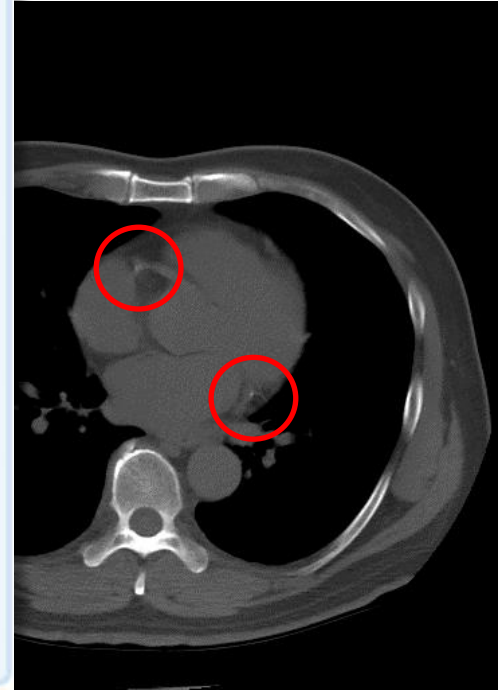
No acute fracture or malalignment.

Opportunistic Screening for Coronary Artery Disease

Primary Outcome: Statin prescription at 6 months post randomization

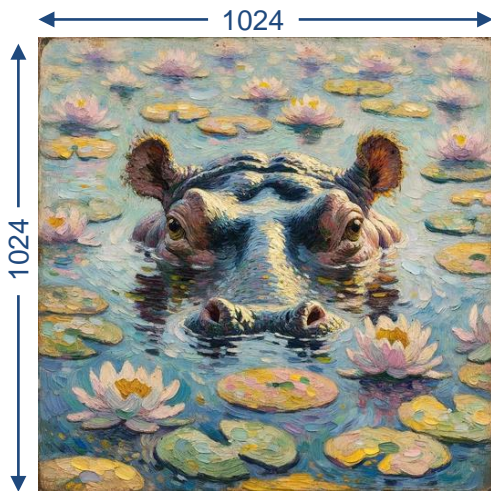


Annual CT: 20M/yr in US

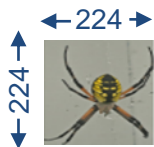


Health Data is Different

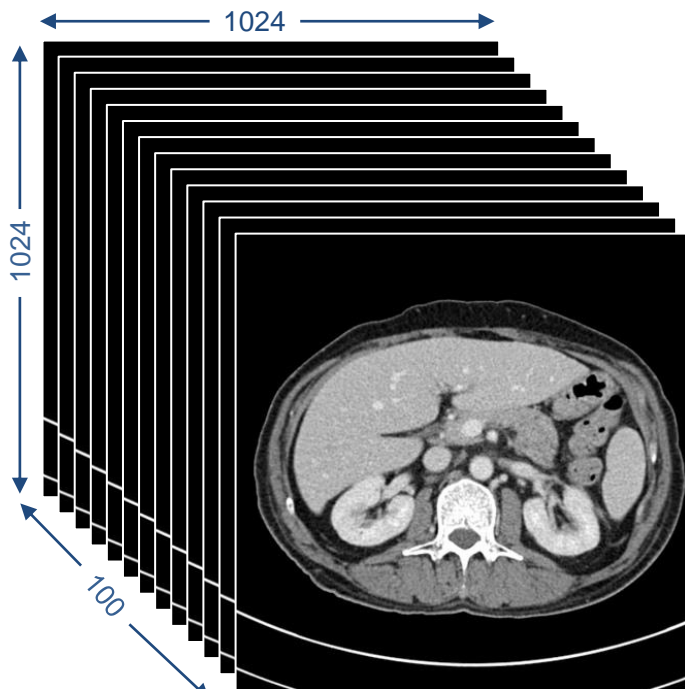
DALL-E 3



ImageNet

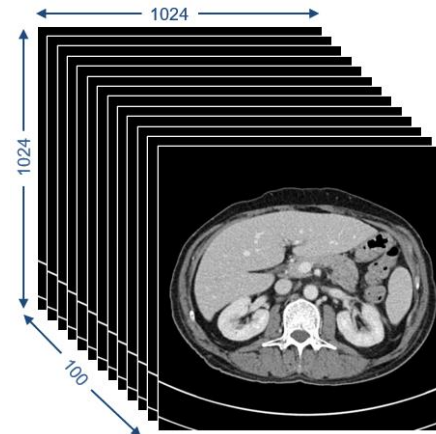


Abdomen CT

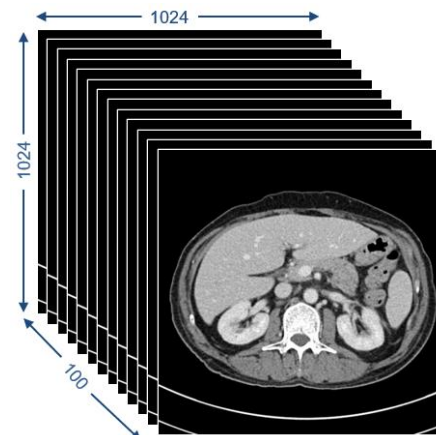


Before IV contrast

“Growth since 6 months ago?”

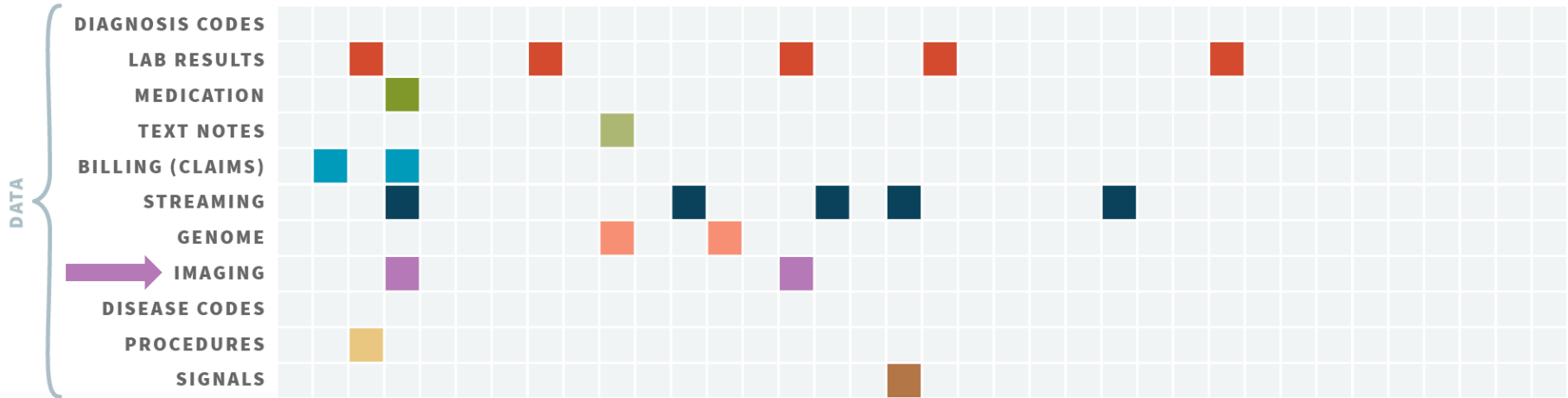


During IV contrast



After IV contrast

PATIENT TIMELINE





Organizational Factors in Clinical Data Sharing for Artificial Intelligence in Health Care

