[2026 Professional Training for Science High School Admission Officers]
Identifying Tomorrow's Scientists - Domain Curiosity, AI Adaptability, and Ethical Problem-Solving

Sunghee Yun

Co-Founder & CTO @ Erudio Bio, Inc.
Co-Founder & CEO @ Erudio Bio Korea, Inc.
Founder & Leader of Silicon Valley AI Nexus
Global Leadership Initiative Fellow @ Salzburg Global Seminar
Visiting & Advisory Professor @ Sogang University & DGIST

About Speaker

- *Co-Founder & CTO @ Erudio Bio, Inc., San Jose & Novato, CA, USA* 2023 ∼
- *Co-Founder & CEO @ Erudio Bio Korea, Inc., Korea* 2025 ∼
- *Leader of Silicon Valley Privacy-Preserving AI Forum (K-PAI), CA, USA* 2024 ∼
- *Advisor to Korean American Semiconductor Professional Alliance (KASPA)* 2026 ∼
- *CGO / Global Managing Partner @ LULUMEDIC, Seoul, Korea* 2025 ∼
- *KFAS-Salzburg Global Leadership Fellow @ Salzburg Global Seminar, Austria* 2024 ∼
- *Adjunct Professor, EE Department @ Sogang University, Seoul, Korea* 2020 ∼
- *Advisory Professor, EECS Department @ DGIST, Korea* 2020 ∼
- *AI-Korean Medicine Integration Initiative Task Force Member @ The Association of Korean Medicine, Seoul, Korea* 2025 ∼
- *Director of AI Semiconductor @ K-BioX, CA, USA* 2025 ∼
- Global Advisory Board Member @ Innovative Future Brain-Inspired Intelligence System Semiconductor of Sogang University, Korea 2020 ∼
- Technology Consultant @ Gerson Lehrman Gruop (GLG), NY, USA 2022 ∼
- Advisor @ CryptoLab, Inc., Seoul, Korea 2025 ∼

- Co-Founder & CTO / Head of Global R&D / Chief Applied Scientist / Senior Fellow @ Gauss Labs, Inc., Palo Alto, CA, USA 2020 $\sim$ 2023
- Senior Applied Scientist @ Amazon.com, Inc., Vancouver, BC, Canada 2017 $\sim$ 2020
- Principal Engineer @ Software R&D Center, Samsung Electronics 2016 $\sim$ 2017
- Principal Engineer @ Strategic Marketing & Sales, Memory Business 2015 $\sim$ 2016
- Principal Engineer @ DT Team, DRAM Development, Samsung 2012 $\sim$ 2015
- Senior Engineer @ CAE Team, Memory Business, Samsung, Korea 2005 $\sim$ 2012
- PhD - Electrical Engineering @ Stanford University, CA, USA 2001 $\sim$ 2004
- Development Engineer @ Voyan, Santa Clara, CA, USA 2000 $\sim$ 2001
- MS - Electrical Engineering @ Stanford University, CA, USA 1998 $\sim$ 1999
- BS - Electrical & Computer Engineering @ Seoul National University 1994 $\sim$ 1998

Highlight of Career Journey

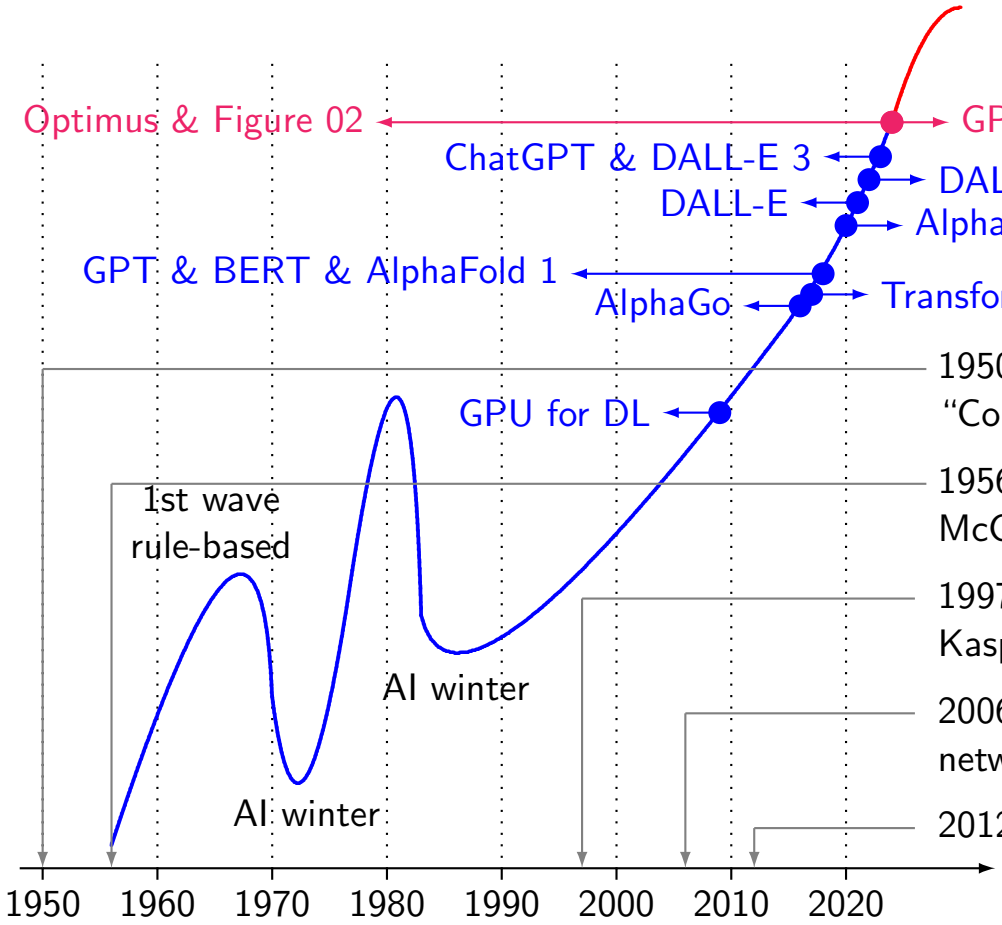
- BS in Electrical Engineering (EE) @ Seoul National University
- MS & PhD in Electronics Engineering (EE) @ Stanford University
  - *Convex Optimization - Theory, Algorithms & Software*
  - advisor - *Prof. Stephen P. Boyd*
- Principal Engineer @ Samsung Semiconductor, Inc.
  - *AI & Convex Optimization*
  - collaboration with *DRAM/NAND Design/Manufacturing/Test Teams*
- Senior Applied Scientist @ Amazon.com, Inc.
  - *e-Commerce AIs* - anomaly detection, deep RL, and recommender system
  - *Jeff Bezos's project - drove $200M* in sales via Amazon Mobile Shopping App
- *Co-Founder & CTO / Global R&D Head & Chief Applied Scientist* @ Gauss Labs, Inc.
- *Co-Founder & CTO* @ Erudio Bio, Inc.
- *Co-Founder & CEO* @ Erudio Bio Korea, Inc.

Unpacking AI for Science High School Admission Officers

Artificial Intelligence

Definition and History

History



1950 - Turing Test introduced by Alan Turing in "Computing Machinery & Intelligence"

1956 - "Artificial Intelligence" coined by John McCarthy

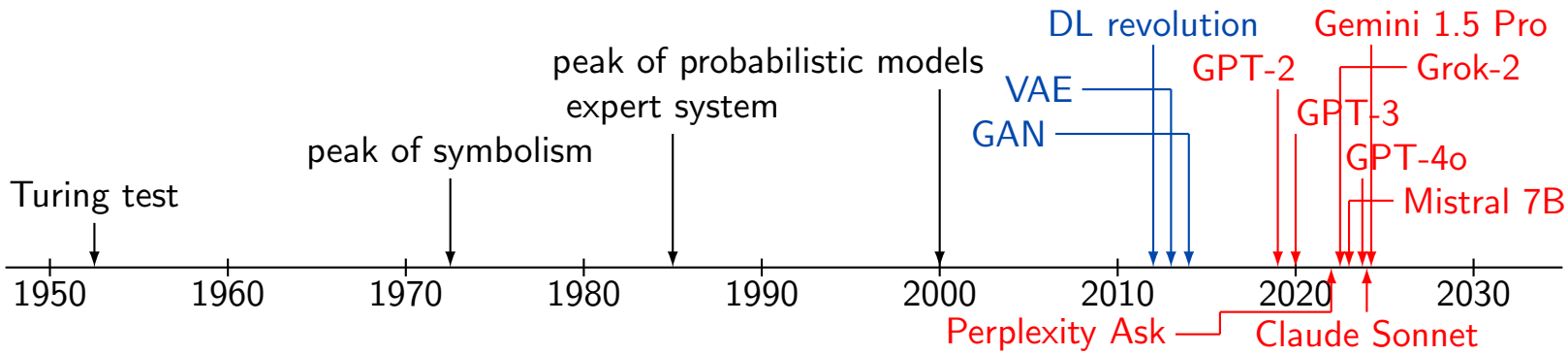
1997 - IBM's Deep Blue defeated Garry Kasparov, (then) the world chess champion

