

improving 

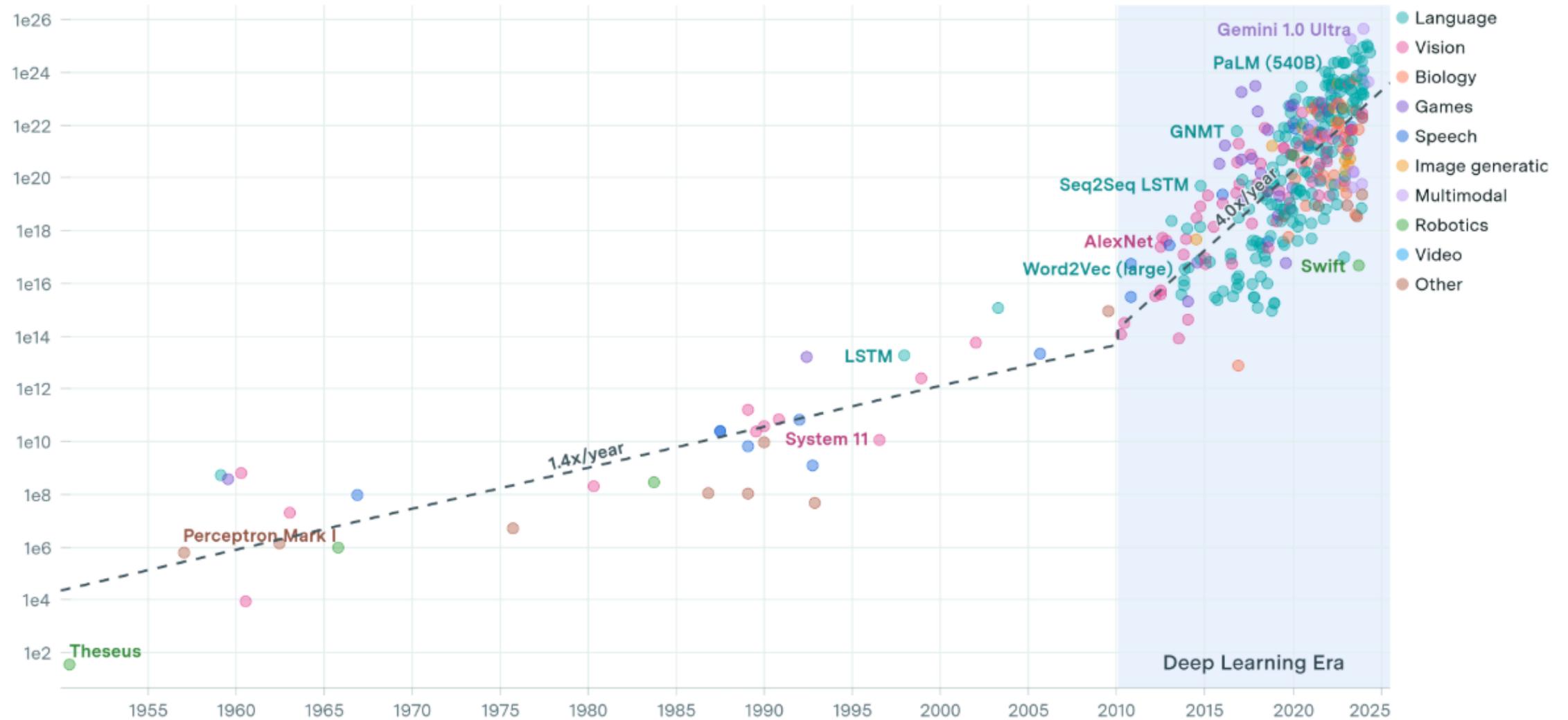
Executive AI Workshop

An Executive's field guide to AI in your business

What do you already
know / believe about AI?

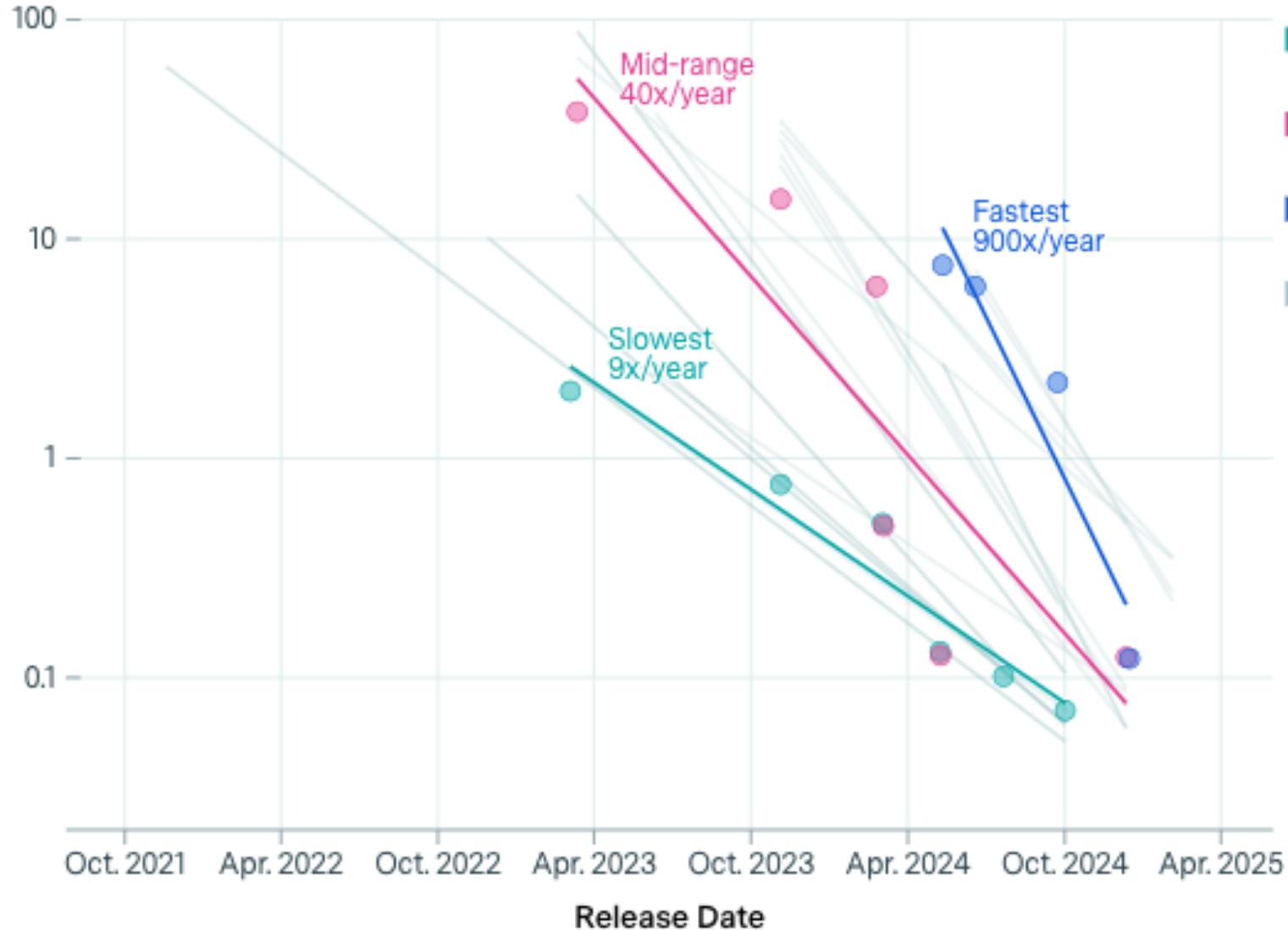
Where are we today?

Training compute (FLOP)



Where are we today?

