Airlines on Autopilot:
From AI to Intelligent Agents

Michael Wu, PhD (@mich8elwu)
chief AI strategist @ PROS

2025.10.07

WORLD AVIATION FESTIVAL

PROS

Michael Wu, PhD (@mich8elwu)
chief AI strategist @ PROS

2025.10.07

5 stages of GenAI

AGI

innovator

agent

problem solver (LRM)

chatbot (LLM)

OpenAI

language model:

attempting to predict the next word w given ***any*** sequence s by modeling

user prompt $= s$ → $P(w|s)$

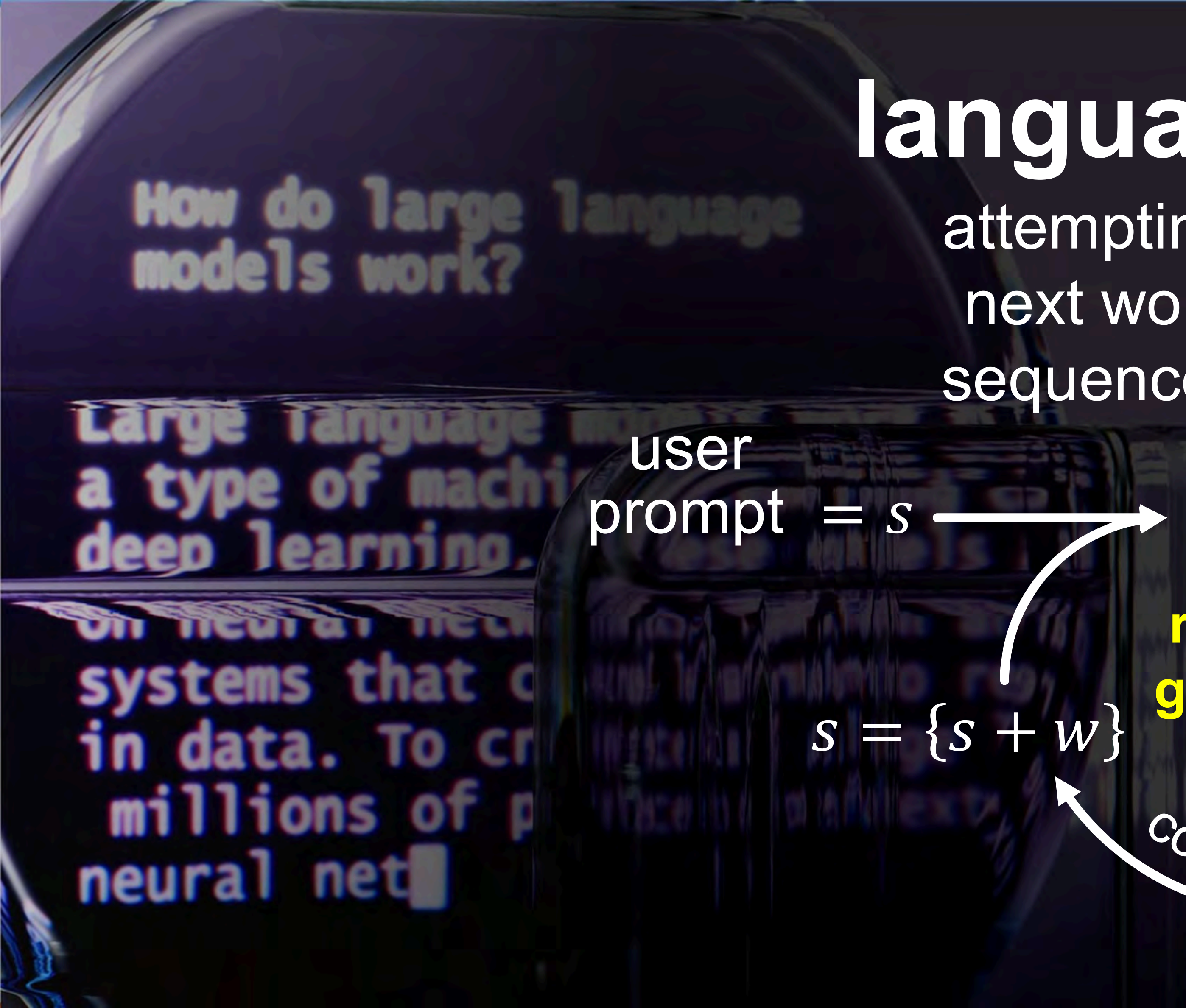
sample

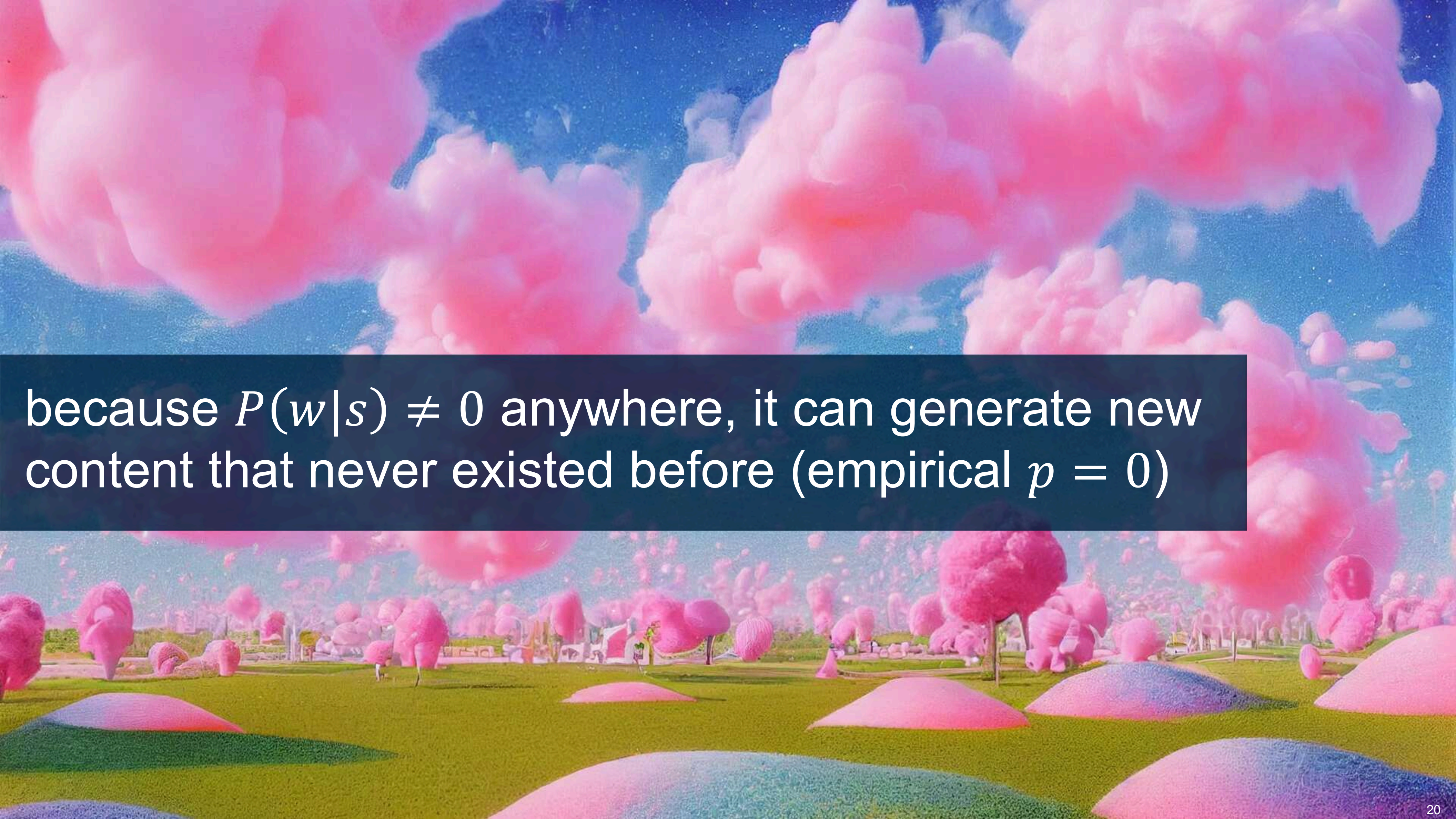
w

concatenate

$s = \{s + w\}$

recursive generative loop





because $P(w|s) \neq 0$ anywhere, it can generate new content that never existed before (empirical $p = 0$)

5 stages of GenAI

AGI

innovator

agent

problem solver (LRM)

chatbot (LLM)

OpenAI


reasoning is fundamental
to all problem solving


chain of
thought
prompting
improves LLM
performance

Chain of Thought (CoT) Prompting Example

prompt:

a farmer has 10 sheep; all but 3 run away, and so he got 2 more; how many sheep does he have now?

LLM response:

9 sheep

INCORRECT

Chain of Thought (CoT) Prompting Example

prompt:

a farmer has 10 sheep; all but 3 run away, and so he got 2 more; how many sheep does he have now?

LLM response:

9 sheep

INCORRECT

prompt:

a farmer has 10 sheep; all but 3 run away, and so he got 2 more; how many sheep does he have now?
please think step by step

LLM response:

1. The farmer starts with 10 sheep
2. The phrase "all but 3 run away" means that 3 sheep did NOT run away
3. So, the farmer still has 3 sheep left
4. he got 2 more, so now he has 3 + 2 = 5 sheep

final answer: 5 sheep

4 “thinking” steps for this simple prompt

CORRECT

why would CoT (simply asking the LLM to think step by step), without any knowledge enhancement (further training or RAG) improve LLM's accuracy?

the effect of CoT is like writing long and detail prompts, except you don't need to write it yourself

user prompt $= s$

longer and more detail prompts → more precise and accurate response

$P(w|s)$

sample

w

concatenate

$s = \{s + w\}$

recursive generative loop

CoT must generate many more tokens as output to explain the steps

if accuracy of a single CoT step is 99%

many real-world problems may involve *tens* of steps

challenging scientific problems may require *hundreds* of steps

accuracy of entire CoT

after 20 steps:

$$(99\%)^{20} = 81.8\%$$

after 50 steps:

$$(99\%)^{50} = 60.5\%$$

after 100 steps:

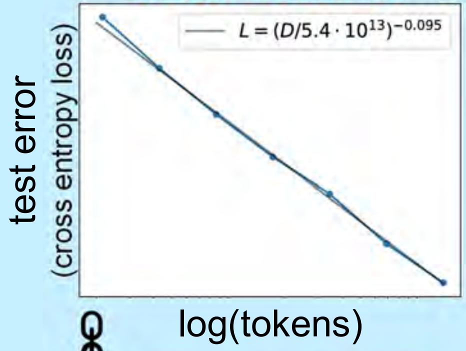
$$(99\%)^{100} = 36.6\%$$

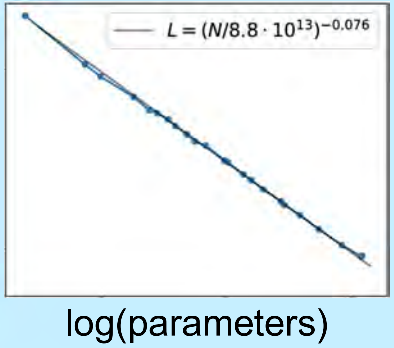
deep reasoning (long CoT) require ~100% LLM accuracy