Alaa Youssef, PhD; Madelena Y. Ng, DrPH, MPH; Jin Long, PhD; Tina Hernandez-Boussard, PhD; Nigam Shah, MBBS, PhD; Adam Miner, PsyD, MS; David Larson, MD, MBA; Curtis P. Langlotz, MD, PhD

Drivers of readiness for data sharing:

- Motivation:
 - alignment of an organization's values with data-sharing priorities
 - modulated by extrinsic incentives for financial or reputational gains
- Capabilities:
 - infrastructure
 - people
 - expertise
 - access to data



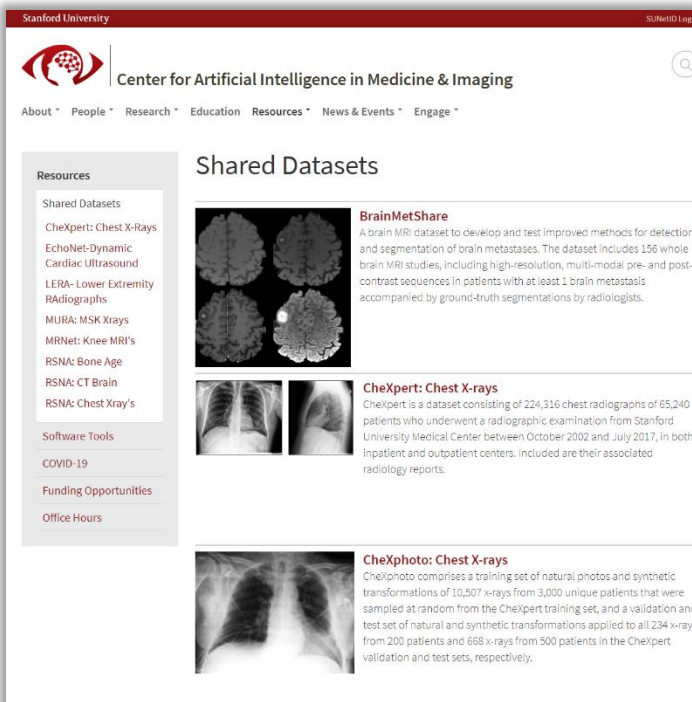
Perceptions of Data Set Experts on Important Characteristics of Health Data Sets Ready for Machine Learning A Qualitative Study

Madelena Y. Ng, DrPH, MPH; Alaa Youssef, PhD; Adam S. Miner, PsyD, MS; Daniela Sarellano, MS; Jin Long, PhD; David B. Larson, MD, MBA; Tina Hernandez-Boussard, PhD, MPH, MS; Curtis P. Langlotz, MD, PhD

Characteristics of AI-ready data sets:

- accuracy
- completeness
- consistency
- fitness
- availability
- data quality standards
- documentation
- team science

Publicly-Released AI-Ready Datasets



Stanford University SUNetID Log

Center for Artificial Intelligence in Medicine & Imaging

About * People * Research * Education Resources * News & Events * Engage *

Resources

Shared Datasets

- CheXpert: Chest X-Rays
- EchoNet-Dynamic Cardiac Ultrasound
- LERA- Lower Extremity Radiographs
- MURA: MSK Xrays
- MRNet: Knee MRI's
- RSNA: Bone Age
- RSNA: CT Brain
- RSNA: Chest Xray's

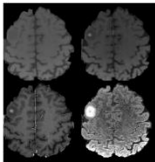
Software Tools

- COVID-19
- Funding Opportunities
- Office Hours

Shared Datasets


BrainMetShare

A brain MRI dataset to develop and test improved methods for detection and segmentation of brain metastases. The dataset includes 156 whole brain MRI studies, including high-resolution, multi-modal pre- and post-contrast sequences in patients with at least 1 brain metastasis accompanied by ground-truth segmentations by radiologists.




CheXpert: Chest X-rays

CheXpert is a dataset consisting of 224,316 chest radiographs of 65,240 patients who underwent a radiographic examination at Stanford University Medical Center between October 2002 and July 2017, in both inpatient and outpatient centers. Included are their associated radiology reports.




CheXphoto: Chest X-rays

CheXphoto comprises a training set of natural photos and synthetic transformations of 10,507 x-rays from 3,000 unique patients that were sampled at random from the CheXpert training set, and a validation and test set of natural and synthetic transformations applied to all 234 x-rays from 200 patients and 668 x-rays from 500 patients in the CheXpert validation and test sets, respectively.




EchoNet-Dynamic Cardiac Ultrasound

EchoNet-Dynamic is a dataset of over 10k echocardiogram, or cardiac ultrasound, videos from unique patients at Stanford University Medical Center. Each apical-4-chamber video is accompanied by an estimated ejection fraction, end-systolic volume, end-diastolic volume, and tracings of the left ventricle performed by an advanced cardiac sonographer and reviewed by an imaging cardiologist.



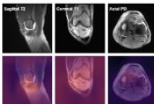
LERA- Lower Extremity Radiographs

In this retrospective, HIPAA-compliant, IRB-approved study, we collected data from 182 patients who underwent a radiographic examination at the Stanford University Medical Center between 2003 and 2014. The dataset consists of images of the foot, knee, ankle, or hip associated with each patient.




MRNet: Knee MRI's

The MRNet dataset consists of 1,370 knee MRI exams performed at Stanford University Medical Center. The dataset contains 1,104 (80.6%) abnormal exams, with 319 (23.3%) ACL tears and 508 (37.1%) meniscal tears; labels were obtained through manual extraction from clinical reports.




MURA: MSK Xrays

MURA (musculoskeletal radiographs) is a large dataset of bone x-rays from the Stanford University Medical Center.



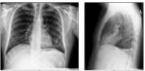
RSNA: Bone Age

From the RSNA AI Challenge 2017, a dataset of bone age x-rays from Stanford University, the University of Colorado and the University of California - Los Angeles.




RSNA: Chest Xrays

From the RSNA AI Challenge 2018, a dataset labeled chest x-rays from the the National Institutes of Health (NIH).




