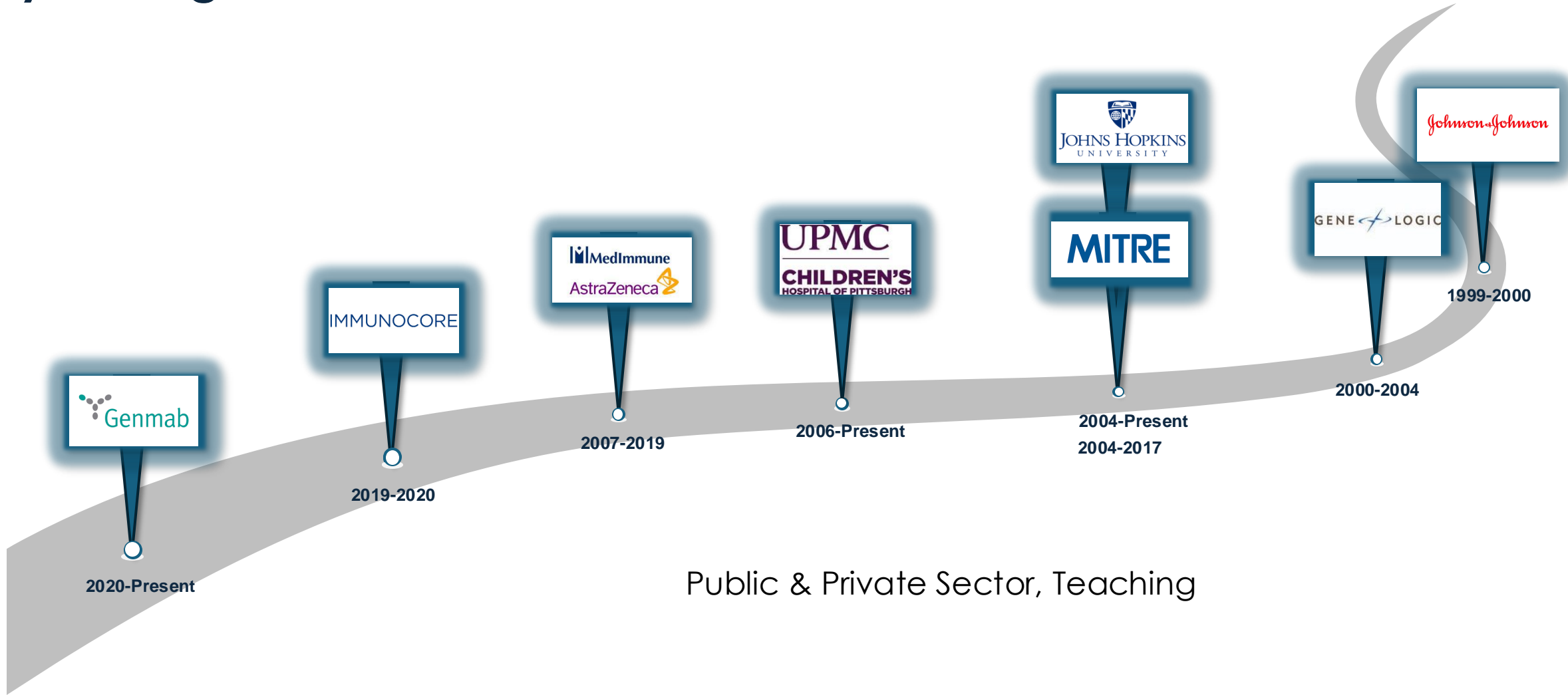


# Artificial Intelligence in Translational and Clinical Applications

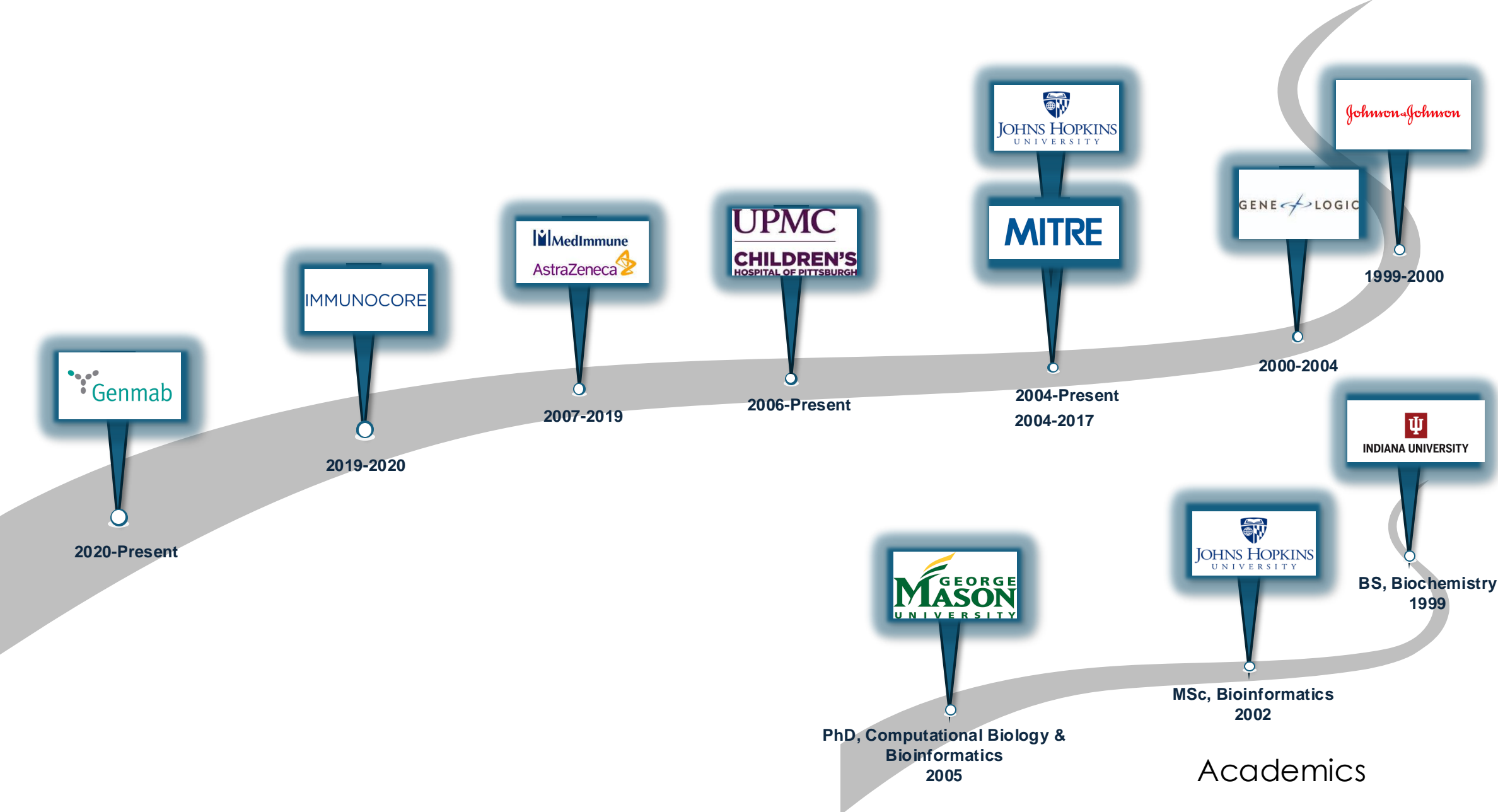
Brandon Higgs, PhD

VP, Global Head of Translational Data Sciences, Genmab

# My background

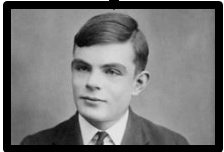


# My background



# A brief history of Artificial Intelligence (AI)

1950



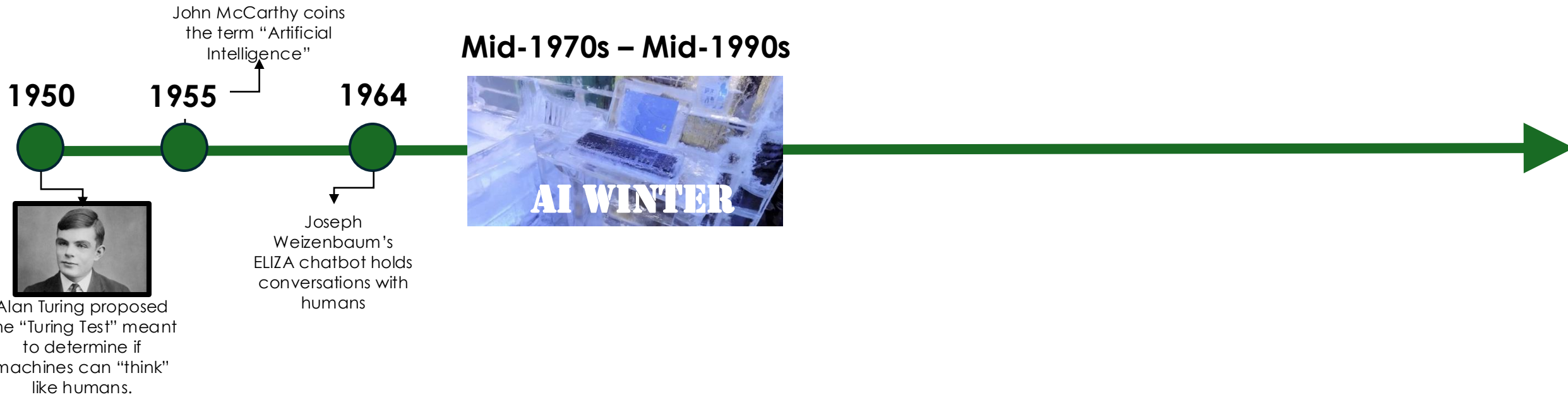
Alan Turing proposed the "Turing Test" meant to determine if machines can "think" like humans.

# A brief history of Artificial Intelligence (AI)

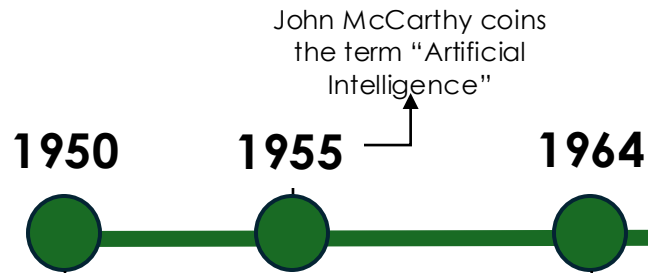




# A brief history of Artificial Intelligence (AI)



# A brief history of Artificial Intelligence (AI)



Alan Turing proposed the "Turing Test" meant to determine if machines can "think" like humans.

John McCarthy coins the term "Artificial Intelligence"


Joseph Weizenbaum's ELIZA chatbot holds conversations with humans

## Mid-1970s – Mid-1990s



 Intel (1968)

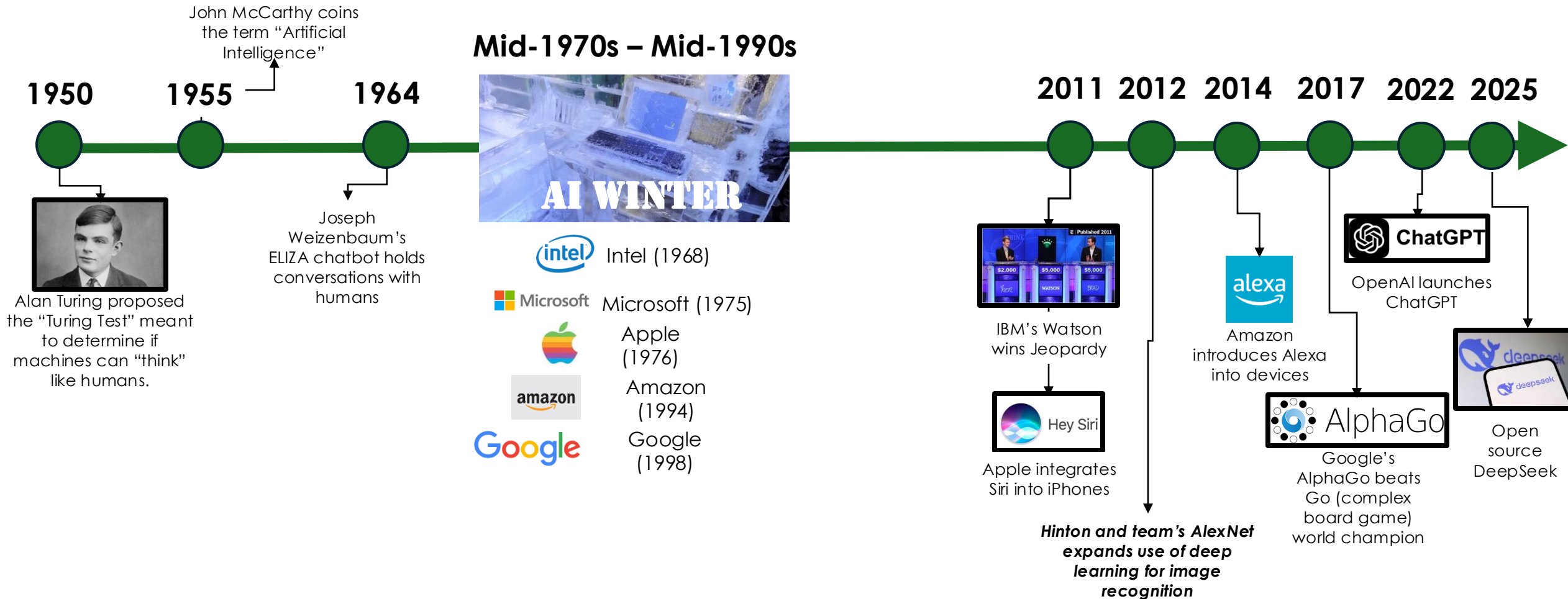
 Microsoft (1975)

 Apple (1976)

 Amazon (1994)