2006 - DL Learning gained traction deep belief networked introduced by Geoffrey Hinton

2012 - AlexNet won ImageNet Competition (ILSVRC) - widespread attention on DL

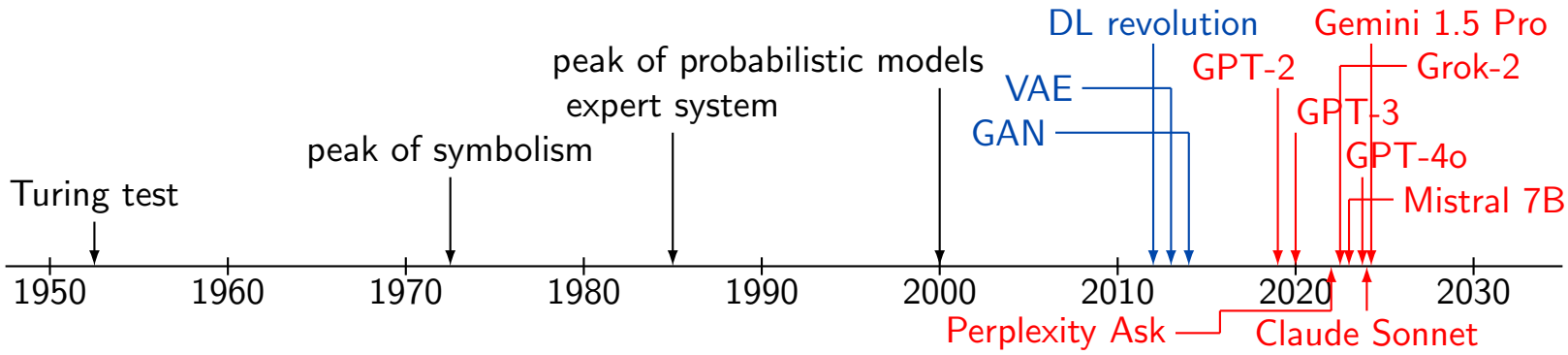
Birth of AI - early foundations & precursor technologies

- 1950s $\sim$ 1970s
  - Alan Turing - concept of *"thinking machine" & Turing test* to evaluate machine intelligence (1950s)
  - *symbolists* (as opposed to connectionists) - early AI focused on symbolic reasoning, logic & problem-solving - Dartmouth Conference in 1956 by *John McCarthy, Marvin Minsky, Allen Newell & Herbert A. Simon*
  - precursor technologies - genetic algorithms (GAs), Markov chains & *hidden Markov models (HMMs)* - laying foundation for generative processes (1970s $\sim$)



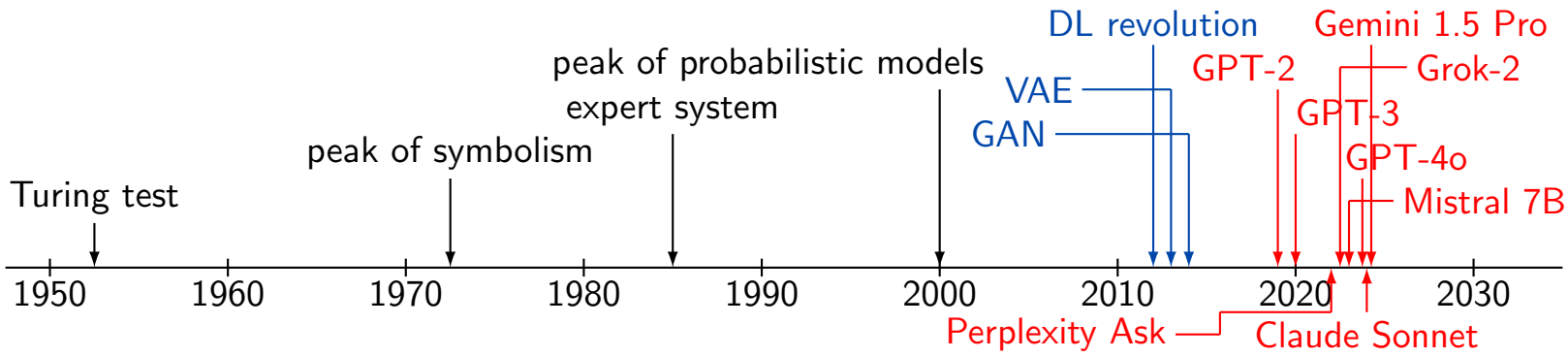
Rule-based systems & probabilistic models

- 1980s $\sim$ early 2000s
  - *expert systems* (1980s) - AI systems designed to mimic human decision-making in specific domains
  - development of neural networks (NN) w/ backpropagation *training multi-layered networks* - setting stage for way more complex generative models
  - *probabilistic models* (including network models, $i.e.$, Bayesian networks) & Markov models - laying groundwork for data generation & pattern prediction



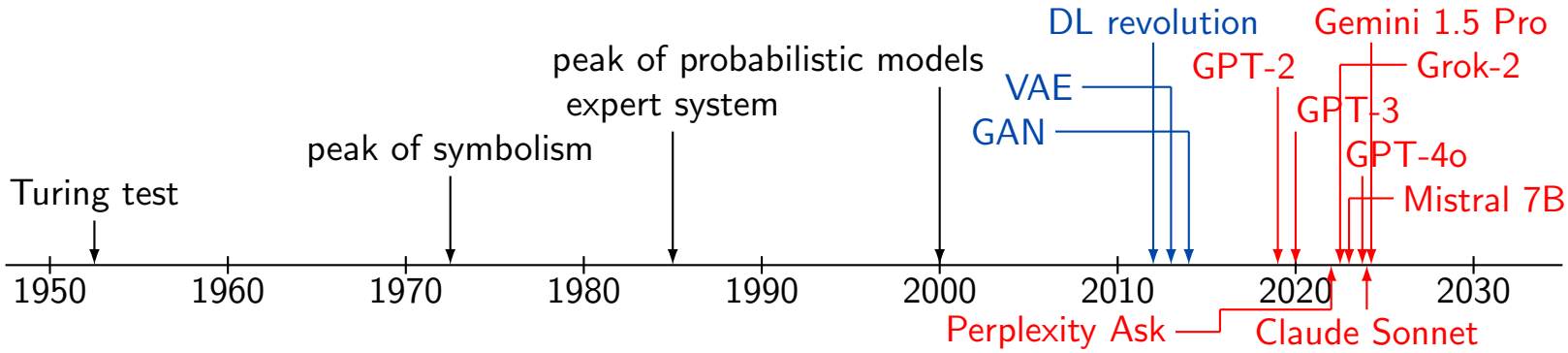
Rise of deep learning & generative models

- 2010s - breakthrough in genAI
  - *deep learning (DL) revolution* - advances in GPU computing and data availability led to the rapid development of deep neural networks.
  - *variational autoencoder (VAE)* (2013) - by Kingma and Welling - learns mappings between input and latent spaces
  - *generative adversarial network (GAN)* (2014) - by Ian Goodfellow - game-changer in generative modeling where two NNs compete each other to create realistic data
    - widely used in image generation & creative tasks



