PROFESSIONAL EDUCATION
Digital Plus+ Programs

Applied Generative-AI for Digital Transformation

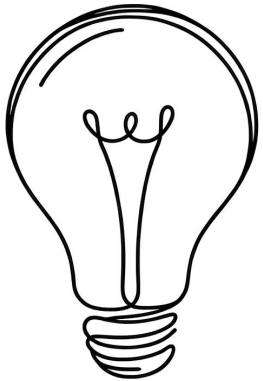
AI Enabled Economy

Abel Sanchez, John R. Williams

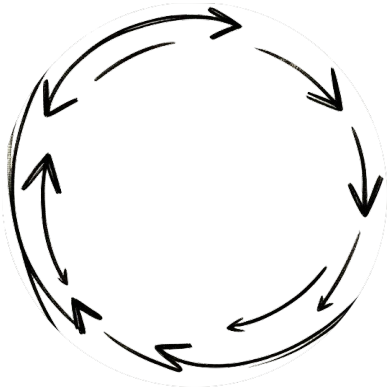
MIT Campus Course | Cambridge, MA

What is it?

Create

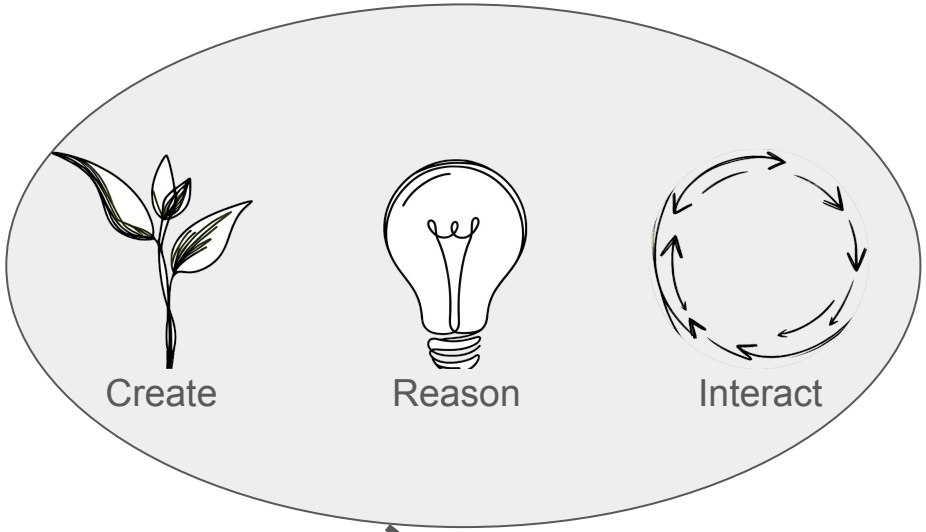


Reason



Interact

Why do we care?