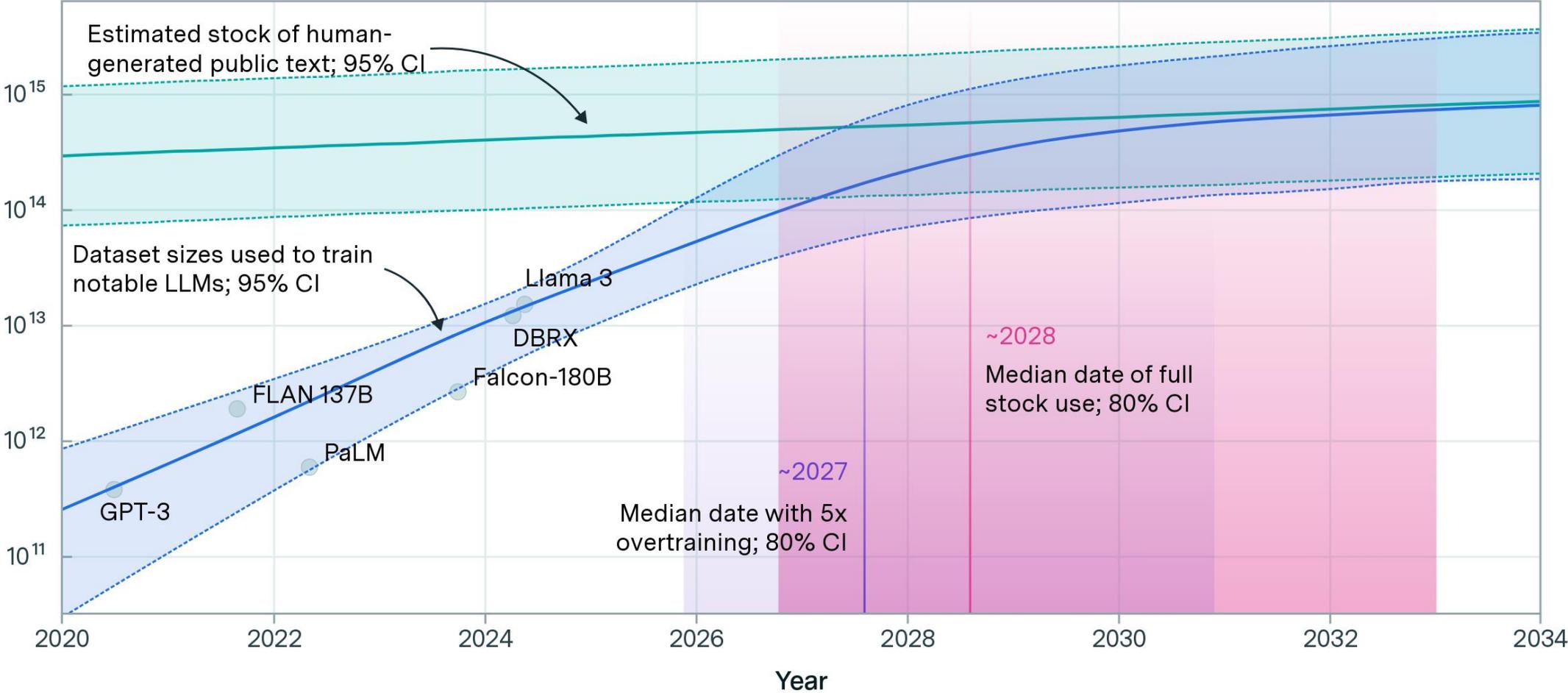
Price (USD per million tokens)



- GPT-3.5 Turbo level or better on general knowledge (MMLU)
- GPT-4 level or better on Ph.D. level science questions (GPQA)
- GPT-4o level or better on Ph.D. level science questions (GPQA)
- Other benchmarks and performance levels

Where are we today?

Effective stock (number of tokens)



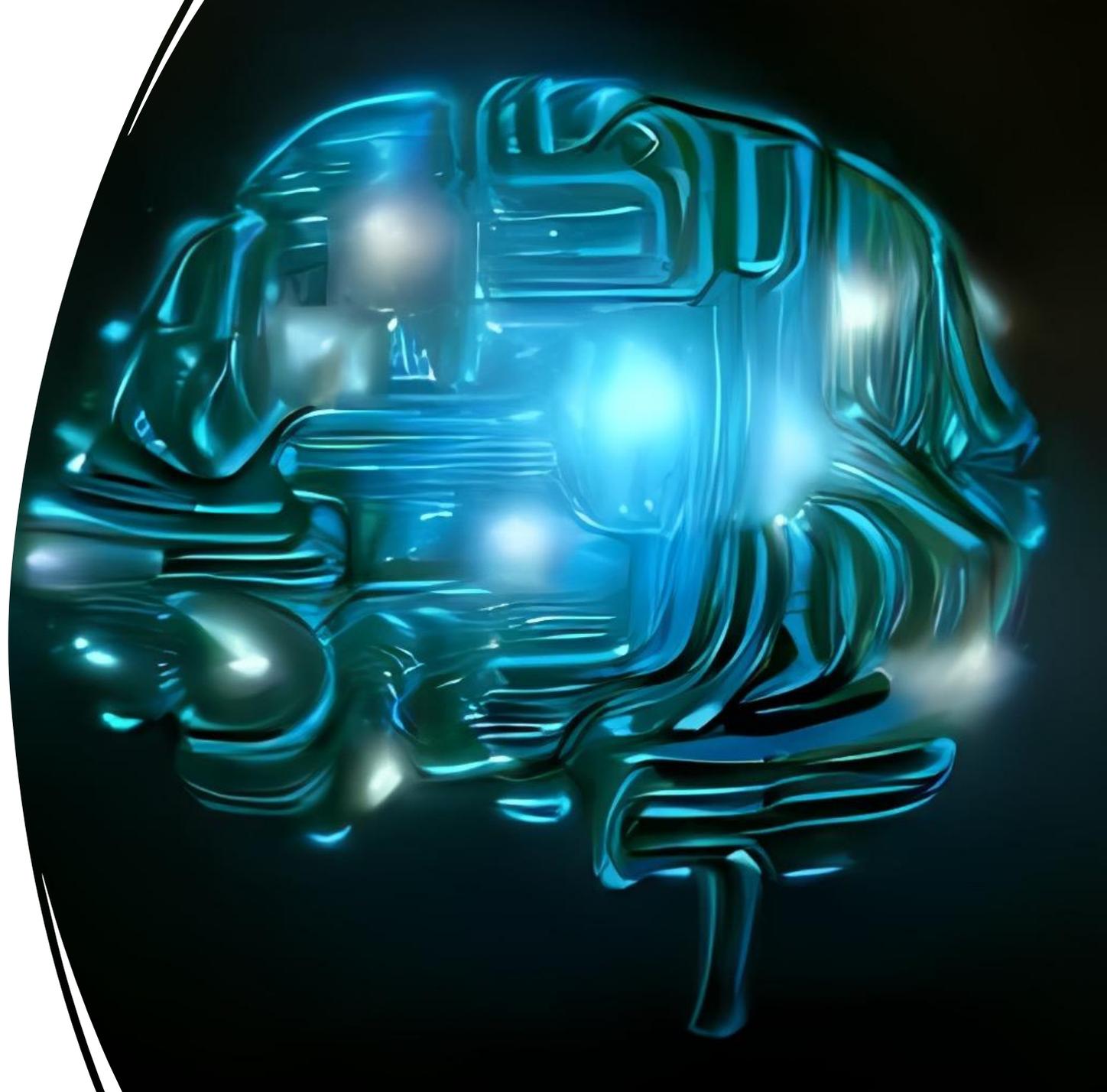
Common perceptions of AI

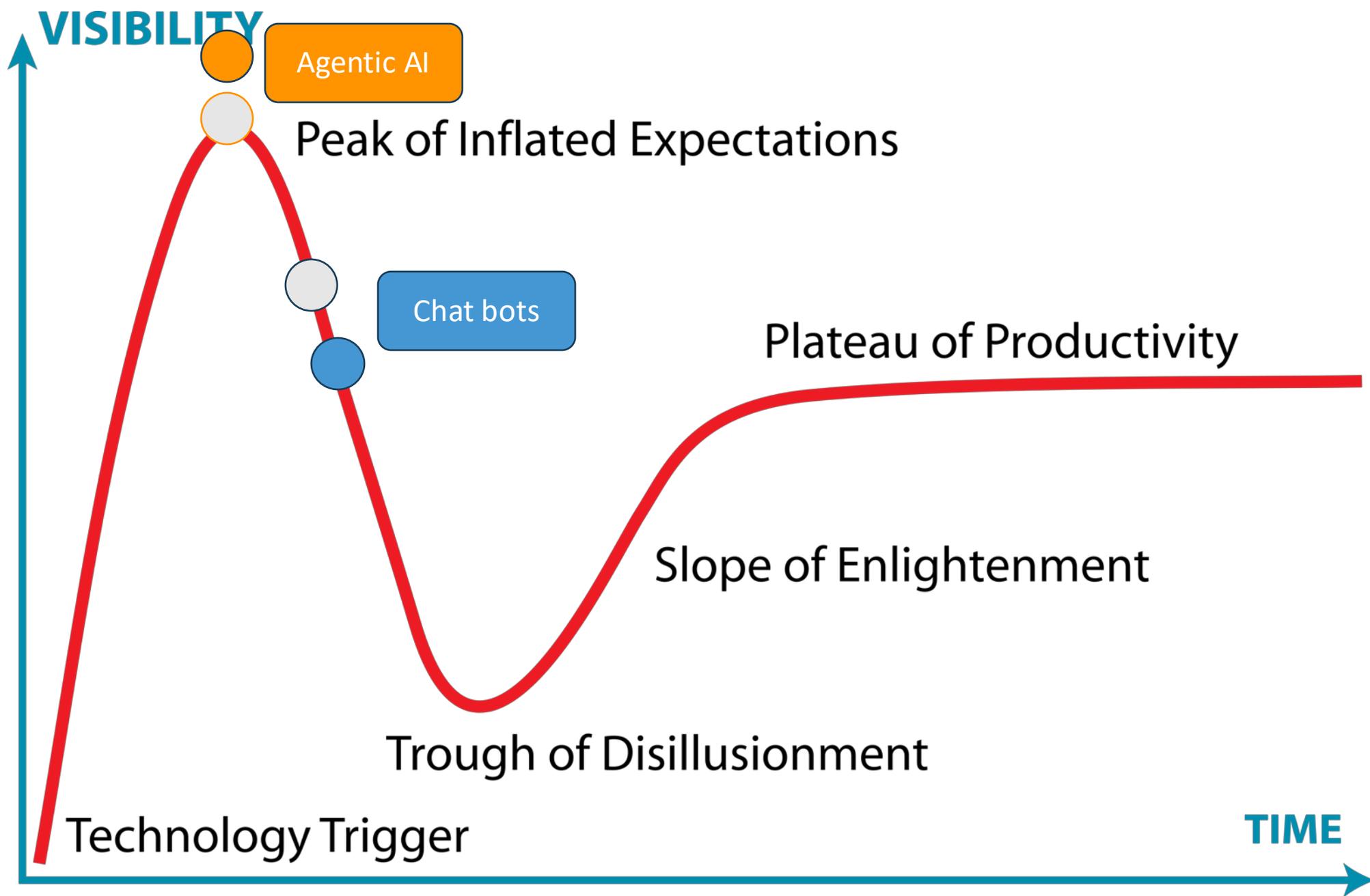


"Greetings, I am your new A.I. assistant...I just sorted all your emails...cleaning up invoices...corporate inefficiencies detected...starting mass layoffs...dumping stocks on the market...poor market reaction, selling all assets...this company is now insolvent, shuttering all operations...you are now fired...goodbye!"

