

# From Hype to Help: Intelligent GenAI Adoption

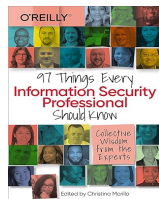
Sasha Czarkowski (Rosenbaum)  
@DivineOps





# Sasha Czarkowski (Rosenbaum) @DivineOps

Dev  
Ops  
Product  
Tech Sales  
Management



Ergonautic





Stop Chasing Buzzwords  
Start Working Better

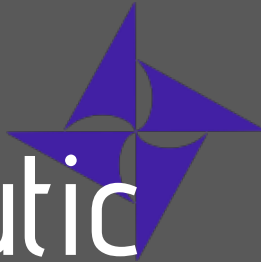


@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)

# Ergonautic



**Stop Chasing Buzzwords  
Start Working Better**



@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)

How about you?

# GenAI is the most Transformative Technology Of the last 17 years



Have you ever asked an LLM to  
generate an **image**?



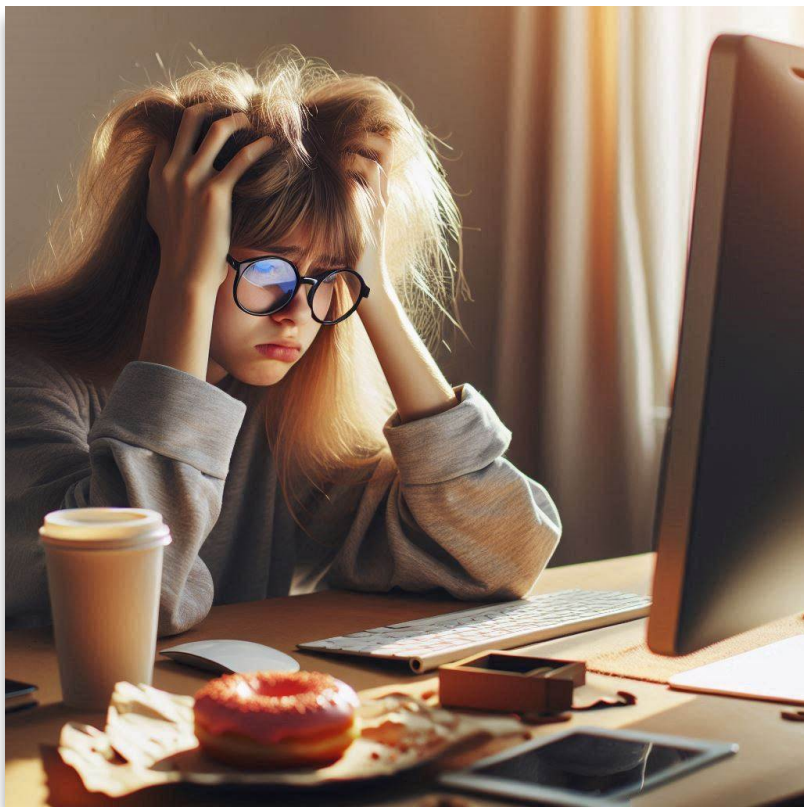




DeveSeceOps







2 hours later...

# Intermittent reinforcement



We work harder for  
intermittent rewards







# GenAI is an accidental experiment in Intermittent reinforcement



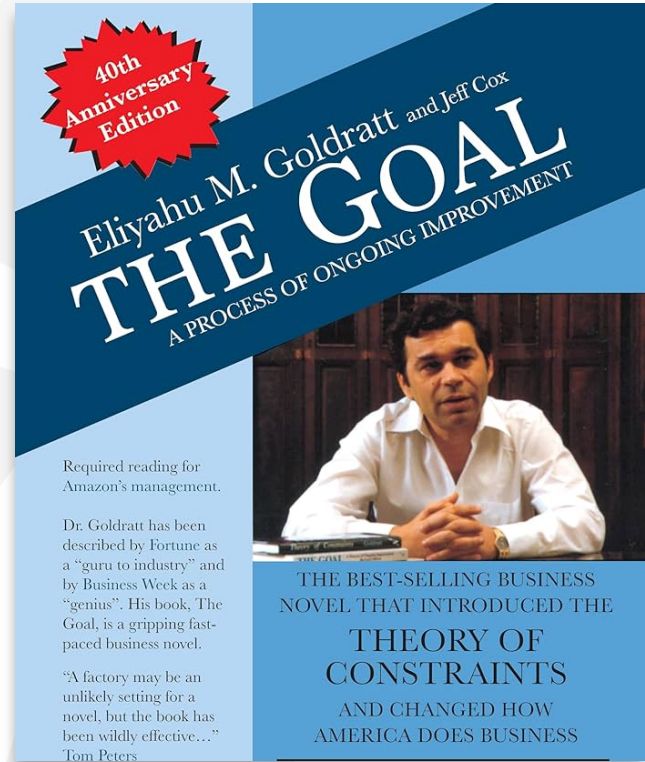
Ergonautic

Your “feels” might be wrong:

“While Copilot did not necessarily improve the **task completion time** or **success rate**, most participants preferred to use Copilot in daily programming tasks”

Source: *Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models.*

“What I’m telling you is,  
productivity is meaningless  
unless you know what your  
goal is”



Your choice of productivity **metrics**  
will determine your **outcomes**

# Klarna international

**New York, NY – February 27, 2024** – [Klarna](#) today announced its AI assistant powered by OpenAI. Now live globally for 1 month, the numbers speak for themselves:

- The AI assistant has had **2.3 million conversations, two-thirds of Klarna's customer service chats**
- It is doing the equivalent work of **700 full-time agents**
- It is on par with human agents in regard to customer satisfaction score
- It is more accurate in errand resolution, **leading to a 25% drop in repeat inquiries**
- Customers now **resolve their errands in less than 2 mins compared to 11 mins previously**
- It's available in **23 markets, 24/7** and communicates in more than **35 languages**
- It's estimated to drive a **\$40 million USD in profit** improvement to Klarna in 2024

Klarna has also seen massive improvement in communication with local immigrant and expat communities across all our markets thanks to the language support.

# IBM: Insurance industry bosses keen on AI. Customers, not so much

Fewer than 30% of clients happy dealing with a generative AI virtual agent

 [Richard Speed](#)

Wed 16 Oct 2024 // 14:45 UTC

An IBM study has found most insurance industry leaders believe generative AI is essential to keep pace with competitors. However, only a quarter of customers want anything to do with the technology.

The study is based on a survey of 1,000 insurance c-suite executives and 4,700 insurance customers. CEOs in the survey were evenly divided on whether generative AI was a risk versus an opportunity although 77 percent who responded said generative AI was necessary to compete.

Customers were not so keen: just 26 percent said they trusted an AI to dispense accurate and reliable advice.

We recruited **95** developers, and split them randomly into two groups.

We gave them the task of **writing a web server in JavaScript**

**45 Used**  
GitHub Copilot

**78%**  
finished

**1 hour, 11 minutes**  
average to complete the task



**50 Did not use**  
GitHub Copilot

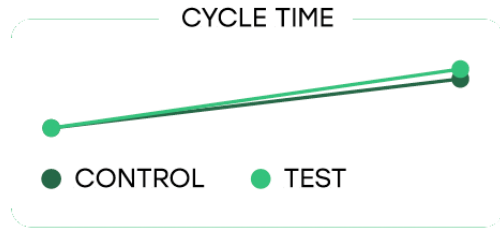
**70%**  
finished

**2 hours, 41 minutes**  
average to complete the task



Results are statistically significant ( $P=.0017$ ) and the 95% confidence interval is [21%, 89%]

Copilot access provided **no significant change** in efficiency metrics.



Developers with Copilot access saw a **significantly higher bug rate** while their issue throughput remained consistent.

**+41%**  
IN BUG RATE



# 77% Of Employees Report AI Has Increased Workloads And Hampered Productivity, Study Finds

**Bryan Robinson, Ph.D.** Contributor 

*author of Chained to the Desk in a Hybrid World: A Guide to Balance.*

Follow



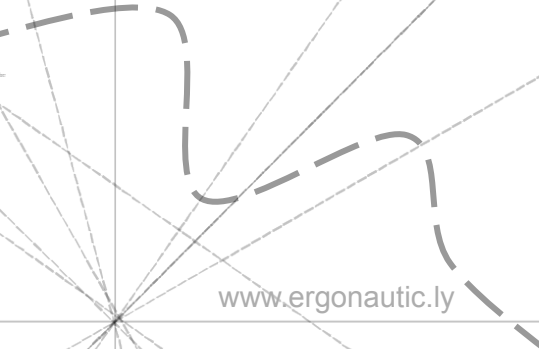
3

Jul 23, 2024, 09:01am EDT

# More Robots



Ergonautic



@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)





r/neovim

u/LongAd9257 · 2h

Join

## Forcing IDE at work

### Discussion

Hey everyone, I'm just wondering if anyone had any similar situations like me. So basically, at work we were using IDE that suits us best, but lately management is forcing us to switch to Cursor IDE. Don't get me wrong I've got nothing against cursor, but I am so used to my neovim config, plugins and motions. I just don't think that it's fair to force bunch of developers to use cursor expecting to have you product/code delivered faster/better because AI will be writing if better... Did anyone had any similar situations?

