

[SNU EE AI Seminar] The AI Landscape - Navigating Technology, Industry Shifts, and Future Trends

Sunghee Yun

Co-founder & CTO - AI Technology & Biz Dev @ [Erudio Bio, Inc.](#)

About Speaker

- *Co-founder & CTO @ Erudio Bio, San Jose & Novato, CA, USA*
- Advisor & Evangelist @ CryptoLab, Inc., San Jose, CA, USA
- Chief Business Development Officer @ WeStory.ai, Cupertino, CA, USA
- Advisory Professor, Electrical Engineering and Computer Science @ DGIST, Korea
- Adjunct Professor, Electronic Engineering Department @ Sogang University, Korea
- Global Advisory Board Member @ Innovative Future Brain-Inspired Intelligence System Semiconductor of Sogang University, Korea
- *KFAS-Salzburg Global Leadership Initiative Fellow @ Salzburg Global Seminar*, Salzburg, Austria
- Technology Consultant @ Gerson Lehrman Group (GLG), NY, USA
- *Co-founder & CTO & Head of Global R&D & Chief Applied Scientist & Senior Fellow @ Gauss Labs, Inc., Palo Alto, CA, USA* 2020 – 2023

- Senior Applied Scientist @ Mobile Shopping Team, Amazon.com, Inc., Vancouver, BC, Canada – 2020
- Principal Engineer @ Software R&D Center of DS Division, Samsung, Korea – 2017
- Principal Engineer @ Strategic Marketing & Sales Team, Samsung, Korea – 2016
- Principal Engineer @ DT Team of DRAM Development Lab, Samsung, Korea – 2015
- Senior Engineer @ CAE Team - Samsung, Korea – 2012
- MS & PhD - Electrical Engineering @ Stanford University, CA, USA – 2004
- Development Engineer @ Vyan, Santa Clara, CA, USA – 2001
- BS - Electrical Engineering @ Seoul National University, Seoul, Korea – 1998

Highlight of Career Journey

- BS in EE @ SNU, MS & PhD in EE @ Stanford University
 - *Convex Optimization - Theory, Algorithms & Software*
 - advised by *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 - time-series anomaly detection, deep reinforcement learning & recommender system
 - Jeff Bezos's project - increase sales by *\$200M* via Amazon Mobile Shopping App
- Co-founder & CTO & Head of Global R&D & Chief Applied Scientist & Senior Fellow @ Gauss Labs, Inc.
- Co-founder & CTO - AI Technology & Business Development @ Erudio Bio, Inc.

Today

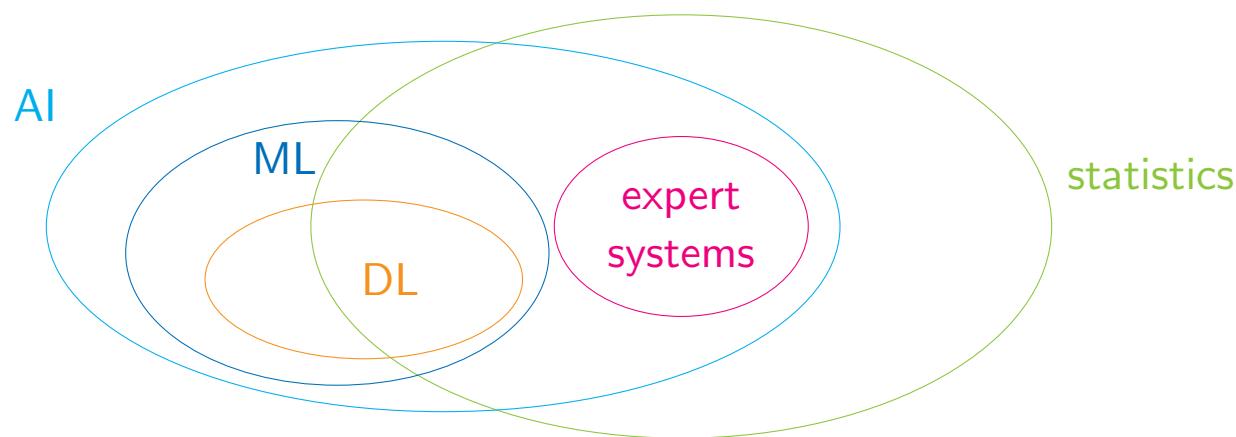
- Artificial Intelligence - 5
 - AI history & recent significant achievements
 - Industry & market indices, is AI hype?
- AI Agents - 25
 - Big Data → DL → LLM & genAI → Agentic AI
 - Future of society powered by AI agents
- Silicon Valley's Cultural Engine of Innovation and Disruption - 32
 - My journey from Samsung, Amazon & Gauss Labs to Erudio Bio
 - Innovation ecosystem & case study - Amazon
- Selected references - 40

Artificial Intelligence

Definition and History

Definition & relation with other technologies

- AI
 - is technology doing tasks requiring human intelligence, such as learning, problem-solving, decision-making & language understanding
 - encompasses *range of technologies, methodologies, applications & products*
- AI, ML, DL, statistics & expert system¹ [?]