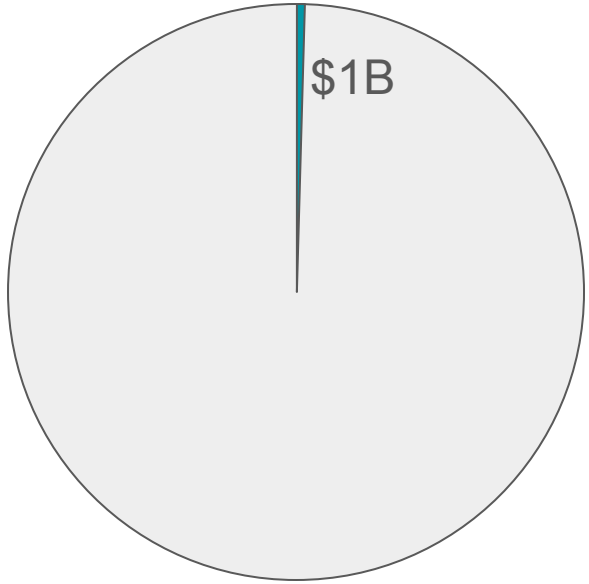


Opportunity to replace services with software

Marketplace

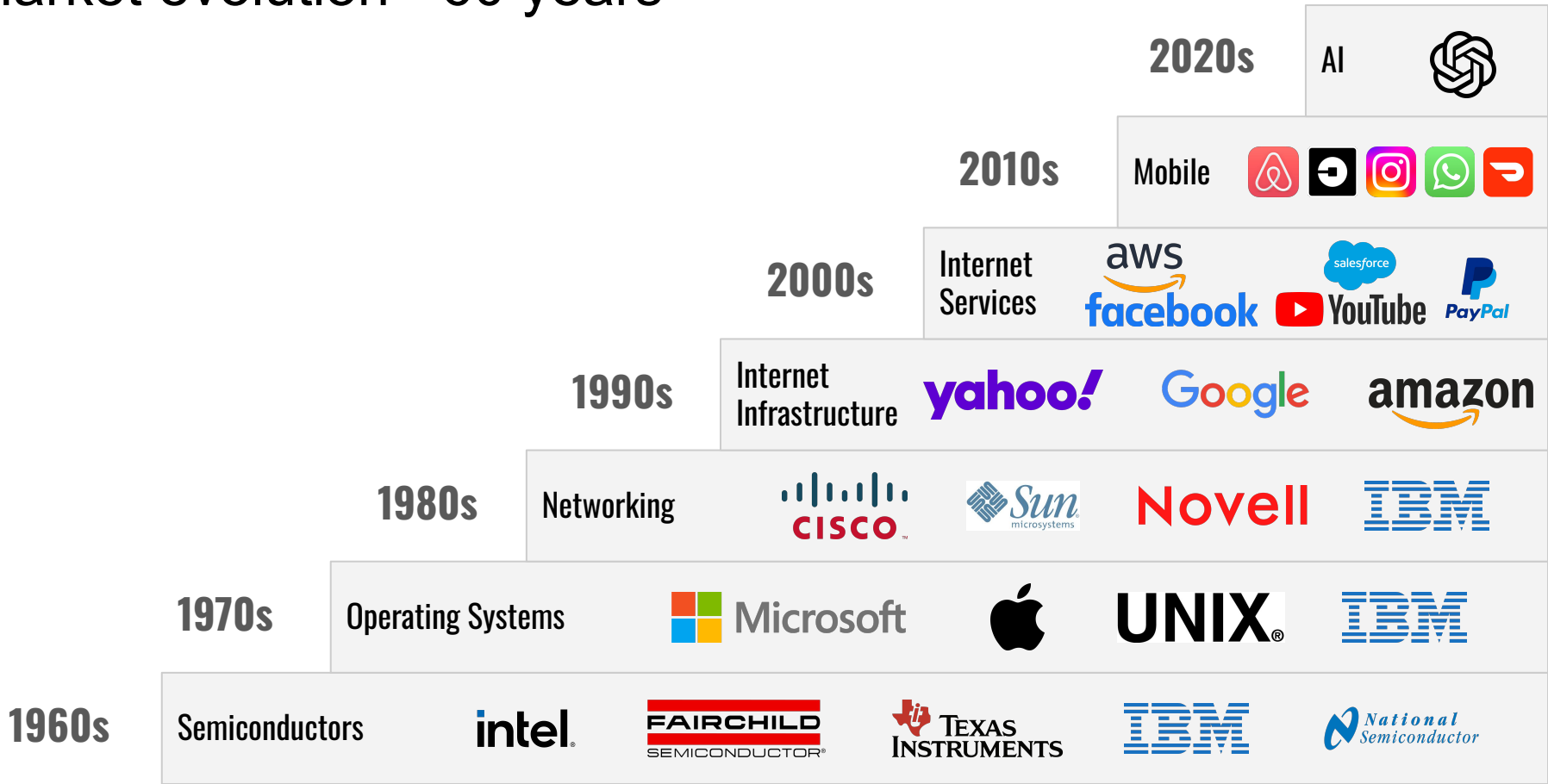
In 2010
$350B all software
$6B cloud

Present
$650B all software
$400B cloud
Over 40% growth/yr



$15,000B Services

Market evolution - 60 years



What is missing?

	Cloud	Mobile	AI
Apps	Dropbox, asana, Notion, box, zoom, DocuSign, salesforce, Adobe		ChatGPT
Developer	GitHub, docker, Unity, twilio		API
Data	splunk>, snowflake, ORACLE, databricks		
Security	okta, paloalto NETWORKS, CROWDSTRIKE		
Infrastructure	Google Cloud, aws, Azure		NVIDIA, Google Cloud, aws, Azure

AI Everywhere

State of the Practice

How enterprises are building and buying Gen-AI

- Generative AI **budgets skyrocketing**
- Leaders **reallocating AI investments** to recurring software budget lines (IT).
- ROI is still an art and a science.
- Implementing and scaling generative **AI requires the right technical talent**, which currently isn't in-house for many enterprises.
- A multi-model future
- **Open source is booming.**
- While cost factored into open source appeal, it ranked below **control and customization as key selection criteria**.
- Desire for control stems from sensitive use cases and enterprise data security concerns.
- Leaders generally customize models through **fine-tuning instead of building models from scratch**.
- Cloud is still highly influential in model purchasing decisions.
- Customers still care about early-to-market features.
- That said, most enterprises think **model performance is converging.**
- Optimizing for optionality.
- **Enterprises are building, not buying, apps**—for now.
- Enterprises are excited about internal use cases but remain more **cautious about external ones** (apps).
- Enterprise spend projected at **$5B in 2024**

