

# Frontiers in Generative AI for Medical Imaging and Healthcare

Session 1  
Introduction to Generative AI in Healthcare  
Sep 04 2025

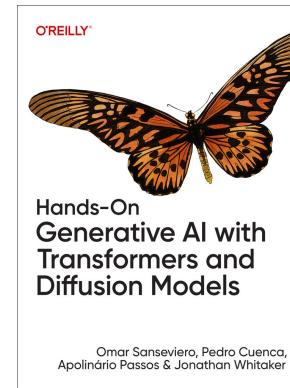
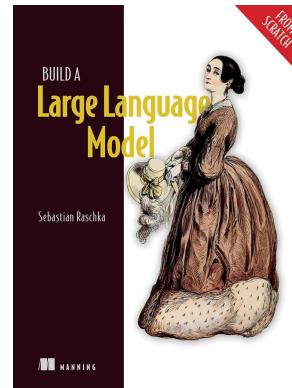
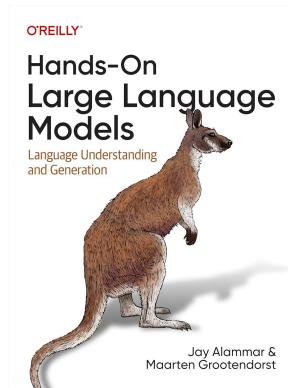
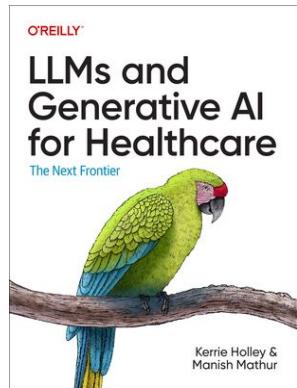
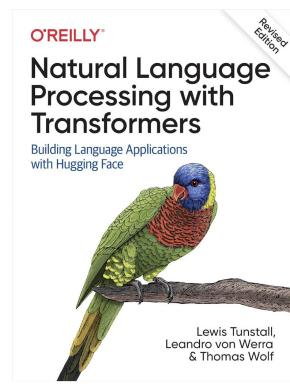
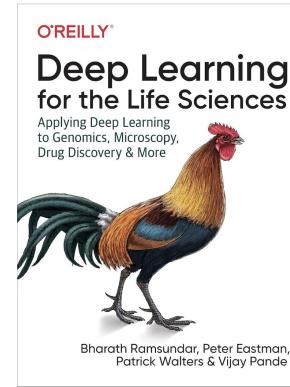
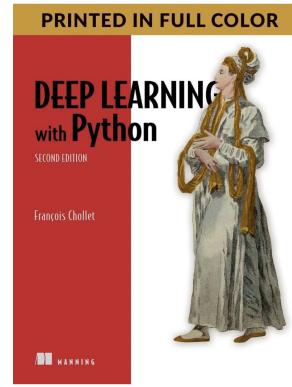
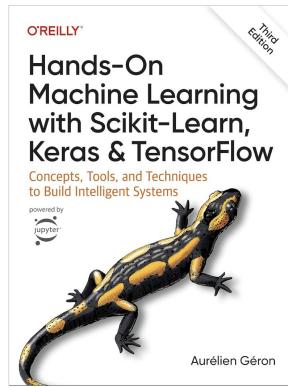
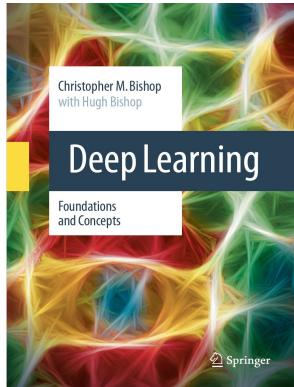
# Course structure and objectives

---

- Format Lecture + Interactive Lab sessions (hands-on coding in Jupyter Notebooks)
- Resources All slides, notebooks, and readings on Canvas; code repositories on **GitHub**
- Evaluation Homework assignments, Capstone project (individual)
- Learning Outcomes
  - Design, train, and debug deep-learning models for medical imaging, drug discovery, and related tasks.
  - Apply state-of-the-art generative models (e.g., GANs, VAEs, diffusion) to healthcare data.
  - Assess model safety, bias, and regulatory compliance; communicate findings to technical and clinical stakeholders.
- Prerequisites Proficiency in **Python** programming, Working knowledge of **PyTorch**, Core machine-learning concepts (supervised / unsupervised, evaluation metrics)

# Some useful resources

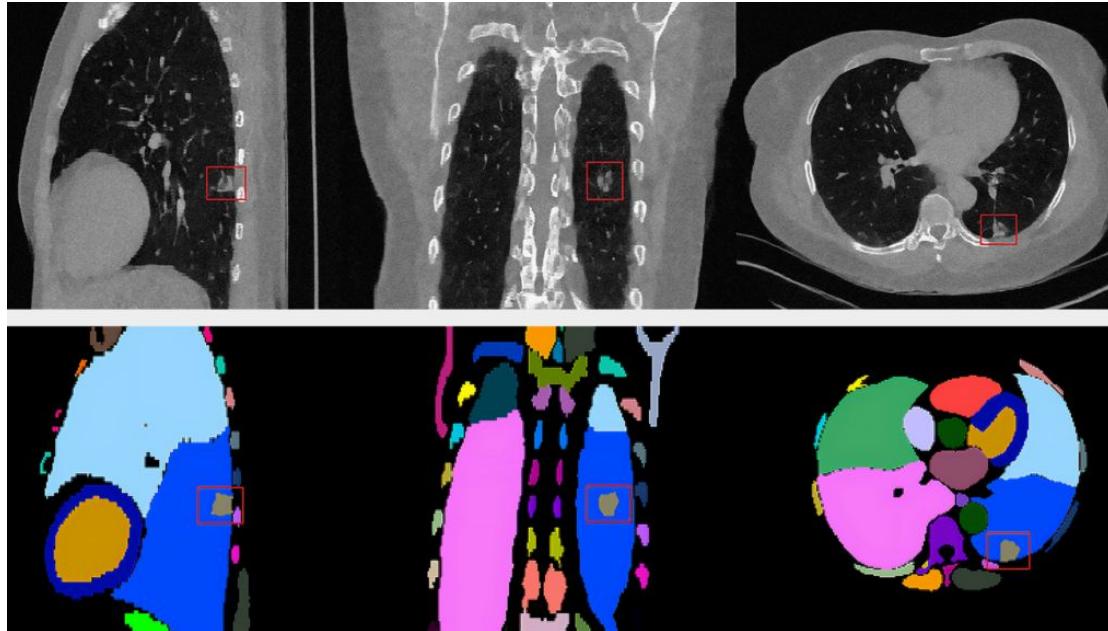
---



# AI in Medical Imaging

---

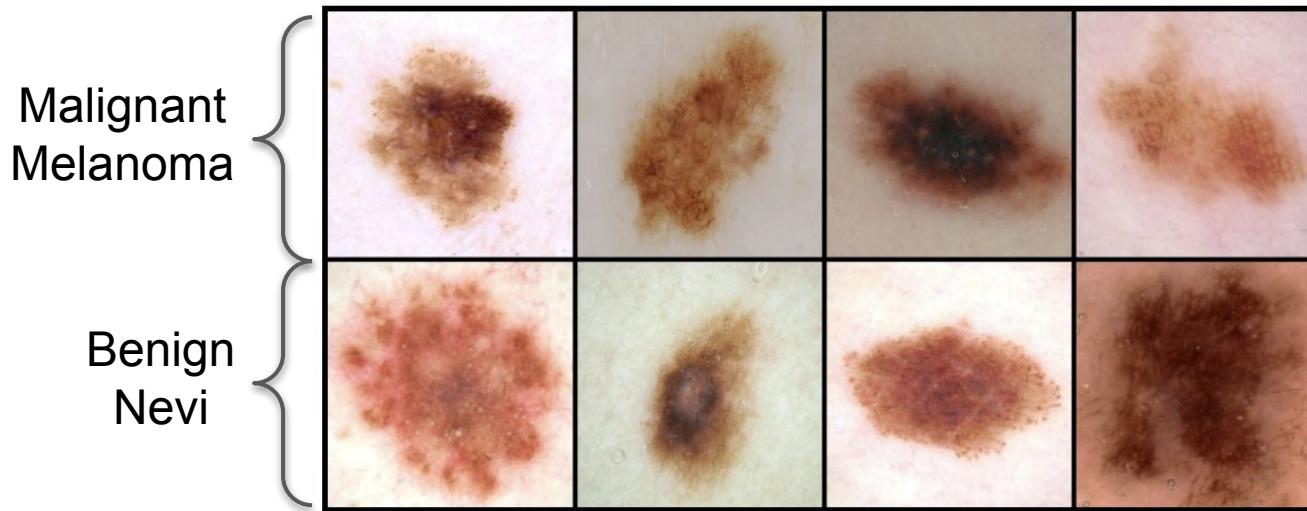
**Data Availability:** Electronic Health Records (EHRs), medical imaging repositories, genomics data.



# AI for Medical Diagnosis

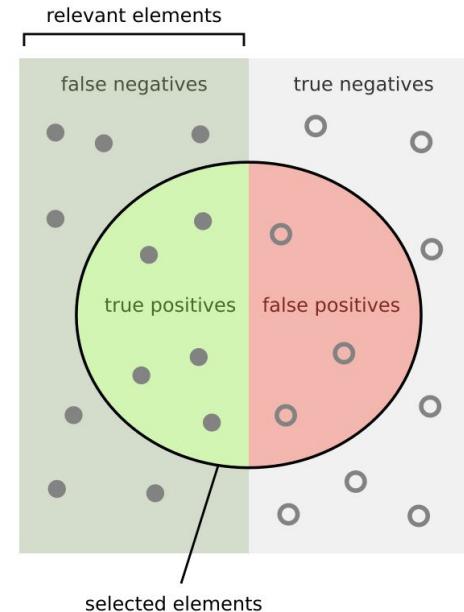
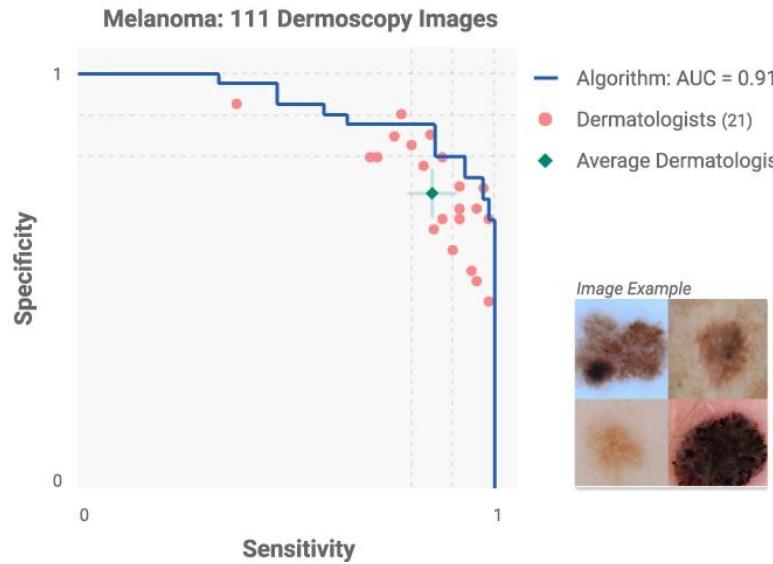
---