¹ML: machine learning & DL: deep learning

History



Significant AI Achievements - 2014 – 2024

Deep learning revolution

- 2012 – 2015 - DL revolution²
 - 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



²CV: computer vision, NN: neural network, CNN: convolutional NN, RL: reinforcement learning

Transformer changes everything

- 2017 – 2018 - Transformers & NLP breakthroughs³
 - *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



³NLP: natural language processing, GPT: generative pre-trained transformer

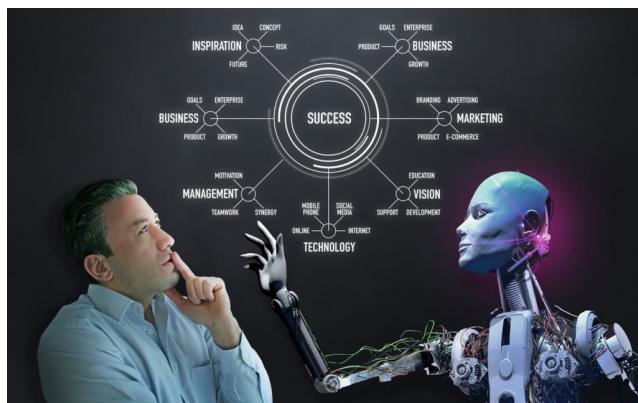
Lots of breakthroughs in AI technology and applications in 2024

- proliferation of advanced AI models
 - GPT-4o, Claude Sonnet, 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

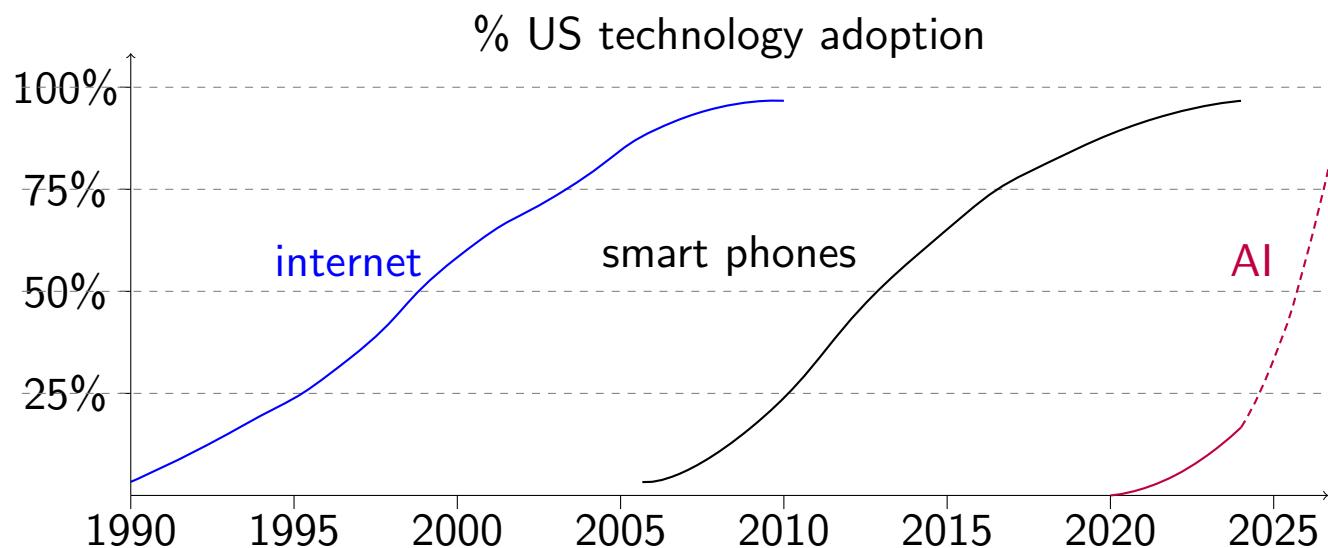
- accelerating human-AI collaboration
 - not only reshaping industries but *altering how humans interact with technology*
 - AI's role as collaborator and augmentor redefines productivity, creativity, the way we address global challenges, e.g., *sustainability & healthcare*
- AI-driven automation *transforms 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