https://a16z.com/generative-ai-enterprise-2024/

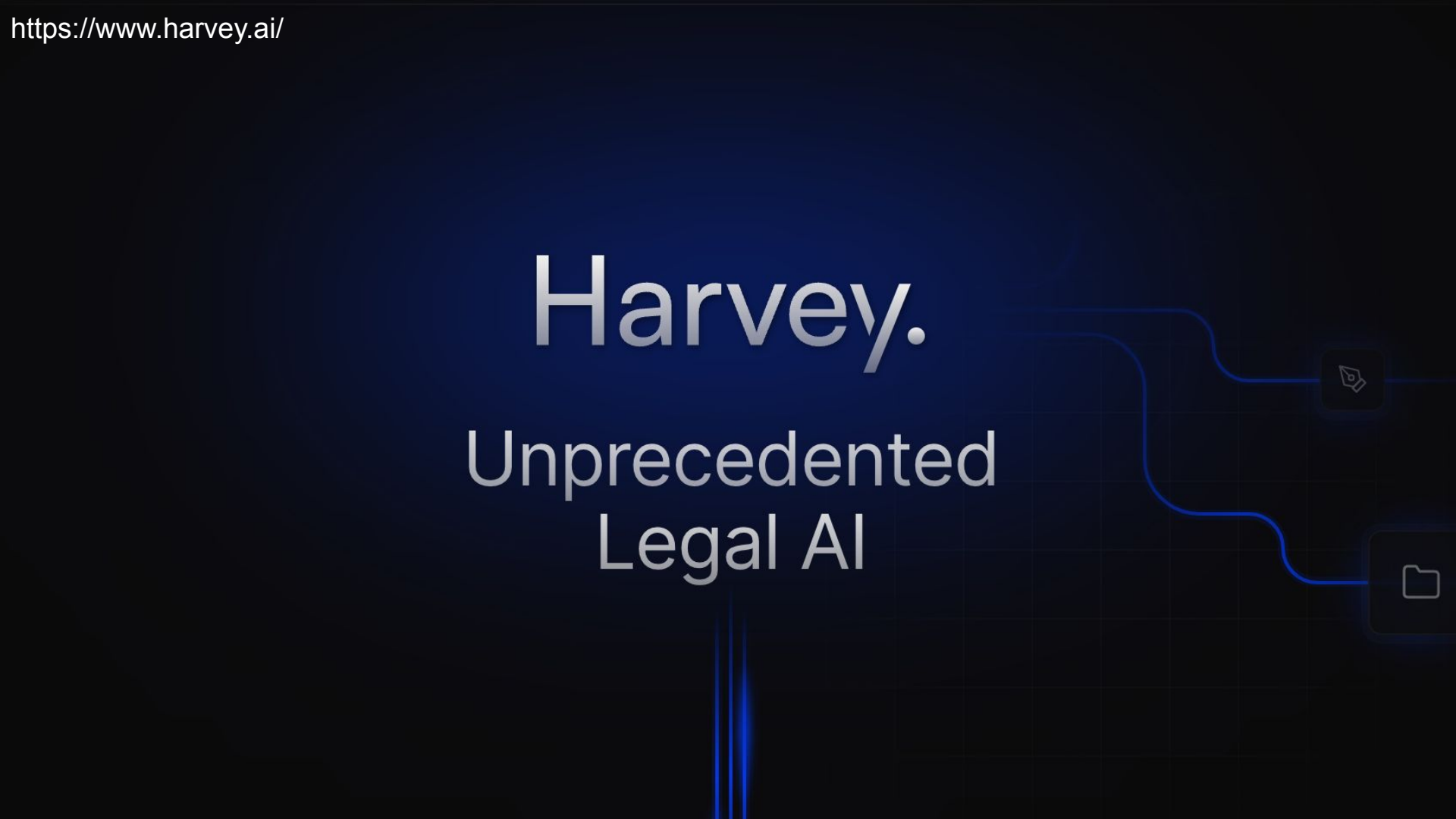
Klarna AI assistant handles two-thirds of customer service chats

- 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, leading to a **25% drop in repeat inquiries**
- Customers resolve errands in less than **2 mins compared to 11 mins** previously
- Available in **23 markets, 24/7** and communicates in more than 35 languages
- Estimated **$40 million USD in profit** improvement to Klarna in 2024

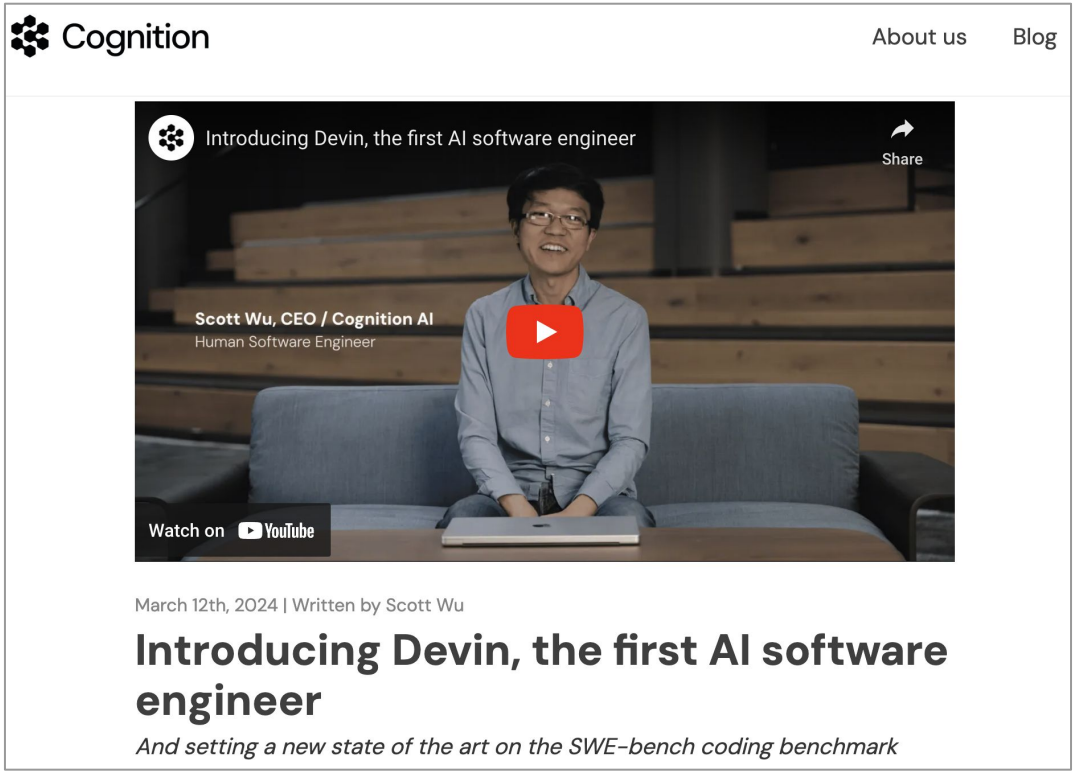
Cost Comparison - Human vs Gen-AI

- Graphic artist image
  - Gen-AI cost: ~$.001, takes ~1 second
  - Human cost: $100s min, takes hours or days
  - 100,000 times cheaper (assuming $100)
  - 3,600 times faster
- Lawyer, summarize and answer questions on a complex legal brief
  - Gen-AI cost: ~$.001, takes ~10 second
  - Human cost: $100s min, takes hours or days
- Therapist
  - Gen-AI cost: ~$.001, takes ~10 second
  - Human cost: $100s, takes 1 hour
  - Availability: scalable and global