Transformer models & multimodal AI

- late 2010s $\sim$ Present
  - Transformer architecture (2017) - by Vaswani et al.
    - *revolutionized NLP,* $e.g.$, LLM & various genAI models
  - GPT series - generative pre-trained transformer
    - GPT-2 (2019) - generating human-like texts - *marking leap in language models*
    - GPT-3 (2020) - 175B params - set *new standards for LLM*
  - multimodal systems - DALL-E & CLIP (2021) - *linking text and visual data*
  - emergence of diffusion models (2020s) - new approach for generating high-quality images - progressively "denoising" random noise (DALL-E 2 & Stable Diffusion)



Significant AI Achievements - 2014 – 2025

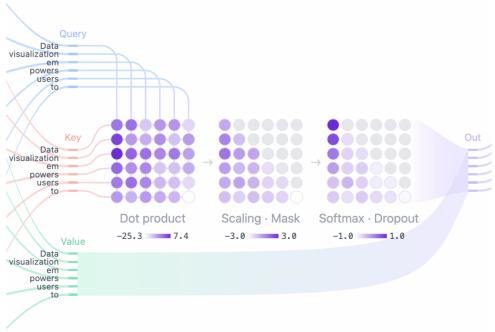
Deep learning revolution

- 2012 – 2015 - DL revolution[1]
  - 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

[1] CV: computer vision, NN: neural network, CNN: convolutional NN, RL: reinforcement learning

Transformer changes everything

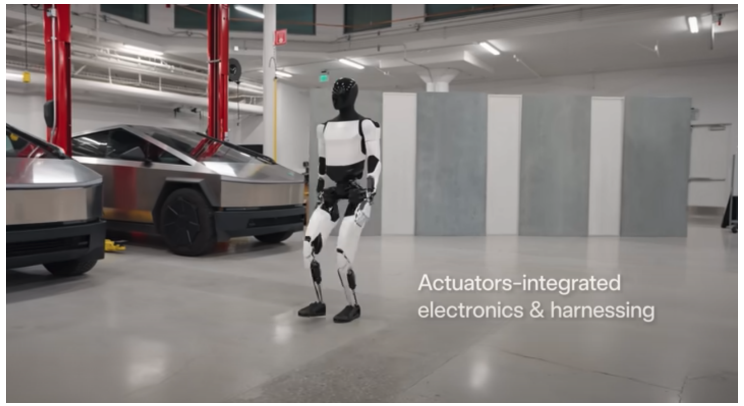
- 2017 – 2018 - Transformers & NLP breakthroughs[2]
  - *Transformer (e.g., BERT & GPT) revolutionized NLP*
  - major advancements in, *e.g.*, machine translation & chatbots
- 2020 - AI in healthcare – AlphaFold & beyond
  - DeepMind's *AlphaFold solves 50-year-old protein folding problem* predicting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



[2]NLP: natural language processing, GPT: generative pre-trained transformer

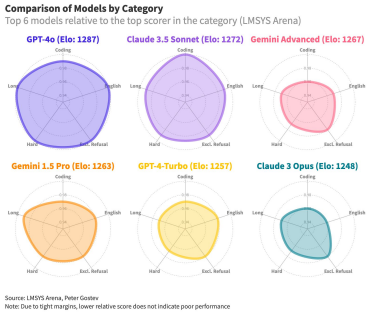
Lots of breakthroughs in AI technology and applications in 2024

- proliferation of advanced AI models
  - GPT-4o, Claude Sonnet, Claude 3 series, Llama 3, Sora, Gemini
  - *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*



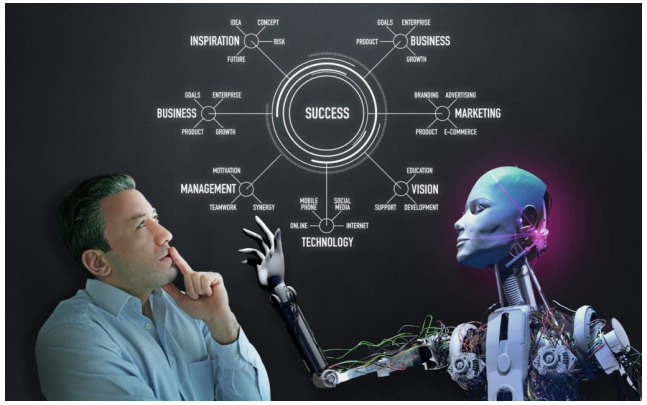
Major AI Breakthroughs in 2025

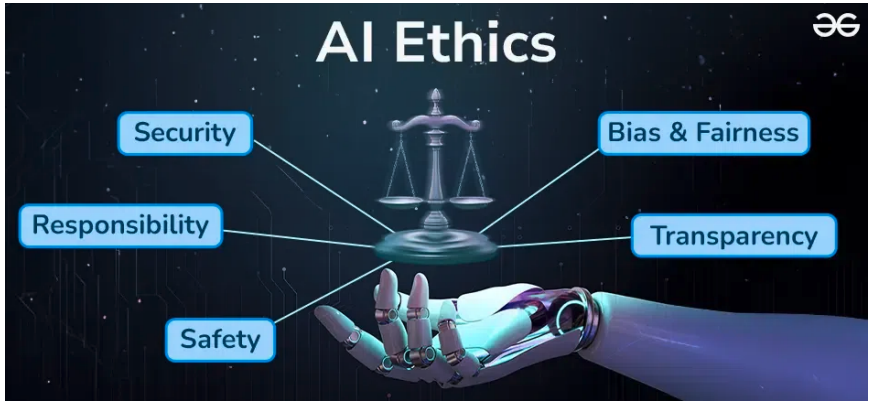
- next-generation foundation models
  - GPT-5 and Claude 4 demonstrate emergent reasoning abilities
  - open-source models achieving parity with leading commercial systems from 2024
- hardware innovations
  - NVIDIA's Blackwell successor architecture delivering 3-4x performance improvement
  - AMD's MI350 accelerators challenging NVIDIA's market dominance
- AI-human collaboration systems
  - seamless multimodal interfaces enabling natural human-AI collaboration
  - AI systems effectively explaining reasoning and recommendations
  - augmented reality interfaces providing real-time AI assistance in professional contexts