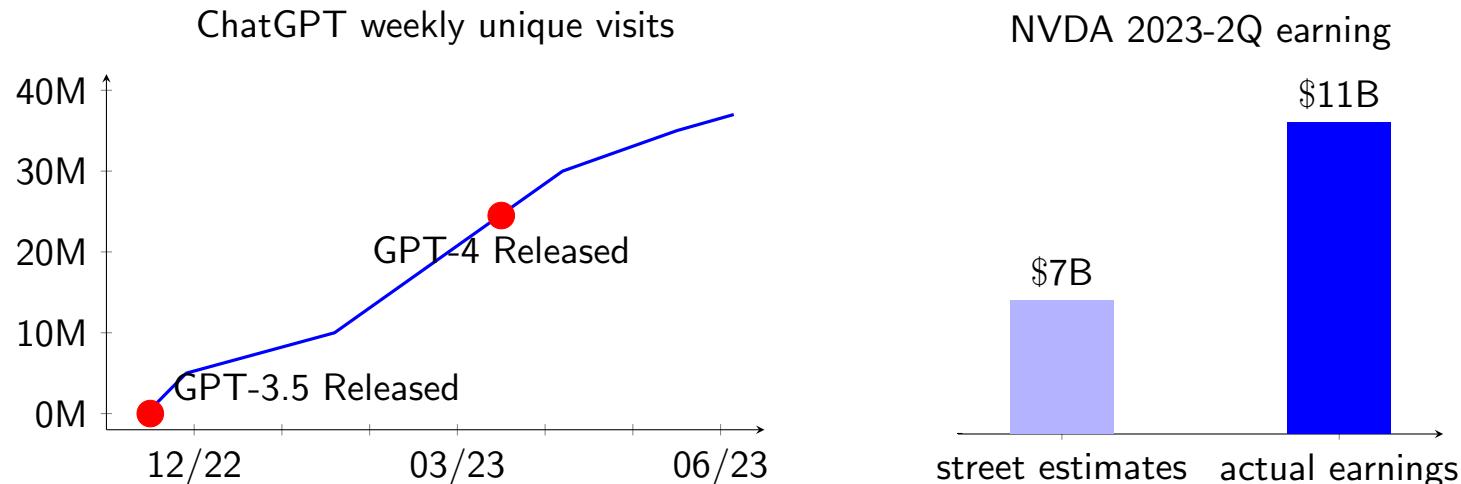
Where are we in AI today?

- 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

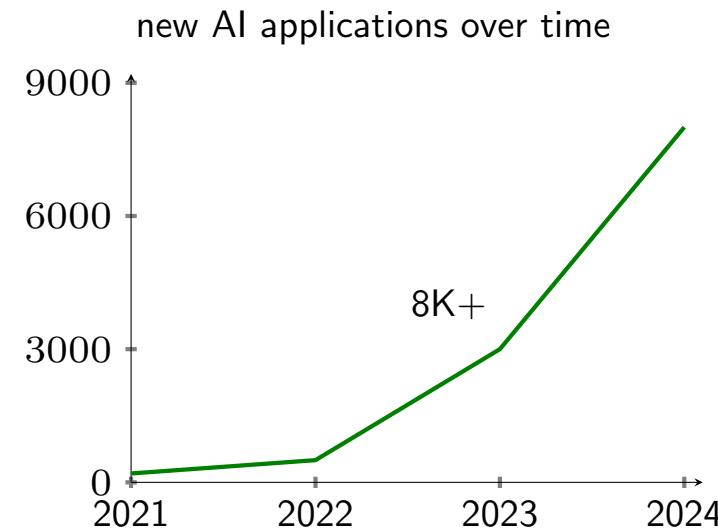
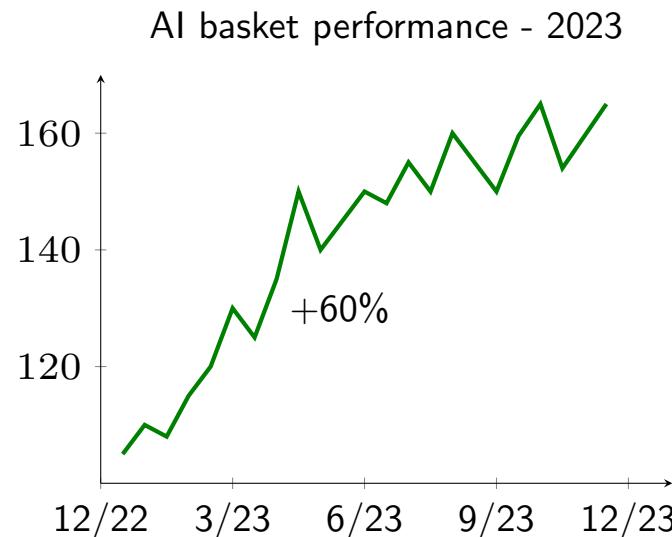
- took only *5 months for ChatGPT users to reach 35M*
- NVIDIA 2023 Q2 earning exceeds market expectation by big margin - \$7B vs \$13.5B
 - surprisingly, *101% year-to-year growth*
 - even more surprisingly *gross margin was 71.2%* - up from 43.5% in previous year⁴