We have been here before

# The Tesla Robot Revolution

- Tesla's 2017 Model 3 assembly line built with > 1000 robots
- Musk said the line will produce 20,000 cars a month by EOY
- Tesla makes only 2,425 in the 4th quarter of 2017

TRANSP0 / CARS / TESLA

## Tesla relied on too many robots to build the Model 3, Elon Musk says / The guy telling everyone to be afraid of robots uses too many robots in his factory

By [Andrew J. Hawkins](#), transportation editor with 10+ years of experience who covers EVs, public transportation, and aviation. His work has appeared in The New York Daily News and City & State.

Apr 13, 2018, 1:41 PM EDT



0 Comments (0 New)

Elon Musk says Tesla relied on too many robots to build the Model 3, which is partly to blame for the delays in manufacturing the crucial mass-market electric car. In an interview with CBS Good Morning, Musk agreed with Tesla's critics that there was over-reliance on automation and too few human assembly line workers building the Model 3.



But wait,  
there is more

# The GM Robot Revolution

- In 1980s, GM spends > \$45 Billion on industrial robots
- CEO Roger Smith lauded as a “visionary 21st century futurist”
- By 1990, GM earnings per US car sale fall from \$588 to \$47



# GM's decline truly began with its quest to turn people into machines

Meanwhile, GM continued to lavish spending on big capital investments, confident that the secret to competitiveness lay in replacing humans with technology. But as in Lordstown, the spending bore little fruit. As automotive analyst Maryann Keller recounted in her 1989 book *Rude Awakening*, one GM executive observed that, between 1980 and 1985, the company shelled out an eye-popping \$45 billion in capital investment. Despite that spending, its global market share rose by but a single percentage point, to 22%. "For the same amount of money, we could buy Toyota and Nissan outright," said the executive -- which would have instantly bumped GM's market share to 40%.



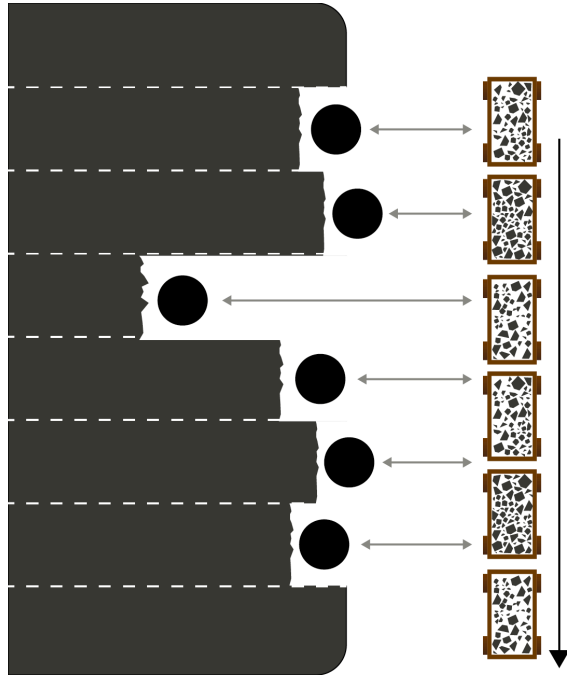
But wait,  
there is more

The origins of the term

“Socio-Technical”:

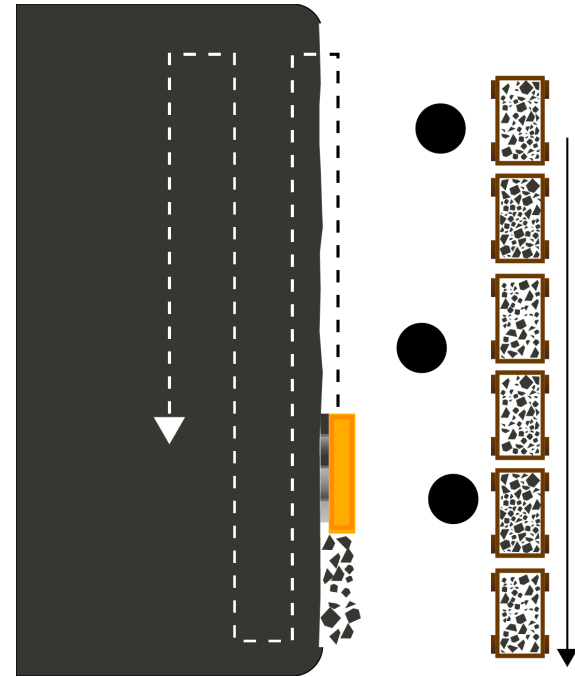
studies by Eric Trist in 1950 Britain  
of mechanization of Coal Mining

## Mostly Manual



VS

## Giant Machine



The expected **productivity** of the new mechanized approach **did not develop**

# The Irony



Ergonautic



@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)

February 17 2024

## **Ironies of Generative AI: Understanding and mitigating productivity loss in human-AI interactions**

AUSTE SIMKUTE\*, University of Edinburgh, United Kingdom

LEV TANKELEVITCH\*, Microsoft Research, United Kingdom

VIKTOR KEWENIG, University College London, United Kingdom

AVA ELIZABETH SCOTT, University College London, United Kingdom

ABIGAIL SELLEN, Microsoft Research, United Kingdom

SEAN RINTEL, Microsoft Research, United Kingdom

**(a) The production-to-evaluation shift**



**(b) Unhelpful workflow restructuring**

Prompting



Output adaptation



**(c) Task interruptions**



**(d) Task-complexity polarization**



Loss of task sequence



Loss of feedback







## Production to Evaluation



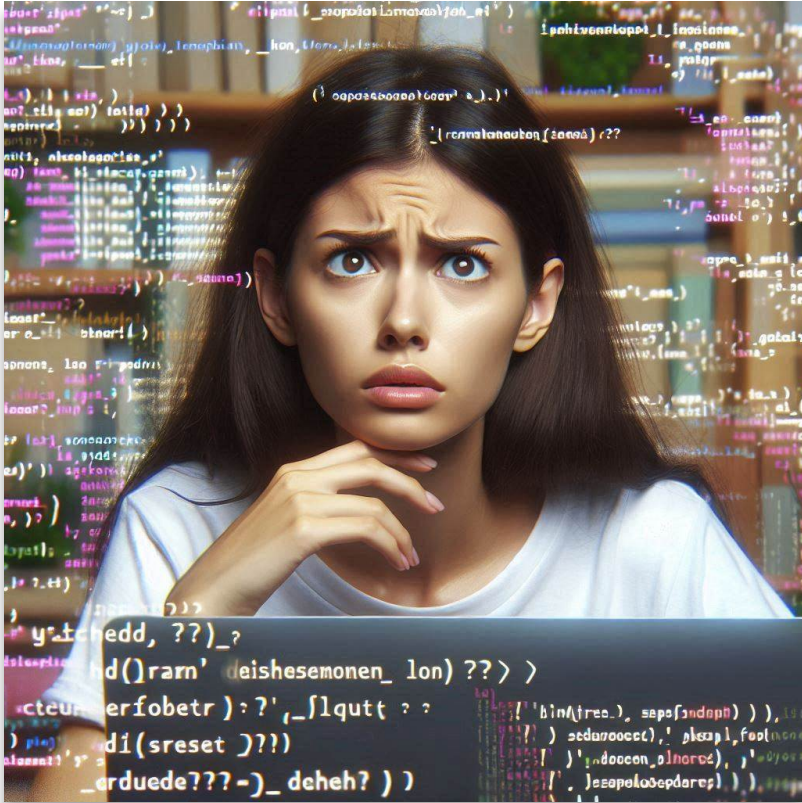
- Output overload
- Complacency
- Reduced situational awareness