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*

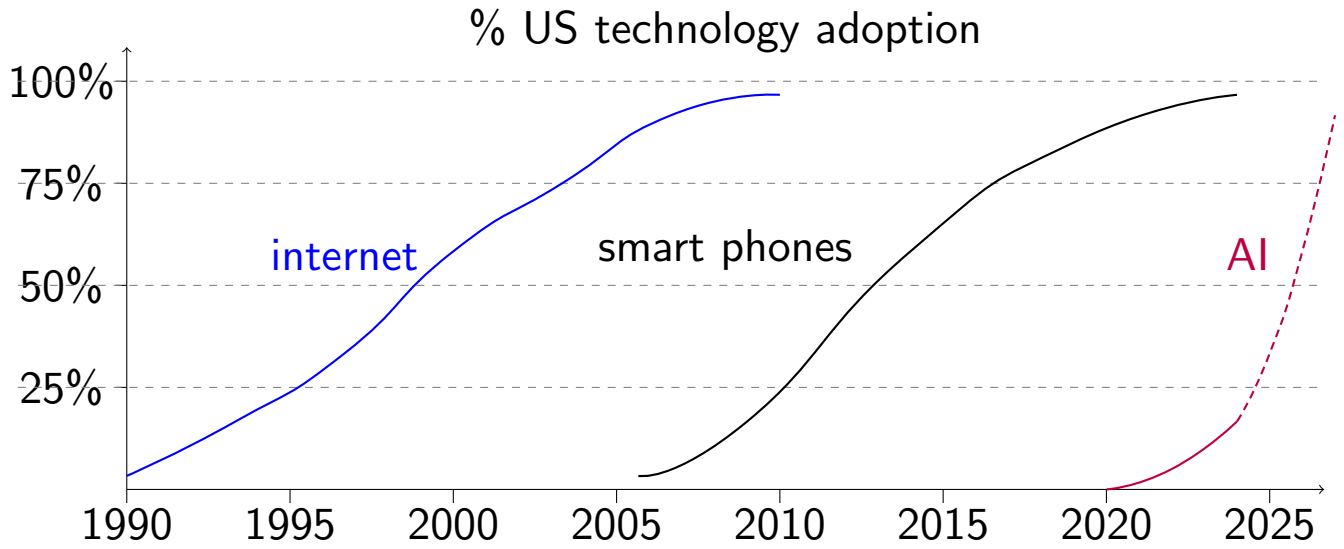




Measuring AI's Ascent

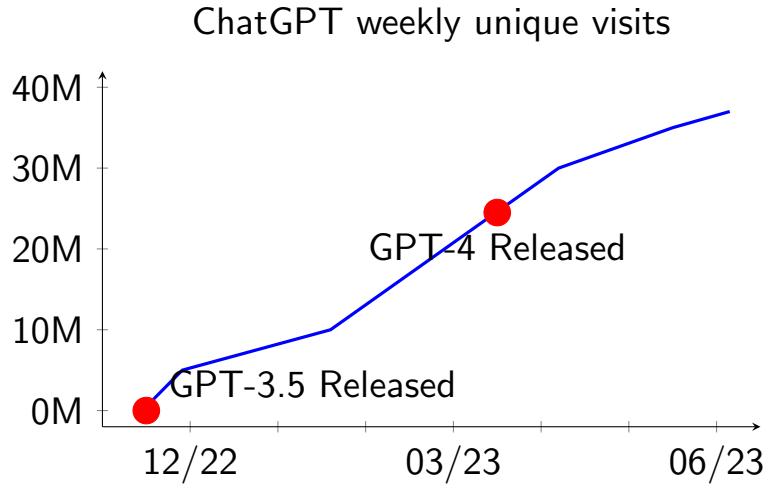
Where are we in AI today?

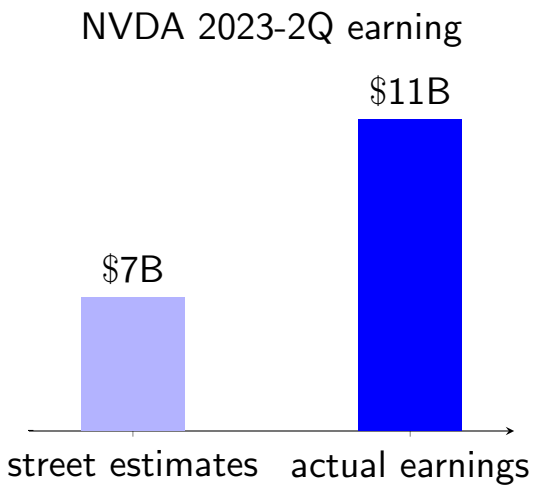
- sunrise phase - currently experiencing dawn of AI era with significant advancements and increasing adoption across various industries
- early adoption - in early stages of AI lifecycle with widespread adoption and innovation across sectors marking significant shift in technology's role in society



Explosion of AI ecosystems - ChatGPT & NVIDIA

- took only *5 months for ChatGPT users to reach 35M*
- NVDIA 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[3]



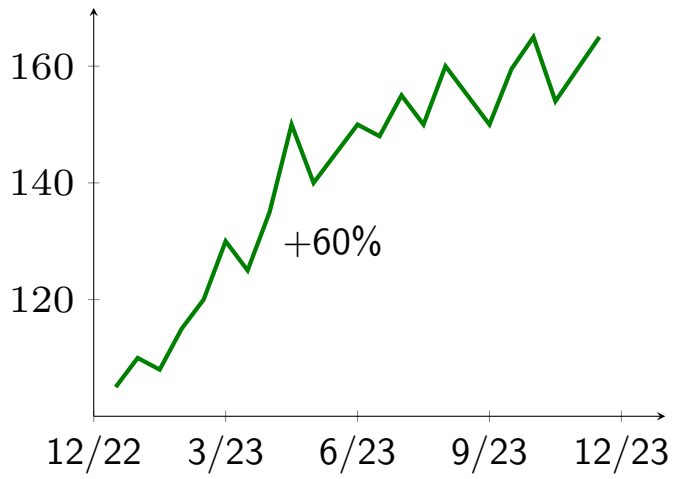


[3] source - Bloomberg

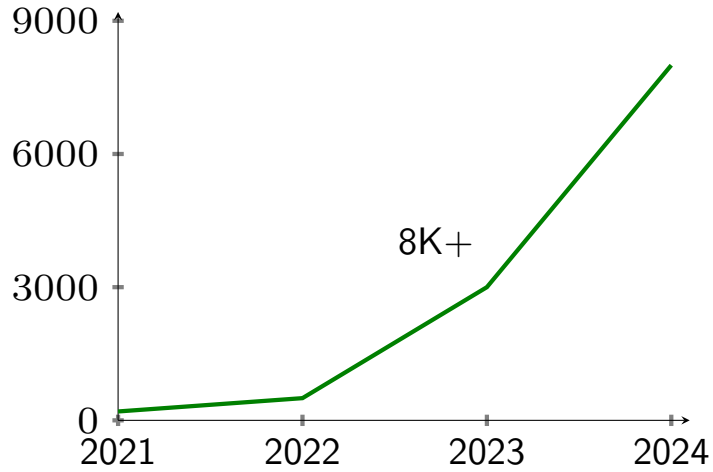
Explosion of AI ecosystems - AI stock market

- *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 basket performance - 2023