How to Improve LLM Accuracy? – LLM Scaling Laws

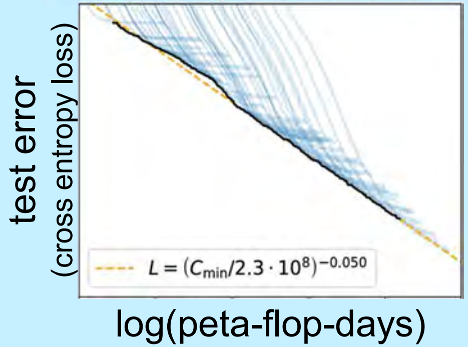
training (pre-training, fitting, etc.) → **LLM** $P(w|s)$ → **inference** (testing, scoring, etc.)

data volume **model size**





compute time



exponential increase in training resources (10x) → ***linear*** improvement in accuracy (~10%)

twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD

How to Improve LLM Accuracy? – LLM Scaling Laws

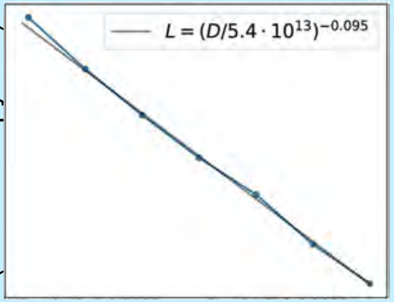
training
(pre-training, fitting, etc.)

→ **LLM** $P(w|s)$ →

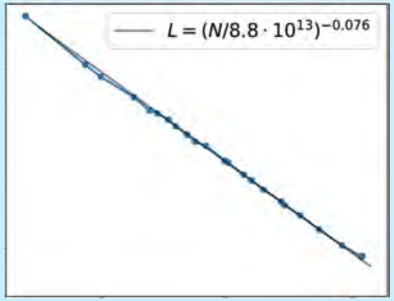
inference
(testing, scoring, etc.)

data volume **model size**

test error
(cross entropy loss)



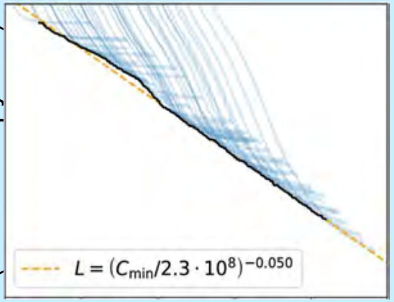
log(tokens)



log(parameters)

compute time

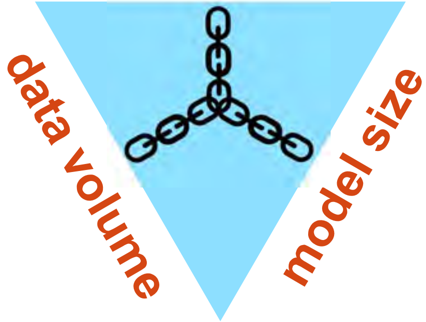
test error
(cross entropy loss)



log(peta-flop-days)

exponential increase in training resources (10x) → ***linear*** improvement in accuracy (~10%)

compute time



we are out of publicly accessible data

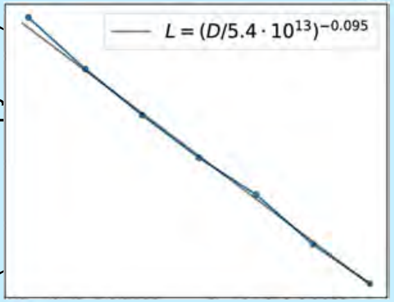
How to Improve LLM Accuracy? – LLM Scaling Laws

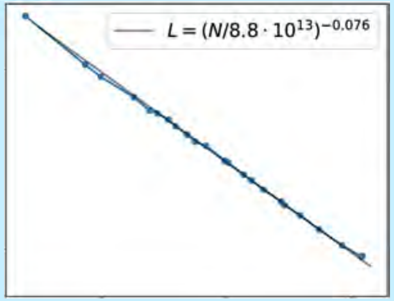
training
(pre-training, fitting, etc.)

→ **LLM** $P(w|s)$ →

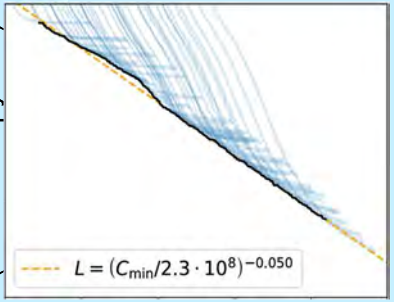
inference
(testing, scoring, etc.)

data volume **model size**

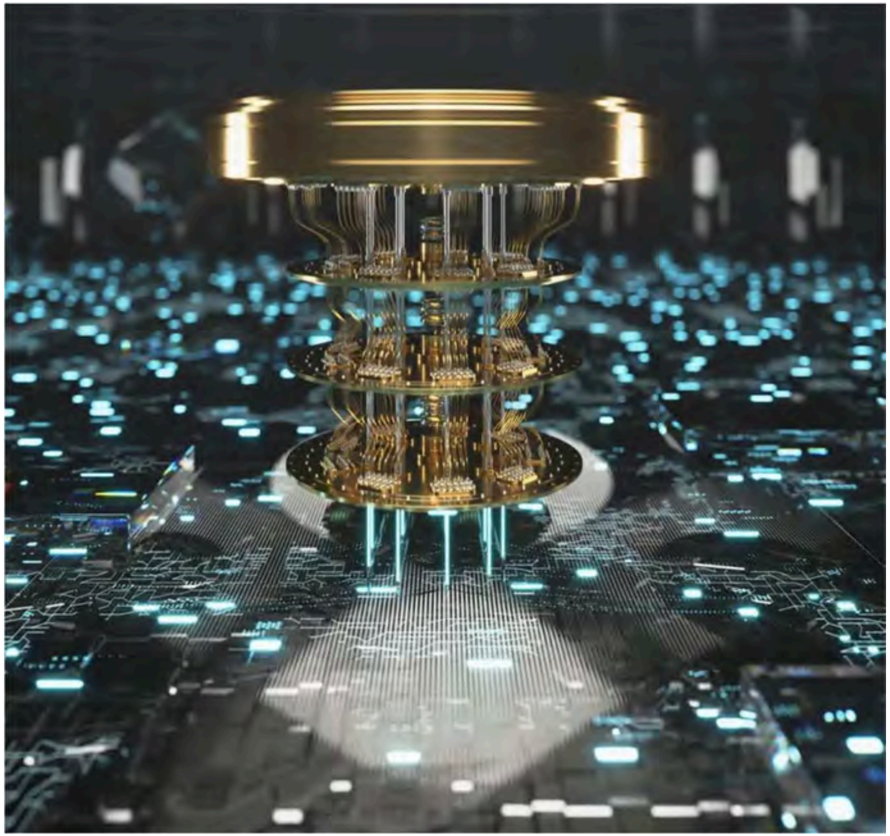




compute time



exponential
increase in
training resources
(10x) → ***linear***
improvement in
accuracy (~10%)



twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD

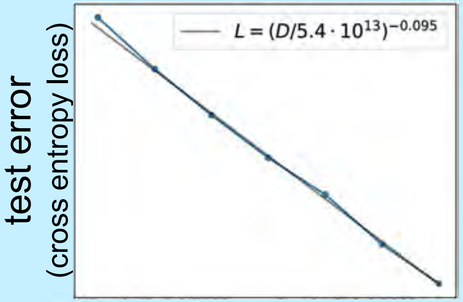
How to Improve LLM Accuracy? – LLM Scaling Laws

training
(pre-training, fitting, etc.)

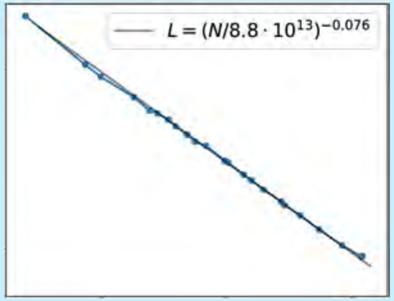
→ **LLM** $P(w|s)$ →

inference
(testing, scoring, etc.)

data volume **model size**

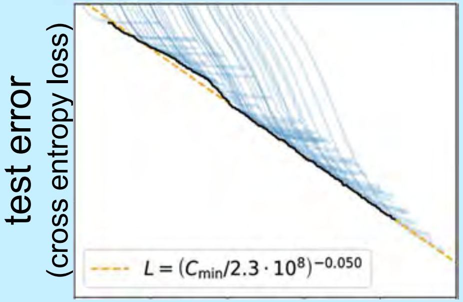


log(tokens)



log(parameters)

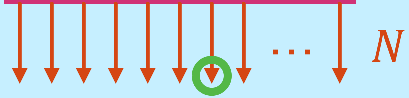
compute time



log(peta-flop-days)

exponential increase in training resources (10x) → ***linear*** improvement in accuracy (~10%)