⁴source - Bloomberg

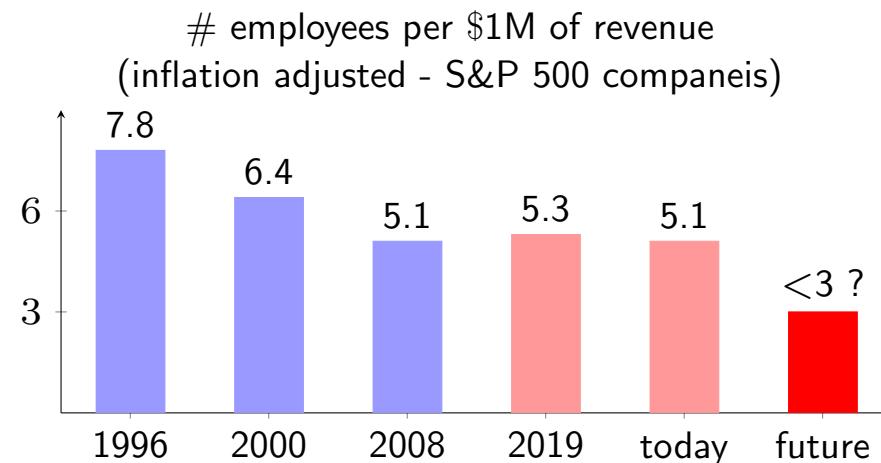
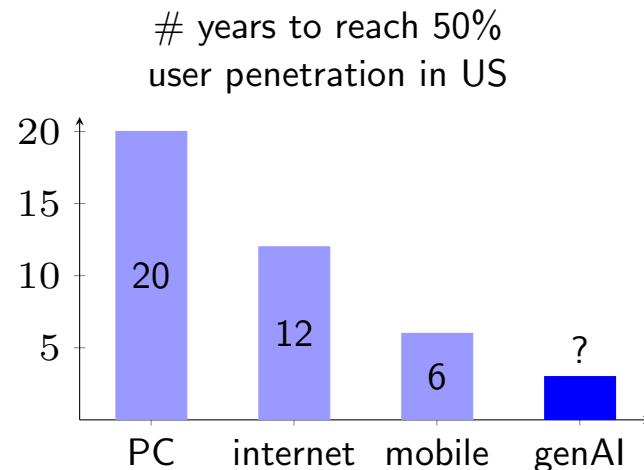
Explosion of AI ecosystems - AI stock market

- *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



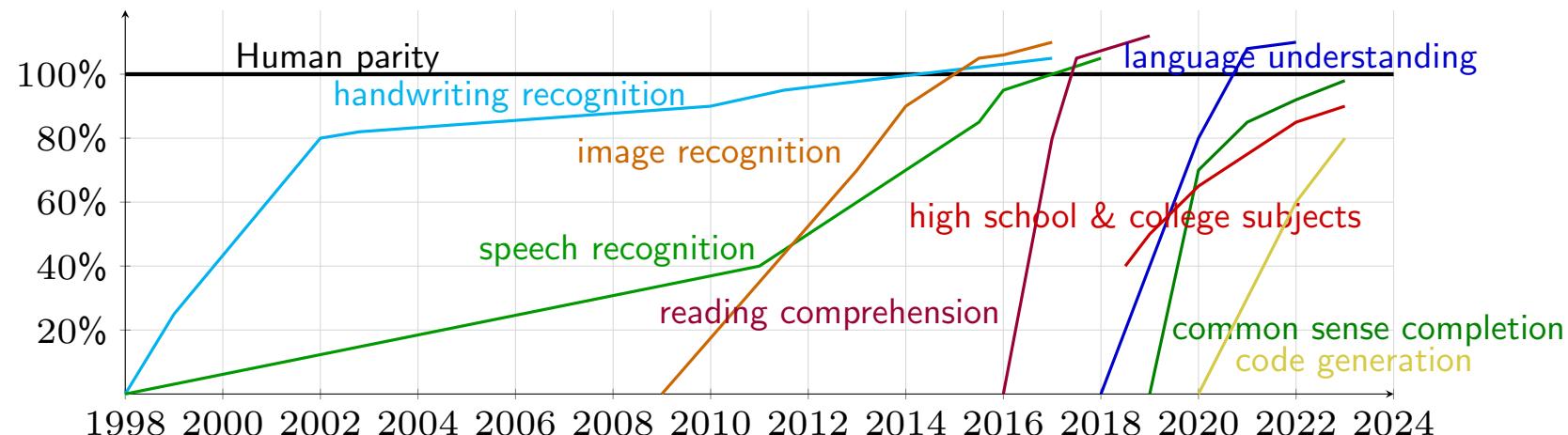
AI's transformative impact - adoption speed & economic potential