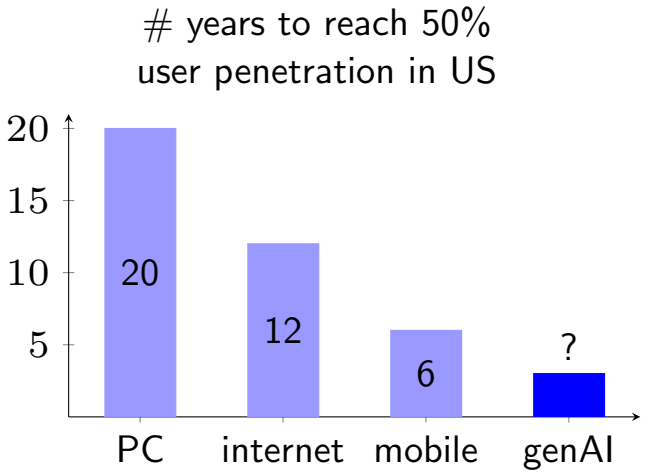


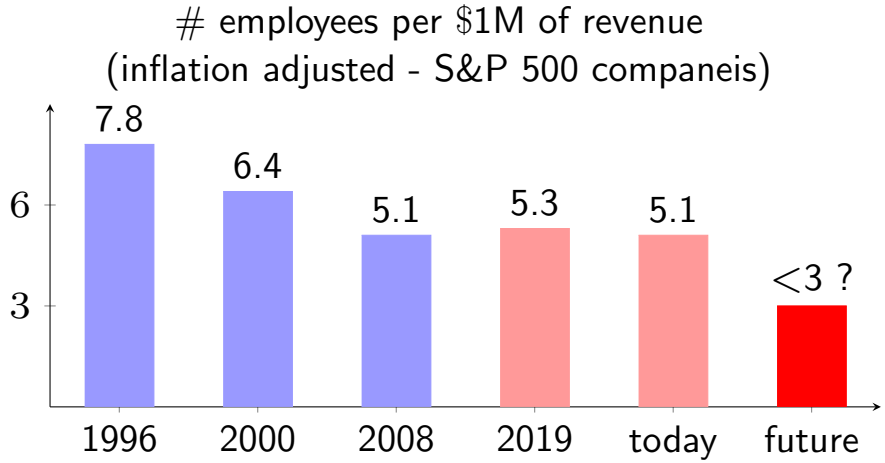
new AI applications over time



AI's transformative impact - adoption speed & economic potential

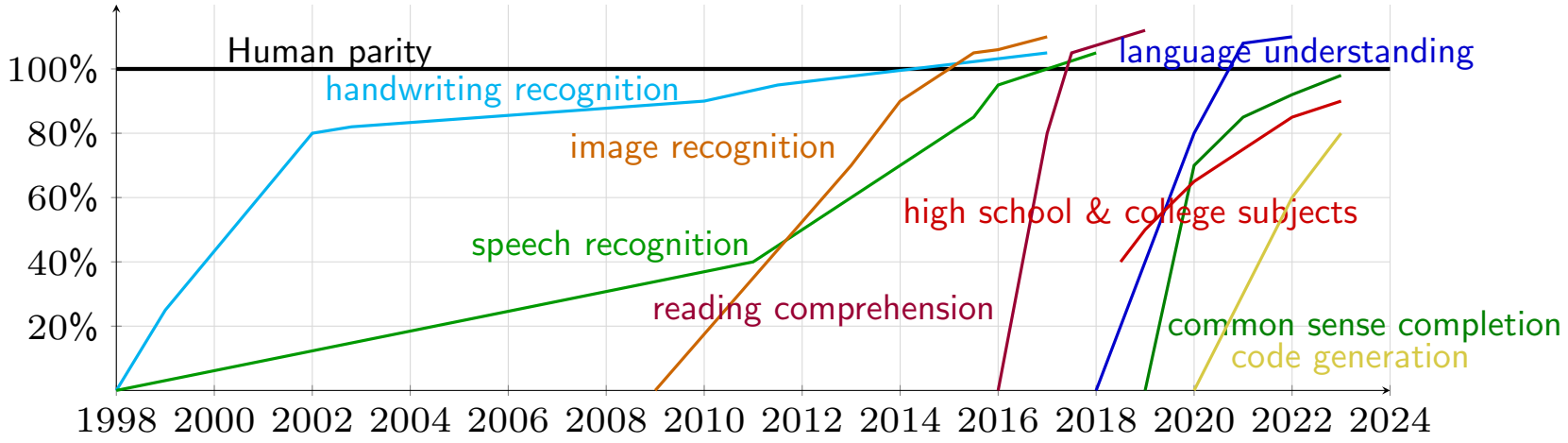
- adoption - has been twice as fast with platform shifts suggesting
  - increasing demand and readiness for new technology improved user experience & accessibility
- AI's potential to drive economy for years to come
  - 35% improvement in productivity driven by introduction of PCs and internet
  - greater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





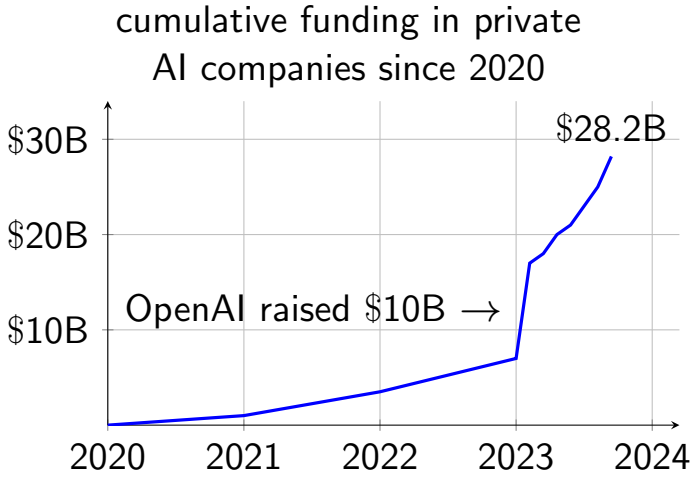
AI getting more & more faster

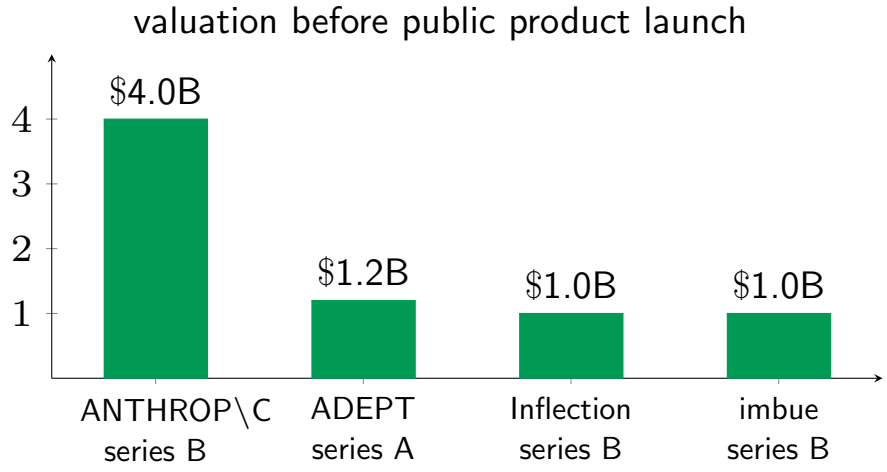
- steep upward slopes of AI capabilities highlight accelerating pace of AI development
  - period of exponential growth with AI potentially mastering new skills and surpassing human capabilities at ever-increasing rate
- closing gap to human parity - some capabilities approaching or arguably reached human parity, while others having still way to go
  - achieving truly human-like capabilities in broad range remains a challenge



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 stella 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

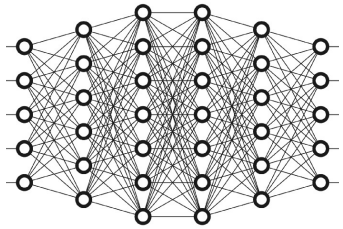


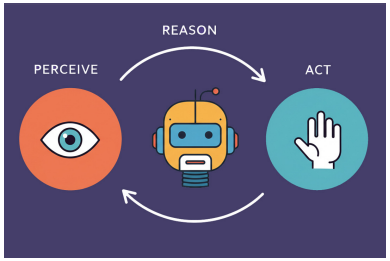


AI Agents

AI progress in 21st century in keywords

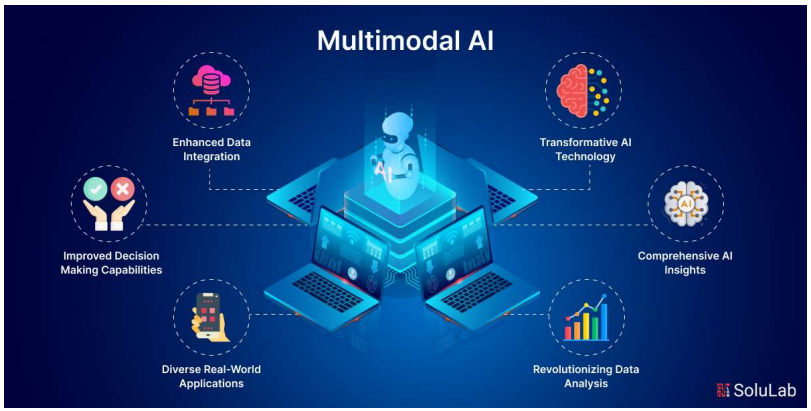
- 2010 $\sim$ Big Data
- 2012 $\sim$ Deep Learning
- 2017 $\sim$ Transformer - Attention is All you need!
- 2022 $\sim$ LLM & genAI
- 2024 $\sim$ AI Agent (Agentic AI)

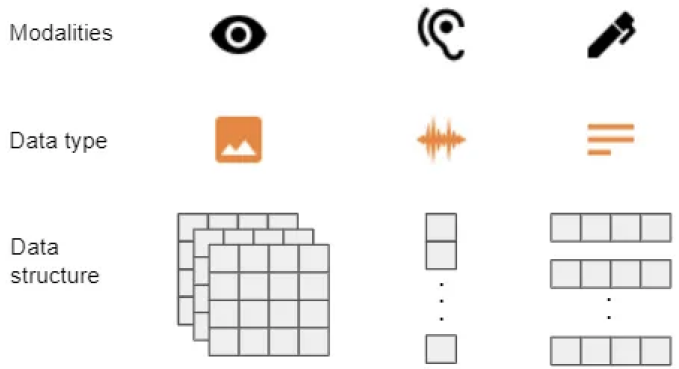




Multimodal learning

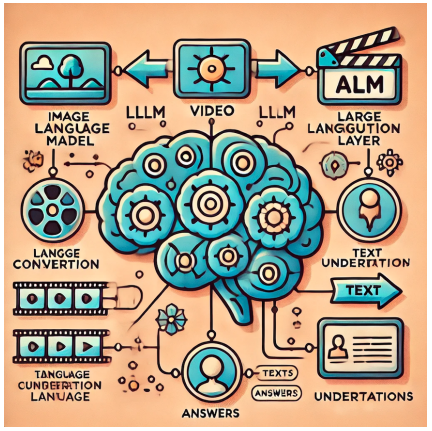
- understand information from multiple modalities, $e.g.$, text, images, audio, video
- representation learning methods
  - combine multiple representations or learn multimodal representations simultaneously
- applications
  - images from text prompt, videos with narration, musics with lyrics
- collaboration among different modalities
  - understand image world (open system) using language (closed system)