paralleled generation



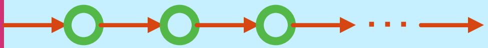
how to generate

- **temperature > 0**
- top p / top k
- min p
- beam search

how to pick the best

- human judge
- reward function
- verifier
- **LLM judge**

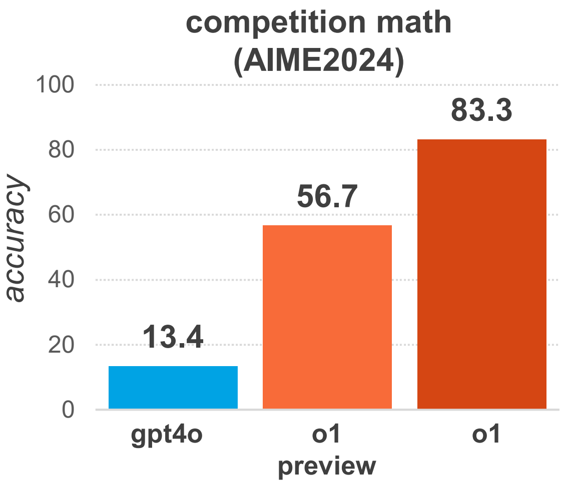
sequential refinement

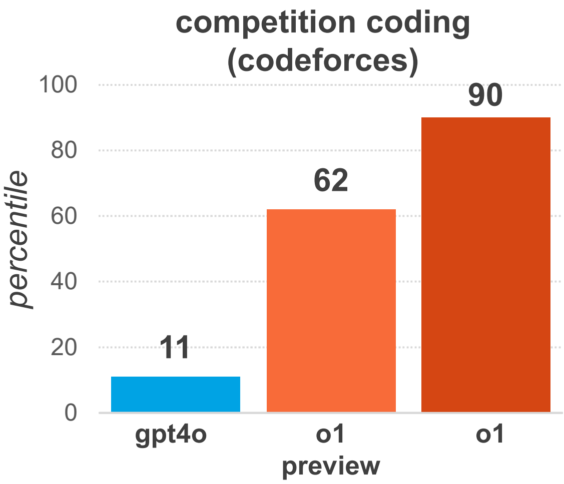


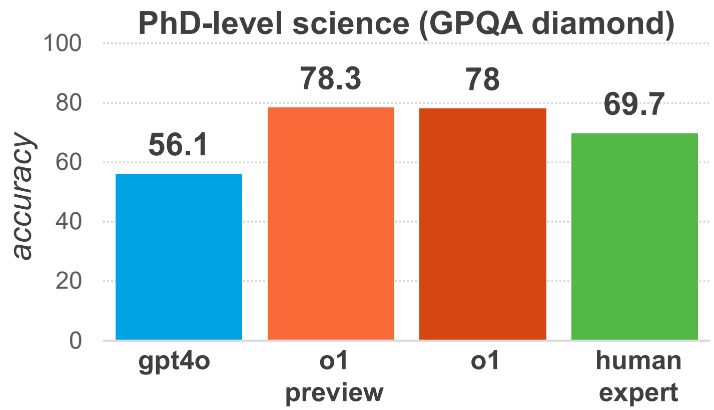
"thinking" behavior

scaling inference time "thinking" requires big working memory (long context window)

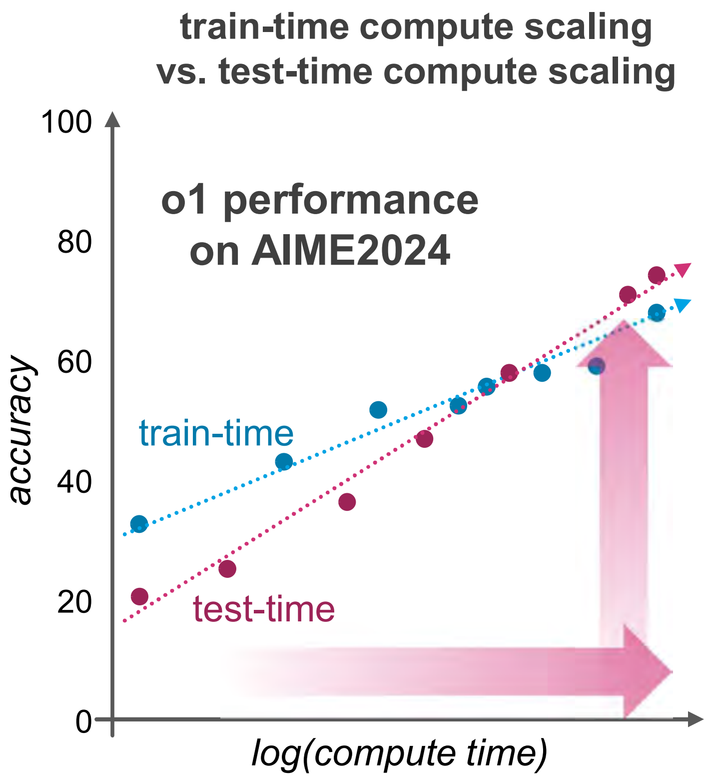
Test-Time Compute (TTC): Performance + Scaling







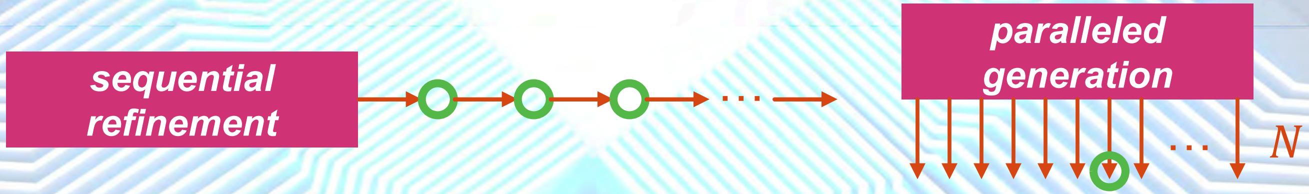
Learning to reason with LLMs — OpenAI



continuing to scale TTC →
continually improving accuracy

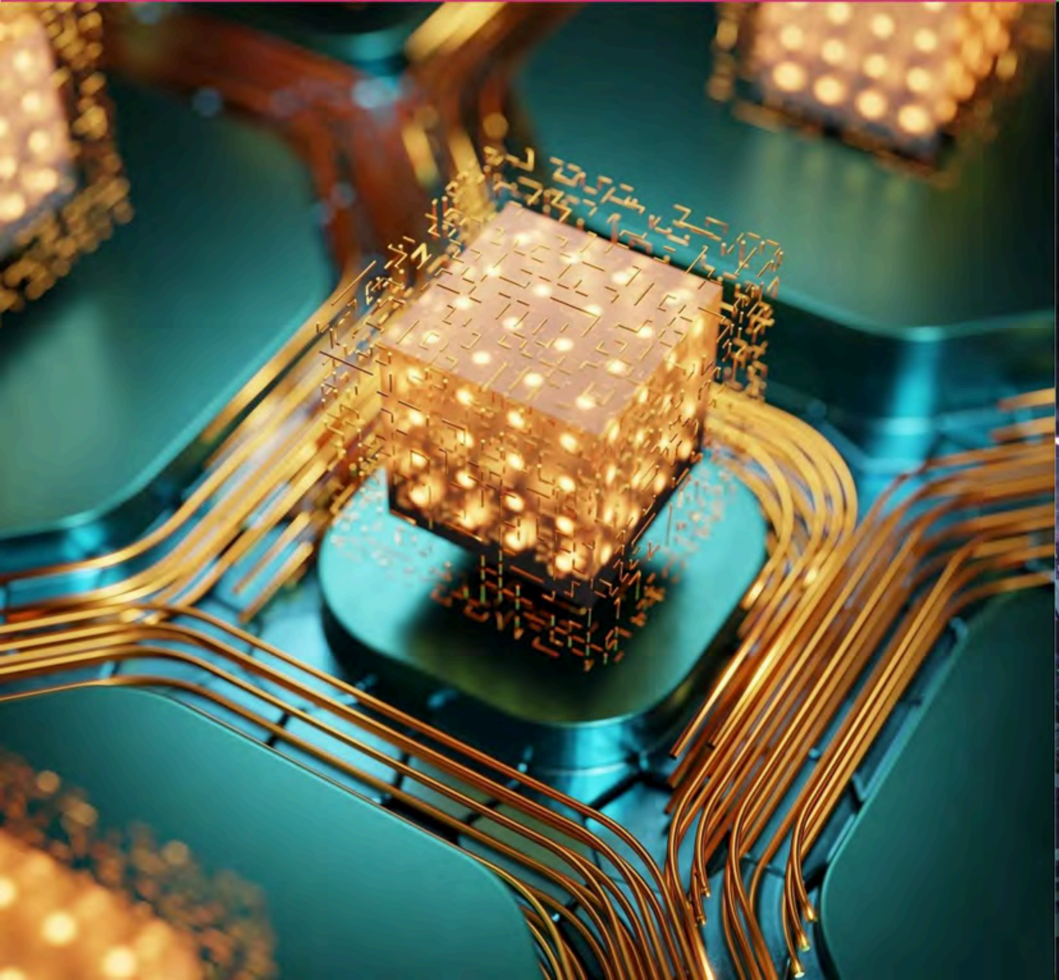

new scaling law:
LLM performance $\propto$ TTC$^{-\alpha}$ (power law)