https://www.harvey.ai/



Introducing Devin, the first AI software engineer



https://youtu.be/fjHtjT7GO1c

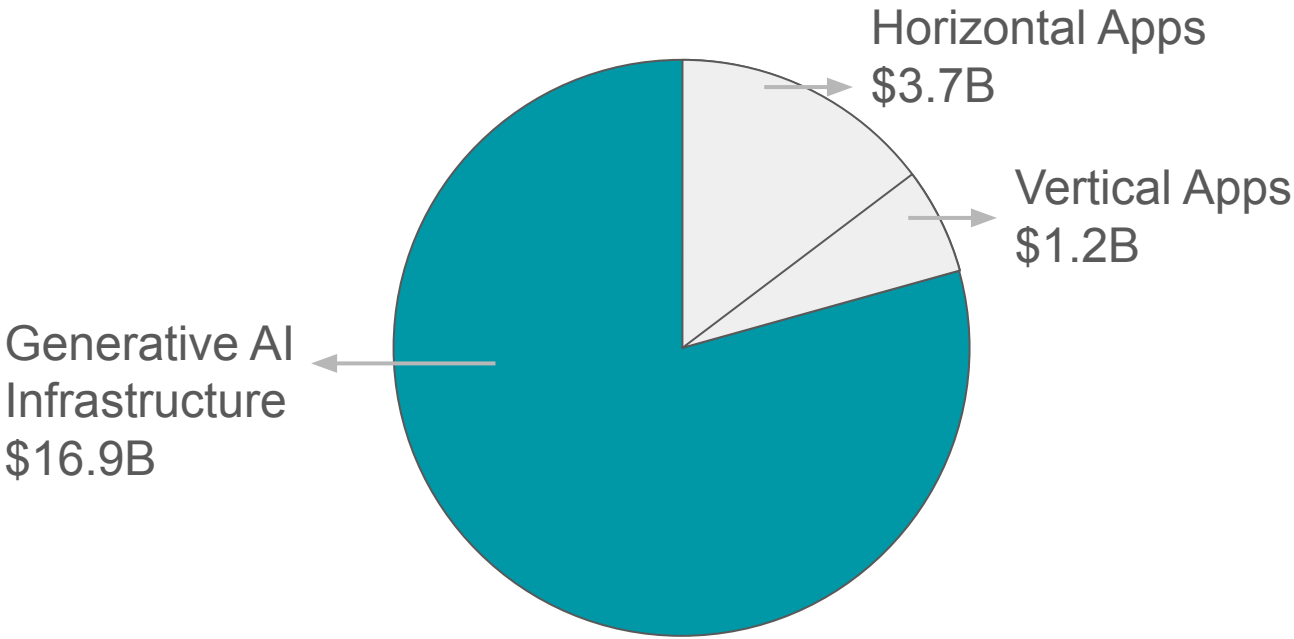
https://www.cognition-labs.com/introducing-devin
https://twitter.com/cognition_labs/status/1767548763134964000

AI 10x faster than SaaS to $3B USD

Caveats

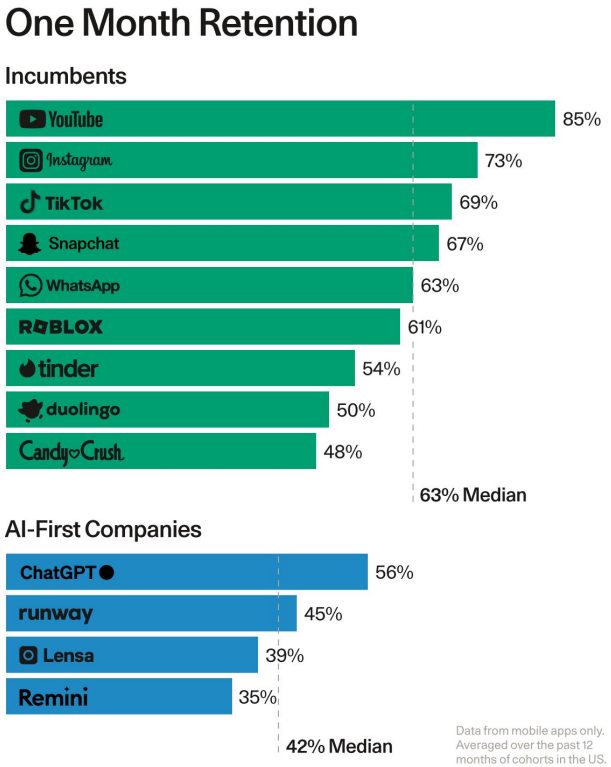
Infrastructure getting more funding than applications - ROI?

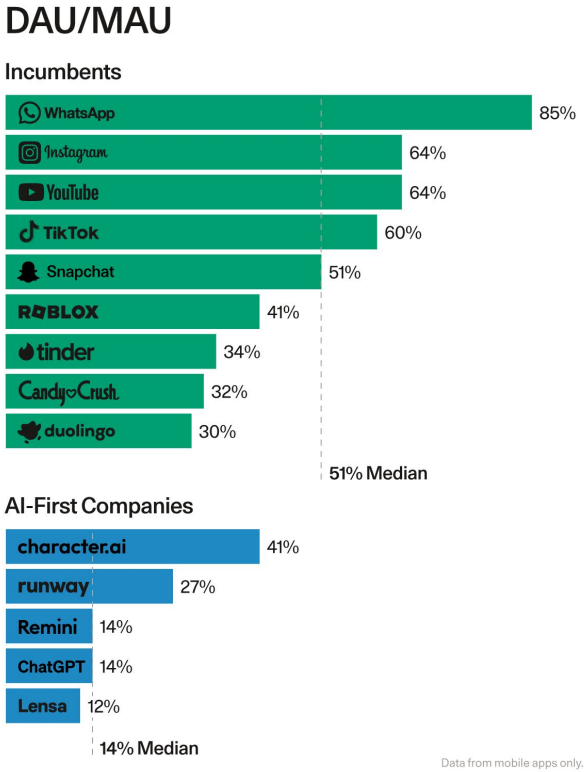
$50B on NVidea GPUs in 2023



CB Insights - Generative AI Funding

Retention - user disillusionment





Daily Active Users (DAU) to Monthly Active Users (MAU)

