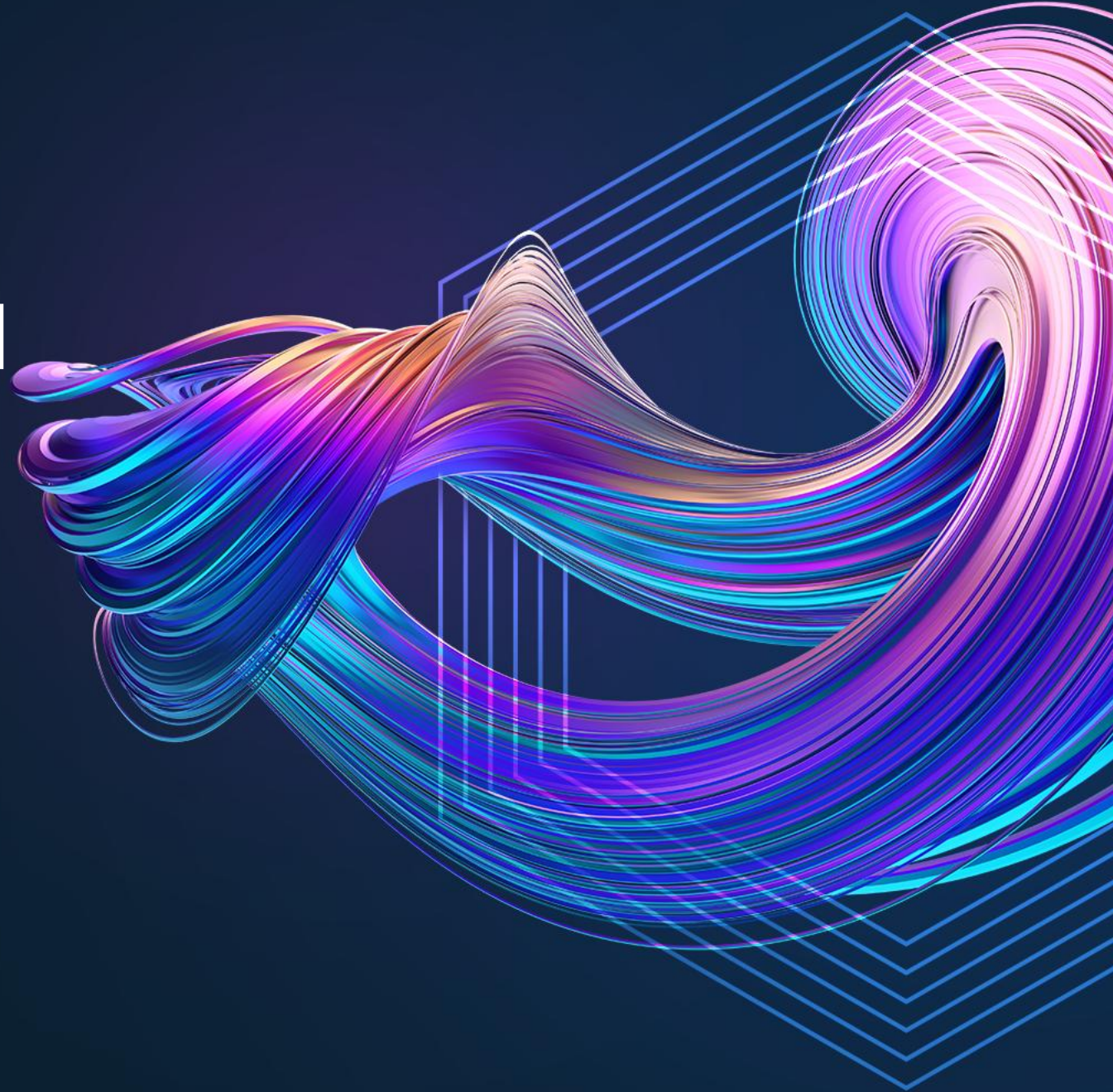


The Evolution of AI and the Path to Partnership

William Chappell



Neither path is enough on its own

Commercial



Scale but
lighter mission-depth

Strong on one dimension
Constrained on the other

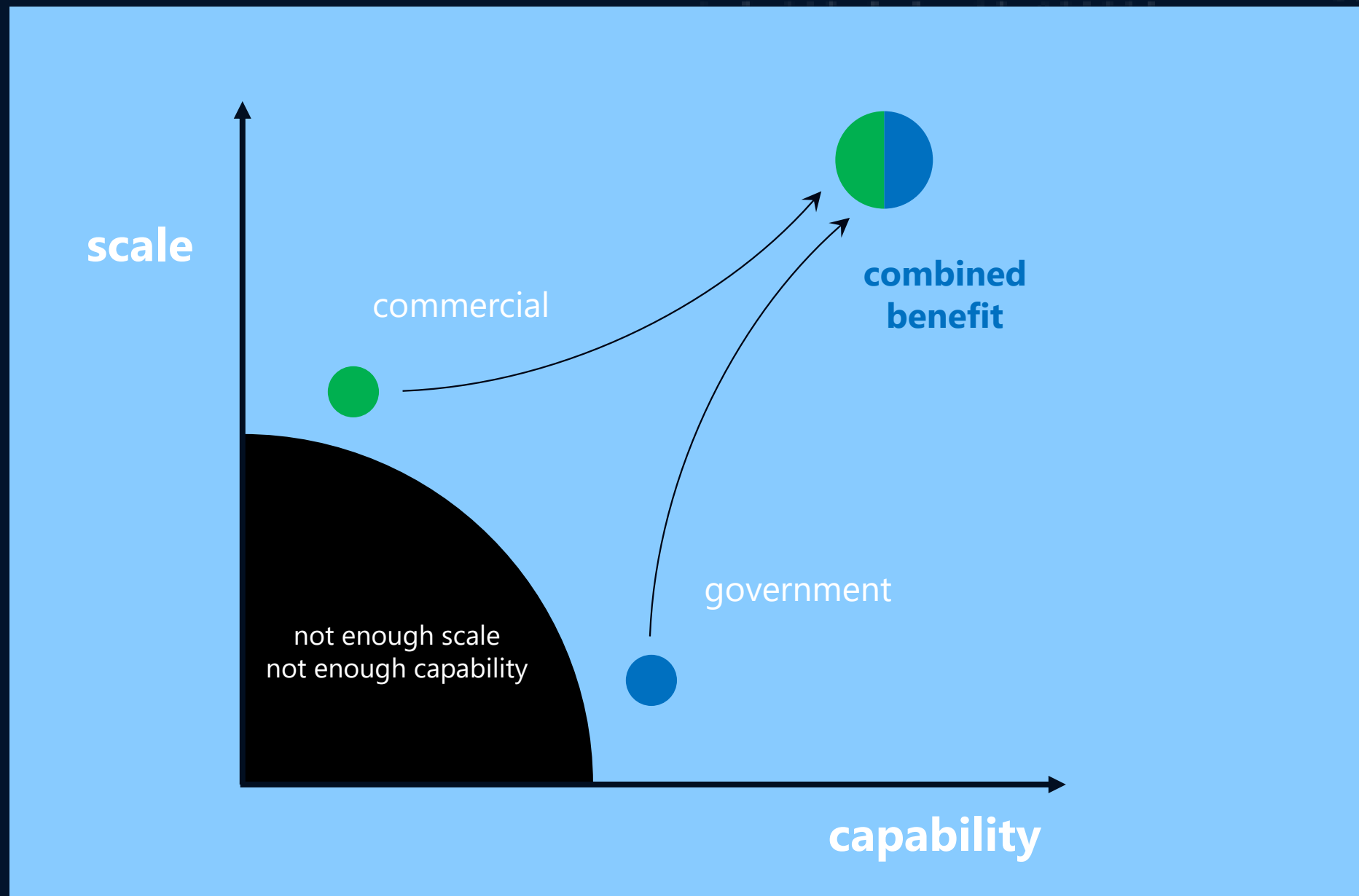
The
opportunity
is integration

Government



Capability without as
broad scale

Strong on one dimension
Constrained on the other



$$\text{Scale} + \text{Capability} = \text{Combined Benefit}$$

A Combined Government and Commercial AI Path

Third Offset Strategy
details AI and Human/
Machine teaming as
differentiation for
applications

- Artificial Intelligence
- Human-machine collaboration
- Advanced cyber capabilities
- Robotics and unmanned systems
- Networked warfare and battle management system
- Autonomous systems
- Machine-speed decision making

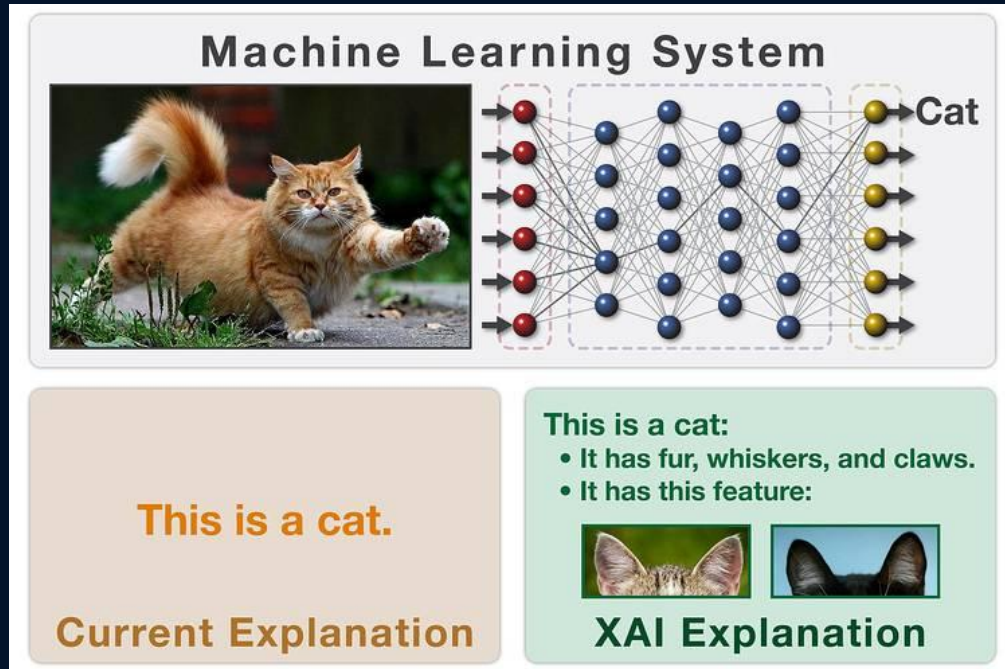
2016

GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

~ 2017 – DARPA Discussions After 3rd Offset Strategy

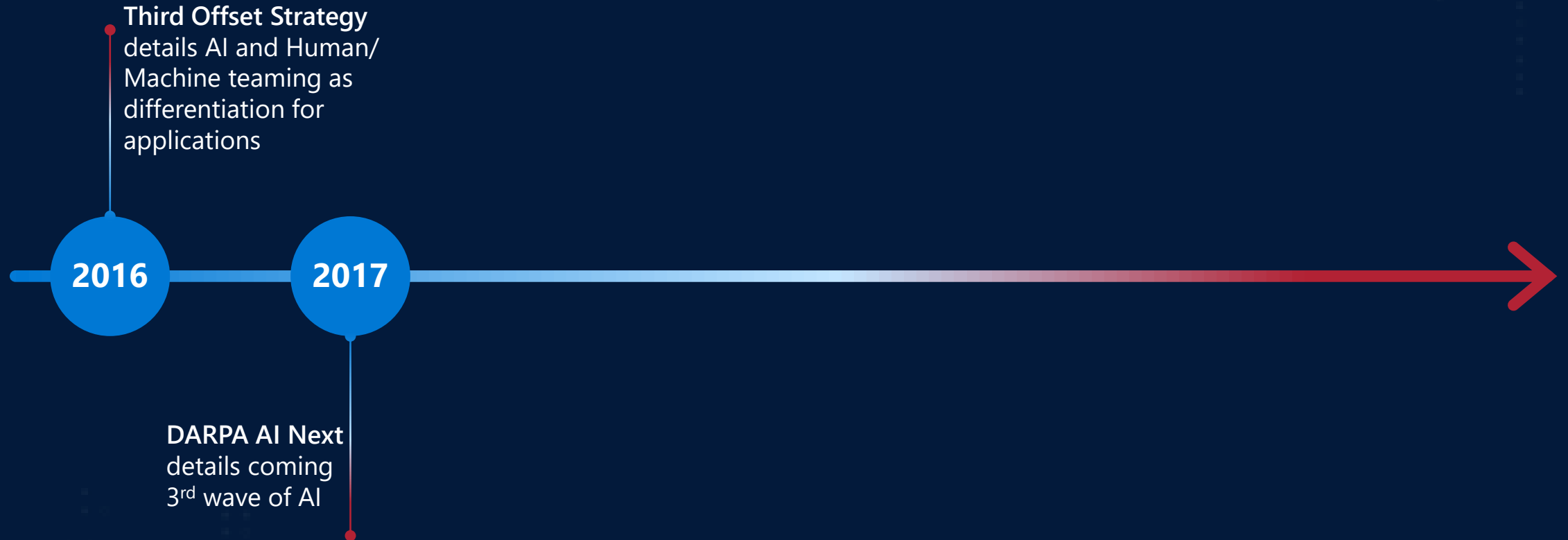


"A young boy is holding a baseball bat." Second wave systems are statistically impressive, but individually unreliable.