all you need is compute

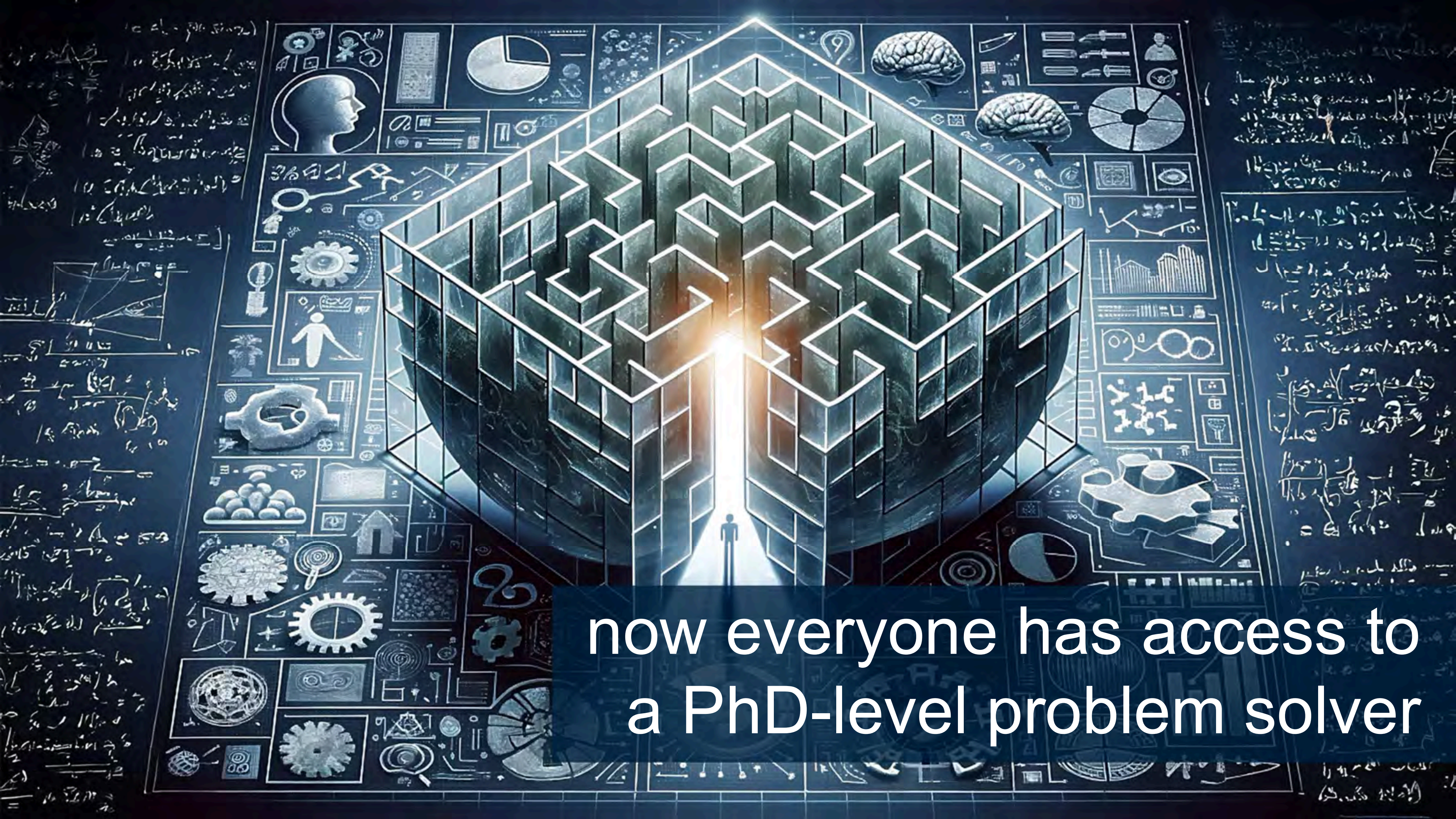


We Just Need More Compute: Future AI Data Center Investment

quantum compute

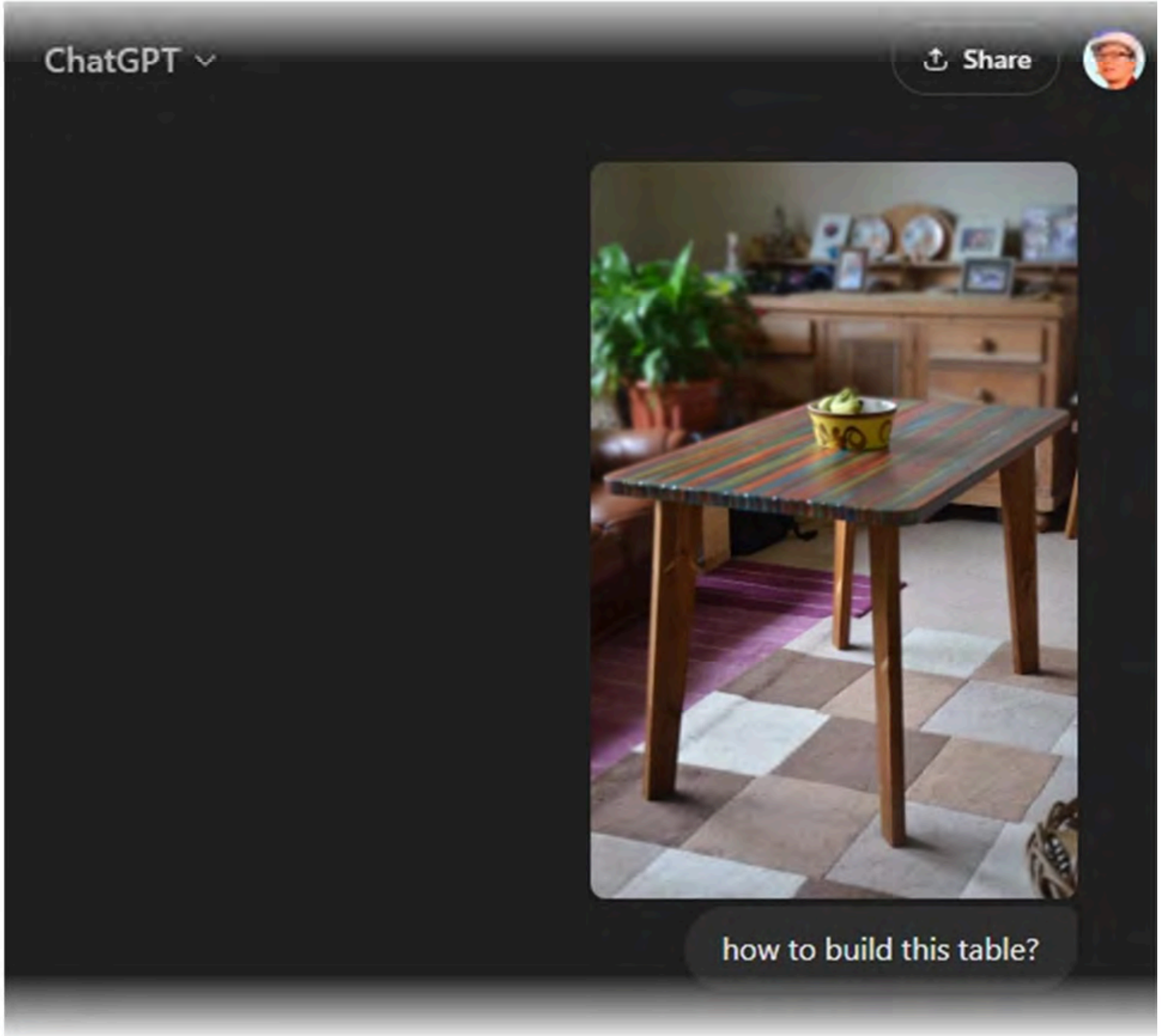


fusion energy



now everyone has access to
a PhD-level problem solver

LLMs are Passive



LLMs are Passive

Building a table like the one in the image involves several steps, including selecting materials, assembling the frame, and finishing the tabletop. Here's a step-by-step guide:

Materials Needed

1. **Wood for the tabletop** – Plywood, MDF, or reclaimed wood.
2. **Wood for the legs** – Solid wood (such as pine or oak).
3. **Wood screws and brackets** – For attaching the legs.
4. **Wood glue** – For extra strength.
5. **Sandpaper** – For smoothing edges.
6. **Primer and paint or resin** – For the colorful finish.
7. **Clear epoxy or varnish** – To seal the painted surface.
8. **Drill and screwdriver** – For assembling the frame.
9. **Saw** – To cut the wood to size.
10. **Clamps** – For securing pieces while gluing.

LLMs are Passive

Step 1: Preparing the Tabletop

- **Measure and cut** the tabletop to the desired size.
- **Sand the edges and surface** for a smooth finish.
- **Paint or apply the colorful design:**
  - Use acrylic paints or wood stains to create the multi-colored striped effect.
  - Allow each layer to dry before adding the next.
  - Optionally, use a brush or sponge for texture.

Step 2: Building the Legs

- Cut **four wooden legs** to your desired height.
- Use **tapered or straight legs** based on preference.
- Sand the legs to ensure smoothness.

LLMs are Passive

Step 3: Attaching the Legs

- **Option 1: Screw and Bracket Method**
  - Attach metal brackets under the tabletop.
  - Secure the legs to the brackets with screws.
- **Option 2: Pocket Hole Joinery**
  - Drill angled pocket holes into the legs.
  - Attach them to the tabletop with screws and wood glue.

Step 4: Finishing the Table

- **Apply a clear coat of epoxy or varnish** to protect the painted surface.
- Let it dry completely before use.

LLMs are Passive

- Drill angled pocket holes into the legs.
- Attach them to the tabletop with screws and wood glue.

Step 4: Finishing the Table

- **Apply a clear coat of epoxy or varnish** to protect the painted surface.
- Let it dry completely before use.

Final Touch

Place the table in your desired space and enjoy your handmade furniture!

Would you like help with a more detailed design plan or measurements?

5 stages of GenAI

AGI

innovator

agent

problem solver (LRM)

chatbot (LLM)

OpenAI

inspired by real-world concepts in software agents back in 1999

goal directed

programmed to keep order in the system (i.e. the Matrix)

context aware

can appear virtually anywhere to monitor anomalies everywhere

autonomy

can do whatever he deem necessary (eliminating humans + rogue programs)

tool usage

can use anything and anyone to give him super-human power

inspired by real-world concepts in software agents back in 1999

what's all the hype about agents now?

inspired by real-world concepts in software agents back in 1999

goal directed

programmed to keep order in the system (i.e. the Matrix)

context aware

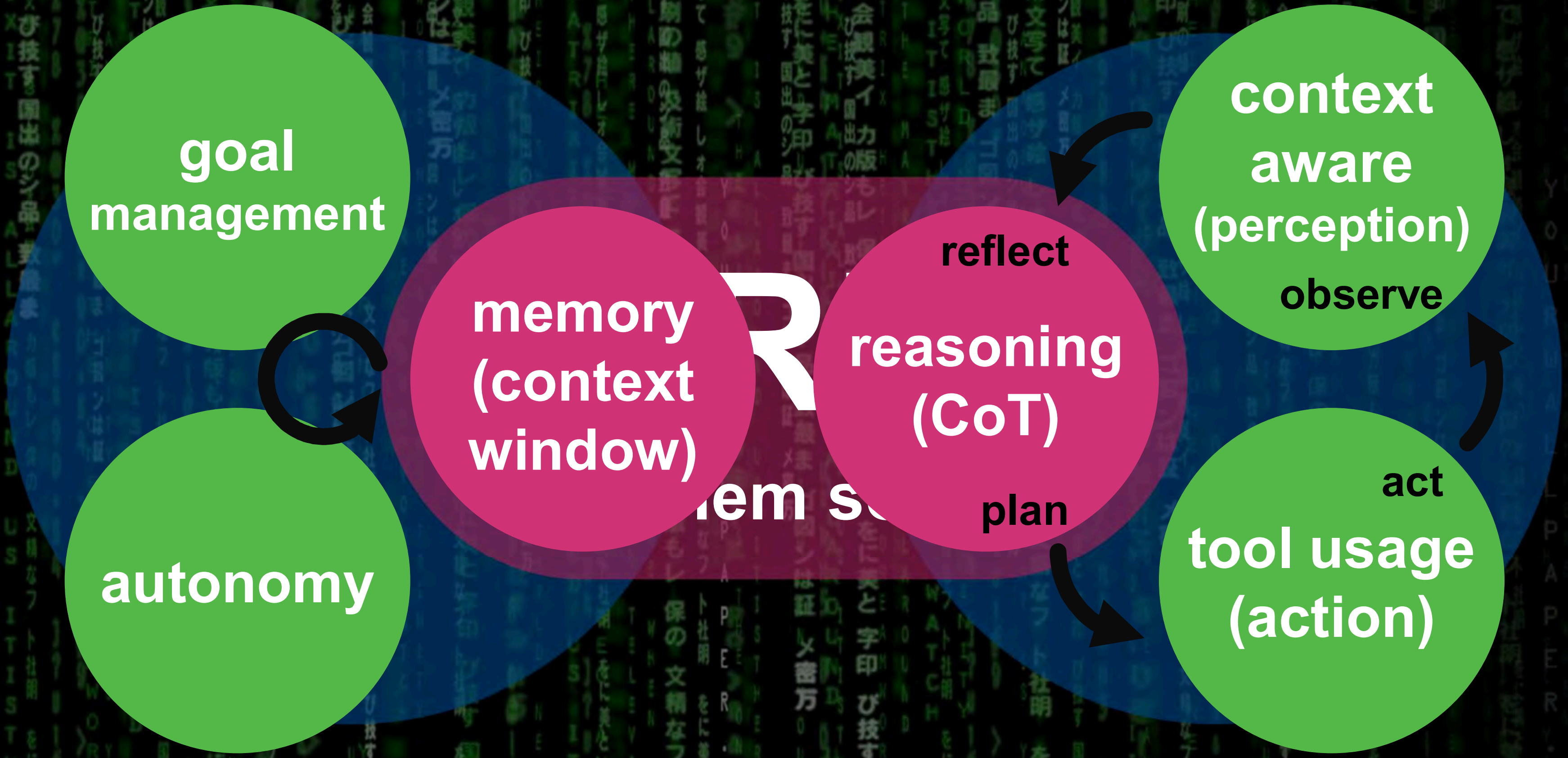
can appear virtually anywhere to monitor anomalies everywhere

