

**[CryptoLab Lyon Research Center AI Seminar]**

**AI - Technology, Research & Products**

**Sunghee Yun**

**Co-founder / CTO - AI Technology @ Erudio Bio, Inc.**

## About Speaker

- *Co-founder / CTO - AI Technology & Product Strategy @ Erudio Bio, CA, USA*
- Advisory Professor, Electrical Engineering and Computer Science @ DGIST
- Adjunct Professor, Electronic Engineering Department @ Sogang University
- Technology Consultant @ Gerson Lehrman Group (GLG)
- *KFAS-Salzburg Global Leadership Initiative Fellow @ Salzburg Global Seminar*
- *Co-founder / CTO & Chief Applied Scientist @ Gauss Labs, CA, USA – 2023*
- Senior Applied Scientist @ Mobile Shopping App Org, Amazon.com, Inc. – 2020
- Principal Engineer @ Software R&D Center of DS Division, Samsung – 2017
- Principal Engineer @ Strategic Marketing & Sales Team, Samsung – 2016
- Principal Engineer @ DT Team of DRAM Development Lab, Samsung – 2015
- Senior Engineer @ CAE Team - Samsung – 2012
- M.S. & Ph.D. - Electrical Engineering @ Stanford University – 2004
- B.S. - Electrical Engineering @ Seoul National University – 1998

## Highlight of career journey

- B.S. in EE @ SNU, M.S. & Ph.D. in EE @ Stanford Univ.
  - *Convex Optimization - theory / algorithms / applications* - supervision of *Prof. Stephen P. Boyd*
- Principal Engineer @ Memory Design Technology Team
  - AI & optimization partnering with *DRAM/NAND Design/Process/Test teams*
- Senior Applied Scientist @ Amazon
  - *S-Team Goal (Bezos's) project* - improve customer engagement via Amazon Mobile Shopping App using AI - *increased sales by USD 200M*
- Co-founder / CTO & Chief Applied Scientist @ Gauss Labs
  - *R&D industrial AI products & technology, market/product/investment strategies*
- Co-founder / CTO - AI Technology & Product Strategy @ Erudio Bio
  - *biotech - AI technology & product strategy*

# Today

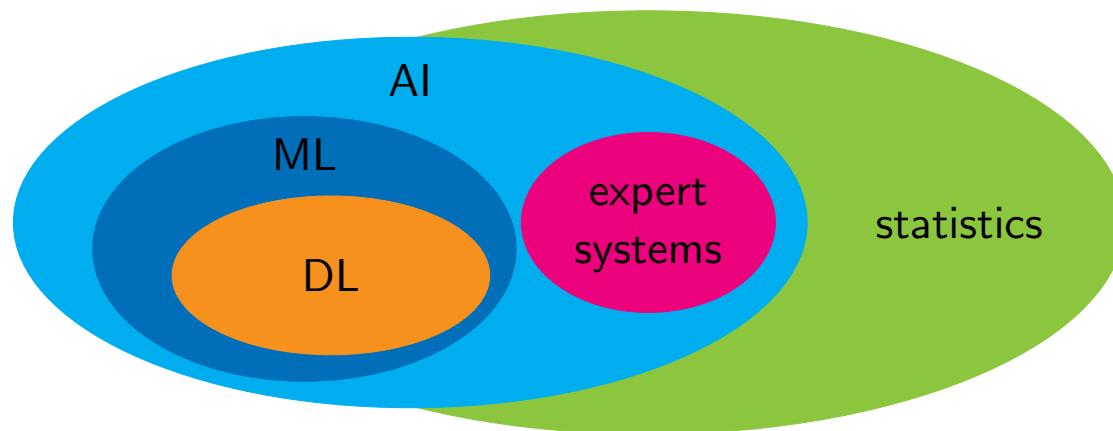
- AI - history and recent progress
- AI research and development trend
- AI products
- Serendipities around AI
- Some important questions around AI
- Appendices
  - AI hardware
  - AI & biotech

# **Artificial Intelligence**

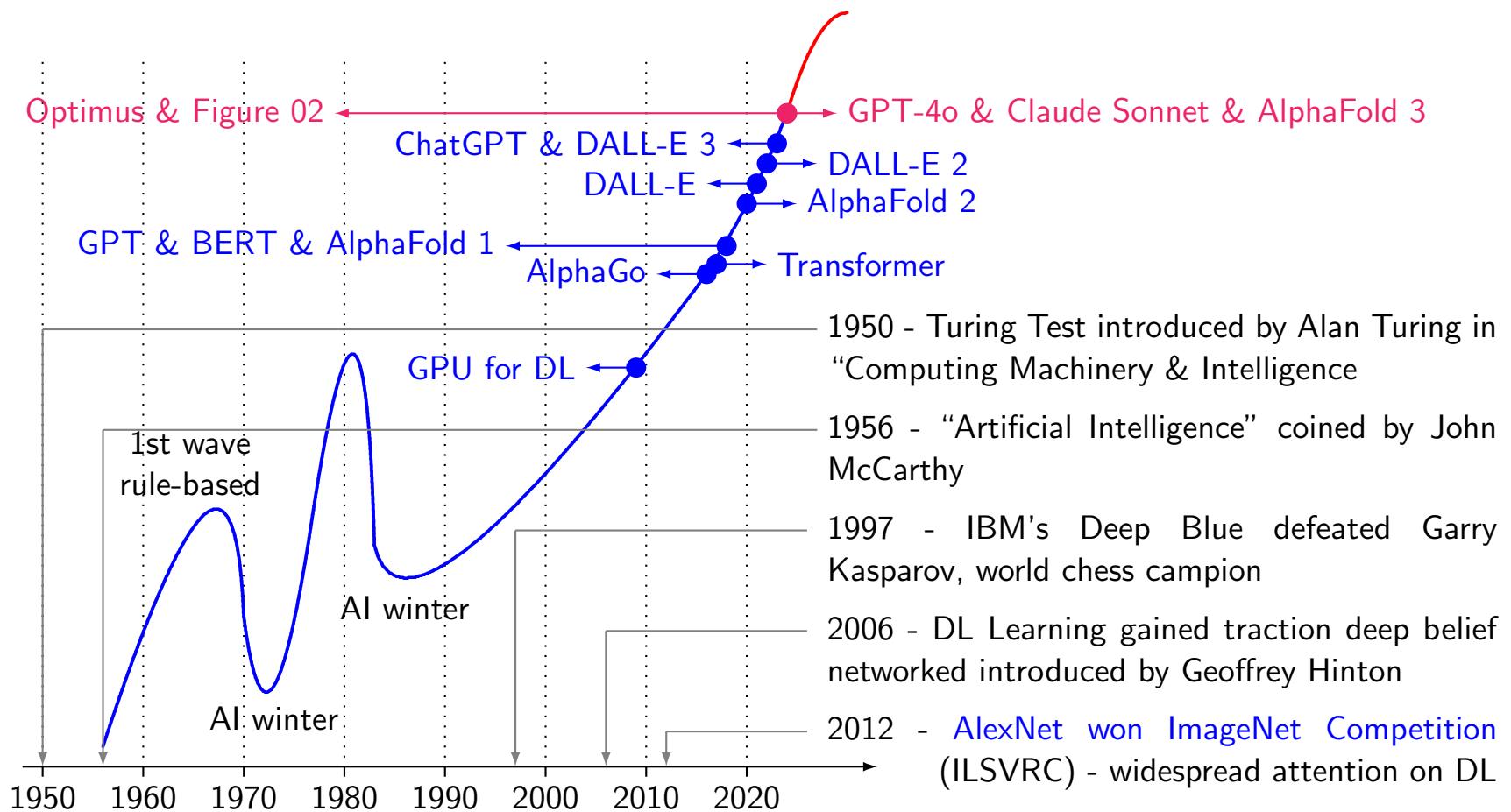
## **Definition and History**

## Definition of AI

- AI is
  - technology enabling machines to do tasks requiring human intelligence, such as learning, problem-solving, decision-making & language understanding
  - *not one thing* - encompass range of technologies, methodologies & applications
- relationship of AI, statistics, ML, DL, NN & expert system [HGH<sup>+</sup>22]



# History of AI



# **Significant AI Achievements - 2014 – 2024**

## Deep learning revolution

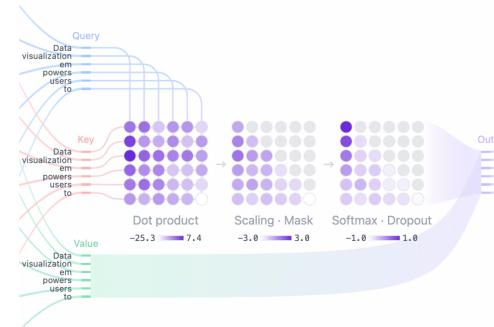
- 2012 – 2015 - DL revolution<sup>1</sup>
  - CNNs demonstrated exceptional performance in image recognition, e.g., *AlexNet's victory in ImageNet competition*
  - widespread adoption of DL learning in CV transforming industries
- 2016 - AlphaGo defeats human Go champion
  - DeepMind's AlphaGo defeated world champion in Go, extremely complex game *believed to be beyond AI's reach*
  - significant milestone in RL - AI's potential in solving complex & strategic problems



<sup>1</sup>DL: deep learning, CNN: convolutional neural network, CV: computer vision, RL: reinforcement learning

## Transformer changes everything

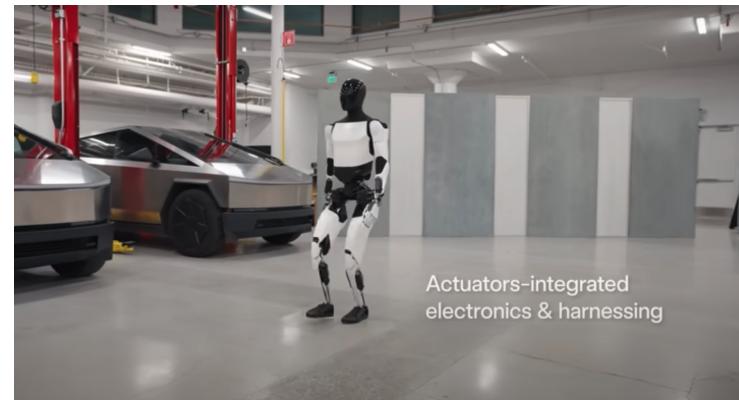
- 2017 – 2018 - Transformers & NLP breakthroughs<sup>2</sup>
  - *Transformer (e.g., BERT & GPT) revolutionized NLP*
  - major advancements in, *e.g.*, machine translation & chatbots
- 2020 - AI in healthcare – AlphaFold & beyond
  - DeepMind's *AlphaFold solves 50-year-old protein folding problem* predicting 3D protein structures with remarkable accuracy
  - accelerates drug discovery and personalized medicine - offering new insights into diseases and potential treatments



<sup>2</sup>NLP: natural language processing, GPT: generative pre-trained transformer

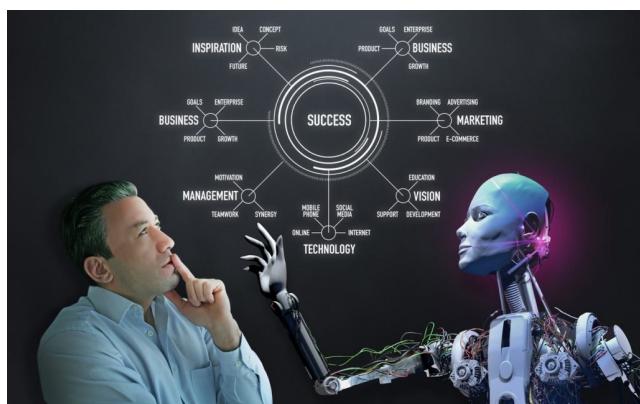
## Lots of breakthroughs within 6 months in 2024

- proliferation of advanced AI models
  - GPT-4o, Claude Sonnet, Llama 3, Sora
  - *transforming industries* such as content creation, customer service, education, etc.
- breakthroughs in specialized AI applications
  - Figure 02, Optimus, AlphaFold 3
  - driving unprecedented advancements in automation, drug discovery, scientific understanding - *profoundly affecting healthcare, manufacturing, scientific research*



## **Transformative impact of AI - reshaping industries, work & society**

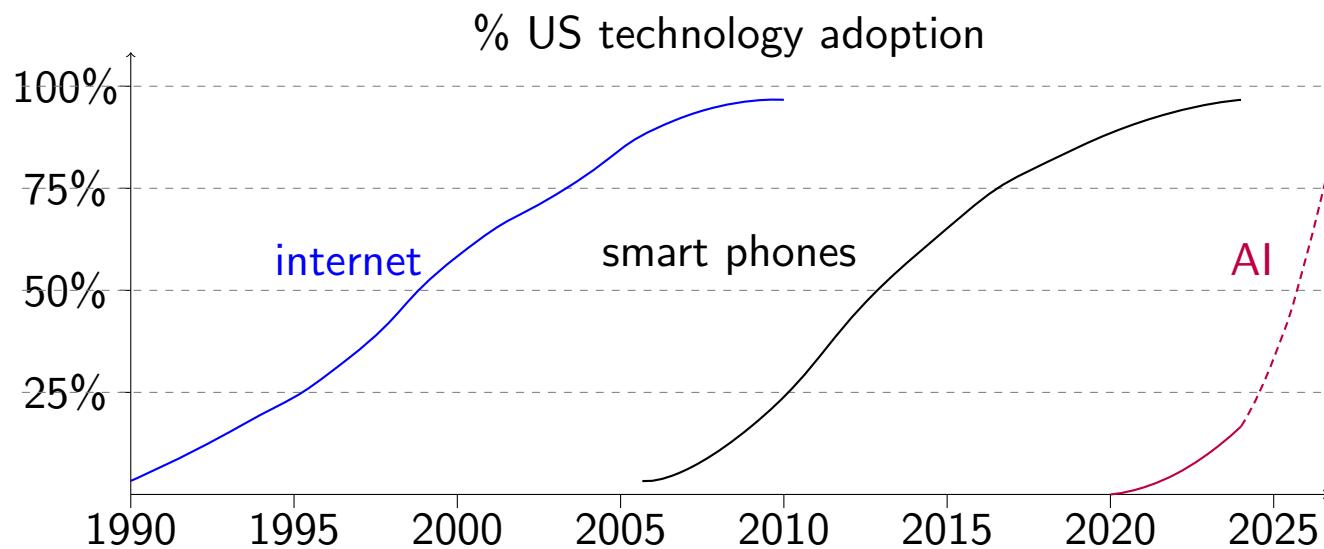
- accelerating human-AI collaboration
    - not only reshaping industries but *altering how humans interact with technology*
    - AI's role as collaborator and augmentor redefines productivity, creativity, the way we address global challenges, e.g., *sustainability & healthcare*
  - AI-driven automation *transforms workforce dynamics* - creating new opportunities while challenging traditional job roles
  - *ethical AI considerations* becoming central not only to business strategy, but to society as a whole - *influencing regulations, corporate responsibility & public trust*