- 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



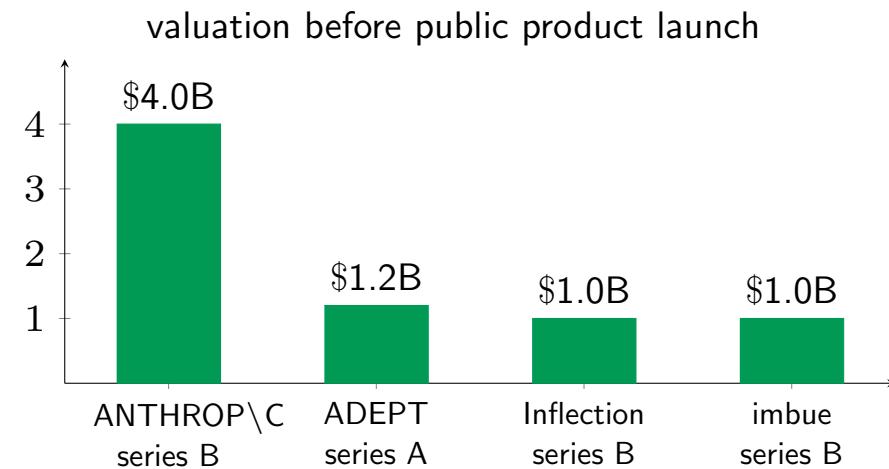
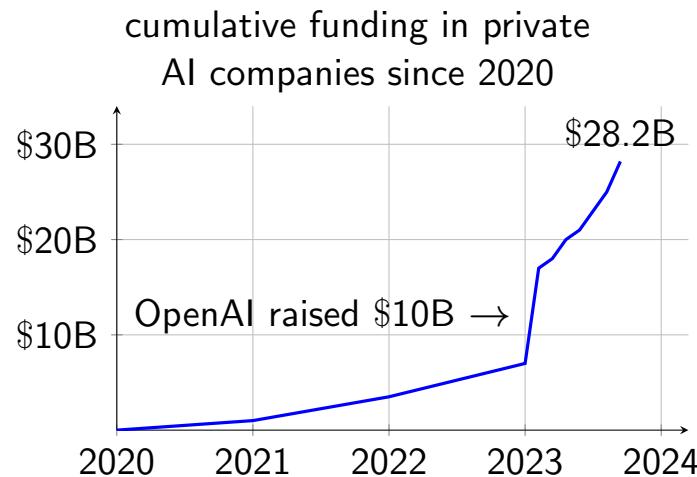
AI getting more & more faster

- 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



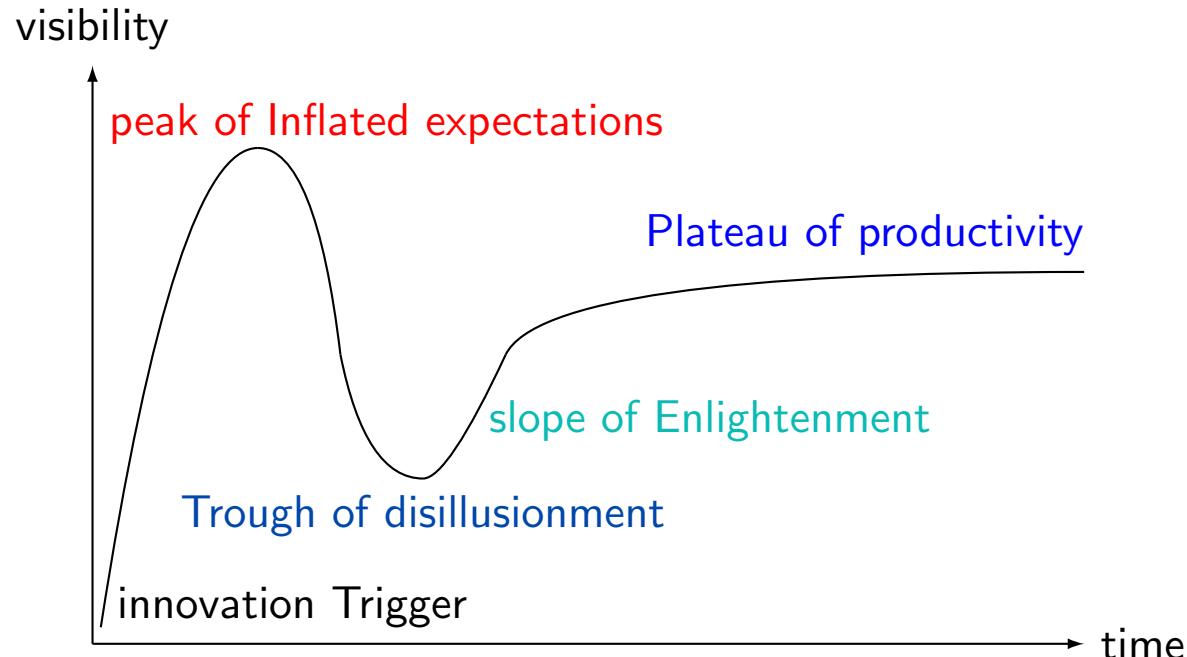
Massive investment in AI

- *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 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*