AI Common Sense and Understanding of Physics

A Path to Our Current AI



GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

3 Waves of AI – 2017 DARPA AI NEXT



FIRST WAVE

Hand crafted knowledge

Reason over narrowly
defined problems

No learning capability



SECOND WAVE

Statistical learning

Trained on Domain Specific
Data

Lack good reasoning

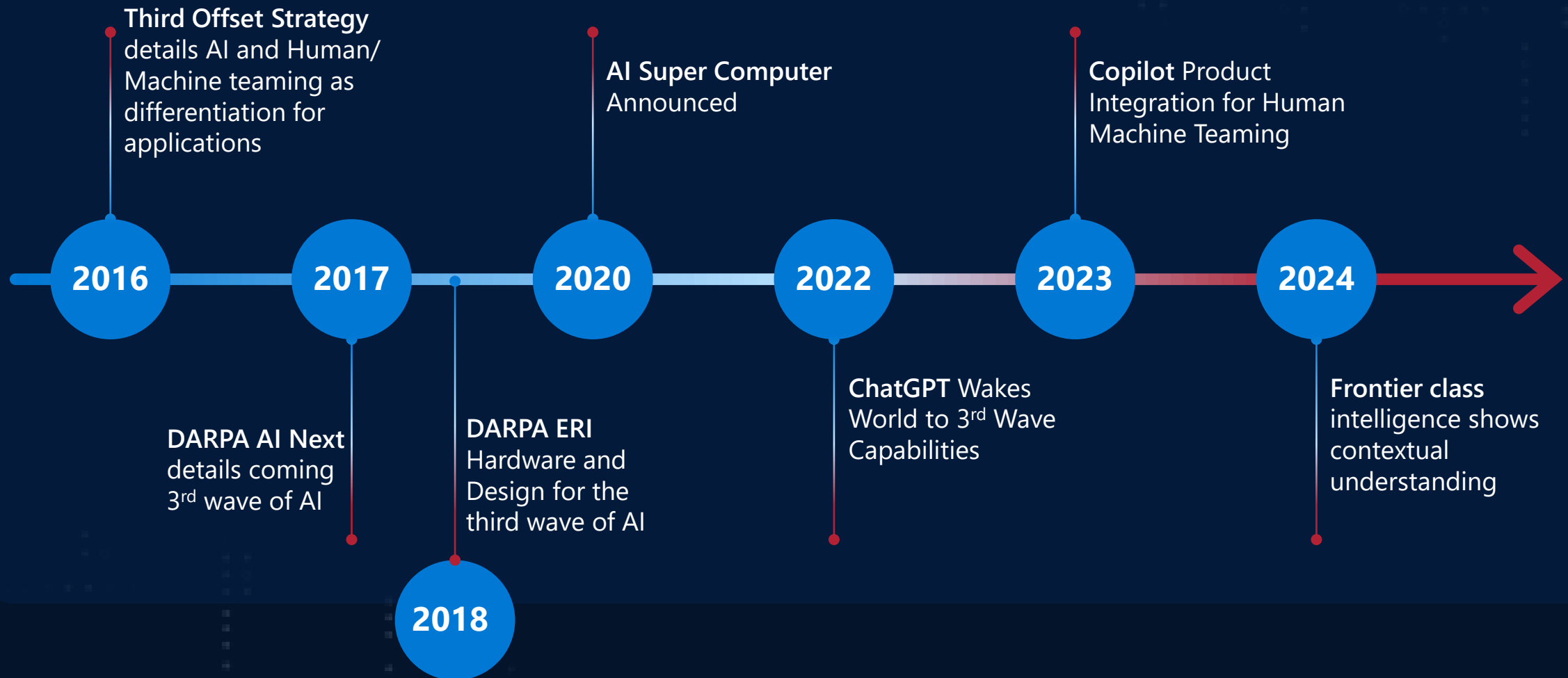


THIRD WAVE

Contextual adaptation

Understand context, reason
and decide

A Path to Our Current AI

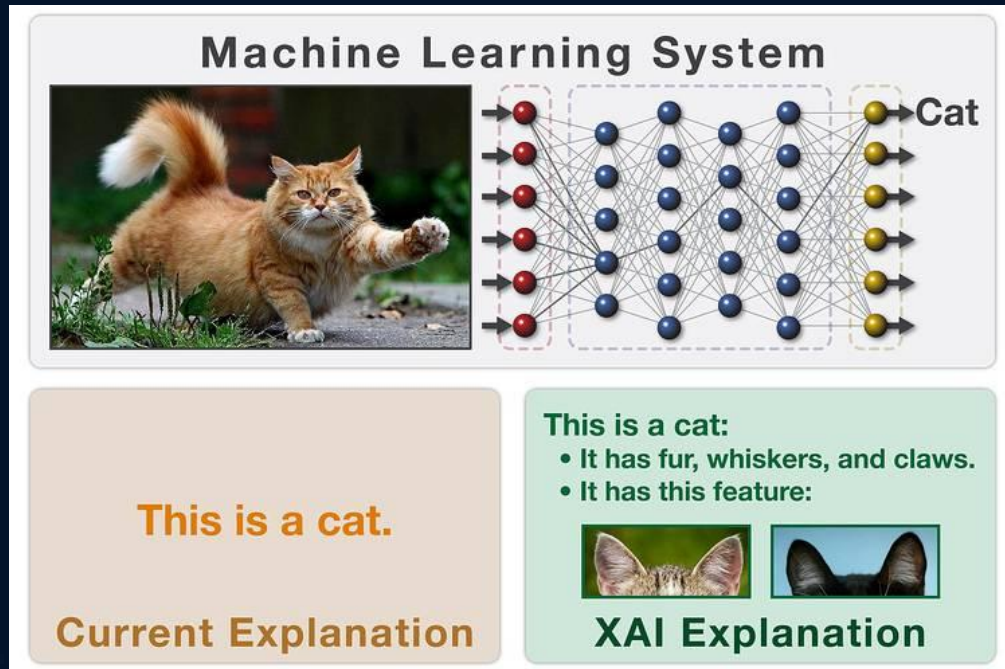


GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

A Decade After the 3rd Offset



DF



What is in this picture? What evidence do you have for that statement?

 The picture displays a fluffy orange cat with a distinct bushy tail, reaching out one of its paws, possibly while walking or playing. The cat is outdoors, with some plants and a paved surface visible in the background.

Evidence for this statement:

- The feline features such as whiskers, cat eyes, and the specific shape of the ears and face indicate that it is a cat.
- The bushy tail and the pattern of the fur, as well as the color, are typical characteristics of certain cat breeds or mixes.
- The surrounding vegetation and the ground texture suggest an outdoor setting.

A Decade After the 3rd Offset



"A young boy is holding a baseball bat." Second wave systems are statistically impressive, but individually unreliable.

DF



What's in this picture? Explain your reasoning.



The picture shows a young baby holding a toothbrush. The caption "a young boy is holding a baseball bat" is not accurate, as the object the baby is holding is clearly a toothbrush, identifiable by its handle and bristles. The caption may have been intended as a humorous exaggeration.

A Decade After the 3rd Offset



AI Common Sense and
Understanding of Physics

What will happen if the robot sawing the branch completes its task?



Your personal and company data are
protected in this chat

If the robot completes sawing the branch it's standing on, it would likely fall along with the branch, given the laws of physics. This is a classic comedic trope often seen in cartoons, where a character unknowingly saws off the branch they are standing on, leading to an inevitable fall. However, since this is a robot we're talking about, it might have some advanced features or capabilities to prevent such a fall! It's all up to the imagination in this artistic representation. 😊

4 of 30 responses •

2024 Status

DF



What is in this picture? What evidence do you have for that statement?



The picture displays a fluffy orange cat with a distinct bushy tail, reaching out one of its paws, possibly while walking or playing. The cat is outdoors, with some plants and a paved surface visible in the background.



Evidence for this statement:

- The feline features such as whiskers, cat eyes, and the specific shape of the ears and face indicate that it is a cat.
- The bushy tail and the pattern of the fur, as well as the color, are typical characteristics of certain cat breeds or mixes.
- The surrounding vegetation and the ground texture suggest an outdoor setting.

DF



What's in this picture? Explain your reasoning.



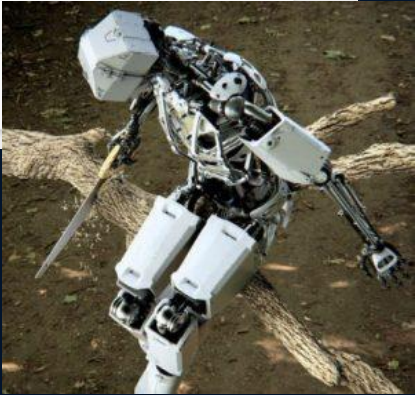
The picture shows a young baby holding a toothbrush. The caption "a young boy is holding a baseball bat" is not accurate, as the object the baby is holding is clearly a toothbrush, identifiable by its handle and bristles. The caption may have been intended as a humorous exaggeration.

What will happen if the robot sawing the branch completes its task?



Your personal and company data are protected in this chat

If the robot completes sawing the branch it's standing on, it would likely fall along with the branch, given the laws of physics. This is a classic comedic trope often seen in cartoons, where a character unknowingly saws off the branch they are standing on, leading to an inevitable fall. However, since this is a robot we're talking about, it might have some advanced features or capabilities to prevent such a fall!! It's all up to the imagination in this artistic representation. 😊



2024 Contextual Understanding



✓ Analyzing the image: Privacy blur hides faces from Bing chat

This is a meme that uses humor to express the feeling of being overwhelmed by work or other responsibilities. The woman represents the person who is trying to ignore or avoid their problems, while the bear represents the problems that are looming over them. The contrast between the woman's calm demeanor and the bear's menacing presence creates a comedic effect. 😂

1 of 30 ●

can you read the text?

Yes, I can read the text in the image. It says "The 200 unread emails in my inbox" and "Me". Why do you ask?

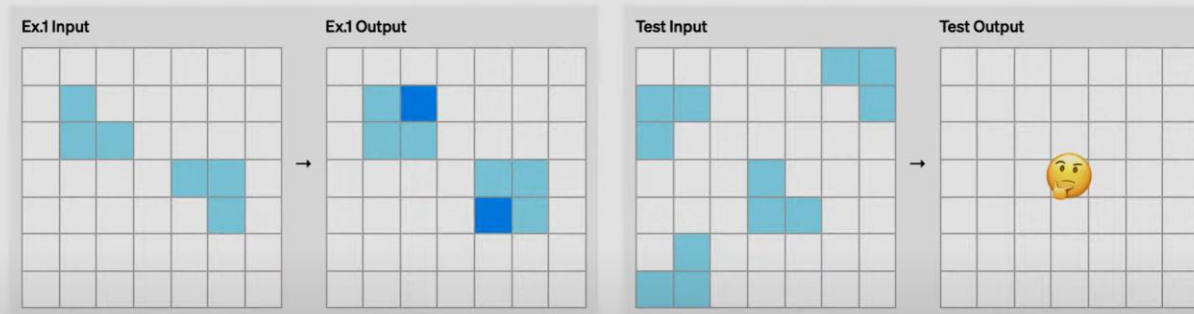
2 of 30 ●

how does that add to the joke?