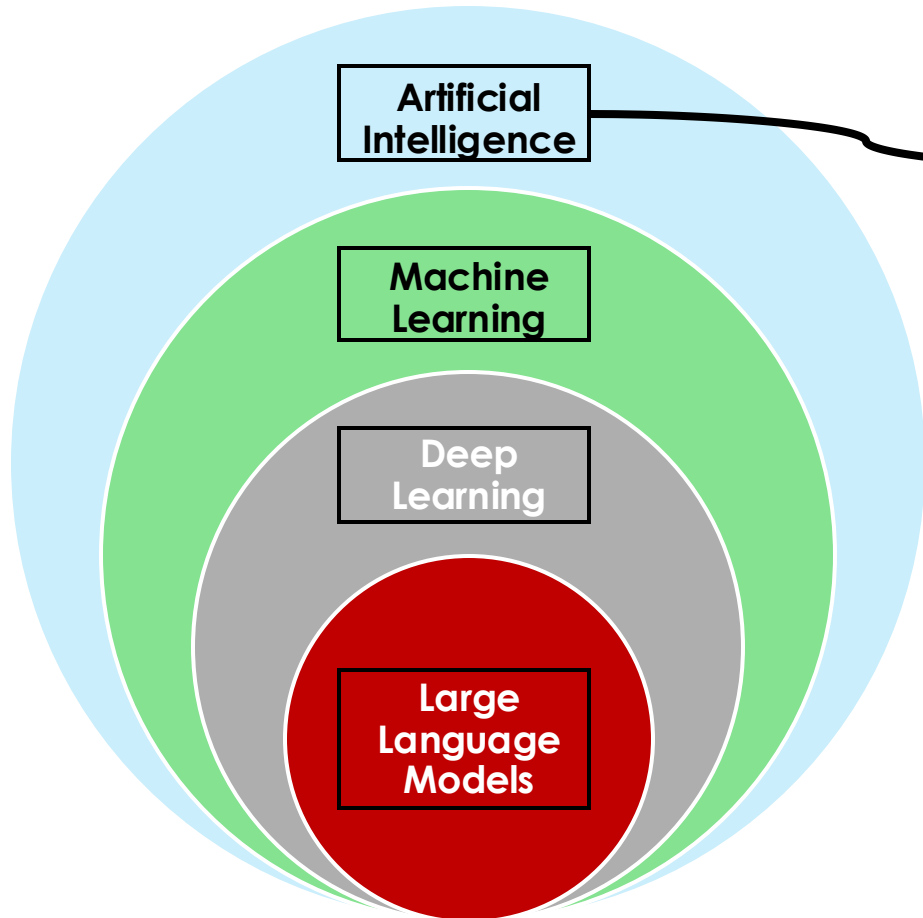
 Google (1998)

# A brief history of Artificial Intelligence (AI)





# Some commonly used buzzwords...



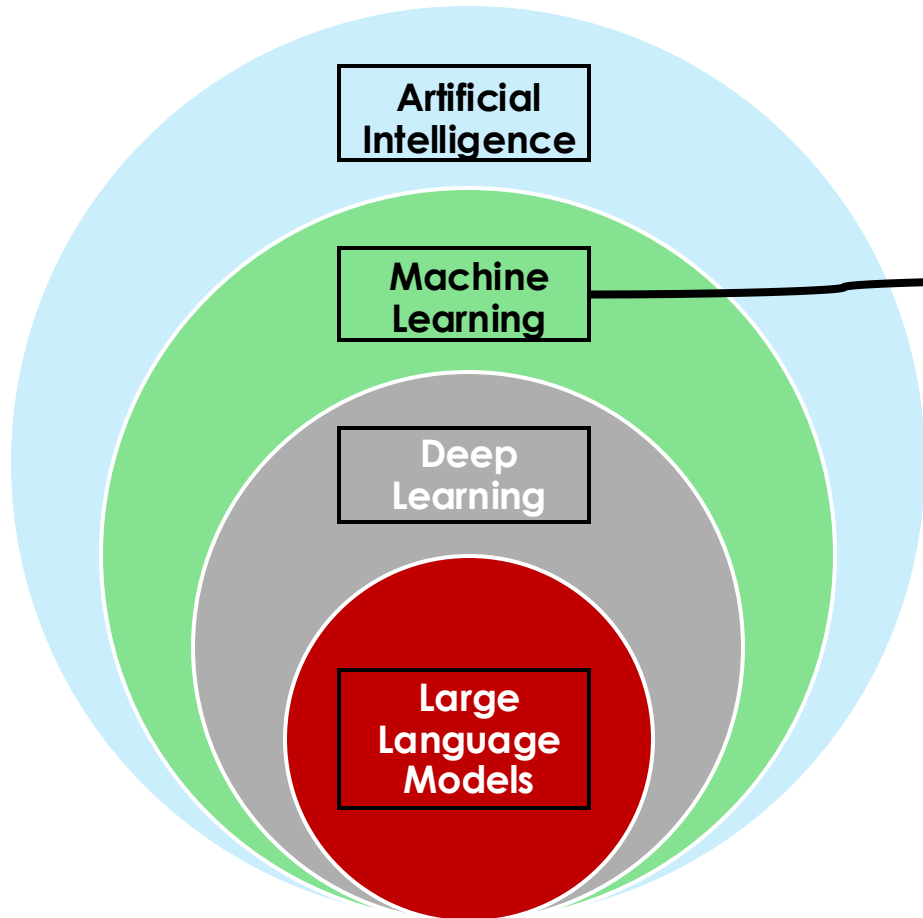
**Artificial Intelligence (AI)** is a branch of computer science that simulates human intelligence using machines or software

**Narrow/Weak AI:** AI systems which can perform a specific task (e.g. Siri, Alexa)

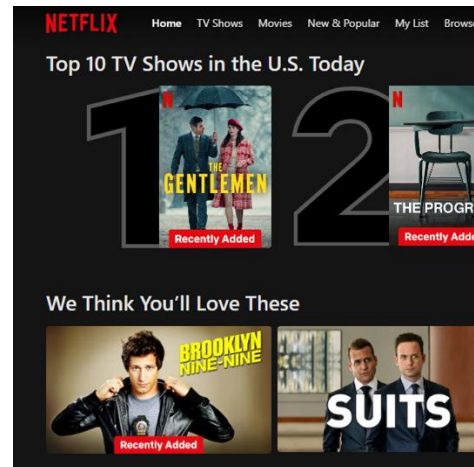
**Generative AI:** AI systems which can *generate* new information based on what it knows (e.g. Claude, ChatGPT)

**Artificial General Intelligence/Strong AI:** AI systems that can perform a broad range of tasks and can reason/problem solve based on what it knows.

# Some commonly used buzzwords...

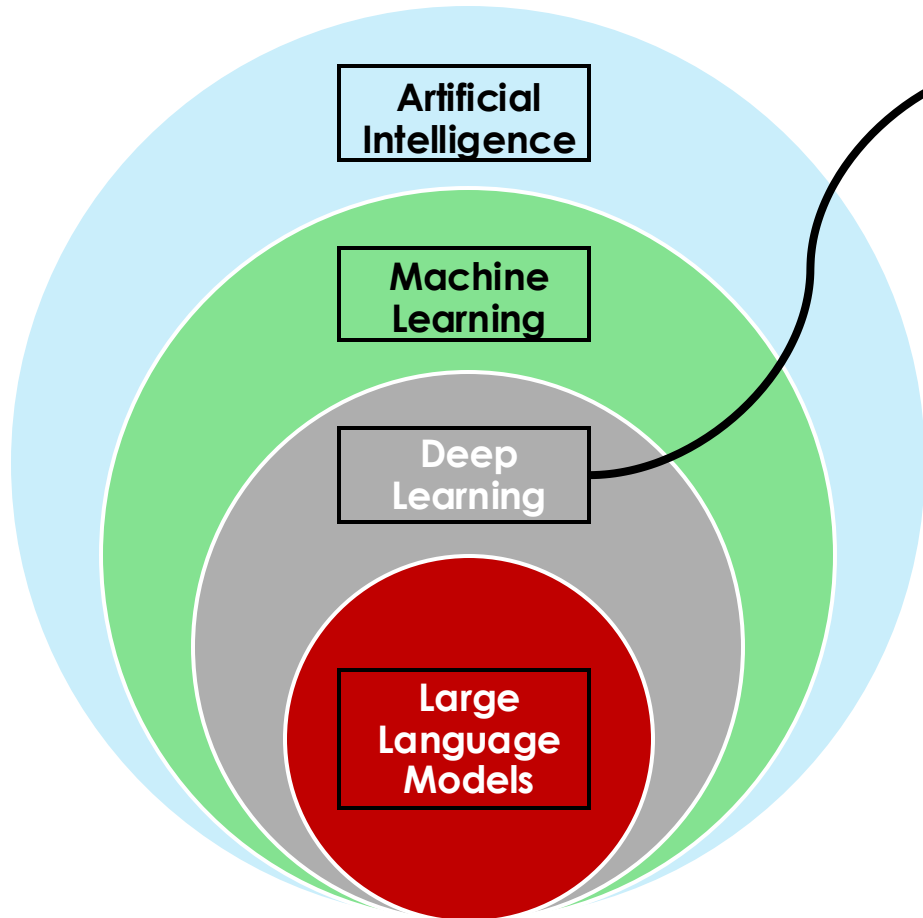


**Machine Learning** involves a machine or software using a process (algorithm) to learn from information (data) and perform complex tasks (e.g. recognize patterns, predict outcomes)



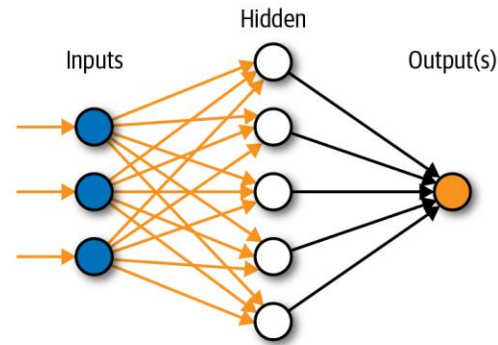
Example:  
Based on what I've watched before, predict what shows I might like

# Some commonly used buzzwords...



**Deep Learning** is a type of machine learning which uses neural networks on large *unstructured* datasets to make connections, find patterns, and make predictions on new data.

Artificial Neural Network



Artificial Neural Network

TECH / ARTIFICIAL INTELLIGENCE / US & WORLD

**Former Go champion beaten by DeepMind retires after declaring AI invincible**

Google Photos

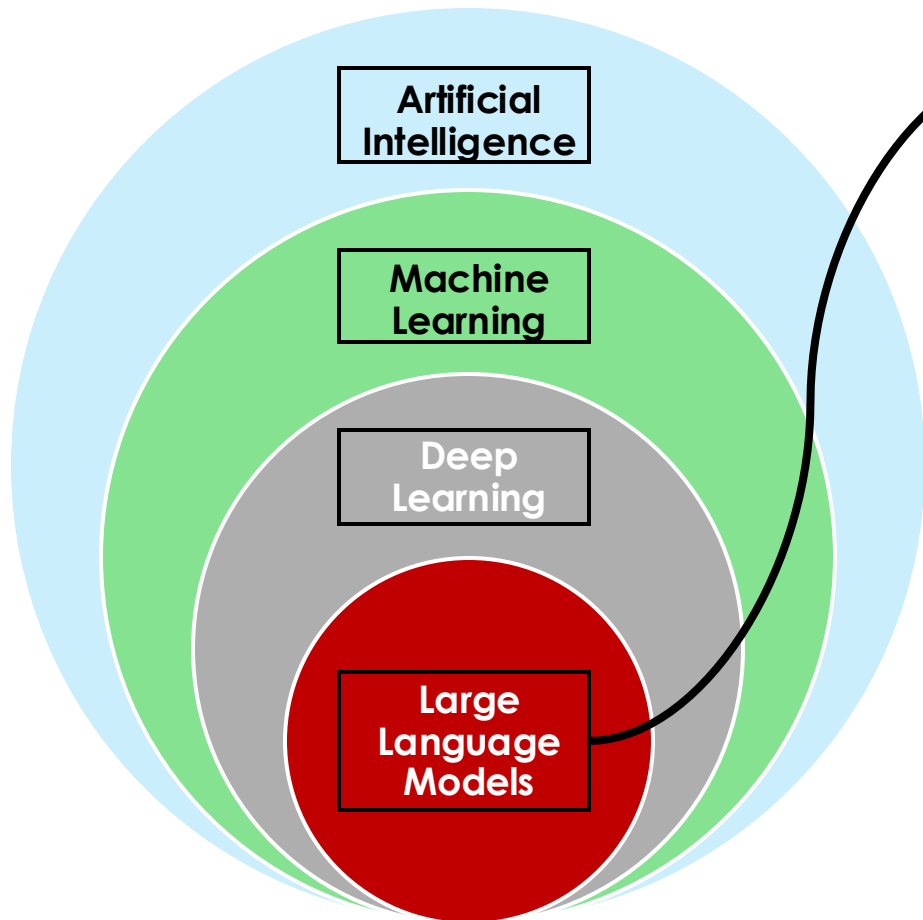
Search tropical beaches



The New York Times

**Computer Wins on 'Jeopardy!': Trivial, It's Not**

# Some commonly used buzzwords...



**Large Language Models** are types of deep learning models focused on *natural language*, which aims to predict and generate human language text



**ChatGPT**

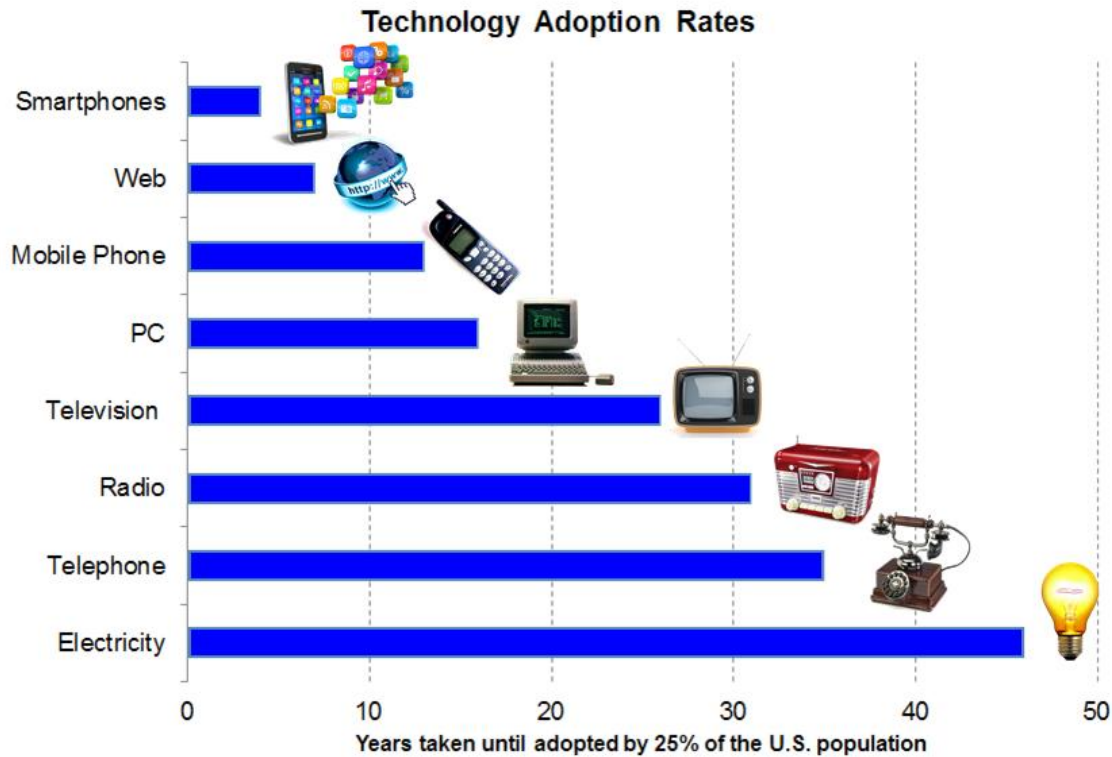
**You**

Tell me tomorrow's weather forecast in New Brunswick NJ, but in the voice of William Shakespeare

**ChatGPT**

In New Brunswick's realm, where Rutgers' towers stand tall, on the morrow, Tuesday, the twelfth of March, the sun shall don its brightest robe, casting rays both warm and broad across the land. The air, as if by gentle hand, shall move from west at twelve miles per hour, with skies mostly clear, save for a few clouds that dare to tread upon the sun's vast stage. From dawn till dusk, from 7:12 in morn to 7:01 in eve, the daylight spans, bearing witness to a play of light and shadow.