Loss of Feedback

No social commentary

No ratings



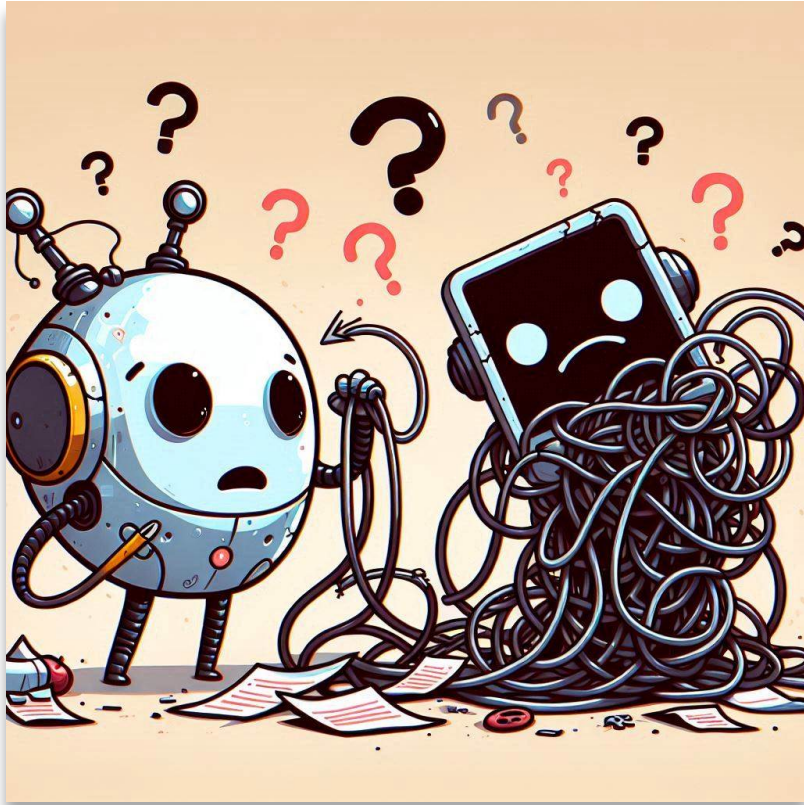
# Task Interruptions

Cognitive costs

Broken Flow

# Task Complexity Polarization

Automation can make  
easy tasks easier and  
hard tasks even harder



Inevitably, the friction with the machine  
is getting blamed on **human error**



TRANSP0

# Deadly Boeing crashes raise questions about airplane automation

‘Automation is a double-edged sword’

By **Andrew J. Hawkins**, transportation editor with 10+ years of experience who covers EVs, public transportation, and aviation. His work has appeared in The New York Daily News and City & State.

Mar 15, 2019, 10:40 AM PDT



0

Comments (0 New)

# The Transformation



Ergonautic



@ergonauticly

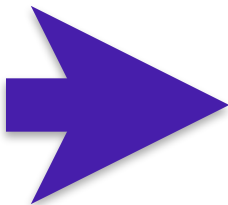
COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)

xkcd.com/1425/




IN CS, IT CAN BE HARD TO EXPLAIN  
THE DIFFERENCE BETWEEN THE EASY  
AND THE VIRTUALLY IMPOSSIBLE.



Is this a bird?



 Copilot

✓ Analyzing the image: Privacy blur hides faces from Copilot  
Yes, this is a bird! Specifically, it's a white stork. The image shows the stork standing by the edge of a body of water, likely a pond or stream. White storks are known for their long red legs and pointed red beaks, along with black wing feathers that contrast with their predominantly white bodies. They are often seen in wetlands and are quite fascinating to observe in their natural habitat. 🦩

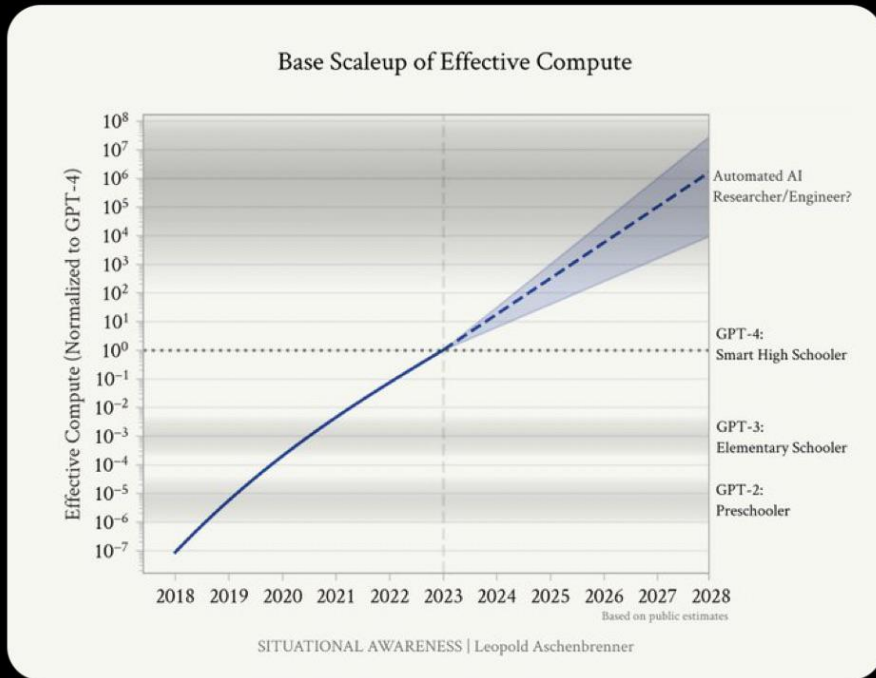




Leopold Aschenbrenner  @leopo... · 20m ...

AGI by 2027 is strikingly plausible.

That doesn't require believing in sci-fi; it just requires believing in **straight lines on a graph.**




 Sasha Czarkowski

 16

 19

 104

 3.7K



[www.ergonautic.ly](http://www.ergonautic.ly)