autonomy

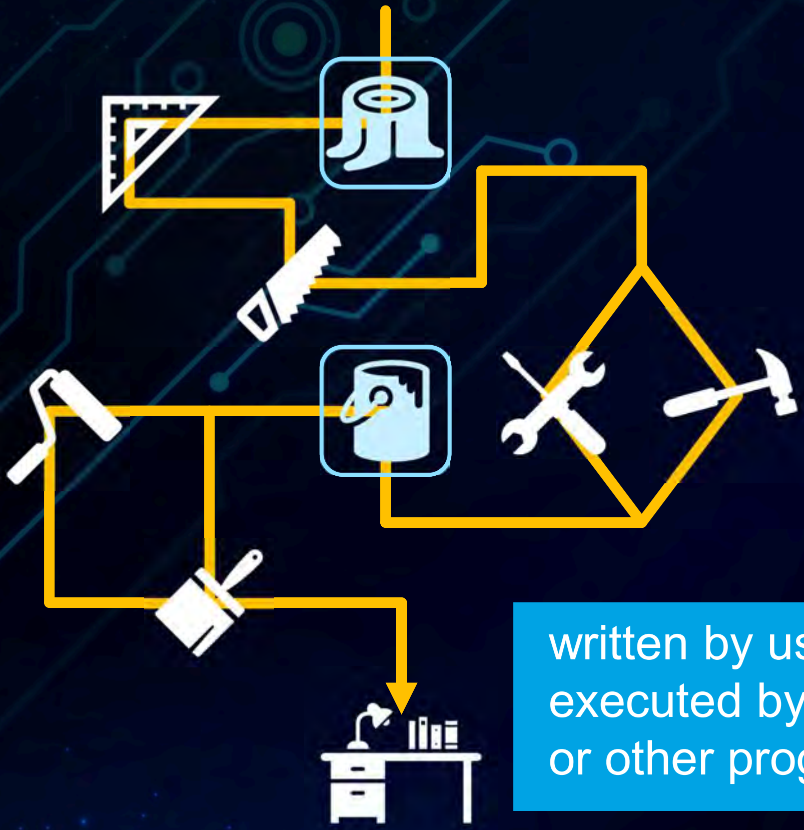
can do whatever he deem necessary (eliminating humans + rogue programs)

tool usage

can use anything and anyone to give him super-human power


goal management
autonomy
memory (context window)
reflect
reasoning (CoT)
plan
context aware (perception)
observe
act
tool usage (action)

ordinary computer program

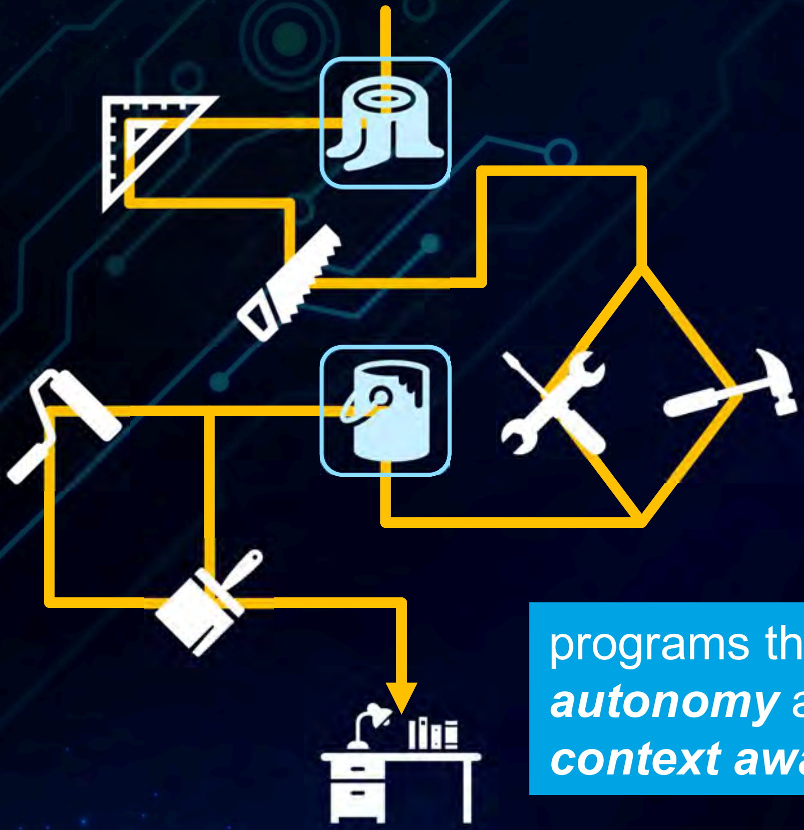


written by users
executed by users
or other programs

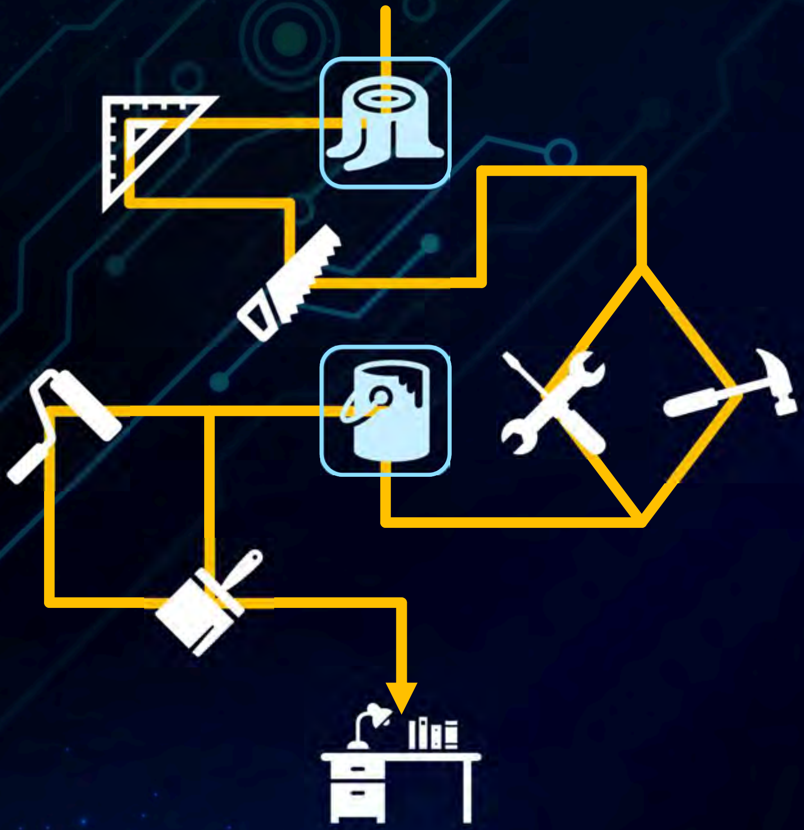
software agent

- can self execute (***autonomy***)
- must be able to observe the environment (***context aware***)
- still task specific (goal directed) + great for automation
- e.g. web crawler, load balancer, network intrusion detection