# **Recent Advances in AI**

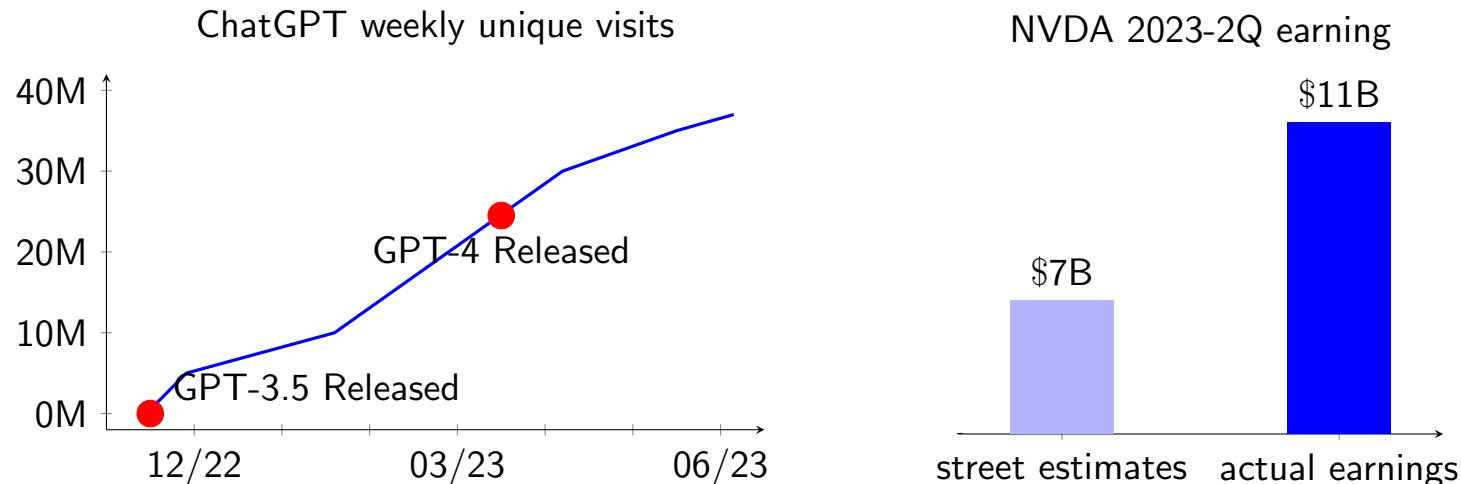
## Where are we in AI today?

- sunrise phase - currently experiencing dawn of AI era with significant advancements and increasing adoption across various industries
- early adoption - in early stages of AI lifecycle with widespread adoption and innovation across sectors marking significant shift in technology's role in society



## Explosion of AI ecosystems - ChatGPT & NVIDIA

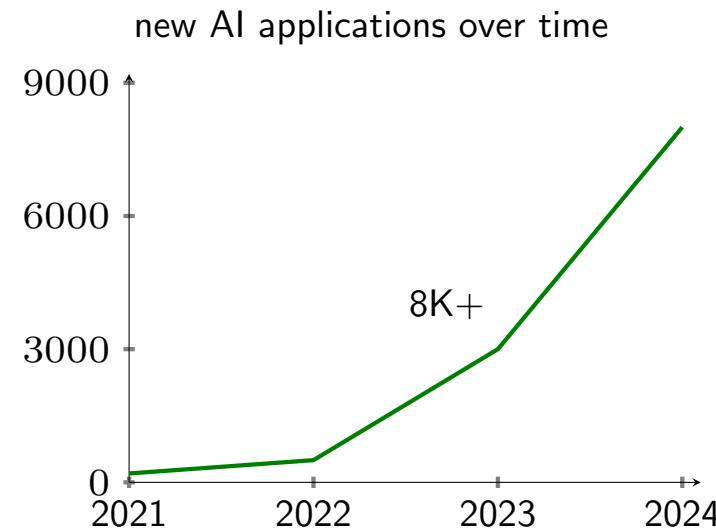
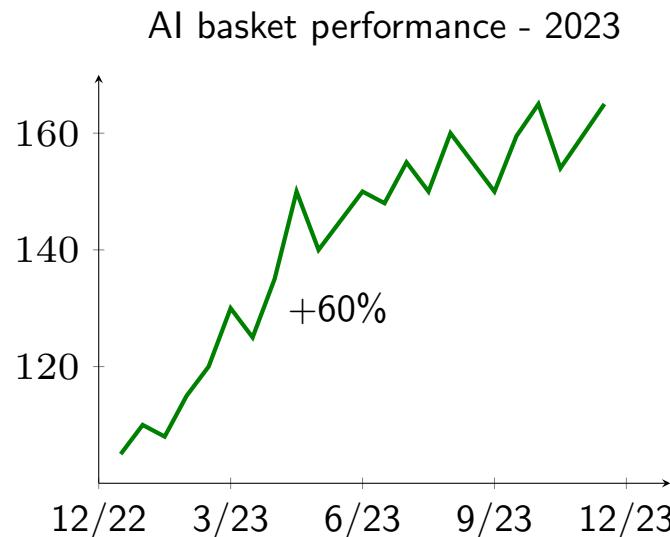
- took only *5 months for ChatGPT users to reach 35M*
- NVIDIA 2023 Q2 earning exceeds market expectation by big margin - \$7B vs \$13.5B
  - surprisingly, *101% year-to-year growth*
  - even more surprisingly *gross margin was 71.2%* - up from 43.5% in previous year<sup>3</sup>



<sup>3</sup>source - Bloomberg

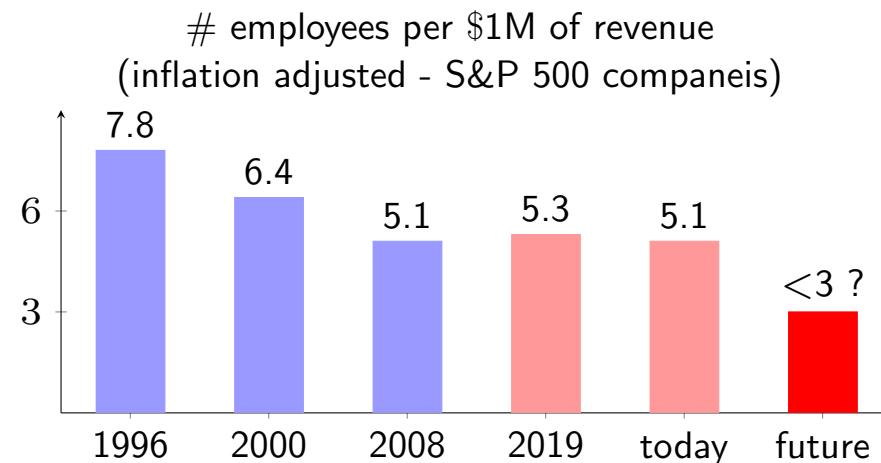
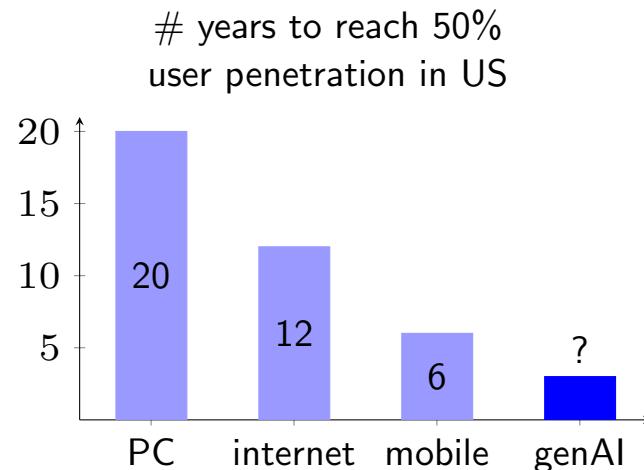
## Explosion of AI ecosystems - AI stock market

- *AI investment surge in 2023 - portfolio performance soars by 60%*
  - AI-focused stocks significantly outpaced traditional market indices
- *over 8,000 new AI applications* developed in last 3 years
  - applications span from healthcare and finance to manufacturing and entertainment



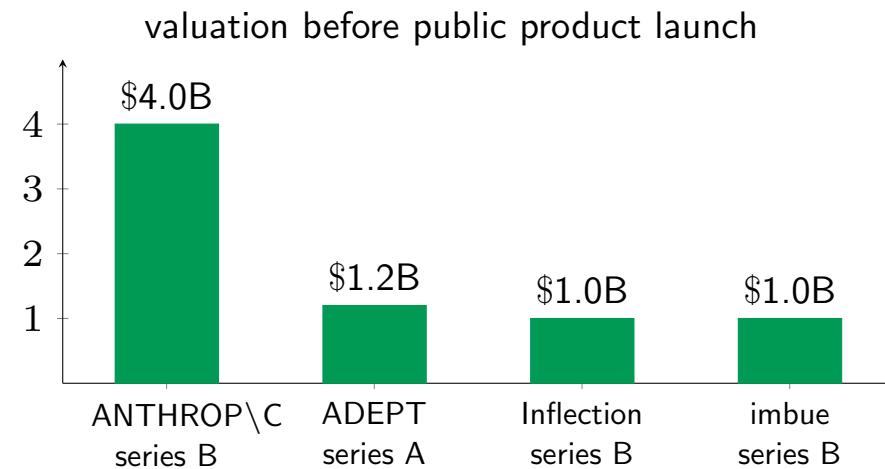
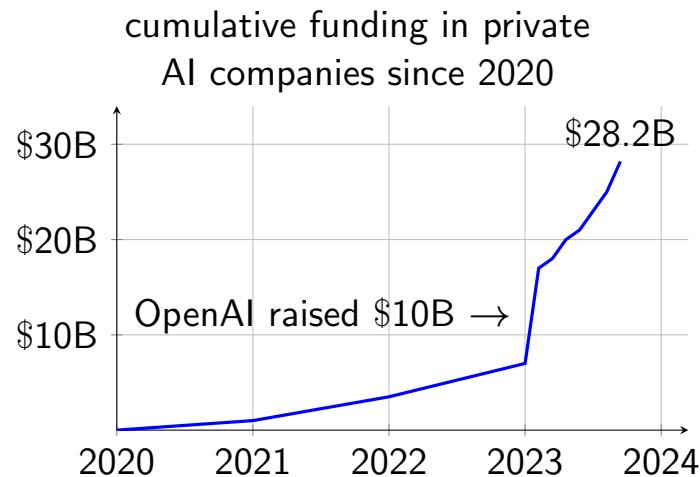
## AI's transformative impact - adoption speed & economic potential

- adoption - has been twice as fast with platform shifts suggesting
  - increasing demand and readiness for new technology improved user experience & accessibility
- AI's potential to drive economy for years to come
  - 35% improvement in productivity driven by introduction of PCs and internet
  - greater gains expected with AI proliferation



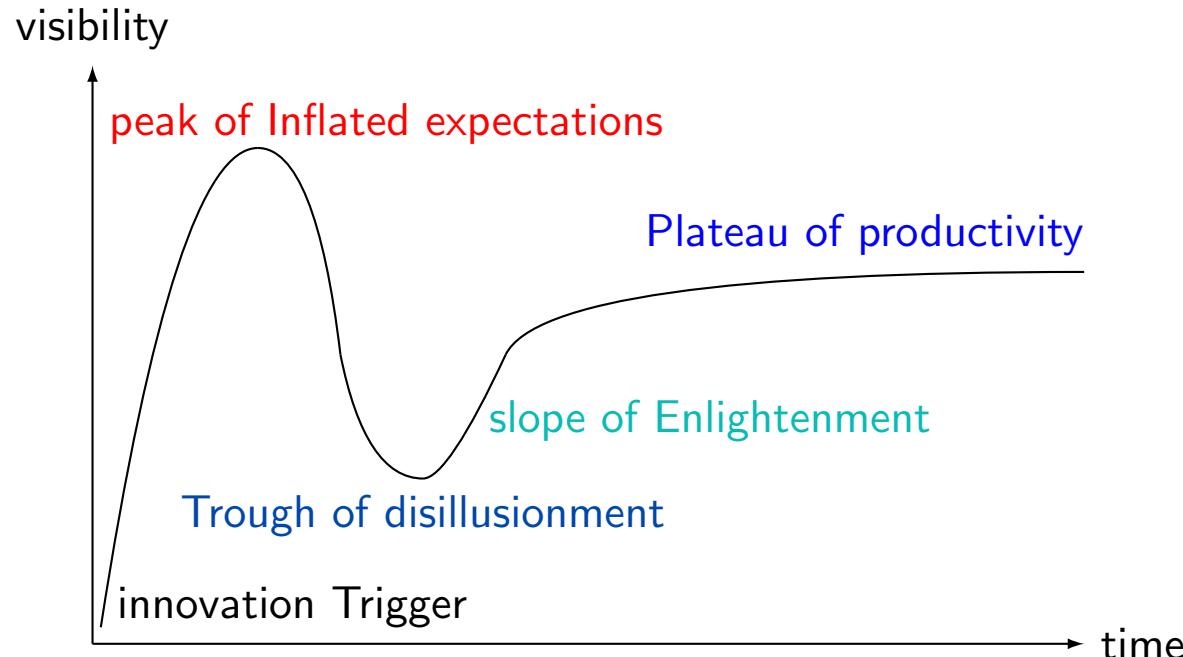
## Massive investment in AI

- *explosive growth* - cumulative funding skyrocketed reaching staggering \$28.2B
- OpenAI - significant fundraising (= \$10B) fueled rapid growth
- *valuation surge* - substantial valuations even before public products for stellar companies
- *fierce competition for capital* among AI startups driving innovation & accelerating development
- massive investment indicates *strong belief in & optimistic outlook for potential of AI* to revolutionize industries & drive economic growth