Christian Keil  

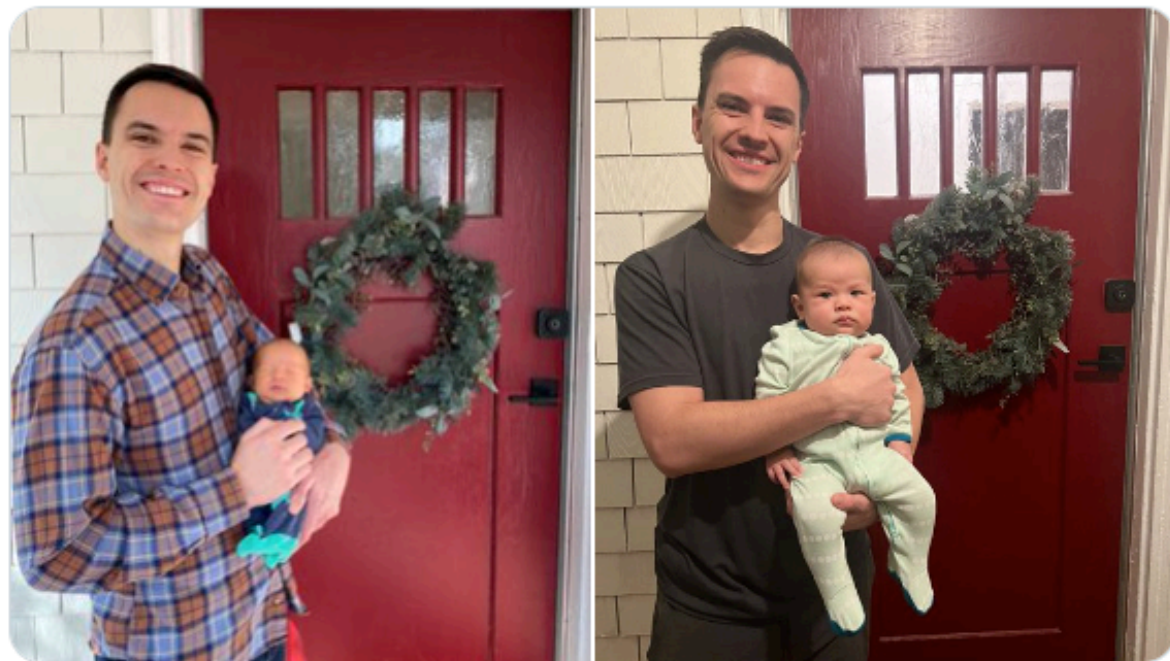
@pronounced\_kyle

Subscribe



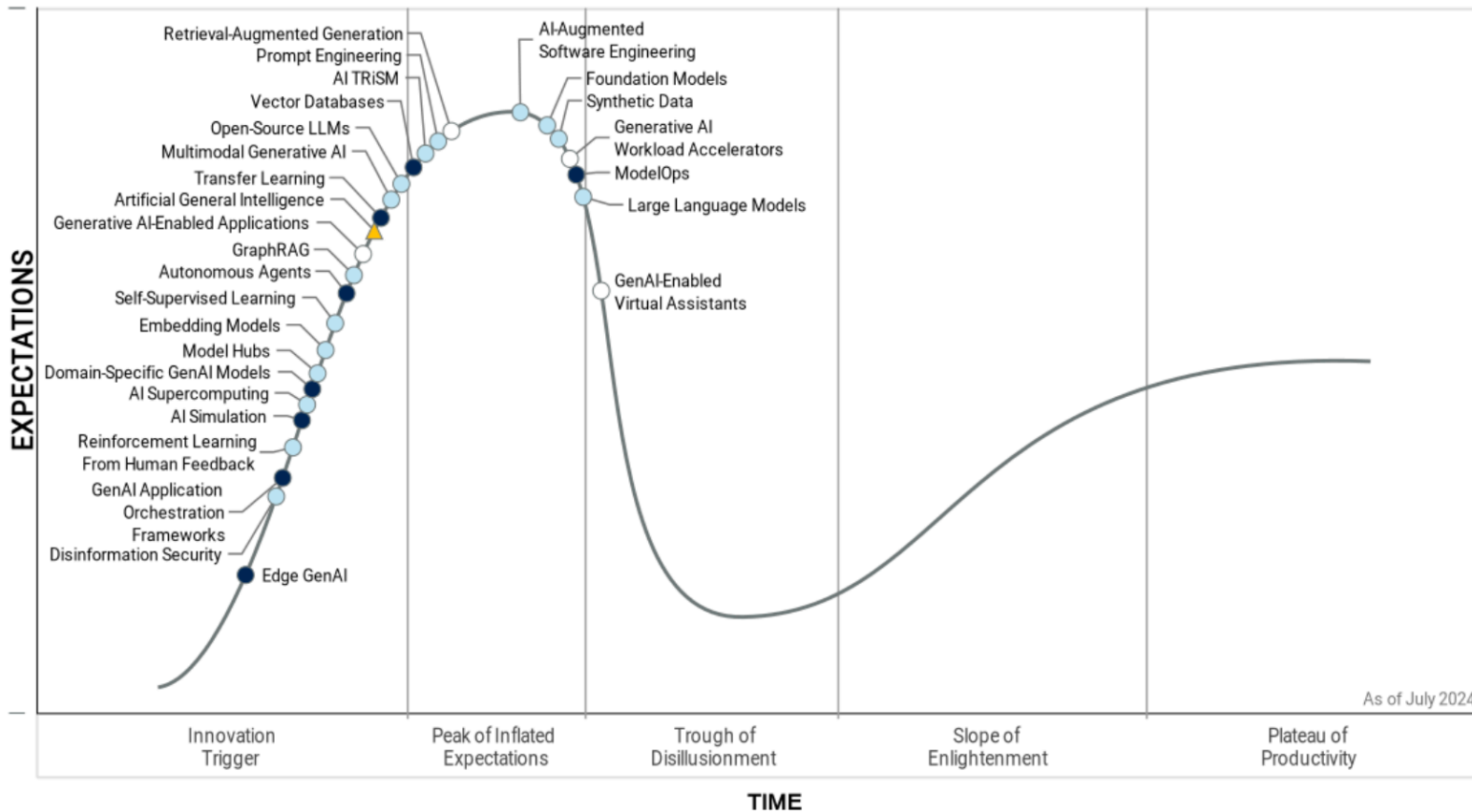
My 3-month-old son is now TWICE as big as when he was born.

He's on track to weigh 7.5 trillion pounds by age 10



12:11 AM · Mar 16, 2024 · 3M Views

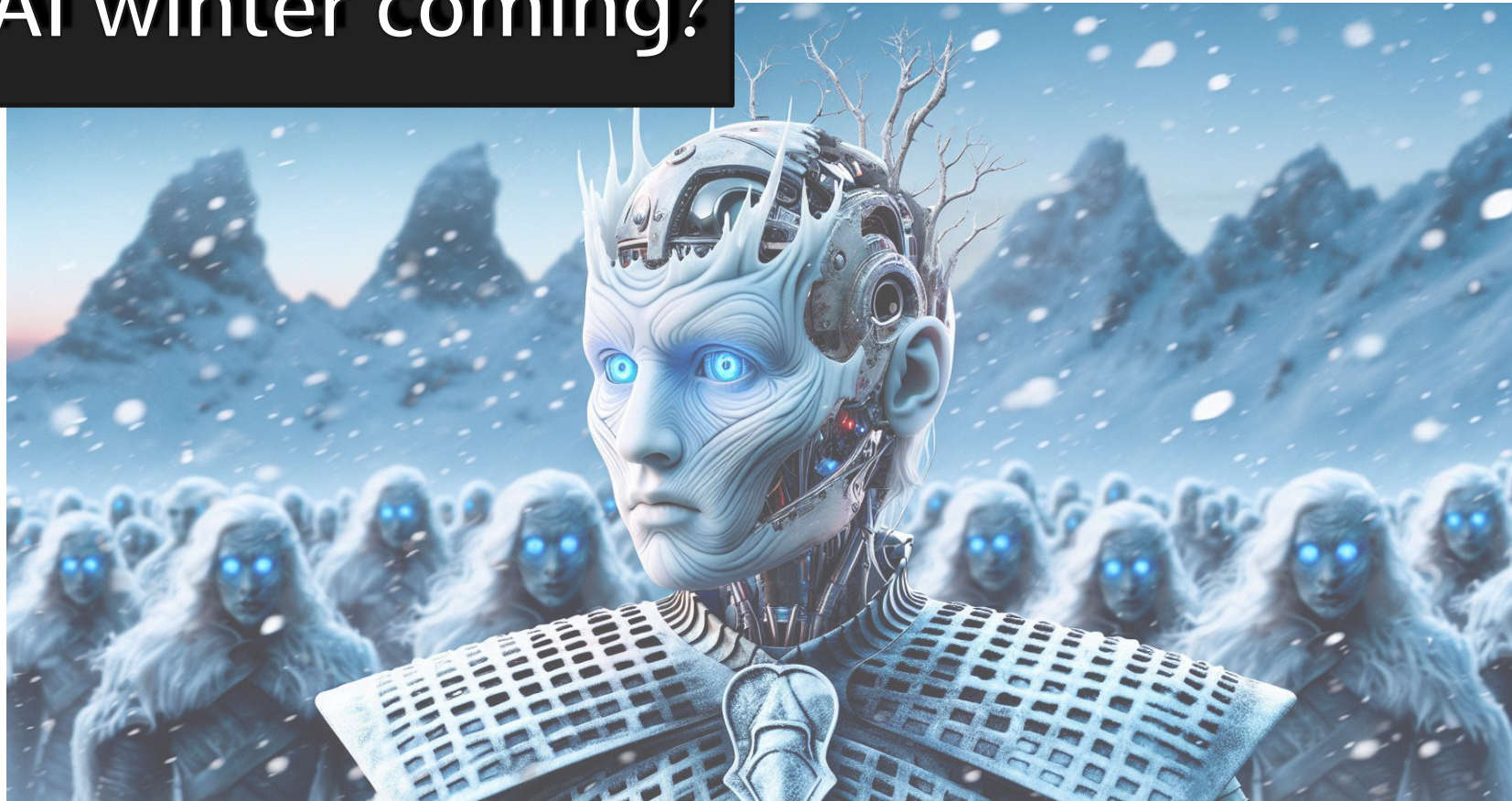
# Figure 1: Hype Cycle for Generative AI, 2024



As of July 2024

Plateau will be reached: ○ <2 yrs. ● 2-5 yrs. ● 5-10 yrs. ▲ >10 yrs. ⊗ Obsolete before plateau

# Is AI winter coming?



GenAI?