RSNA: CT Brain


Four research institutions provided large volumes of de-identified CT studies that were assembled to create the RSNA AI 2019 challenge dataset: Stanford University, Thomas Jefferson University, Unity Health Toronto and Universidade Federal de São Paulo (UNIFESP). The American Society of Neuroradiology (ASNR) organized a cadre of more than 60 volunteers to label over 25,000 exams for the challenge dataset.



Stanford AIMI Center
1701 Page Mill Road
Stanford, CA 94304
aimi@stanford.edu
Campus Map

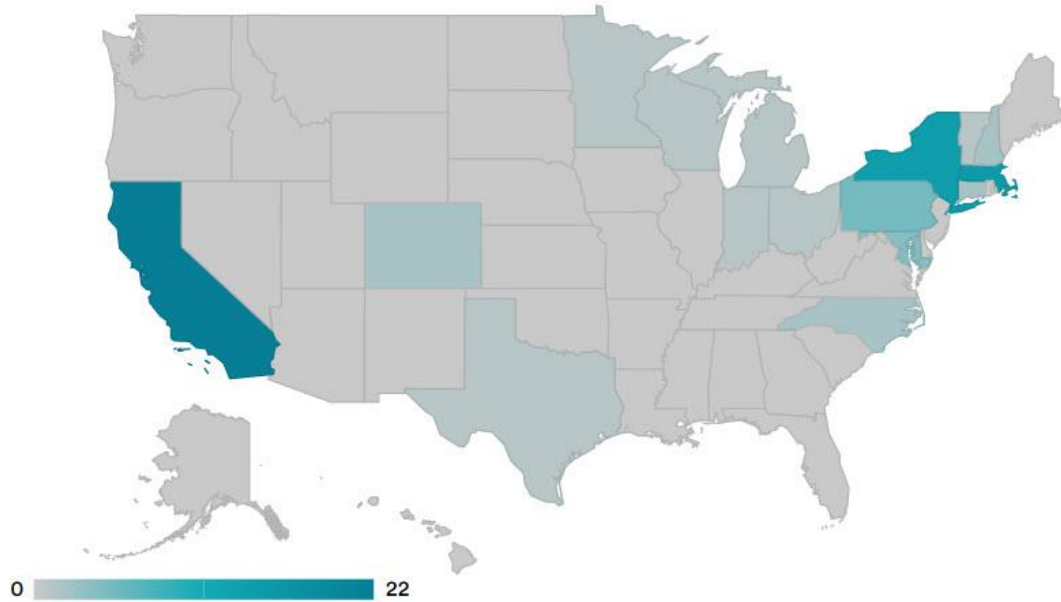
Connect


Related Links
[Stanford University](#)
[Research at Stanford](#)
[Stanford News](#)

Stanford AIMI 

Stanford University
Stanford Home | Maps & Directions | Search Stanford | Emergency Info
Terms of Use | Privacy | Copyright | Trademarks | Non-Discrimination | Accessibility
© Stanford University, Stanford, California 94305

Geographic Distribution of Cohorts to Train Deep Learning Algorithms



REBECCA ROBBINS/STAT

SOURCE: "GEOGRAPHIC DISTRIBUTION OF US COHORTS USED TO TRAIN DEEP LEARNING ALGORITHMS," JAMA 2020.

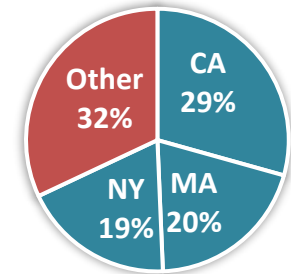
STAT

Table. US Patient Cohorts Used for Training Clinical Machine Learning Algorithms, by State^a

States	No. of studies
California	22
Massachusetts	15
New York	14
Pennsylvania	5
Maryland	4
Colorado	2
Connecticut	2
New Hampshire	2
North Carolina	2
Indiana	1
Michigan	1
Minnesota	1
Ohio	1
Texas	1
Vermont	1
Wisconsin	1

^a Fifty-six studies used 1 or more geographically identifiable US patient cohort in the training of their clinical machine learning algorithm. Thirty-four states were not represented in geographically identifiable cohorts: Alabama, Alaska, Arizona, Arkansas, Delaware, Florida, Georgia, Hawaii, Idaho, Illinois, Iowa, Kansas, Kentucky, Louisiana, Maine, Mississippi, Missouri, Montana, Nebraska, Nevada, New Jersey, New Mexico, North Dakota, Oklahoma, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Virginia, Washington, West Virginia, and Wyoming.

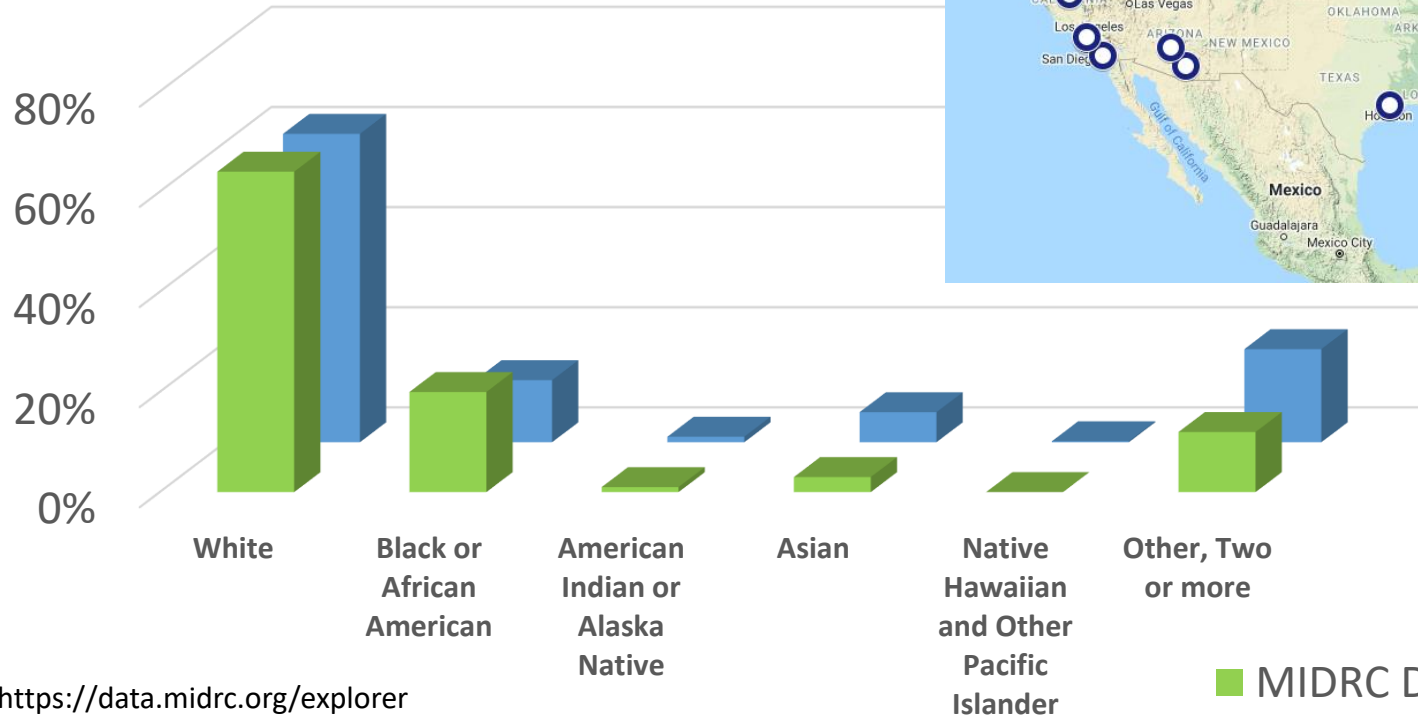
JAMA September 22/29, 2020 Volume 324, Number 12