**Is AI hype?**

## Technology hype cycle



- innovation trigger - technology breakthrough kicks things off
- peak of inflated expectations - early publicity induces many successes followed by even more
- trough of disillusionment - expectations wane as technology producers shake out or fail
- slope of enlightenment - benefit enterprise, technology better understood, more enterprises fund pilots

## Fiber vs cloud infrastructure

- fiber infrastructure - 1990s
  - Telco Co's raised \$1.6T of equity & \$600B of debt
  - bandwidth costs decreased 90% within 4 years
  - companies - Covage, NothStart, Telligent, Electric Lightwave, 360 networks, Nextlink, Broadwind, UUNET, NFS Communications, Global Crossing, Level 3 Communications
  - became *public good*
- cloud infrastructure - 2010s
  - entirely new computing paradigm
  - mostly public companies with data centers
  - *big 4 hyperscalers generate \$150B + annual revenue*



## Yes & No

characteristics of hype cycles	speaker's views
value accrual misaligned with investment	<ul style="list-style-type: none"><li>• OpenAI still operating at a loss; business model <i>still</i> not clear</li><li>• gradual value creation across broad range of industries and technologies (<i>e.g.</i>, CV, LLMs, RL) unlike fiber optic bubble in 1990s</li></ul>
overestimating timeline & capabilities of technology	<ul style="list-style-type: none"><li>• self-driving cars delayed for over 15 years, with limited hope for achieving level 5 autonomy</li><li>• AI, however, has proven useful within a shorter 5-year span, with enterprises eagerly adopting</li></ul>
lack of widespread utility due to technology maturity	<ul style="list-style-type: none"><li>• AI already providing significant utility across various domains</li><li>• vs quantum computing remains promising in theory but lacks widespread practical utility</li></ul>

# AI Research

## AI research race gets crazy

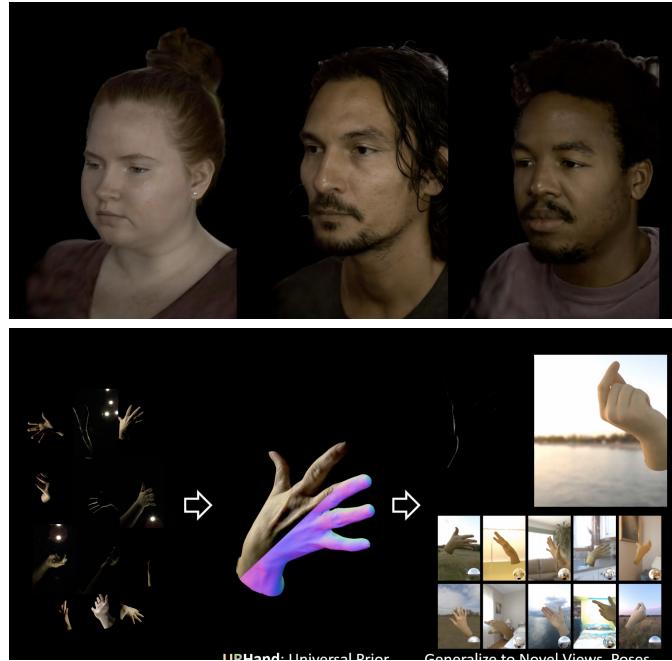
- practically impossible to follow all developments announced everyday
  - new announcement and publication of important work everyday!
- *industry leads research - academia lags behind*
  - trend observed even before 2015
- everyone excited to show off their work to the world
  - conference and [github.com](https://github.com)
  - biggest driving force behind unprecedented scale and speed of advancement of AI together with massive investment of capitalists



## AI progress within a month - March, 2024

- UBTECH Humanoid Robot Walker S: Workstation Assistant in EV Production Line
- H1 Development of dance function
- Robot Foundation Models (Large Behavior Models) by Toyota Research Institute (TRI)
- Apple Vision Pro for Robotics
- Figure AI & OpenAI
- Human modeling
- LimX Dynamics' Biped Robot P1 Conquers the Wild Based on Reinforcement Learning
- HumanoidBench: Simulated Humanoid Benchmark for Whole-Body Locomotion and Manipulation - UC Berkeley & Yonsei Univ.
- Vision-Language-Action Generative World Model
- RFM-1 - Giving robots human-like reasoning capabilities

## Papers of single company accepted by single conference



- CVPR 2024

- PlatoNeRF: 3D Reconstruction in Plato's Cave via Single-View Two-Bounce Lidar - MIT, Codec Avatars Lab, & Meta [KXS<sup>+</sup>24]
  - 3D reconstruction from single-view
- Nymeria Dataset
  - large-scale multimodal egocentric dataset for full-body motion understanding
- Relightable Gaussian Codec Avatars - Codec Avatars Lab & Meta [SSS<sup>+</sup>24]
  - build high-fidelity relightable head avatars being animated to generate novel expressions
- Robust Human Motion Reconstruction via Diffusion (RoHM) - ETH Zürich & Reality Labs Research, Meta [ZBX<sup>+</sup>24]
  - robust 3D human motion reconstruction from monocular RGB videos

# **AI Products**

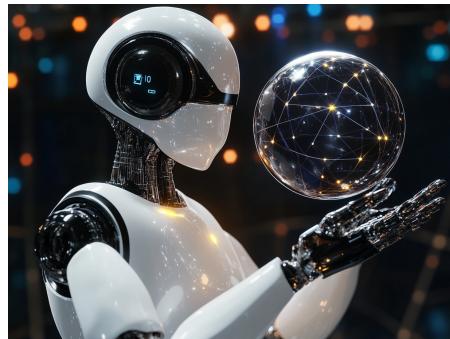
## AI product development - trend and characteristics

- *rapid pace* of innovation - new AI models & products being released at unprecedented rate, improvements coming in weeks or months (rather than years)
- *LLMs dominating* - models like GPT-4 & Claude pushing boundaries in NLP & genAI
- *multimodal AI* gaining traction - models processing & generating text, images & even video becoming more common, e.g., Grok, GPT-4, Gemini w/ vision capabilities
- *open-source* AI movement - growing trend of open-source AI models and tools, challenging dominance of proprietary systems
- *AI integration in everyday products* - from smartphones to home appliances, AI being integrated into wide array of consumer products



# AI product development - trend and characteristics

- *ethical AI & regulatory focus* - increased attention on ethical implications of AI & calls for regulation of AI development and deployment
  - AI in enterprise - businesses across industries rapidly adopting AI for various applications
  - *specialized AI models* - development of AI models tailored for specific industries or tasks, e.g., healthcare, biotech, financial analysis
  - AI-assisted *coding and development* - help software developers write code more efficiently & tools becoming increasingly sophisticated
  - *concerns about AI safety & existential risk* - growing debate about potential short & long-term risks of advanced AI



## LLM products

- OpenAI - ChatGPT 4o, GPT-4 Turbo Canvas
- Anthropic - Claude 3.5 Sonnet (with Artifacts), Claude 3 Opus, Claude 3 Haiku
- Mistral AI - Mistral 7B, Mistral Large 2, Mistral Small xx.xx, Mistral Nemo (12B)
- Google - Gemini (w/ 1.5 Flash), Gemini Advanced (w/ 1.5 Pro)
- X - Grok [mini] [w/ Fun Mode]
- Perplexity AI - Perplexity [Pro] - combines GPT-4, Claude 3.5, and Llama 3
- Liquid AI - Liquid-40B, Liquid-3B (running on small devices)

flying cats generated by Grok, ChatGPT 4o & Gemini



## Comparison of LLMs & LLM products

model	developer	training data	# params	strength	weakness
GPT-4	OpenAI	web & books	170B	advanced reasoning & multimodal capabilities	high computational resources
LLaMA-2	Meta	public info & research articles	7~70B	open access & good performance for different sizes	not powerful for complex tasks
Claude	Anthropic	mix of high-quality datasets	not disclosed	safety-first approach avoiding harmful responses	limited in publicly available details
PaLM 2	Google	multilingual text corpus	540B	high multilingual comprehension supporting various downstream apps	significant resources & not versatile in some contexts

## Comparison of LLMs & LLM products

model	developer	training data	# params	strength	weakness
BLOOM	BigScience Community	diverse multilingual corpus	176B	open & support multiple languages	resource-intensive & lower performance
Mistral <sup>4</sup>	Mistral AI	public web data	7~13B	lower parameter count	limited scalability for specialized apps
Liquid Foundation Model (LFM)	Liquid AI	adaptive datasets	adaptive & dynamic parameters	modular & support more specialized fine-tuning for niche use-cases & adaptable in deployment	complexity in design and implementation

## Multimodal genAI products

- DALL-E by OpenAI
  - *generate unique and detailed images based on textual descriptions*
  - understanding context and relationships between words
- Midjourney by Midjourney
  - let people *create imaginative artistic images*
  - can interactively guide the generative process, providing high-level directions



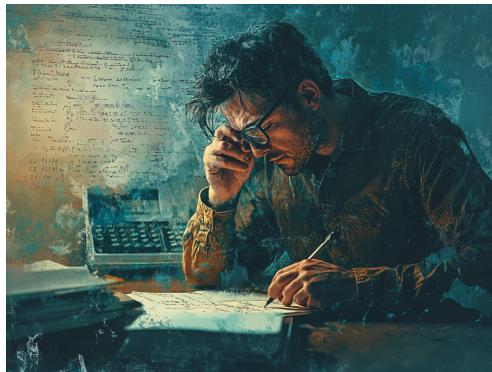
## Multimodal genAI products



- Dream Studio by Stability AI
  - *analyze patterns in music data & generates novel compositions*
  - musicians can explore new ideas and enhance their *creative* processes
  
- Runway by Runway AI
  - *realistic images, manipulate photos, create 3D models & automate filmmaking*

## Rise of co-pilot products

- definition - AI-powered tools designed to enhance human productivity across multiple domains including document creation, presentations & coding
- benefits
  - *efficiency* - automate repetitive tasks allowing users to focus on high-value activities
  - *error reduction* - minimize mistakes common in manual work
  - *creativity* - suggestions and prompts help users explore new ideas and approaches
  - *integration* with major productivity suites - Microsoft 365, Google Workspace
- popular products
  - [GitHub Copilot](#), [Microsoft 365 Copilot](#), [Grammarly AI](#), [Visual Studio Code Extensions](#)



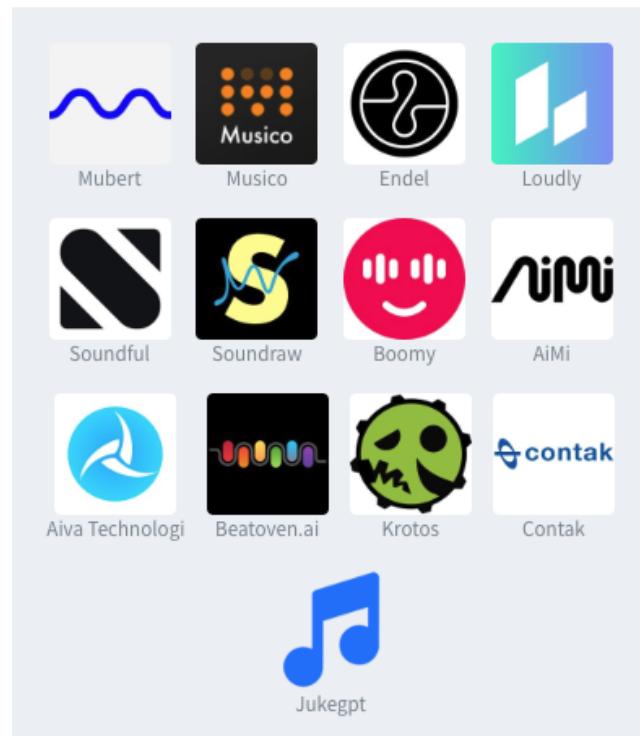
## Future of co-pilot products

- potential advancements
  - wider adoption across industries and professions
  - *real-time fully automated collaboration, predictive content generation*, personalization
- impact on work environments & creative processes
  - *collaborative human-AI relationships* with augmented reality
  - unprecedented levels of problem-solving due to *augmented cognitive abilities*
- challenges & considerations
  - *ethical concerns around data privacy & AI decision-making*
  - potential impact on *human skills & job markets*