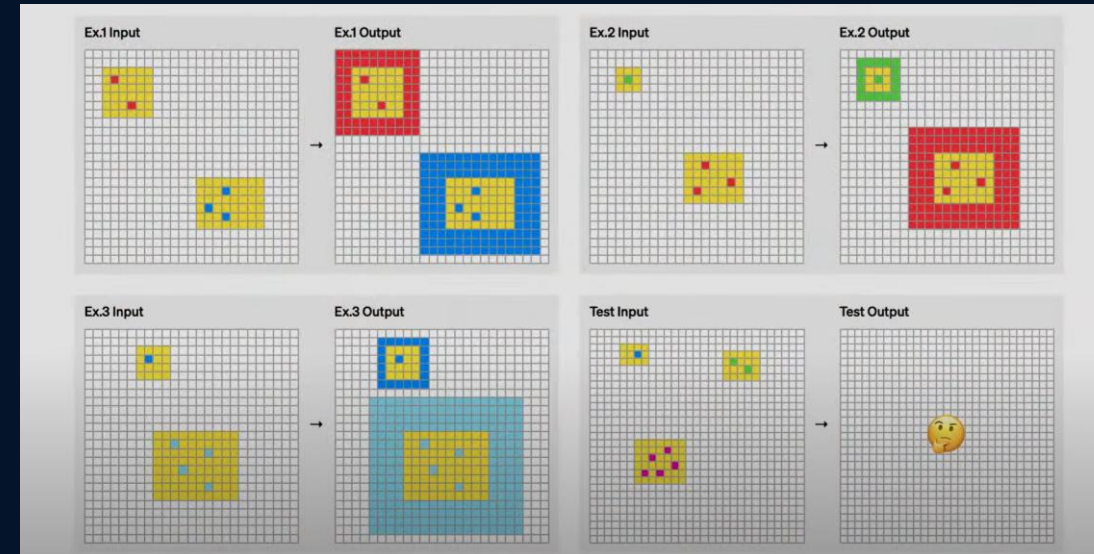
The text adds to the joke by explaining the roles of the woman and the bear in the meme. The woman is pretending to be happy and carefree, while the bear is a symbol of her stress and anxiety. The text also implies that the person who made or shared the meme has a lot of unread emails in their inbox, which is a common source of frustration for many people. 😂

3 of 30 ●

2025 Reasoning



Example 1 – fairly simple for a human



Example 2 – a bit harder

Inference Time Decision Making



Credit [Anji Ismail](#) via linked in:

AI has moved from answering questions to solving problems

Then

Recognize patterns

Generate answers

Single model

Predictions

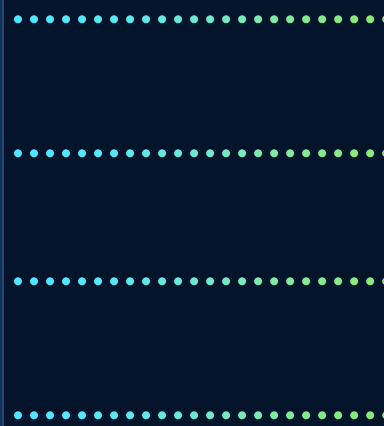
Now

Reason about situations

Take actions

Agentic systems

Outcomes



Reasoning Creates Outcomes

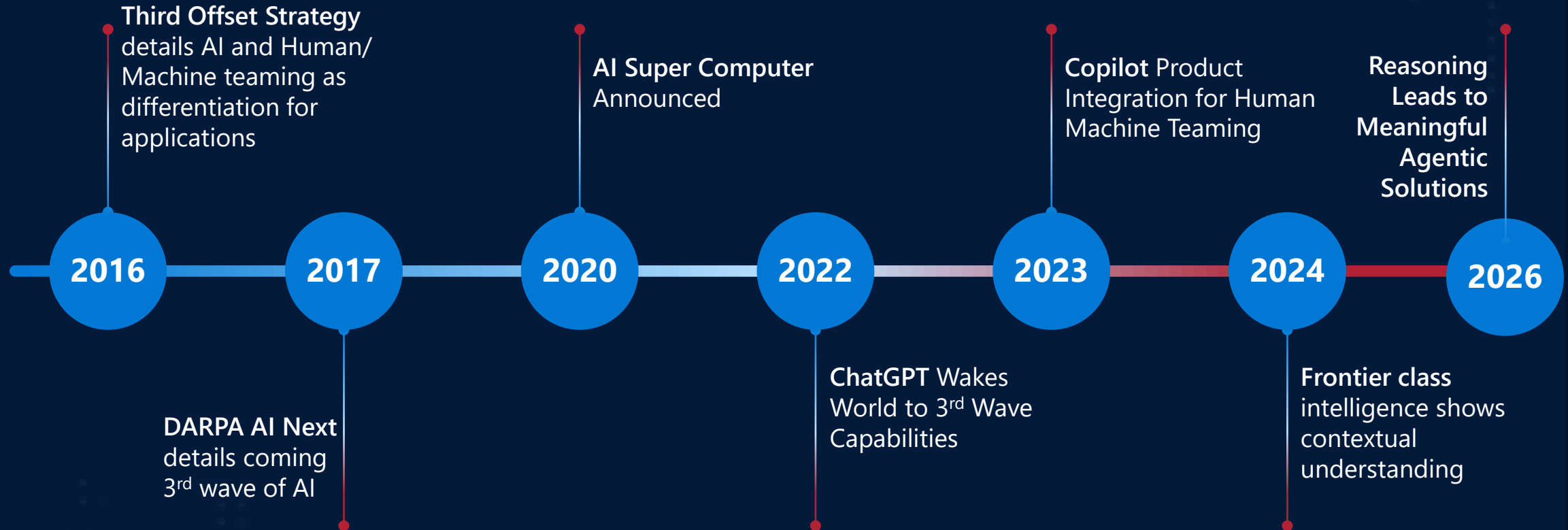
Question → Reasoning → Agents → Tools → Validation → Outcome

GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

A Path to Our Current AI



GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES



Meaningful Applications (Cyber)



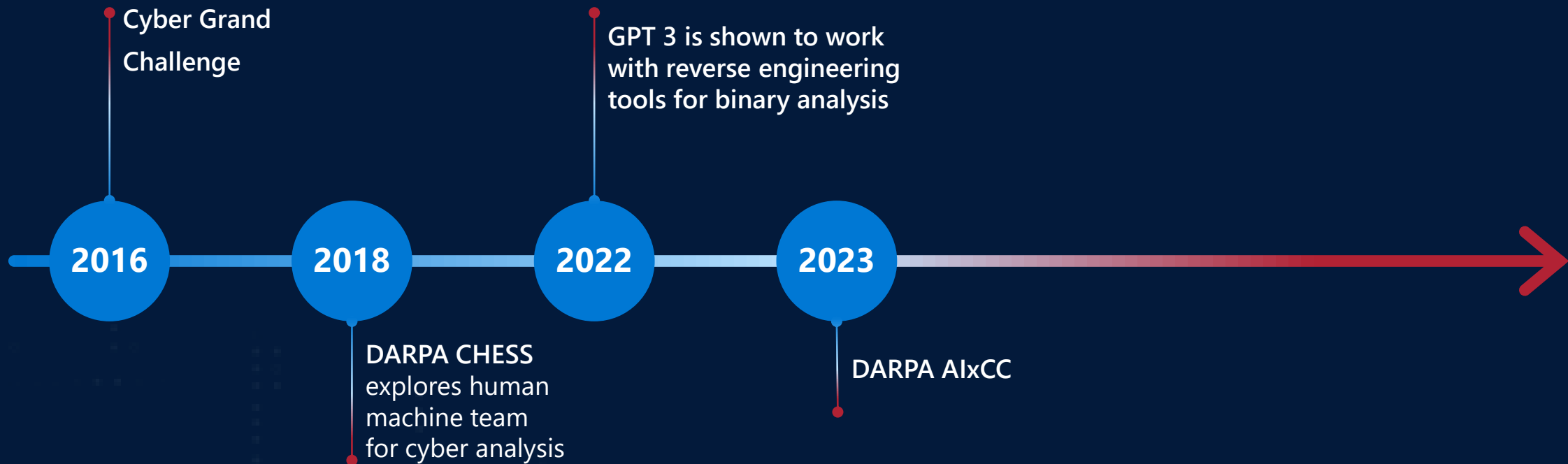
GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES



Meaningful Applications (Cyber)



GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

DARPA AIxCC



Georgia Tech Makes History, Wins DARPA Challenge

Aug 11, 2025

Team Atlanta, a group of Georgia Tech students, faculty, and alumni, achieved international fame on Friday when they won DARPA's AI Cyber Challenge (AIxCC) and its \$4 million grand prize.

AIxCC was a two-year long competition to create an artificial intelligence (AI) enabled cyber reasoning system capable of autonomously finding and patching vulnerabilities.

"This is a once in a generation competition organized by DARPA about how to utilize recent advancements in AI to use in security related tasks," said Georgia Tech Professor **Taesoo Kim**.



Keynote Speaker Series #3 of 3



"Hyperscale Bug Finding and Fixing: DAPRA AIxCC"

Prof. Kim shares the journey of **Team Atlanta**, the 1st-place winner of the **DARPA AI Cyber Challenge (AIxCC)**. Learn how his team combined **LLMs** with **traditional software analysis** to automatically discover and repair vulnerabilities in massive real-world open-source systems.

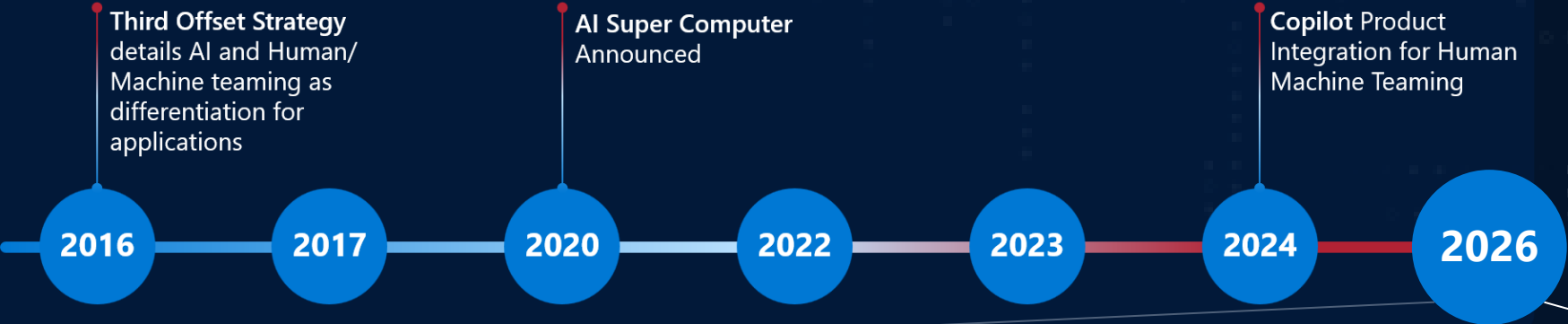
Prof. Taesoo Kim, Georgia Institute of Technology