Introduction to AI

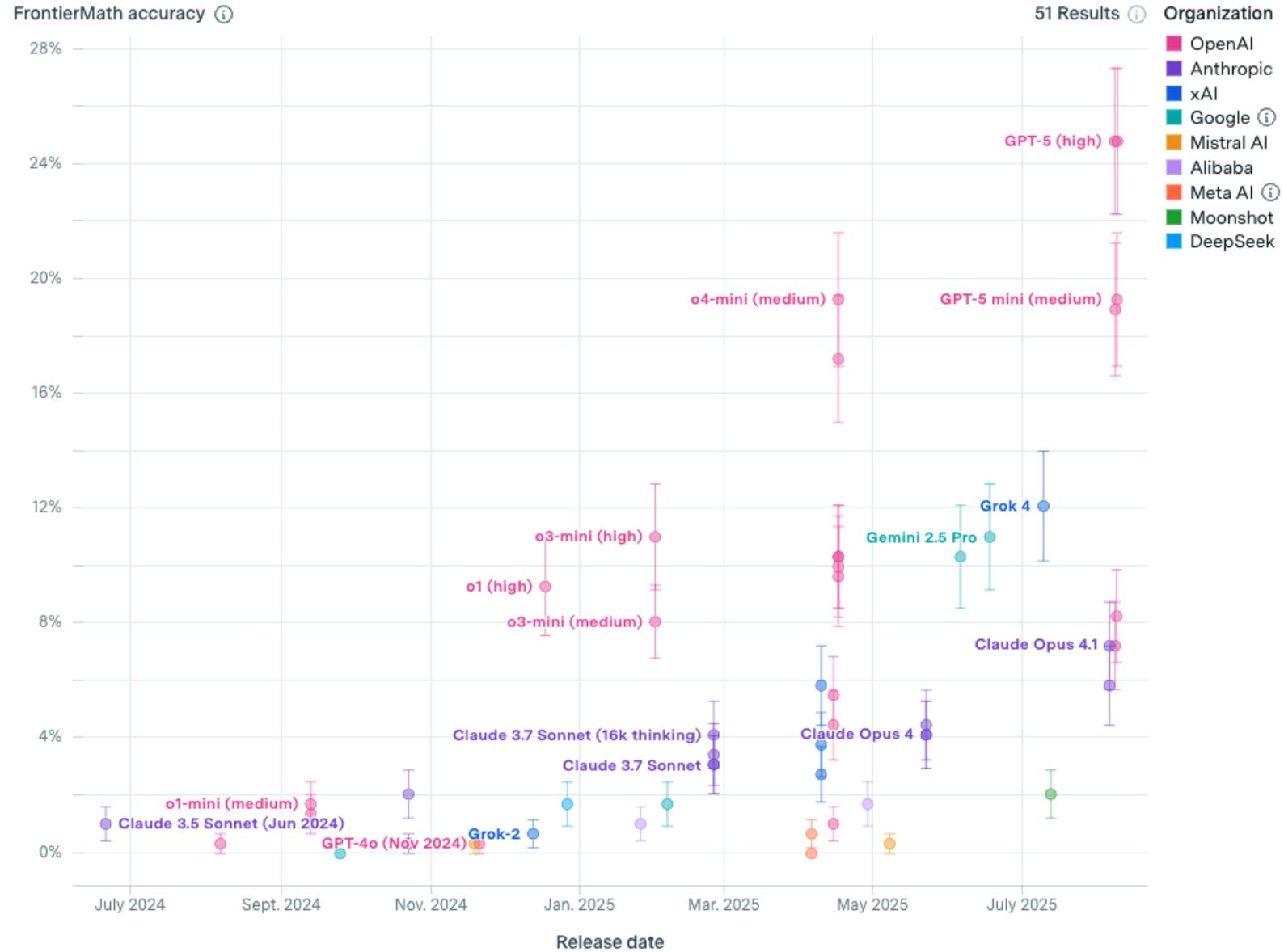
- Historical Context
- Impact on the Business
- Potential Good
- Potential Bad
- Hands on Examples