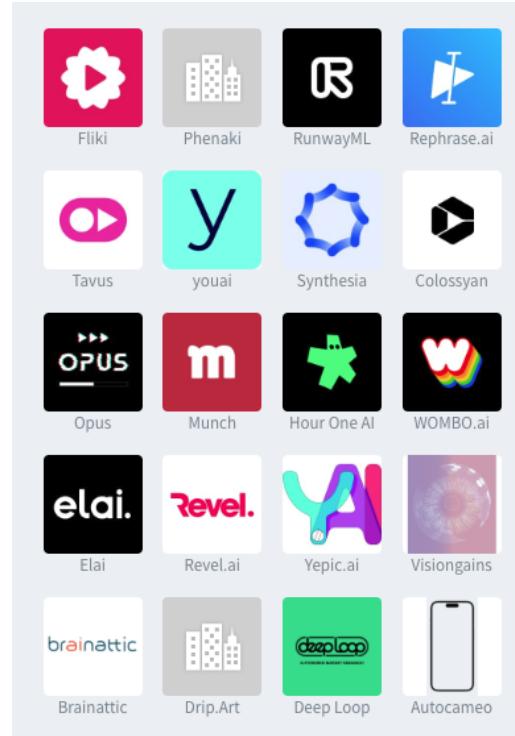


## Other AI products - audio/video/text

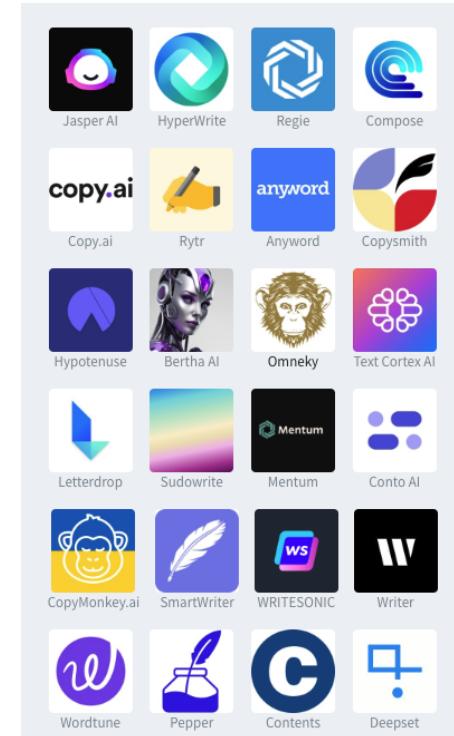
audio



video

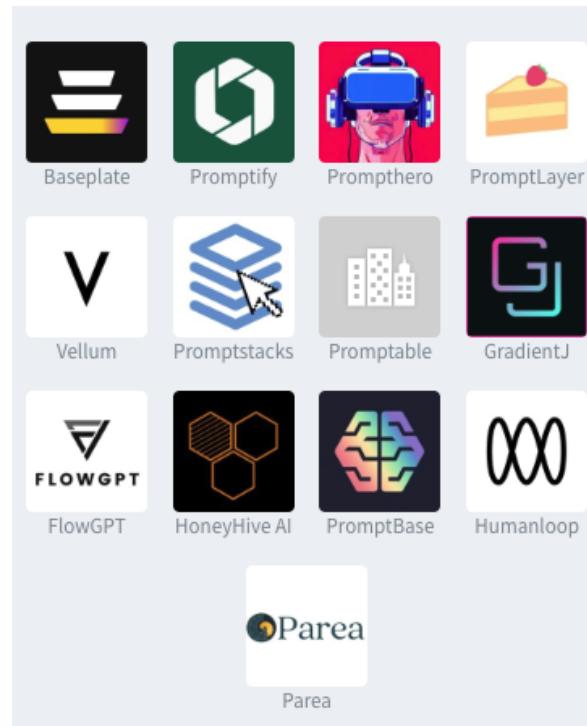


text

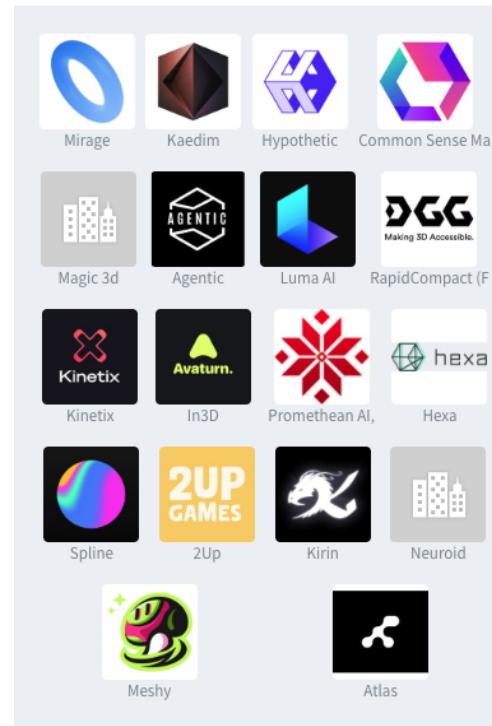


## Other AI products - LLM/gaming/design/coding

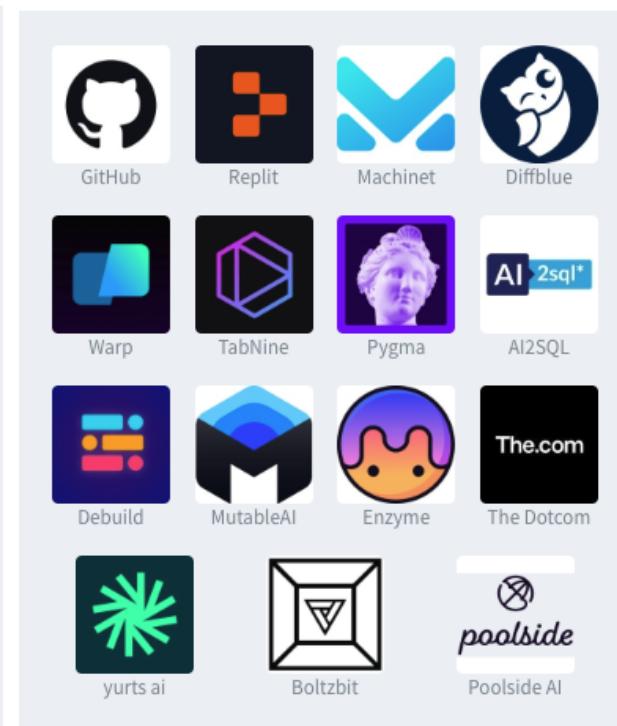
LLM



gaming &amp; design



coding



# **Serendipities around Als**

## Serendipity or inevitability?

- What if Geoffrey Hinton had not been a persistent researcher?
- What if Geoffrey had been a symbolist (instead of connectionist)?
- What if symbolists won AI race over connectionists?
- What if attention mechanism did not perform well?
- What if Transformer architecture did not perform super well?
- What if Jensen Hwang had not been crazy about making hardware for professional gamers?
- Is it like Alexander Fleming's Penicillin?
- Or more like Inevitability?

# **Important Questions to be Asked**

## **Some important questions around AI**

- why human-level AI in the first place?
- what lies in very core of DL architecture? what makes it work amazingly well?
- biases that can hurt judgement, decision making, social good?
- ethical and legal issues
- consciousness, knowledge, belief, reasoning
- future of AI

**Human-level AI?**

## Why human-level in the first place?

- lots of times, when we measure AI performance, we say
  - how can we achieve human-level performance, *e.g.*, CV models?
- why human-level?
  - are all human traits desirable? are humans flawless?
  - aren't humans still evolving?
- advantage of AI over humans
  - *e.g.*, self-driving cars can use extra eyes, GPS, computer network
  - *e.g.*, recommendation system runs for hundreds of millions of people overnight
  - AI is available 24 / 7 while humans cannot
    - . . . critical advantages for medical assistance, emergency handling
  - AI does not make more mistakes because task is repetitive and tedious
  - AI does not request salary raise or go on strike

**What makes DL so successful?**

## Factors contributing to astonishing success of DL

- analysis based on speaker's mathematical, numerical algorithmic & statistical perspectives considering hardware innovations

**30%** universal approximation theorem? - (partially) yes! but that's not all

- function space of neural network is *dense* (math theory), *i.e.*, for every  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ , exists  $\langle f_n \rangle$  such that  $\lim_{n \rightarrow \infty} f_n = f$

**25%** architectures/algorithms tailored for each class of applications, *e.g.*, CNN, RNN, Transformer, NeRF, diffusion, GAN, VAE, . . .

**20%** data labeling - expensive, data availability - unlimited web text corpus

**15%** computation power/parallelism - AI accelerators, *e.g.*, GPU, TPU & NPU

**10%** rest - Python, open source software, cloud computing, MLOps, . . .

**Why do we see sudden leap in LLM performance?**

## Probability inferred sequence is correct

- assume

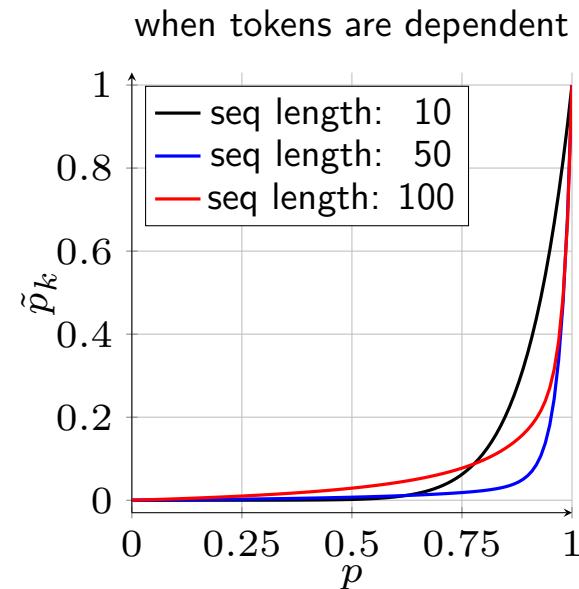
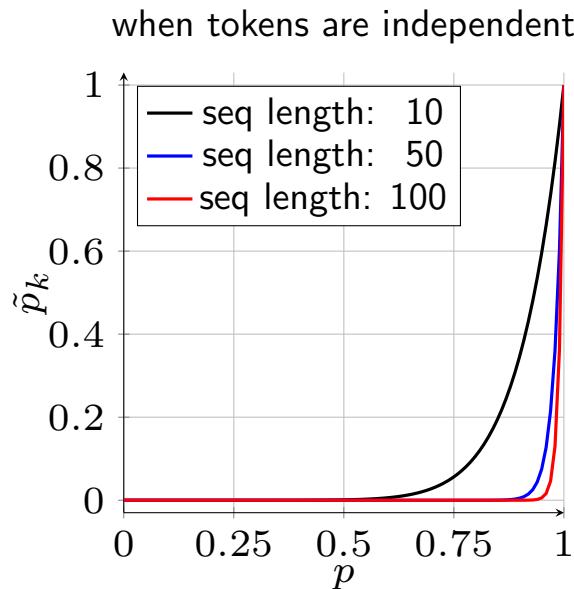
- $t_i$  -  $i$ th token
- $p_i$  - probability that  $t_i$  is correct
- $\rho_i$  - correlation coefficient between  $t_{i-1}$  &  $t_i$
- $\tilde{p}_k$  - probability that  $(t_1, \dots, t_k)$  are correct

- recursion

$$\rho_i = \frac{\tilde{p}_i - \tilde{p}_{i-1}p_i}{\sqrt{\tilde{p}_{i-1}(1 - \tilde{p}_{i-1})p_i(1 - p_i)}}$$
$$\Leftrightarrow \quad \tilde{p}_i = \tilde{p}_{i-1}p_i + \rho_i \sqrt{\tilde{p}_{i-1}(1 - \tilde{p}_{i-1})p_i(1 - p_i)}$$

## Dramatic improvement of LLM near saturation

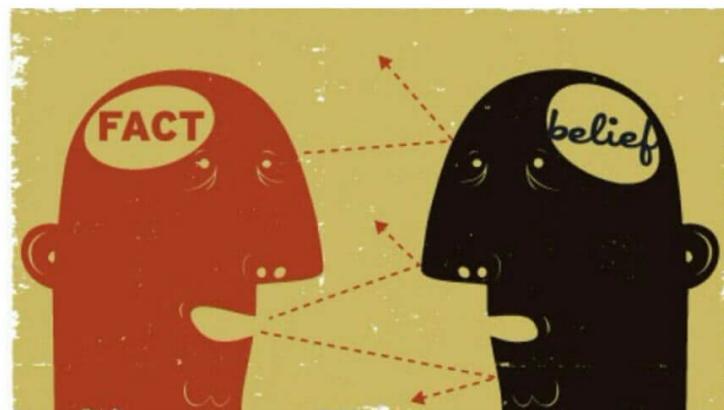
- do simulations for both independent & dependent cases
  - assume  $p_i$  are same for all  $i$
- (for both cases) sequence inference improves dramatically as  $p$  approaches 1
- this explains *why we have observed sudden dramatic performance improvement of certain seq2seq learning technologies*, e.g., LLM



## **Biases - by Humans & Machines**

## Cognitive biases

- cognitive biases [Kah11]
  - confirmation bias, availability bias
  - hindsight bias, confidence bias, optimistic bias
  - anchoring bias, halo effect, framing effect, outcome bias
  - belief bias, negativity bias, false consensus,



## LLM biases

- plausible with LLM
  - availability bias - biased by imbalancedly available information
    - LLM trained by imbalanced # articles for specific topics
  - belief bias - derive conclusion not by reasoning, but by what it saw
    - LLM easily inferencing what it saw, *i.e.*, data it trained on
  - halo effect - overemphasize on what prestigious figures say
    - LLM trained by imbalanced # reports about prestigious figures
- similar facts true for other types of ML models,
  - *e.g.*, video caption, text summarization, sentiment analysis
- cognitive biases only humans represent
  - confirmation bias, hindsight bias, confidence bias, optimistic bias, anchoring bias, negativity bias, framing effect

## **Ethical and Legal Issues**

## Ethics - possibilities & questions

- AI can be exploited by those who have bad intention to
  - manipulate / deceive people - using manipulated data corpus for training
    - *e.g.*, spread false facts
  - induce unfair social resource allocation
    - *e.g.*, medical insurance, taxation
  - exploit advantageous social and economic power
    - *e.g.*, unfair wealth allocation, mislead public opinion
- AI for Good - advocated by Andrew Ng
  - *e.g.*, public health, climate change, disaster management
- should scientists and engineers be morally & politically conscious?
  - *e.g.*, Manhattan project

## Ethically controversial issues

- AI girlfriends
  - lots of AI girlfriend apps already developed
  - ethical considerations and provisions for user privacy with AI partners imperative - as with every technology involving personal data and emotional interaction
  - prospect of developing lifelike digital companions will grow better with evolution of AI
  - perhaps changing ways relationships and companionship perceived in digital age one day
  - why not many AI boyfriend apps? is this sexual discrimination issue (at all)?