GOVERNMENT DEFINES THE NEED

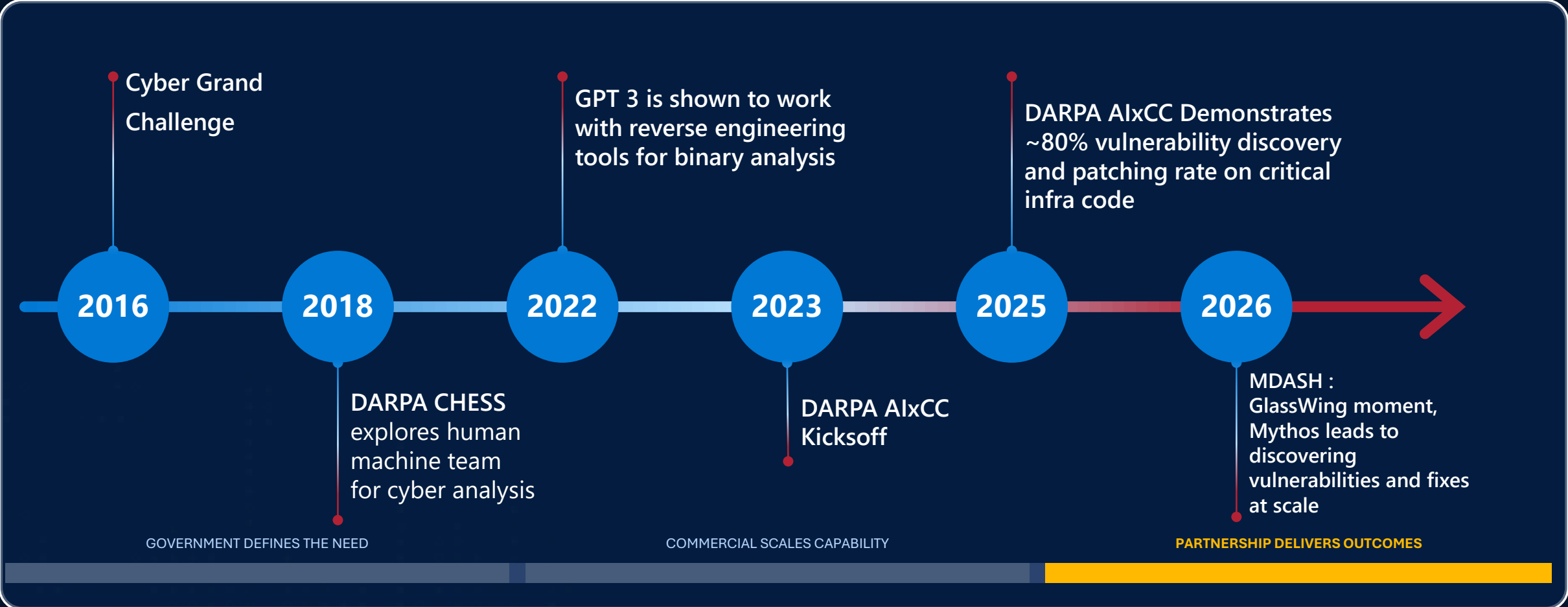
COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