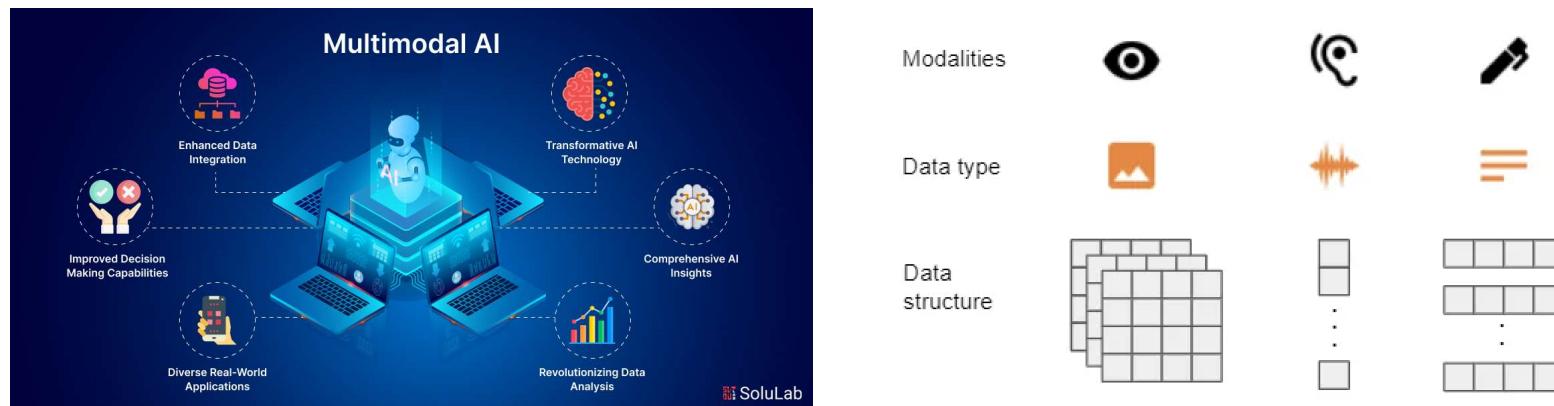
Yes & No

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

AI Agents

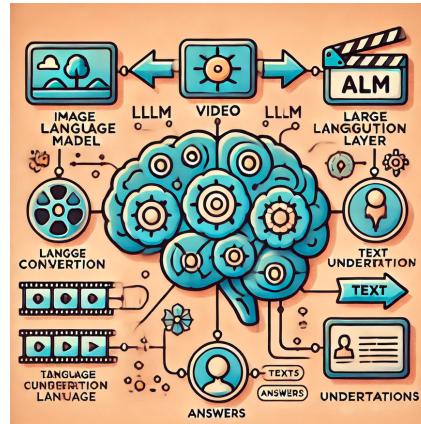
Multimodal learning

- 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)



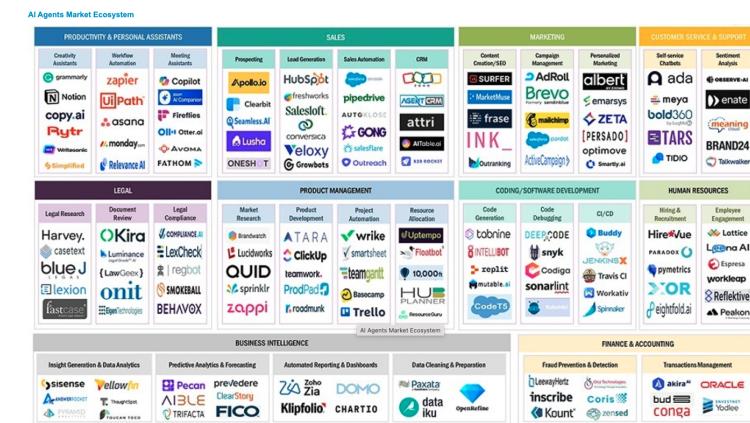
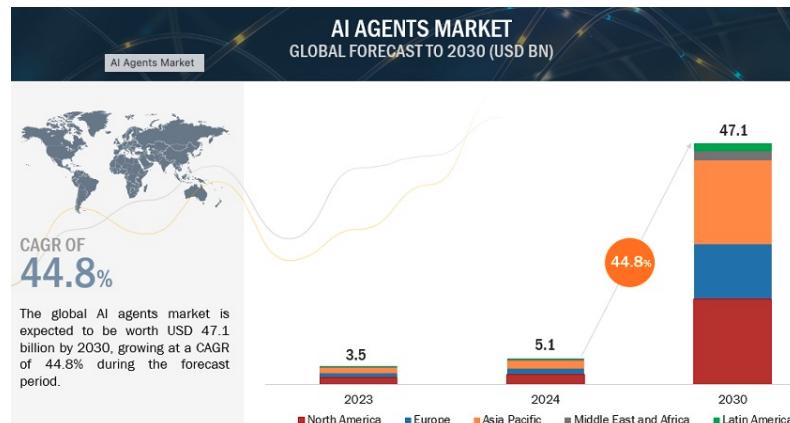
Implications of success of LLMs

- many researchers change gears towards LLM
 - from computer vision (CV), speech, 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) - definition & history

- 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

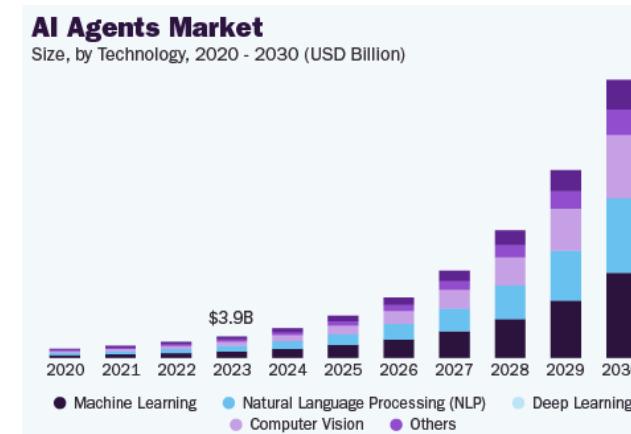
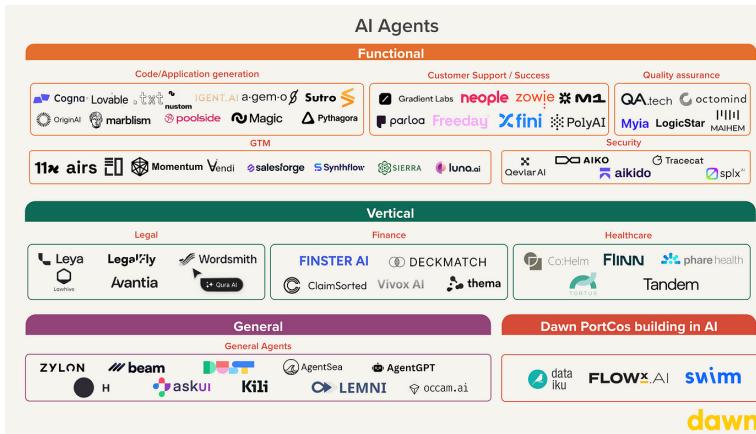