software agent

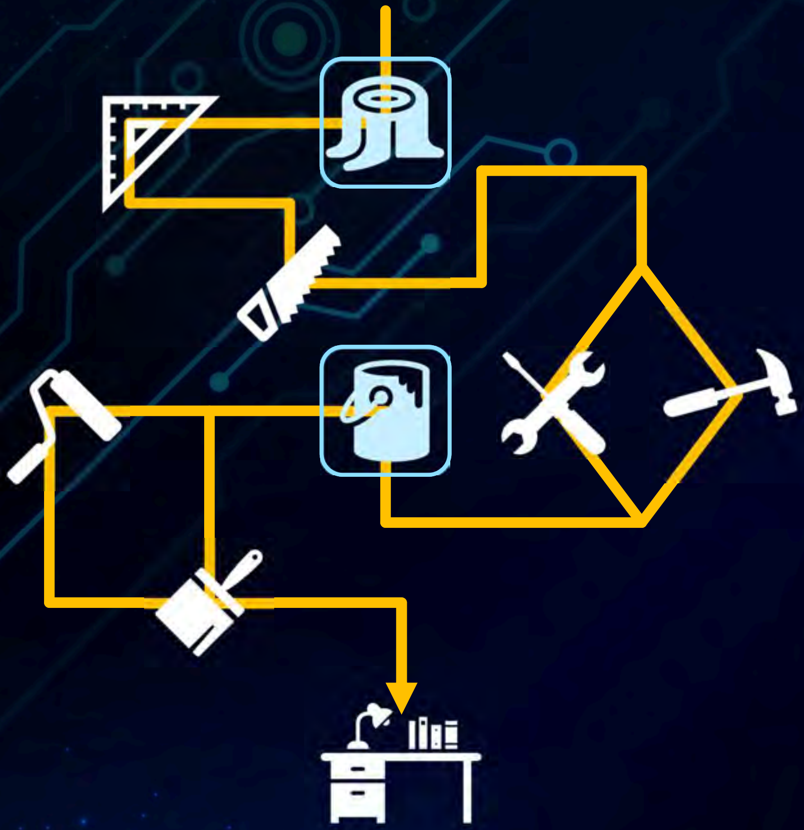


programs that has ***autonomy*** and are ***context aware*** software agent



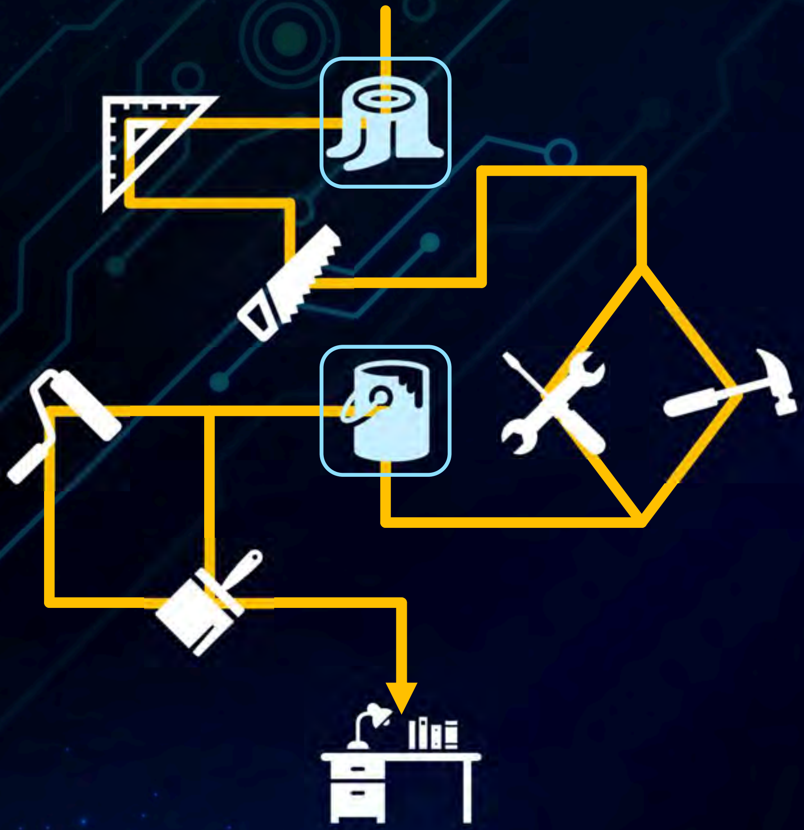
LLM agent

software agent

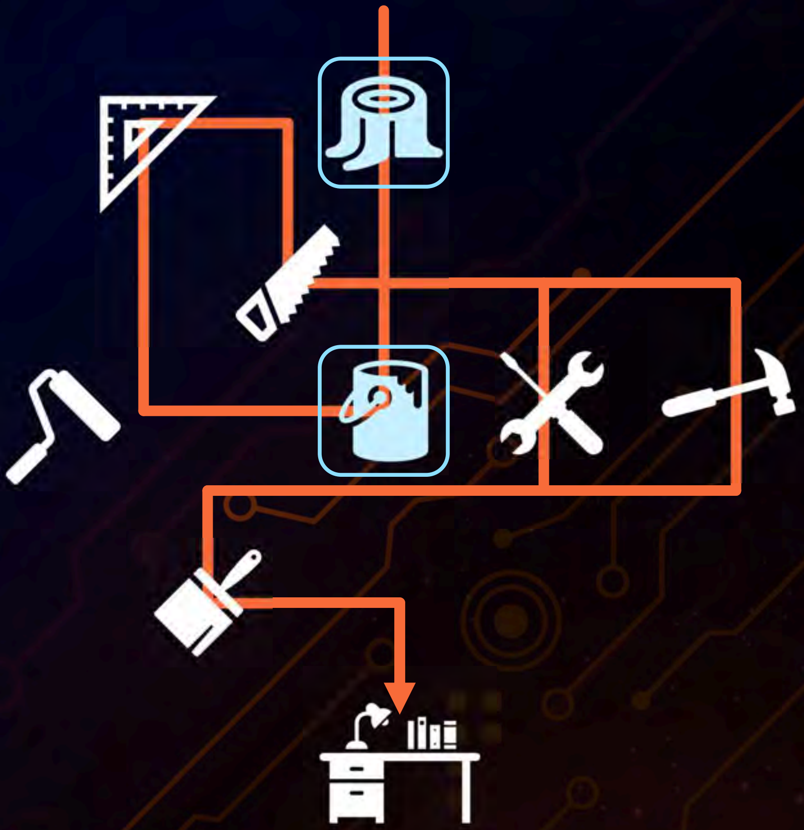


LLM agent

software agent



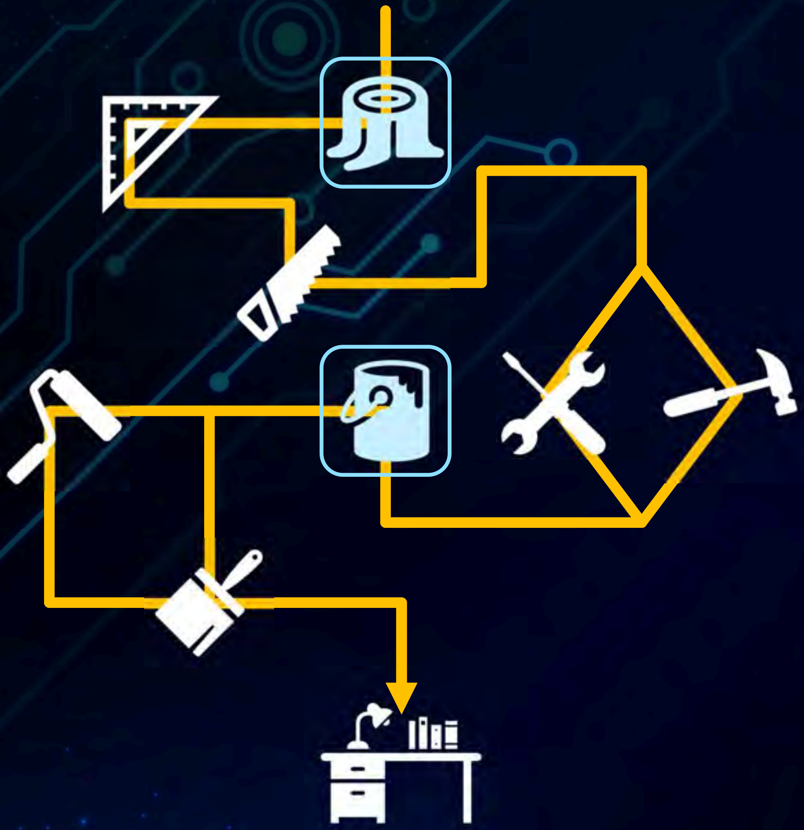
LLM agent



software agent

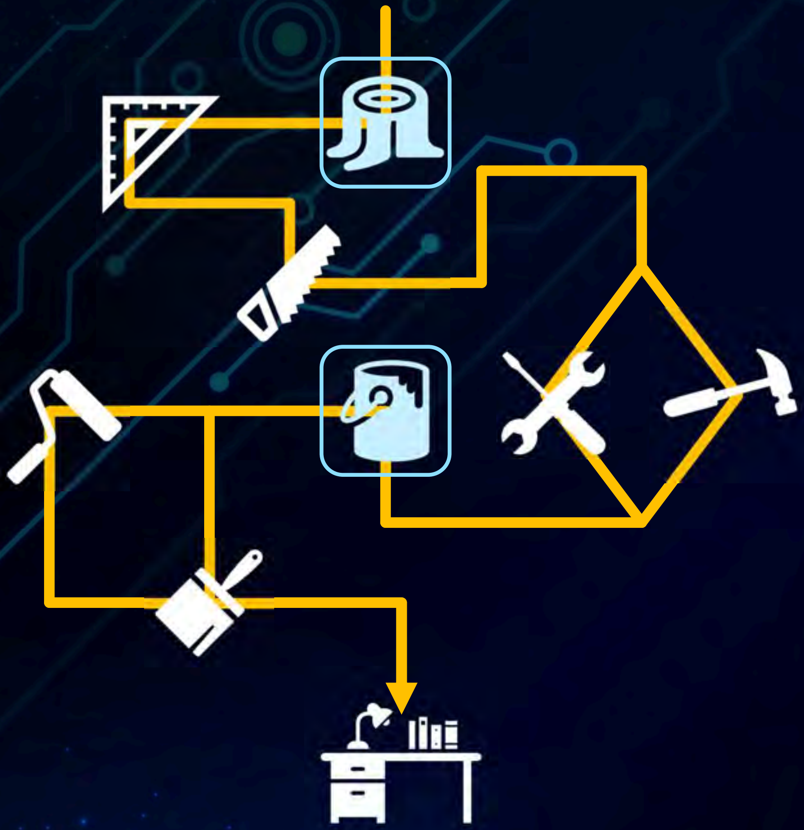
LLM agent

software agent



LLM agent

software agent



LLM agent

LLMs must use tools to act

LLMs must use tools to act, and in the digital world, tools = programs, API, or other agents

agents are a major step toward AGI
$\triangleq$ AI that has human-level performance
at most ***economically valuable work***


can GenAI
generate
revenue?
not by itself
37.71
+3.94%
35,294.76
+1.09%
16%
44%

an agent is only as good as
the ***tools*** it has access to

Future AI-Augmented Airline Retail

right time

conventional RM + WTP forecaster

model-based reinforcement learning

right channel

offer marketing

impossible to be everywhere → funnel them in

hyper-personalization

right customer

request-specific pricing

double debiased neural network-based robust estimation

right offer

ancillary AI algorithms

ranking
bundling
pricing

reinforcement learning

right price

continuous pricing + WTP model

online learning w/ variational inference

Future AI-Augmented Airline Retail

right time

conventional RM + WTP forecaster

model-based reinforcement learning

right channel

offer marketing

impossible to be everywhere → funnel them in

hyper-personalization

avg ROI

~192x

right customer

request-specific pricing

double debiased neural network-based robust estimation

>4.7%

revenue lift

right offer

ancillary AI algorithms

dynamic ancillary pricing

ranking bundling pricing

reinforcement learning

6%

revenue lift

right price

continuous pricing + WTP model

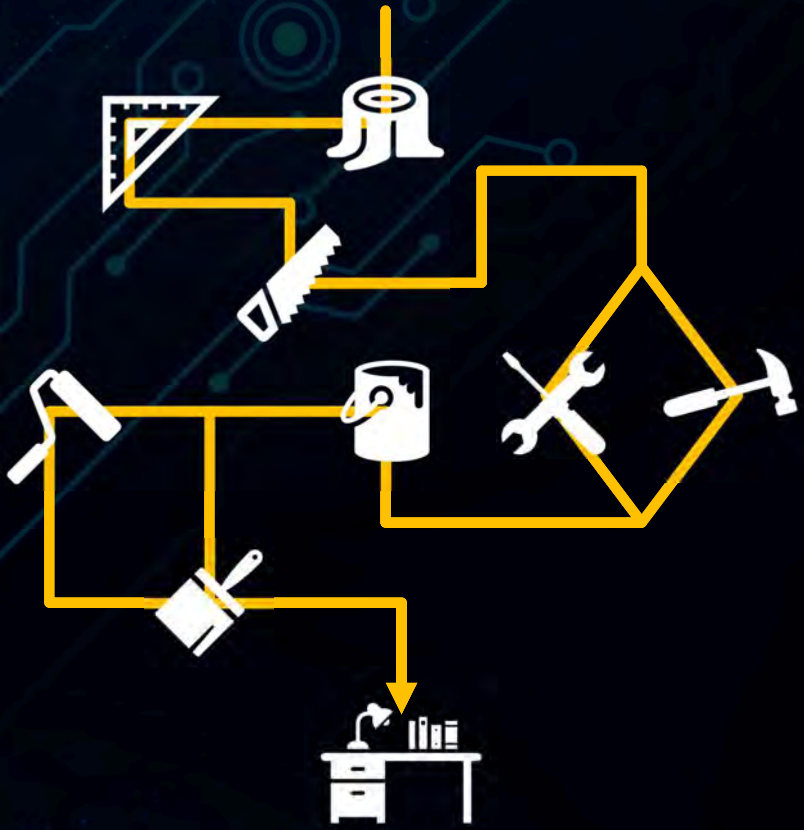
online learning w/ variational inference