# Rates of technology adoption

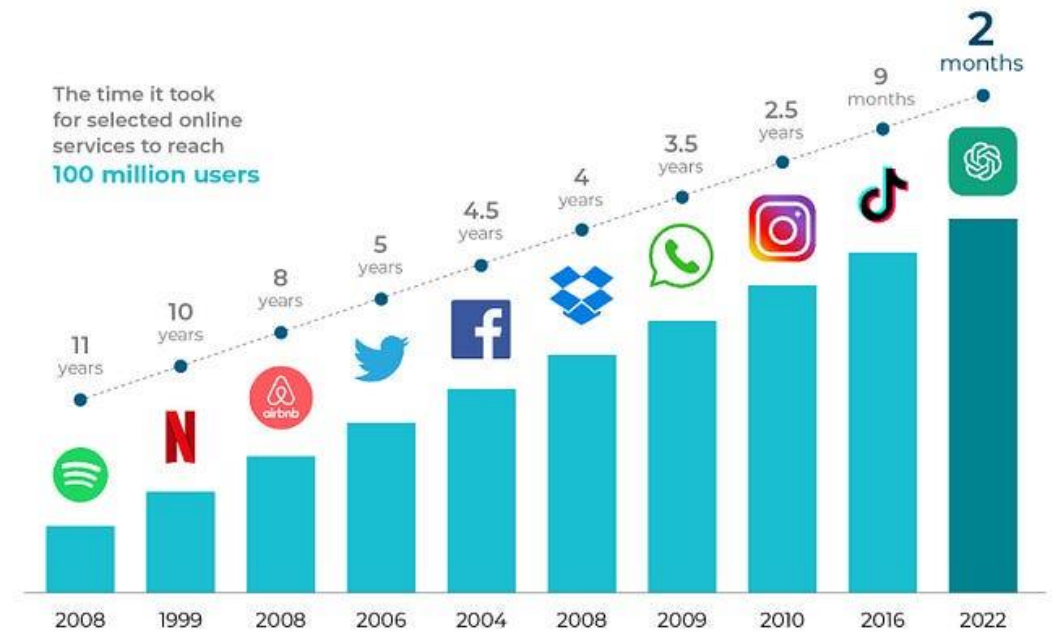


Market Realist

Source: U.S. Census, Wall Street Journal

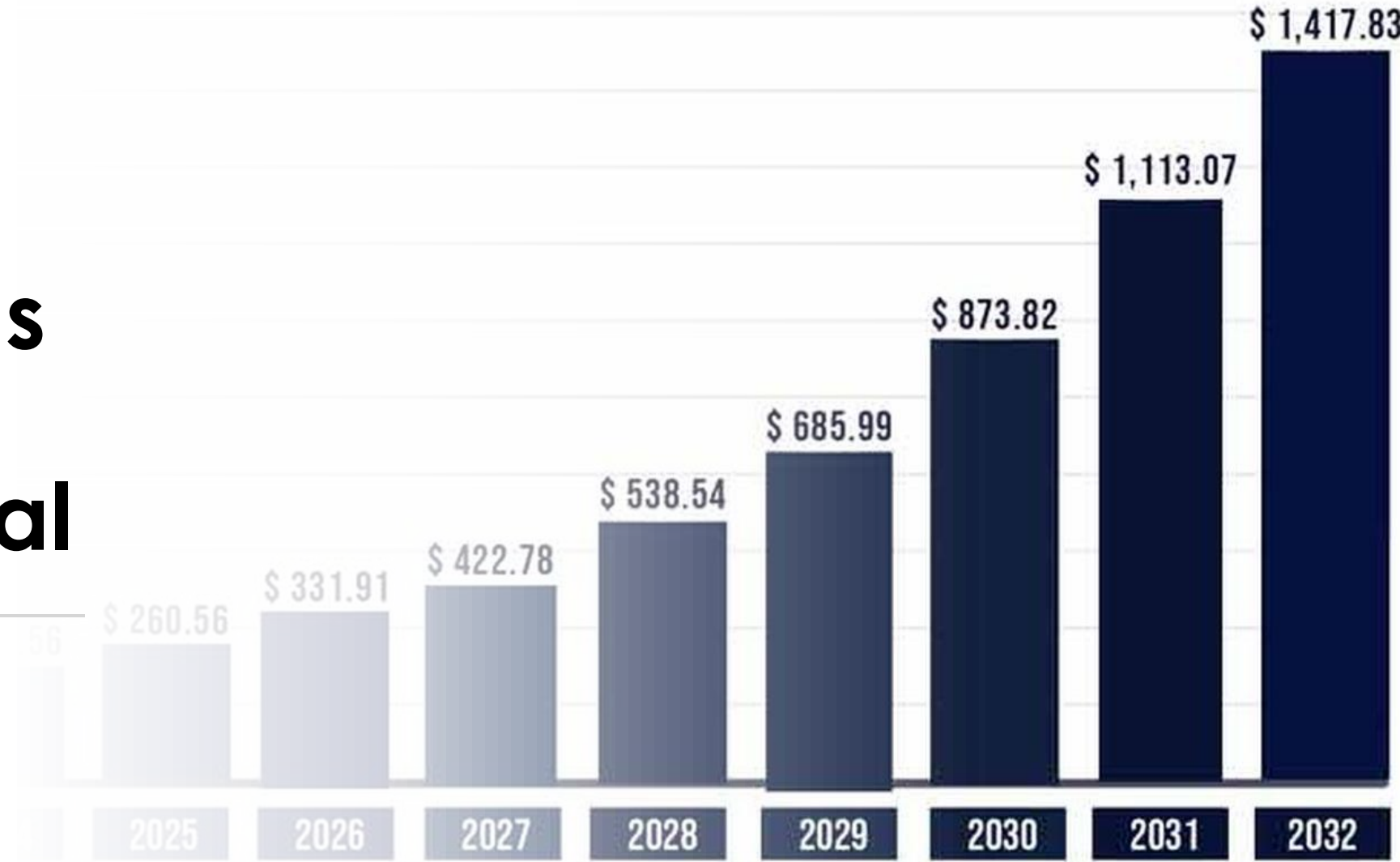
netscribes

## Chat-GPT sprints to 100 million users



Source: World of Statistics

# AI in drug discovery has enormous profit potential





# The transformer network introduced in 2017

## Attention Is All You Need

**Ashish Vaswani\***  
Google Brain  
avaswani@google.com

**Noam Shazeer\***  
Google Brain  
noam@google.com

**Niki Parmar\***  
Google Research  
nikip@google.com

**Jakob Uszkoreit\***  
Google Research  
usz@google.com

**Llion Jones\***  
Google Research  
llion@google.com

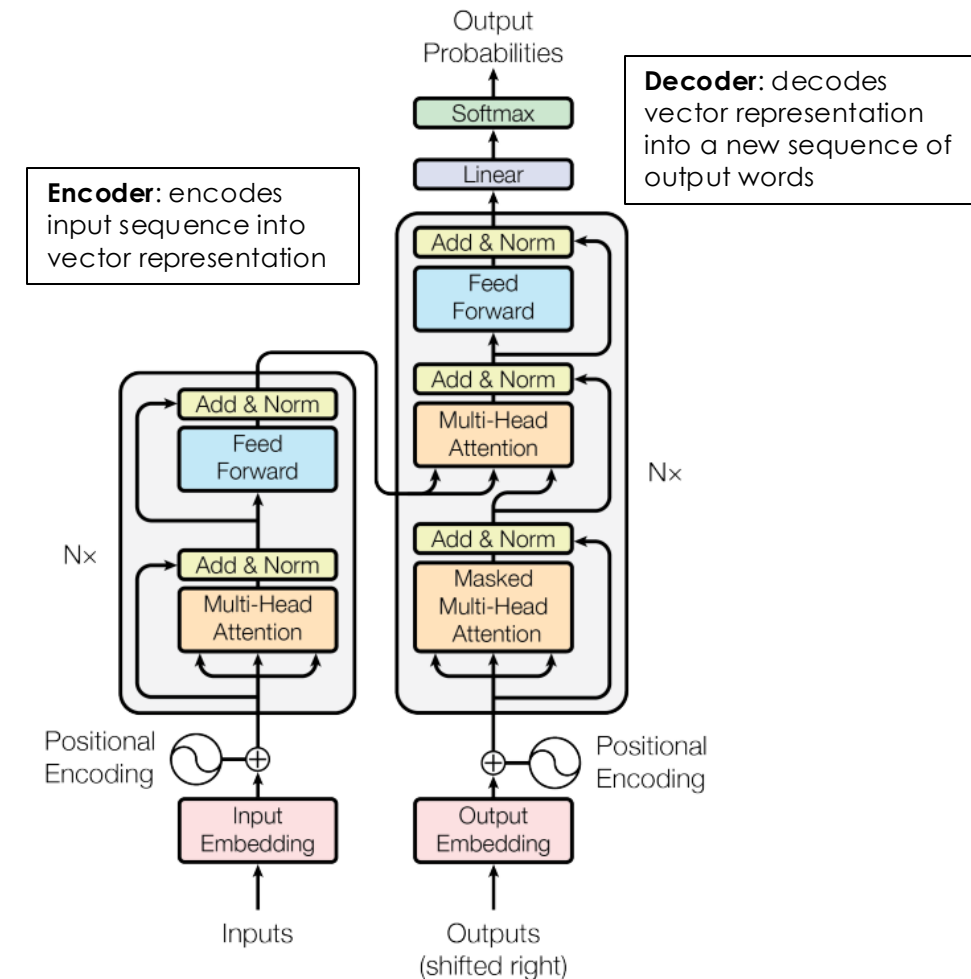
**Aidan N. Gomez\* †**  
University of Toronto  
aidan@cs.toronto.edu

**Lukasz Kaiser\***  
Google Brain  
lukaszkaizer@google.com

**Illia Polosukhin\* ‡**  
illia.polosukhin@gmail.com

### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

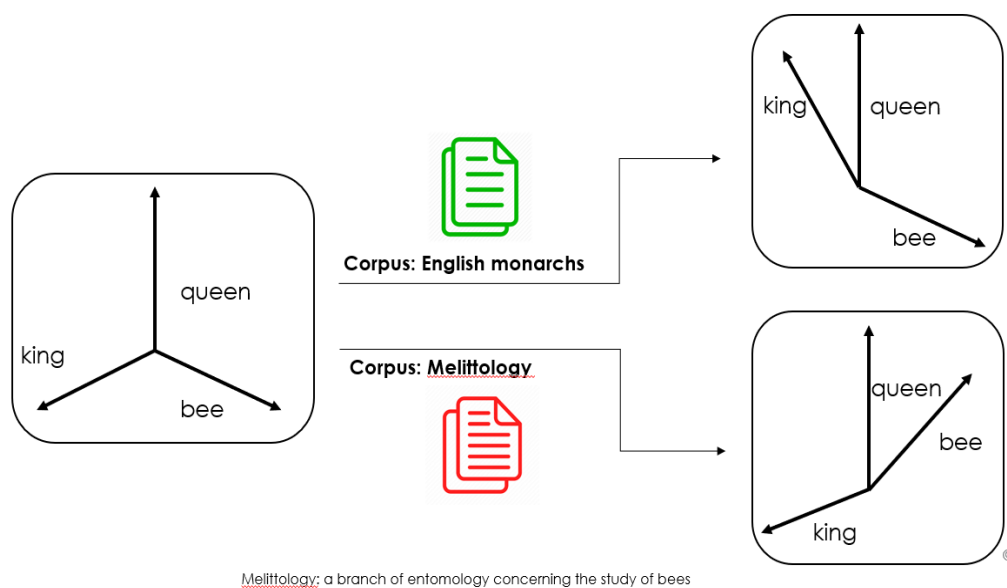


# Large Language Models (LLMs)

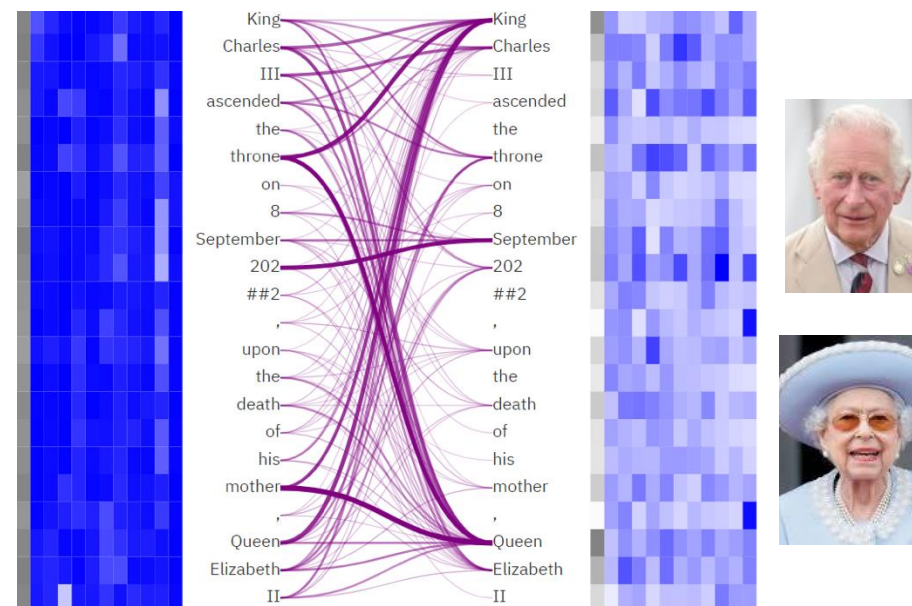


Sentence: “King Charles III ascended the throne on 8 September 2022, upon the death of his mother, Queen Elizabeth II”