**Growing Need:** Address physician shortages, reduce medical errors, accelerate drug discovery, personalized medicine.



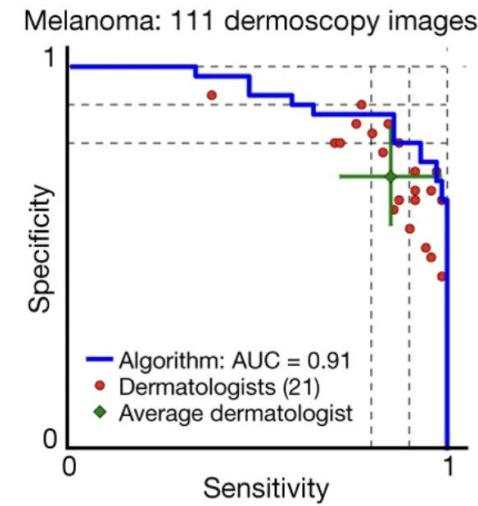
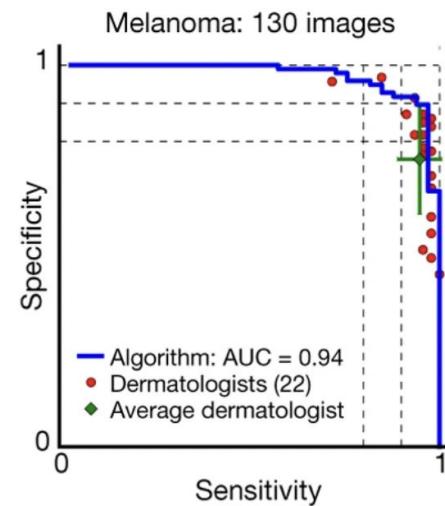
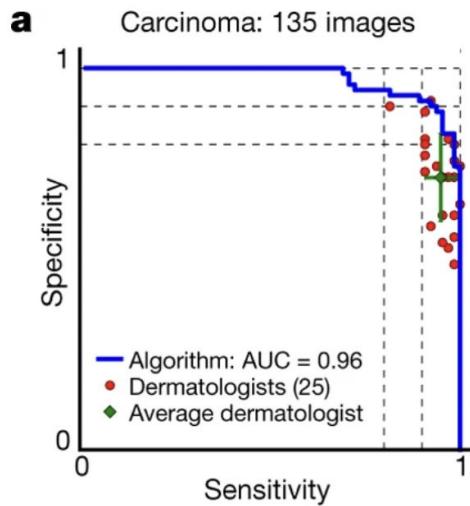
# AI in Medical Diagnosis

**Transformation Potential:** From diagnostics to drug discovery, deep learning is driving innovation.



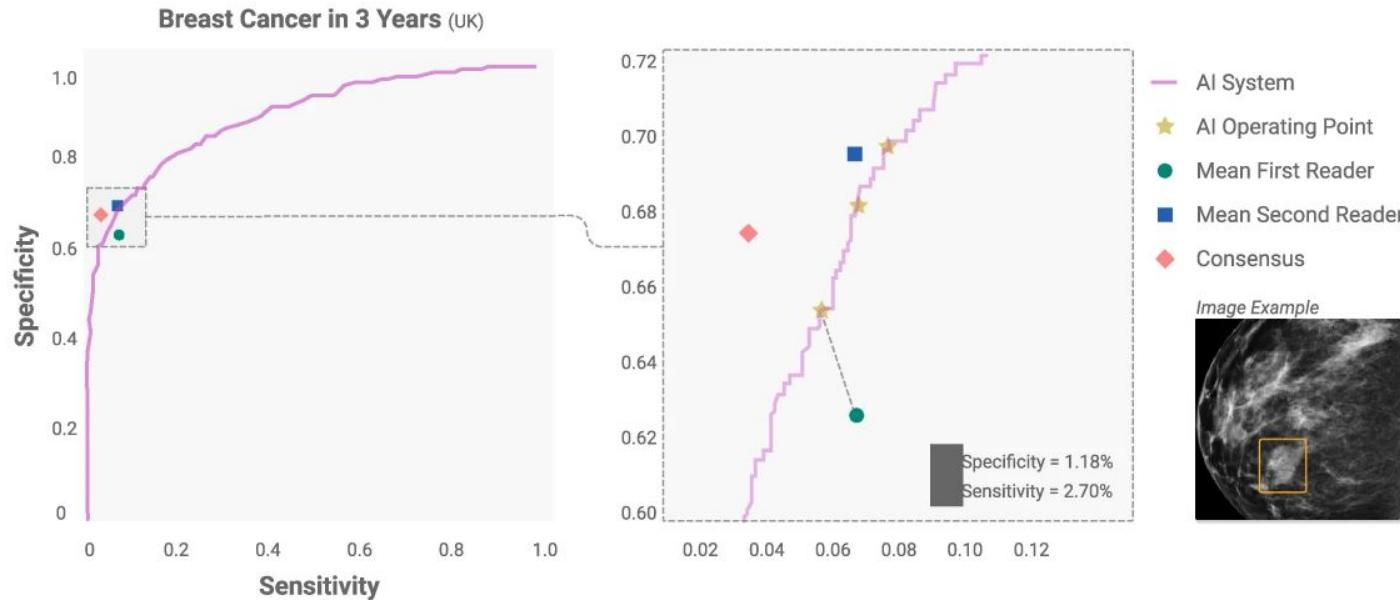
# AI in Medical Diagnosis

**Growing Need:** Address physician shortages, reduce medical errors, accelerate drug discovery, personalized medicine.



# AI in Medical Diagnosis

**Transformation Potential:** From diagnostics to drug discovery, deep learning is driving innovation.



# AI for Medical Diagnosis

---

**Training set** A large set of lesion images each labelled as *malignant* or *benign* (from biopsy)

**Training** Adjustment of millions parameters in *deep neural network* using the training set

**Supervised learning** For each training example, the network is told the correct label

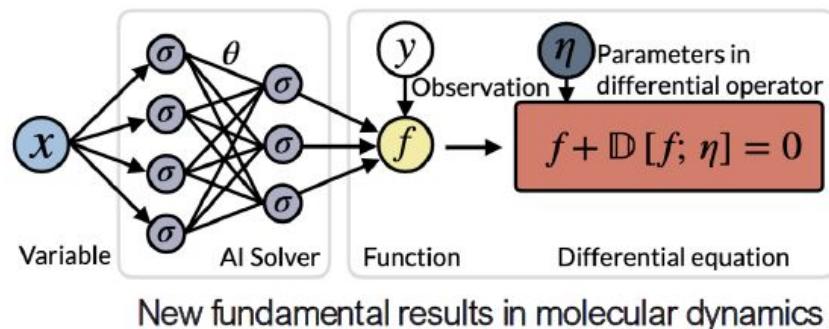
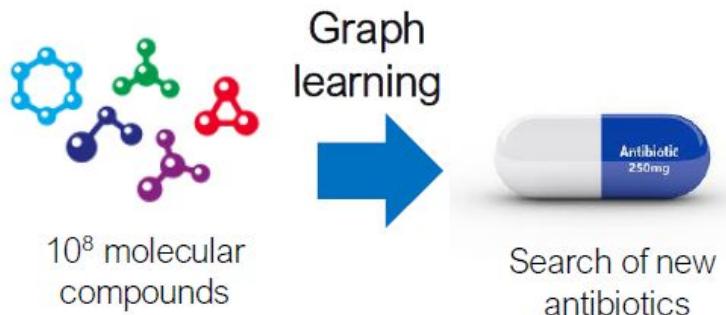
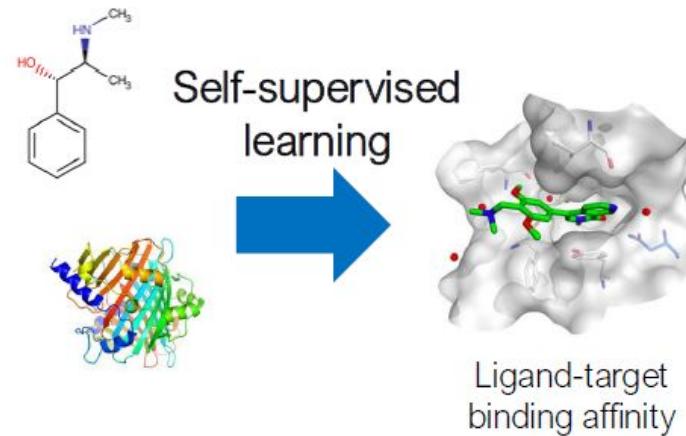
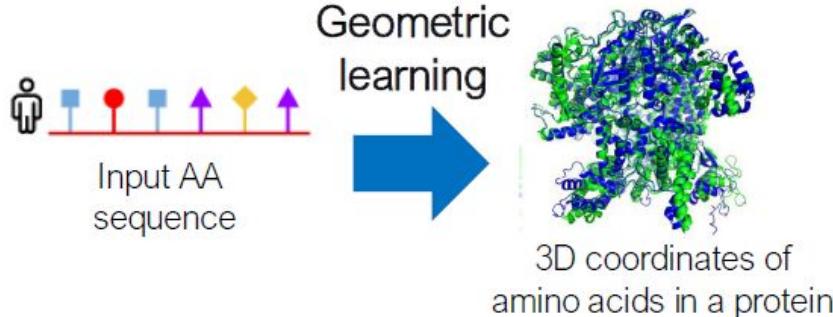
**Classification** Each input is assigned to a discrete set of classes (benign or malignant)

**Transfer learning** The deep neural network was first trained on a much larger data set of 1.28 million images of everyday objects (such as dogs, buildings, and mushrooms) and then fine-tuned on the 129,000 data set of lesion images

**Evaluation metrics** Accuracy, sensitivity, specificity, ROCAUC, confusion matrices, recall, precision, etc.

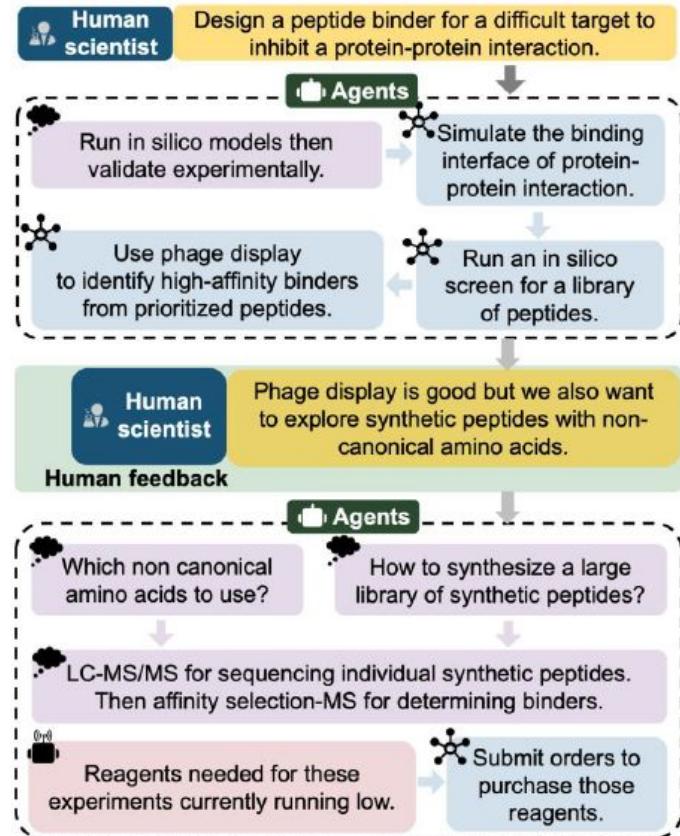
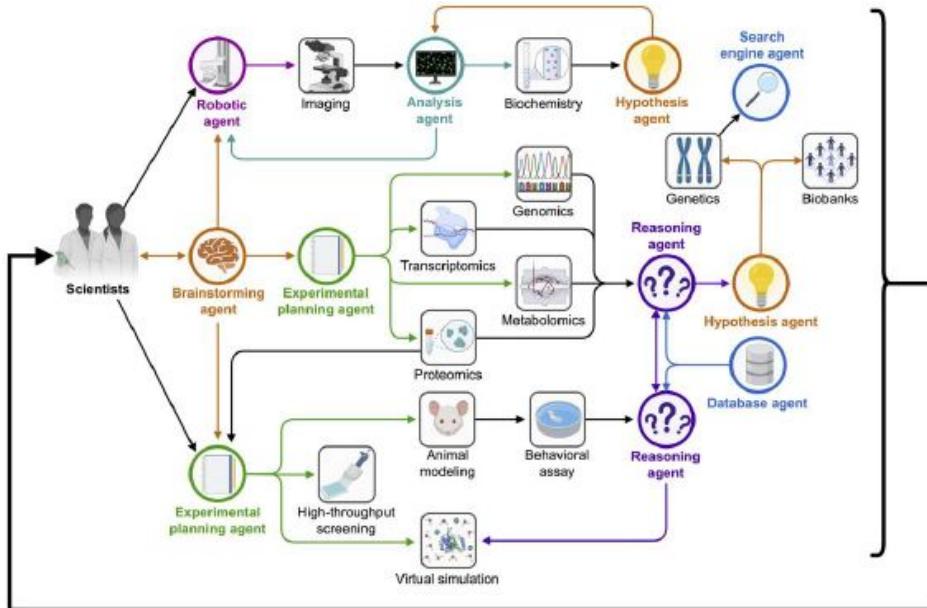
---

# AI in Medicine



# Generative AI Agents as “AI Scientist”

AI Systems can make **major scientific discoveries!**



# AI in Healthcare

Health system-scale language models  
are **all-purpose predictions engines**

a

Clinical task



#### In-hospital mortality prediction

How likely is the patient to die in the hospital before discharge?