Sequoia, Gen-AI's Act Two, https://www.sequoiacap.com/article/generative-ai-act-two/

Good news, lots of competition, capabilities increasing

	Claude 3 Opus	Claude 3 Sonnet	Claude 3 Haiku	GPT-4	GPT-3.5	Gemini 1.0 Ultra	Gemini 1.0 Pro
Undergraduate level knowledge *MMLU*	**86.8%** 5 shot	79.0% 5-shot	75.2% 5-shot	86.4% 5-shot	70.0% 5-shot	83.7% 5-shot	71.8% 5-shot
Graduate level reasoning *GPQA, Diamond*	**50.4%** 0-shot CoT	40.4% 0-shot CoT	33.3% 0-shot CoT	35.7% 0-shot CoT	28.1% 0-shot CoT	—	—
Grade school math *GSM8K*	**95.0%** 0-shot CoT	92.3% 0-shot CoT	88.9% 0-shot CoT	92.0% 5-shot CoT	57.1% 5-shot	94.4% Maj1@32	86.5% Maj1@32
Math problem-solving *MATH*	**60.1%** 0-shot CoT	43.1% 0-shot CoT	38.9% 0-shot CoT	52.9% 4-shot	34.1% 4-shot	53.2% 4-shot	32.6% 4-shot
Multilingual math *MGSM*	**90.7%** 0-shot	83.5% 0-shot	75.1% 0-shot	74.5% 8-shot	—	79.0% 8-shot	63.5% 8-shot
Code *HumanEval*	**84.9%** 0-shot	73.0% 0-shot	75.9% 0-shot	67.0% 0-shot	48.1% 0-shot	74.4% 0-shot	67.7% 0-shot
Reasoning over text *DROP, F1 score*	**83.1** 3-shot	78.9 3-shot	78.4 3-shot	80.9 3-shot	64.1 3-shot	82.4 Variable shots	74.1 Variable shots
Mixed evaluations *BIG-Bench-Hard*	**86.8%** 3-shot CoT	82.9% 3-shot CoT	73.7% 3-shot CoT	83.1% 3-shot CoT	66.6% 3-shot CoT	83.6% 3-shot CoT	75.0% 3-shot CoT
Knowledge Q&A *ARC-Challenge*	**96.4%** 25-shot	93.2% 25-shot	89.2% 25-shot	96.3% 25-shot	85.2% 25-shot	—	—
Common Knowledge *HellaSwag*	**95.4%** 10-shot	89.0% 10-shot	85.9% 10-shot	95.3% 10-shot	85.5% 10-shot	87.8% 10-shot	84.7% 10-shot

Success takes time - same with internet (e.g. Google)