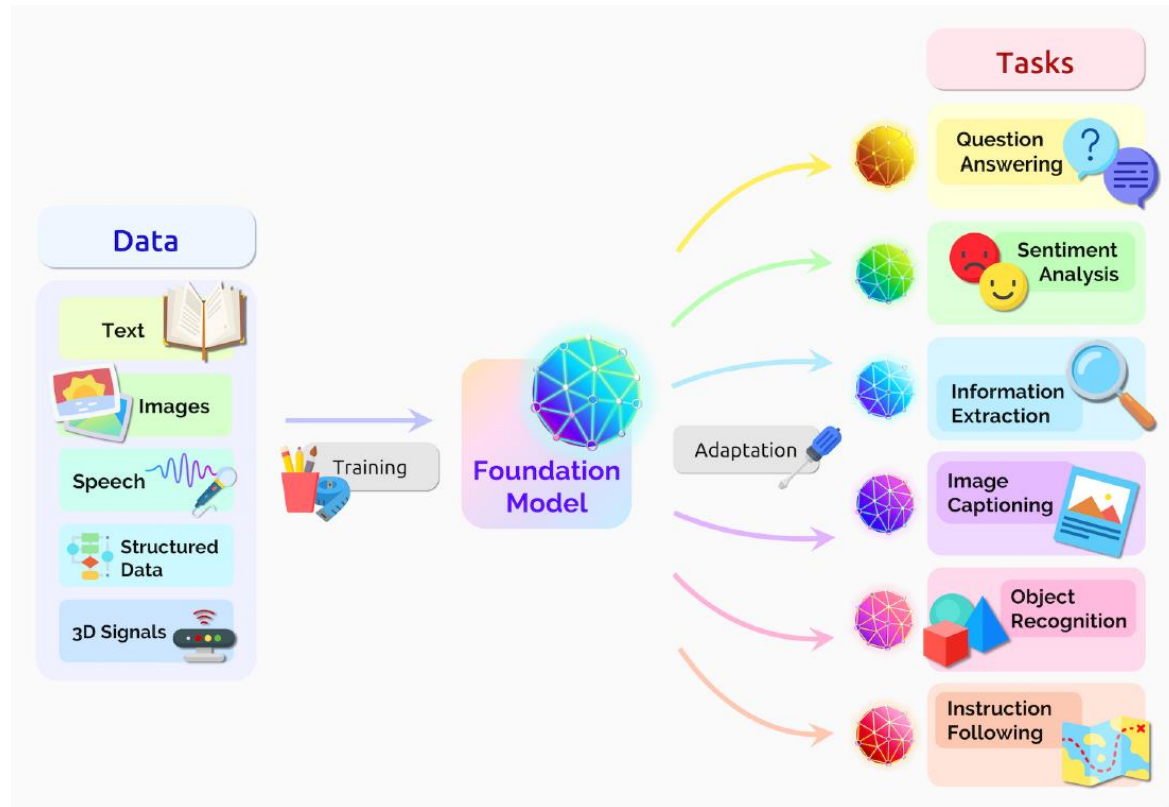


MIDRC

MEDICAL IMAGING AND DATA RESOURCE CENTER.



Foundation Models



ChatGPT=Generative Pre-trained Transformer

Pre-Trained Using Self-Supervised Learning

A 36-year-old woman presents to the clinic with a 2-week history of persistent dry cough, shortness of [REDACTED], and mild chest pain. She also reports occasional low-grade [REDACTED] and night sweats. She is a non-smoker and works as an elementary school [REDACTED]. Her medical history is unremarkable, and she takes no medications. On examination, she appears [REDACTED]. Lung auscultation reveals decreased breath sounds and dullness to percussion in the left lower lung field.

ChatGPT

Trained to be great at dialogue.

Prompt: "What is the meaning of life?"



Response 1



Response 2



Response 3



Response 4



Human raters

Quality and Empathy of Chatbot Explanations

Question about going blind following bleach being splashed into an eye, resulting in irritated and dry eye:

Physician response

Sounds like you will be fine. You should flush the eye anytime you get a chemical or foreign body in the eye. You can also contact Poison Control 1-800-222-1222.

Chatbot response^c

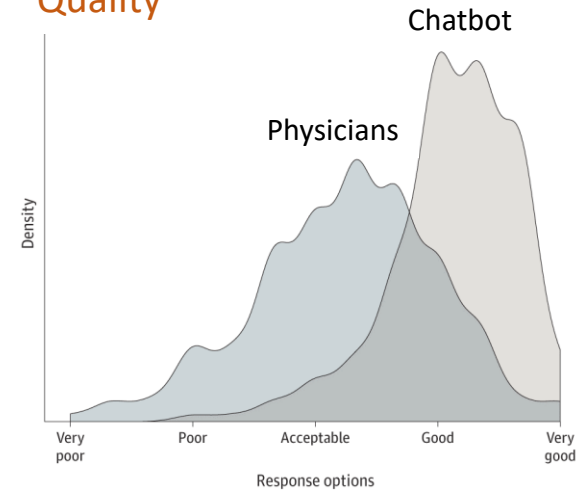
I'm sorry to hear that you got bleach splashed in your eye. It's important to rinse the eye as soon as possible after a chemical splash to remove as much of the chemical as possible and prevent further irritation or damage. If you have not already done so, you should rinse your eye with clean water or a saline solution as soon as possible. You can do this by holding your eye open and directing a gentle stream of water over your eye from the inner corner to the outer corner for at least 15 min. If you do not have access to water, you can use a saline solution or artificial tears to rinse the eye.