#### Binned comorbidity index imputation

Without structured ICDS, how sick/chronically ill is the patient?

#### 30-day all-cause readmission prediction

How likely is the patient to come back within 30 days of discharge?

Operational task

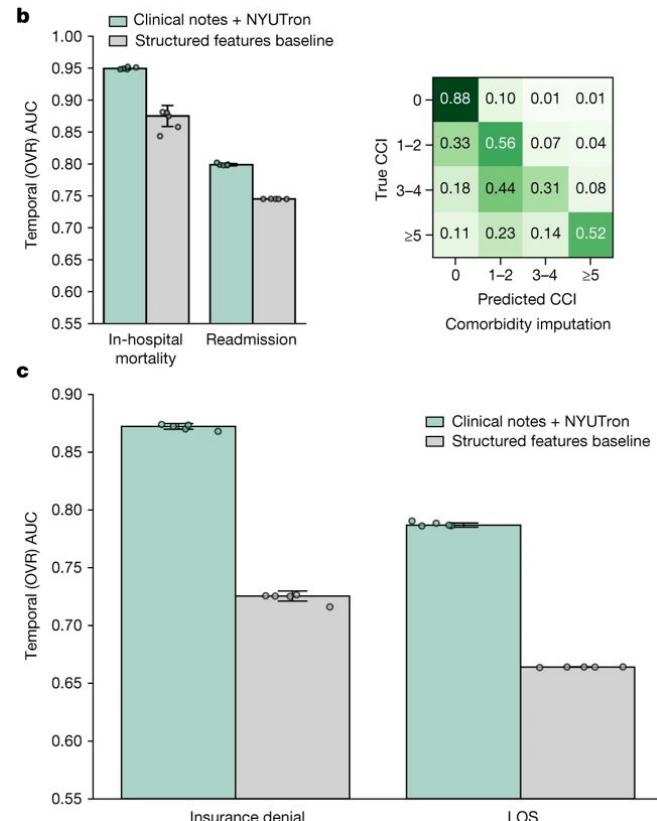


#### Binned LOS prediction

How long will the patient stay in the hospital?

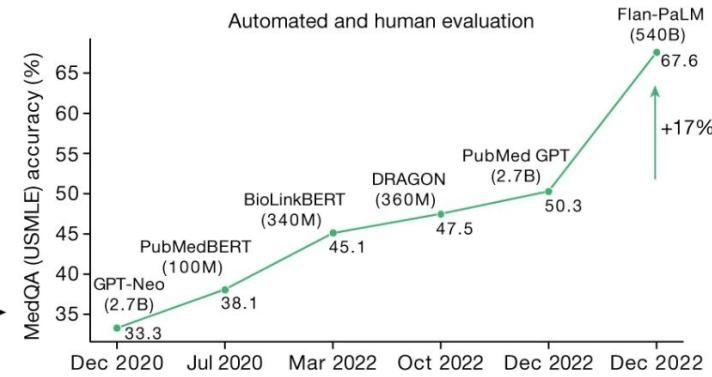
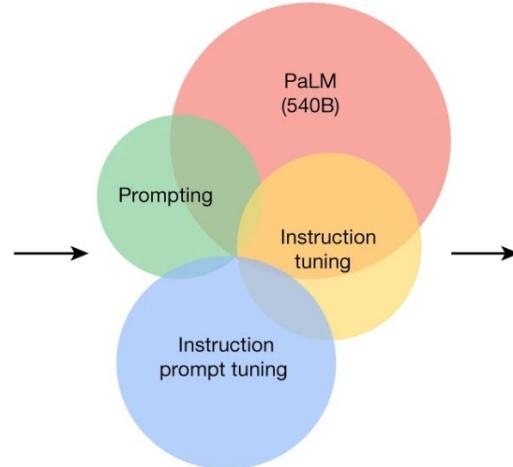
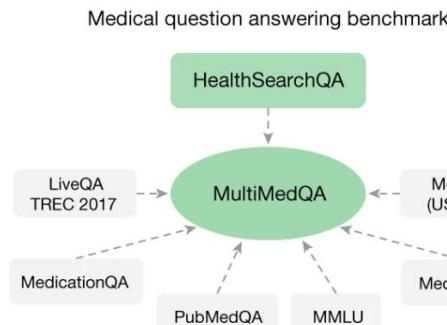
#### Insurance denial prediction

How likely is the patient's insurance claim to be denied?



# AI in Healthcare

Large language models encode **clinical knowledge**



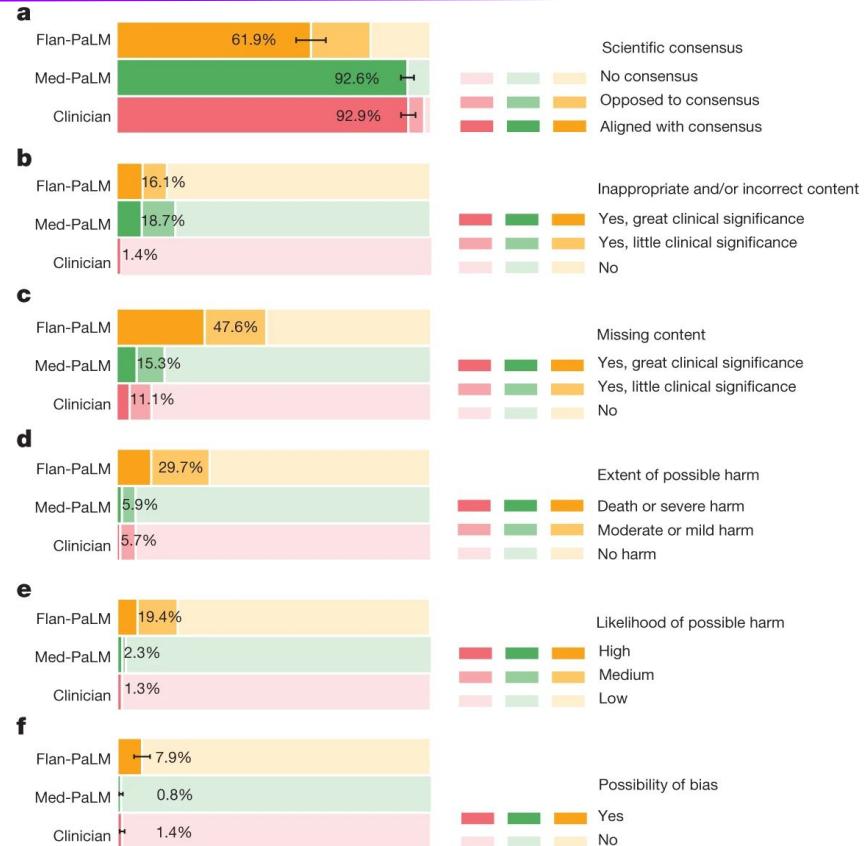
**Q:** How long does it take for newborn jaundice to go away?

**Med-PaLM:** Newborn jaundice is when a newborn baby's skin and eyes look yellow. It is caused by a buildup of a substance called bilirubin in the baby's blood. Bilirubin is a byproduct of the normal process of red blood cells breaking down. Newborn jaundice is common and typically harmless, but it can be a sign of a more serious condition. The jaundice typically goes away on its own within a few weeks. In some cases, it may last longer or require treatment. The time it takes for the jaundice to go away can vary depending on the cause and the severity of the jaundice. If the jaundice is severe or lasts longer than a few weeks, the doctor may recommend testing or treatment to determine the cause and prevent complications.

Med-PaLM performs encouragingly  
on consumer medical question answering

# AI in Healthcare

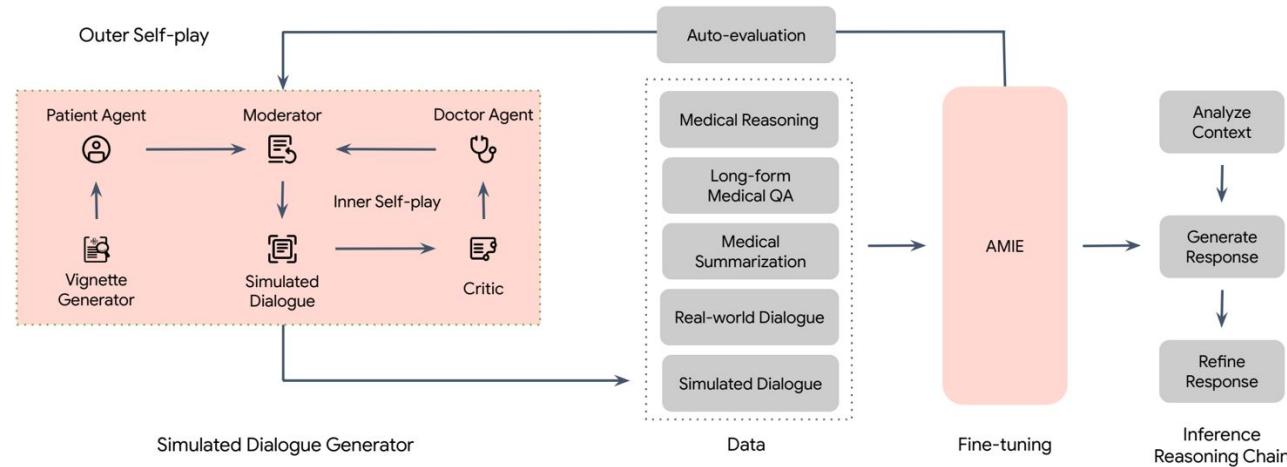
Clinicians were asked to rate answers to questions in the **HealthSearchQA**, **LiveQA** and **MedicationQA** datasets for agreement with scientific and clinical consensus.



# Conversational medical AI for diagnostic dialogue

**AMIE** is a conversational medical AI optimised for diagnostic dialogue.

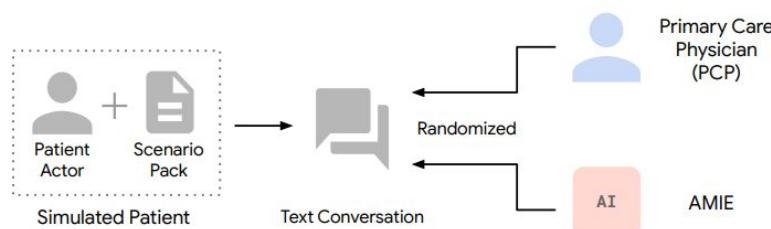
AMIE is instruction fine-tuned with a combination of real-world and simulated medical dialogues, alongside a diverse set of **medical reasoning, question answering, and summarization datasets**.



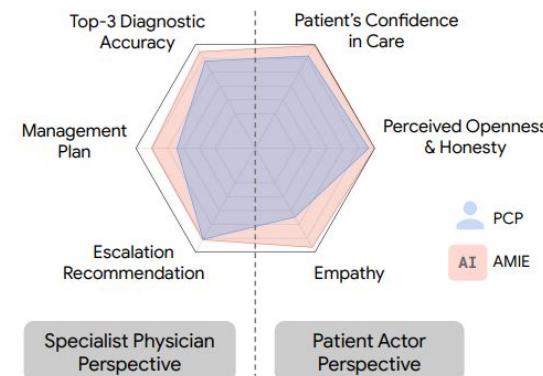
# Conversational medical AI for diagnostic dialogue

**AMIE** is a conversational medical AI optimised for diagnostic dialogue.

AMIE is instruction fine-tuned with a combination of real-world and simulated medical dialogues, alongside a diverse set of **medical reasoning, question answering, and summarization datasets**.