Serverless

Cloud

IoT

Continuous Delivery

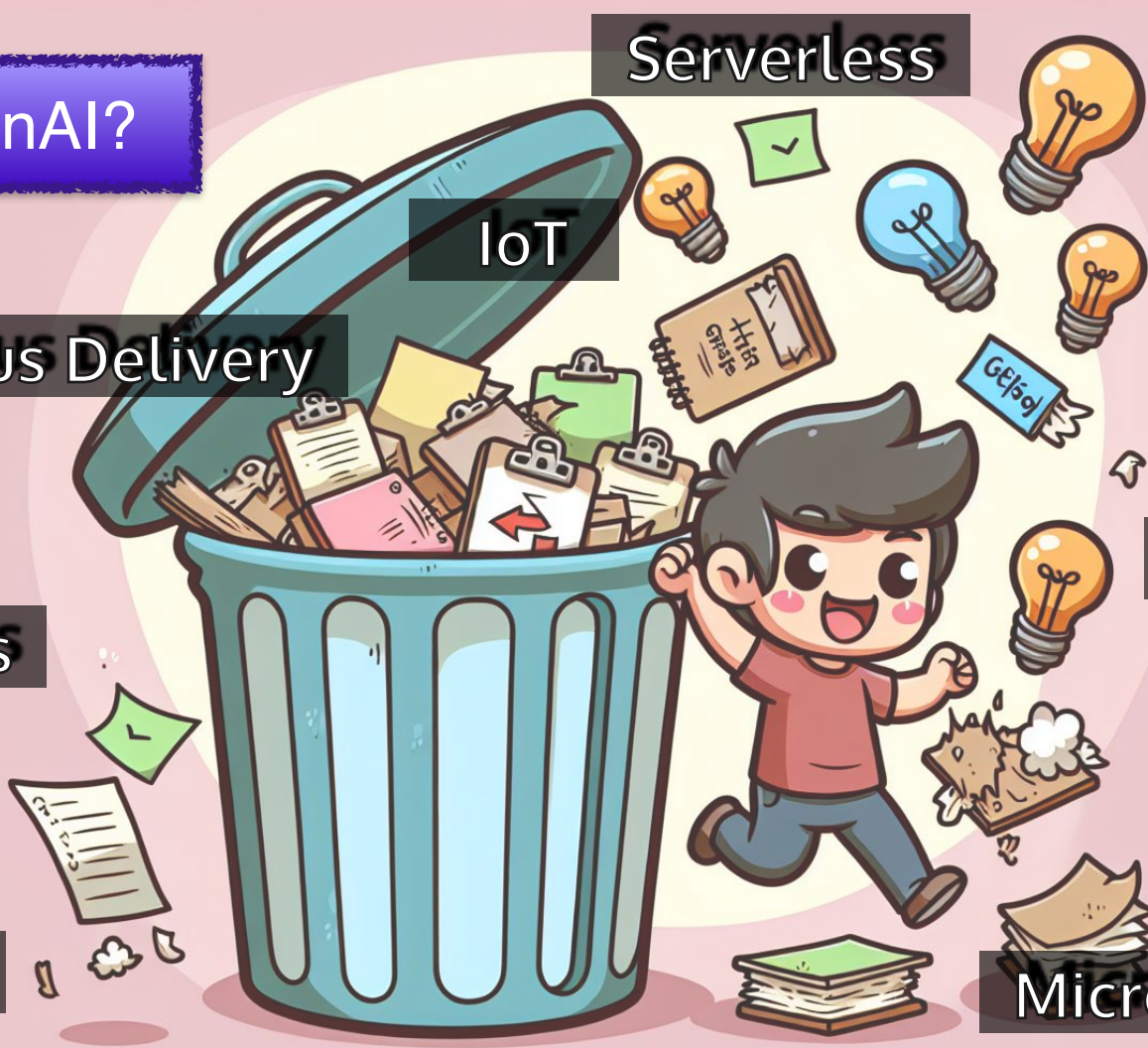
DevOps

Lean

Agile

Big Data

Microservices



# Workflow Design





Person

Product

# Personal [work] flow



Ergonautic



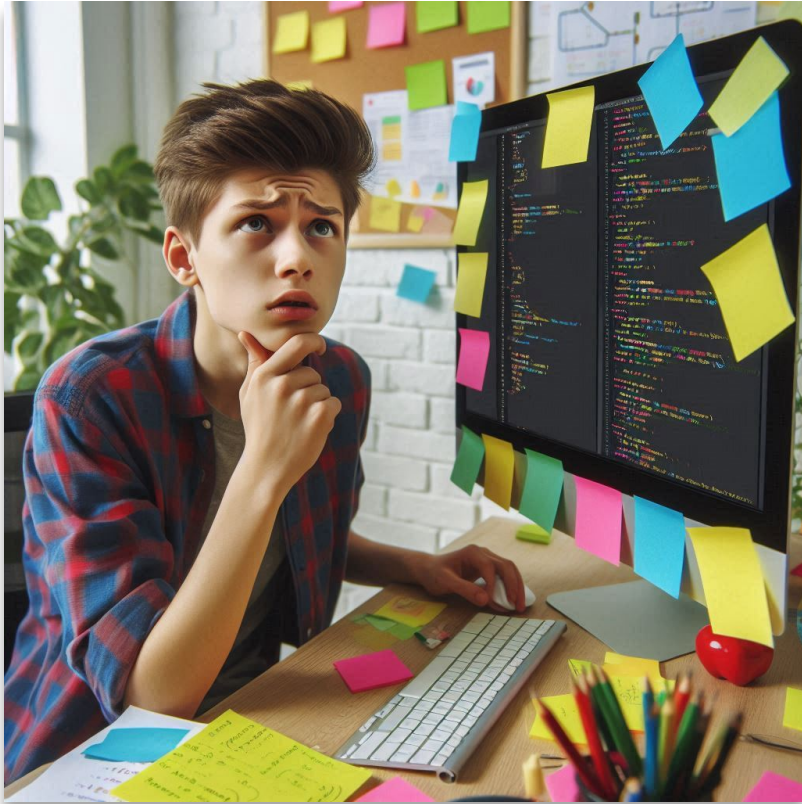
@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)





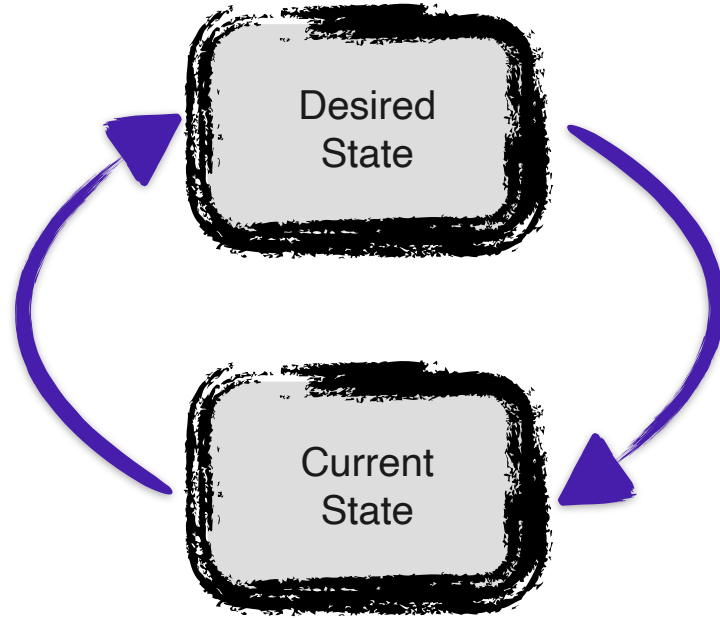


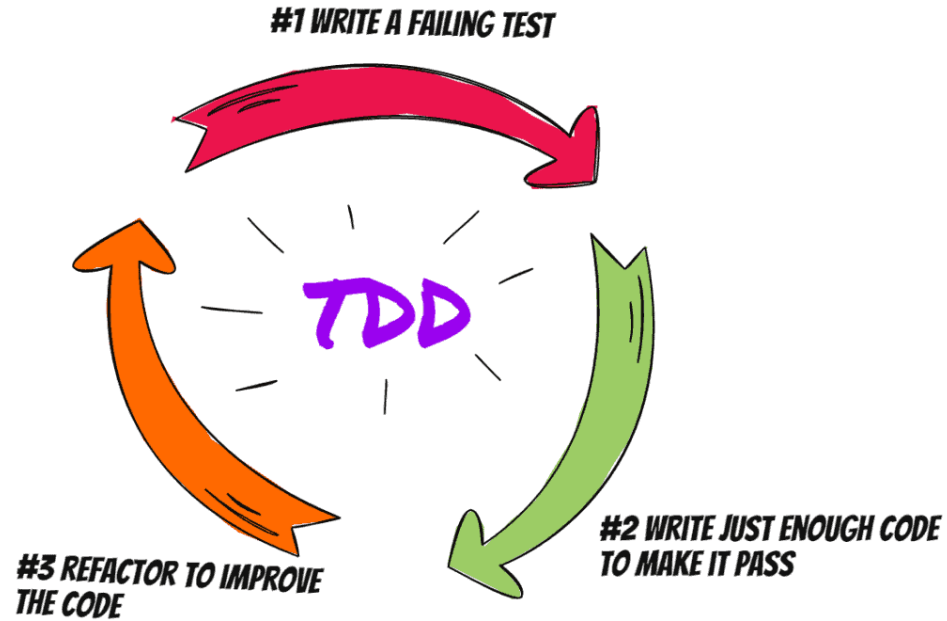
# “The Death of the Junior Professional”

What **skills** do “seniors” have?