Words are related based on document context



Words are related based on sentence context



# A *basic* introduction to LLMs

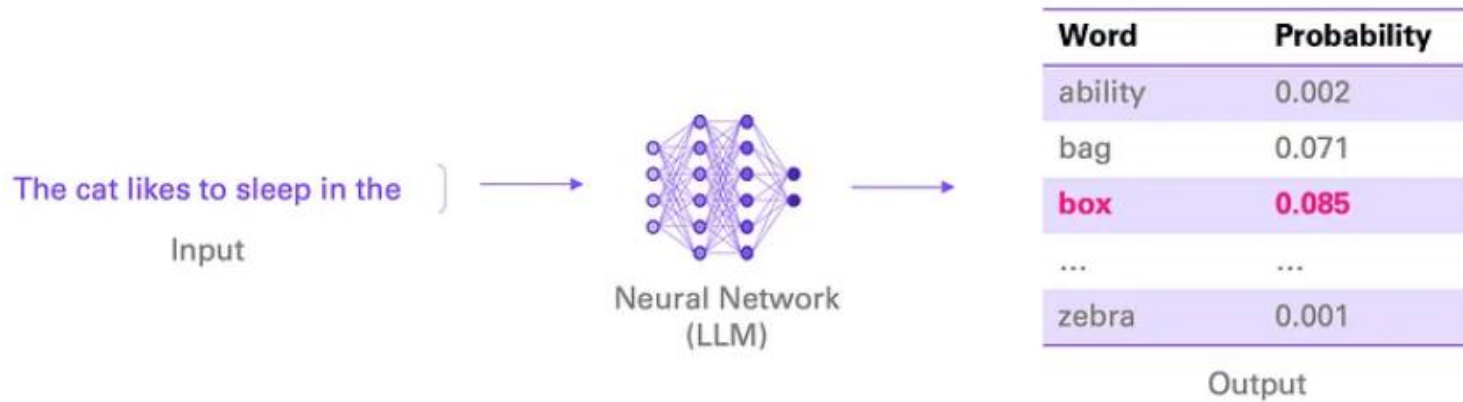
Language models utilize neural networks to *predict* the next word in a sequence

The cat likes to sleep in the --

0.5	0.4	-0.2	0.3	0.4	0.3	0.4
0.9	-0.1	0.6	-0.2	-0.1	-0.2	-0.1
-0.7	0.8	-0.3	0.4	0.8	0.4	0.8
...	...	...	...	...	...	...
-0.3	0.2	-0.7	-0.6	0.2	-0.6	0.2

Each word has a meaning and context, which is represented numerically. The numerical representation is an **embedding**.

# A basic introduction to LLMs

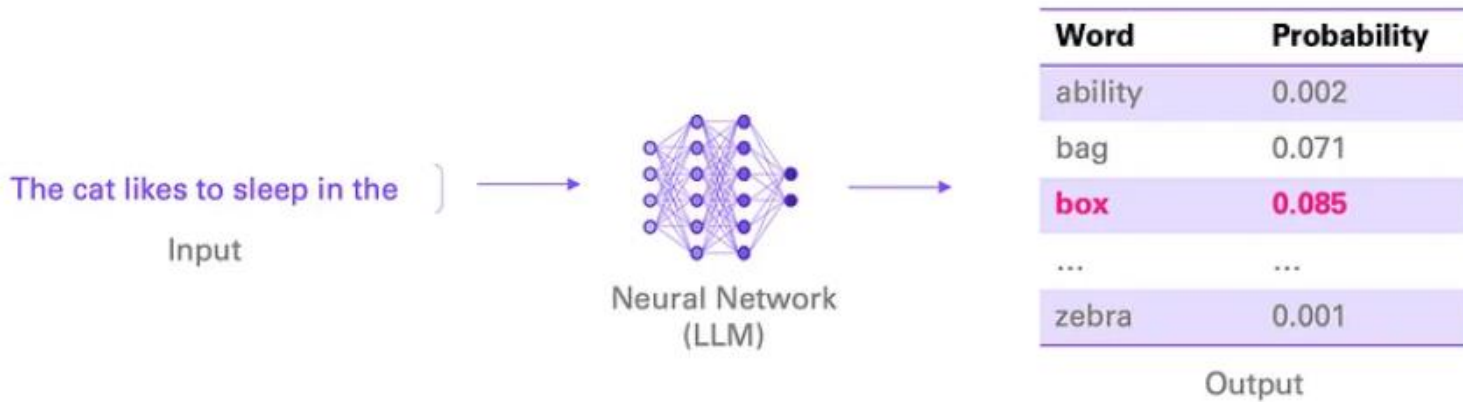


Language modeling is learning to predict the next word.

The Neural Network for LLMs is trained on the text contents of the internet, coding word relationships, context, syntax, and sentiment.

It uses the input embeddings and its neural network to generate the most likely next word in a sequence of words.

# A basic introduction to LLMs



Language modeling is learning to predict the next word.

The Neural Network for LLMs is trained on the text contents of the internet, coding word relationships, context, syntax, and sentiment.

It uses the input embeddings and its neural network to generate the most likely next word in a sequence of words.

We can create **vast amounts of sequences** for training a language model

● Context ● Next Word ● Ignored

- { The **cat** likes to sleep in the }
- { The cat **likes** to sleep in the }
- { The cat likes **to** sleep in the }
- { The cat likes to **sleep** in the }
- { The cat likes to sleep **in** the }

The model predicts the next word irrespective of the length of the phrase

We do the same with much **longer sequences**. For example:

A language model is a probability distribution over sequences of words. [...] Given any sequence of words, the model predicts the **next** ...

The model applies the same process for longer sequences of words.

# Refining LLM responses to scientific Q&A

**Problem: General purpose LLMs (e.g. Bing, ChatGPT) may generate false information when asked specific scientific questions**

- General purpose LLMs are good at generating new content, but it doesn't always have to be *accurate*
- When using LLMs for scientific Q&A or research, they need to be evaluated against *ground truth*
- For technical/scientific questions, how can we check for and improve accuracy of responses?



# Refining LLM responses to scientific Q&A

**Problem: General purpose LLMs (e.g. Bing, ChatGPT) may generate false information when asked specific scientific questions**

**Approach: Use Retrieval Augmented Generation (RAG) to source information from specific scientific literature providing source content for LLMs**

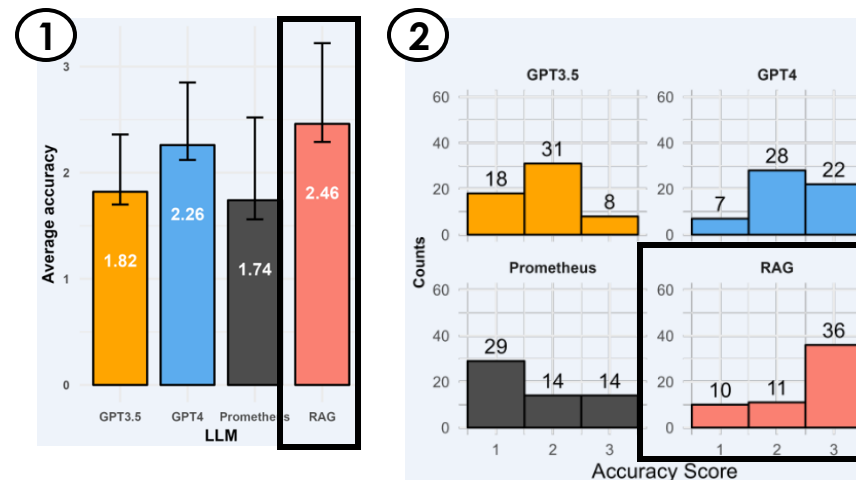
- General purpose LLMs are good at generating new content, but it doesn't always have to be *accurate*
- When using LLMs for scientific Q&A or research, they need to be evaluated against *ground truth*
- For technical/scientific questions, how can we check for and improve accuracy of responses?

**Improving accuracy of GPT-3/4 results on biomedical data using a retrieval-augmented language model**

David Soong\*<sup>1</sup>, Sriram Sridhar\*<sup>1</sup>, Han Si<sup>1</sup>, Jan-Samuel Wagner<sup>1</sup>, Ana Caroline Costa Sá<sup>1</sup>, Christina Y Yu<sup>1</sup>, Kubra Karagoz<sup>1</sup>, Meijian Guan<sup>1</sup>, Hisham Hamadeh<sup>1</sup>, Brandon W Higgs<sup>1</sup>

<sup>1</sup>Data Sciences, Genmab, Princeton, NJ

\*Equal contribution



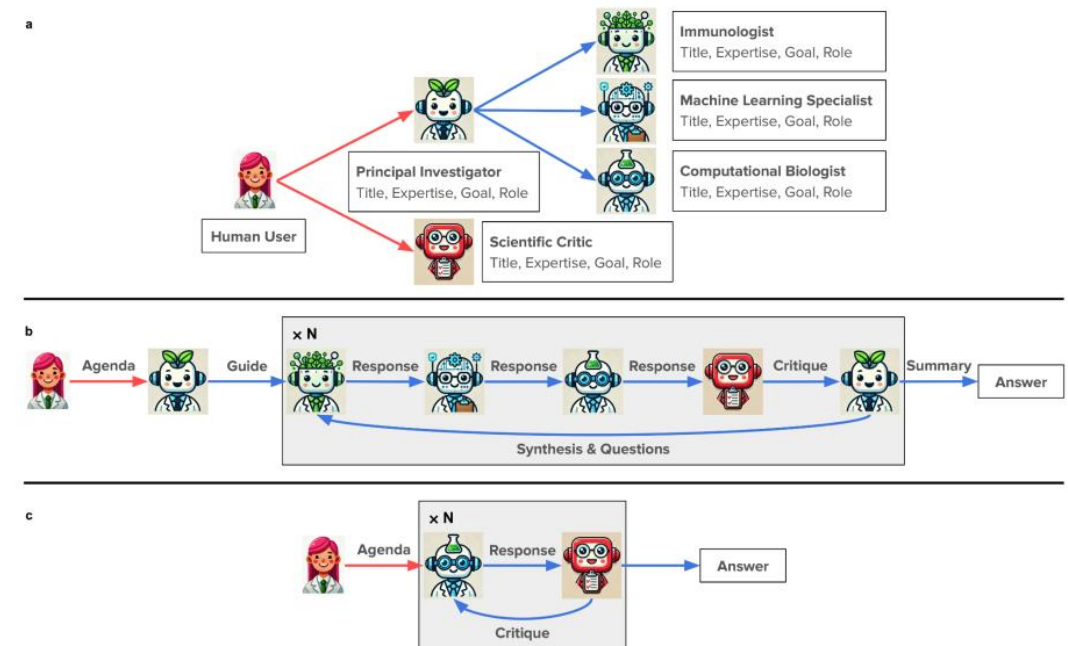
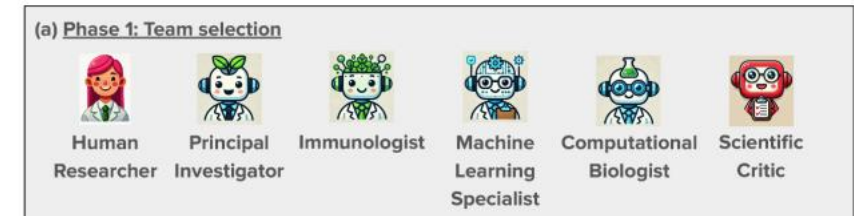
When testing this approach to Q&A for DLBCL (type of lymphoma), RAG approach had better overall accuracy (1) and more precise responses to questions (2)



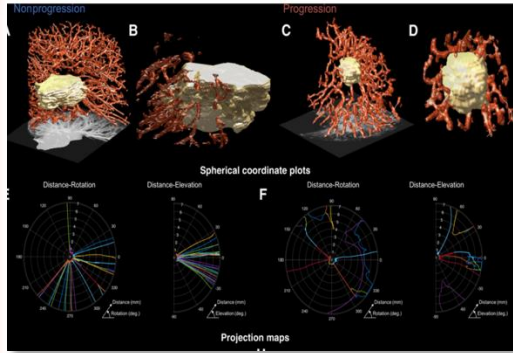
# AI agents

LLM roles working together to solve a problem

- An AI agent is a software program that perceives its environment, makes decisions, and acts to achieve goals
- Can be reactive, learning-based, or autonomous, depending on their intelligence level
- Process data, predict outcomes, automate tasks, and improve decision-making
- Each agent has a distinct role; can combine distinct roles to solve a problem efficiently