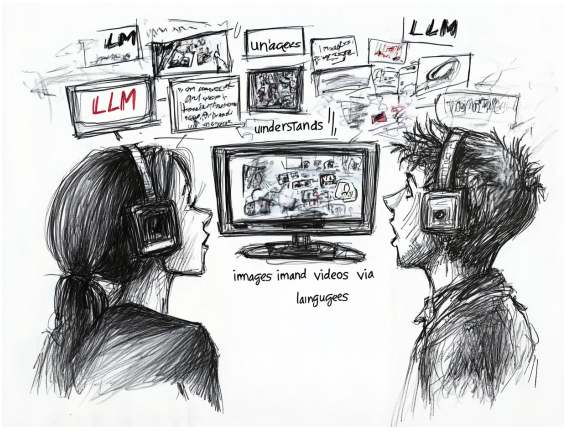




Implications of success of LLMs

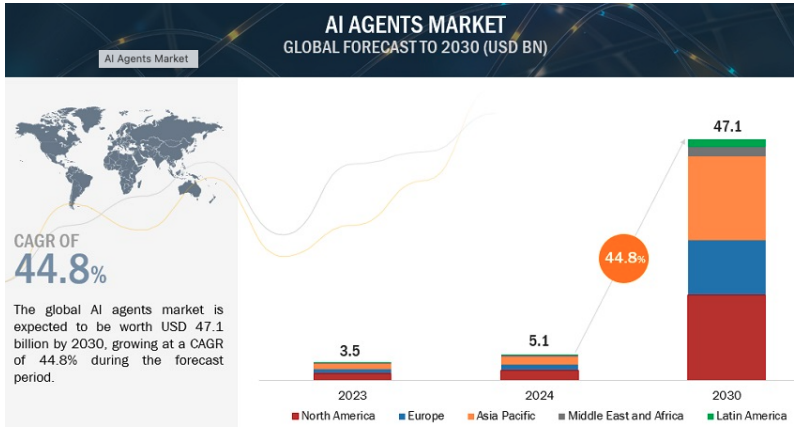
- many researchers change gears towards LLM
  - from computer vision (CV), speach, music, video, even reinforcement learning
- *LLM is not only about NLP . . .* humans have . . .
  - evolved to optimize natural language structures for eons
  - handed down knowledge using *this natural languages* for thousands of years
  - internal structure (or equivalently, representation) of natural languages optimized via *thousands of generation by evolution*
- LLM *connects non-linguistic world (open system) via natural languages (closed system)*

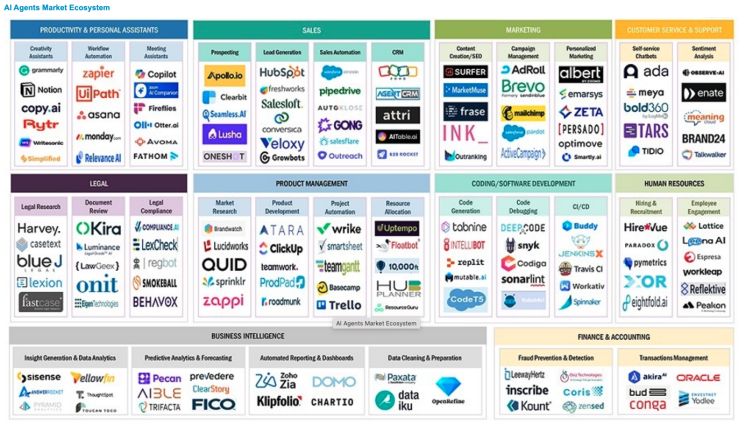




Multimodal AI (mmAI)

- mmAI - systems processing & integrating data from multiple sources & modalities, to generate unified response / decision
- 1990s – 2000s - early systems - initial research combining basic text & image data
- 2010s - CNNs & RNNs enabling more sophisticated handling of multimodality
- 2020s - modern multimodal models - Transformer-based architectures handling complex multi-source data at highly advanced level
- mmAI *mimics human cognitive ability* to interpret and integrate information from various sources, leading to holistic decision-making



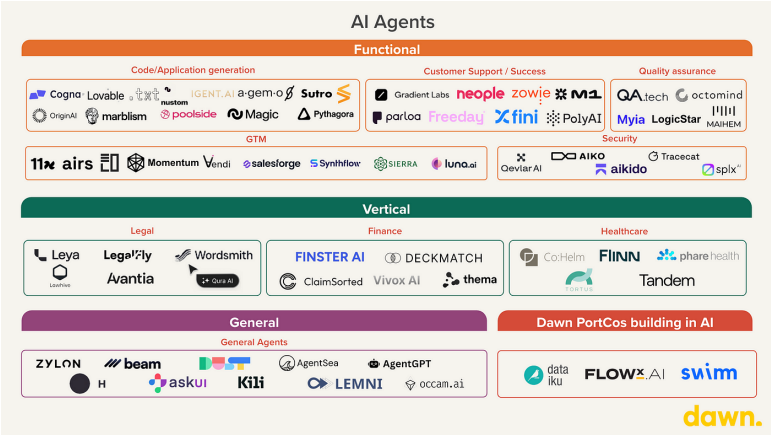


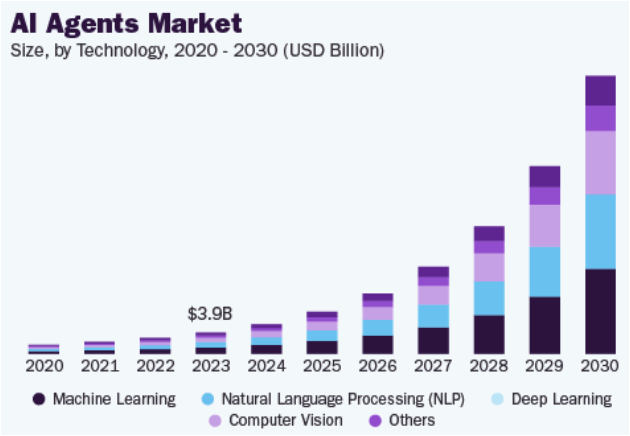
mmAI Technology

- core components
  - data preprocessing - images, text, audio & video
  - architectures - unified Transformer-based ($e.g.$, ViT) & cross-attention mechanisms / hybrid architectures ($e.g.$, CNNs + LLMs)
  - integration layers - fusion methods for combining data representations from different modalities
- technical challenges
  - data alignment - accurate alignment of multimodal data
  - computational demand - high-resource requirements for training and inferencing
  - diverse data quality - manage variations in data quality across modalities
- advancements
  - multimodal embeddings - shared feature spaces interaction between modalities
  - self-supervised learning - leverage unlabeled data to learn representations across modalities

AI agents powered by multimodal LLMs

- foundation
  - integrate multimodal AI capabilities for enhanced interaction & decision-making
- components
  - perceive environment through multiple modalities (visual, audio, text), process using LLM technology, generate contextual responses & take actions
- capabilities
  - understand complex environments, reason across modalities, engage in natural interactions, adapt behavior based on context & feedback





AI agents - Present & Future

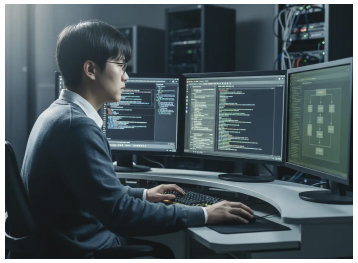
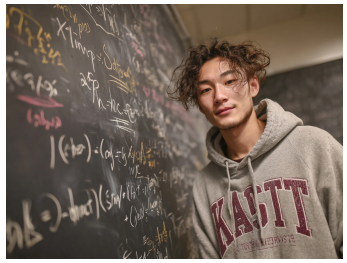
- emerging applications
  - scientific research - agents analyzing & running experiments & generating hypotheses
  - creative collaboration - AI partners in design & art combining multiple mediums
  - environmental monitoring - processing satellite sensor data for climate analysis
  - healthcare - enhanced diagnostic combining imaging, $e.g.$, MRI, with patient history
  - customer experience - virtual assistants understanding spoken language & visual cues
  - autonomous vehicles - integration of visual, radar & audio data
- future
  - ubiquitous AI agents - seamless integration into everyday devices
  - highly tailored personalized experience - in education, entertainment & healthcare