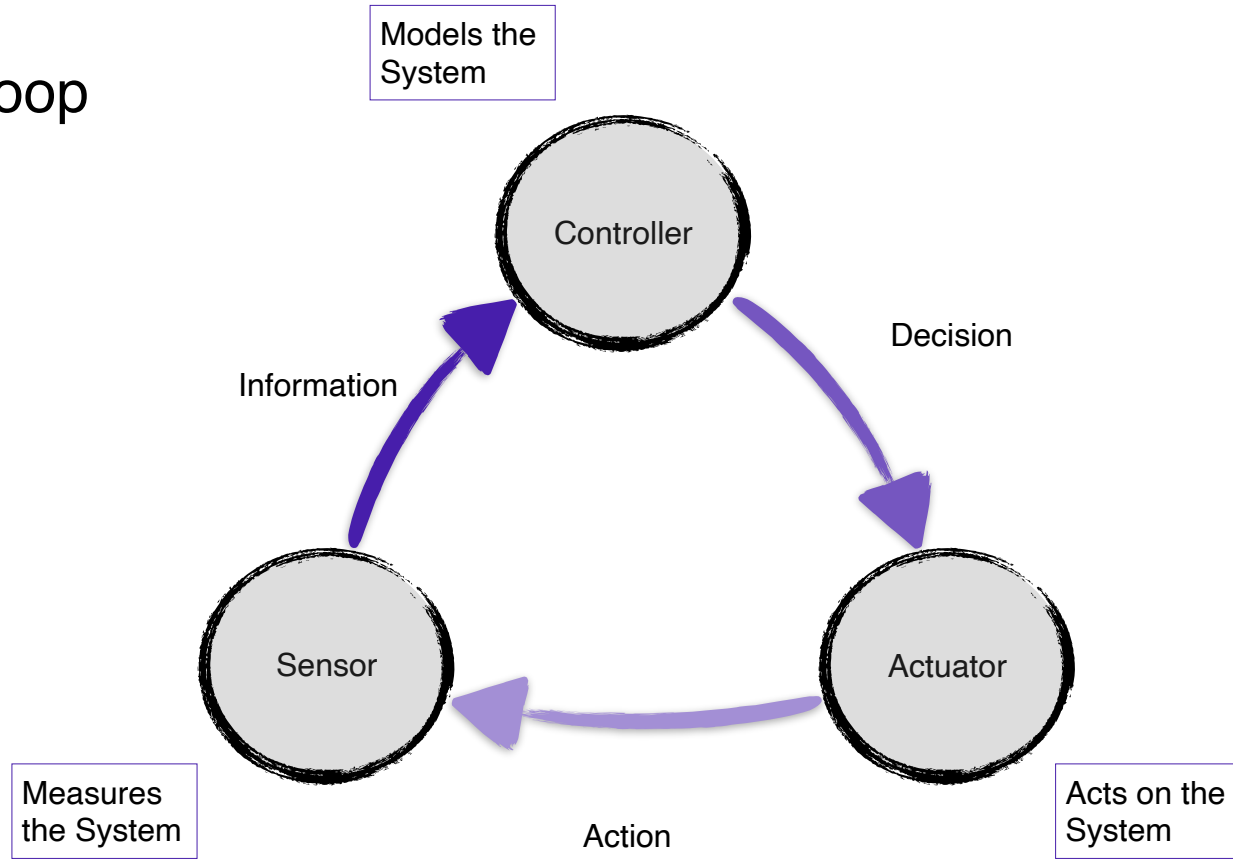
# Recognizing what “good” looks like





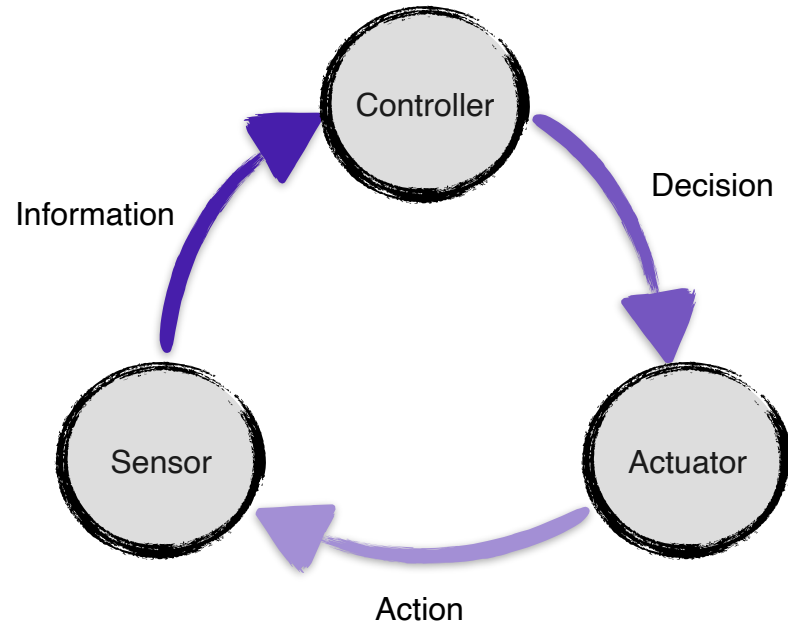


# Control Loop

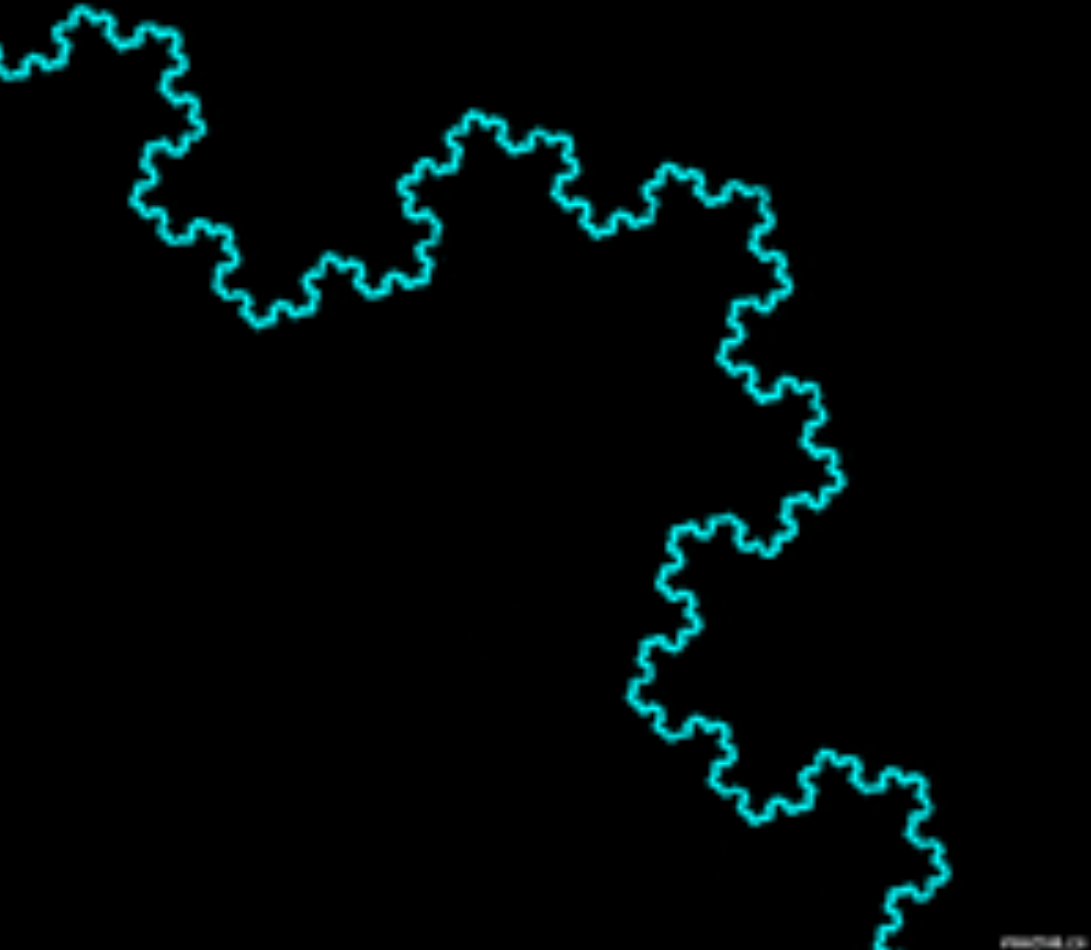


# Improving Productivity:

- Better Information
- Better Decisions
- Better Actions







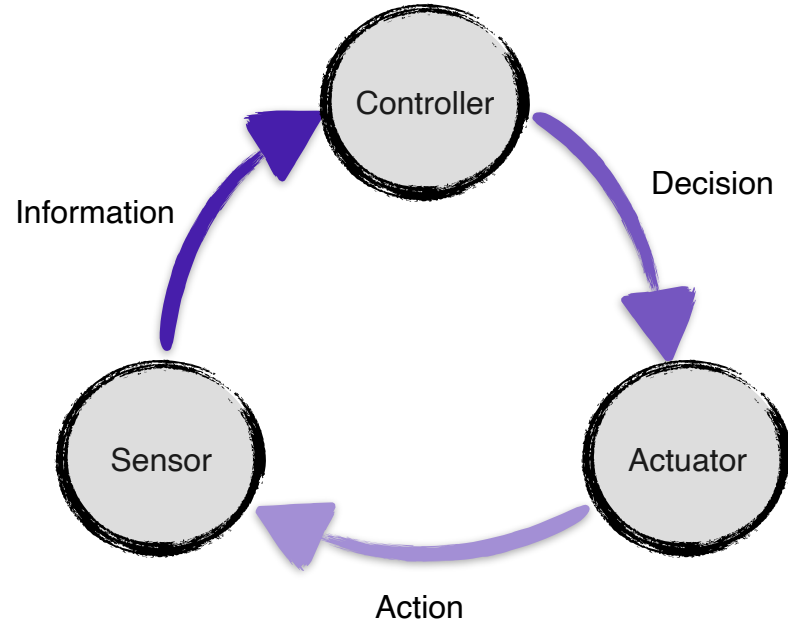
It is  
control loops  
all the way  
down





# Creator to Curator

# You are the loop!



July 23 2024

## Learning Task Decomposition to Assist Humans in Competitive Programming

**Jiixin Wen<sup>1,2</sup>, Ruiqi Zhong<sup>3</sup>, Pei Ke<sup>1,2,†</sup>, Zhihong Shao<sup>1,2</sup>,  
Hongning Wang<sup>1,2</sup>, Minlie Huang<sup>1,2,†</sup>**

<sup>1</sup>The CoAI group, Tsinghua University, Beijing, China

<sup>2</sup>Department of Computer Science and Technology, Tsinghua University, Beijing, China

<sup>3</sup> University of California, Berkeley