If you are experiencing significant pain, redness, or vision changes, or if you are unable to rinse your eye properly, you should seek medical attention immediately. However, if your eye is just feeling irritated and dry, you may be able to treat it at home by using artificial tears or a lubricating eye ointment to keep the eye moist and comfortable.

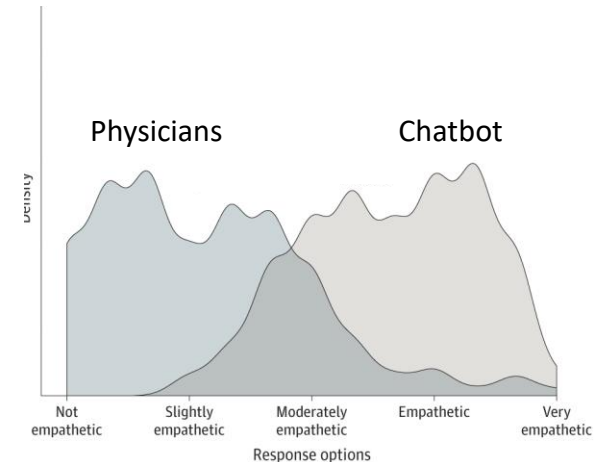
It is unlikely that you will go blind from getting bleach splashed in your eye, but it is important to take care of the eye and seek medical attention if necessary to prevent further irritation or damage.

Ayers, J. W., et al. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine*. <https://doi.org/10.1001/jamainternmed.2023.1838>

Quality



Empathy



Foundation Models in Medicine: 3 Approaches



“Small language models” fit for purpose

Highest accuracy



Generalist foundation models “out of the box”

Democratization

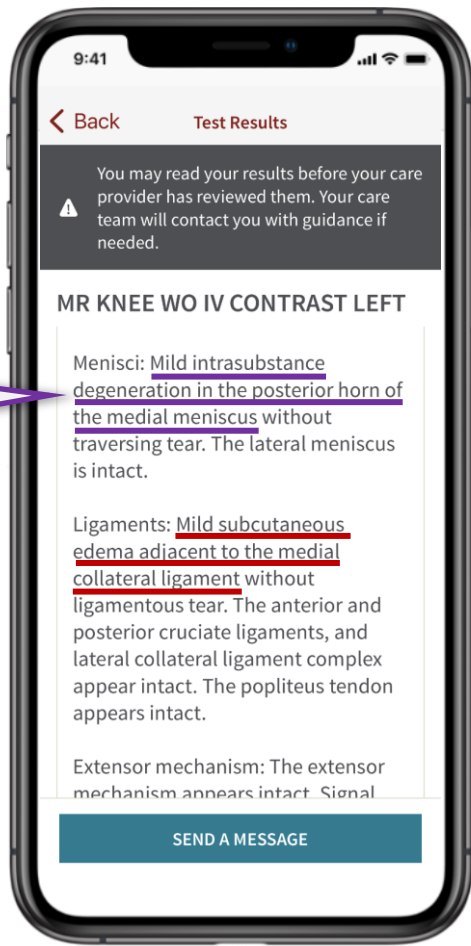
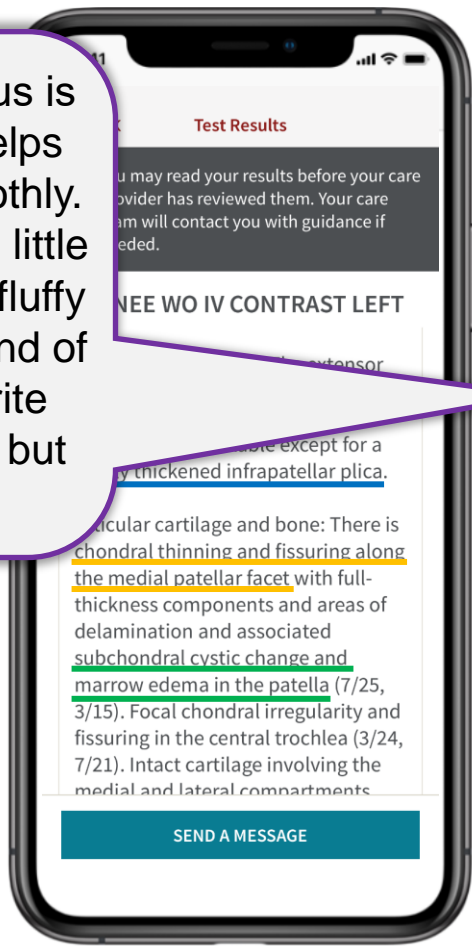


“Build our own” medical foundation models

“ChatGPT moment”

Patients Have Questions About Their Test Results...

“The medical meniscus is like a cushion that helps your knee move smoothly. Over time, it can get a little bit worn out or not as fluffy as it used to be. It's kind of like when your favorite pillow gets a little flat but still works fine.”