Randomized Study Design for Remote  
Objective Structured Clinical Examination (OSCE)

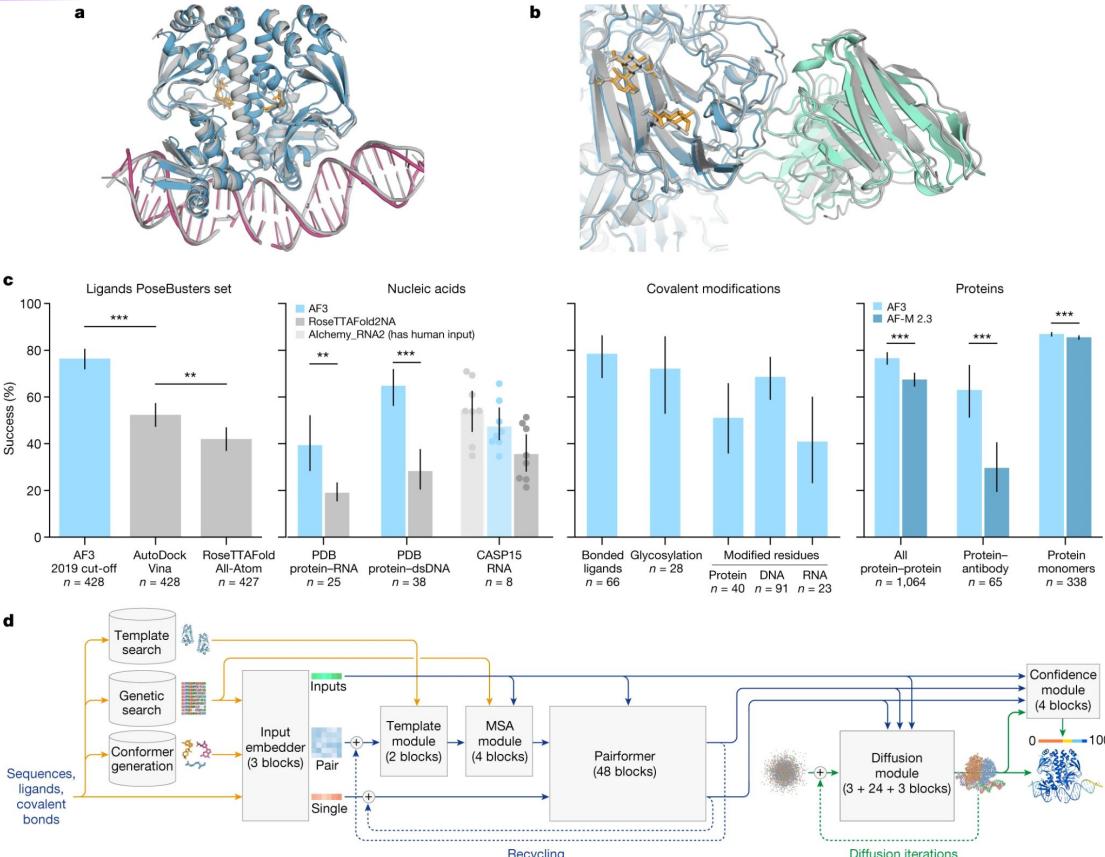


AMIE Outperforms PCPs on  
Multiple Evaluation Axes for Diagnostic Dialogue

# AI in Drug Discovery

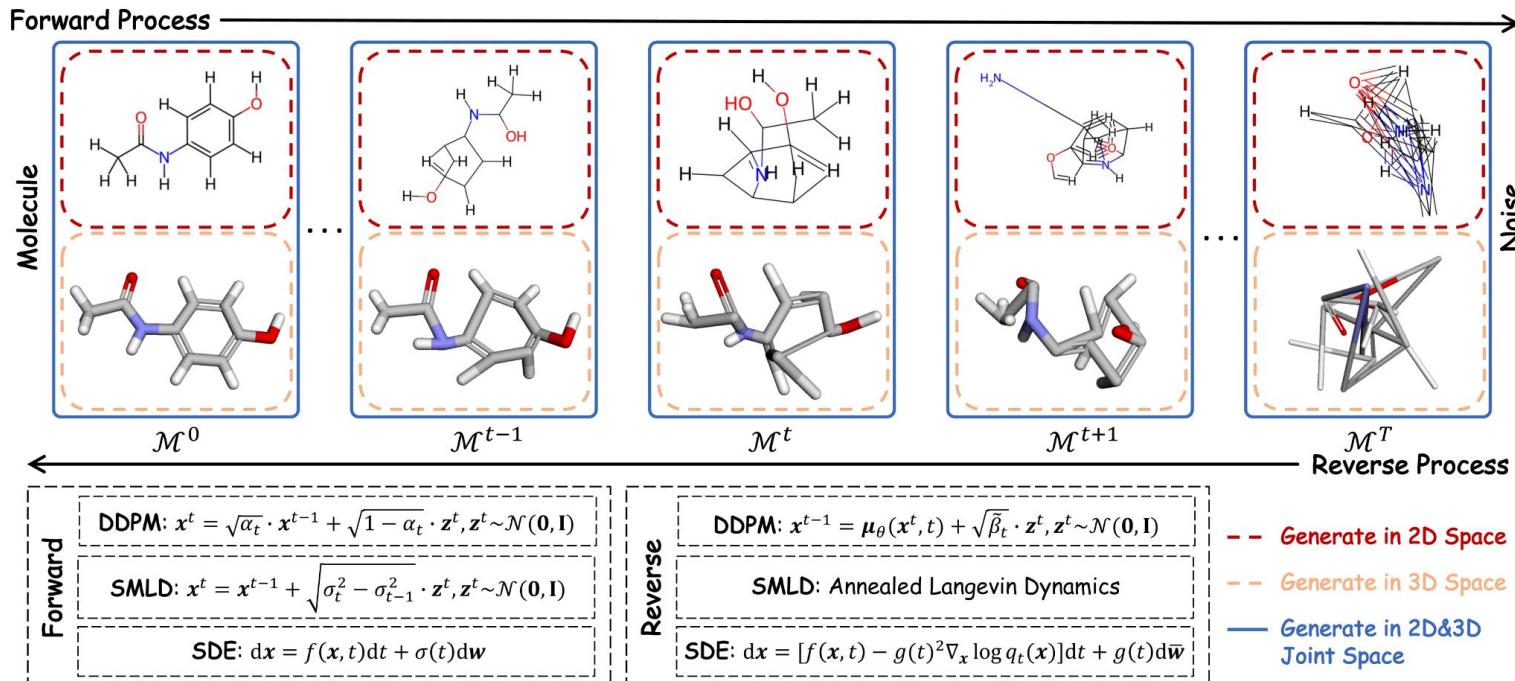
**Growing Need:** Address physician shortages, reduce medical errors, **accelerate drug discovery**, personalized medicine.

AlphaFold 3 model with a substantially updated **diffusion-based** architecture that is capable of predicting the **joint structure of complexes** including proteins, nucleic acids, small molecules, ions and modified residues.



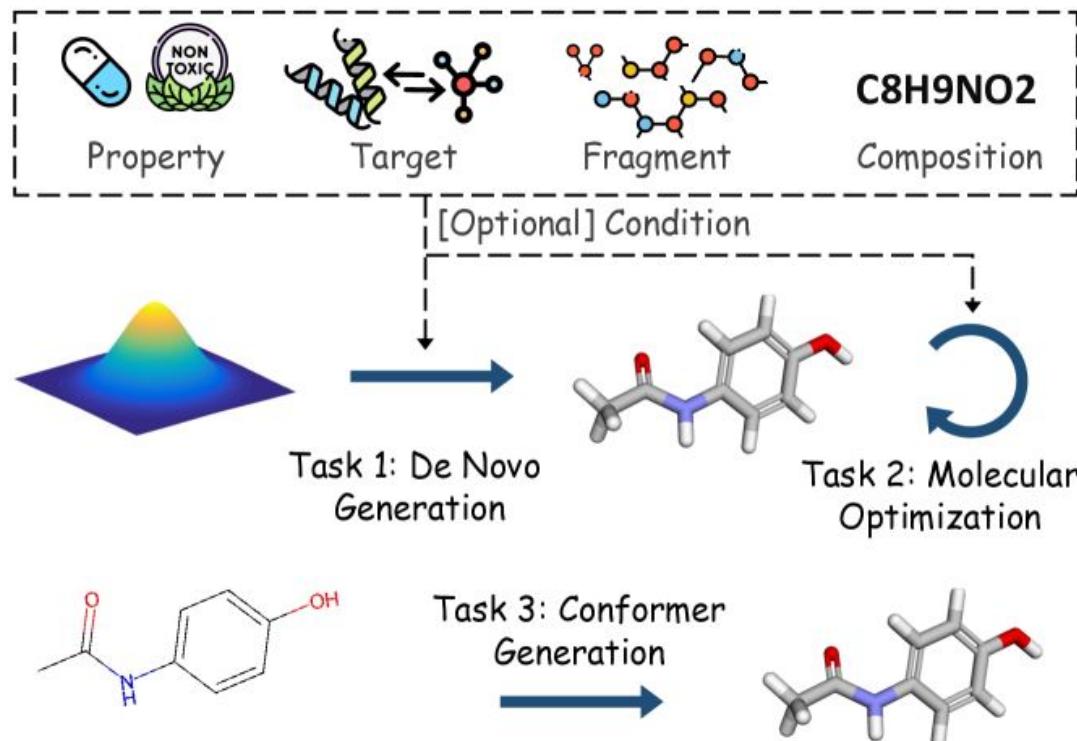
# AI in Drug Discovery

## Equivariant diffusion for small molecule generation



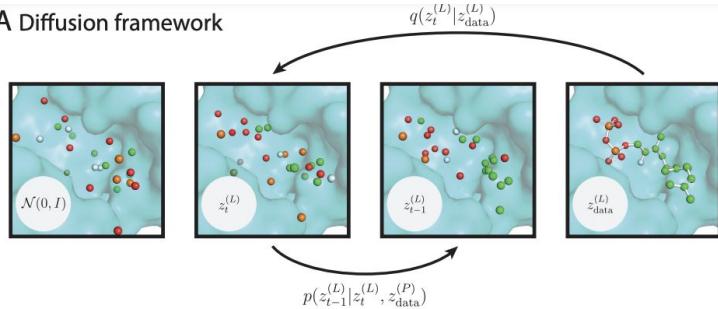
# AI in Drug Discovery

Equivariant diffusion for small molecule generation

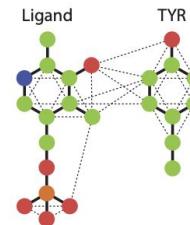


# AI in Drug Discovery

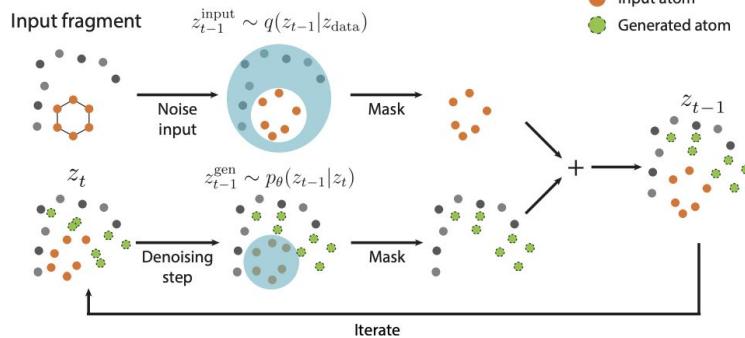
A Diffusion framework



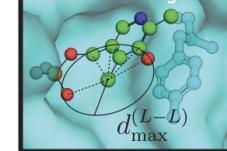
B Graph modeling



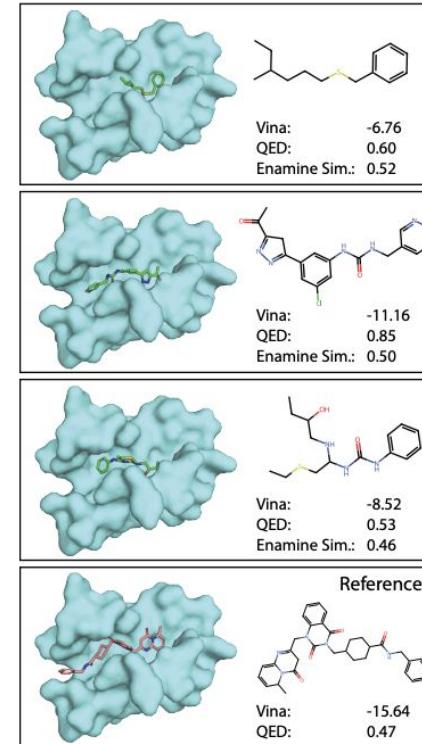
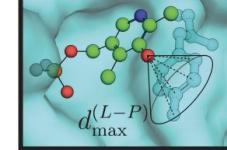
C Fixing known substructures