Math Proficiency

AI performance on a set of expert-level mathematics problems



2024 ● 2025

Science Proficiency

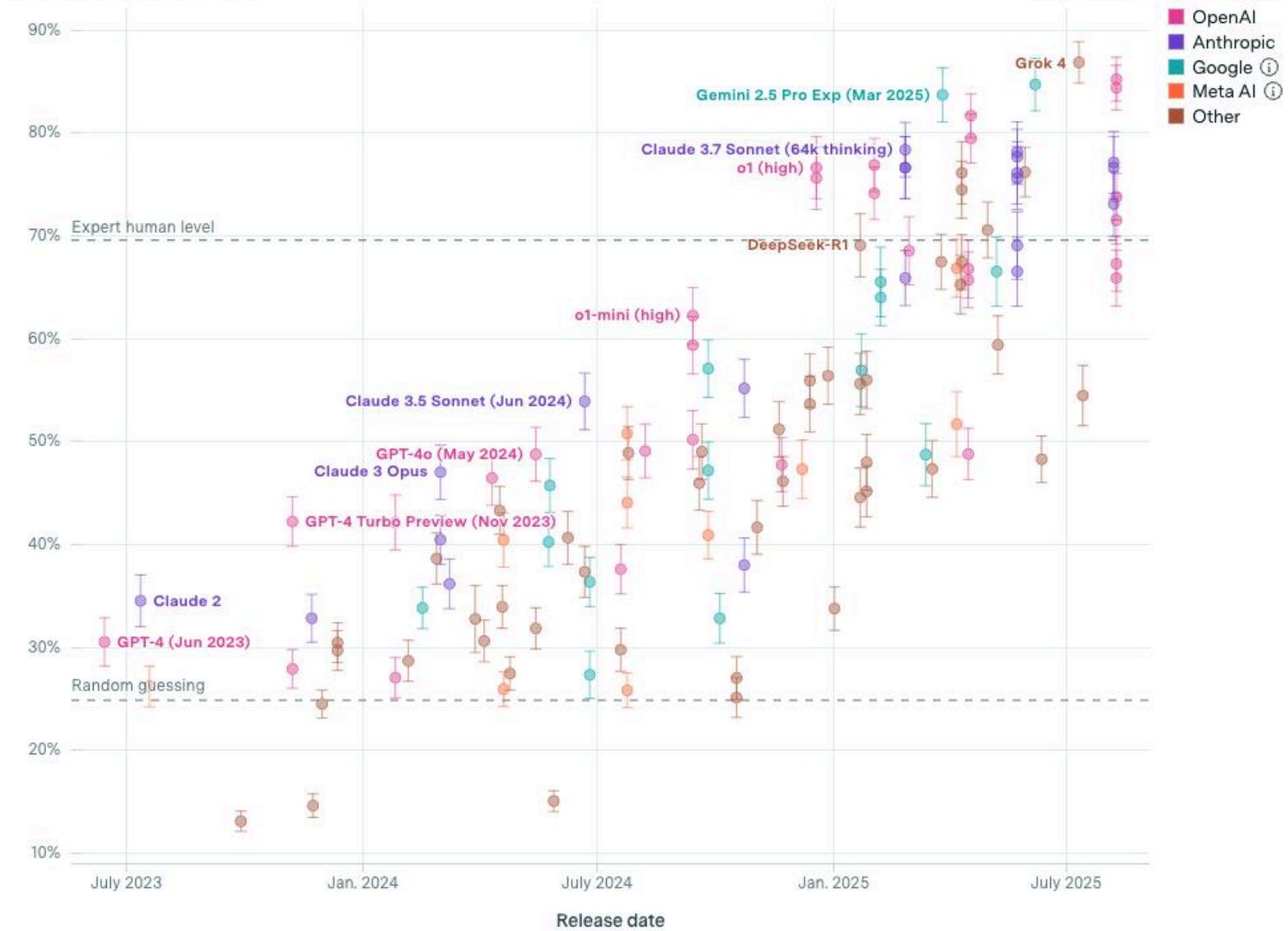
Ph.D. Non-Research

AI performance on a set of Ph.D.-level science questions

GPQA Diamond accuracy ⓘ

122 Results ⓘ

Organization ⓘ



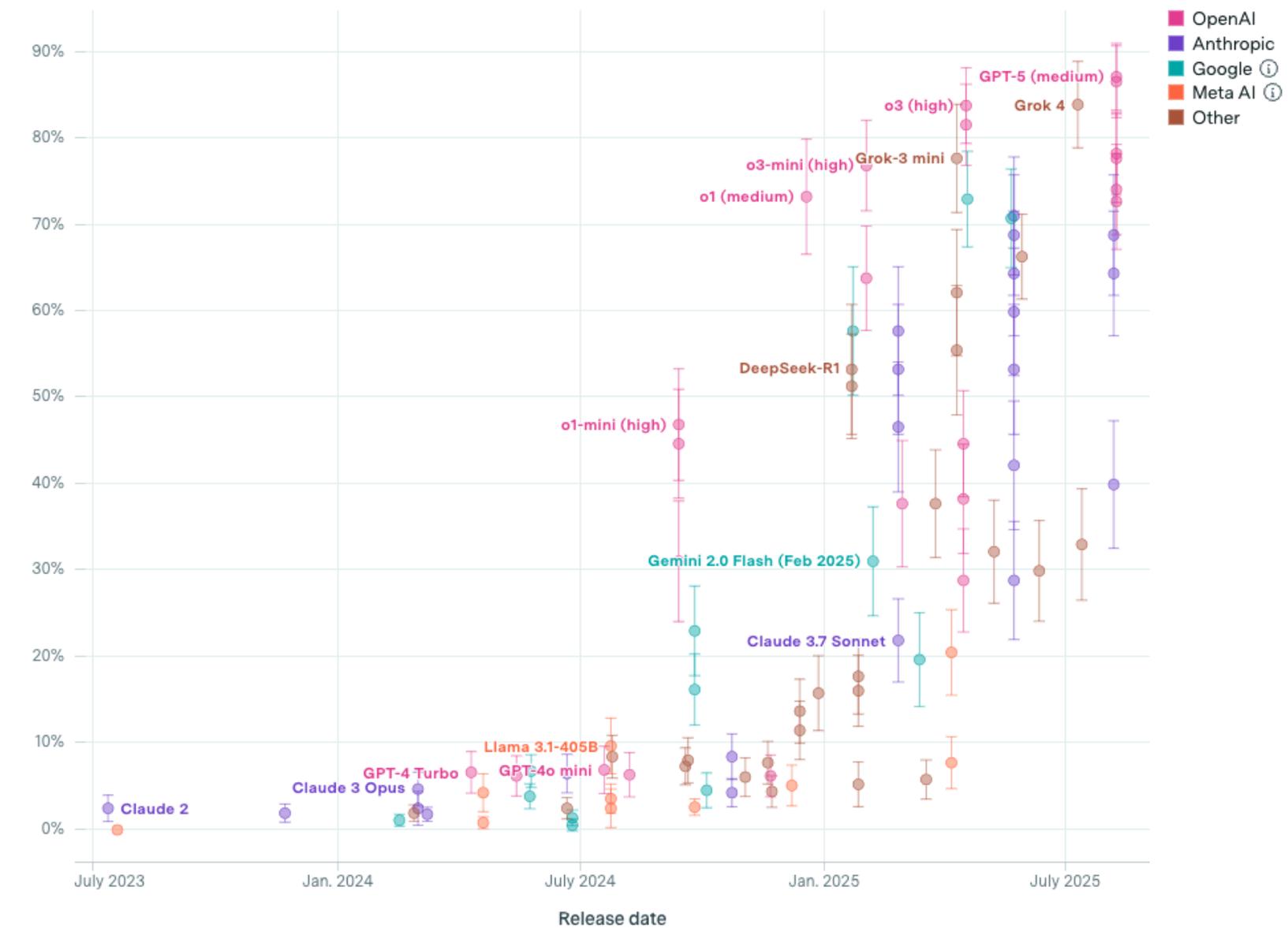
High School Math

AI performance on a challenging high-school level math contest