1–3%

revenue lift

twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD

ordinary computer program



LLM agent

maximize total revenue

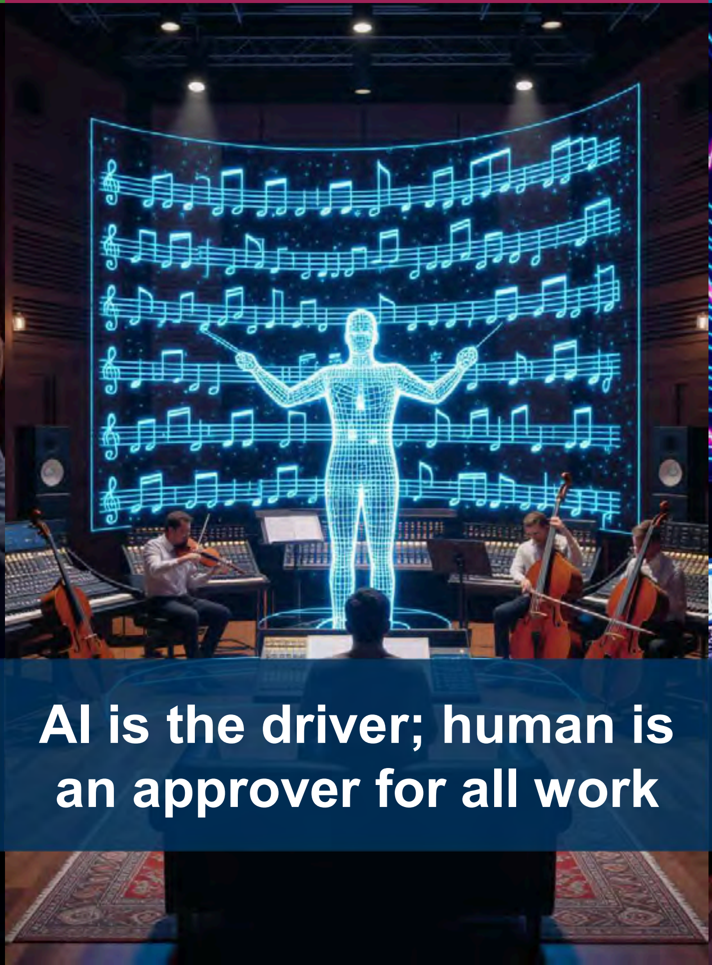

who’s ready to let agents make ***all*** their
business decisions ***fully*** autonomously?

AI Automation Roadmap

AI assisted

human is the driver; AI is an assistant ready to help

AI led

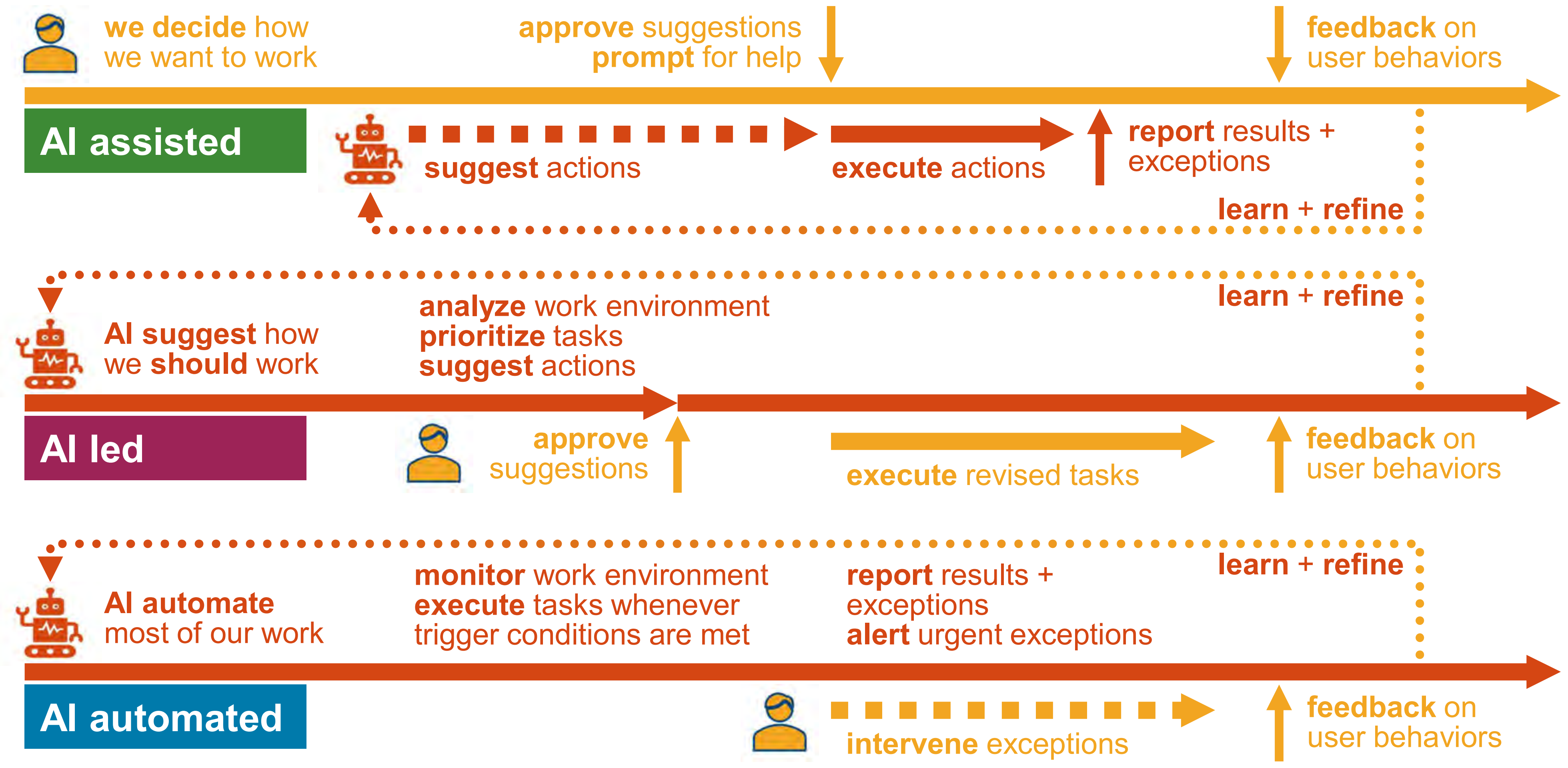


AI is the driver; human is an approver for all work

AI automated

AI is the driver; human is an intervenor as needed

AI Automation Roadmap



twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD

AI Automation Roadmap

human: start the work and working continuously

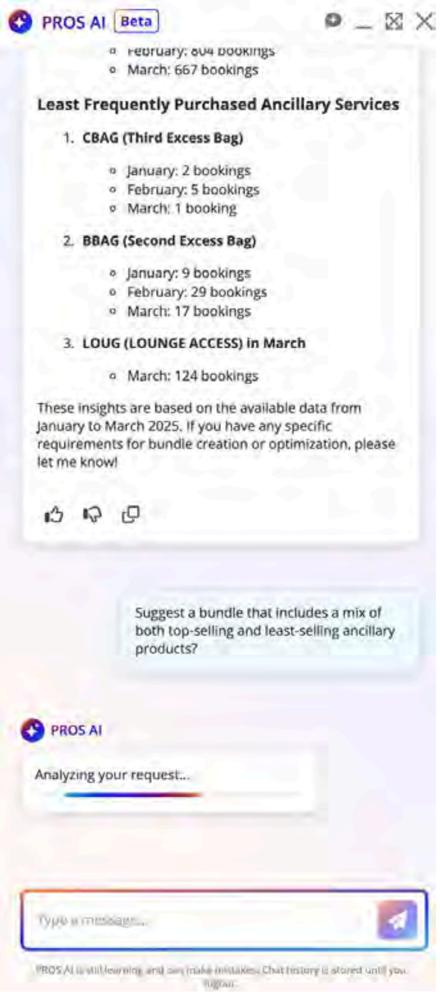
AI: working intermittently and help complete the work as needed

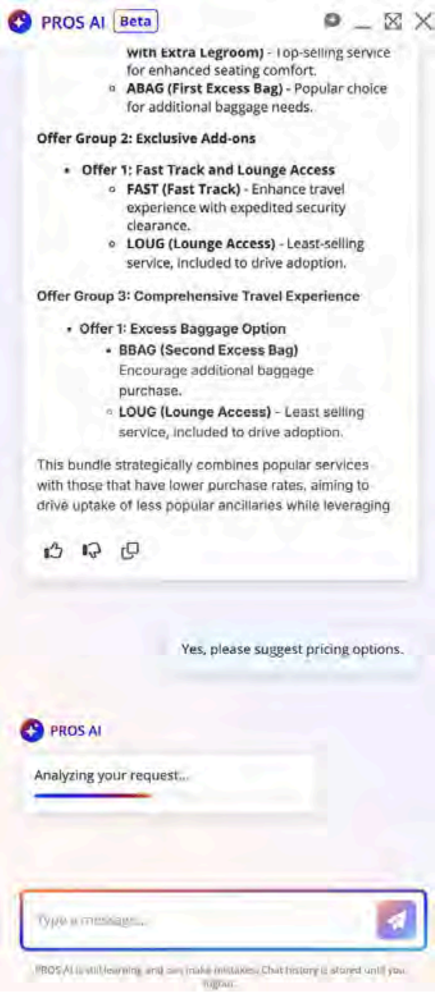
AI: start the work and working continuously