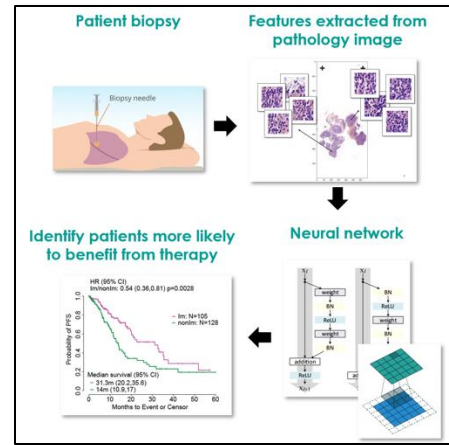


# AI applications at Genmab



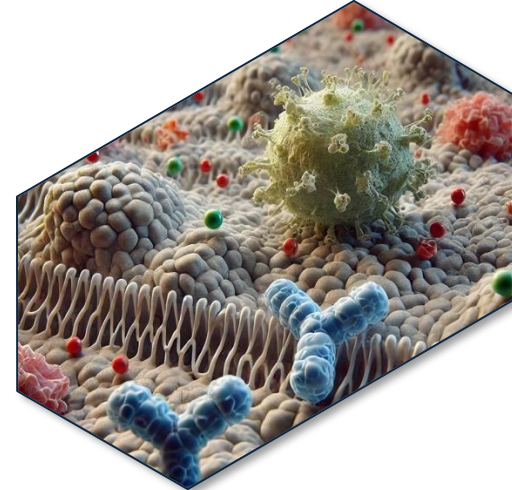
## Radiomics

Use CT patient scans to identify those who will respond to treatment



## Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



## Drug combinations

Predicting synergistic drug combinations in oncology



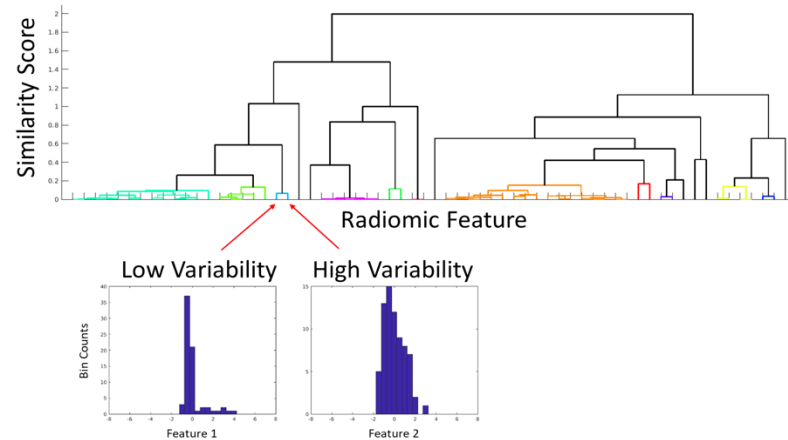
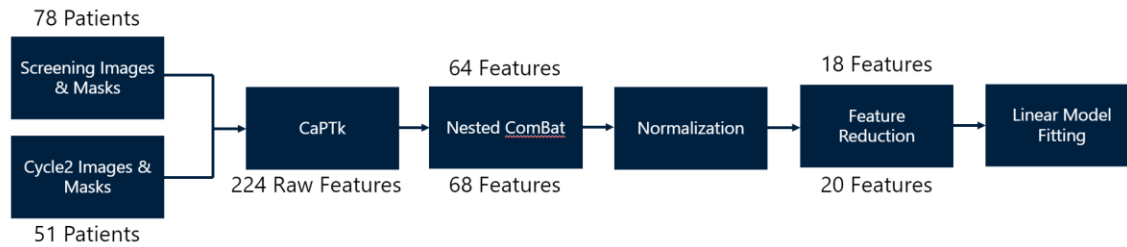
## Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics

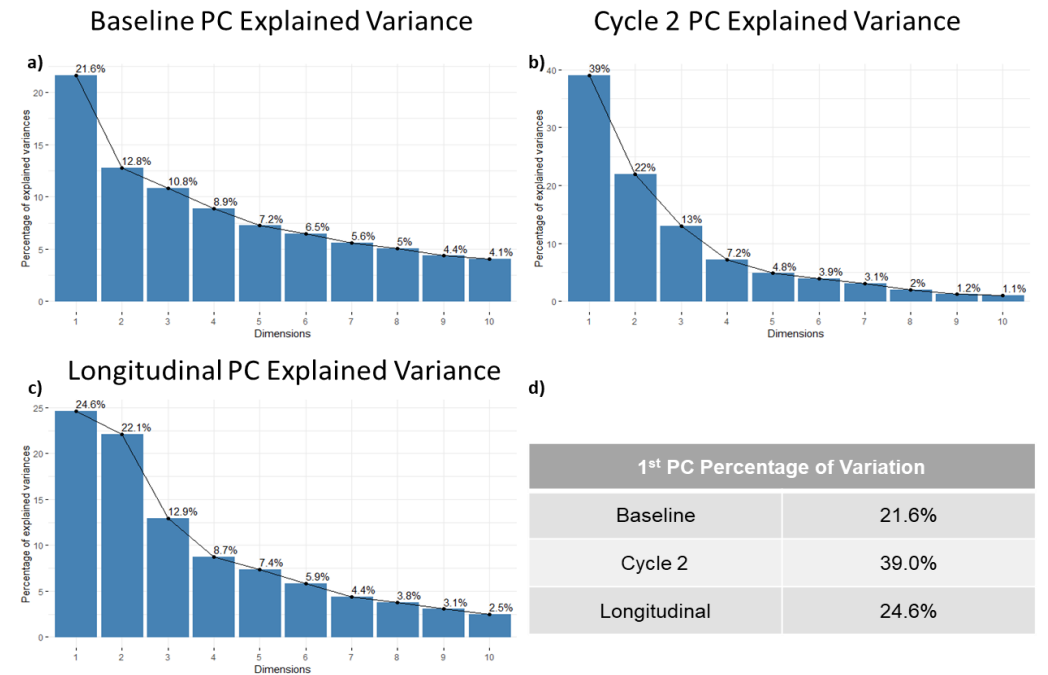
# Evaluating the potential of radiomics features to predict outcome to an antibody-drug conjugate therapy in NSCLC



## Data flow from raw images to model build

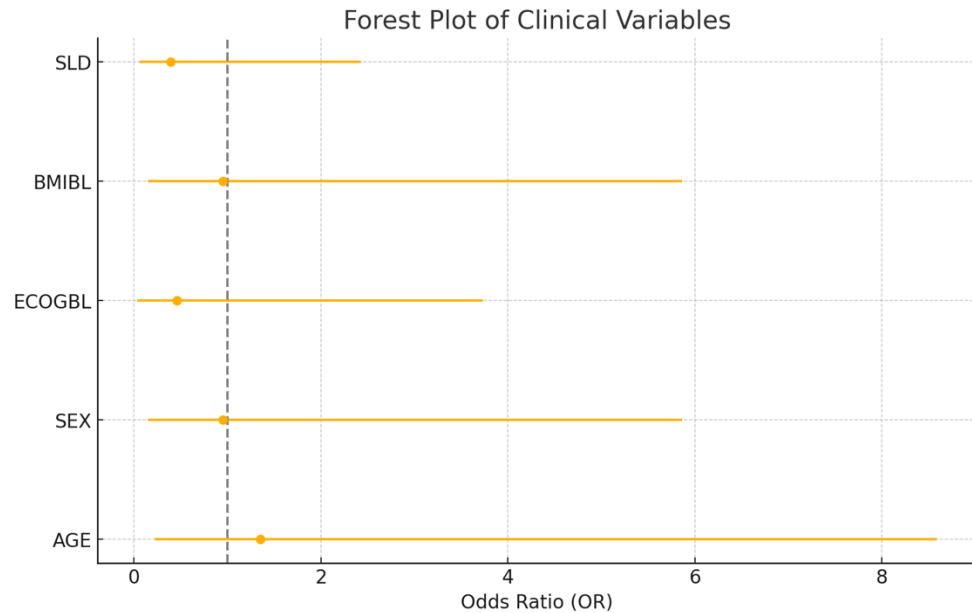


## Radiomics feature explained variability

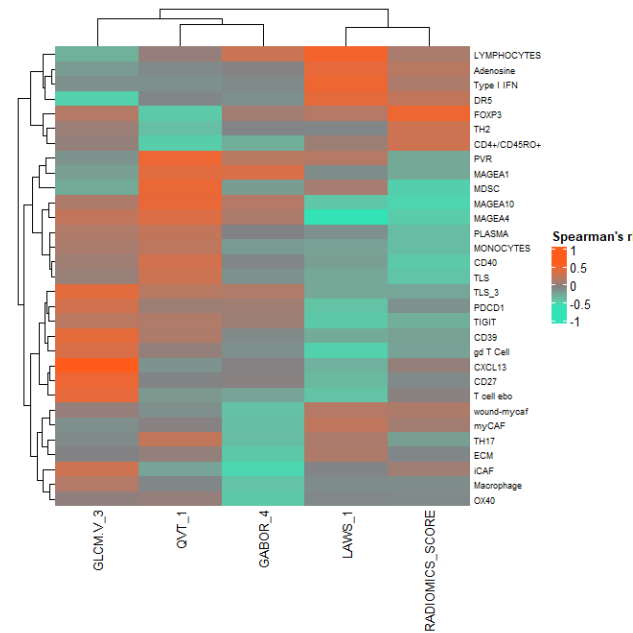


# Molecular profiling of NSCLC patients associates CT-based radiomic features with signaling pathways indicating a favorable tumor microenvironment for treatment with immunotherapy

Radiomics signature is independent of clinical factors



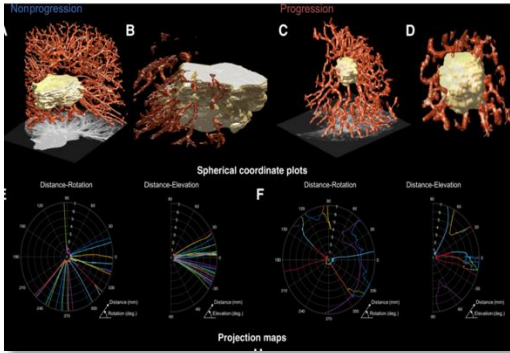
Radiomics features associated with mechanism



Radiomics Feature	Biological Functional Classifications
QVT_1	Immune Suppression, Tumor-associated Antigens, Cell Adhesion and Migration, Hypoxia and Angiogenesis
GLCM.V_3	B Cell Function, Immune Checkpoint Regulation, Immune Evasion and Suppression, Immune Activation and Inflammation, Cancer-Associated Fibroblasts
LAWS_1	Immune Response and Activation, Immune Suppression, Apoptosis, Cell Cycle and Proliferation, Angiogenesis
GABOR_4	Tumor-associated Antigens, Immune Regulation, Cell Cycle and Proliferation, Immune Response and Activation
RADIOMIC S_SCORE	Regulatory T Cells and Immune Suppression, Immune Regulation, Antigen Presentation, Apoptosis, Cell Cycle and Proliferation

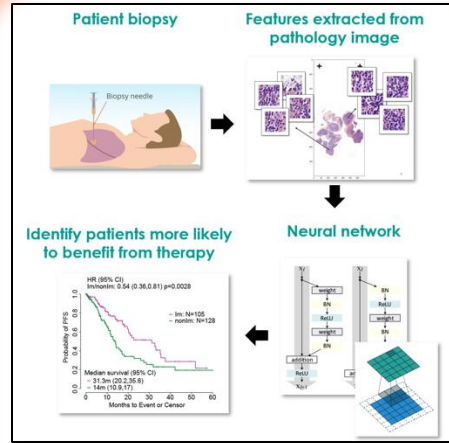


# AI applications at Genmab



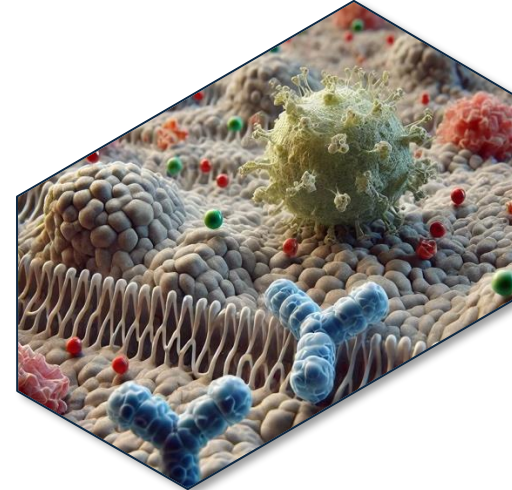
## Radiomics

Use CT patient scans to identify those who will respond to treatment



## Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



## Drug combinations

Predicting synergistic drug combinations in oncology



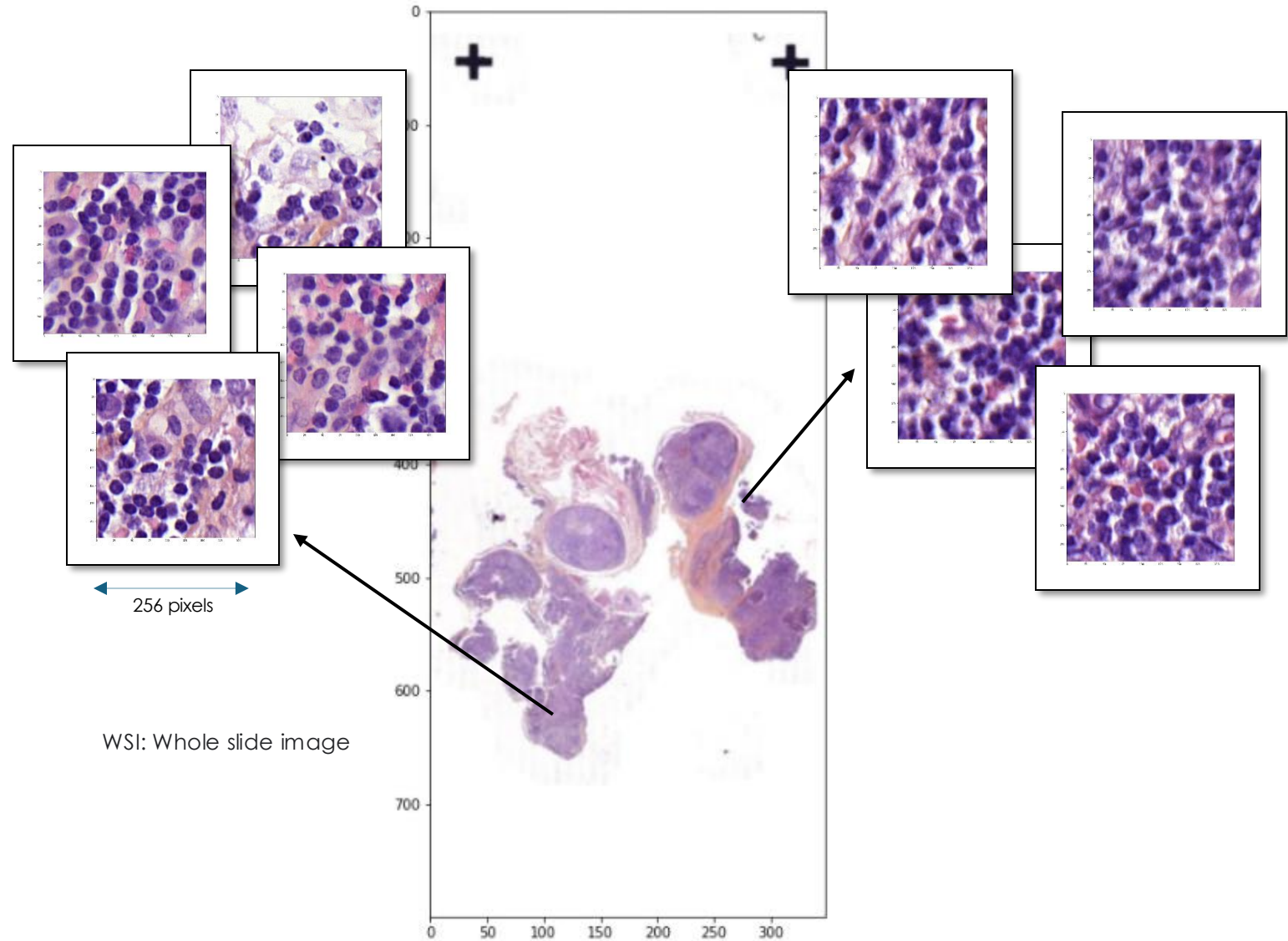
## Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics



# Tile/Patch extraction from whole slide image

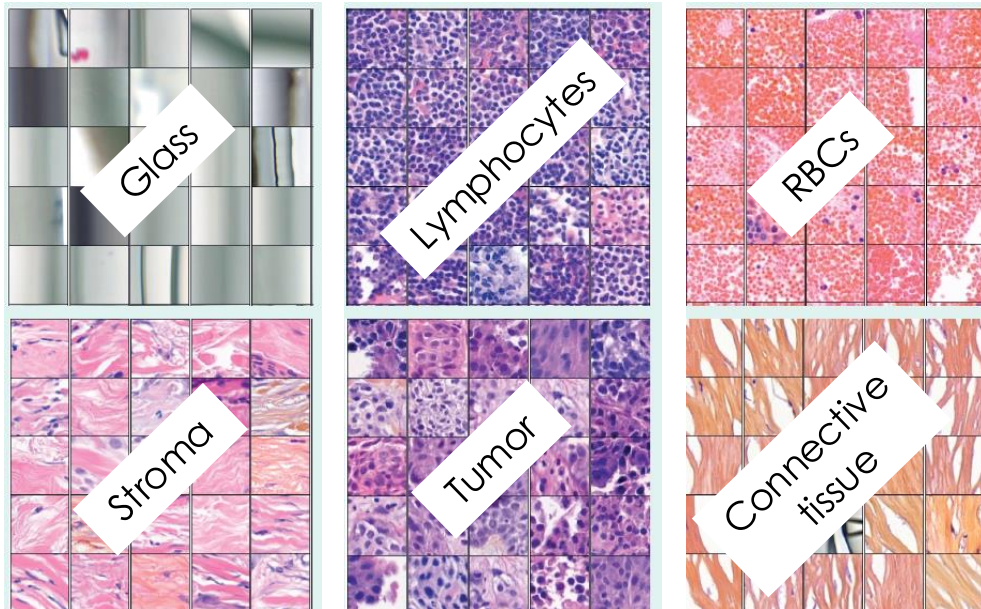
- Tissue contexture is heterogeneous and varies by region
  - WSI is prohibitively large search space
- Can divide the WSI into thousands of small image tile representations
- Use these tiles as micro-representations of the tissue
- Size of WSI versus tile
  - Single WSI: 1-3 GB
  - Tile size:  
 $[(256 \times 256 \text{ pixels}) \times 24 \text{ bits}] / 8 = 197 \text{ KB}$



# Automated histologic, morphologic, and artifact detection

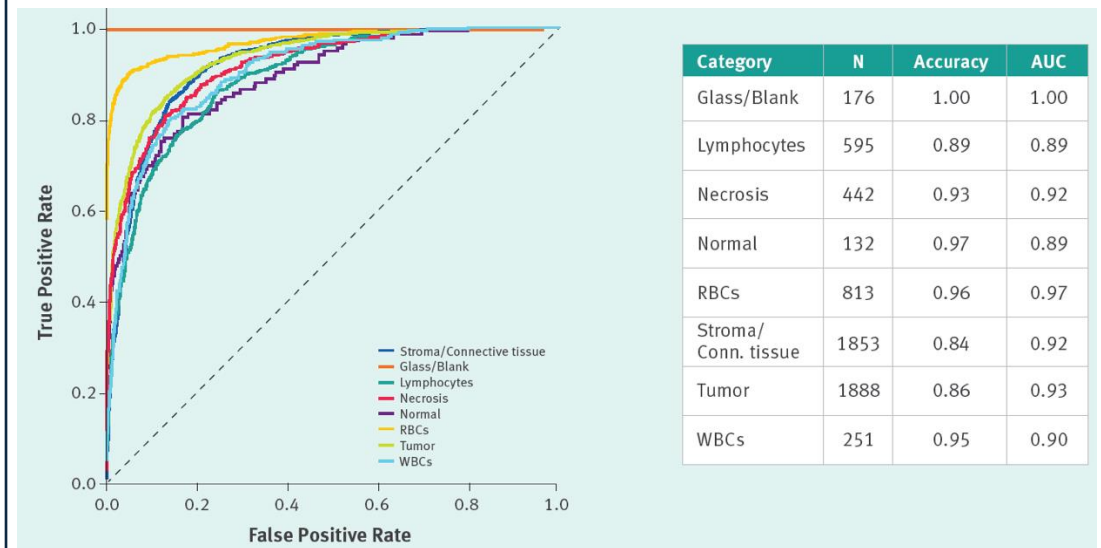
*ResNet-50 CNN used to learn pathologists' annotations*

Example image tiles with different morphologic features



Soong et al, SITC 2021

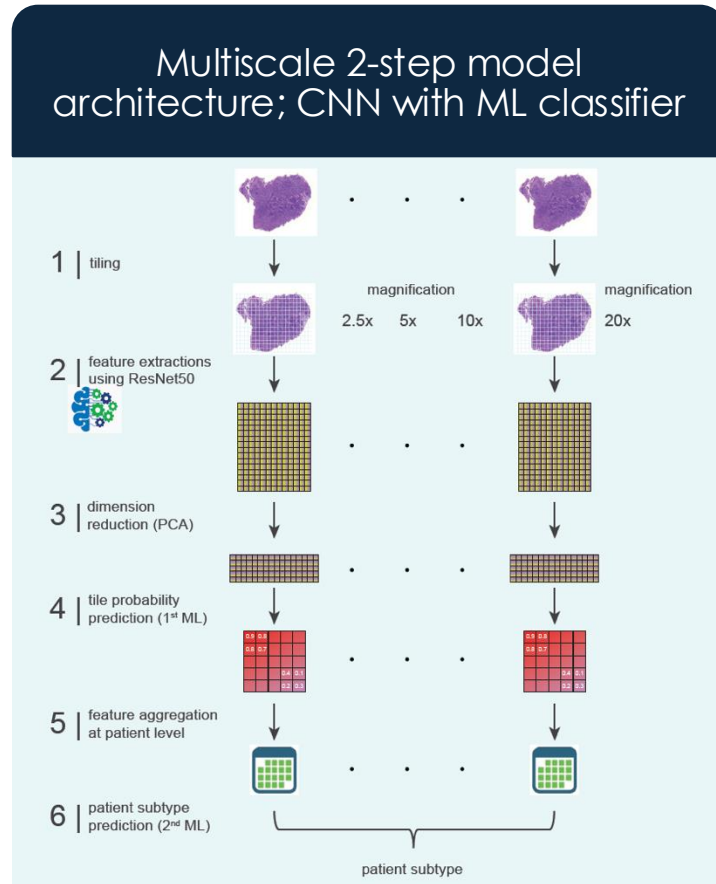
Deep neural network classifier performance (five-fold cross validation)



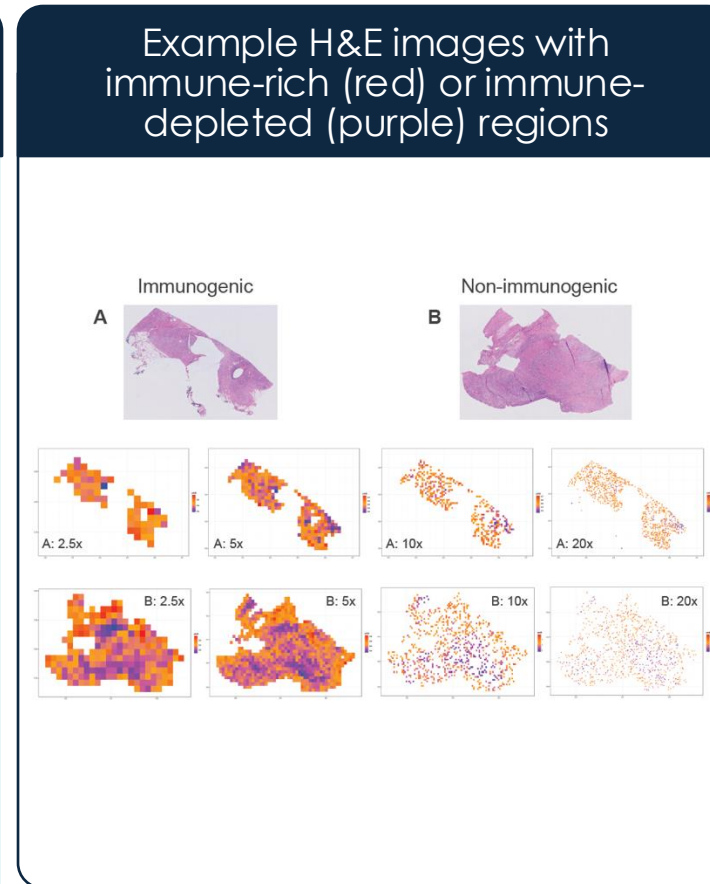
H&E: Hematoxylin-and-Eosin (stain)

# Image-based prediction of immunogenic subtype in pancreatic cancer

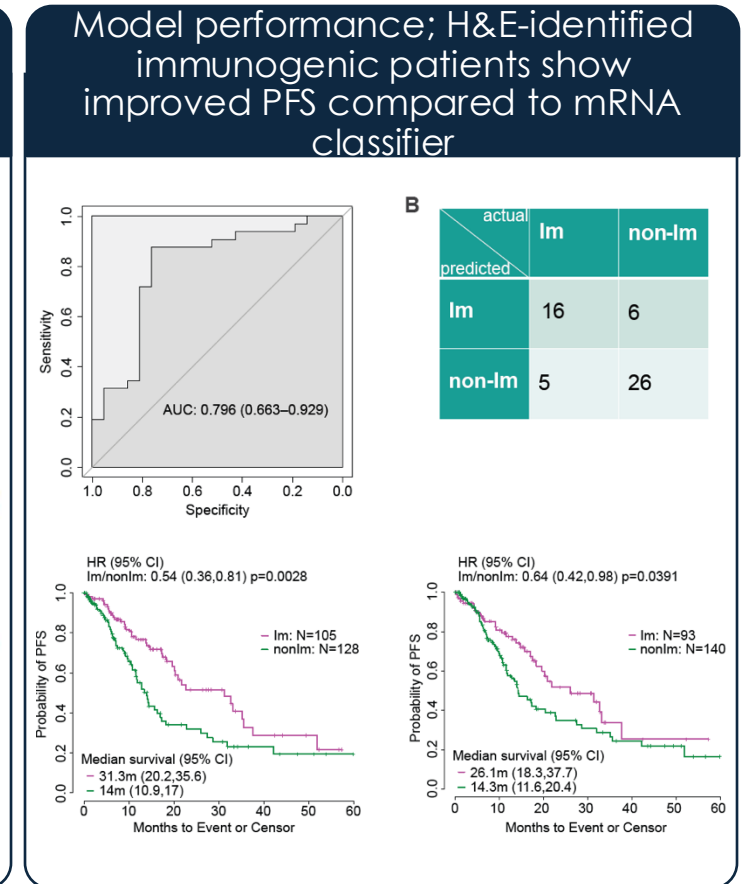
- Identifying patients with immunogenic phenotype using whole slide H&E images
- ResNet-50 CNN architecture implemented



Si et al, AACR 2022

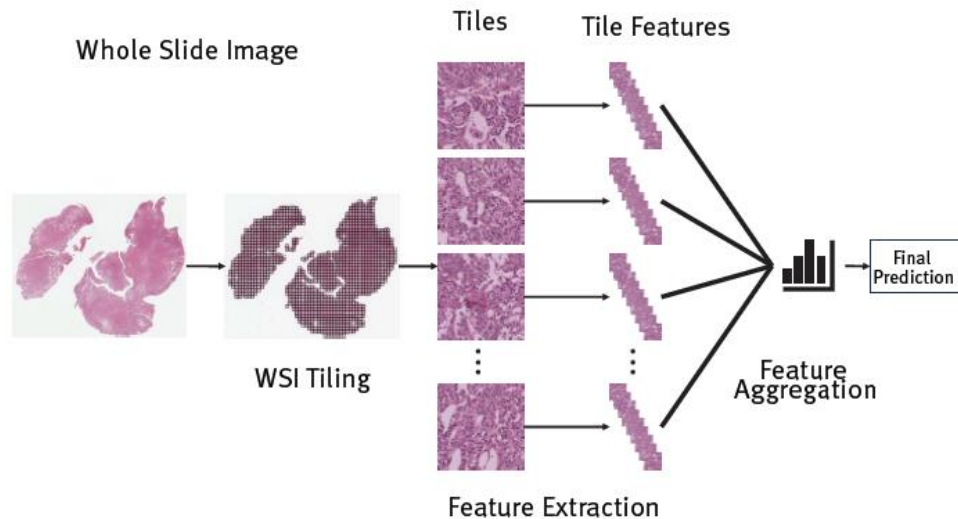


H&E: Hematoxylin-and-Eosin (stain)