Web Busts

yahoo!
Netscape
AOL
excite
LYCOS

1990s

2007

Top Ten Apps

2008

Winners

2010

First sucesses - they may pale in the future

Customer
Support

Interactive
AI browser
Chatbot

Enterprise
Knowledge

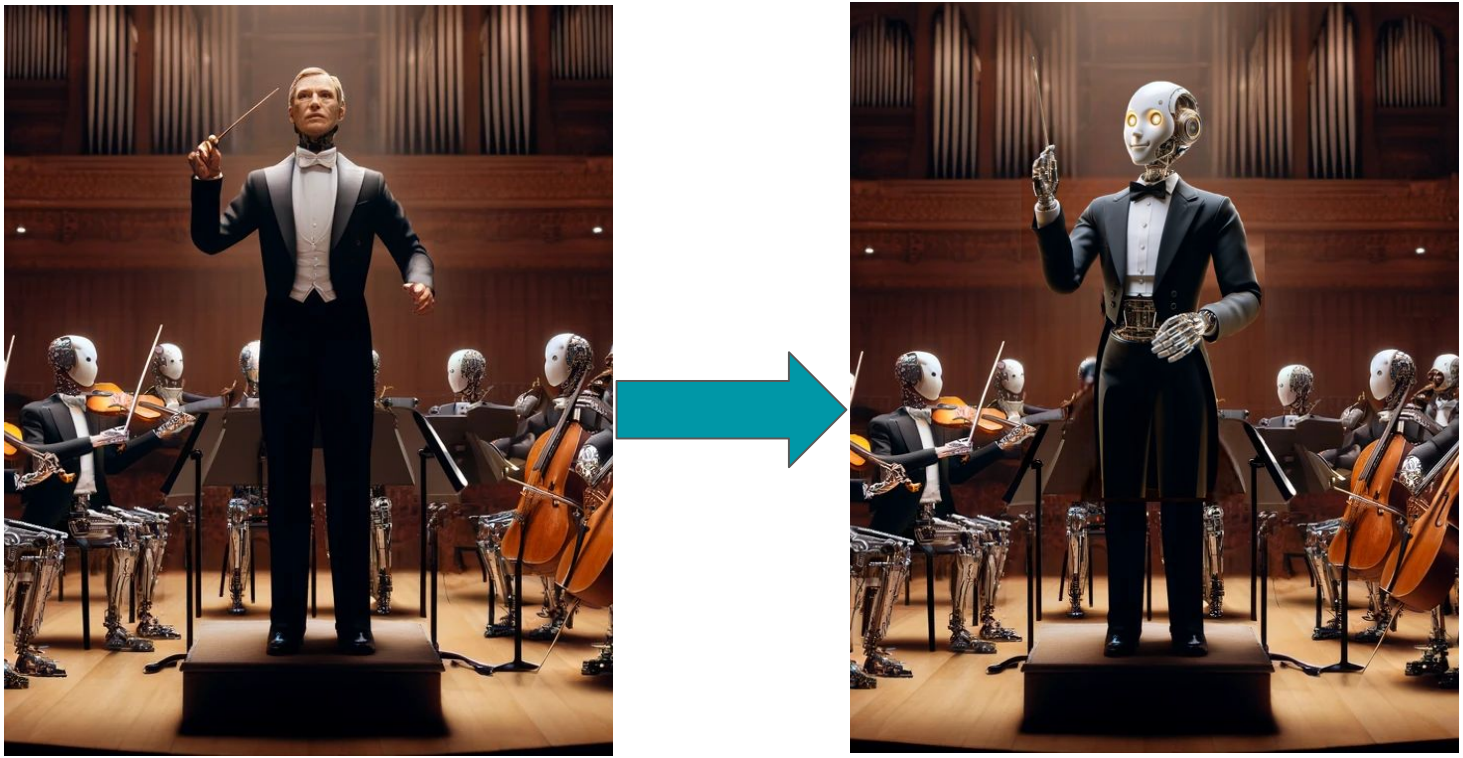
Predictions

Agentic Resoning

Prototyping
to production

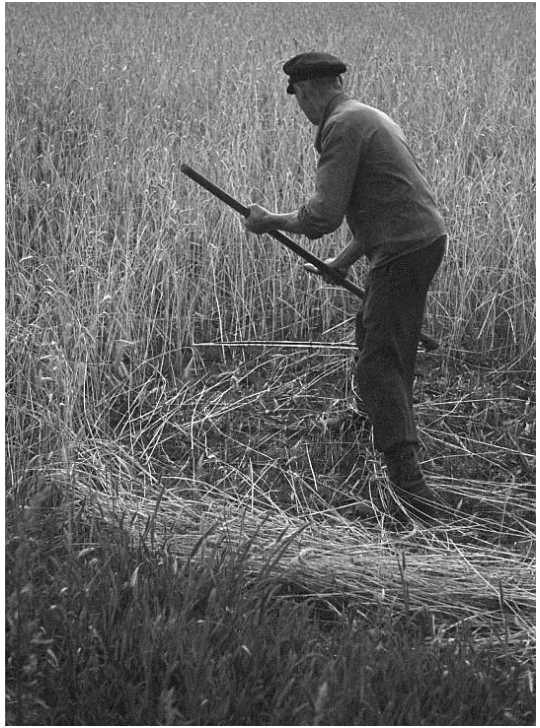
Reliability and
robustness

Predictions - human coordination to agents (AI) coordination. Agentic Reasoning.