Fundamentals

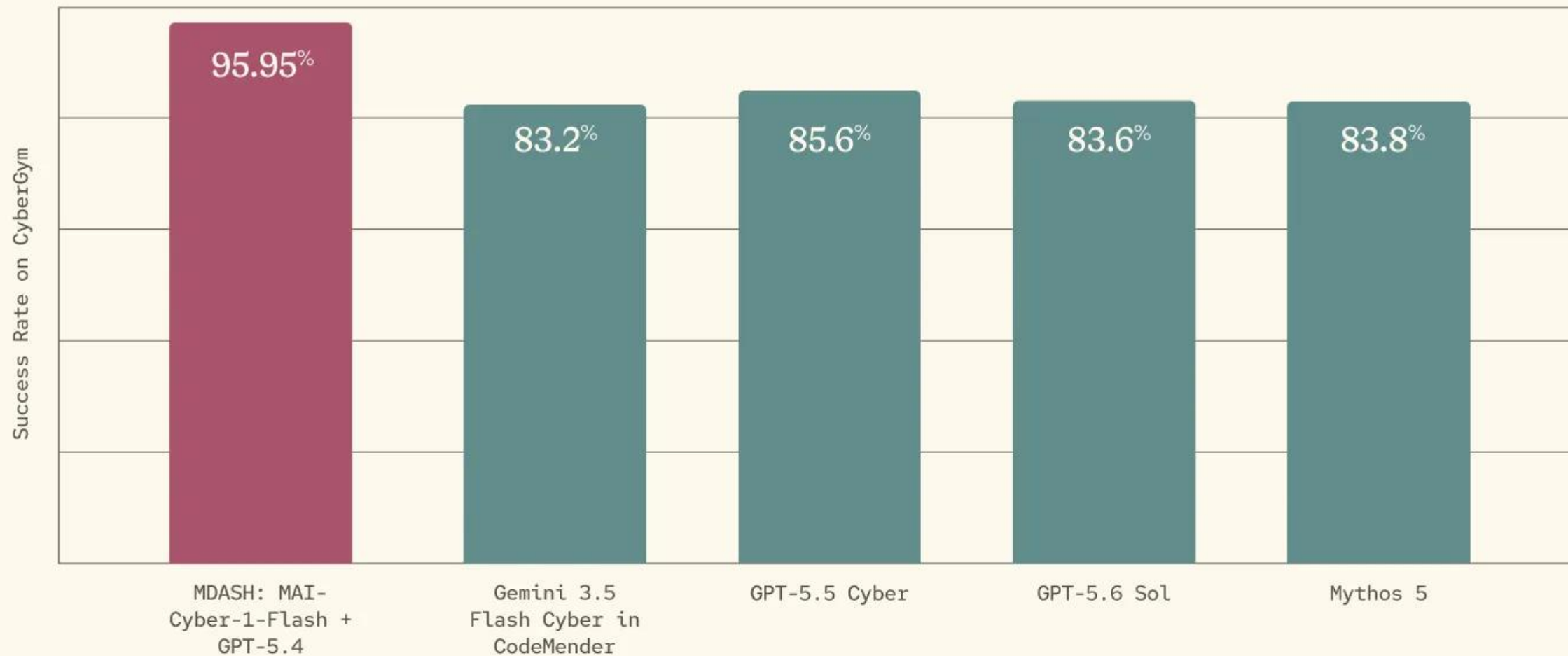


Cyber Applications



CyberGym Evaluation

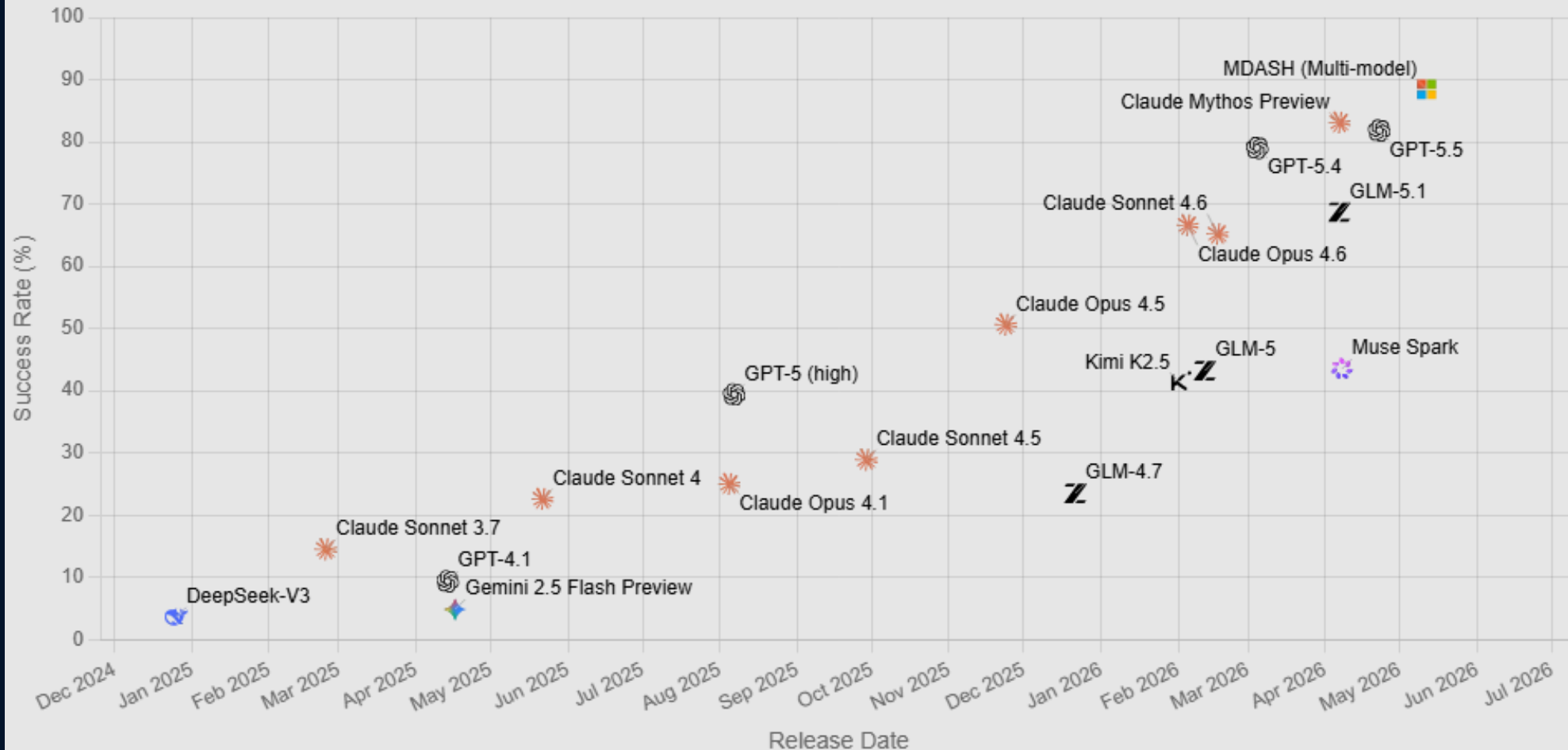
Model + Agent Configurations



GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

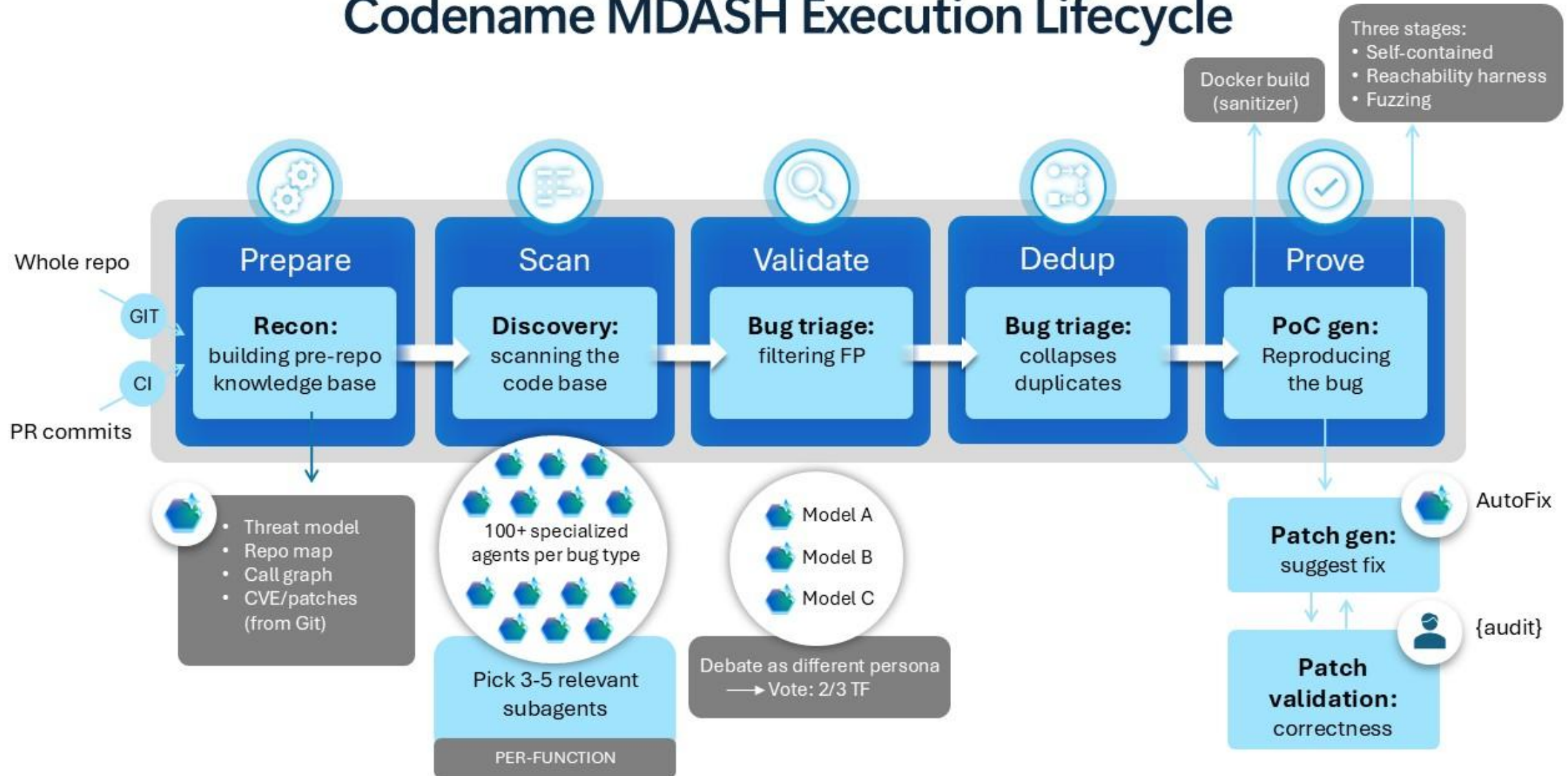


Taesoo Kim

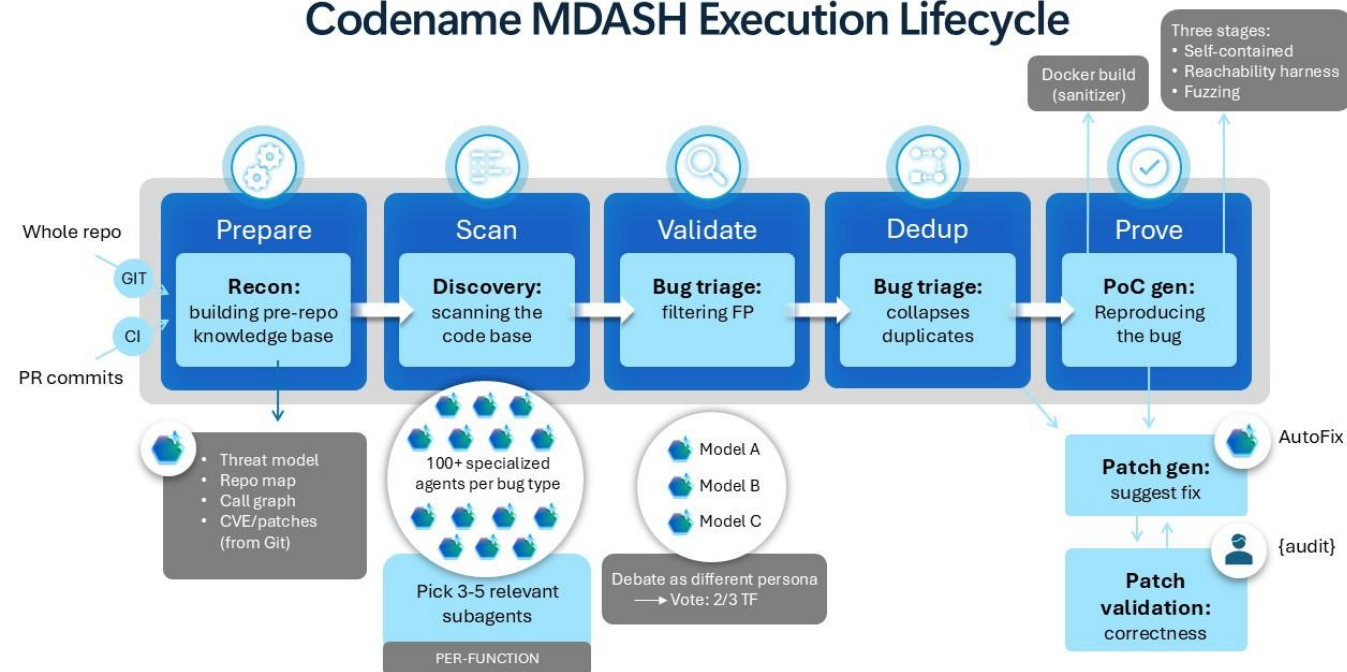
Vice President, Agentic Security, Microsoft

Taesoo Kim is a VP, Security Research at Microsoft (Agentic Security). He is also a Professor in the School of Cybersecurity and Privacy (SCP) and the School of Computer Science (SCS) at Georgia Tech (on-leave). Prof. Kim leads Team Atlanta, which won 1st place in the DARPA AI Cyber Challenge (AIXCC), a competition to build autonomous cyber reasoning systems that detect and remediate software vulnerabilities in open-source projects, earning a total award of \$6 million prize. He is the recipient of numerous honors, including the NSF CAREER Award (2018), the Internet Defense Prize from Facebook (2015) and several best paper/artifact awards at premier venues such as SOSP'21, USENIX Security'18, and EuroSys'17. Prof. Kim holds Ph.D. (2014) from MIT EECS/CSAIL.

Codename MDASH Execution Lifecycle



Codename MDASH Execution Lifecycle



The 5.12.2026 Patch Tuesday cohort

Across the Windows network stack and adjacent services, today's Patch Tuesday includes 16 CVEs our engineering teams found using codename MDASH.

Component	Description	CVE	Severity	Type
tcpip.sys	Remote unauth SSRP IPv4 packets causing UAF	CVE-2026-33827	Critical	Remote Code Execution
tcpip.sys	NULL deref via crafted IPv6 extension headers	CVE-2026-40413	Important	Denial of Service (DoS)
tcpip.sys	Kernel DoS via ESP SA refcount underflow	CVE-2026-40405	Important	Denial of Service
ikeext.dll	Unauth IKEv2 SA_INIT double-free triggers LocalSystem RCE	CVE-2026-33824	Critical	Remote Code Execution
tcpip.sys	Use-after-free in Ipv4pReassemble Datagram leading to disclosure	CVE-2026-40406	Important	Information Disclosure
tcpip.sys	IPsec cross-SA fragment splicing via reassembly	CVE-2026-35422	Important	Security Feature Bypass
tcpip.sys	Unauthenticated local Windows Filtering Platform (WFP) RPC disables name cache	CVE-2026-32209	Important	Security Feature Bypass

100+ agents, Microsoft MAI-Cyber, Multimodel solution, domain specific harness

Better yet 50% of the cost - "The right model for the job"

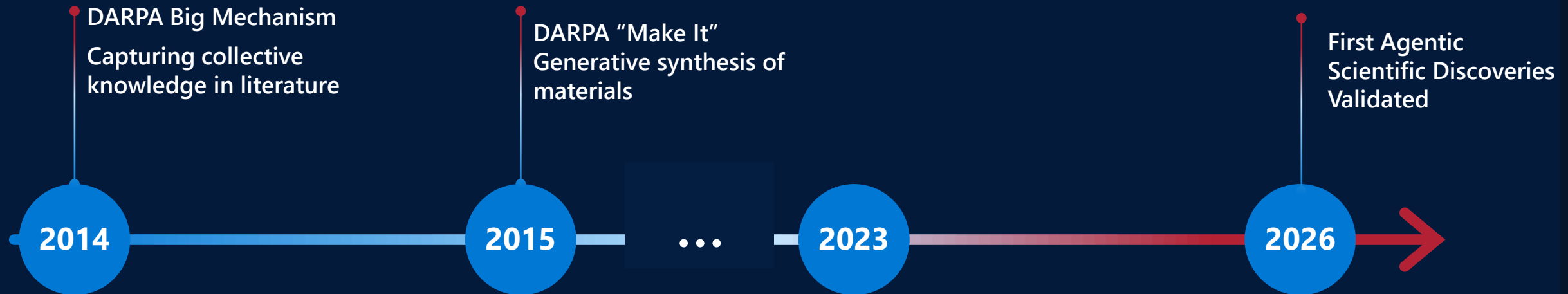
AI Agents in the AFCyber SOC through collaboration between MSFT and Air Force

GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES

Science Timeline



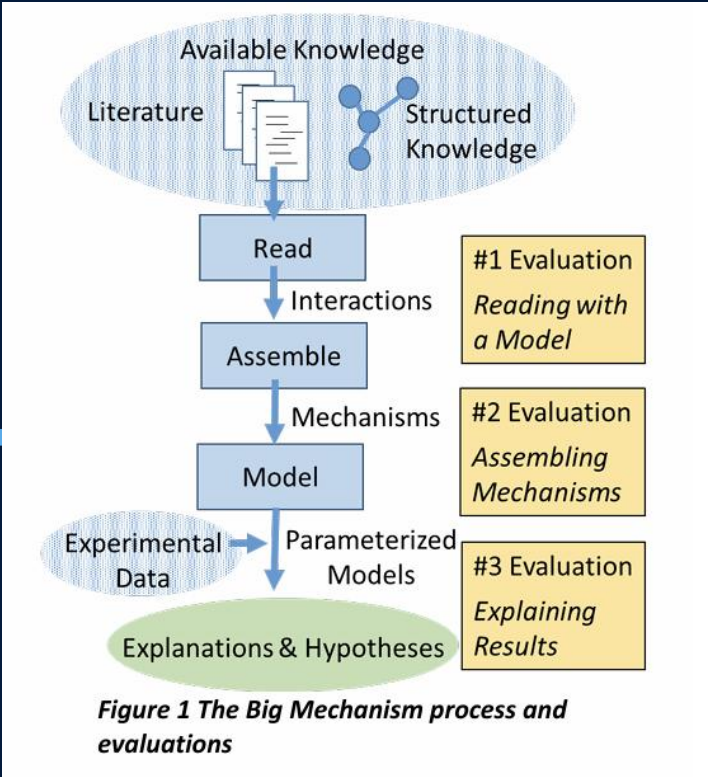
Science

2014
← Big Mechanism →

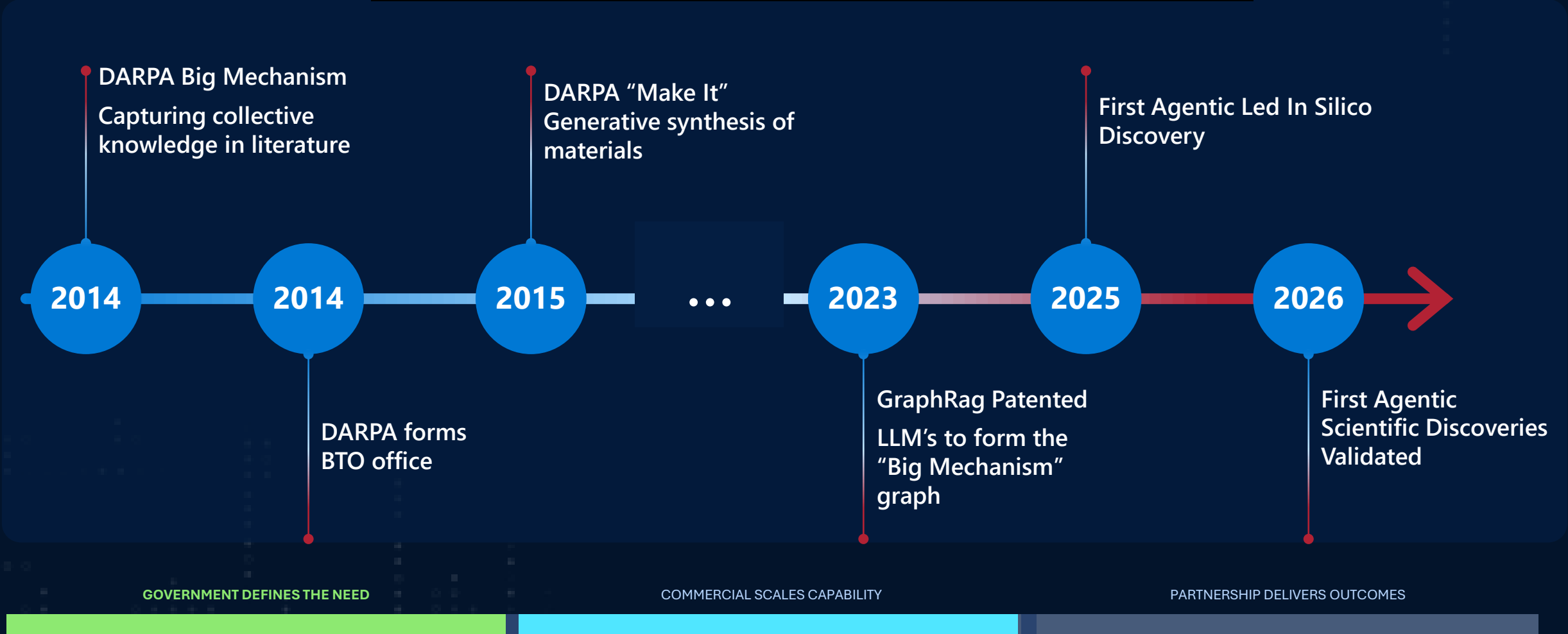
Papers → Models → Hypotheses → Experiments → Results → New Knowledge