Mock AIME 24-25 pass@1 accuracy ⓘ

94 Results ⓘ

Organization



Web Research



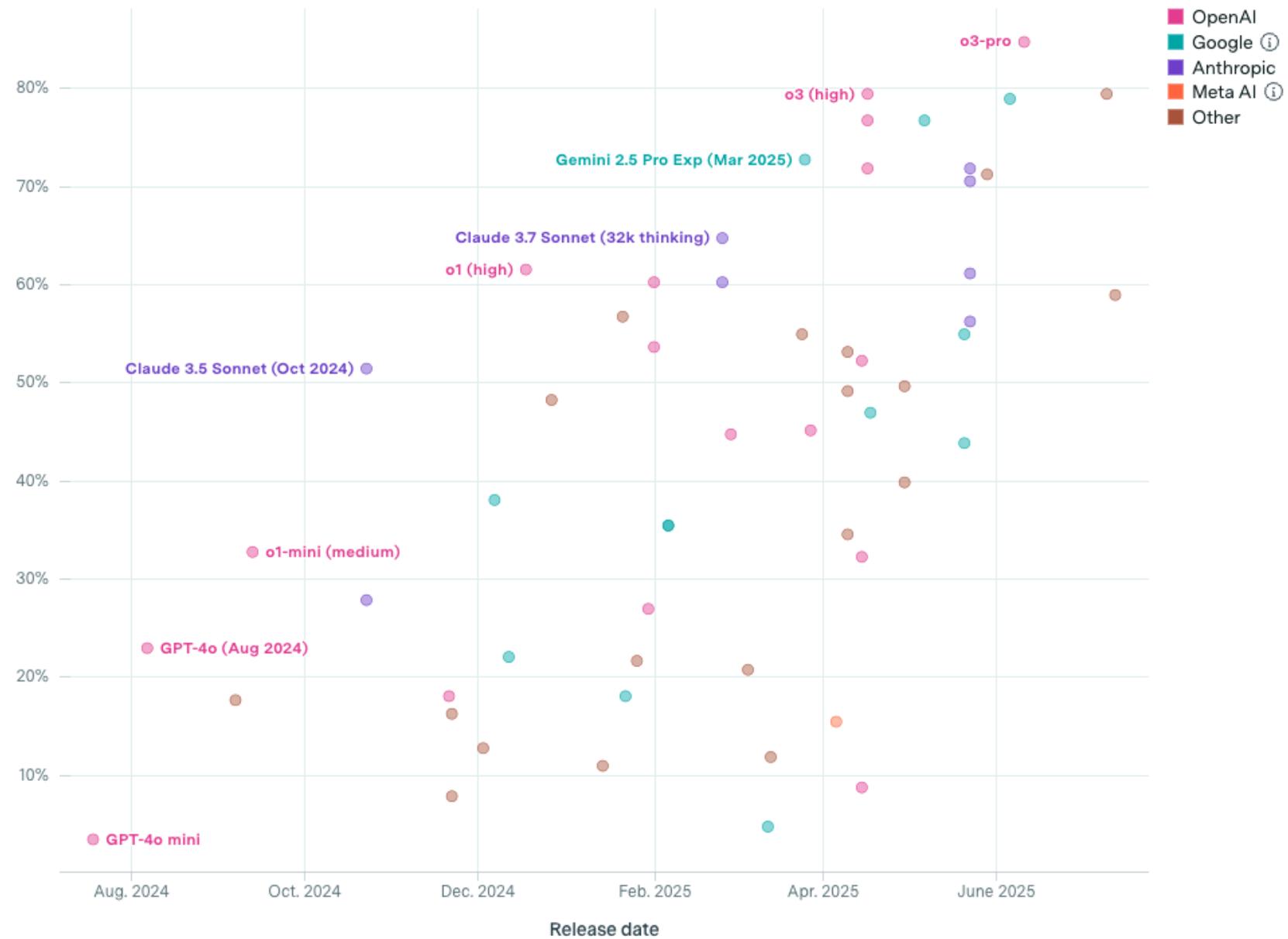
Software Development

AI performance on code editing tasks in multiple programming languages

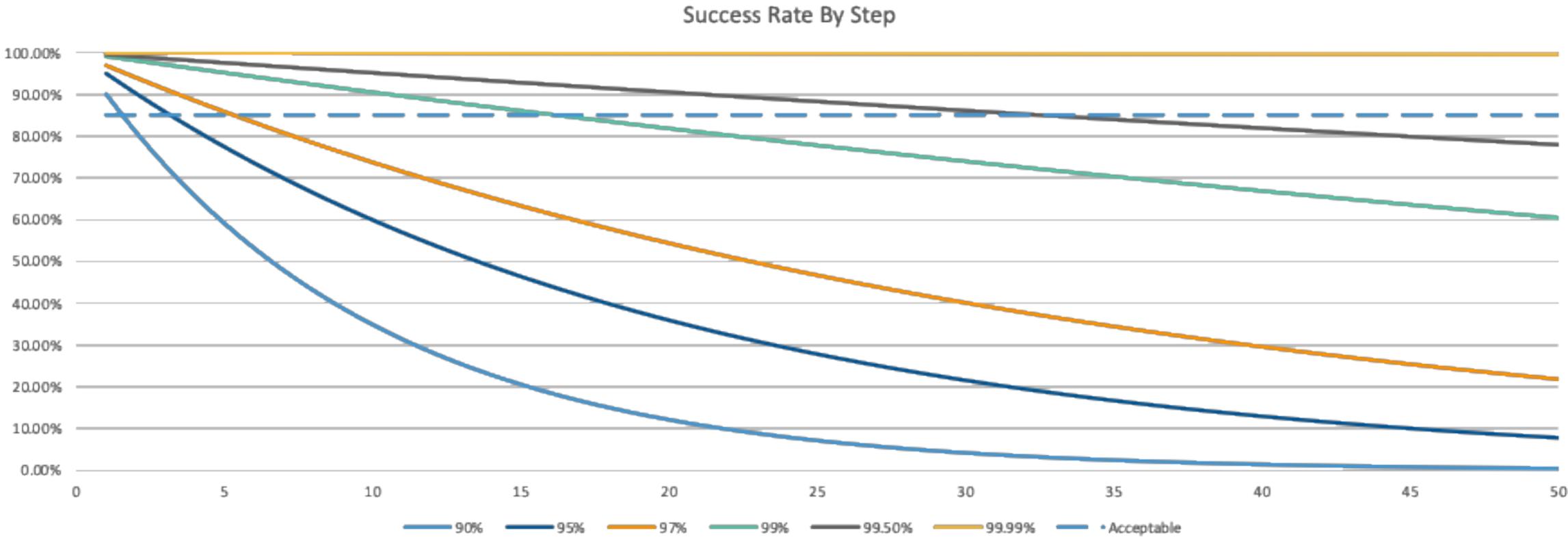
Aider Polyglot accuracy ⓘ

57 Results ⓘ

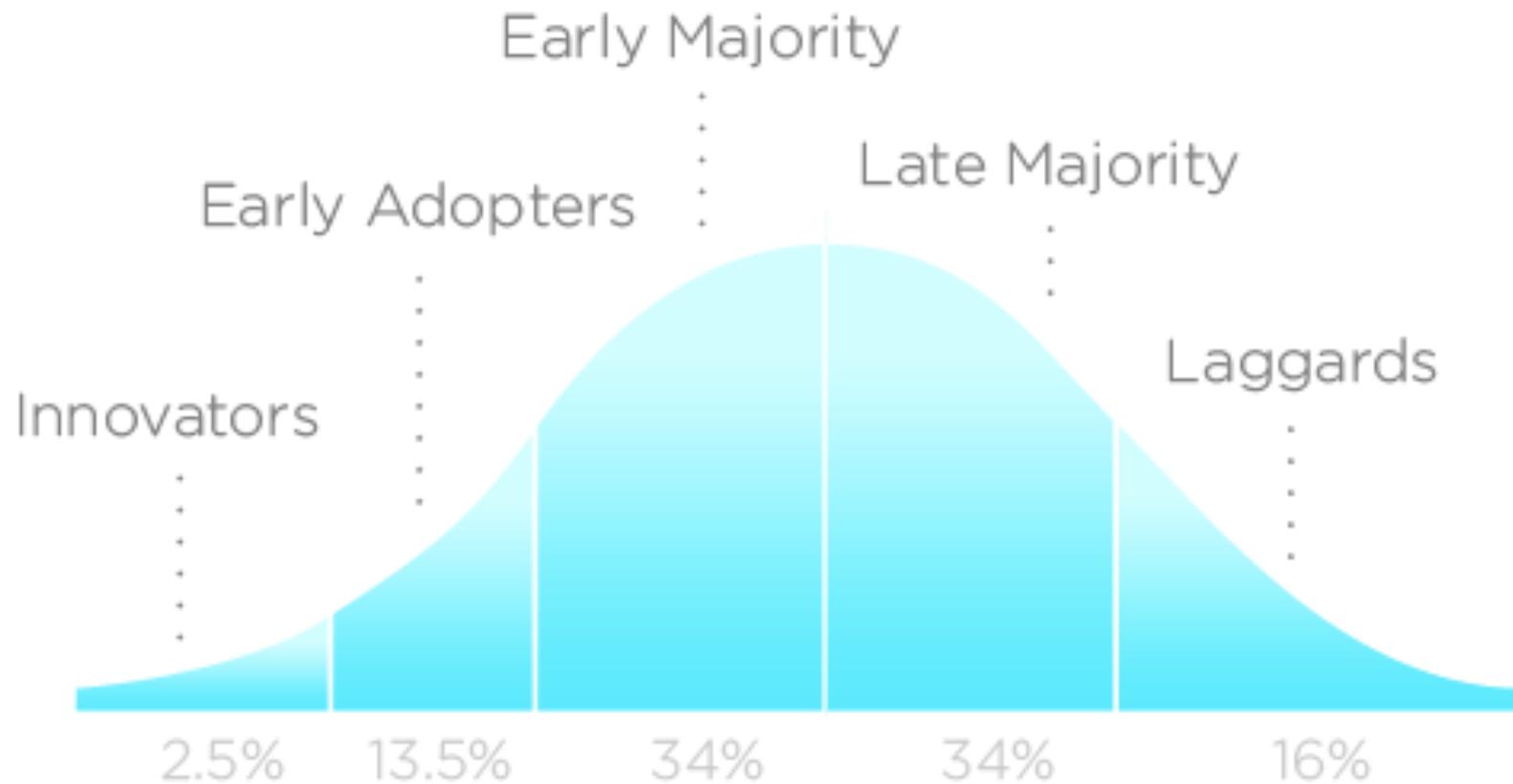
Organization ⓘ



Autonomous Agents – Stacking Errors