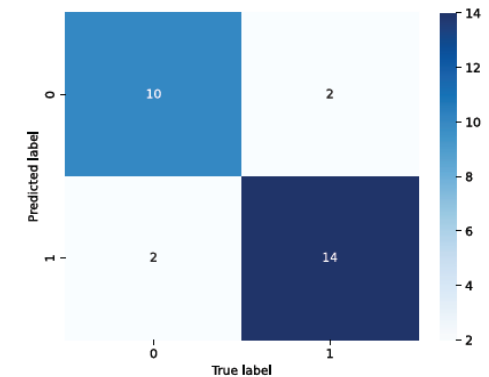
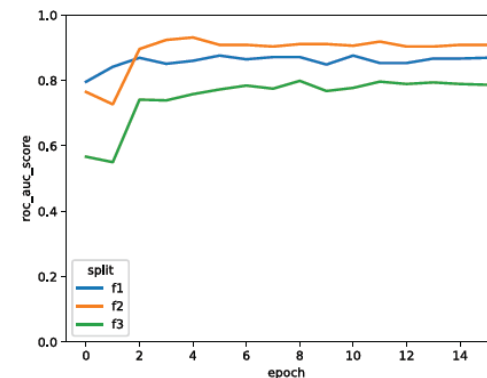


# Identification of tertiary lymphoid structures (TLS) from H&E images in pancreatic cancer using a deep learning approach

A pretrained transformer-based unsupervised contrastive learning model was used to extract 768 image features from each tile



TLS identification achieved an aAUC of 0.86 with the pathologists-generated TLS status as ground truth (N=130)

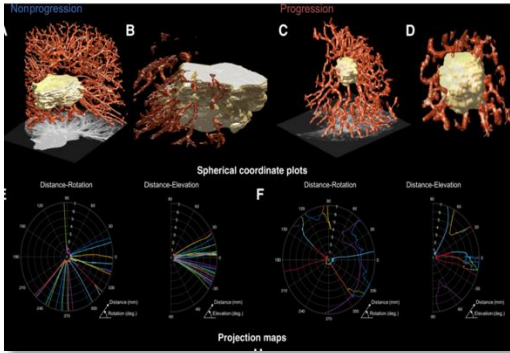


WSI: Whole slide image

H&E: Hematoxylin-and-Eosin (stain)

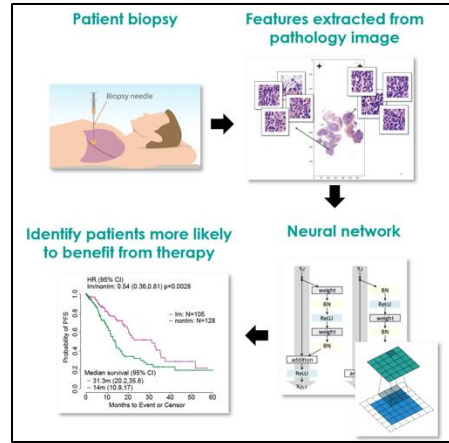


# AI applications at Genmab



## Radiomics

Use CT patient scans to identify those who will respond to treatment



## Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



## Drug combinations

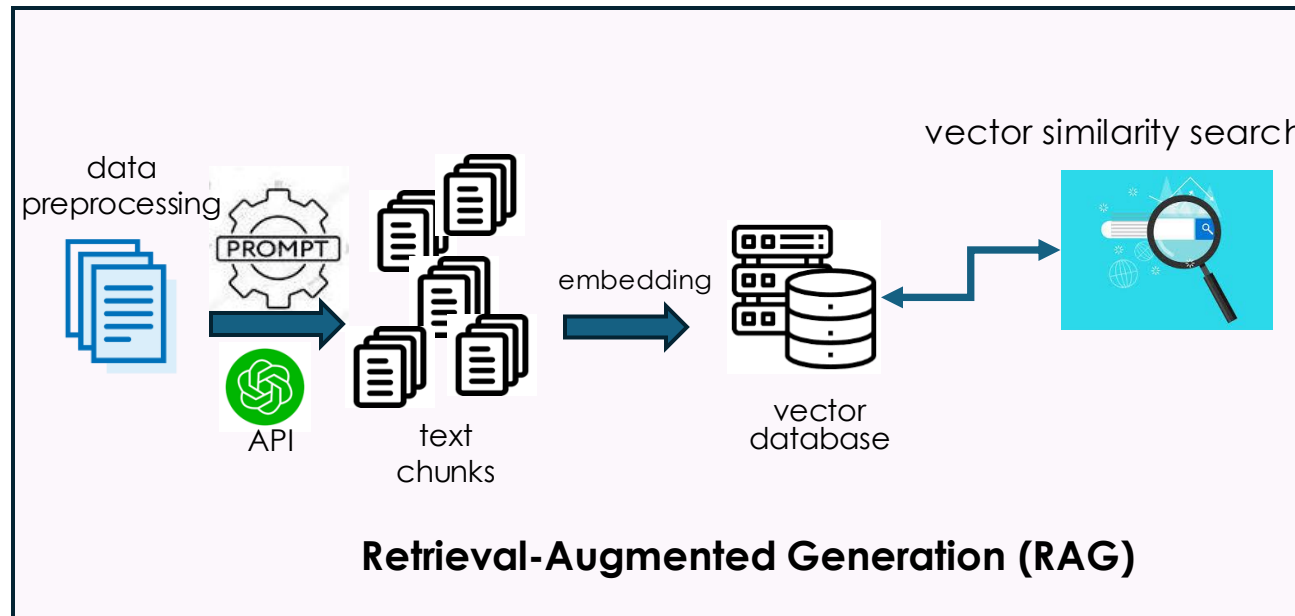
Predicting synergistic drug combinations in oncology



## Prompt-driven analysis

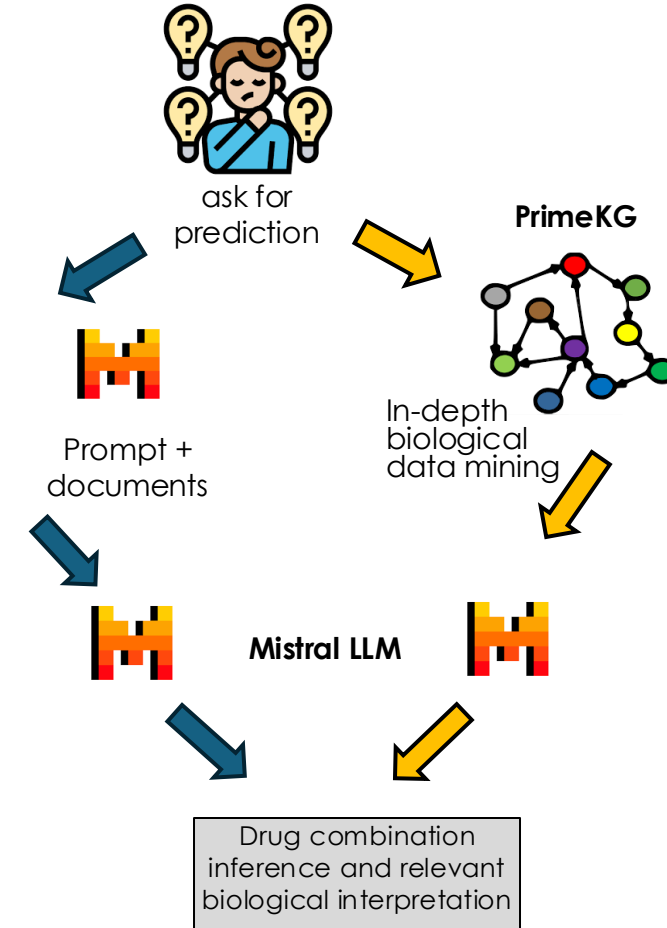
Biomarker and clinical correlative analysis using generative AI to derive code and graphics

# LLM with retrieval augmentation and a knowledge graph to predict drug combinations



Soong et al, PLoS Digital Health 2024

Query: Protein/Gene targets,  
Drugs, Indications



Si et al, AACR 2025



# Explainability test case unseen to model: true positive

ADC-Targeted drug combination



## Explanation of prediction

The synergistic effect of Mirvetuximab soravtansine (Drug A) and Bevacizumab (Drug B) in treating platinum-resistant ovarian cancer is predicted to be positive due to their distinct mechanisms of action targeting different molecular pathways involved in cancer progression. Mirvetuximab soravtansine targets FOLR1 and tubulin, which are essential for cell division and platinum resistance in ovarian cancer. FOLR1 overexpression is linked to platinum resistance, and tubulin is a component of microtubules, which are crucial for mitosis and meiosis. By targeting FOLR1, Mirvetuximab soravtansine can disrupt the folate metabolism pathway and inhibit cell division, making the cancer cells more susceptible to chemotherapy. Simultaneously, Bevacizumab targets VEGF, a protein that plays a crucial role in angiogenesis, the process of forming new blood vessels. VEGF is overexpressed in many cancers, including ovarian cancer, and contributes to tumor growth, progression, and metastasis. By inhibiting VEGF, Bevacizumab can prevent the formation of new blood vessels, starving the tumor of nutrients and oxygen, and enhancing the efficacy of chemotherapy. The combination of these two drugs can lead to a synergistic effect by targeting different molecular pathways and enhancing the overall therapeutic effect against platinum-resistant ovarian cancer.



Drug combo: Mirvetuximab soravtansine, Bevacizumab



Targets: FOLR1 + tubulin, VEGFA

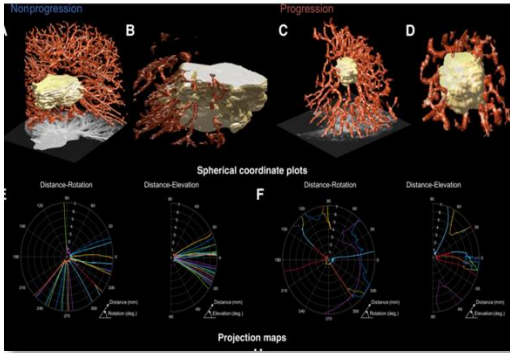


Disease: Platinum-resistant ovarian cancer



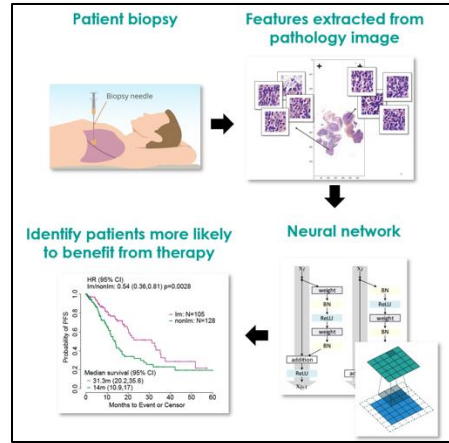
Inferred synergy: Yes  
(True positive)

# AI applications at Genmab



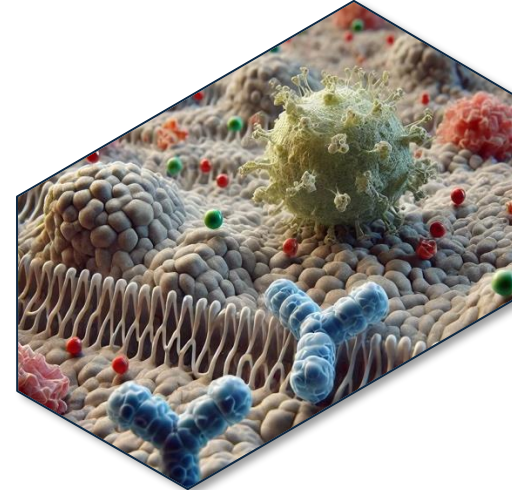
## Radiomics

Use CT patient scans to identify those who will respond to treatment



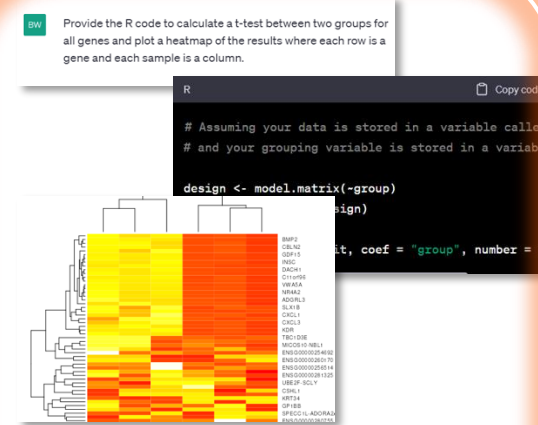
## Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



## Drug combinations

Predicting synergistic drug combinations in oncology



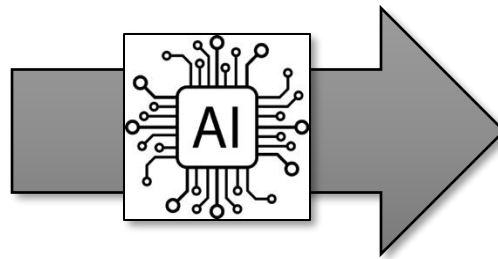
## Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics

# Writing prompts to create programming code and graphics

can you provide R code to calculate this using the method in the uploaded paper? BH

Sure, here is some R code to calculate the sample size for the example I walked through, using the approach from the paper you provided:



```
r Copy code  
  
# Required packages  
library(stats)  
  
# Fatigue effect size  
p_fatigue = 0.6 # Proportion  
  
# Pain effect size  
percent_pain = 0.1 # 10% dec  
p0_pain = 0.5  
p1_pain = p0_pain * (1 - per  
  
# Null probabilities  
p0_fatigue = 0.5  
p0_pain = 0.5  
  
# Alternative probabilities  
p1_fatigue = p_fatigue  
p1_pain = p0_pain * (1 - percent_pain)  
  
# Effect sizes  
delta_fatigue = p1_fatigue - p0_fatigue  
delta_pain = p1_pain - p0_pain
```

The graph shows the relationship between power and intervention sample size for a control sample size of 100. The x-axis represents the intervention sample size (5 to 40), and the y-axis represents power (0.2 to 1.0). Six lines represent different scenarios based on control sample size (N=100), odds ratio, and intervention proportion (Int. Prop.).