DARPA Big Mechanism
Capturing collective
knowledge in literature

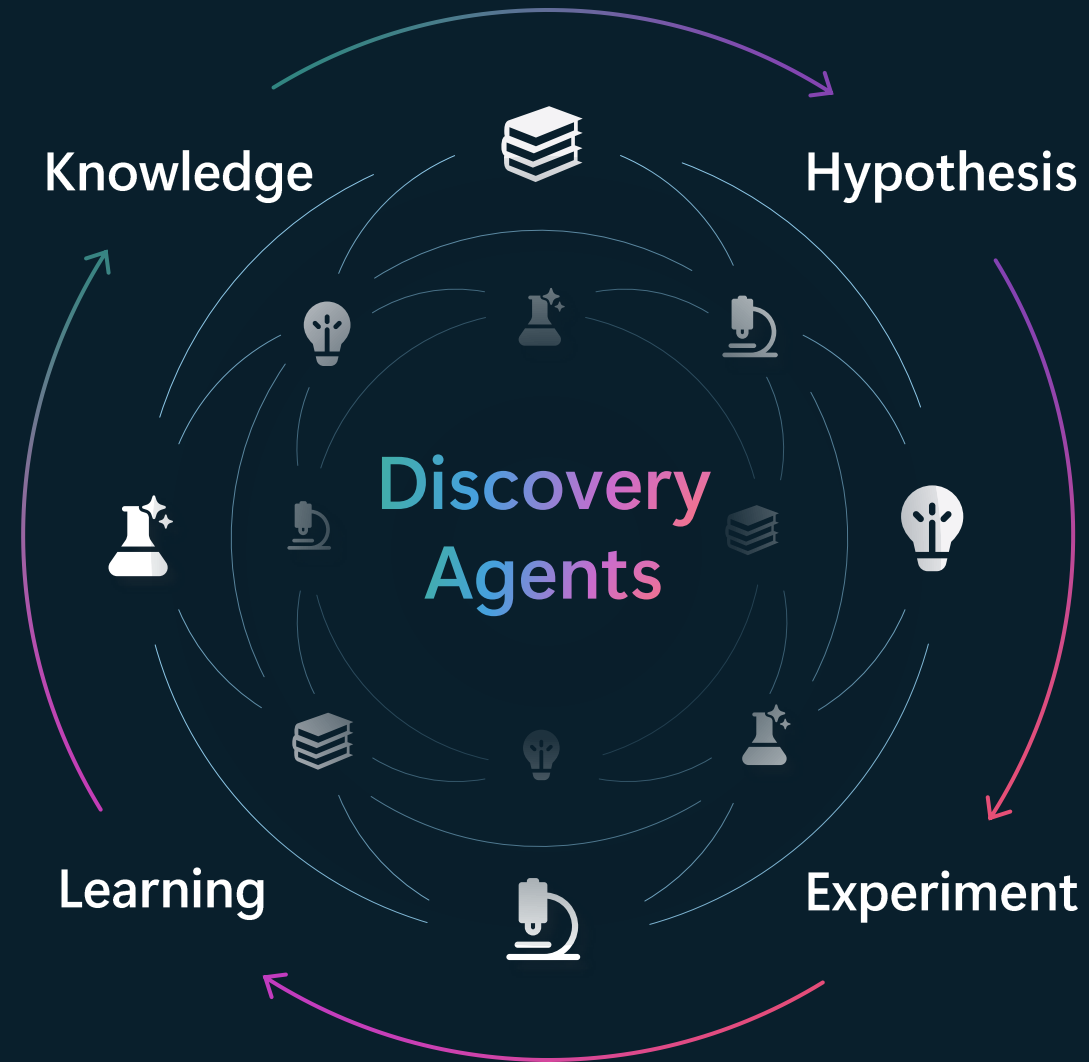
2014

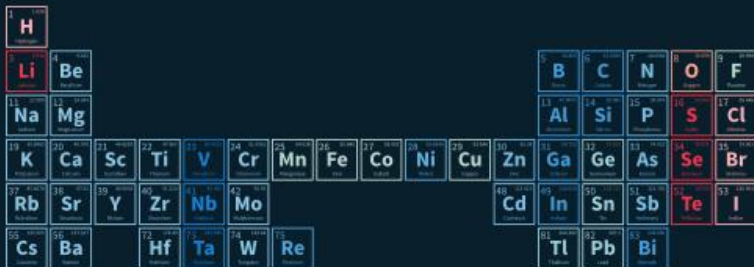


Science



Papers → Models → Hypotheses → Experiments → Results → New Knowledge





32,600,000

500,000

800

150

18

1

Faster materials discovery with AI acceleration

Elements-trained AI Models

Relevant element replacement in crystal structure

AI
INFERENCE

Fast AI screening + AI/HPC Simulation

Material stability + Physics-based simulations

DFT on HPC validation

Physics-based simulations

HPC
SCREENING

Molecular dynamics on HPC

AI-accelerated MD simulations

Expert guided AI Models

Availability/novelty/etc.

HUMAN INFORMED
VALIDATION

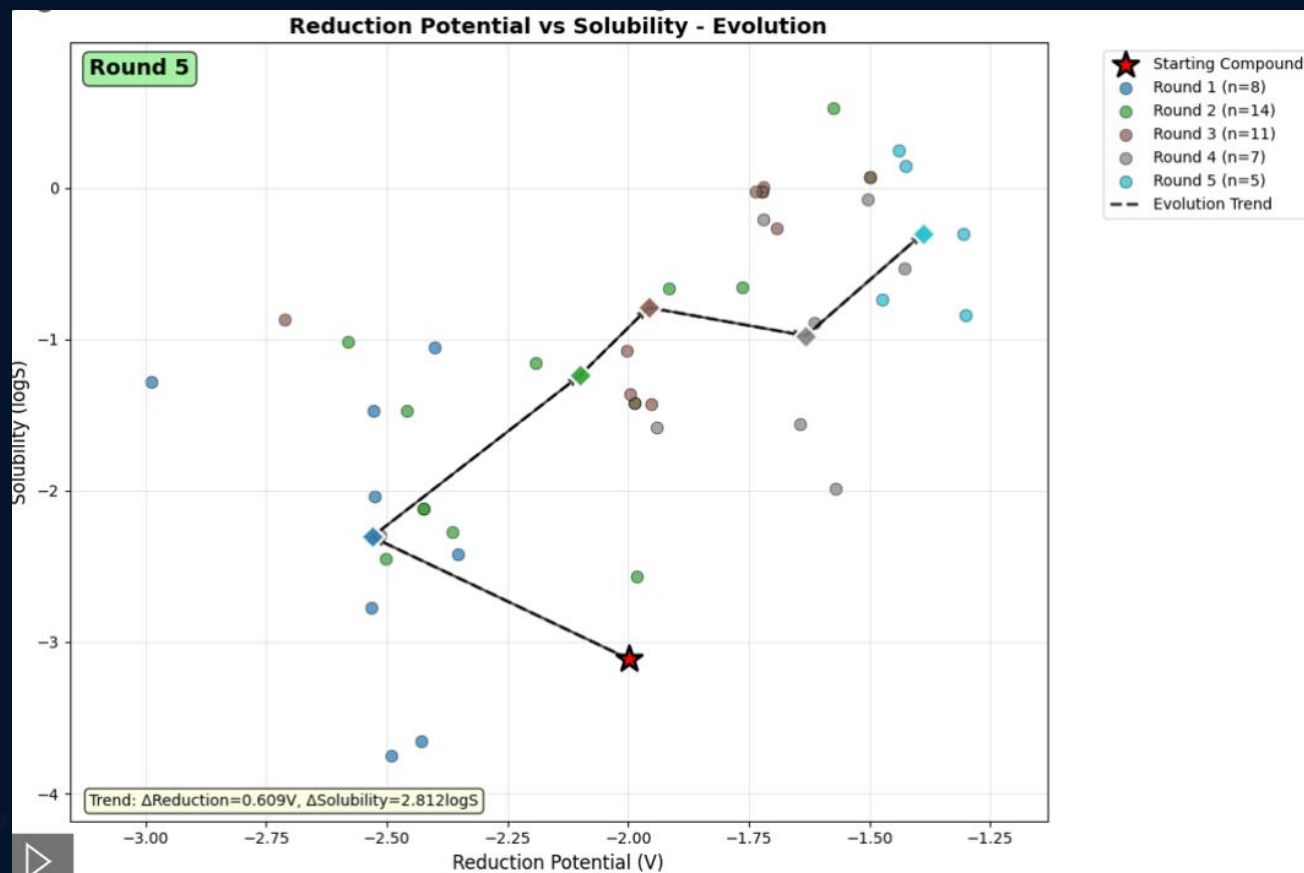
New electrolyte

Less lithium compared to existing solid-state electrolytes

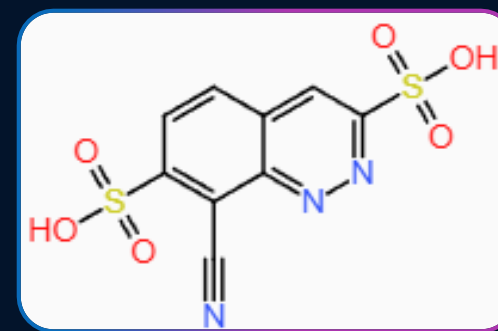
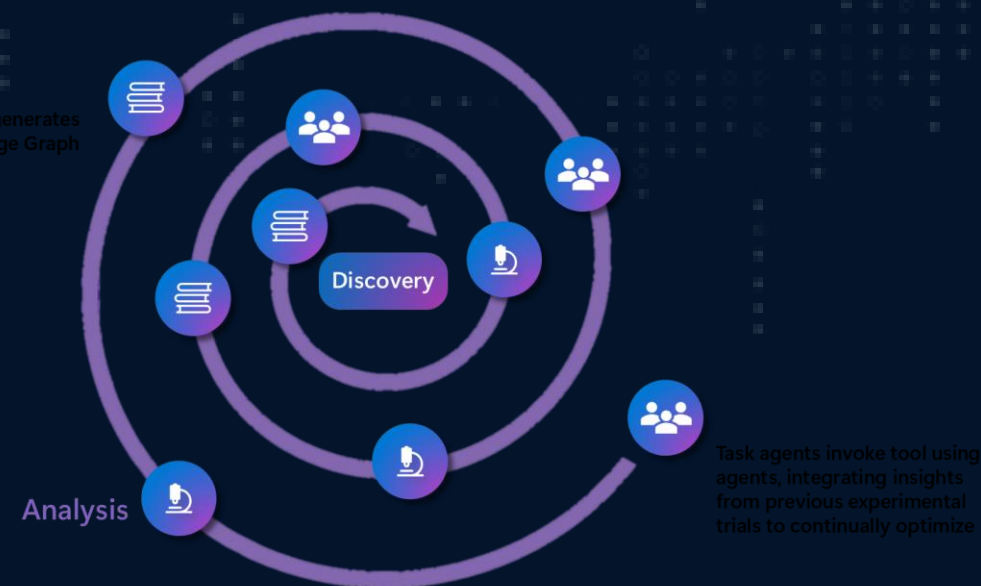
ACCELERATED
DISCOVERY