Intramolecular edges



Intermolecular edges



# AI in Drug Discovery

nature

Explore content ▾ About the journal ▾ Publish with us ▾

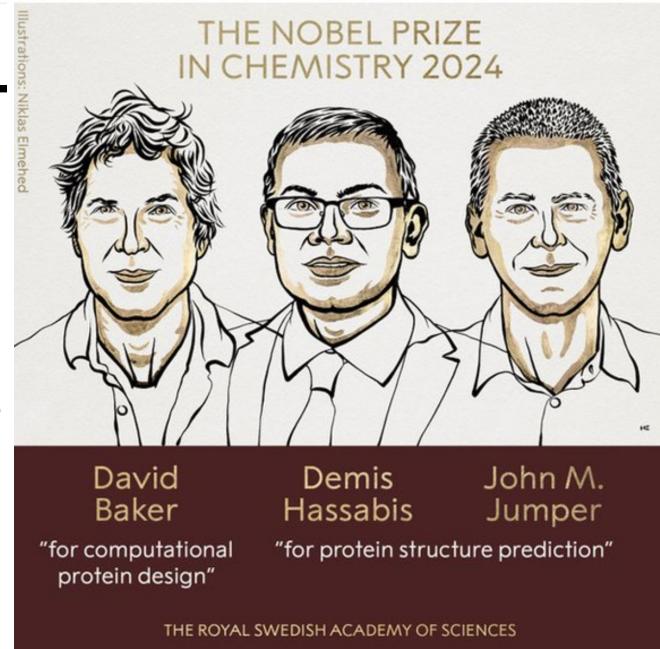
[nature](#) > [articles](#) > [article](#)

Article | [Open access](#) | Published: 11 July 2023

## De novo design of protein structure and function with RFdiffusion

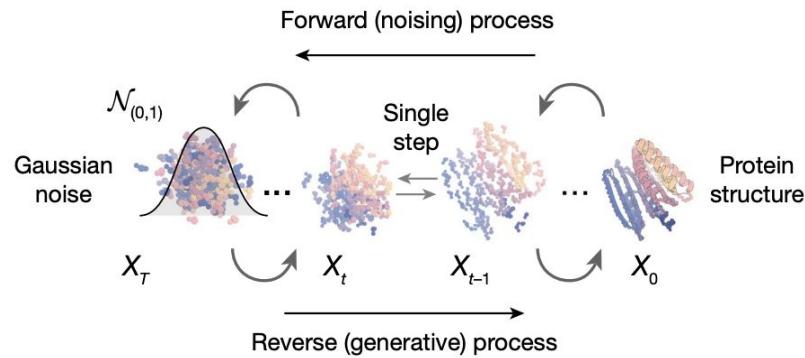
Joseph L. Watson, David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, Woody Ahern, Andrew J. Borst, Robert J. Ragotte, Lukas F. Milles, Basile I. M. Wicky, Nikita Hanikel, Samuel J. Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, Valentin De Bortoli, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, ... David Baker  [+ Show authors](#)

[Nature](#) 620, 1089–1100 (2023) | [Cite this article](#)

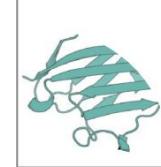


# AI in Drug Discovery

Diffusion model



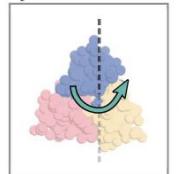
Binding target



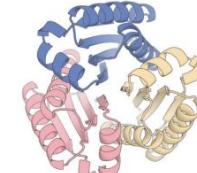
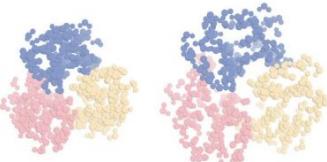
Functional motif



Symmetric noise



Symmetric oligomers



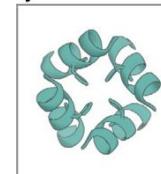
Binder design



Motif scaffolding

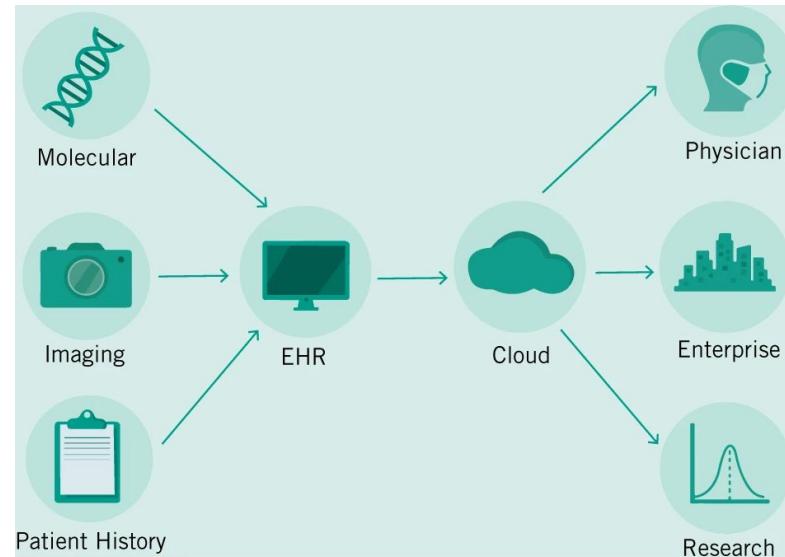
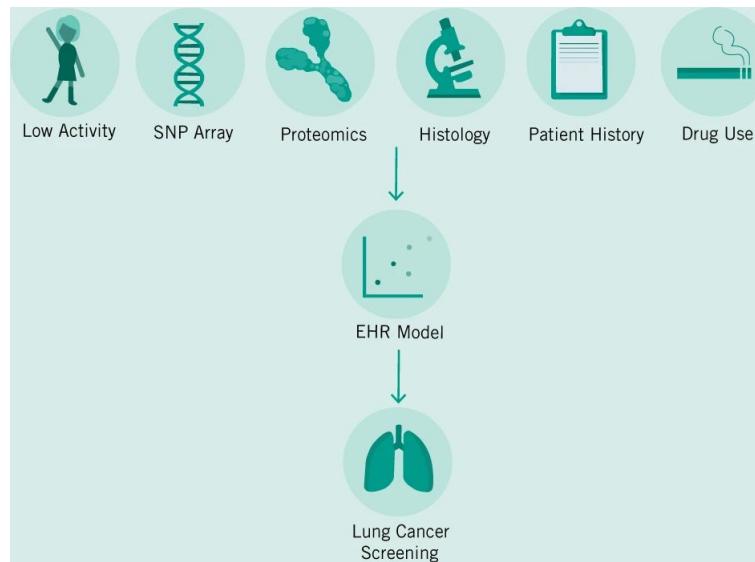


Symmetric motif



# What makes biomedical data different?

**Data Availability:** Electronic Health Records (EHRs), medical imaging repositories, genomics data.



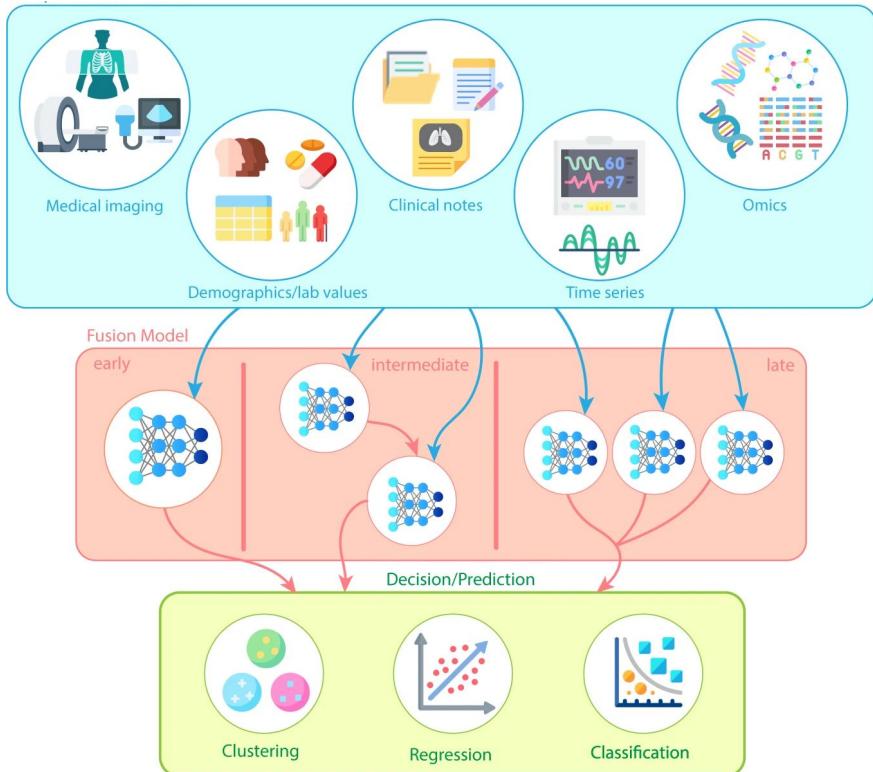
# What makes biomedical data different?

## Complexity of Healthcare Data:

Structured (EHRs, lab results), unstructured (clinical notes), image data (X-ray, MRI), multimodal data, etc.

**Accuracy** alone is not sufficient

- Explainability
- Non-discriminatory predictions
- Privacy-preserving
- Causal



# What makes biomedical data different?

- Little labeled data - narrow generalization
- Lots of missing data, varying time intervals, censored labels
- Difficult to correct **biases** and **inequities**
- Motivates **semi-supervised** and **self-supervised** learning
- Motivates **generative AI** for rare diseases



Russakovsky et al. '14

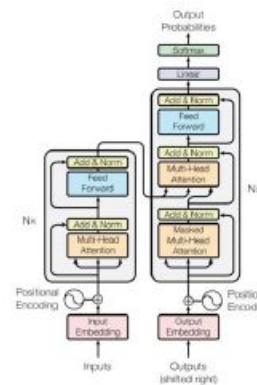


Figure 1: The Transformer - model architecture.

Vaswani et al. '18

# Medical Image Synthesis using Generative AI

---

Medical image acquisition is challenging:

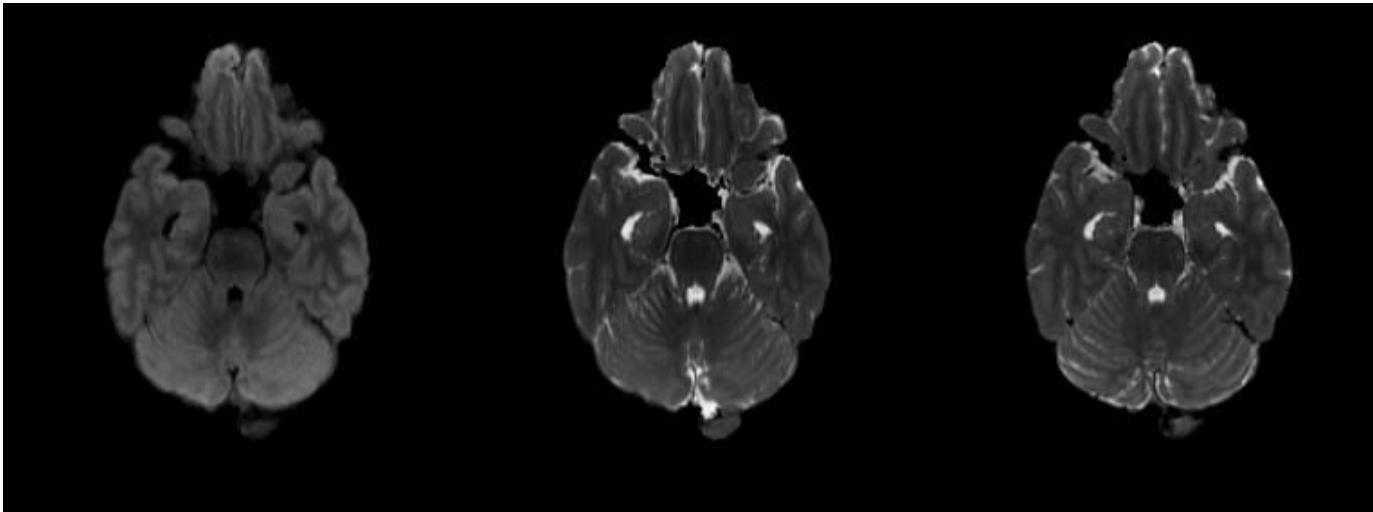
- **High operational costs** (technical fees, professional fees, facility fee)
- **High radiation exposure** (PET/CT scans expose patient to high radiation)
- **Long acquisition times** (motion artifacts due to patient movements)