Looking forward

Agriculture evolution



Human + Tool

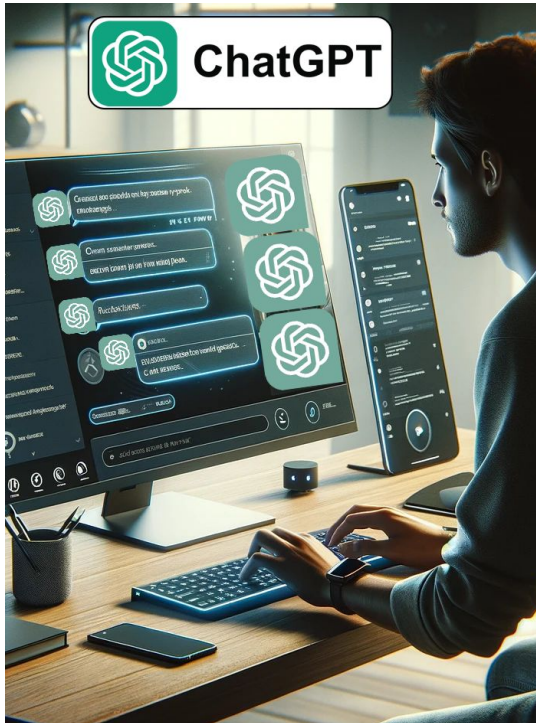
Human + Machine

Data, AI, Network

Computation evolution

Human + Tool

Human + Machine



Data, AI, Network

AI Transformation

Tool

Human plus the tool

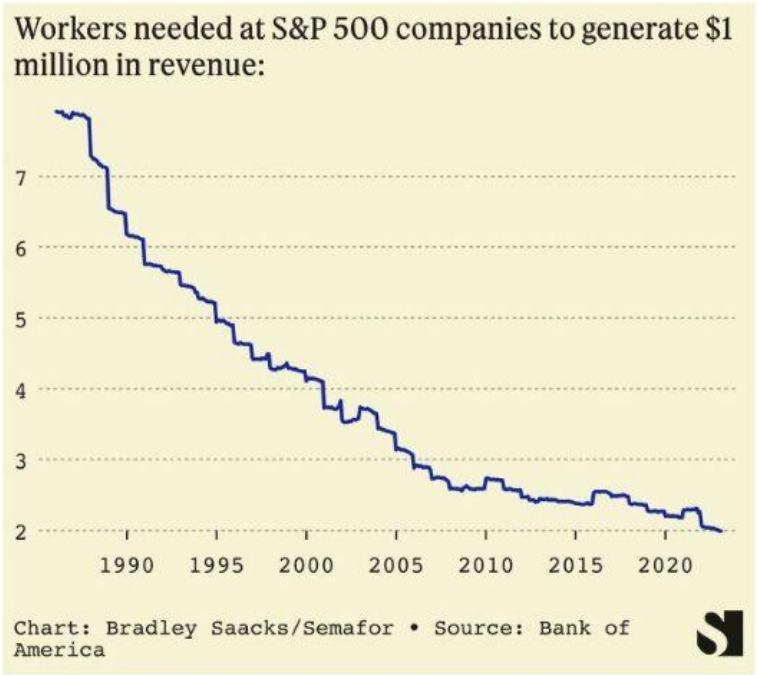
Machine

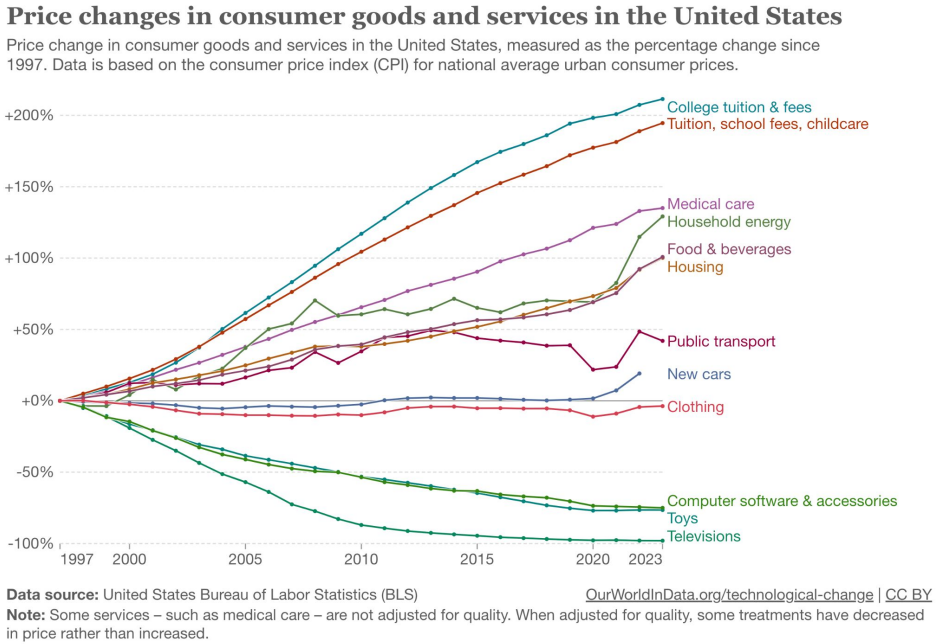
Human plus a collection of tools

Data, AI, Networks

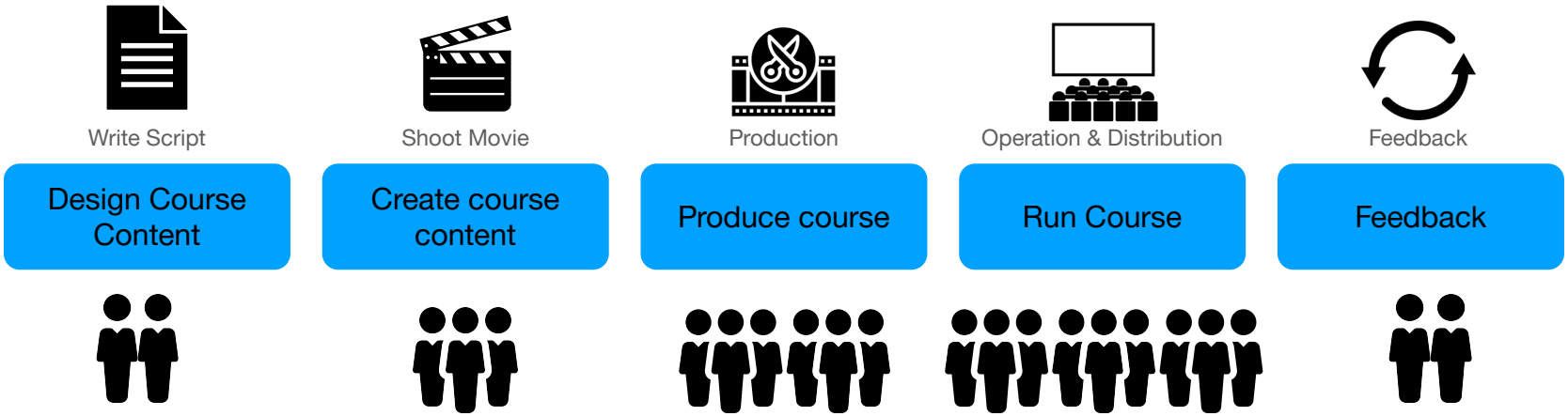
Collection of tools plus data, plus networks (llms,cloud)