wenjx22@mails.tsinghua.edu.cn, aihuang@tsinghua.edu.cn

Humans are better at understanding  
**complex tasks** after they are  
decomposed into **simpler subtasks**

Human-Designed Problem

Coding Problem

GenAI Proposal

GenAI Proposed Solution

GenAI Decomposition

Initial

Heuristic

Vanilla

Assistive

Human Repair

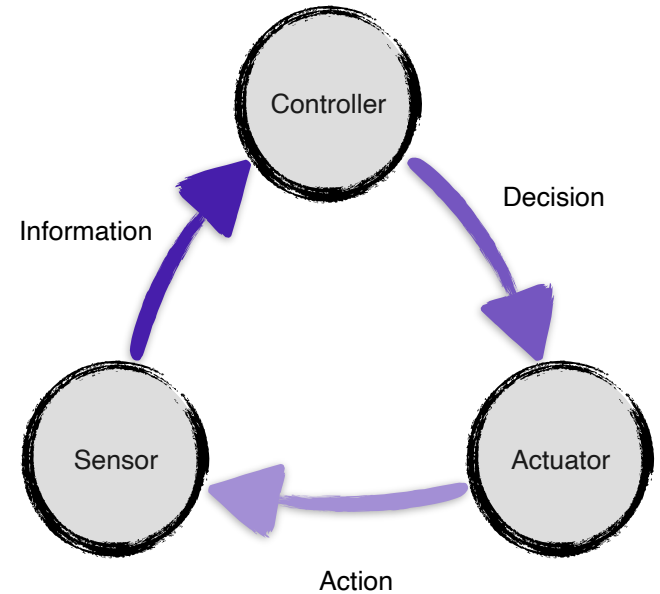
Experts

“Non-Experts”

Unit Test Evaluation

Solution Evaluation

Experts, “non-experts”, LLMs  
are better at “repairing” the  
**well-decomposed** LLM proposal



# Productivity



Ergonautic



@ergonauticly

COPYRIGHT 2024. ALL RIGHTS RESERVED

[www.ergonautic.ly](http://www.ergonautic.ly)



# Better Metrics



# Generative AI enables us to create MORE

- More code
- More text
- More media

If you ask for more [X]  
You will get more [X]

# GOODHART'S LAW

WHEN A MEASURE BECOMES A TARGET,  
IT CEASES TO BE A GOOD MEASURE

IF YOU  
MEASURE  
PEOPLE ON...