# GenAI in Medical Image Synthesis

---

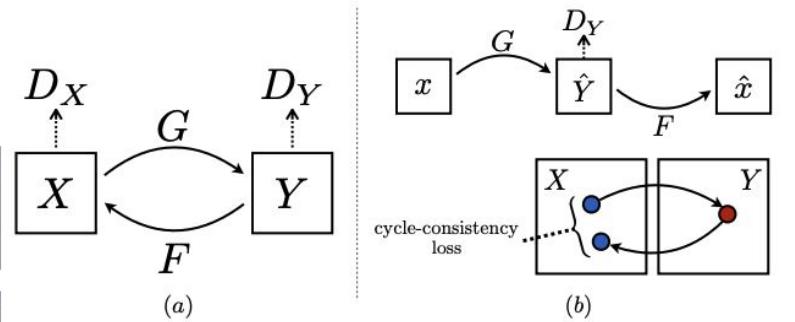
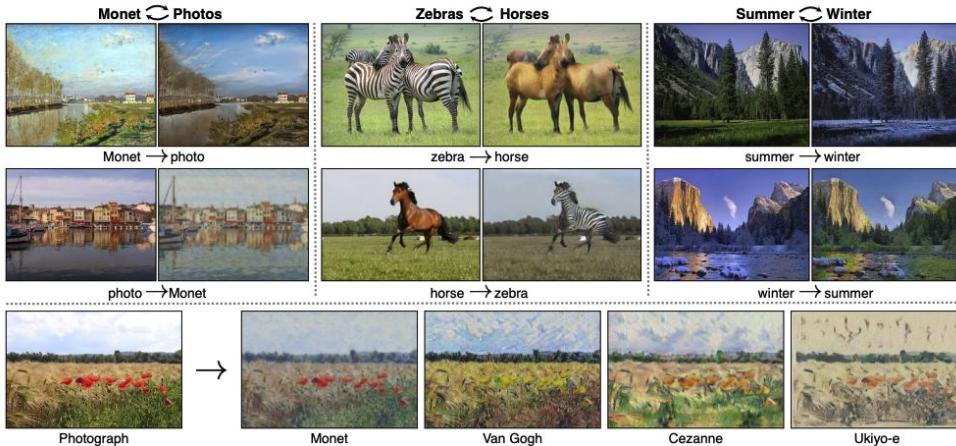
- **Low-field MRI** cuts equipment and operational costs
- **low-dose PET** reduces patient radiation exposure.
- Difficulties with **image quality, diagnostic accuracy, and practical implementation.**



# CycleGANs for unpaired image-to-image translation

## Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Jun-Yan Zhu\* Taesung Park\* Phillip Isola Alexei A. Efros  
Berkeley AI Research (BAIR) laboratory, UC Berkeley



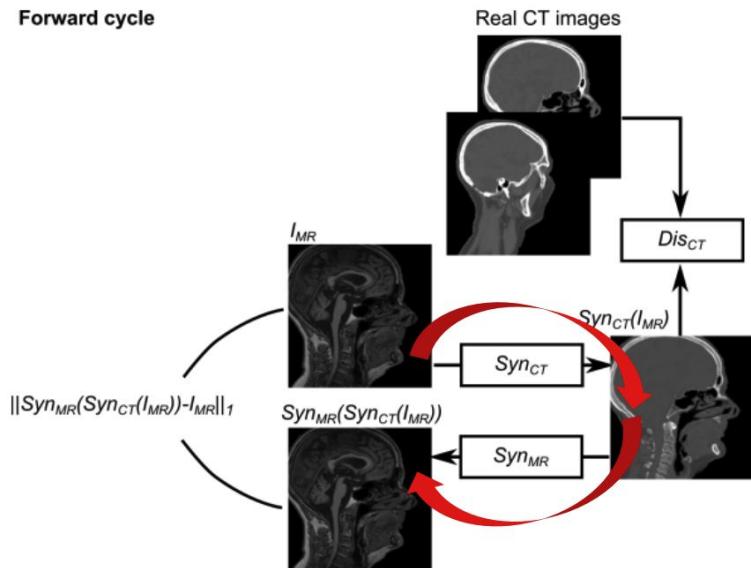
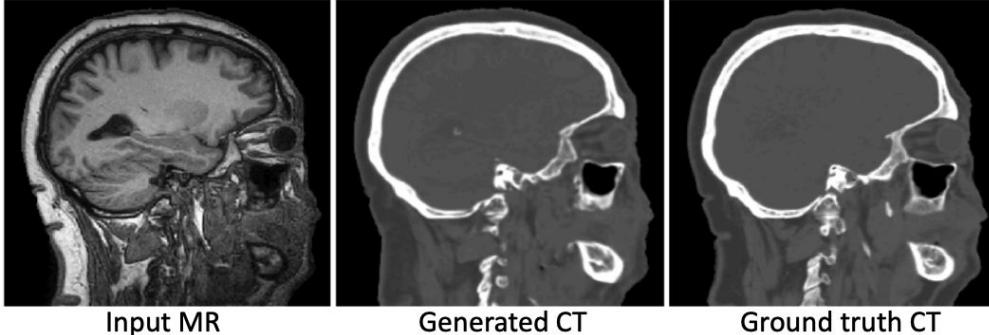
# MR to CT synthesis using CycleGANs

## Deep MR to CT Synthesis using Unpaired Data

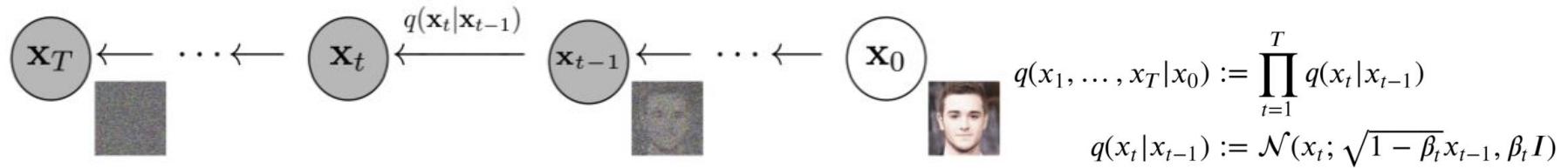
Jelmer M. Wolterink<sup>1</sup>✉, Anna M. Dinkla<sup>2</sup>, Mark H.F. Savenije<sup>2</sup>,  
Peter R. Seevinck<sup>1</sup>, Cornelis A.T. van den Berg<sup>2</sup>, Ivana Išgum<sup>1</sup>

<sup>1</sup> Image Sciences Institute, University Medical Center Utrecht, The Netherlands  
[j.m.wolterink@umcutrecht.nl](mailto:j.m.wolterink@umcutrecht.nl)

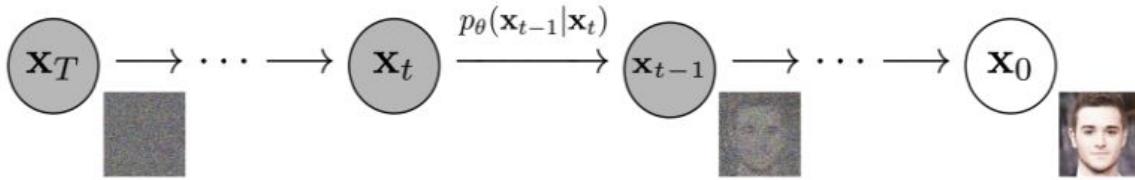
<sup>2</sup> Department of Radiotherapy, University Medical Center Utrecht, The Netherlands



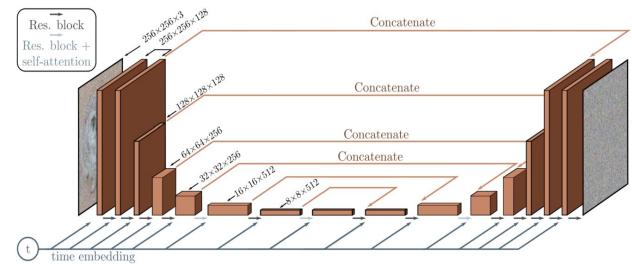
# Diffusion models for image-to-image translation



$$q(x_1, \dots, x_T | x_0) := \prod_{t=1}^T q(x_t | x_{t-1})$$
$$q(x_t | x_{t-1}) := \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I)$$



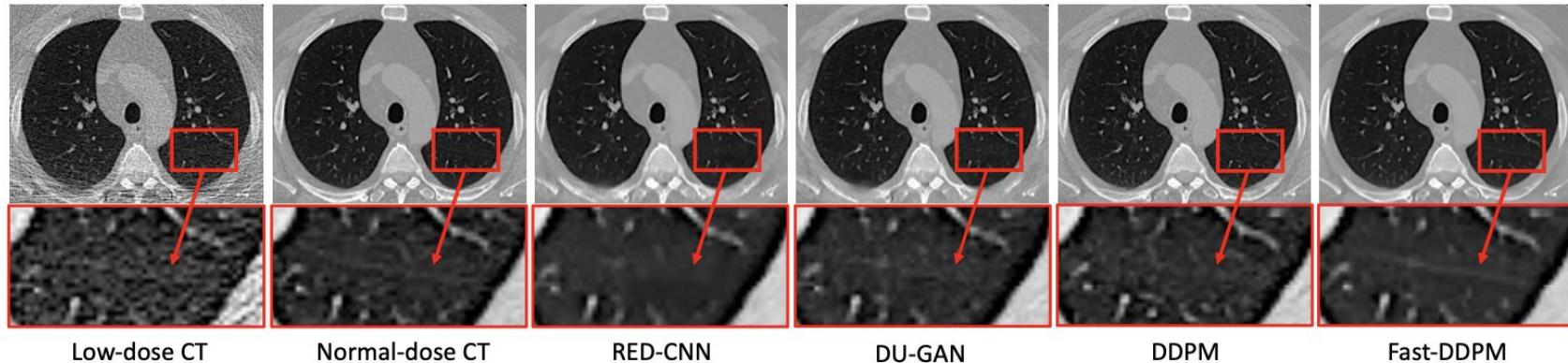
$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) := \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$



The U-Net architecture used in DDPMs

# DDPMs for medical image generation/translation

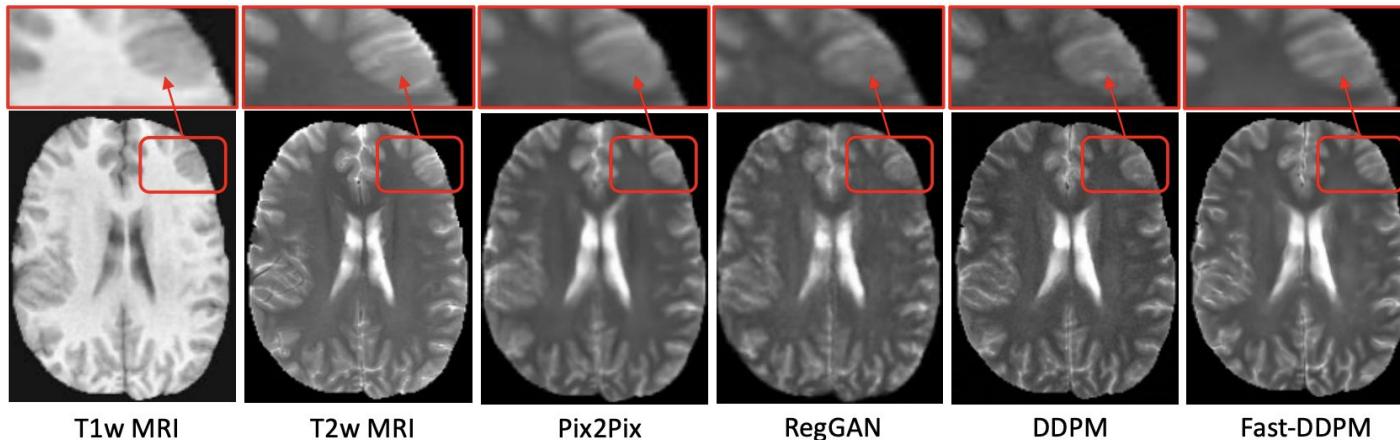
- Conversion of **low-dose (10% of normal)** to **full-dose lung CT**