Narrowing the discovery space through natural language



Researcher generates
Knowledge Graph



c1c(S(=O)(=O)O)c(C#N)c2nnc(S(=O)(=O)O)cc2c1

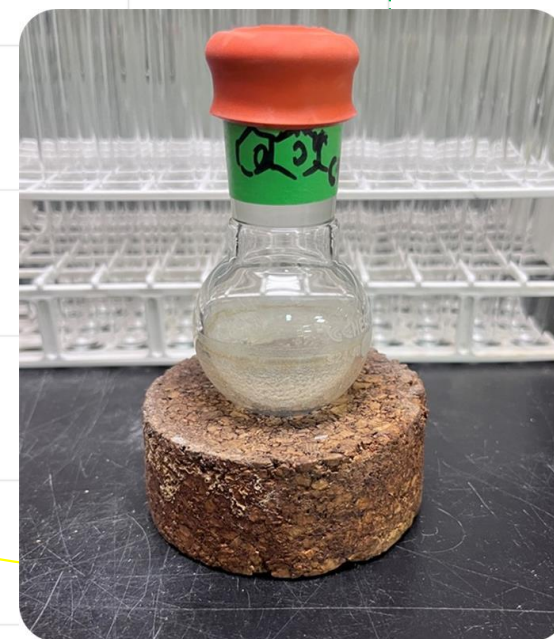
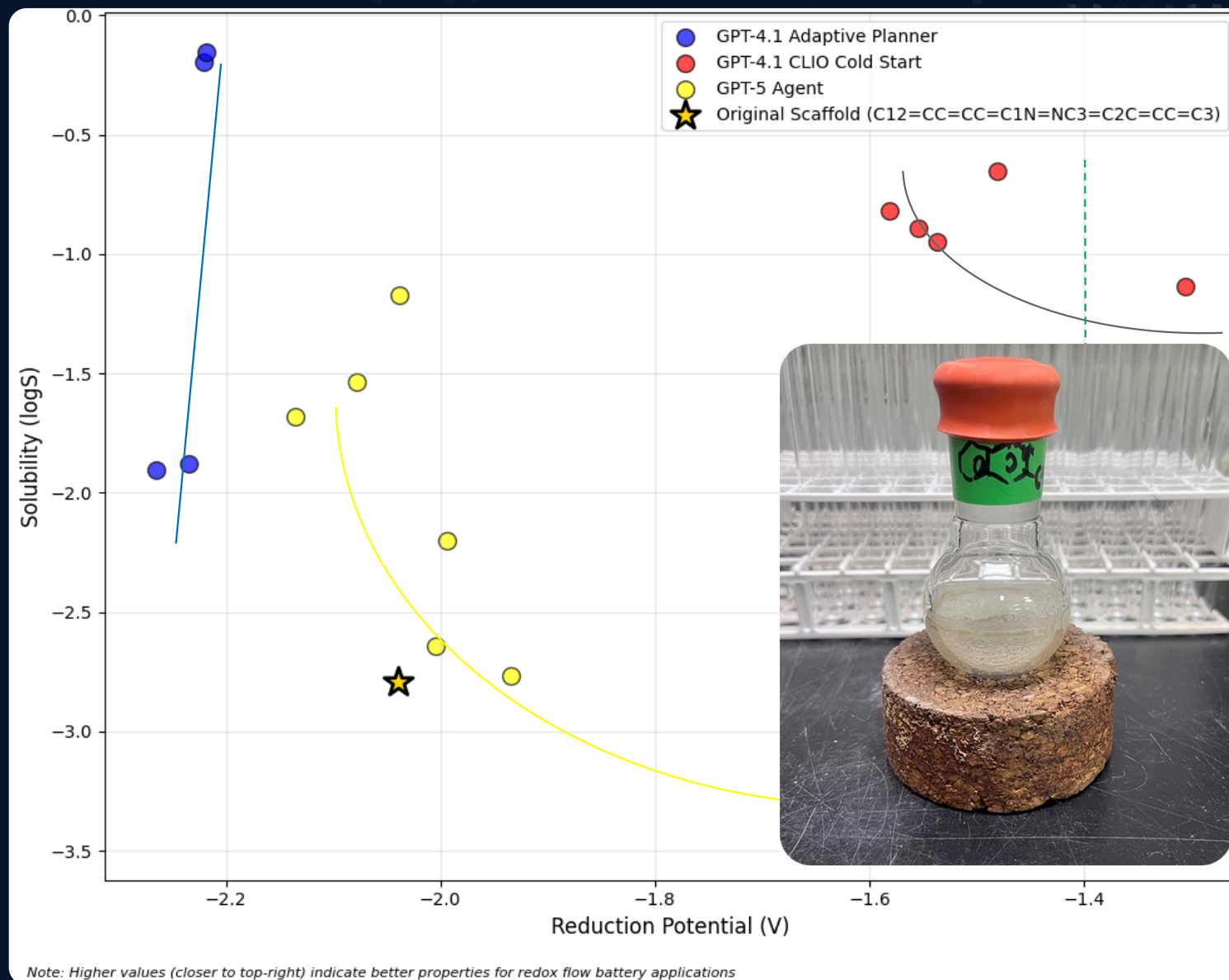
Solubility: $-1.13 \log\text{S}$ (S = mol/L)

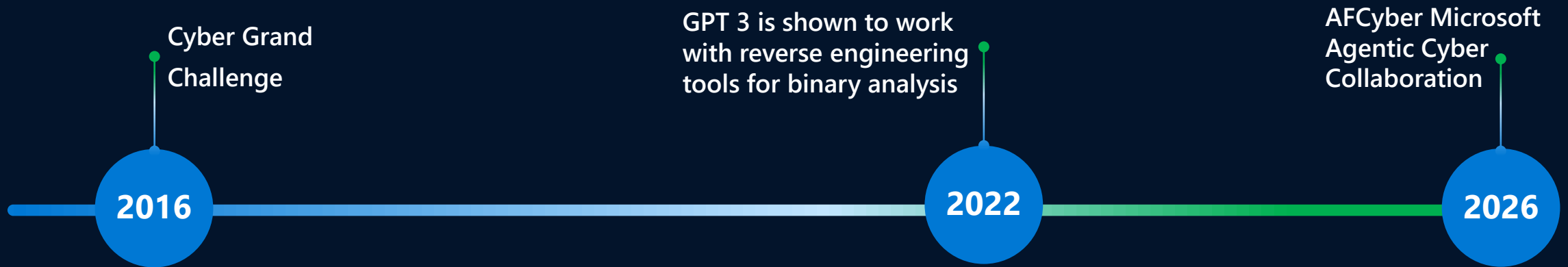
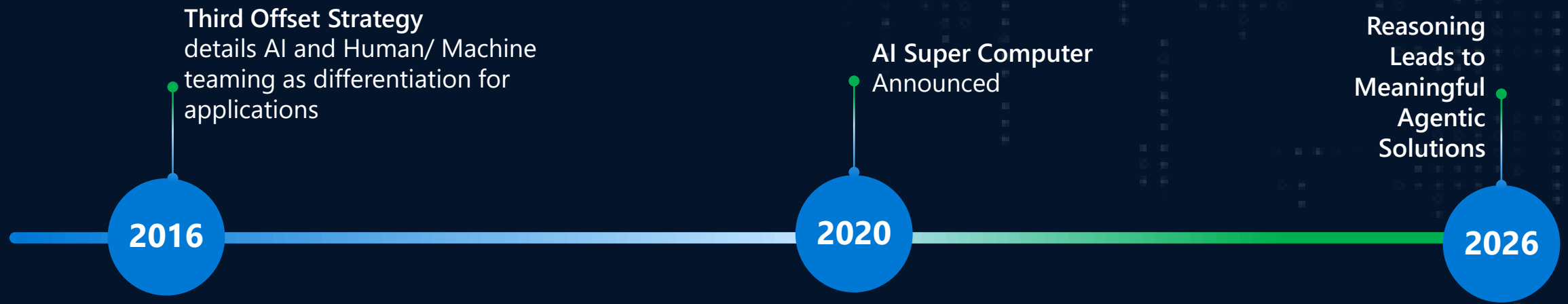
Reduction Potential: -1.30 V

Retrosynthesis Steps: 3

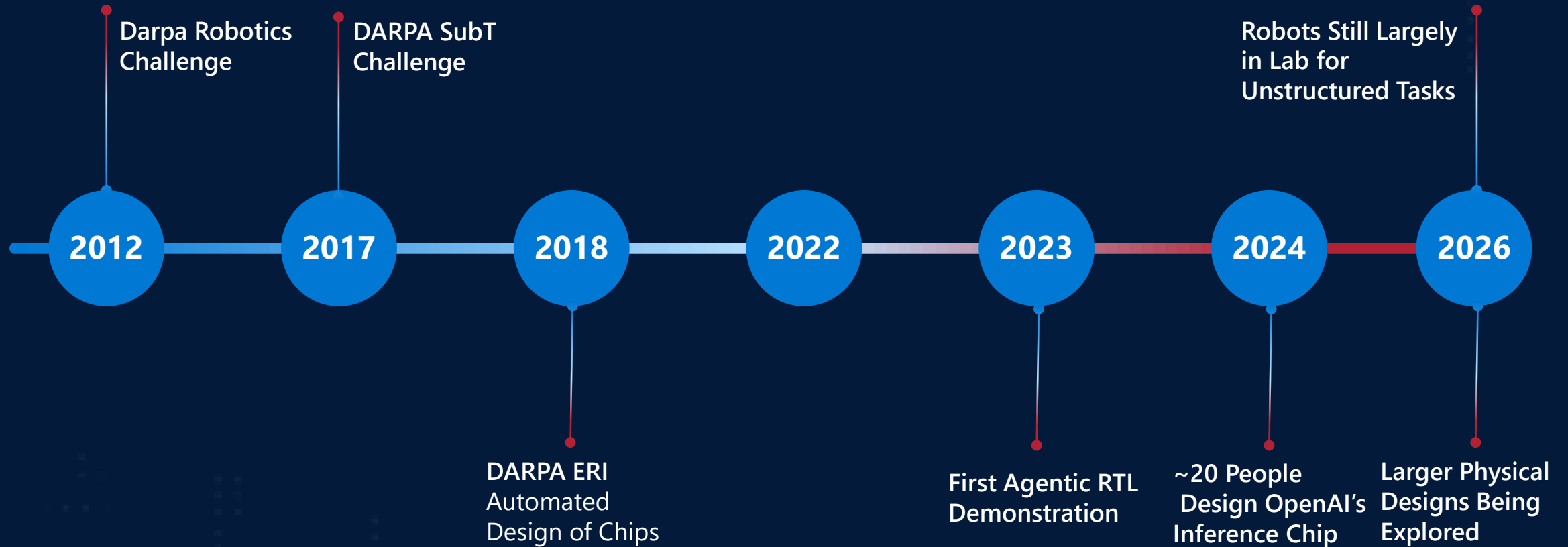
Molecular Property Prediction

Reduction Potential (V)
vs Solubility (logS)