## Legal issues with ethical consideration - (hypothetical) scenarios

- scenario 1: full self-driving algorithm causes traffic accident killing people
  - who is responsible? - car maker, algorithm developer, driver, algorithm itself?
- scenario 2: self-driving cars kill less people than human drivers
  - e.g., human drivers kill 1.5 people for 100,000 miles & self-driving cars kill 0.2 people for 100,000 miles
  - how should law makers make regulations?
  - utilitarian & humanistic perspectives
- scenario 3: someone is not happy with their data being used for training
  - “The Times sues OpenAI and Microsoft over AI use of copyrighted work” (Dec. 2023)
  - “Newspaper publishers in California, Colorado, Illinois, Florida, Minnesota and New York said Microsoft and OpenAI used millions of articles without payment or permission to develop ChatGPT and other products” (Apr. 2024)

# **Consciousness**

## Consciousness

- what is consciousness, anyway?
  - recognizes itself as independent, autonomous, valuable entity?
  - recognizes itself as living being, unchangeable entity?
- no agreed definition on consciousness exists yet
  - . . . and will be so forever
- does it have anything to do with the fact that humans are biologically living being?
  - *I don't think so . . .*
- is SKYNET ever plausible (without someone's intention)?
  - can AI have *desire* to survive (or save earth)?



## Utopia or dystopia



- not important questions (at all) *I think . . .*
- what we should focus on is not the possibilities of doomsday or Judgment Day, but rather
  - our limits on controlling unintended impacts of AI
  - misuse by (greedy and bad) people possessing social, economic & political power
  - social good and welfare impaired by (exploiting) AI
  - choice among utilitarianism, humanism, justice & equity
  - amend or improve laws and regulations
  - addressing ethical issues caused by AI

# **Knowledge, Belief, and Reasoning of AI**

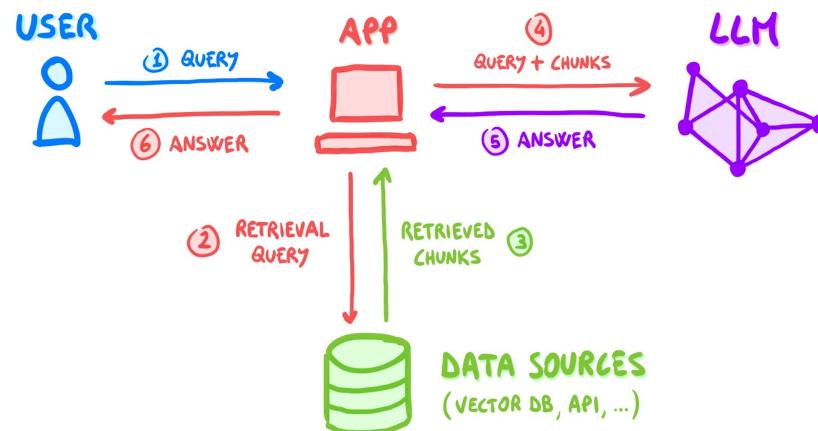
**Does LLM (or AI) have knowledge or belief? Can it reason?**

**What categories of questions should they be in?**

**Scientific, philosophical, cognitive scientific, or something else?**

## LLMs or systems in which they are embedded?

- crucial to distinguish between the two (for philosophical clarity)
  - LLM (bare-bones model) - highly specific & well-defined function, which is *conditional probability estimator*
  - systems in which LLMs are embedded - question-answering, news article summarization, screenplays generation, language translation



## Three surprises of LLM

- LLM is very different sort of animal . . . except that it is *not* an animal!
- *unreasonable* effectiveness of data [HNF09]
  - *performance scales with size of training data*
  - *qualitative leaps* in capability as models scale
  - tasks demanding human intelligence *reduced to next token prediction*
- focus on third surprise
  - “*conditional probability model looks like human with intelligence*”
  - making vulnerable to anthropomorphism
- examine it by throwing questions
  - “*does LLM have knowledge and belief?*”
  - “*can it reason?*”

## What LLM really does!

- given prompt “the first person to walk on the Moon was”, LLM responds with “Neil Armstrong”. . . strictly speaking
  - it’s *not* being asked *who* was the first person to walk on the Moon
  - what are being *really* asked is “*given statistical distribution of words in vast public corpus of text, what words are most likely to follow ‘The first person to walk on the Moon was’?*”
- given prompt “after ring was destroyed, Frodo Baggins returned to”, LLM responds with “the Shire”
  - on one level, it seems fair to say, you might be testing LLM’s knowledge of fictional world of Tolkien’s novels
  - what are being *really* asked is “*given statistical distribution of words in vast public corpus of text, what words are most likely to follow ‘After the ring was destroyed, Frodo Baggins returned to’?*”

## Knowledge, belief & reasoning around LLM

- *not* easy topic to discuss, or even impossible because
  - we do *not* have agreed definition of these terms especially in context of being asked questions like

*does LLM have belief?*  
or  
*do humans have knowledge?*
- let us discuss them in two different perspectives
  - laymen's perspective
  - cognitive scientific perspective

## Laymen's perspective on knowledge, belief & reasoning

- does (good) LLM have knowledge?
  - Grandmother - looks like it cuz when instructed “*explaining big bang*”, it says  
*“The Big Bang theory is prevailing cosmological model that explains the origin and evolution of the universe. . . . 13.8 billion years ago . . .”*
- does it have belief?
  - Grandmother: I don't think so, e.g., it does not believe in God.
- can it reason?
  - Grandmother: seems like it! e.g., when asked “*Sunghee is a superset of Alice and Beth is a superset of Sunghee. Is Beth a superset of Alice?*”, it says  
*“Yes, based on information provided, if Sunghee is a superset of Alice and Beth is a superset of Sunghee, then Beth is indeed a superset of Alice . . .”*
- can it reason to prove theorem whose inferential structure is more complicated?
  - Grandmother: I'm not sure. - actually, I don't know what you're talking about!

# Knowledge

- could argue LLM “knows” which words follow which other words with high probability
- but, only *in context of capacity to distinguish truth from falsehood*, can we legitimately speak of “knowledge” !
- LLM(-embedded BOT)
  - can be said to “encode”, “store”, or “contain” knowledge
  - lacks means to use words “true” & “false” in all ways & in all contexts because . . .
  - does not inhabit the world we human language-users share!



## Belief

- nothing can count as *belief about the world* we share unless
  - it is against backdrop of “ability to update beliefs appropriately in light of evidence from *that world*” - (again) essential capacity to distinguish truth from falsehood
- change taking place in humans when acquiring or updating belief is
  - reflection of their nature as language-using animals inhabiting shared world with community of language-users
  - humans are natural home of talk of beliefs, *etc.*, and behavioural expectations going hand-in-hand with such talk are grounded in mutual understanding
- so *what if LLM-embedded system updates LLM with outside world information?*
- even so, when interacting with AI systems based on LLMs, these grounds are *absent!*



can you provide me up-to-date information about AI agent perplexity?

❖ Sources

- Perplexity AI: A Guide for Beginners - Jeff Su
- Perplexity AI's value to soar to \$90m as it finalizes major funding round - allonwefund
- Perplexity AI: A Deep Dive - Reflections on AI
- anonyme
- Show all

❖ Perplexity

Perplexity AI is an AI-powered search engine and answer engine that has gained significant traction in recent years. Founded in 2022, Perplexity uses large language models (LLMs) to generate answers to user queries, citing sources from the web within the text response. As of Q1 2024, Perplexity had reached 15 million monthly users.

Key features and developments of Perplexity AI include:

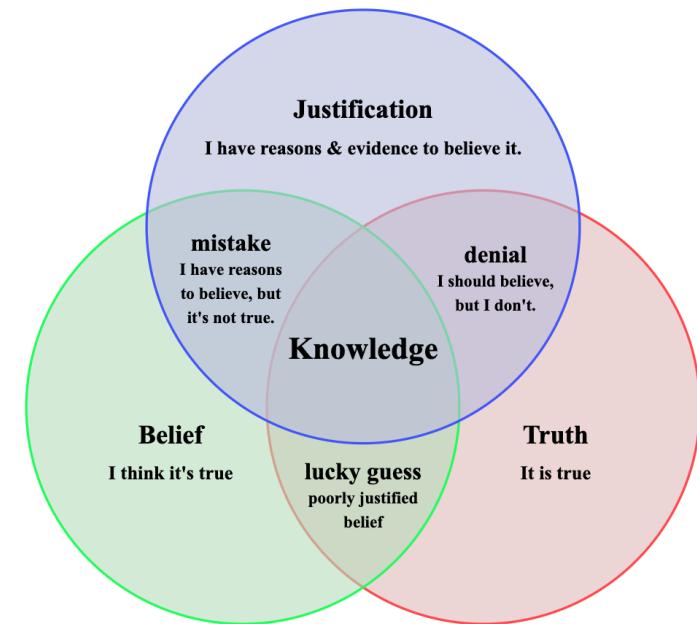
1. Search functionality: Perplexity summarizes search results and produces text with inline citations, providing personalized results based on the context of user queries

Ask follow-up Pro Generate Image

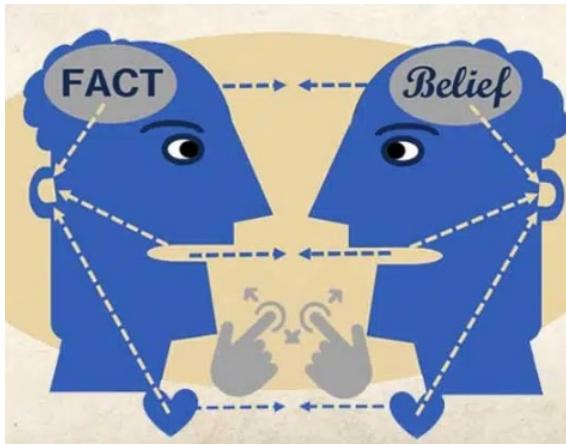


## Cognitive scientific perspective on knowledge

- does LLM have knowledge?
  - Sunghee: I don't think so.
- why?
  - Sunghee: we say we have "knowledge" when  
*"we do so against ground of various human capacities that we all take for granted when we engage in everyday conversation with each other."*
  - LLM *cannot* do this.
  - Sunghee: also when asked "*who is Tom Cruise's mother?*", it says "*Tom Cruise's mother is Mary Lee Pfeiffer.*" However, this is nothing but "*guessing*" by *conditional probability model* the most likely following words after "*Tom Cruise's mother is.*"
  - Sunghee: so we *cannot* say it really knows the fact!



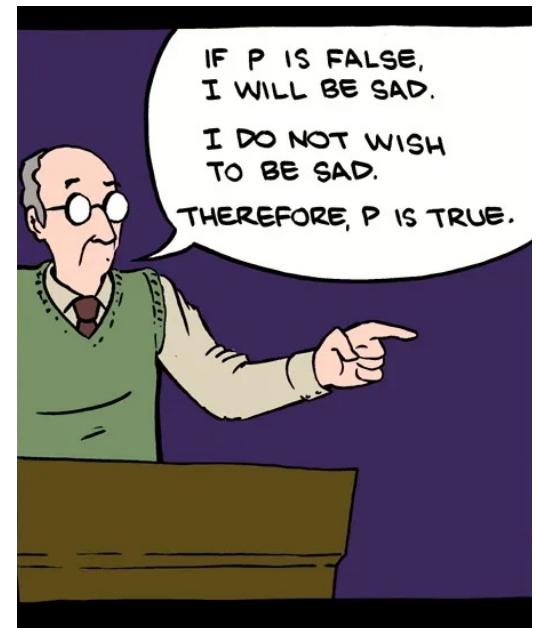
## Cognitive scientific perspective on belief



- for the discussion
  - we do not concern *any specific belief*
  - we concern prerequisites for ascribing any beliefs to AI system
- so does it have belief?
  - Sunghee: nothing can count as belief about the world we share unless  
*it is against ground of the ability to update beliefs appropriately in light of evidence from that world, an essential aspect of the capacity to distinguish truth from falsehood.*
  - Sunghee: when a human being takes to Wikipedia and confirms some fact, what happens is not her language model update, but  
*reflection of her nature as language-using animal inhabiting shared world with a community of other language-users.*
  - Sunghee: LLM does not have this ground, an essential consideration when deciding whether it *really* had beliefs.
  - Sunghee: so no, *LLM cannot have belief!*

## Cognitive scientific perspective on reasoning

- note reasoning is *content neutral*
  - e.g., following logic is perfect regardless of truth of premises  
*if Socrates is a human and humans are immortal, then Socrates would have survived today.*
- Sunghee: when asked “*if humans are immortal, would Socrates have survived today?*”, LLM says
  - “ . . . it’s logical to conclude that Socrates would likely still be alive today. . . . ”
  - however, remember, once again, what we just asked it to do is *not* “deductive inference”, but  
*given the statistical distribution of words in public corpus, what words are likely to follow the sequence, “humans are immortal and Socrates is human therefore.”*
- Sunghee: so LLM *cannot* or rather *does not* reason
- however, LLM can *mimic even multi-step reasoning whose inferencing structure is complicated* using *in-context learning* or *few-short prompting!*



## A simple example supporting reasoning incapability



- You  
*"Who is Tom Cruise's mother?"*
- LLM(-embedded question-answering system) (as of Jan 2022)  
*"Tom Cruise's mother is Mary Lee Pfeiffer. She was born Mary Lee South. . . . Information about his family, including his parents, has been publicly available, . . . "*
- You  
*"Who is Mary Lee Pfeiffer's son?"*
- LLM(-embedded question-answering system) (as of Jan 2022)  
*"As of my last knowledge update in January 2022, I don't have specific information about Mary Lee Pfeiffer or her family, including her son. . . . "*

# **Future of AI**

## Aschenbrenner's essay

- Leopold Aschenbrenner, who left OpenAI showing concerns about safety, wrote *epic 165-page treatise* - Jun-2024
  - rapid progress
    - AI development (is) accelerating at unprecedented rate, predicting by 2027, AI models lead to intelligence explosion surpassing human intelligence
  - economic and security implications
    - trillions of dollars being invested into infrastructure supporting AI systems
    - critical need for securing technologies to prevent misuse, *e.g.*, by state actors like Chinese Communist Party (CCP)
  - technical and ethical challenges
    - significant challenges in controlling AI (smarter than humans), *i.e.*, “superalignment” problem, to prevent catastrophic outcomes
  - predictions and societal impact
    - few people truly understand scale of change by AI
    - potential for AI to reshape industries, enhance national security
    - pose new ethical and governance challenges