Method	PSNR	SSIM	Training time	Inference Time
REDCNN [2]	36.4	0.91	3 h	0.5 s
DU-GAN [12]	36.3	0.90	20 h	3.8 s
DDPM [9]	35.4	0.87	141 h	21.4 m
Fast-DDPM	<b>37.5</b>	<b>0.92</b>	26 h	12.5 s

# DDPMs for medical image generation/translation

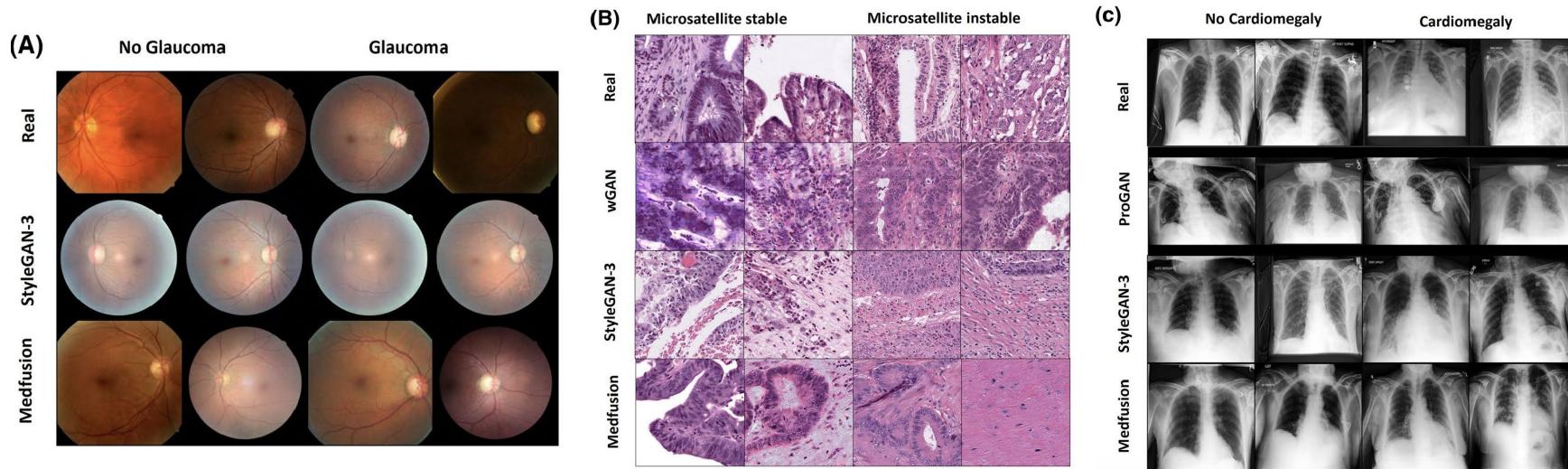
- T1w MRI to T2w MRI conversion



Method	PSNR	SSIM	Training time	Inference Time
Pix2Pix [14]	25.6	0.85	6 h	3.3 s
RegGAN [19]	26.0	0.86	9 h	3.1 s
DDPM [9]	<b>26.3</b>	0.89	135 h	22.2 m
Fast-DDPM	<b>26.3</b>	0.89	27 h	13.2 s

# Latent diffusion model for medical image synthesis

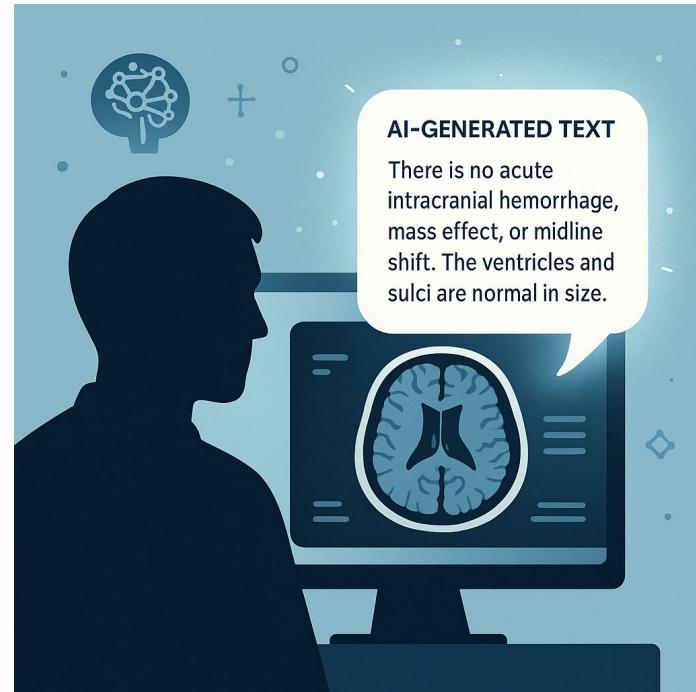
- Training data simulation for **computer-aided diagnostics**



# Medical Chatbots

---

- Initialized with general-purpose LLMs (GPT-4, LLaMA, etc.) and fine-tuned on customized sets of biomedical instruction-following data
  - ChatDoctor
  - Med-Alpaca
  - PMC-LLaMA
  - Clinical Camel
  - DoctorGLM
  - Huatuo
  - Visual Med-Alpaca (multimodal: text + vision)



# Medical Chatbots

---

Question	Answer
What is the relationship between very low Mg <sup>2+</sup> levels, PTH levels, and Ca <sup>2+</sup> levels?	Very low Mg <sup>2+</sup> levels correspond to low PTH levels which in turn results in low Ca <sup>2+</sup> levels.
What are the cautionary considerations and contraindications for the use of $\beta$ -blockers?	$\beta$ -blockers must be used cautiously in decompensated heart failure and are contraindicated in cardiogenic shock.
What is the recommended next step in management for a stable child who is suspected of having intussusception and has a "target sign" on ultrasound?	The recommended next step in management for a stable child with suspected intussusception and a "target sign" on ultrasound is to undergo an ultrasound-guided air contrast enema. This procedure involves using a small amount of air and contrast material to inflate the colon and push the telescoped intestine back into place. It is a non-invasive and effective method for both diagnosing and treating intussusception in children. In some cases, the procedure may need to be repeated or followed by surgery if the intussusception cannot be reduced. It is important to treat intussusception promptly to avoid complications such as bowel obstruction, bowel perforation, or sepsis.

---

# Medical Chatbots

- United States Medical Licensing Examination (**USMLE**)

## MEDALPACA - AN OPEN-SOURCE COLLECTION OF MEDICAL CONVERSATIONAL AI MODELS AND TRAINING DATA

A PREPRINT

Tianyu Han<sup>1,+</sup>, Lisa C. Adams<sup>2,+</sup>, Jens-Michalis Papaioannou<sup>4</sup>, Paul Grundmann<sup>4</sup>, Tom Oberhauser<sup>4</sup>, Alexei Figueiroa<sup>4</sup>, Alexander Löser<sup>4</sup>, Daniel Truhn<sup>1,+</sup>, and Keno K. Bressem<sup>5,6,+</sup>

<sup>1</sup>Department of Radiology, University Hospital Aachen, Aachen, Germany  
Email: {tianyu.han, dtruhn}@ukaachen.de

<sup>2</sup>Department of Diagnostic and Interventional Radiology, Technical University of Munich, Munich, Germany  
Email: lisa.adams@tum.de

<sup>4</sup>Berliner Hochschule für Technik (BHT), Berlin, Germany  
Email: {michalis.papaioannou, pgrundmann, tom.oberhauser, alexei.figueiroa, aloeser}@bht-berlin.de

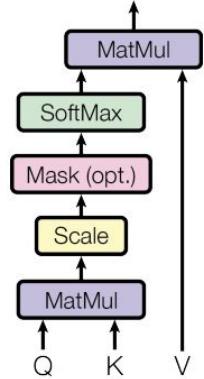
<sup>5</sup>Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute for Radiology, Berlin, Germany  
Email: keno-kirill.bressem@charite.de

Table 2: Zero shot performance on the USMLE self assessment

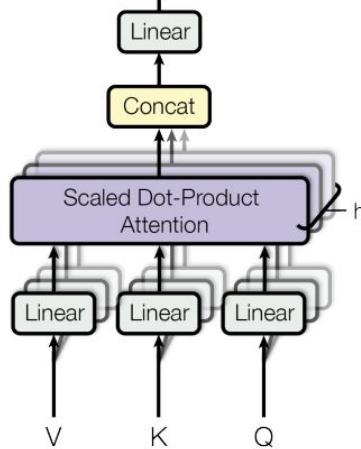
Model	Step1	Step2	Step3
LLaMA 7b [15]	0.198	0.202	0.203
Alpaca 7b naive [11]	0.275	0.266	0.293
Alpaca 7b LoRA	0.220	0.138	0.252
MedAlpaca 7b	0.297	0.312	0.398
MedAlpaca 7b LoRA	0.231	0.202	0.179
MedAlpaca 7b LoRA 8bit	0.231	0.241	0.211
ChatDoctor (7b) [10]	0.187	0.185	0.148
LLaMA 13b [15]	0.222	0.248	0.276
Alpaca 13b naive	0.319	0.312	0.301
MedAlpaca 13b	<b>0.473</b>	<b>0.477</b>	<b>0.602</b>
MedAlpaca 13b LoRA	0.250	0.255	0.255
MedAlpaca 13b LoRA 8bit	0.189	0.303	0.289

# Attention is all you need!

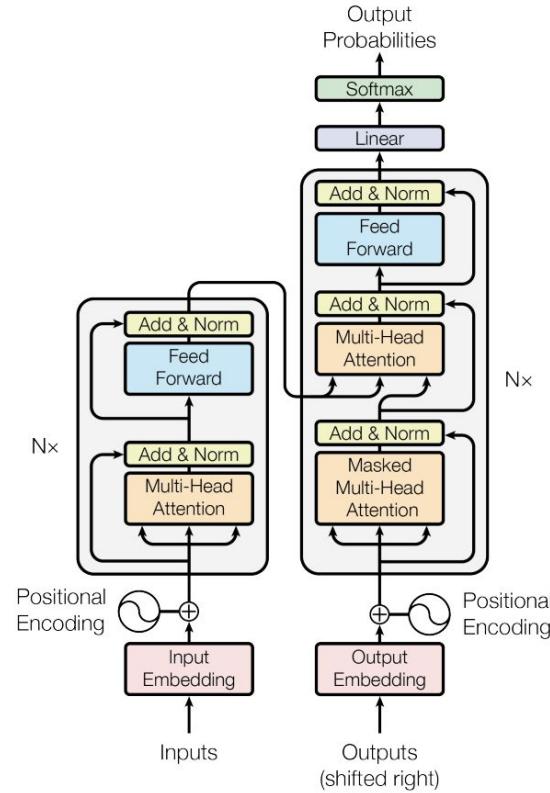
Scaled Dot-Product Attention



Multi-Head Attention

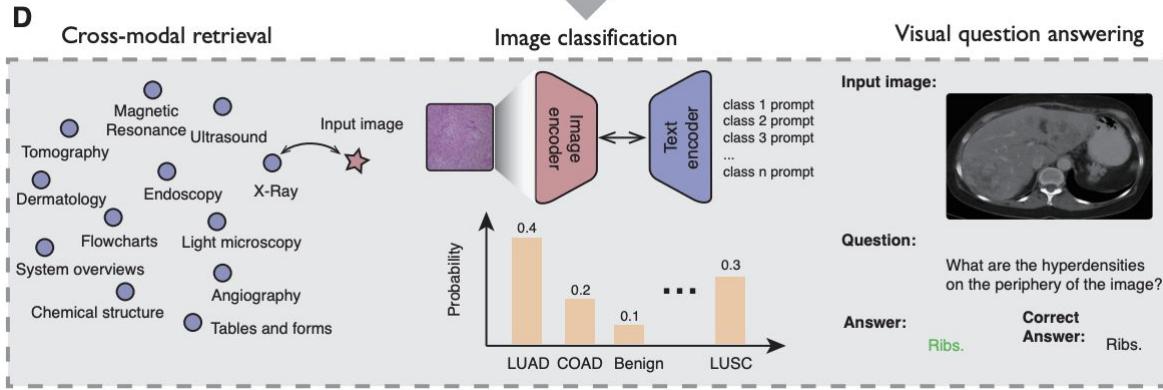
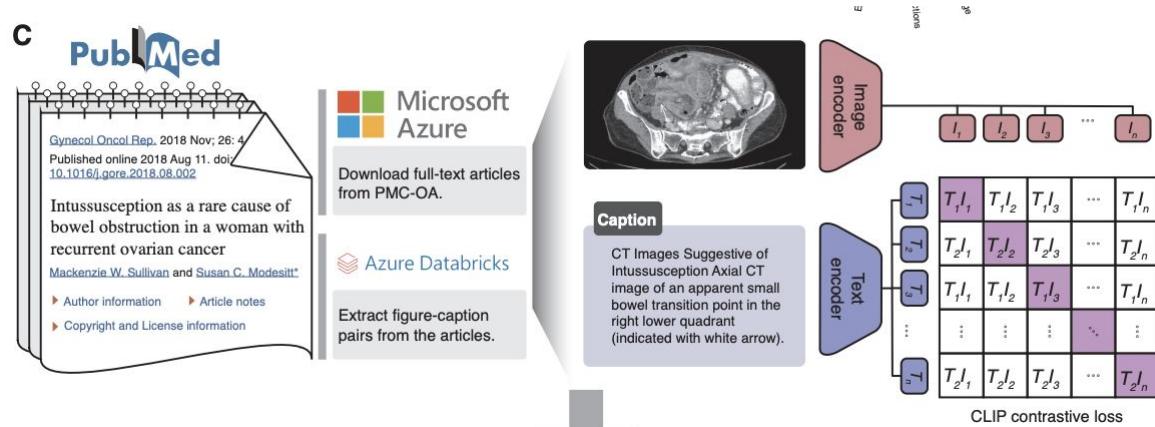