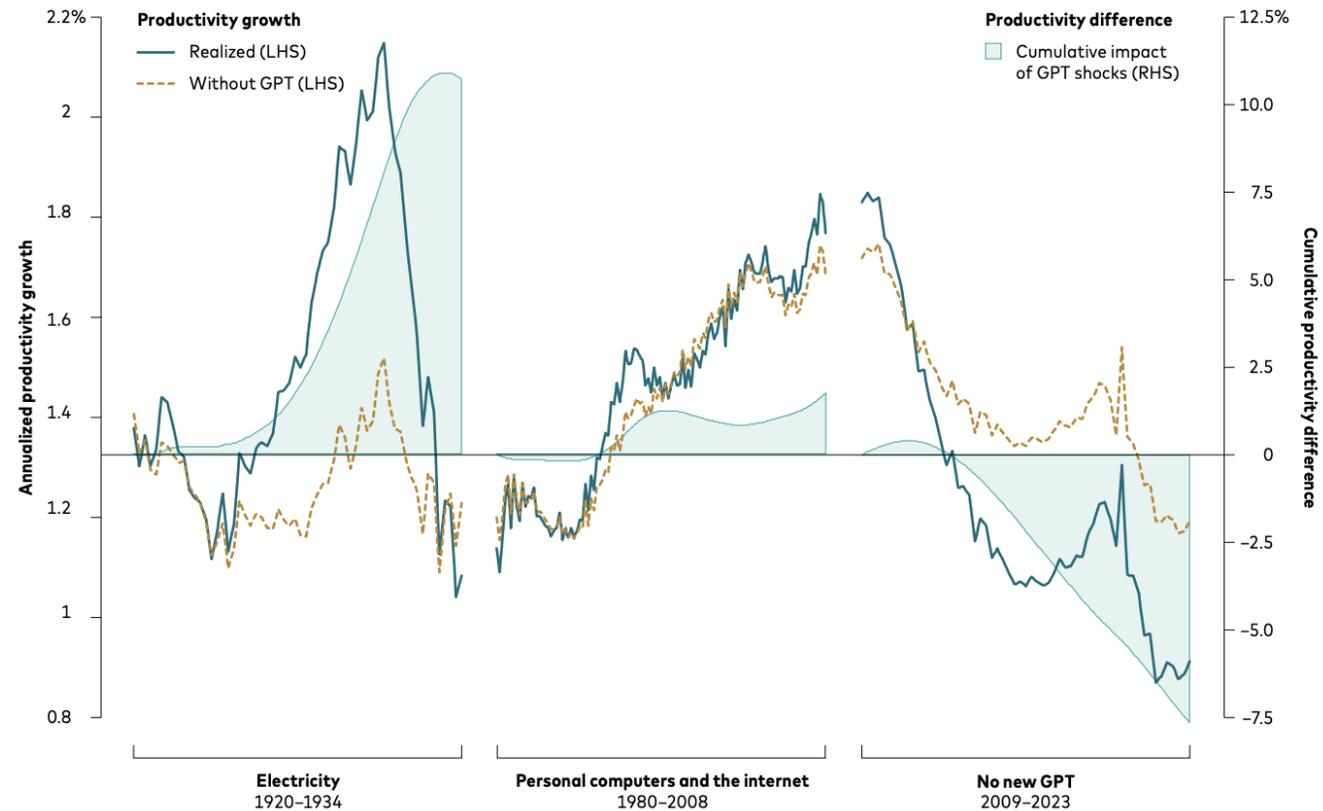


Diffusion of Innovations



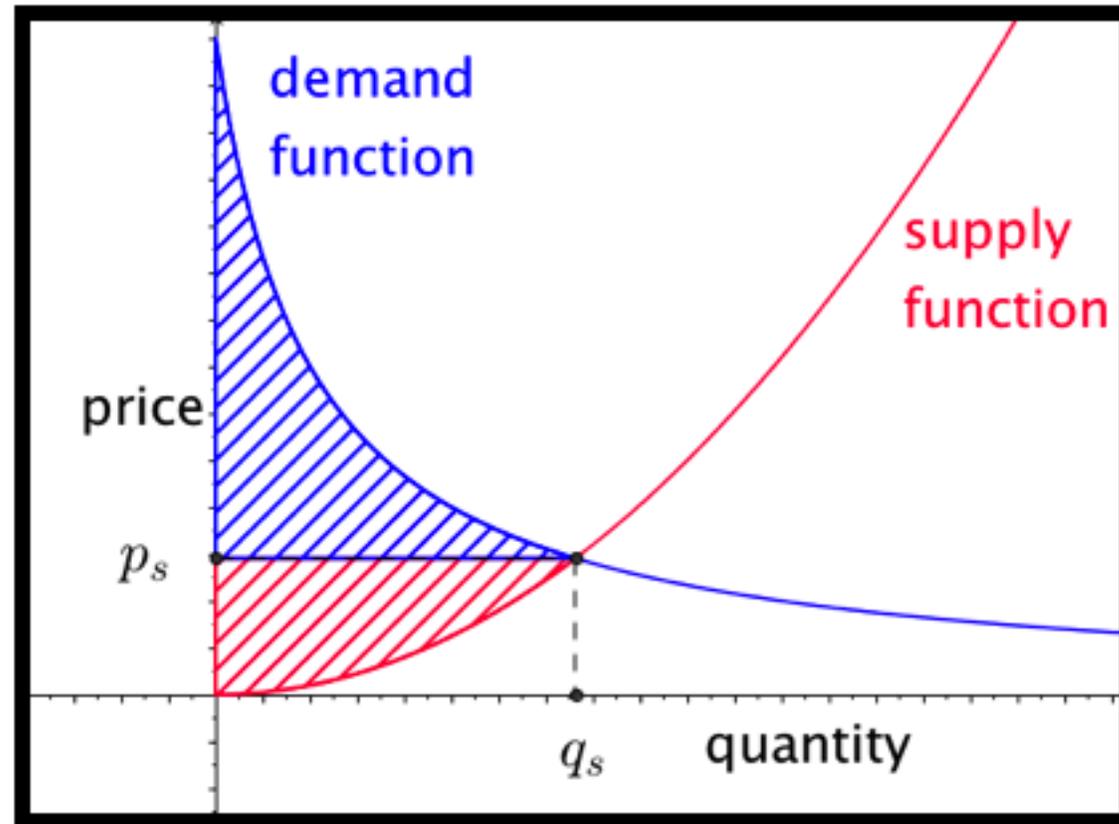
Not that GPT

FIGURE 6
GPTs: The good (electricity), the OK (computers and the internet), and the ugly (no new GPT)



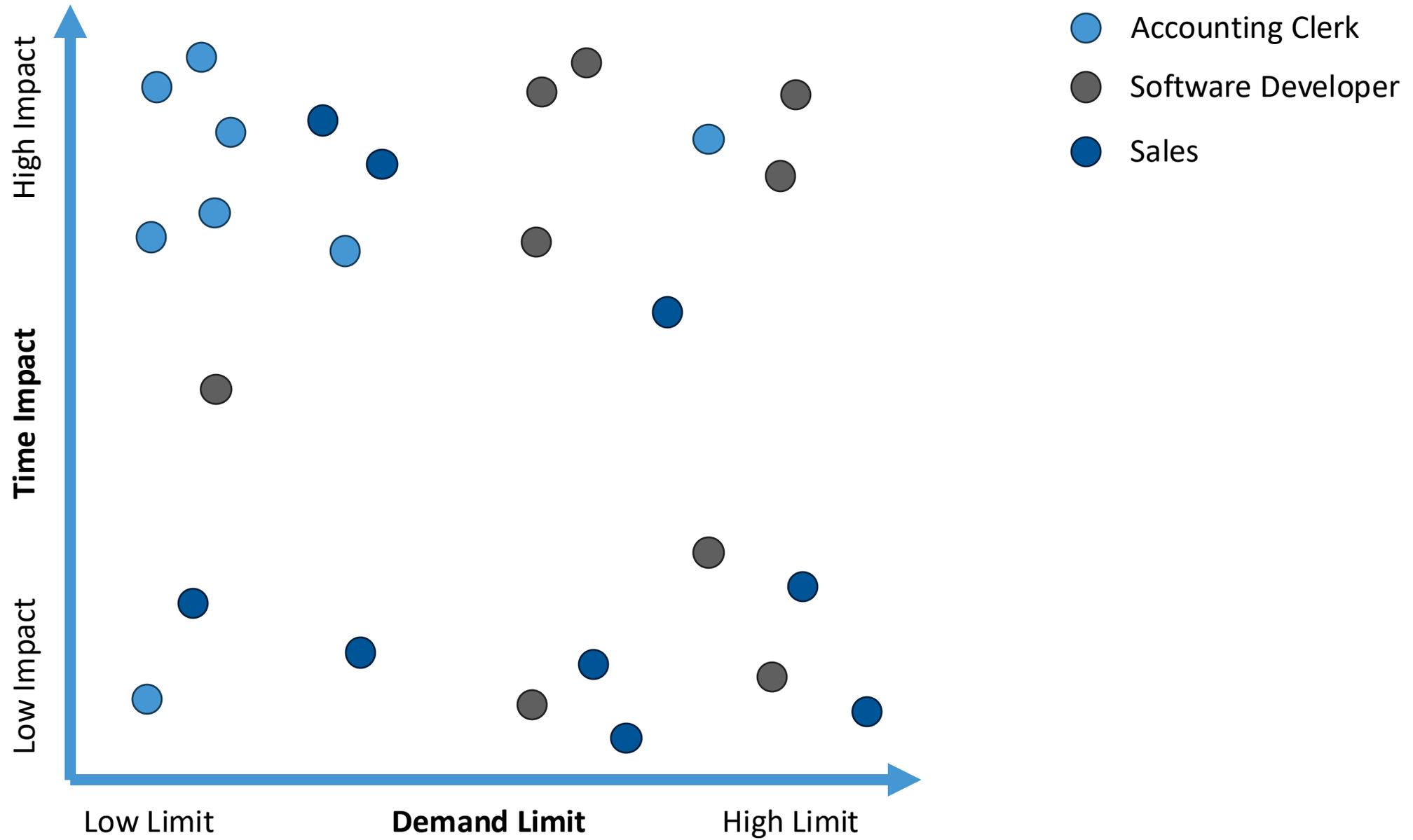
How to map the impact?

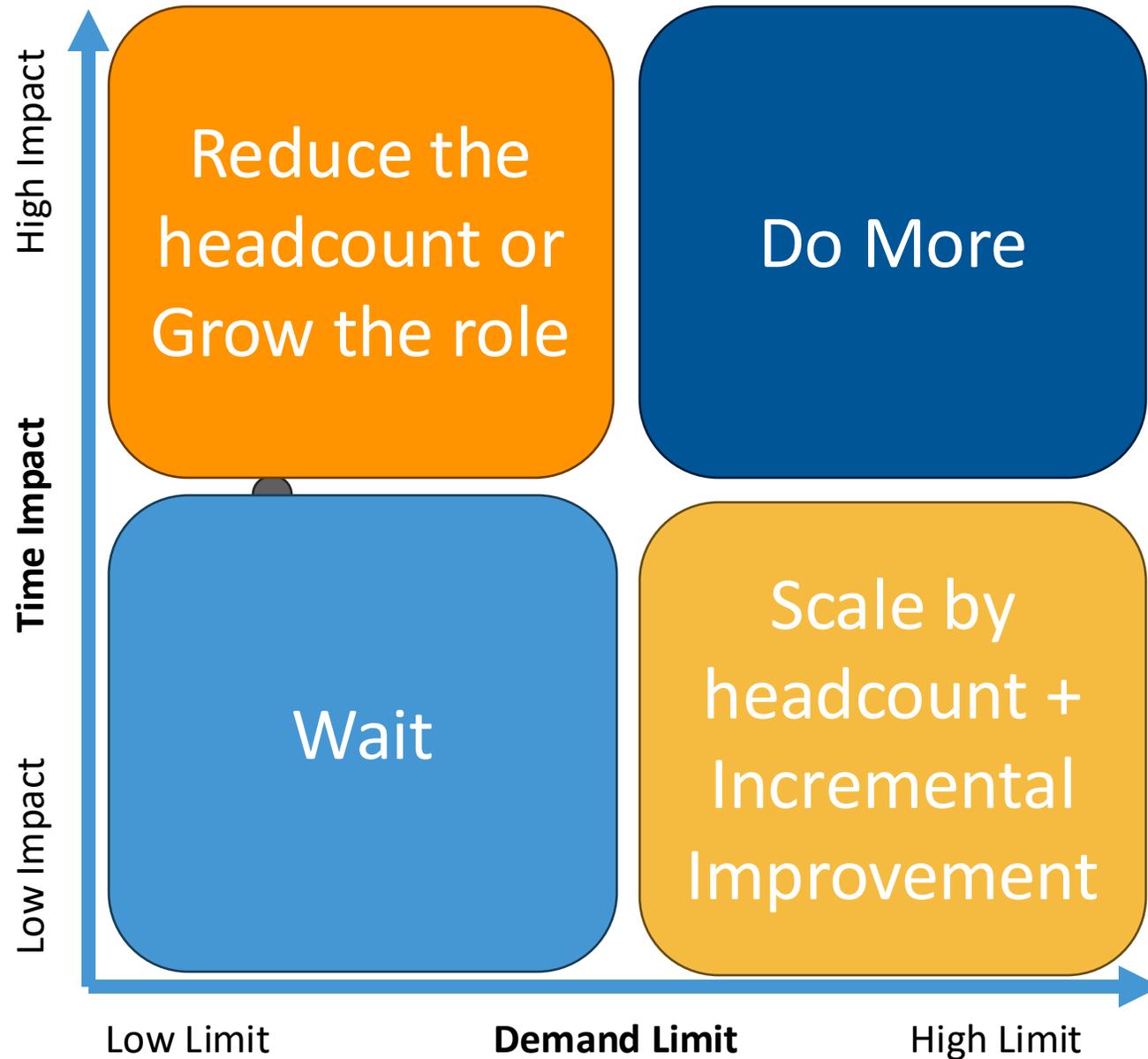
AI impact on tasks that are demand limited