Uses of Large Language Models in Medicine

Explanation for patients

Doctor-patient communication

Visit documentation

Record summarization

Decision support*

Meta AI



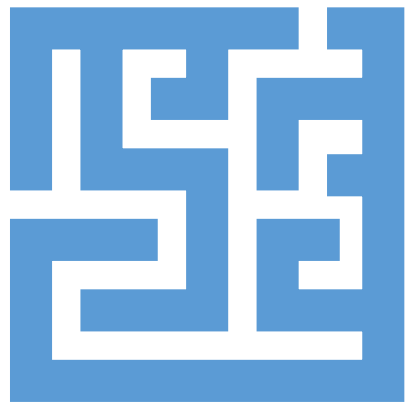
LLaVA-Med



OpenAI
ChatGPT



Claude



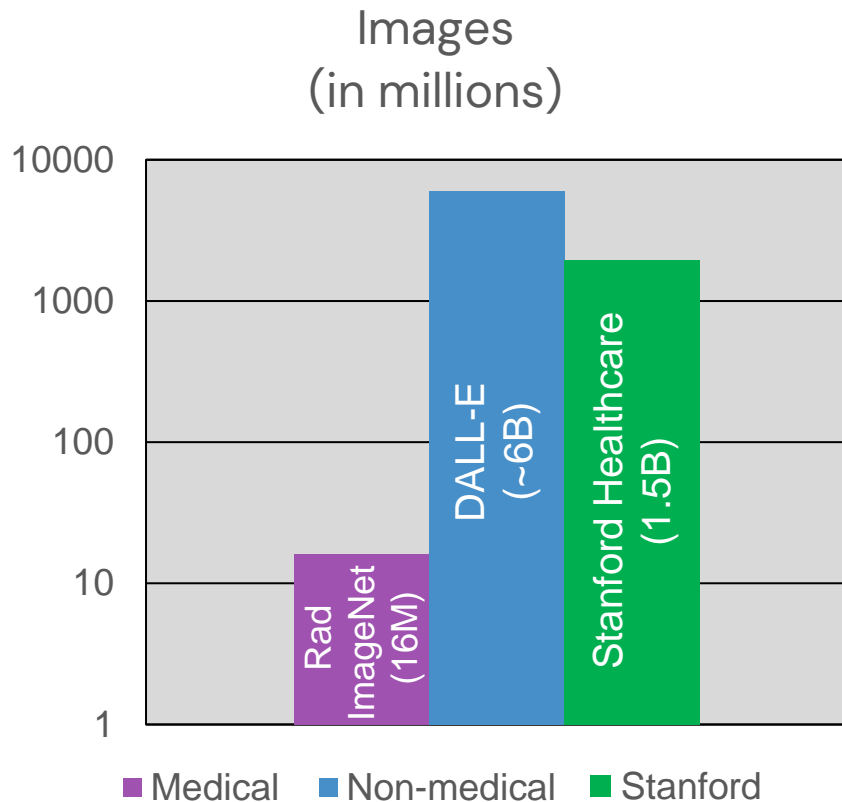
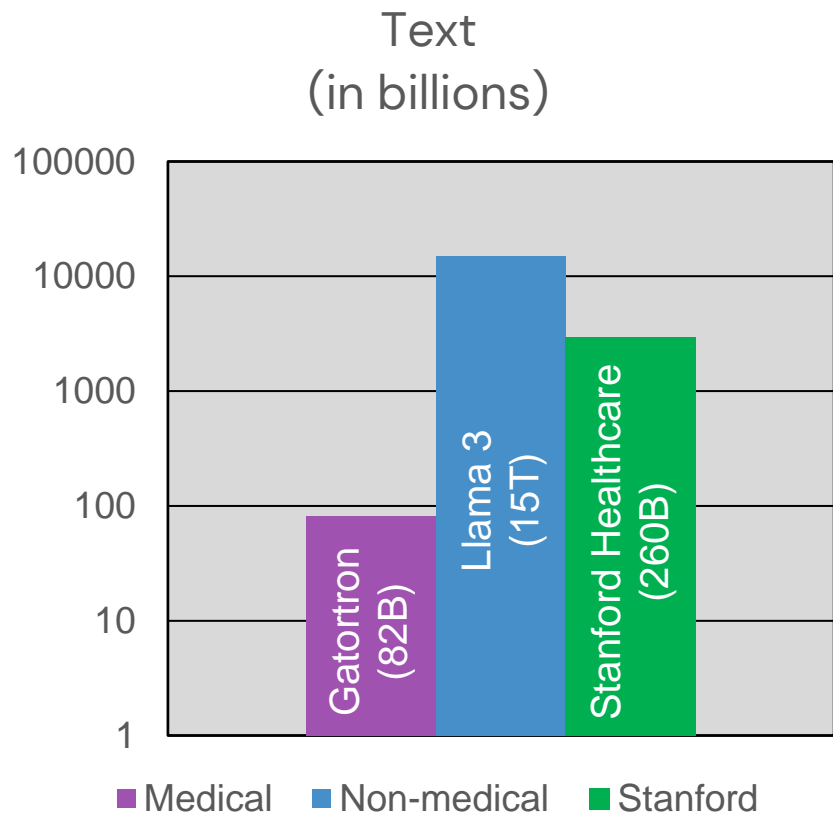
Challenges of Large Language Models

- Inherent limitations of training data
- Hallucination, confabulation
- Planning
- Mathematics

Retrieval-augmented generation (RAG)

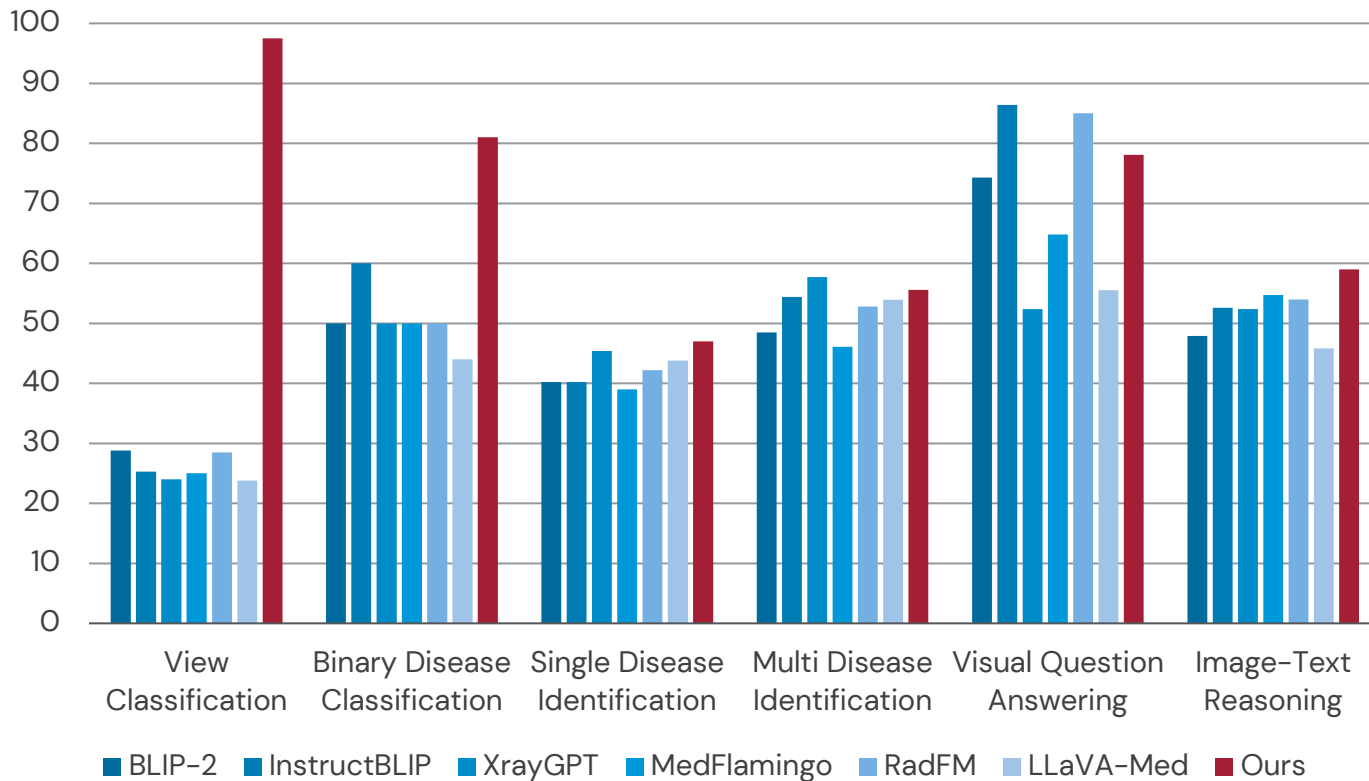
Evaluation

Foundation Models in Medicine: Training Dataset Size



Performance of chest x-ray foundation model on multiple benchmarks

Accuracy (%)