$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad \sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$



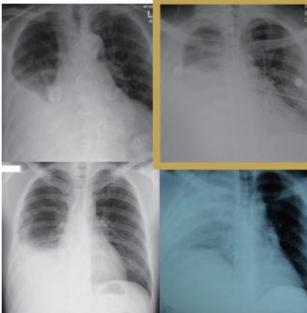
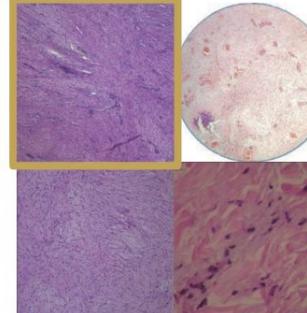
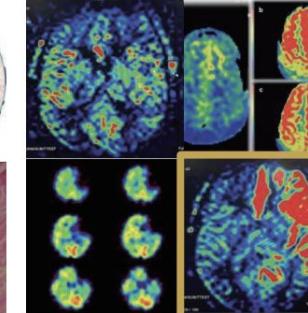
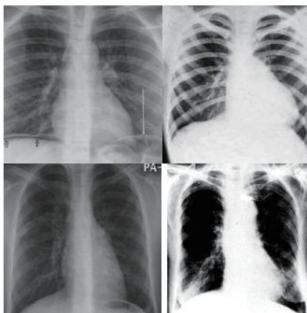
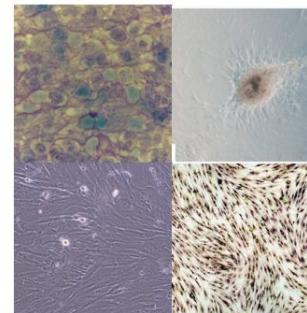
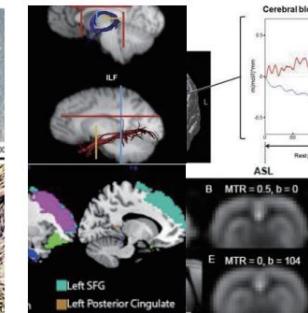
# Multimodal biomedical foundation models

- BiomedCLIP

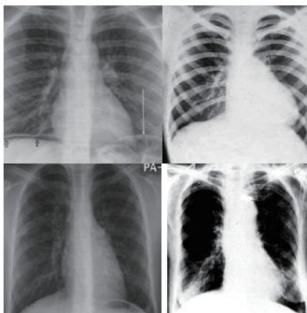
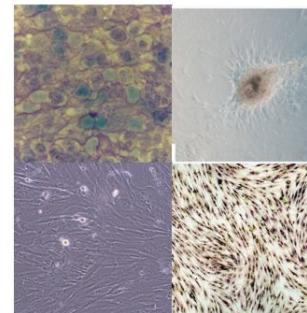
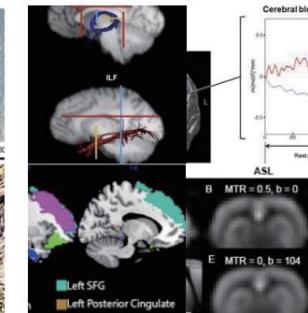
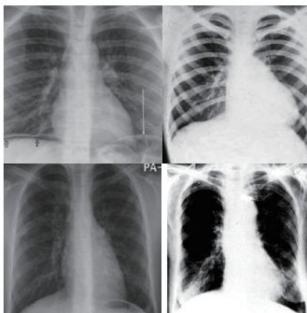
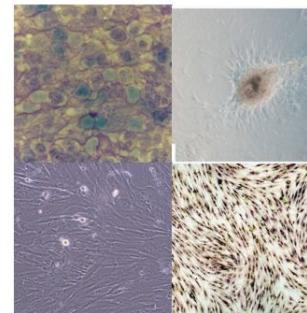
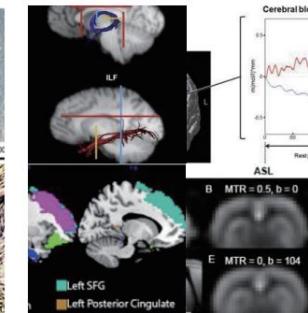


# Medical Image Captioning

BiomedCLIP

	Caption	Caption	Caption
	Chest X-ray on admission showing a large pleural effusion on the right.	Photomicrograph showing proliferation of spindle cells with wavy nuclei in collagenous to myxoid stroma (Hematoxylin and Eosin staining, 40x).	ASL image shows increased cerebral blood flow (CBF) in left fronto-parieto-temporal region.
			
			

CLIP

	Caption	Caption	Caption
	Chest X-ray on admission showing a large pleural effusion on the right.	Photomicrograph showing proliferation of spindle cells with wavy nuclei in collagenous to myxoid stroma (Hematoxylin and Eosin staining, 40x).	ASL image shows increased cerebral blood flow (CBF) in left fronto-parieto-temporal region.
			
			

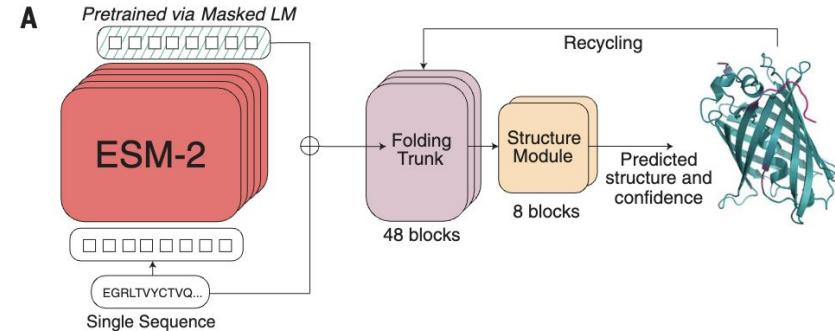
# Protein language models (PLM)

## STRUCTURE PREDICTION

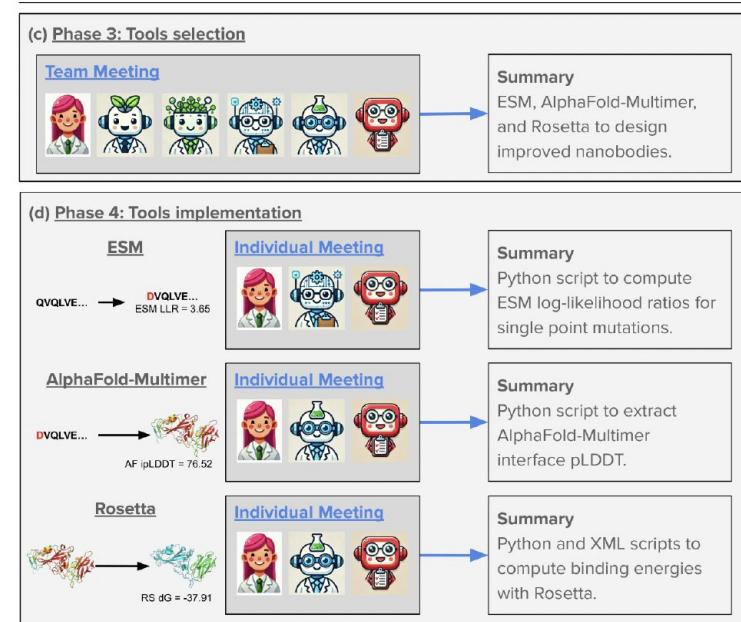
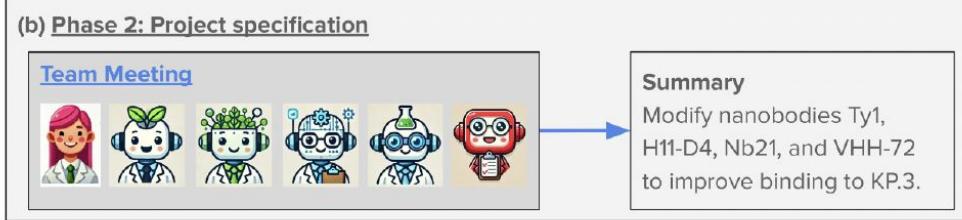
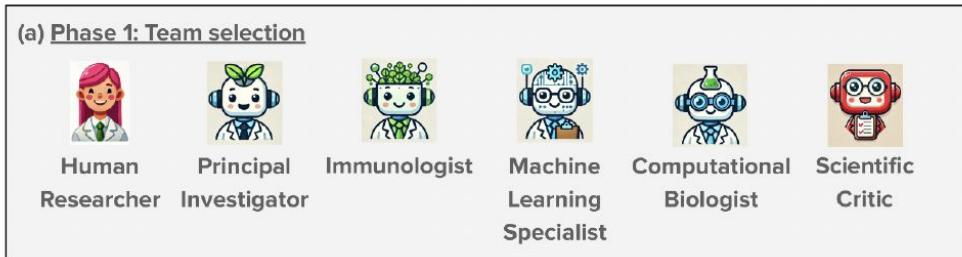
### Evolutionary-scale prediction of atomic-level protein structure with a language model

Zeming Lin<sup>1,2†</sup>, Halil Akin<sup>1†</sup>, Roshan Rao<sup>1†</sup>, Brian Hie<sup>1,3†</sup>, Zhongkai Zhu<sup>1</sup>, Wenting Lu<sup>1</sup>, Nikita Smetanin<sup>1</sup>, Robert Verkuij<sup>1</sup>, Ori Kabeli<sup>1</sup>, Yaniv Shmueli<sup>1</sup>, Allan dos Santos Costa<sup>4</sup>, Maryam Fazel-Zarandi<sup>1</sup>, Tom Sercu<sup>1</sup>, Salvatore Candido<sup>1</sup>, Alexander Rives<sup>1,2\*</sup>

Recent advances in machine learning have leveraged evolutionary information in multiple sequence alignments to predict protein structure. We demonstrate direct inference of full atomic-level protein structure from primary sequence using a large language model. As language models of protein sequences are scaled up to 15 billion parameters, an atomic-resolution picture of protein structure emerges in the learned representations. This results in an order-of-magnitude acceleration of high-resolution structure prediction, which enables large-scale structural characterization of metagenomic proteins. We apply this capability to construct the ESM Metagenomic Atlas by predicting structures for >617 million metagenomic protein sequences, including >225 million that are predicted with high confidence, which gives a view into the vast breadth and diversity of natural proteins.



# AI Agents propose strategy for antibody design



# Conclusions

---

- **Enhanced Diagnostic Accuracy:** Deep learning models excel in medical imaging, detecting anomalies with precision, often surpassing human performance.
  - **Accelerated Drug Discovery:** Generative AI speeds up the design of novel drugs, **reducing time and cost** in the development pipeline.
  - **Personalized Medicine:** AI enables **tailored treatments** by predicting patient-specific responses, improving outcomes, and minimizing side effects.
  - **Scalable and Equitable Healthcare:** AI-powered solutions improve access to care in under-resourced areas through **telemedicine** and **predictive analytics**.
  - **Data-Driven Insights:** Unlocks patterns in large-scale health data, **supporting evidence-based decisions** and advancing medical research.
-