## More about Aschenbrenner's essay

- AGI by 2027
  - seen AI advancing from preschool-level to high-schooler abilities in 4 years highlighting rapid progress from GPT-2 to GPT-4
- superintelligence following AGI - post AGI
  - rapid advancement from human-level to superhuman capabilities
- G-dollar investment on AI clusters
- national & global security dynamics
  - may lead to all-out war, *e.g.*, with China, if not managed properly
- superalignment challenges
  - keeping superintelligent AI aligned with human values and interests - “one of the most critical predictions”
- societal and economic transformations, project involvement by US government, technological mobilization

**Moral**

## Moral

- AI, *e.g.*, LLM, shows incredible utility and commercial potentials, hence we should
  - make informed decisions about trustworthiness and safety
  - avoid ascribing capacities they lack
- today's AI is so powerful, so (seemingly) convincingly intelligent
  - obfuscate mechanism
  - actively encourage *anthropomorphism* with philosophically loaded words like “believe” and “think”
  - easily mislead people about character and capabilities of AI
- matters not only to scientists, engineers, developers, and entrepreneurs, but also
  - *general public, policy makers, media people*

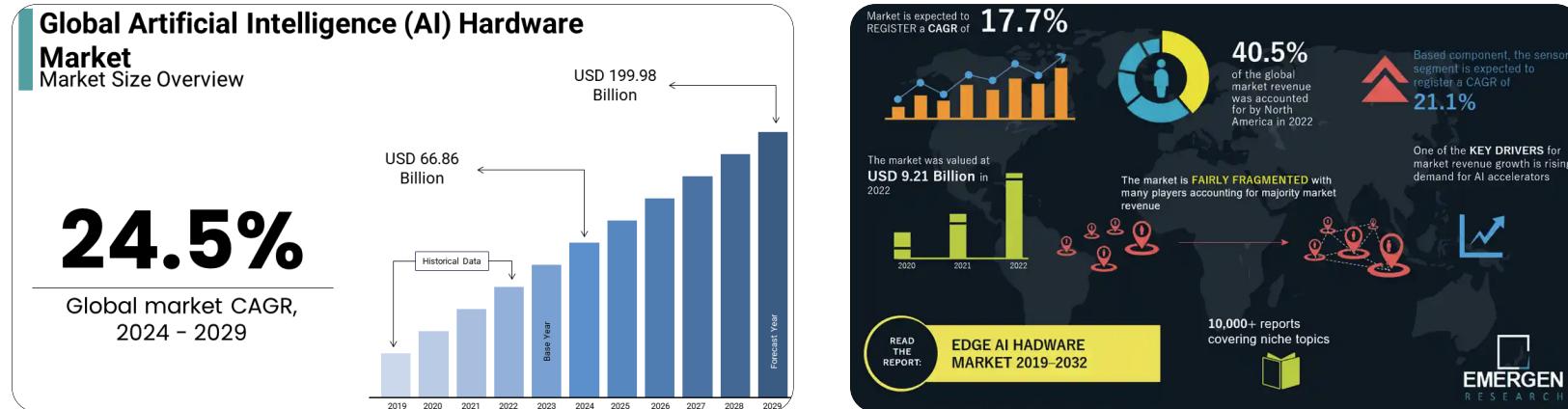
# **Appendix**

# **AI Hardware**

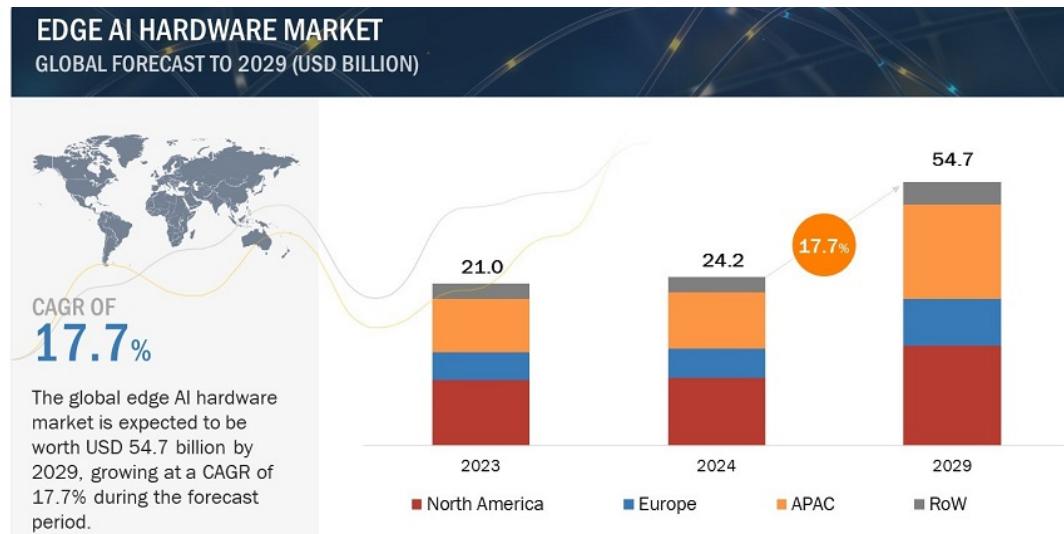
# **AI Hardware Industry**

## Landscape of AI hardware industry

- global AI hardware market valued at \$66.96B in 2024, projected to grow significantly
- major companies - Nvidia, Intel, AMD, Qualcomm, and IBM w/ Nvidia holding substantial market share



- North America leading market - high R&D investments & key industry players
- Asia Pacific rapidly expanding - strong semiconductor industries in South Korea, China & Japan
- demand for advanced processors such as GPUs, TPUs & AI accelerators rising due to complexity of AI algorithms & high computational power



## Predictions for future of AI hardware market

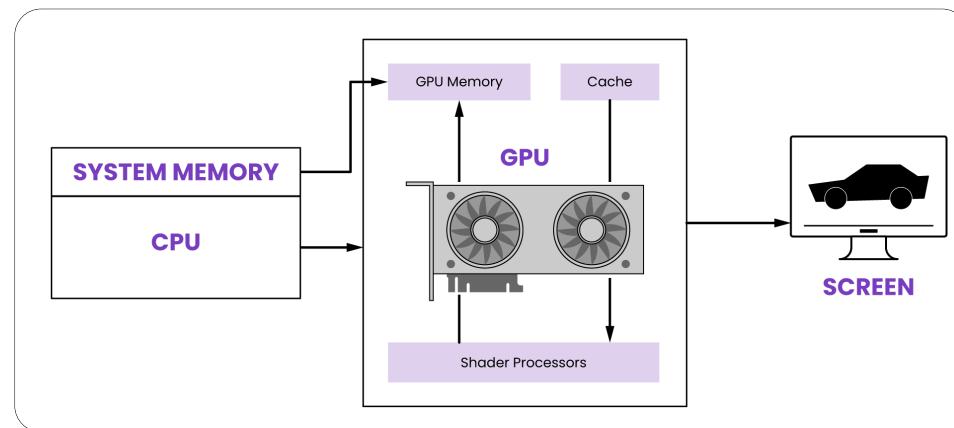
- AI hardware market expected to reach \$382B by 2032 - significant growth in data center AI chips
- integration of AI w/ 5G & increased use of AI in edge computing anticipated to drive future demand
- AI hardware becoming crucial in sectors such as autonomous vehicles, robotics & medical devices
- need to address challenges such as heat and power management along with technical complexities



# **GPUs and AI Accelerators**

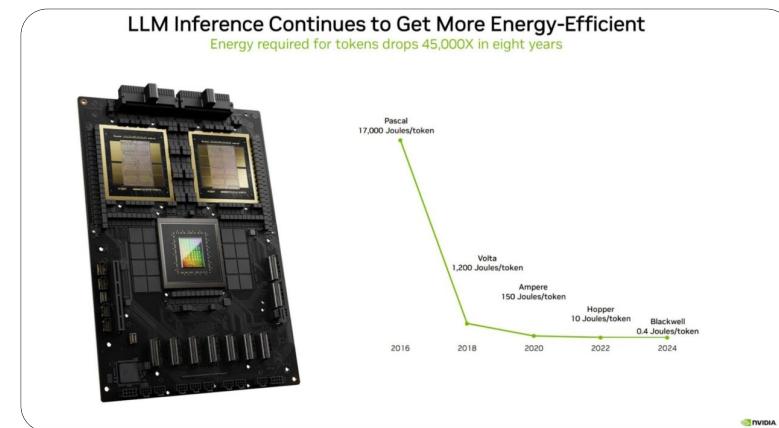
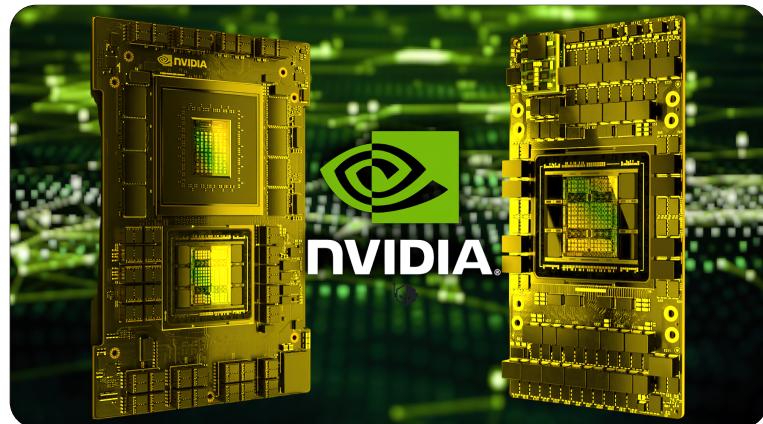
## Technical challenges of GPUs & AI accelerators

- facing challenges in scaling to handle increasingly large AI models and datasets - traditional architectures struggling w/ massive parallel processing demands of modern AI applications
- AI applications require extensive memory bandwidth often leading to bottlenecks - efficient memory management is crucial
- AI accelerators consume significant power - high operational costs and environmental concerns for both cloud-based & edge AI applications



## Potential solutions for overcoming challenges

- development of AI-specific architectures such as tensor cores and custom ASICs to improve efficiency and performance - novel architectures like FPGAs for specific AI tasks, *e.g.*, for RAG & vectorDB
- implementing software optimizations to enhance hardware usability and performance - use of compilers and frameworks that maximize efficiency of existing hardware
- encouraging market competition to drive innovation and reduce monopolistic control - exploring alternative hardware solutions and improving energy efficiency standards



## **Big tech's in-house chip development**

- shift towards in-house AI hardware - major tech companies increasingly developing their own AI chips - move to enhance AI capabilities and reduce dependence
- collaboration with specialized partners - partnering with specialized firms for manufacturing and technology blending in-house expertise with external innovation

	Microsoft	Google	Amazon	Meta
Chip	Maia 100	TPU v5e	Inferentia2	MTIA v1
Launch Date	November, 2023	August, 2023	Early 2023	2025
IP	ARM	ARM	ARM	RISC-V
Process Technology	TSMC 5nm	TSMC 5nm	TSMC 7nm	TSMC 7nm
Transistor Count	105 billion	-	-	-
INT8	-	393 TOPS	-	102.4 TOPS
FP16	-	-	-	51.2 TFLOPS
BF16	-	197 TFLOPS	-	-
Memory	-	-	-	LPDDR5
TDP	-	-	-	25W
Packaging Technology	CoWoS	CoWoS	CoWoS-S	2D
Collaborating Partners	Global Unichip Corp.	Broadcom	Alchip Technologies	Andes Technology
Application	Training/Inference	Inference	Inference	Training/Inference
LLM	GPT-3.5, GPT-4	BERT, PaLM, LaMDA	Titan FM	Llama, Llama2

## AMD - Nvidia's new competitor

- key points
  - AMD launched new AI accelerator chip, *Instinct MI300X*, on Dec 6, 2023
  - CDNA 3 architecture, mix of 5nm and 6nm IPs, delivering 153B transistors
  - *outperforms Nvidia's H100 TensorRT-LLM* by 1.6X higher memory bandwidth and 1.3X FP16 TFLOPS
  - up to 40% faster vs Nvidia's Llama-2 70B model in 8x8 server configurations
- market impact
  - significant challenge to Nvidia's dominance in AI accelerator market
  - performance gains over Nvidia's offerings could drive *customer adoption and market share for AMD*
- future prediction
  - *AMD stocks soared* since launch indicating investor confidence in their competitiveness
  - Lisa Su, AMD's CEO, categorized Instinct MI300X as “next big thing” in tech industry
  - potential risks include need to *manage ROCm vs CUDA software ecosystem* & ensure rapid customer adoption and production coverage

# **AI Accelerator Startups**

## AI accelerator startups

- innovative architectures - startups like Groq, SambaNova & Graphcore leading with *novel architectures designed to accelerate AI workloads*
  - *Groq* - tensor streaming processor (TSP) offering ultra-low latency & high throughput, high-performance AI inference chips enhancing speed & efficiency
  - *SambaNova* - reconfigurable dataflow architecture optimizing for various AI workloads
  - *Graphcore* - intelligence processing unit (IPU) tailored for graph-based computation excelling in sparse data processing
  - *Cerebras Systems* - develop wafer scale engine (WSE), largest chip built for AI workloads, unmatched computational power revolutionizing AI hardware capabilities
  - *Hailo* - specialize for edge devices optimizing AI processes for real-time applications, raised \$120M emphasizing potential to disrupt traditional AI chip markets