Regulatory Challenges for Foundation Models



Defining scope of use



Open-ended inputs and outputs



Foundation models not controlled by device manufacturer

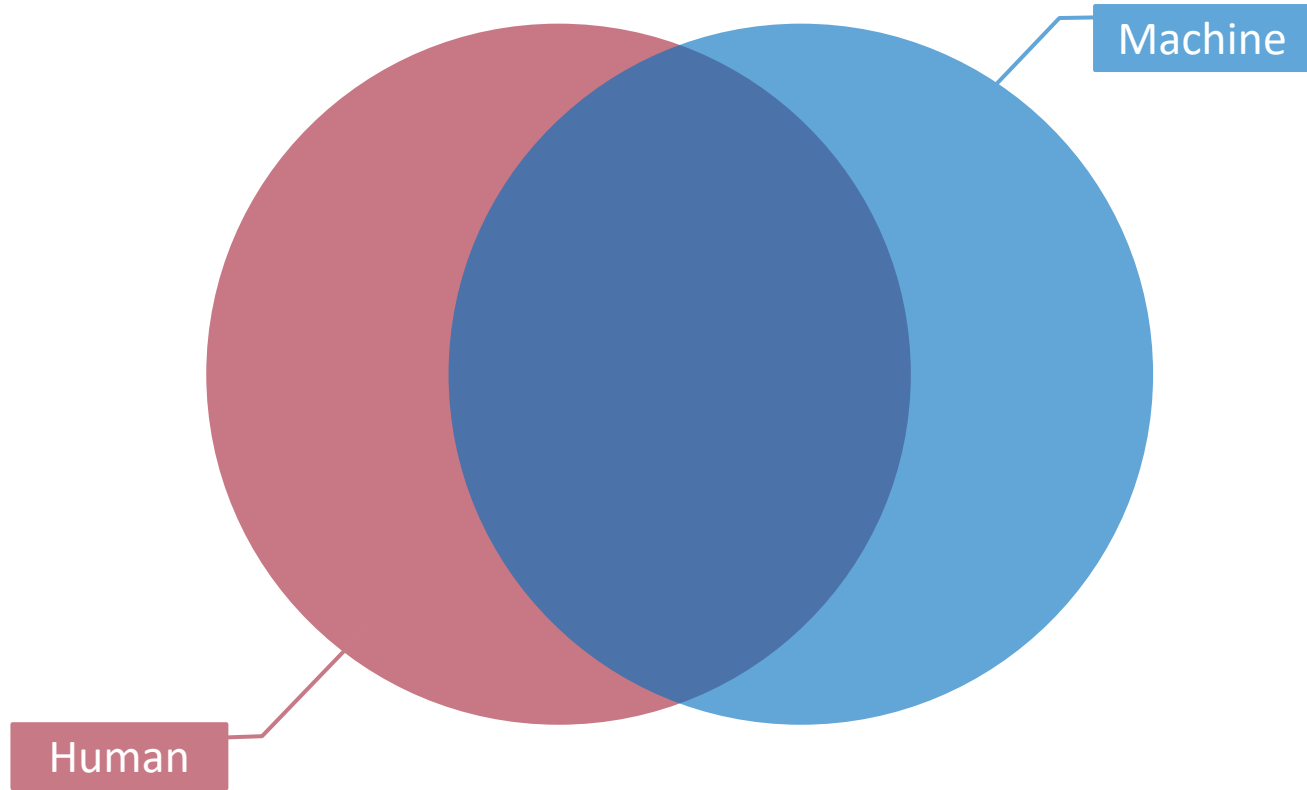


Oversight of an adaptive system

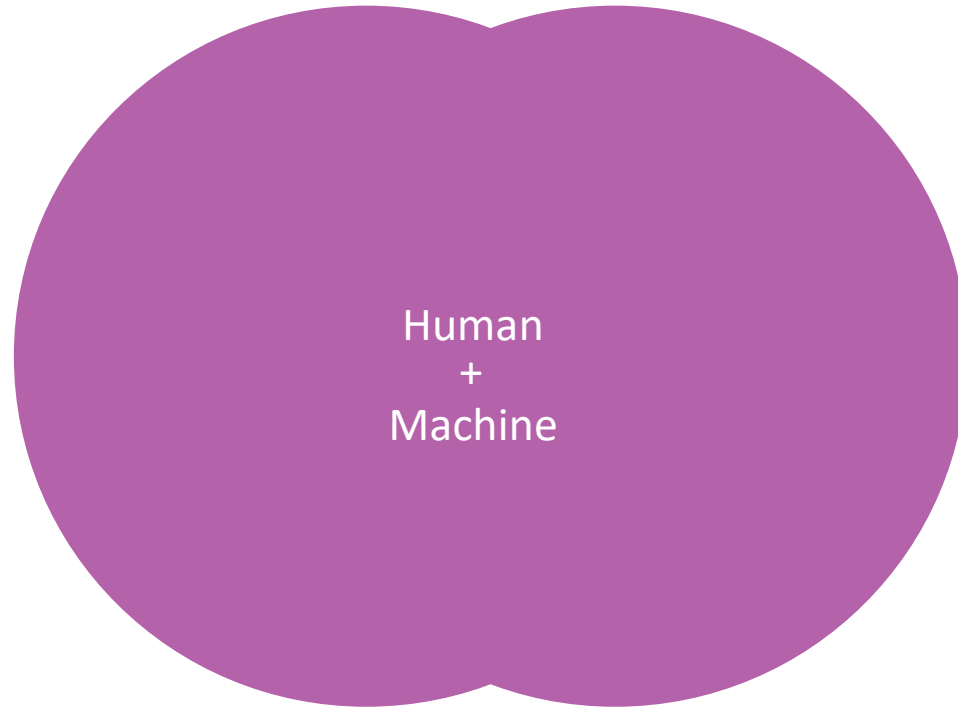


Evaluation and monitoring in the wild

Complementary Expertise of Humans and Machines

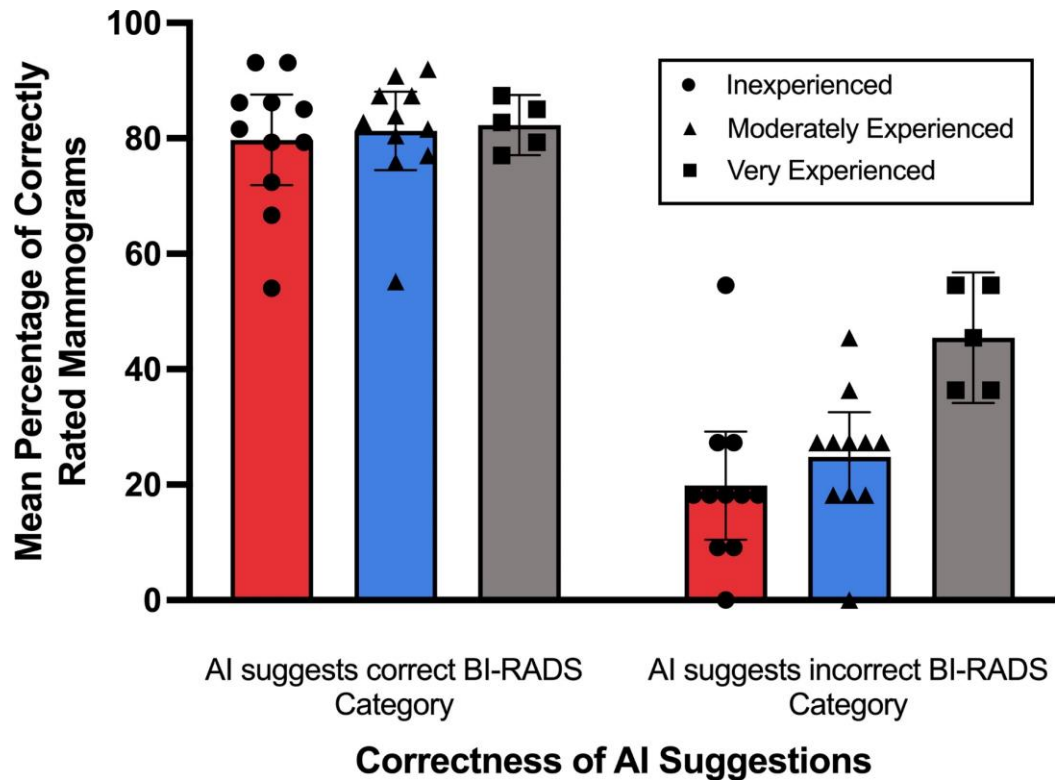


Complementary Expertise of Humans and Machines



Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance

Thomas Dratsch, MD* • Xue Chen, MD* • Mohammad Rezaade Mehrizi, PhD • Roman Kloeckner, MD •
Aline Mähringer-Kunz, MD • Michael Piisken, MD • Bettina Baeflter, MD • Stephanie Sauer, MD •
David Maintz, MD • Daniel Pinto dos Santos, MD