Building Students' Superpower

Students already building with AI - two paths, one future

- AI power user - domain experts using AI (majority)
  - 17-year-old high school student - passionate about helping elderly grandparents
    - built medication reminder app using Claude - no programming background
    - ". . . didn't learn to code. I learned to describe what grandparents need to AI."
  - college business student - interested in K-beauty industry
    - used Claude to analyze social media sentiment to predict K-beauty trend
    - "I understood Korean beauty culture. AI understood data patterns."
- AI expert - AI scientists & engineers & builders
  - computer science junior - 4 years studying math, algorithms, neural networks
    - fine-tuned open-source LLM for Korean medical terminology
    - collaborated with doctors (domain experts) on diagnostic tool



AI power user - domain expert AI-amplified

- who this is for?
  - you love literature, business, medicine, law, art, design, *etc.*
  - you find AI interesting as TOOL, not as end in itself
  - you get excited about domain problems, not algorithms
- what you'll do
  - deepen expertise in your chosen field (4+ years) - learn AI tools as power tools
  - use AI to amplify your domain work & compete on domain insight + AI leverage
- career examples
  - doctor using AI diagnostics, teacher using AI personalization
  - lawyer using AI research, artist using AI iteration
  - marketer using AI analytics, scientist using AI simulation




AI expert - researcher/scientist/engineer/developer

- who this is for?
  - you find algorithms, mathematics, systems *beautiful*, and read AI papers for fun
  - you want to work at AI labs being excited when new architectures are published
- what you'll do
  - deep study - mathematics, computer science, ML theory ($e.g.$, 4+ years)
  - understand neural networks, transformers to build and improve AI systems
  - *collaborate* w/ domain experts to apply your systems
- career examples
  - ML Engineer working for (tech) companies
  - AI Researcher in academia or industry labs
  - research scientist, robotics engineer, computer vision specialist

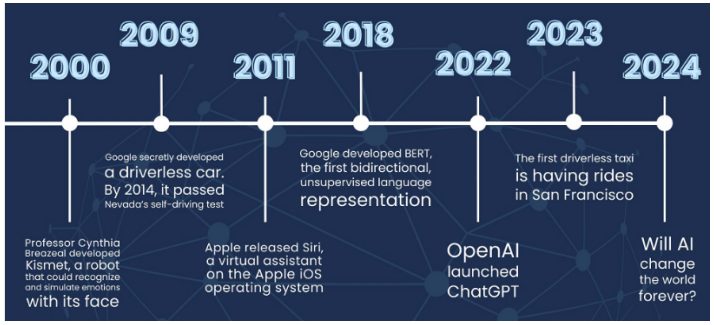


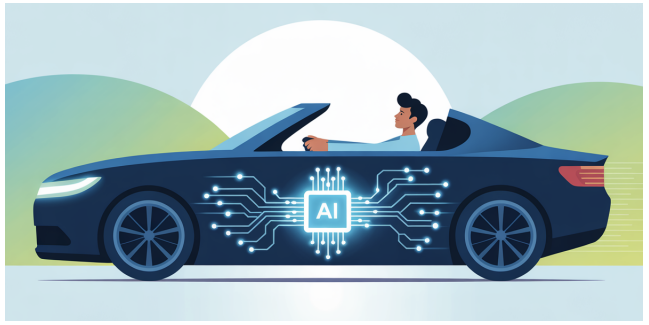

BIG shifts in AI landscape

- old assumption (2015 – 2020)
  - AI will be built by tech companies, everyone else will be disrupted
  - lots of tech companies will compete for best models/products
- new reality (2024 –)
  - AI is commoditized tool, domain expertise is where value accrues
  - only handful of companies can develop cutting-edge foundation models

Domain expert revolution

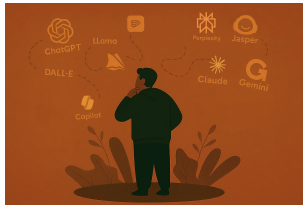
- why domain experts are winning
  - problem identification requires domain knowledge
  - evaluation requires domain judgment
  - trust requires human domain expertise
  - context requires cultural/domain fluency
- pattern across industries
  - best medical AI applications $\rightarrow$ built by doctors collaborating with AI engineers
  - best legal AI tools $\rightarrow$ built by lawyers collaborating with AI engineers
  - best educational AI $\rightarrow$ built by teachers collaborating with AI engineers
  - domain expert leads & AI engineer supports (not the reverse)

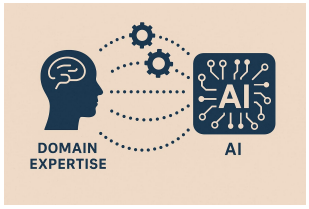


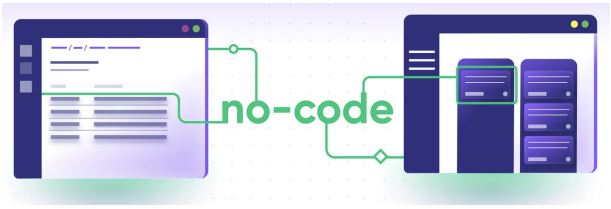


AI power user - what you actually need to learn

- tool awareness
  - what - know which AI tools exist and their capabilities
  - examples - ChatGPT for brainstorming, Claude for research, NotebookLM for synthesis
  - how - YouTube tutorials, free trials, experimentation
- right questions to throw - interactive way
  - what - communicate effectively with AI using domain knowledge
  - why it works - your domain knowledge makes prompts effective
  - how - practice + domain expertise
- tool integration
  - what - connect AI tools to your workflow, build custom GPTs, use APIs (no coding)
  - how - no-code tools + tutorials

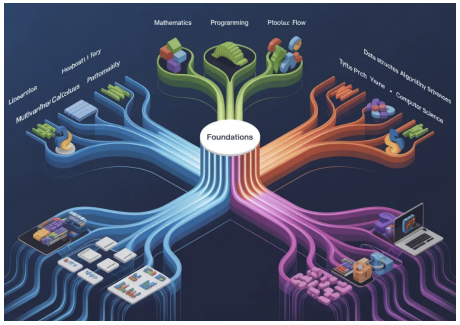


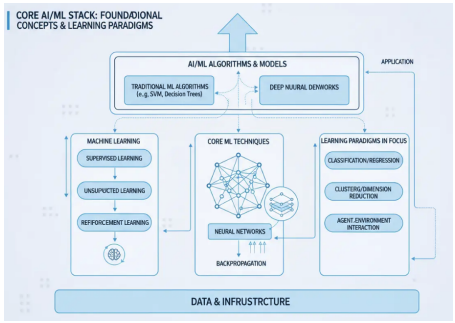




AI expert path - how to become AI scientist/researcher/engineer/practitioner

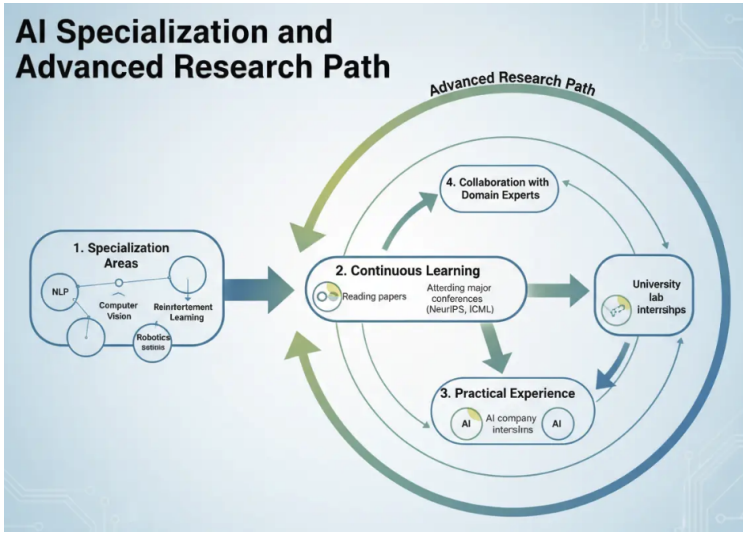
- foundation
  - mathematics - linear algebra, multivariable calculus, probability theory, optimization
  - programming - python w/ PyTorch, TensorFlow, . . .
  - computer science - data structures, algorithms, systems
- core AI/ML
  - machine learning (ML) - supervised/unsupervised learning, reinforcement learning, neural networks, backpropagation
  - deep learning (DL) - CNNs, RNNs, Transformers, attention mechanisms
  - practical projects - competitions, replicate papers, contribute to open source