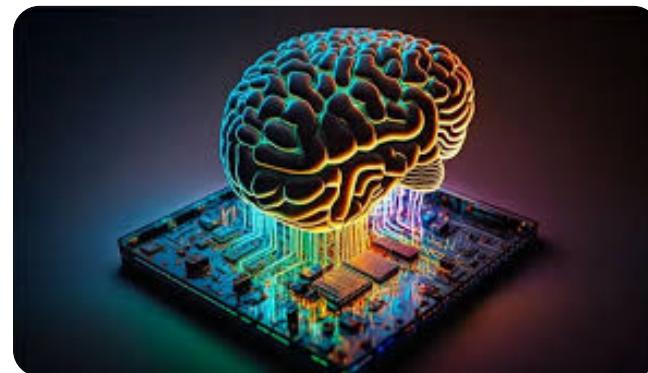


## Technological competitiveness

- energy efficiency
  - energy-efficient designs crucial for scalability in data centers and edge devices
  - startups developing solutions significantly reducing power consumption without compromising performance
- customization & flexibility
  - AI accelerators from startups often offer greater customization options for specific AI tasks compared to traditional GPUs
  - flexibility in hardware allows for tailored solutions that can outperform general-purpose accelerators in certain applications
- software integration
  - robust software ecosystems critical - startups investing in developing software stacks that optimize performance for their hardware
  - compatibility with existing AI frameworks is competitive advantage, *e.g.*, TensorFlow & PyTorch

## Industry and market influence

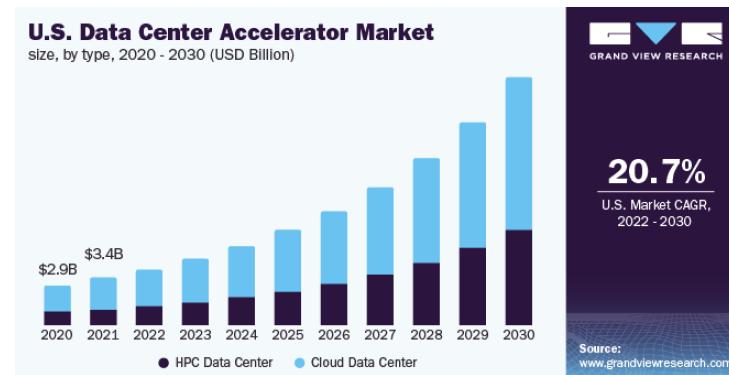
- disruption of traditional players
  - challenging dominance of established players like NVIDIA & Intel
  - unique architectures providing specialized solutions traditional GPUs and CPUs cannot efficiently handle
- driving down costs
  - offering competitive alternatives pushing down cost of AI computation
  - could lead to democratization of AI w/ more companies affording high-performance AI capabilities



- accelerating AI innovation
  - contributing to rapid innovation providing hardware that can handle emerging AI models & workloads
  - adaptability and specialization enable advancements in AI research & faster development cycles
- strategic partnerships & acquisitions
  - big techs increasingly forming strategic partnerships or acquiring startups to stay competitive
  - collaborations can speed up integration of advanced AI hardware into mainstream products



- market growth & opportunities
  - AI accelerator market expected to grow significantly driven by demand in data centers, edge computing & autonomous systems
  - startups well-positioned to capture significant share of growing market particularly in niche applications
- future outlook
  - dependency on Asia for fabrication might lead to strategic shifts in global tech policies and investments in local manufacturing
  - increasing demand for efficient AI processing on edge devices and in data center.



**AI & Biotech**

## AI in biology

- AI has been used in biological sciences, and science in general
- AI's ability to process large amounts of raw, unstructured data (*e.g.*, DNA sequence data)
  - reduces time and cost to conduct experiments in biology
  - enables others types of experiments that previously were unattainable
  - contributes to broader field of engineering biology or biotechnology
- AI increases human ability to make direct changes at cellular level and create novel genetic material (*e.g.*, DNA and RNA) to obtain specific functions.

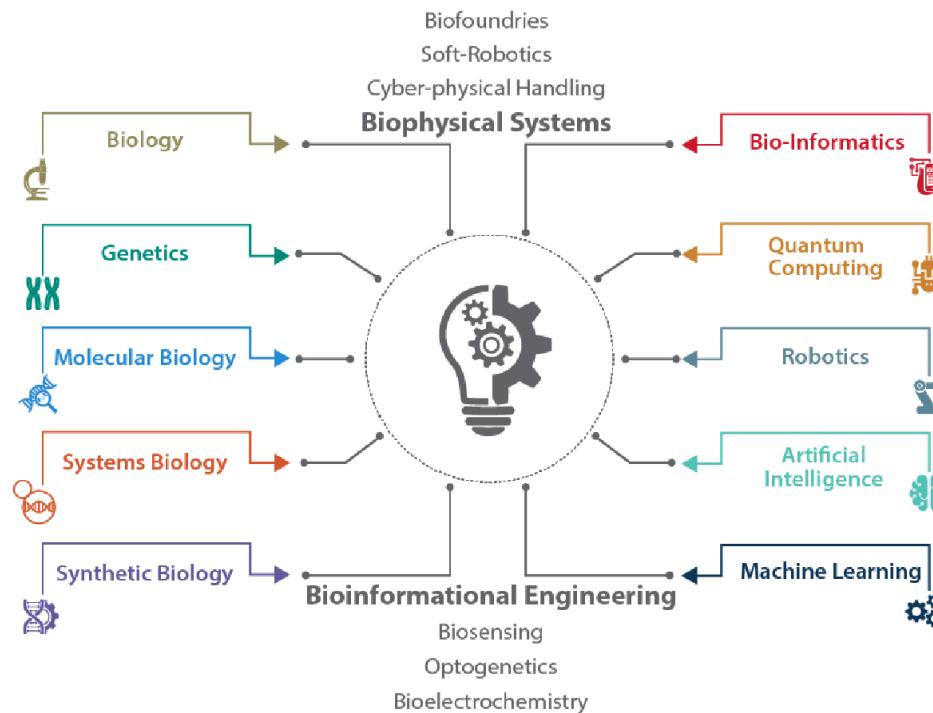
**Biotech**

## Biotech

- biotechnology
  - is multidisciplinary field leveraging broad set of sciences and technologies
  - relies on and builds upon advances in other fields such as nanotechnology & robotics, and, increasingly, AI
  - enables researchers to read and write DNA
    - sequencing technologies “read” DNA while gene synthesis technologies takes sequence data and “write” DNA turning data into physical material
- 2018 National Defense Strategy & senior US defense and intelligence officials identified emerging technologies that could have disruptive impact on US national security [[Say21](#)]
  - artificial intelligence, lethal autonomous weapons, hypersonic weapons, directed energy weapons, *biotechnology*, quantum technology
- other names for biotechnology are engineering biology, synthetic biology, biological science (when discussed in context of AI)

## biotech - multidisciplinary field

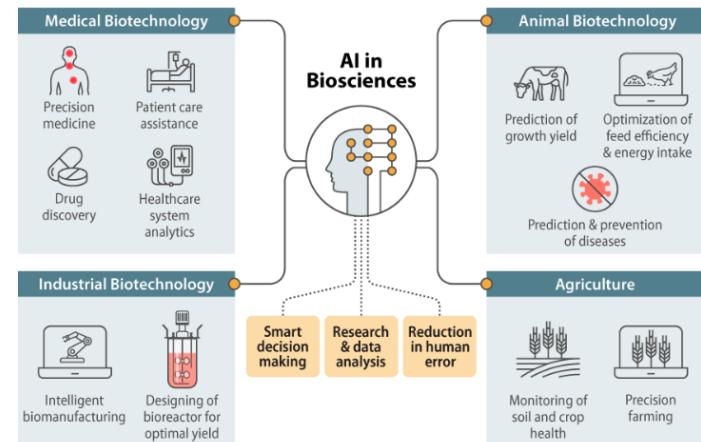
- sciences and technologies enabling biotechnology include, but not limited to,
  - (molecular) biology, genetics, systems biology, synthetic biology, bio-informatics, quantum computing, robotics [DFJ22]



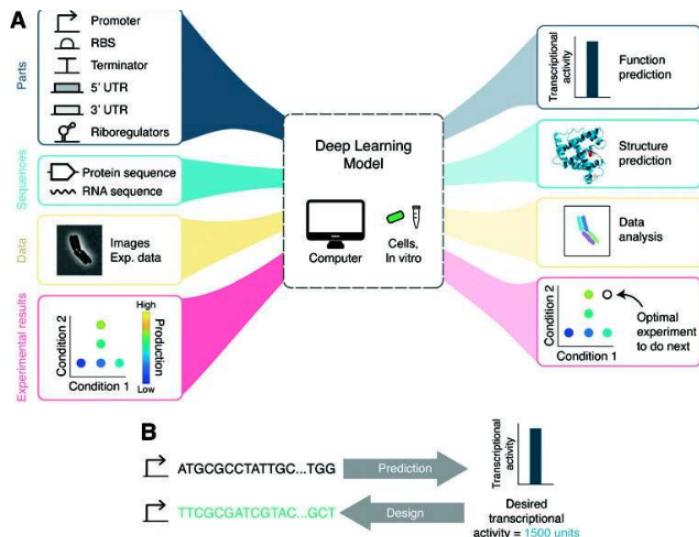
# Convergence of AI and biological design

- both AI & biological sciences increasingly converging [BKP22]
  - each building upon the other's capabilities for new research and development across multiple areas
- Demo Hassabis, CEO & cofounder of DeepMind, said of biology [Toe23]
 

“. . . biology can be thought of as information processing system, albeit extraordinarily complex and dynamic one . . . just as mathematics turned out to be the right description language for physics, biology may turn out to be *the perfect type of regime for the application of AI!*”
- Both AI & biotech rely on and build upon advances in other scientific disciplines and technology fields, such as nanotechnology, robotics, and increasingly big data (e.g., genetic sequence data)
  - each of these fields itself convergence of multiple sciences and technologies
- so *their impacts can combine to create new capabilities*



# Multi-source genetic sequence data



- AI is essential to analyzing exponential growth of genetic sequence data
  - "AI will be essential to fully understanding how genetic code interacts with biological processes"
  - US National Security Commission on Artificial Intelligence (NSCAI)
- process huge amounts of biological data, e.g., genetic sequence data, coming from different biological sources for understanding complex biological systems
  - sequence data, molecular structure data, image data, time-series, omics data
- e.g., analyze genomic data sets to determine the genetic basis of particular trait and potentially uncover genetic markers linked with that trait

## Quality & quantity of biological data

- limiting factor, however, is quality and quantity of the biological data, *e.g.*, DNA sequences, that AI is trained on
  - *e.g.*, accurate identification of particular species based on DNA requires reference sequences of *sufficient quality* to exist and be available
- databases have varying standards - access, type and quality of information
- design, management, quality standards, and data protocols for reference databases can affect utility of particular DNA sequence

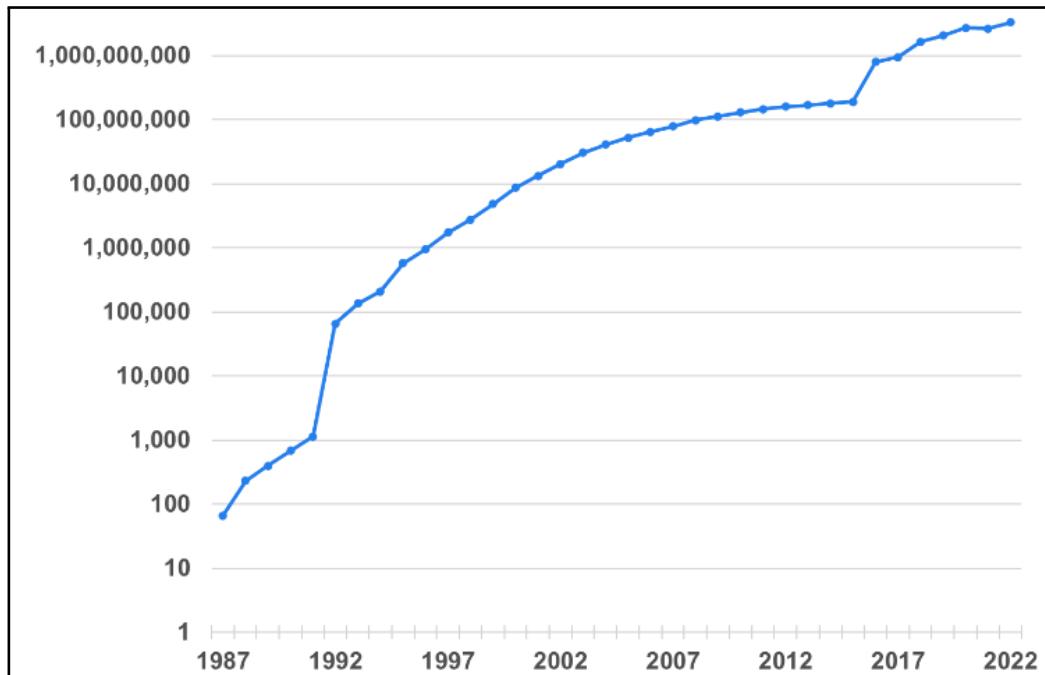
## Rapid growth of biological data

- volume of genetic sequence data grown exponentially as sequencing technology has evolved
- more than 1,700 databases incorporating data on genomics, protein sequences, protein structures, plants, metabolic pathways, *etc.*, *e.g.*
  - open-source public database
    - Protein Data Bank, US-funded data center, contains more than *terabyte of three-dimensional structure data* for biological molecules, including proteins, DNA, and RNA
  - proprietary database
    - Gingko Bioworks - possesses more than *2B protein sequences*
  - public research groups
    - Broad Institute - produces roughly *500 terabases of genomic data per month*
- great potential value in aggregate volume of genetic datasets that can be collectively mined to discover and characterize relationships among genes