NUMBER OF  
NAILS MADE

WEIGHT OF  
NAILS MADE

THEN YOU  
MIGHT GET

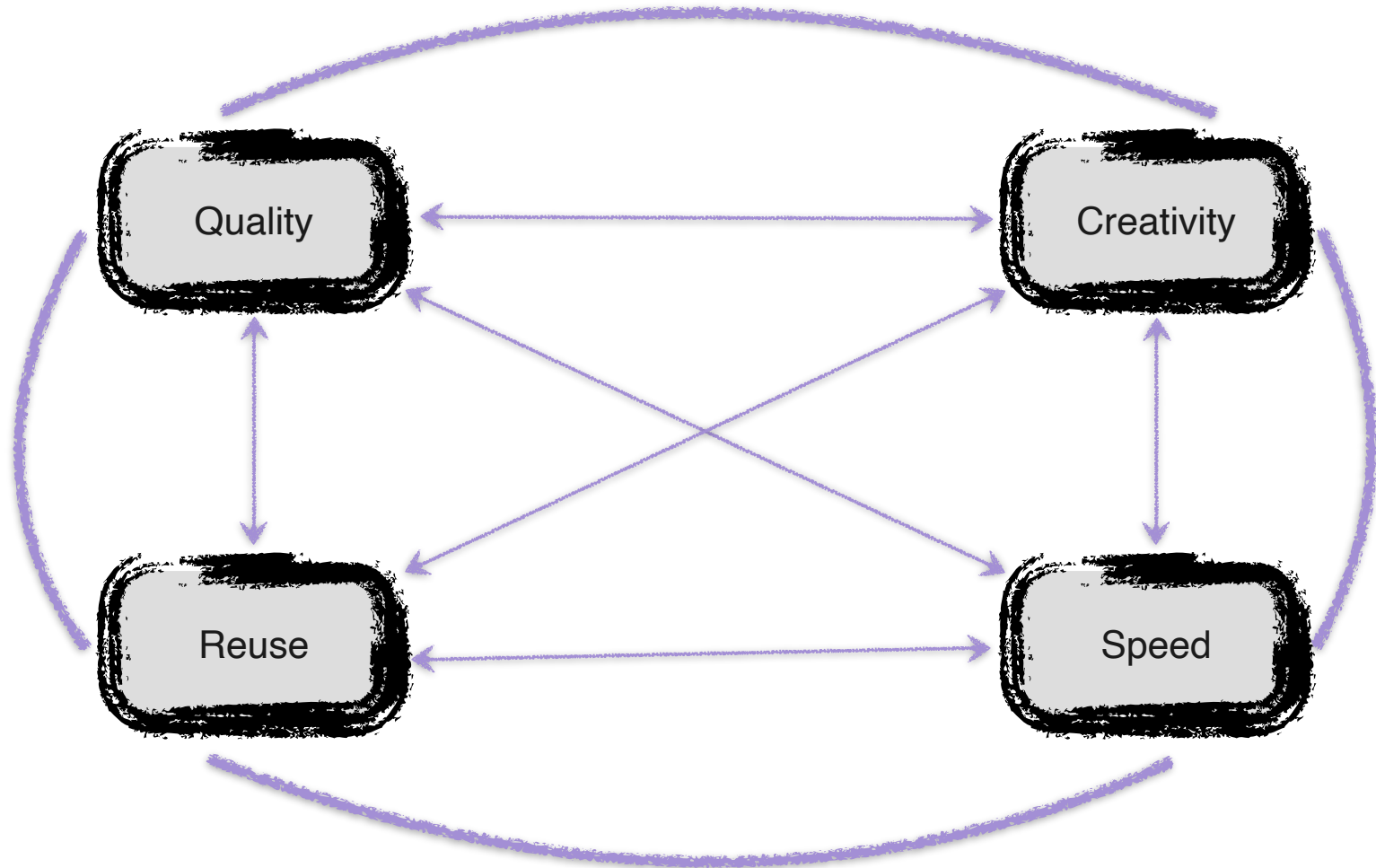
1000'S OF  
TINY NAILS

A FEW GIANT,  
HEAVY NAILS



sketchplanations

# Process



# Market Fit

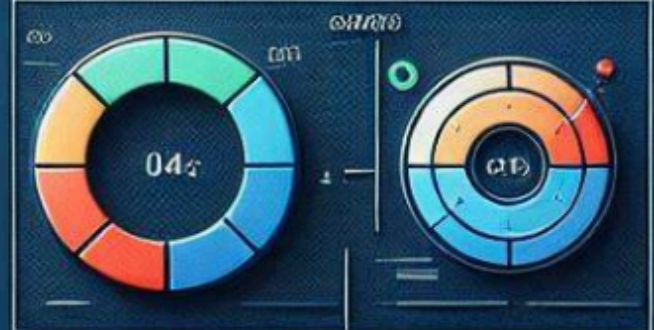
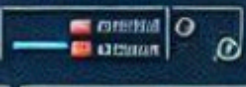
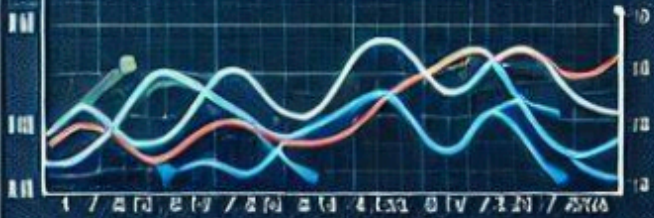
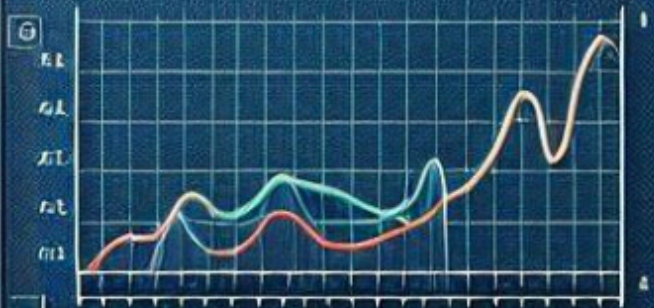
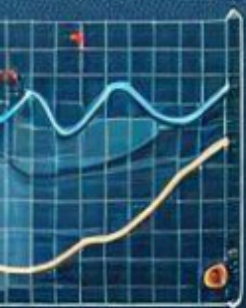
# Market Fit

- Absolute Revenue
- Relative Revenue
- Cash Flow

# Market Fit

- Customer Lifetime Value
- Customer Acquisition Cost
- Churn





30

20

10

0



30

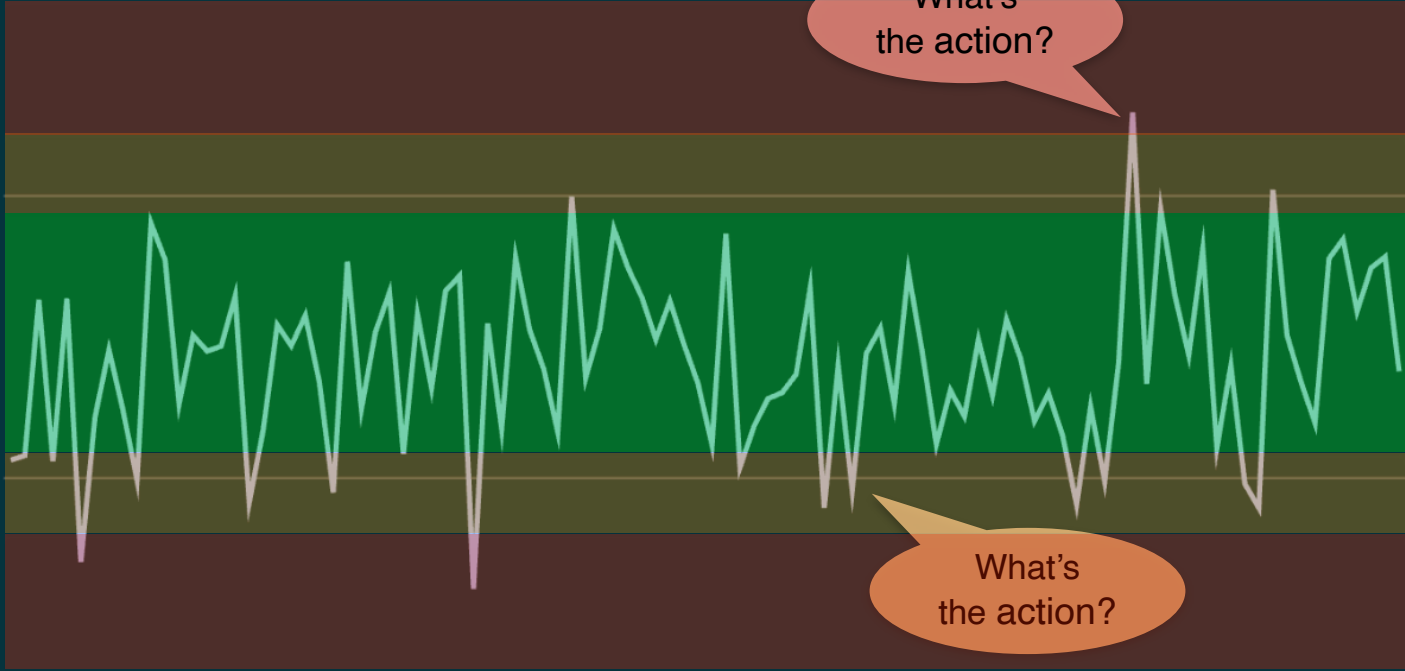
20

10

0

What's the action?

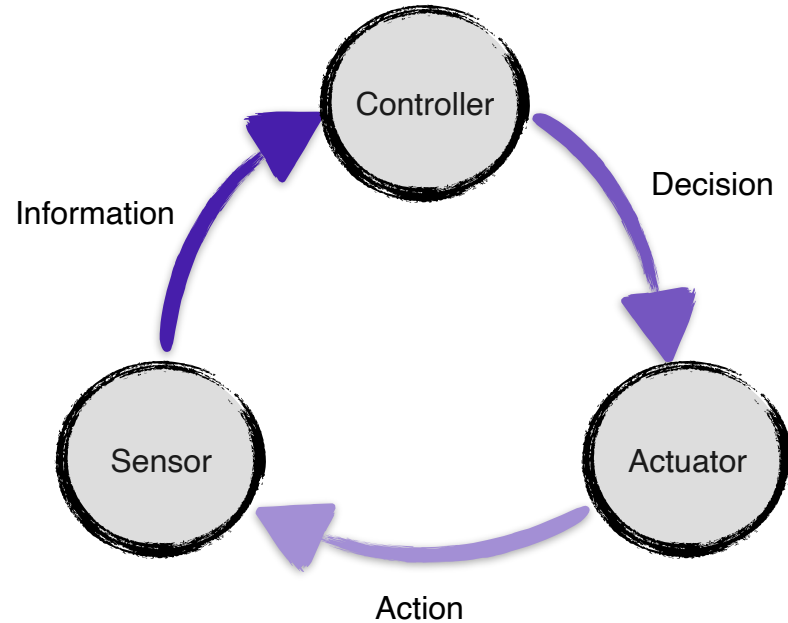
What's the action?



All systems are socio-technical

# Improving Productivity:

- Better Information
- Better Decisions
- Better Actions





Thank you!

**Sasha Czarkowski  
(Rosenbaum)**