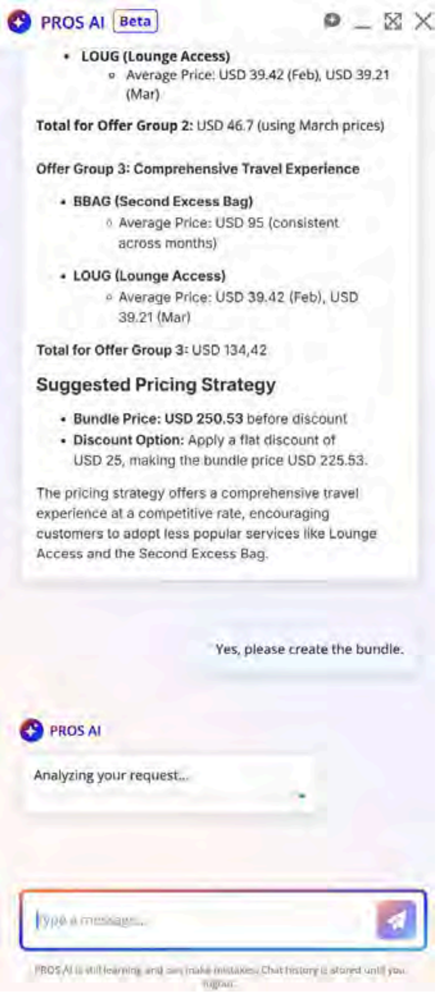
human: working intermittently and help complete the work as needed

Some Examples of PROS Agents

AI assisted: merchandising bundling agent



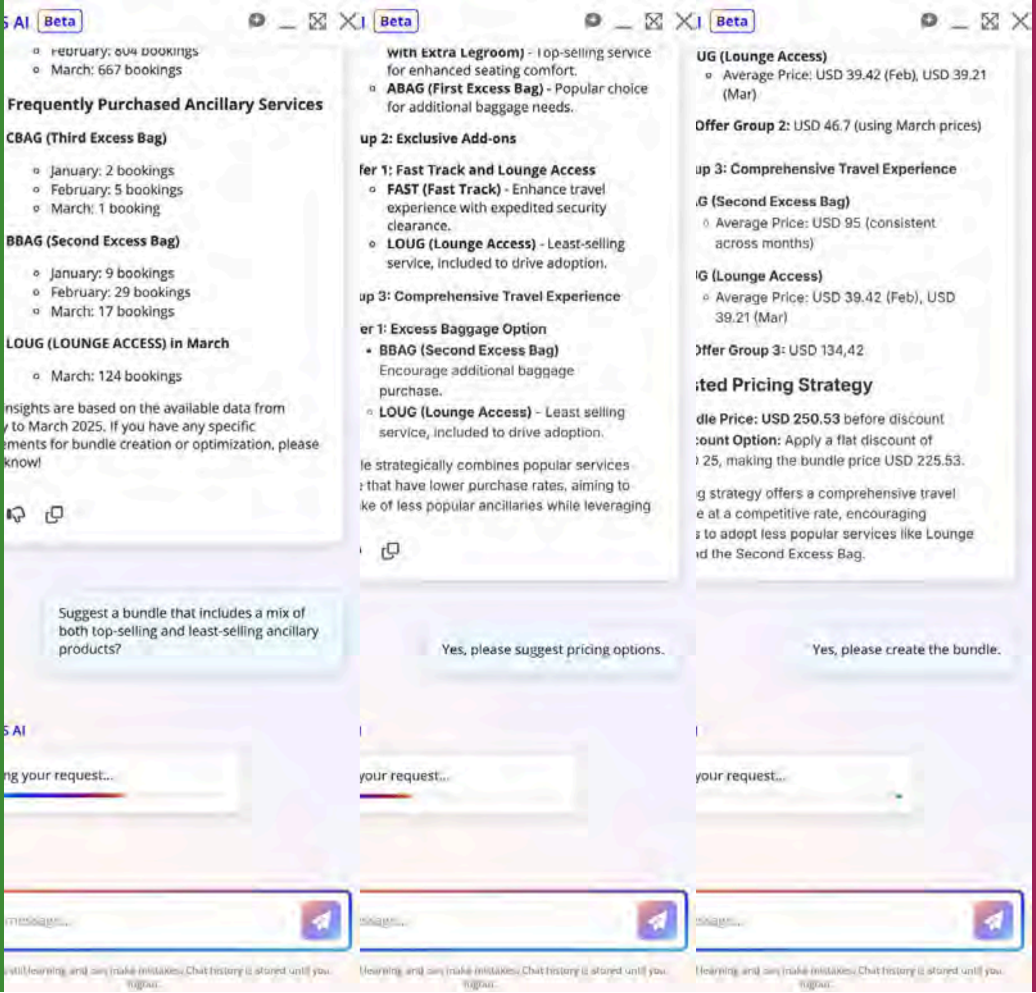




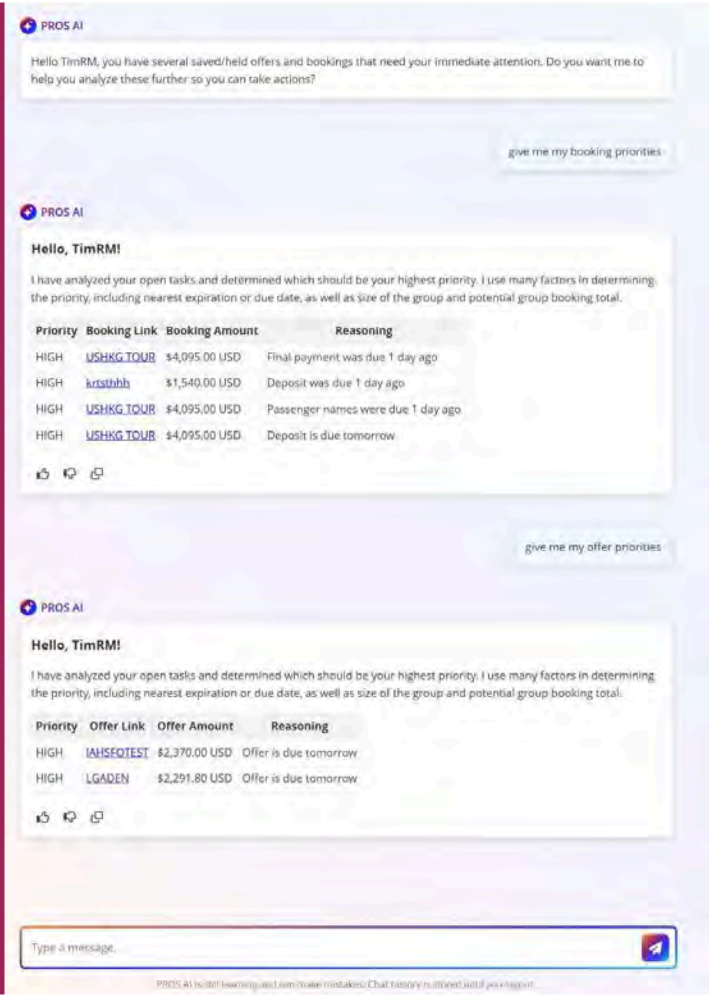
twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD

Some Examples of PROS Agents

AI assisted: merchandising bundling agent

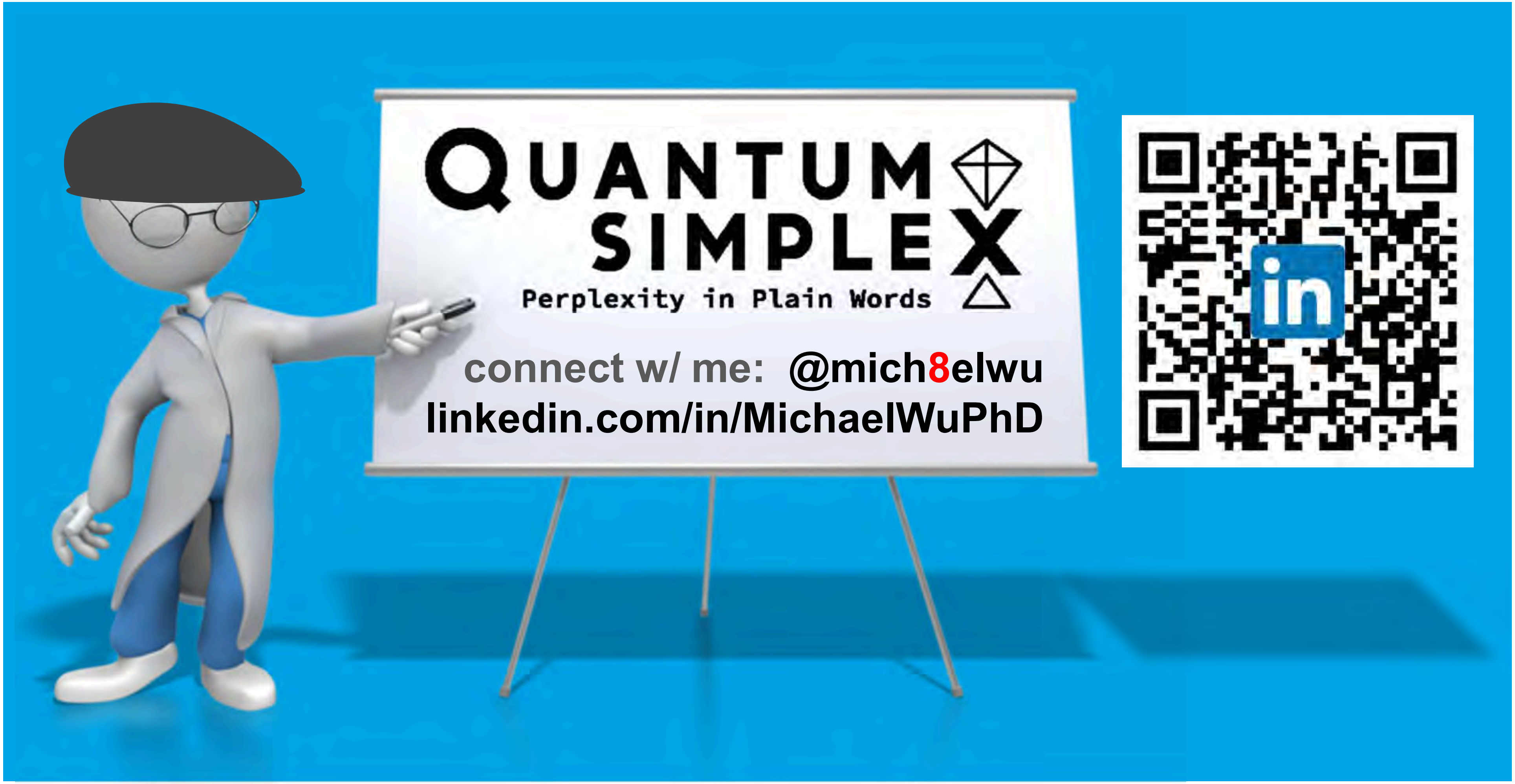


AI led: GSO task agent



twitter: @mich8elwu
linkedin.com/in/MichaelWuPhD


just as the autopilot, having agents
doesn’t mean we don’t need humans


QUANTUM
SIMPLEX
Perplexity in Plain Words
connect w/ me: @mich8elwu
linkedin.com/in/MichaelWuPhD
in