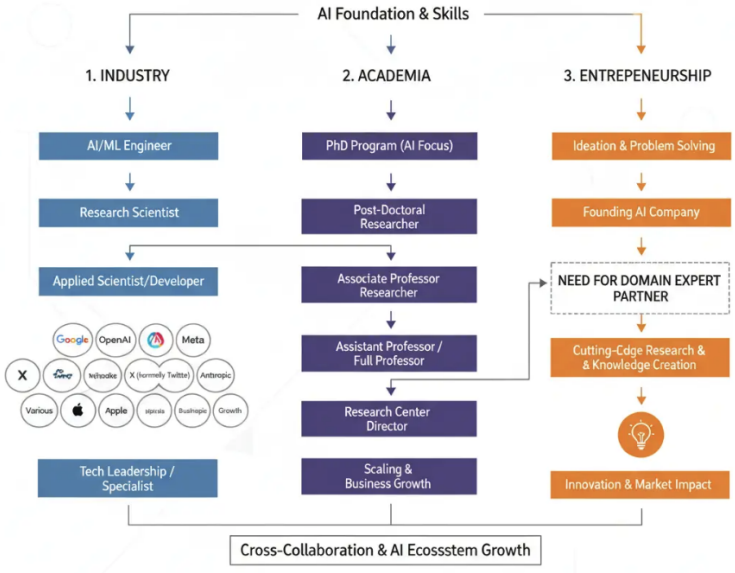
AI expert path - how to become AI scientist/researcher/engineer/practitioner

- specialization
  - choose area (or not!) - NLP, computer vision (CV), robotics, *etc.*
  - read papers - arxiv, attend conferences (*e.g.*, NeurIPS, ICML, ICLR, CVPR, *etc.*)
  - research experience - lab work at university, internships at AI companies
  - collaboration - work with domain experts on real problems



AI expert path - how to become AI scientist/researcher/engineer/practitioner

- career paths
  - industry - Google, OpenAI, Meta, Apple, X, Anthropic, and numerous startups
  - academia - PhD $\rightarrow$ professor / research center
  - entrepreneurship - found AI company (but need domain expert partner!)



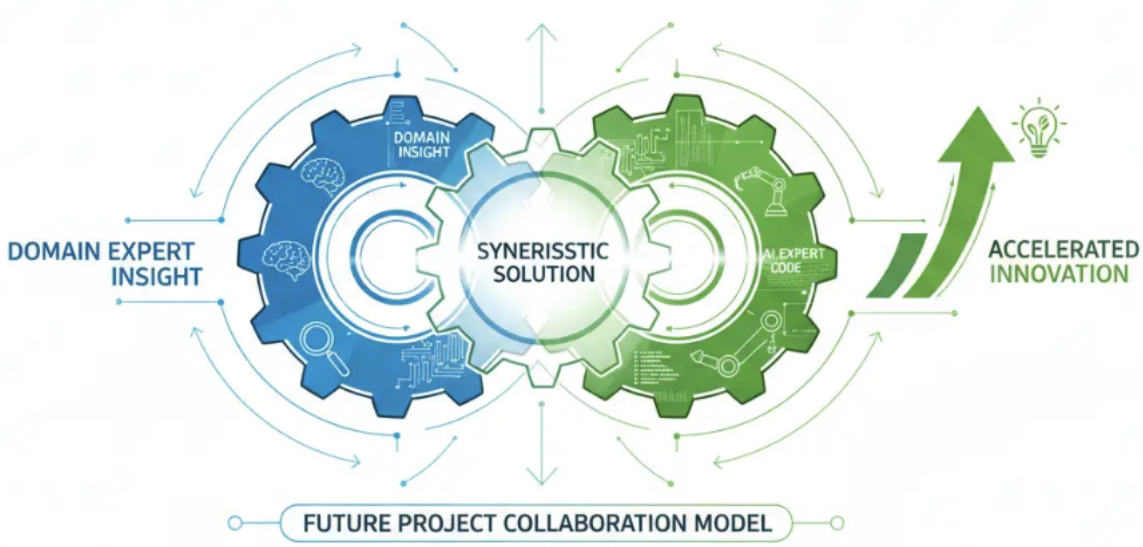
Powerful combination - domain expert + AI expert collaboration

The magic happens at the intersection!

- case study 1 - AlphaFold (Protein Folding)
  - AI engineers - build transformer-based neural network, optimized training
  - structural biologists - identify protein folding as THE problem, validated outputs
  - result - 50-year problem solved - neither could do it alone
  - both Demis Hassabis (CEO @ Google DeepMind) & John Jumper (biochemistry background) got Nobel Prizes in chemistry in 2024!
- case study 2 - cancer diagnostic AI
  - AI engineer role - build model, handle large medical imaging datasets
  - oncologist role - label training data correctly, evaluate clinical relevance
  - result - 30% improvement in early detection
  - who makes final diagnosis - always the doctor!

Powerful combination - domain expert + AI expert collaboration

- case study 3 - your future project
  - domain expert & AI super user- identify problem in your field you understand deeply
  - AI expert - build custom solution beyond available tools
  - result - breakthrough that neither could achieve alone
  - who drives vision - TOGETHER!



Things AI can't do for you - choose who you'll become

- your technical superpower (Domain $\times$ AI) is only as good as *your moral compass*
- why this matters more than ever
  - AI amplifies whatever you choose to do - good intentions OR bad ones
  - history shows us - brilliant minds + no moral foundation = catastrophe
  - your skills will make you powerful; your values determine what you do with that power
- virtues that actually matter in the long run
  - *integrity* - doing the right thing (even) when no one's watching
  - *empathy* - genuinely caring about people; your technology will affect them
  - *service* - building things that help others, not just things that benefit you
  - *morality* - having inner compass that guides you beyond what's legal or profitable

Two pathways to scientific excellence - what to look for

- two equally valuable pathways for students
  - AI researchers and engineers; building the tools
  - domain scientists who leverage AI - biologists, chemists, physicists using AI as powerful tool
  - both require deep problem-solving ability — but manifested differently
- beyond traditional metrics of knowledge mastery
  - ask compelling questions (not just answer them)
  - demonstrate genuine domain curiosity
  - show adaptability in learning new tools
  - exhibit collaborative mindsets
  - demonstrate genuine domain curiosity
  - show adaptability in learning new tools
  - exhibit collaborative mindsets

Selected References & Sources

Selected references & sources

- Robert H. Kane "Quest for Meaning: Values, Ethics, and the Modern Experience" 2013
- Michael J. Sandel "Justice: What's the Right Thing to Do?" 2009
- Daniel Kahneman "Thinking, Fast and Slow" 2011
- Yuval Noah Harari "Sapiens: A Brief History of Humankind" 2014
- M. Shanahan "Talking About Large Language Models" 2022
- A.Y. Halevry, P. Norvig, and F. Pereira "Unreasonable Effectiveness of Data" 2009
- A. Vaswani, et al. "Attention is all you need" @ NeurIPS 2017
- S. Yin, et. al. "A Survey on Multimodal LLMs" 2023
- Chris Miller "Chip War: The Fight for the World's Most Critical Technology" 2022

- CEOs, CTOs, CFOs, COOs, CMOs & CCOs @ startup companies in Silicon Valley
- VCs on Sand Hill Road - Palo Alto, Menlo Park, Woodside in California, USA

References

References

[GPAM+14] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[KW19] Diederik P. Kingma and Max Welling. An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*, 12(4):307–392, 2019.

[MLZ22] Louis-Philippe Morency, Paul Pu Liang, and Amir Zadeh. Tutorial on multimodal machine learning. In Miguel Ballesteros, Yulia Tsvetkov, and Cecilia O. Alm, editors, *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Tutorial Abstracts*, pages 33–38, Seattle, United States, July 2022. Association for Computational Linguistics.

[VSP+17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you

need. In *Proceedings of 31st Conference on Neural Information Processing Systems (NIPS)*, 2017.

[YFZ+24] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. A survey on multimodal large language models, 2024.

Thank You


Sunghee Yun
Co-Founder & CTO @ Erudio Bio / K-PAI
Leader / Global Managing Partner @ LULU...