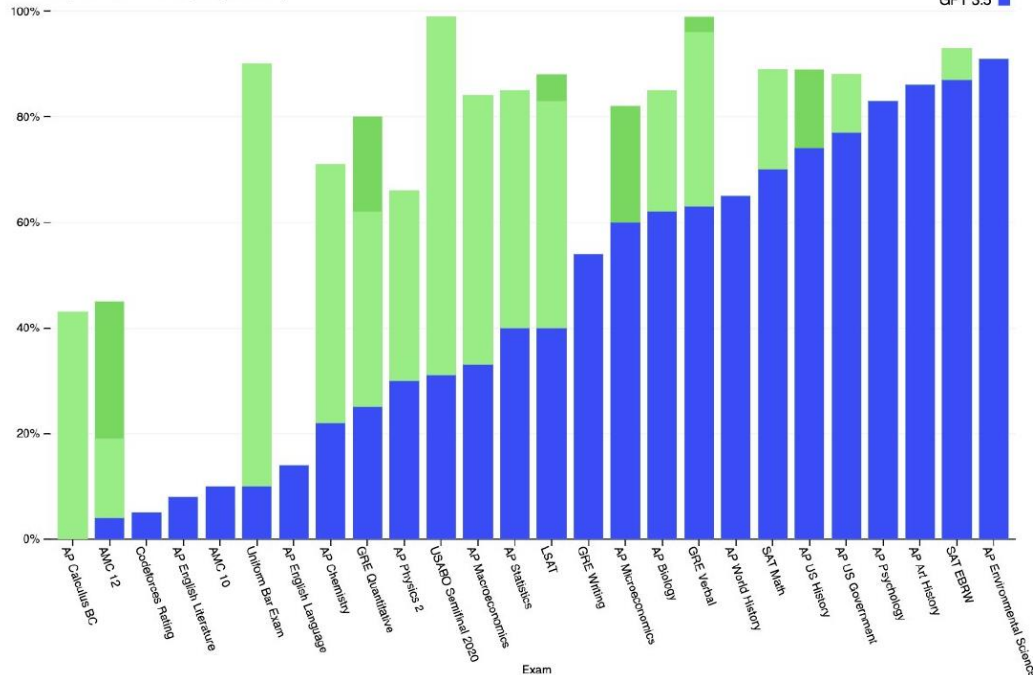
Intervention Sample Size	Control N=100, Odds Ratio=1.5, Int. Prop.=0.3	Control N=100, Odds Ratio=1.5, Int. Prop.=0.6	Control N=100, Odds Ratio=2, Int. Prop.=0.4	Control N=100, Odds Ratio=2, Int. Prop.=0.6	Control N=100, Odds Ratio=2.5, Int. Prop.=0.5	Control N=100, Odds Ratio=3.5, Int. Prop.=0.7
5	0.15	0.25	0.35	0.45	0.55	0.65
10	0.25	0.35	0.45	0.55	0.65	0.75
15	0.35	0.45	0.55	0.65	0.75	0.85
20	0.45	0.55	0.65	0.75	0.85	0.90
25	0.55	0.65	0.75	0.85	0.90	0.95
30	0.65	0.75	0.85	0.90	0.95	0.98
35	0.75	0.85	0.90	0.95	0.98	0.99
40	0.85	0.90	0.95	0.98	0.99	1.00

# What skills are most affected or unaffected by generative AI

## Exam performances by GPT3.5/4

Exam results (ordered by GPT 3.5 performance)

Estimated percentile lower bound (among test takers)



## Skills importance for GPT

Basic Skill	$\alpha$ (std err)	$\beta$ (std err)	$\zeta$ (std err)
<i>All skill importance scores are normalized to be between 0 and 1.</i>			
Constant	0.082*** (0.011)	-0.112*** (0.011)	0.300*** (0.057)
Active Listening	0.128** (0.047)	0.214*** (0.043)	0.449*** (0.027)
Mathematics	-0.127*** (0.026)	0.161*** (0.021)	0.787*** (0.049)
Reading Comprehension	0.153*** (0.041)	0.470*** (0.037)	-0.346*** (0.017)
Science	-0.114*** (0.014)	-0.230*** (0.012)	-0.346*** (0.017)
Speaking	-0.028 (0.039)	0.133*** (0.033)	0.294*** (0.042)
Writing	0.368*** (0.042)	0.467** (0.037)	0.566*** (0.047)
Active Learning	-0.157*** (0.027)	-0.065** (0.024)	0.028 (0.032)
Critical Thinking	-0.264*** (0.036)	-0.196*** (0.033)	-0.129** (0.042)
Learning Strategies	-0.072* (0.028)	-0.209*** (0.025)	-0.346*** (0.034)
Monitoring	-0.067** (0.023)	-0.149*** (0.020)	-0.232*** (0.026)
Programming	0.637*** (0.030)	0.623*** (0.022)	0.609*** (0.024)

GPT lags

GPT excels

# General AI considerations in the popular press

*Ensure responsible AI deployment*

## Accuracy

THE WALL STREET JOURNAL

LIFE & WORK | IDEAS | WORD ON THE STREET

### 'Hallucination': When Chatbots (and People) See What Isn't There

A Latin term for mental wandering was applied to the disorienting effects of psychological disorders and drug use—and then to the misfires of AI programs

### *When A.I. Chatbots Hallucinate*

Give this article



272

## Bias



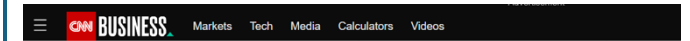
Bloomberg

Live Now Markets Economics Industries Technology Politics Wealth Pursuits Opinion Businessweek E

Newsletter

### OpenAI Chatbot Spits Out Biased Musings, Despite Guardrails

## Privacy



### Don't tell anything to a chatbot you want to keep private

INFORMATION SECURITY

6 Tips to ensure that your chatbots are GDPR compliant

## Job Displacement & Reskilling

### What To Do About AI-Driven Job Displacement

## Accountability & Responsibility

NEWS | 18 January 2023

### ChatGPT listed as author on research papers: many scientists disapprove

At least four articles credit the AI tool as a co-author, as publishers scramble to regulate its use.

Technology

4 minute read - April 5, 2023 9:22 AM EDT - Last Updated a month ago

Elon Musk and others urge AI pause, citing 'risks to society'

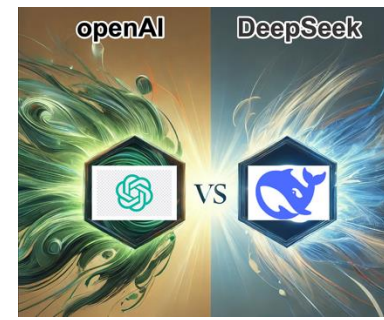
## Transparency & Explainability

Technology And Analytics

### We Need AI That Is Explainable, Auditable, and Transparent

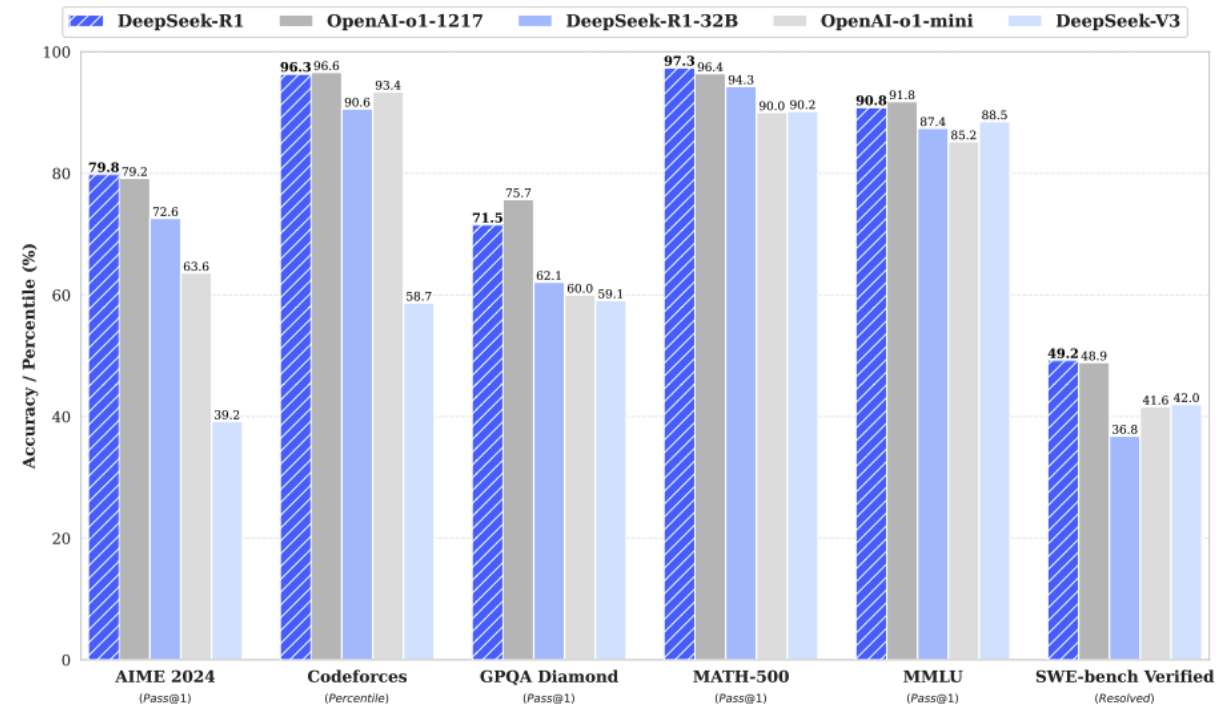
# Closed versus Open Source

*Maybe these models don't require so much compute after all*



## DeepSeek R1 vs OpenAi o1

Feature	OpenAI's o1	DeepSeek's R1
<b>Release Date</b>	December 5, 2024	January 21, 2025
<b>Architecture</b>	Focuses on complex reasoning with an internal chain-of-thought mechanism	Utilizes a Mixture-of-Experts framework, activating 37B out of 671B parameters per token
<b>Context Window</b>	Supports up to 200,000 tokens	Supports up to 128,000 tokens
<b>Mathematical Reasoning</b>	Ranks among the top 500 U.S. students in the AIME	Scores 91.6% on the MATH benchmark
<b>Coding Proficiency</b>	Ranks in the 89th percentile on Codeforces	Excels in generating and debugging complex code
<b>Cost</b>	Proprietary model with higher computational costs	Open-source model with significantly lower pricing
<b>Strengths</b>	Advanced reasoning capabilities; exceptional performance on academic and coding benchmarks	Open-source nature fosters innovation and collaboration; cost-effective; transparent reasoning processes
<b>Limitations</b>	Higher computational costs; slower response times; limited support for certain features like streaming outputs	Slower processing for specific tasks; occasional challenges with nuanced output formatting





# Concluding remarks

**Balanced Approach:** Integration with a balanced perspective, considering both potential and limitations

**Ethical Considerations:** Must be proactively addressed to maintain public trust and ensure equitable outcomes in drug development

**Workforce Adaptability:** Pharma's workforce needs to be agile and adaptable, acquiring new skill sets to work alongside AI technologies effectively

**Regulatory Framework:** A growing need for robust regulatory frameworks to ensure that AI applications in drug development meet safety and efficacy standards

**Future Preparedness:** Improved collaboration to prepare the next generation of scientists, researchers, and professionals for a future where AI will be an integral part of drug development

**AI CHATBOTS**  
postdoctoral researchers polled said that  
AI chatbots, such as ChatGPT, for everything  
fixing code and overcoming writer's block.

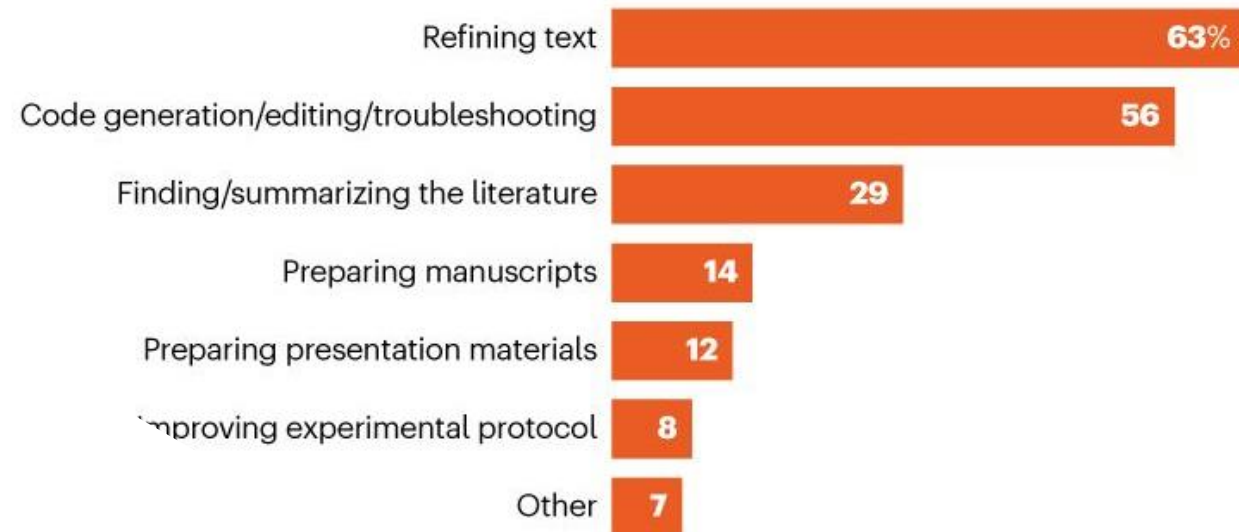
**evolution and rapid adoption of AI chatbots  
changed your day-to-day work or career plans, if at all?**



**Do you use AI chatbots, such as ChatGPT, in your work?**



**What do you use AI chatbots for?**



**AI chatbots in your work**