AI impact on tasks that are collapsable







- Accounting Clerk
- Software Developer
- Sales

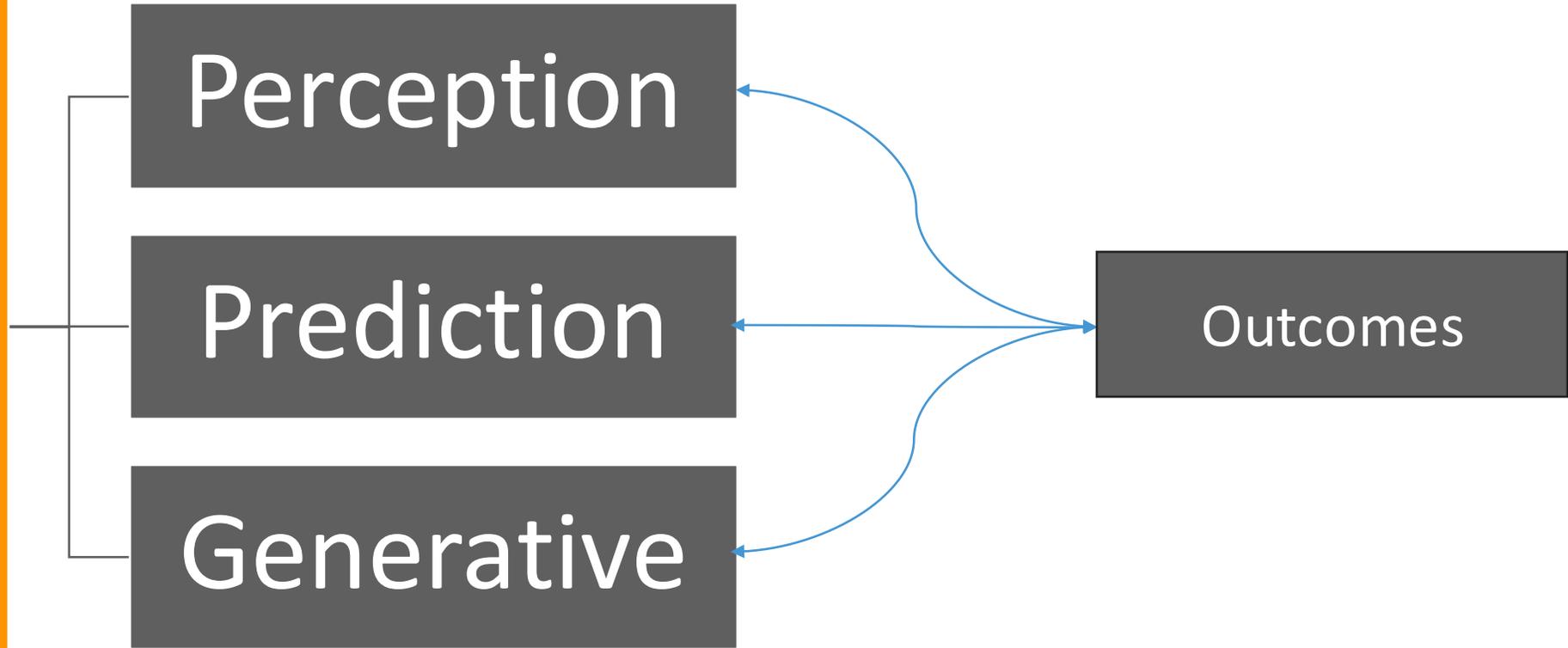
Defining AI

What is your definition?

AI and ML at the simplest



Artificial Intelligence



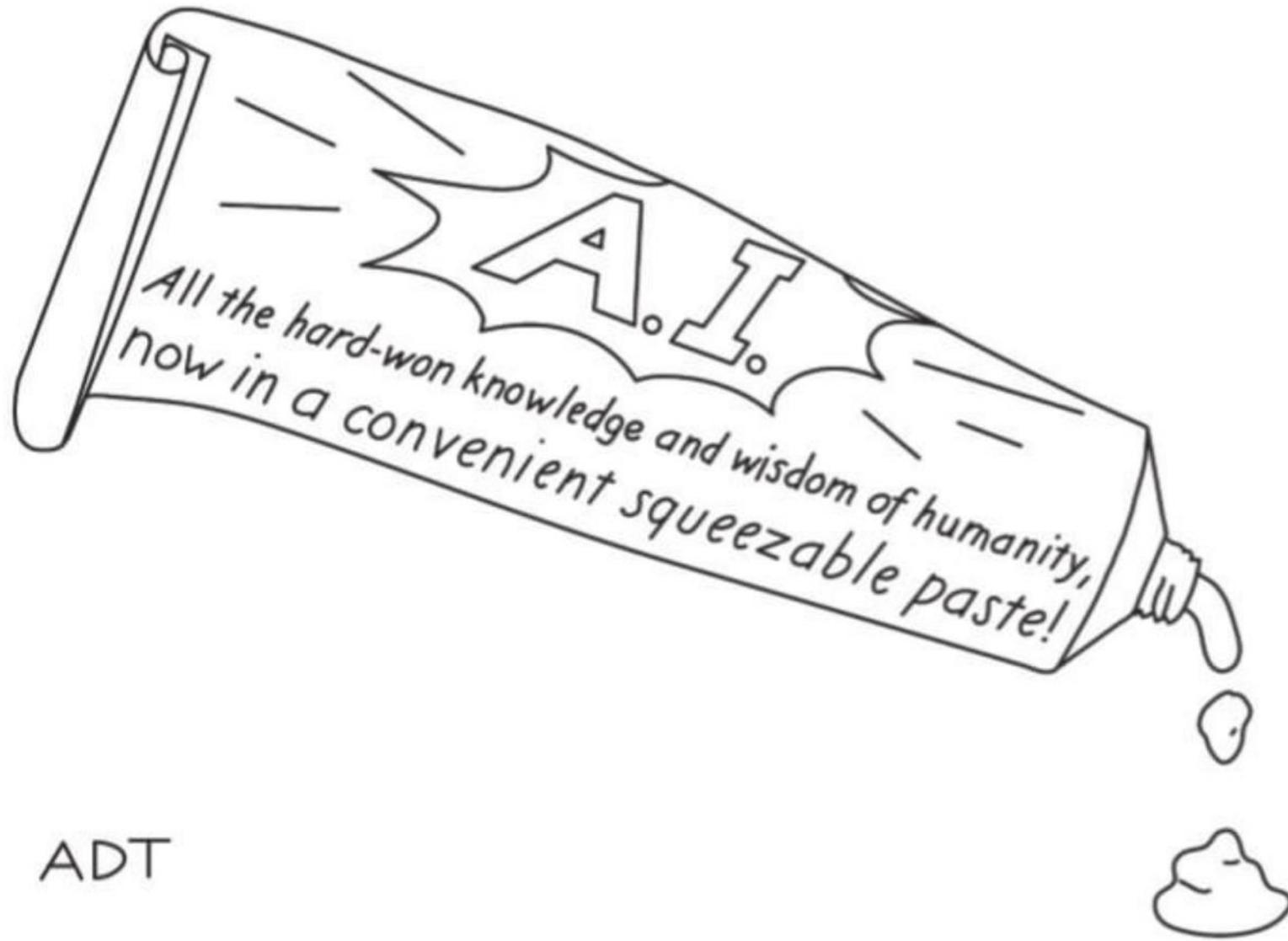


Complexity

Lies, Lies, and Statistics

- **Mistruth of commission**
 - Purposeful lying or creation of knowingly fraudulent data
- **Mistruth of omission**
 - Every stated fact is true, but some facts are unstated
- **Mistruth of perspective**
 - Data and facts are true from a perspective, but combining them doesn't yield 100% truth
- **Mistruth of bias**
 - Data and facts were available, but unable to be considered due to focus or experience
- **Mistruth from Frame of Reference**
 - A lack of experience with an event makes understanding basic data nearly impossible.