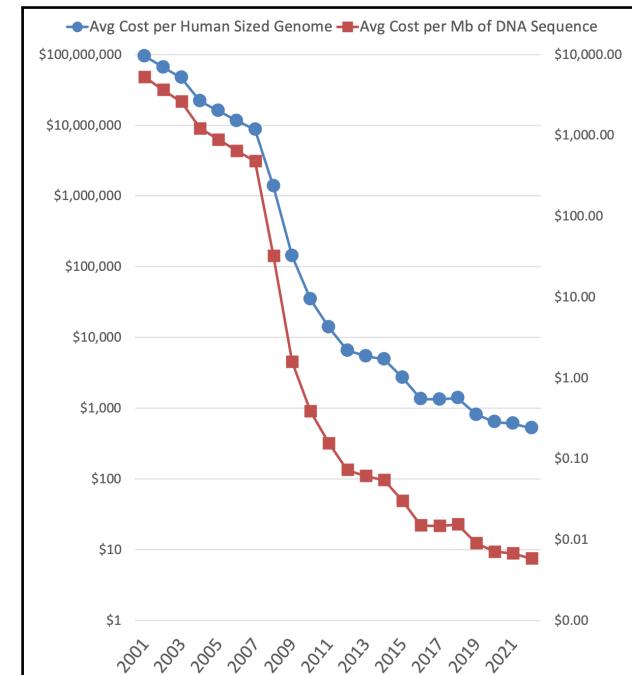
## Volume and sequencing cost of DNA over time

- volume of DNA sequences & DNA sequencing cost
  - data source: National Human Genome Research Institute (NHGRI) [[Wet23](#)] & International Nucleotide Sequence Database Collaboration (INSDC)

# sequences in INSDC



DNA sequencing cost



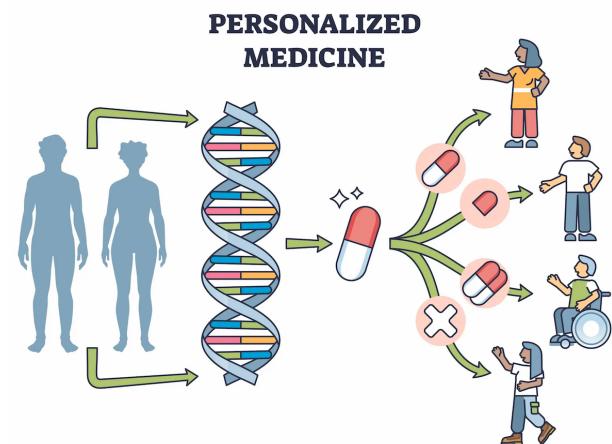
## Bio data availability and bias

- US National Security Commission on Artificial Intelligence (NSCAI) recommends
  - US fund and prioritize development of a biobank containing "*wide range of high-quality biological and genetic data sets securely accessible by researchers*"
  - establishment of database of broad range of human, animal, and plant genomes would
    - *enhance and democratize biotechnology innovations*
    - *facilitate new levels of AI-enabled analysis of genetic data*
- bias - availability of genetic data & decisions about selection of genetic data can introduce bias, e.g.
  - training AI model on datasets emphasizing or omitting certain genetic traits can affect how information is used and types of applications developed - *potentially privileging or disadvantaging certain populations*
  - access to data and to AI models themselves may impact communities of differing socioeconomic status or other factors unequally

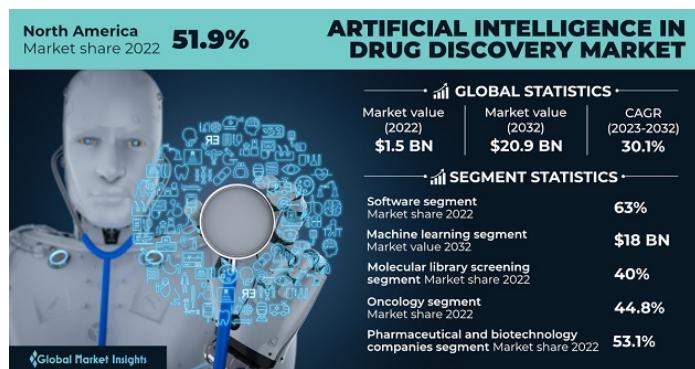
# **Emerging Trends in Biotech**

## Personalized medicine

- *shift from one-size-fits-all approach to tailored treatments*
- based on individual genetic profiles, lifestyles & environments
- AI enables analysis of vast data to predict patient responses to treatments, thus enhancing efficacy and reducing adverse effects
- e.g., custom cancer therapies, personalized treatment plans for rare diseases & precision pharmacogenomics.
- companies - Tempus, Foundation Medicine, etc.



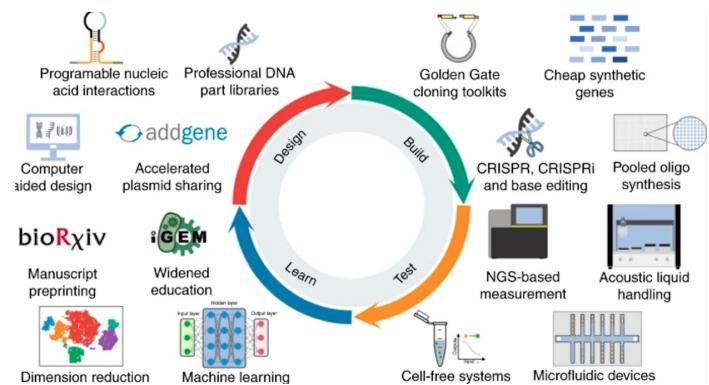
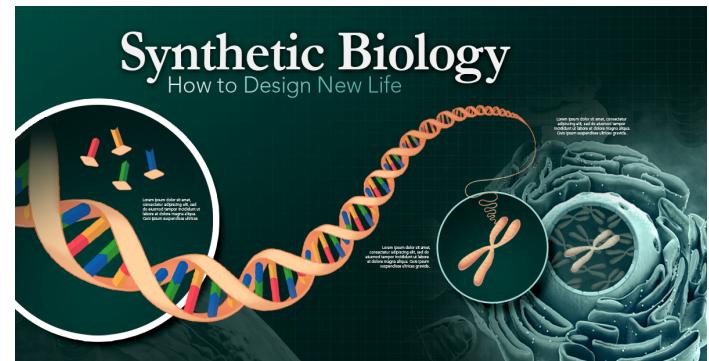
# AI-driven drug discovery



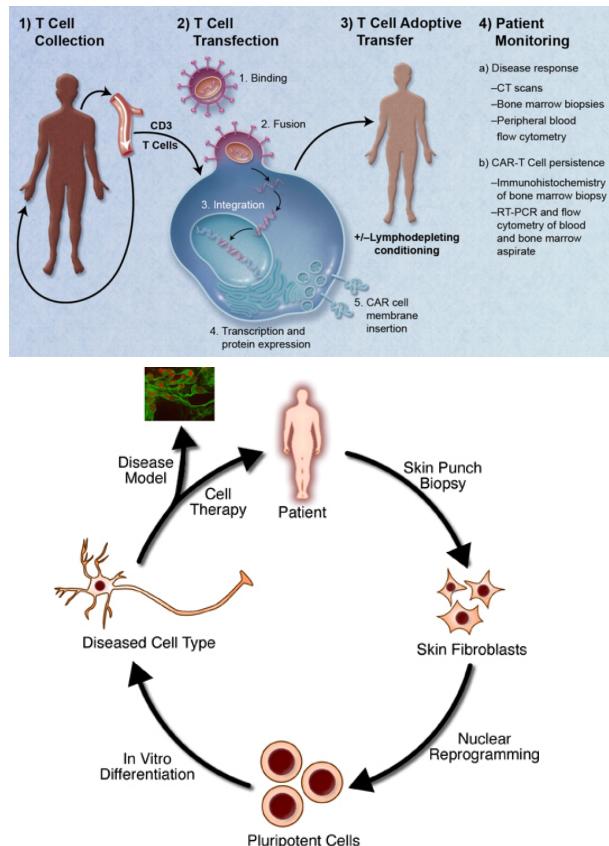
- traditional drug discovery process - time-consuming and costly often taking decades and billions of dollars
- AI streamlines this process by predicting the efficacy and safety of potential compounds with more speed and accuracy
- AI models analyze chemical databases to identify new drug candidates or repurpose existing drugs for new therapeutic uses
- companies - Insilico Medicine, Atomwise.

## Synthetic biology

- use AI for gene editing, biomaterial production and synthetic pathways
- combine principles of biology and engineering to design and construct new biological entities
- AI optimizes synthetic biology processes from designing genetic circuits to scaling up production
- company - Ginkgo Bioworks uses AI to design custom microorganisms for applications ranging from pharmaceuticals to industrial chemicals



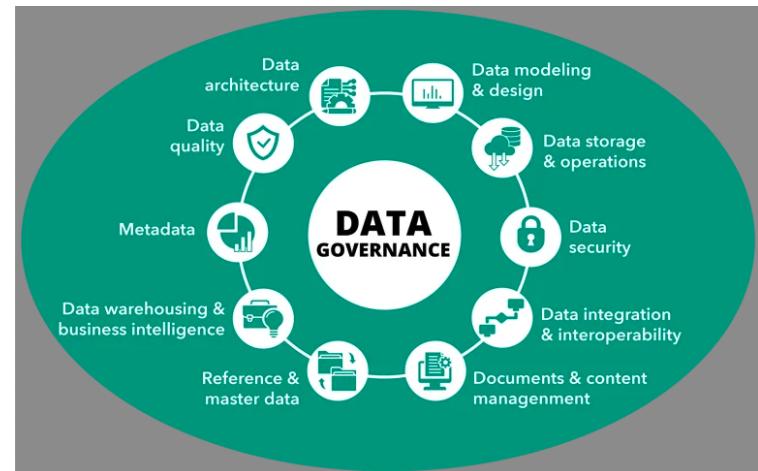
# Regenerative medicine



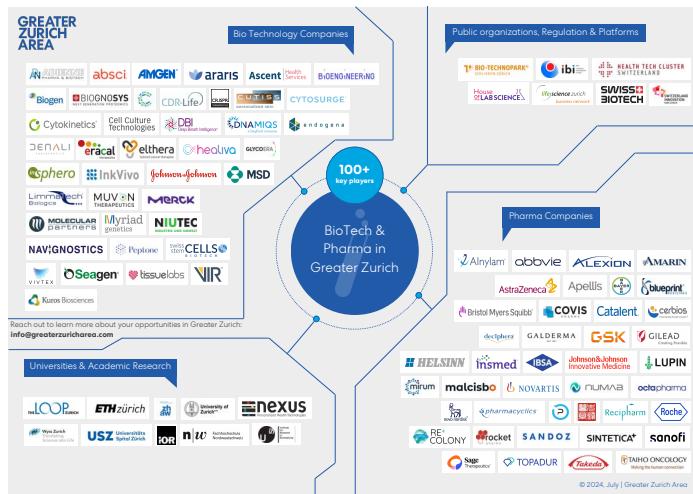
- AI advances development of stem cell therapies & tissue engineering
- AI algorithms assist in identifying optimal cell types, predicting cell behavior & personalized treatments
- particularly for conditions such as neurodegenerative diseases, heart failure and orthopedic injuries
- company - Organovo leverages AI to potentially improve the efficacy and scalability of regenerative therapies, developing next-generation treatments

## Bio data integration

- integration of disparate data sources, including genomic, proteomic & clinical data - one of biggest challenges in biotech & healthcare
- AI delivers meaningful insights *only when* seamless data integration and interoperability realized
- developing platforms facilitating comprehensive, longitudinal patient data analysis - vital enablers of AI in biotech
- company - Flatiron Health working on integrating diverse datasets to provide holistic view of patient health



# Biotech companies



- Atomwise - small molecule drug discovery
- Cradle - protein design
- Exscientia - precision medicine
- Iktos - small molecule drug discovery and design
- Insilico Medicine - full-stack drug discovery system
- Schrödinger, Inc. - use physics-based models to find best possible molecule
- Absci Corporation - antibody design, creating new from scratch antibodies, *i.e.*, “*de novo* antibodies”, and testing them in laboratories

# **References**

## References

- [BKP22] Abhaya Bhardwaj, Shristi Kishore, and Dhananjay K. Pandey. Artificial intelligence in biological sciences. *Life*, 12(1430), 2022.
- [DFJ22] Thomas A. Dixon, Paul S. Freemont, and Richard A. Johnson. A global forum on synthetic biology: The need for international engagement. *Nature Communications*, 13(3516), 2022.
- [HGH<sup>+</sup>22] Sue Ellen Haupt, David John Gagne, William W. Hsieh, Vladimir Krasnopolksky, Amy McGovern, Caren Marzban, William Moninger, Valliappa Lakshmanan, Philippe Tissot, and John K. Williams. The history and practice of AI in the environmental sciences. *Bulletin of the American Meteorological Society*, 103(5):E1351 – E1370, 2022.
- [HM24] Guadalupe Hayes-Mota. Emerging trends in AI in biotech. *Forbes*, June 2024.
- [HNF09] Alon Halevy, Peter Norvig, and Nanediri Fernando. The unreasonable effectiveness of data. *Intelligent Systems, IEEE*, 24:8 – 12, 05 2009.

- [Kah11] Daniel Kahneman. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, 2011.
- [Kui23] Todd Kuiken. Artificial intelligence in the biological sciences: Uses, safety, security, and oversight. *Congressional Research Service*, Nov 2023.
- [KXS<sup>+</sup>24] Tzofi Klinghoffer, Xiaoyu Xiang, Siddharth Somasundaram, Yuchen Fan, Christian Richardt, Ramesh Raskar, and Rakesh Ranjan. Platonerf: 3D reconstruction in Plato's cave via single-view two-bounce lidar, 2024.
- [Say21] Kelley M. Sayler. Defense primer: Emerging technologies. *Congressional Research Service*, 2021.
- [Sha23] Murray Shanahan. Talking about large language models, 2023.
- [SSS<sup>+</sup>24] Shunsuke Saito, Gabriel Schwartz, Tomas Simon, Junxuan Li, and Giljoo Nam. Relightable Gaussian codec avatars, 2024.
- [Toe23] Rob Toews. The next frontier for large language models is biology. *Forbes*, July 2023.
- [Wet23] Kris A. Wetterstrand. Dna sequencing costs: Data, 2023.
- [ZBX<sup>+</sup>24] Siwei Zhang, Bharat Lal Bhatnagar, Yuanlu Xu, Alexander Winkler, Petr

Kadlecik, Siyu Tang, and Federica Bogo. Rohm: Robust human motion reconstruction via diffusion, 2024.

**Thank You**