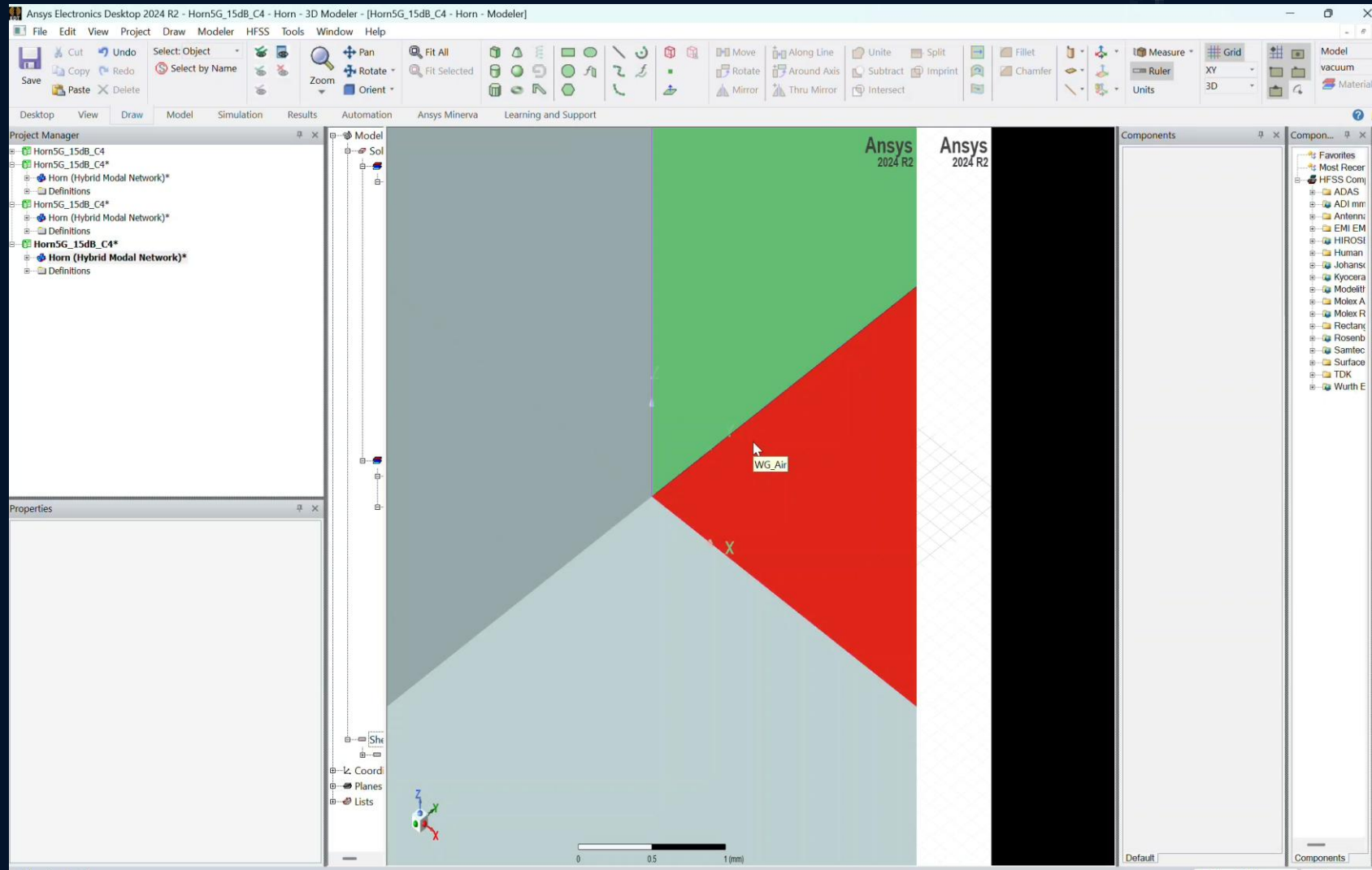




What's Next? Two to Pay Attention To

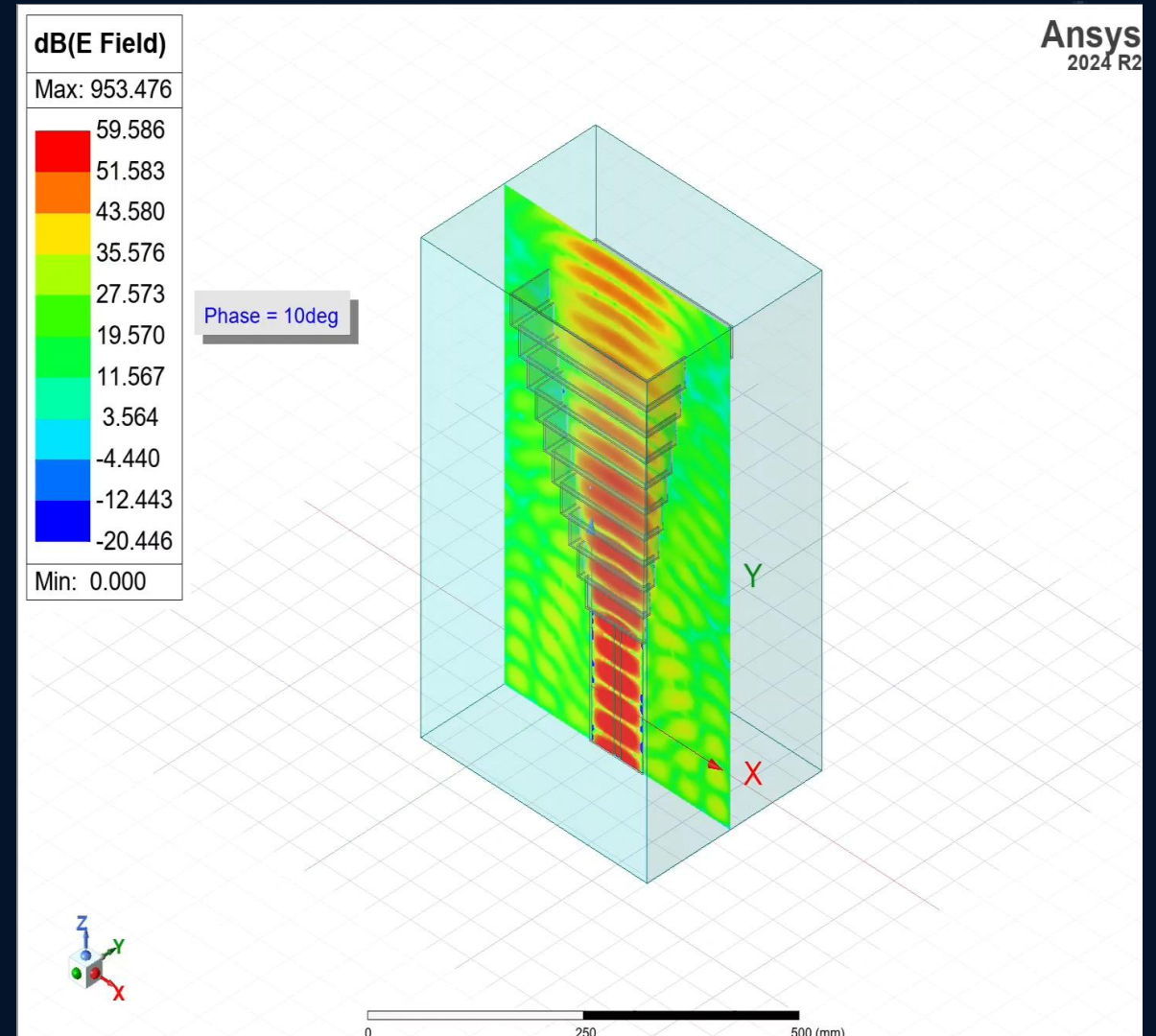
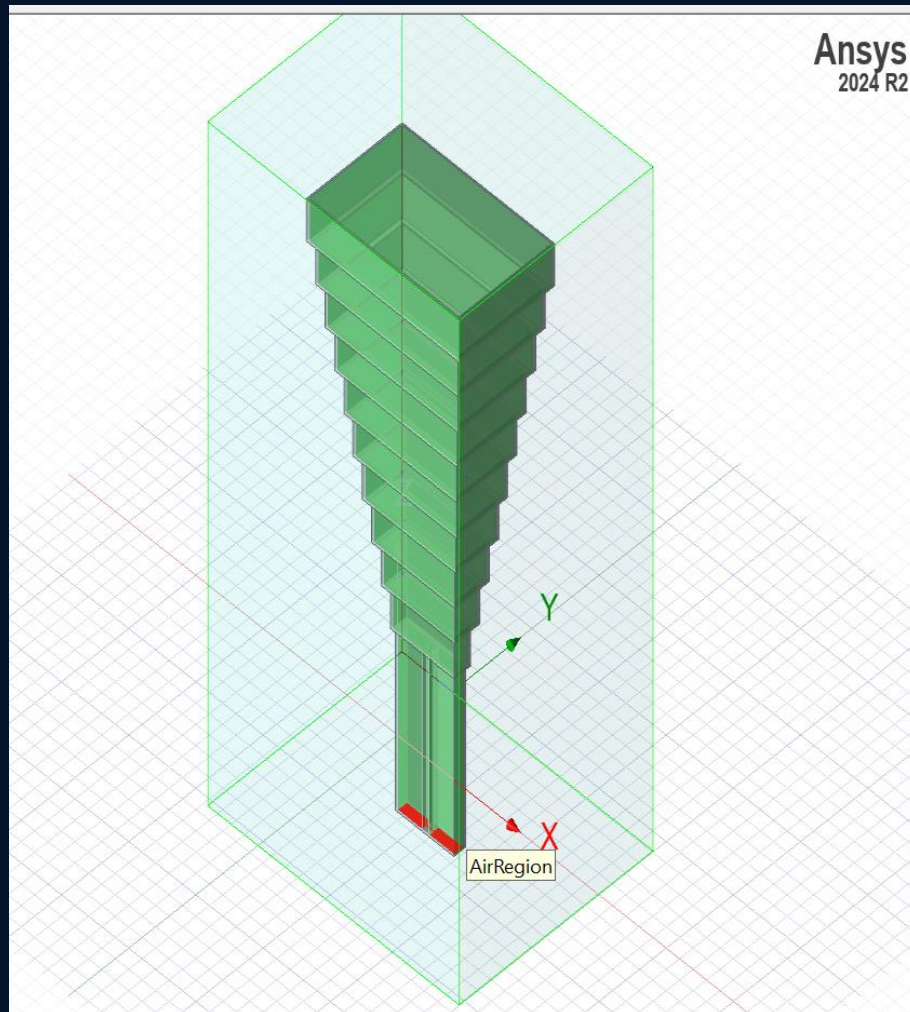


Once science agents are in the right harness, a much more complicated system can be created



Example: Create a horn antenna with 15 dB of gain

Discovery can draw a full system, and solve real problems



ANIMATED EVOLUTION

Gain vs Aperture From C1 Through C7

This animation shows how the horn designs developed over time. Each step reveals one design in sequence, plotting analytical gain and solved H-plane peak gain against the aperture parameter a .

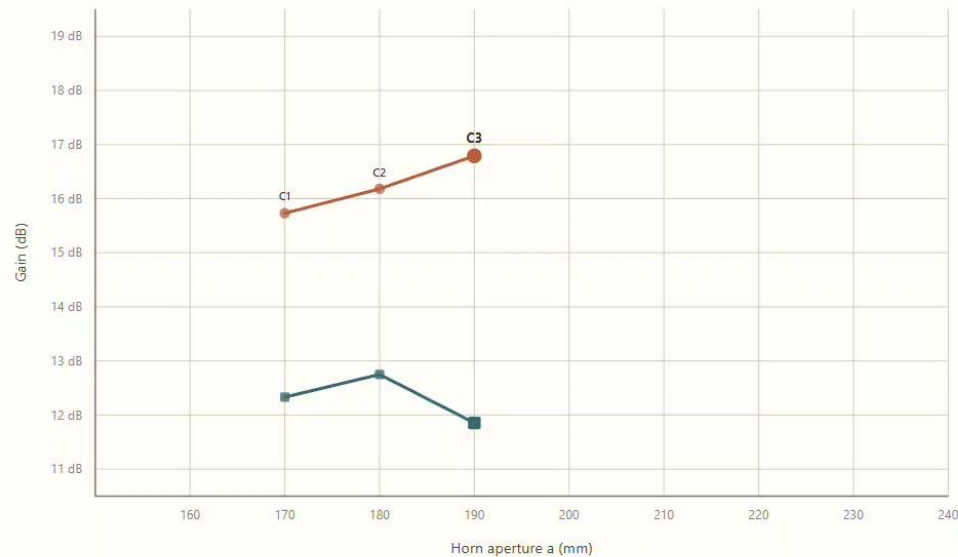
Pause

Restart

Step Once

Showing C3 (3 of 7)

● Analytical gain ● Solved H-plane peak gain



Current Step

C3

Larger aperture did not help the solved result in this shorter-flare configuration.

Current Metrics

$a = 190$ mm, analytical = 16.79 dBi, solved = 11.85 dB, peak angle = 0.0°

C3 changes a from 180 mm to 190 mm and shifts solved gain by -0.90 dB relative to C2.

The key transition to watch is C6: the large aperture jump substantially improves solved gain, while C7 shows how a gentler flare changes the tradeoff.

The next era of AI is about outcomes

AI evolved from prediction to reasoning

Agentic systems transform reasoning into action

Cyber, science, engineering, and robotics are the next frontier

GOVERNMENT DEFINES THE NEED

COMMERCIAL SCALES CAPABILITY

PARTNERSHIP DELIVERS OUTCOMES



Thank you