ADT

AI Problems and How we solve them



Rules of Thumb for AI Development today

1. Stay as close to off the shelf as possible to start
2. Look for 5-10x returns minimum in picking the problems
3. Define well the problem space before starting



Picking the
right problem is
crucial

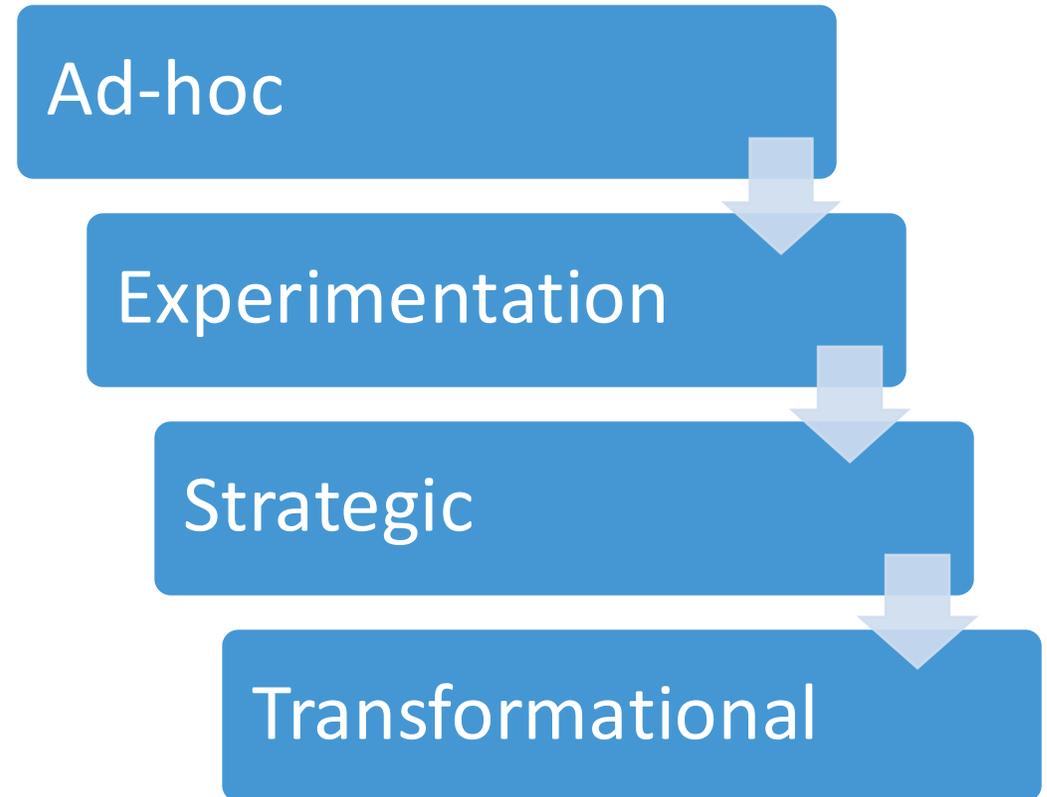


From the Past to the Future



What is the AI Adoption Framework?

AI Adoption Framework is a framework to help enterprises evolve their ML/AI project adoption to follow best practices and eliminate potential areas of unnecessary cost. Proper implementation of this framework is necessary to ensure a profitable adoption of AI.



How to get started?

Starting tomorrow – I want to be a superhuman

Problem Selection

- Create a list of the top 10 things you do each week by the amount of time they take
- Put stars next to the ones you HATE doing
- Pick the highest amount of time starred item

Exploration

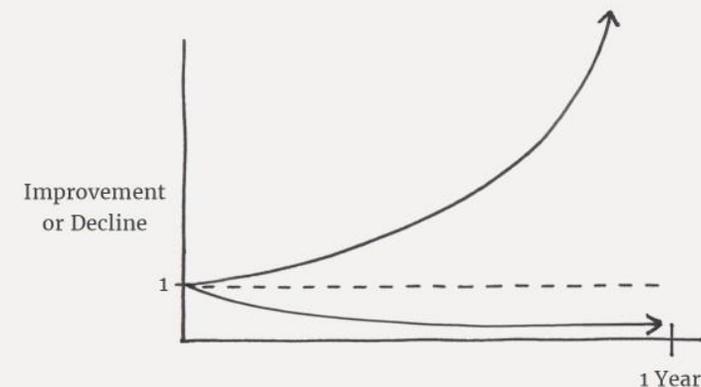
- Spend 30 minutes thinking through how to automate significant portions of that work for yourself
- Test it – Did it save time and make you feel better about your job?
- If yes to both, share it so that the rest of the organization can benefit

Guidelines

- New tools – Make sure they get reviewed BEFORE you use them
- New Knowledge – Aim for 10% understanding to start, not 50%
- Share – Talk about your experiments, others may be doing similar, and you can go farther together

The Power of Tiny Gains

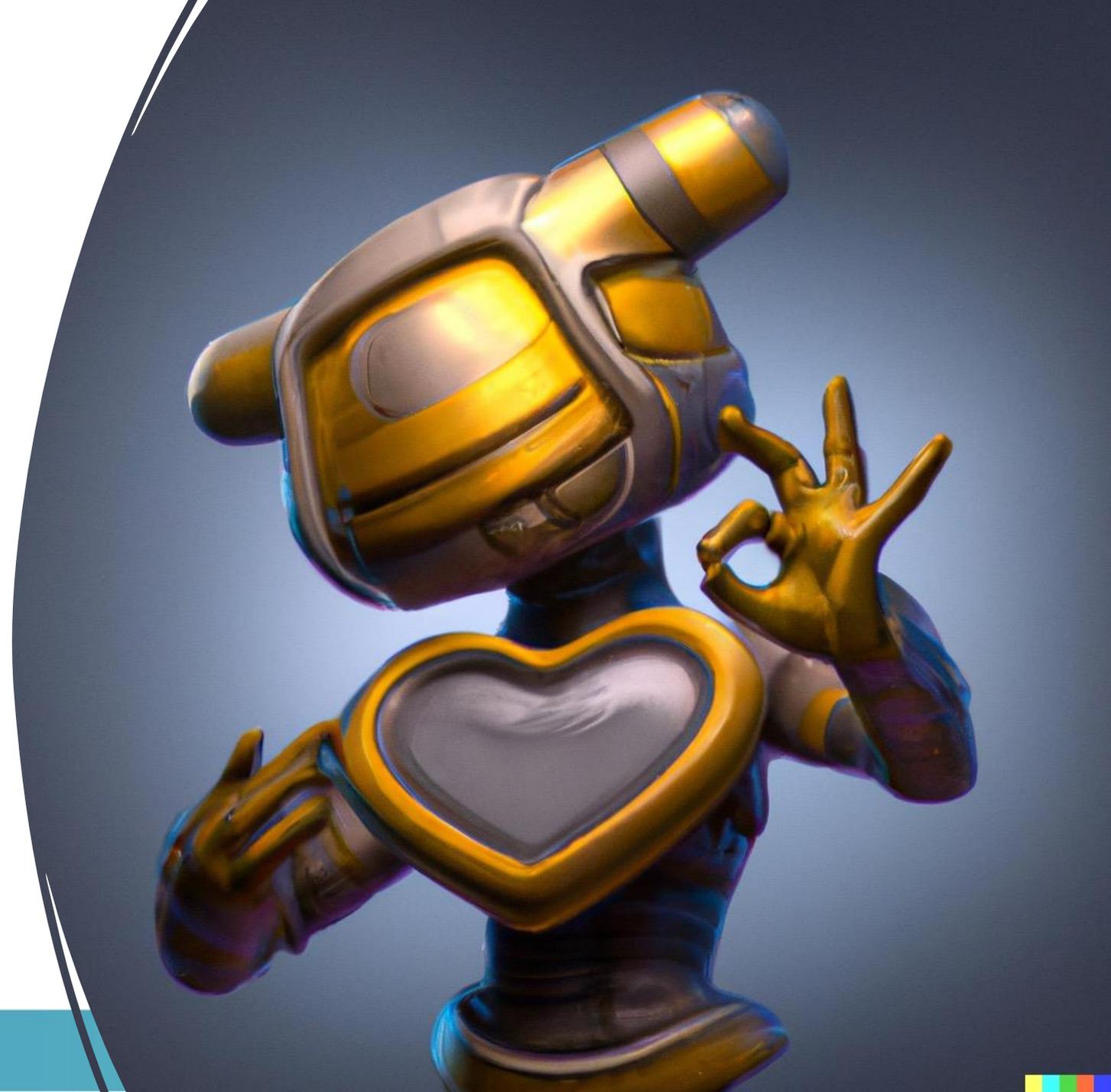
$$\begin{aligned} 1\% \text{ better every day} & \quad 1.01^{365} = 37.78 \\ 1\% \text{ worse every day} & \quad 0.99^{365} = 0.03 \end{aligned}$$



JamesClear.com

AI for Good

- Healthcare Advancements
- Environmental Sustainability
- Education and Personalized Learning
- Humanitarian Aid and Disaster Response
- Accessibility and Inclusivity



AI for Bad

- Autonomous Weapons
- Bias and Discrimination
- Privacy and Surveillance
- Misinformation and Deepfakes
- Cybersecurity Threats



Questions