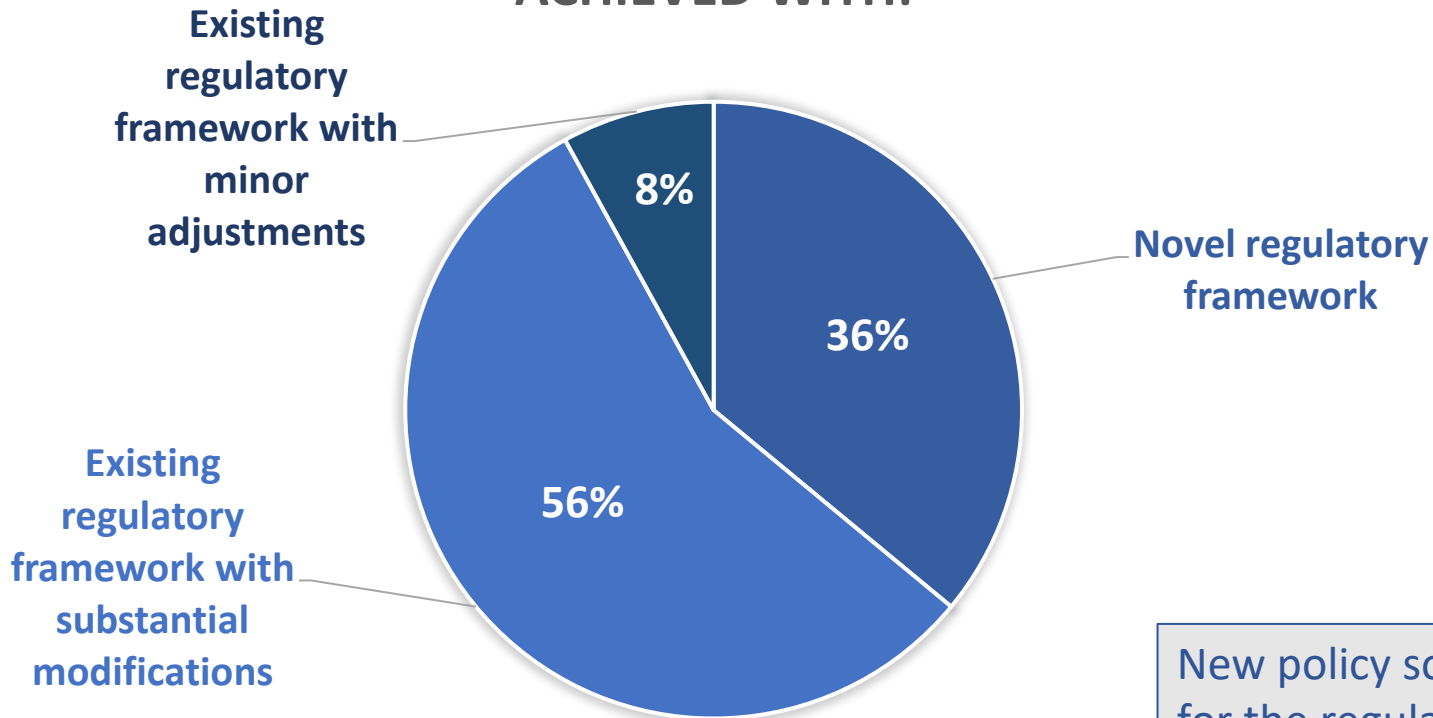


HAI Health Policy Meeting Poll Results



Stanford University
Human-Centered
Artificial Intelligence

EFFECTIVE GOVERNANCE OF HEALTHCARE AI CAN BE ACHIEVED WITH:



New policy solutions are needed for the regulation of health AI.



Recommendations

- Universal electronic data exchange
- Make patient data donation easy
- Research on explainable AI and human-computer interaction
- Improve transparency of AI systems
- “Consumer reports” for health AI
- Monitor system performance over time



Stanford AIMI

Center for Artificial Intelligence
in Medicine and Imaging

Engage:

@curtlanglotz

@StanfordAIMI

@StanfordHAI

LinkedIn: aimi-stanford

Join the mailing list: <https://aimi.stanford.edu/engage/subscribe>