mmAI Technology

- 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



Silicon Valley's Cultural Engine of Innovation and Disruption

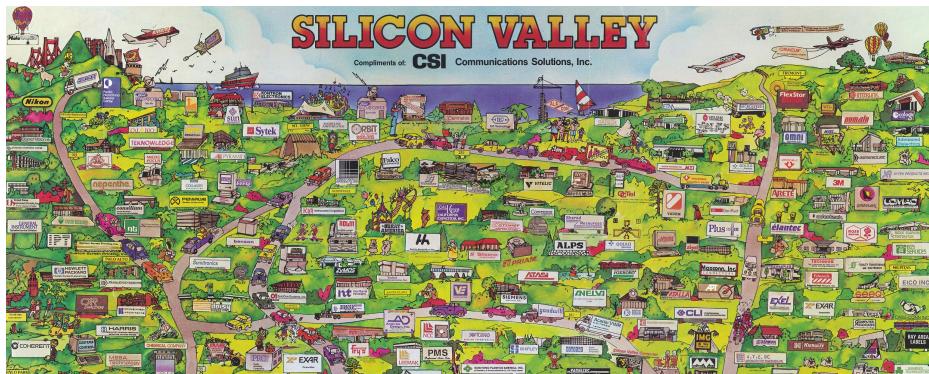
My journey - from Samsung, Amazon & Gauss Labs to Erudio Bio

- Samsung Semiconductor, Inc.
 - inception into industry from academia
 - work with amazing engineers and leaders of the world's best memory chip maker!
- Amazon.com, Inc.
 - experience so-called Silicon Valley big tech culture and technology
 - set tone for my future career trajectory!
- Gauss Labs, Inc.
 - found & operate AI startup, shaping corporate culture & spearheading R&D as CTO
 - inherent challenges of Korean conglomerate spin-off startup - cultural constraints, over-capitalization, and leadership limitations
- Erudio Bio, Inc.
 - concrete & tangible bio-technology in addition to AI
 - good decisions about business models, market fit, go-to-market (GTM) strategies



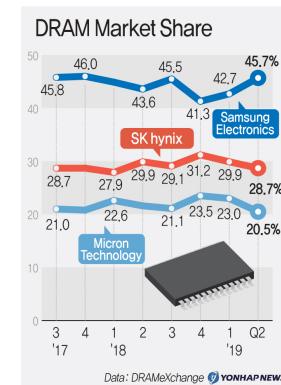
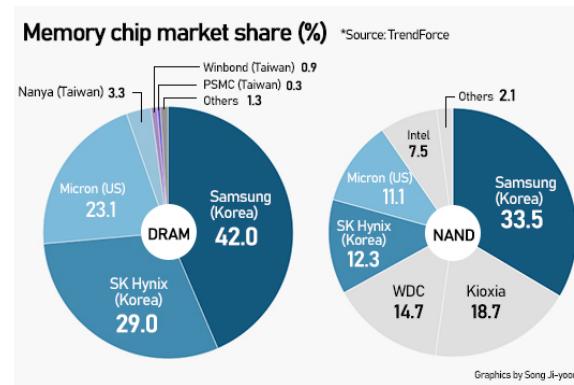
Innovation ecosystem of Silicon Valley

- key characteristics
 - risk-taking culture, trust in technology
 - easy access to huge capital - VCs, angel investors alike
 - talent density - engineers, entrepreneurs, researchers, scientists
 - diversity, “collision density” of ideas
 - ecosystem of collaboration and competition - startups, academia, industry leaders
- what they mean for global big tech
 - set trends in AI, software & hardware innovation
 - act as testing ground for disruptive ideas



Case study: Amazon - amazing differentiators of big techs

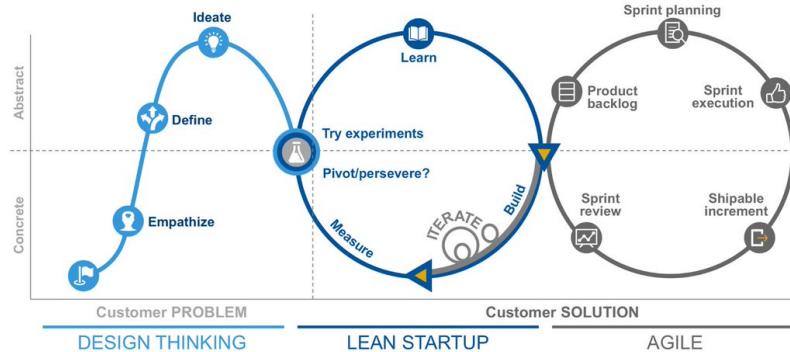
- Amazon's culture & leadership principles
 - customer obsession as driver of innovation
 - high standards & ownership culture
 - bias for action and long-term thinking - sounds contradictory?
 - mechanisms like “two-pizza teams” & “Day One” for scalability
- lessons for Samsung
 - applying customer-centric innovation in hardware & AI, e.g., on-device AI
 - balancing agility with long-term R&D
 - *build / adapt / apply on the core strength of Samsung that no other company has!*



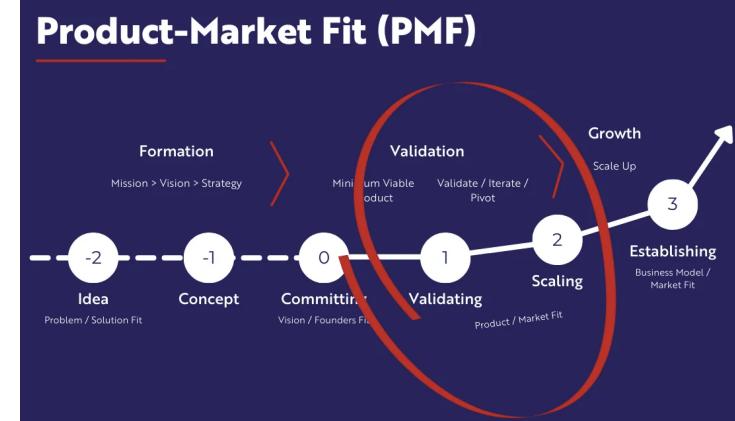
Founding and scaling startups

- challenges
 - competence of and chemistry among co-founders crucial
 - technology & great team are *necessary*, but *not sufficient (at all!)* for success
 - business models, market fit, timing, agility, flexibility for pivoting and perseverance
- insight
 - importance of domain expertise in addition to AI
 - balancing innovation with good business decisions

Combine Design Thinking, Lean Startup and Agile

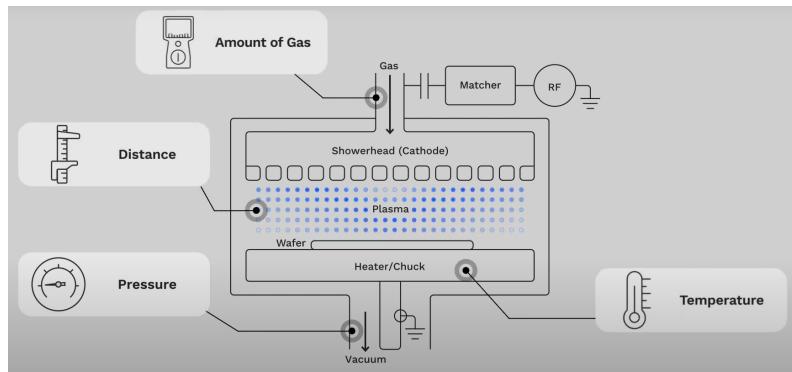


Product-Market Fit (PMF)



AI in shaping future of semiconductor and beyond

- opportunities
 - computer vision applications, *e.g.*, defect inspection, (failure) pattern classification
 - AI-driven optimization in manufacturing, quality control
- challenges
 - data quality, accessibility, *e.g.*, integration with legacy systems
 - demand for extreme accuracy, concept drift & shift
- *bad* examples
 - predictive maintenance - extremely hard (or impossible) problem



Bridging Silicon Valley & Korea

- cultural differences
 - risk appetite & failure tolerance
 - decision-making speed and hierarchy
 - innovation vs execution focus
- opportunities for collaboration
 - leveraging Korea's manufacturing expertise with Silicon Valley's software/AI strengths
 - building global teams with diverse perspectives



To be successful . . .

- embrace customer-centric mindset in innovation and for business decisions
- balance agility with long-term vision
- foster cross-cultural collaboration for global impact
- ((very) strategically and carefully) leverage AI to solve real-world industrial challenges



Selected References & Sources

Selected references & sources

- Chris Miller “Chip War: The Fight for the World’s Most Critical Technology” (2022)
- Daniel Kahneman “Thinking, Fast and Slow” (2011)
- M. Shanahan “Talking About Large Language Models” (2022)
- A.Y. Halevy, 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)
- I.J. Goodfellow, ..., Y. Bengio “Generative adversarial networks (GAN)” (2014)
- T. Kuiken “Artificial Intelligence in the Biological Sciences: Uses, Safety, Security, and Oversight” (2023)
- Stanford Venture Investment Groups
- CEOs & CTOs @ startup companies in Silicon Valley
- VCs on Sand Hill Road - Palo Alto, Menlo Park, Woodside in California, USA

Thank You