Price Reductions

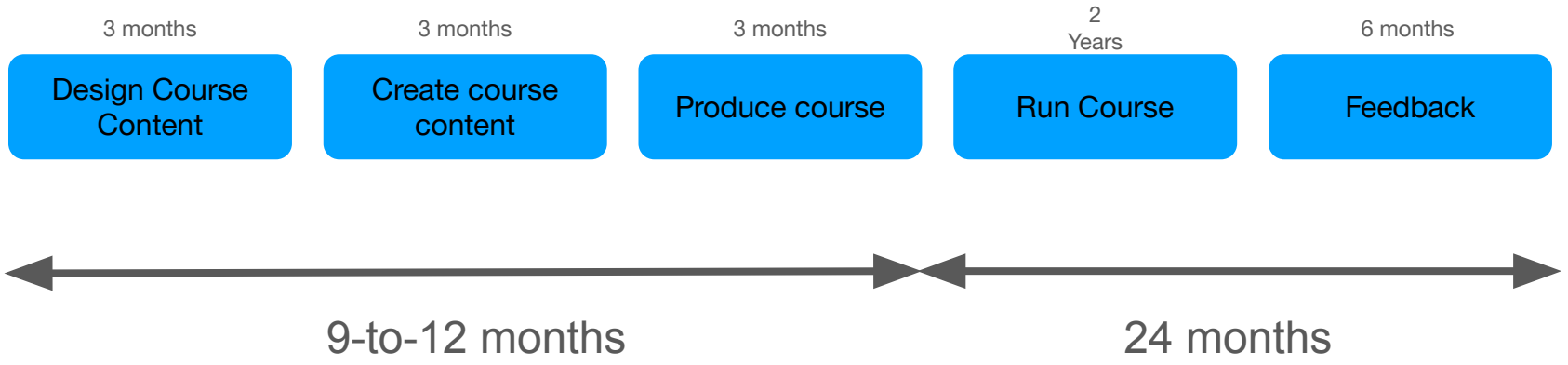


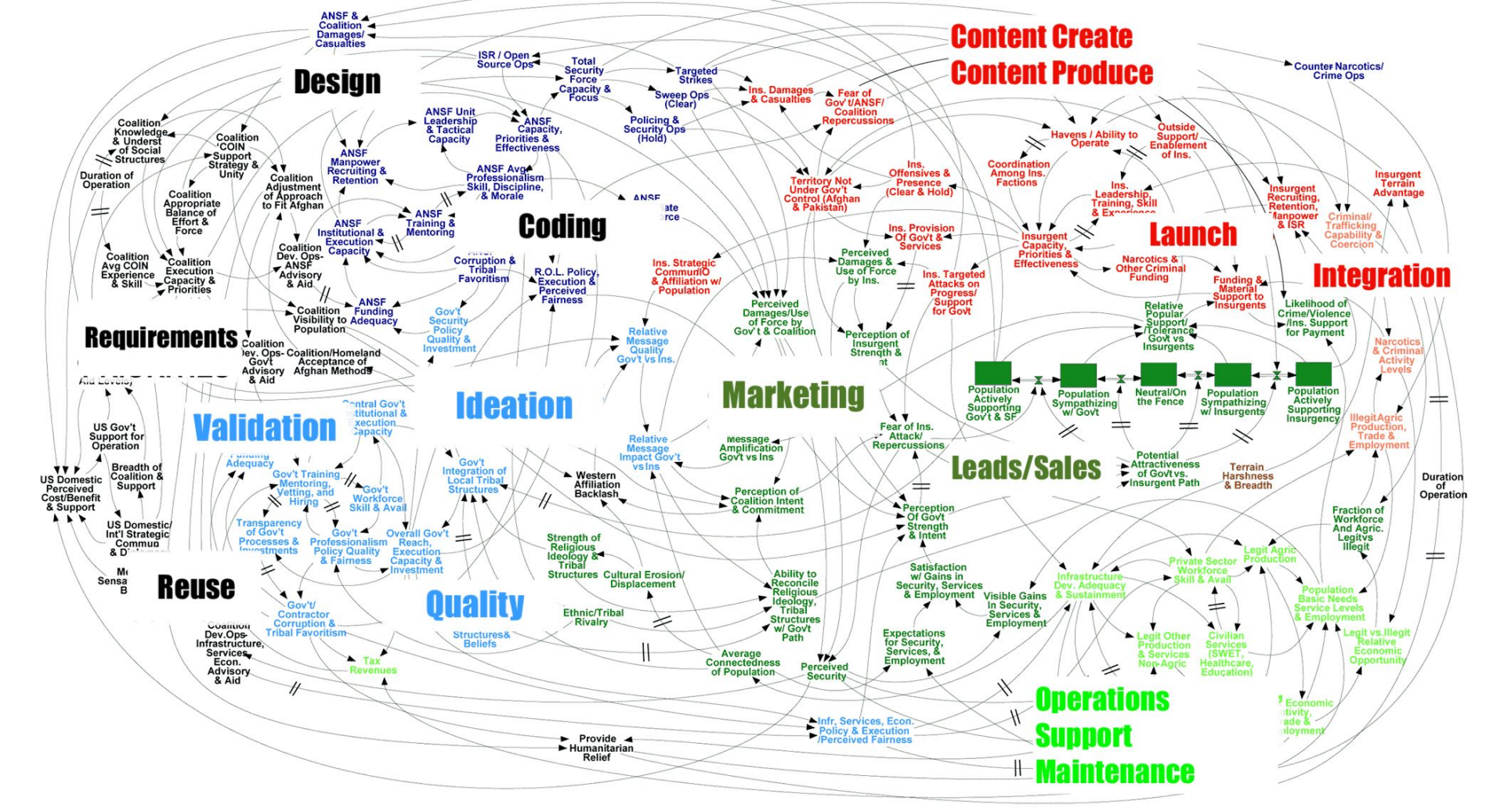


You can think of course production like film making



Timeline


3 months
Design Course Content
3 months
Create course content
3 months
Produce course
2 Years
Run Course
6 months
Feedback
9-to-12 months
24 months



Adapted from, https://www.nytimes.com/2010/04/27/world/27powerpoint.html

AI Process Transformation

Amplify

Increase Process Productivity

Tool: chat bot

Reimagine

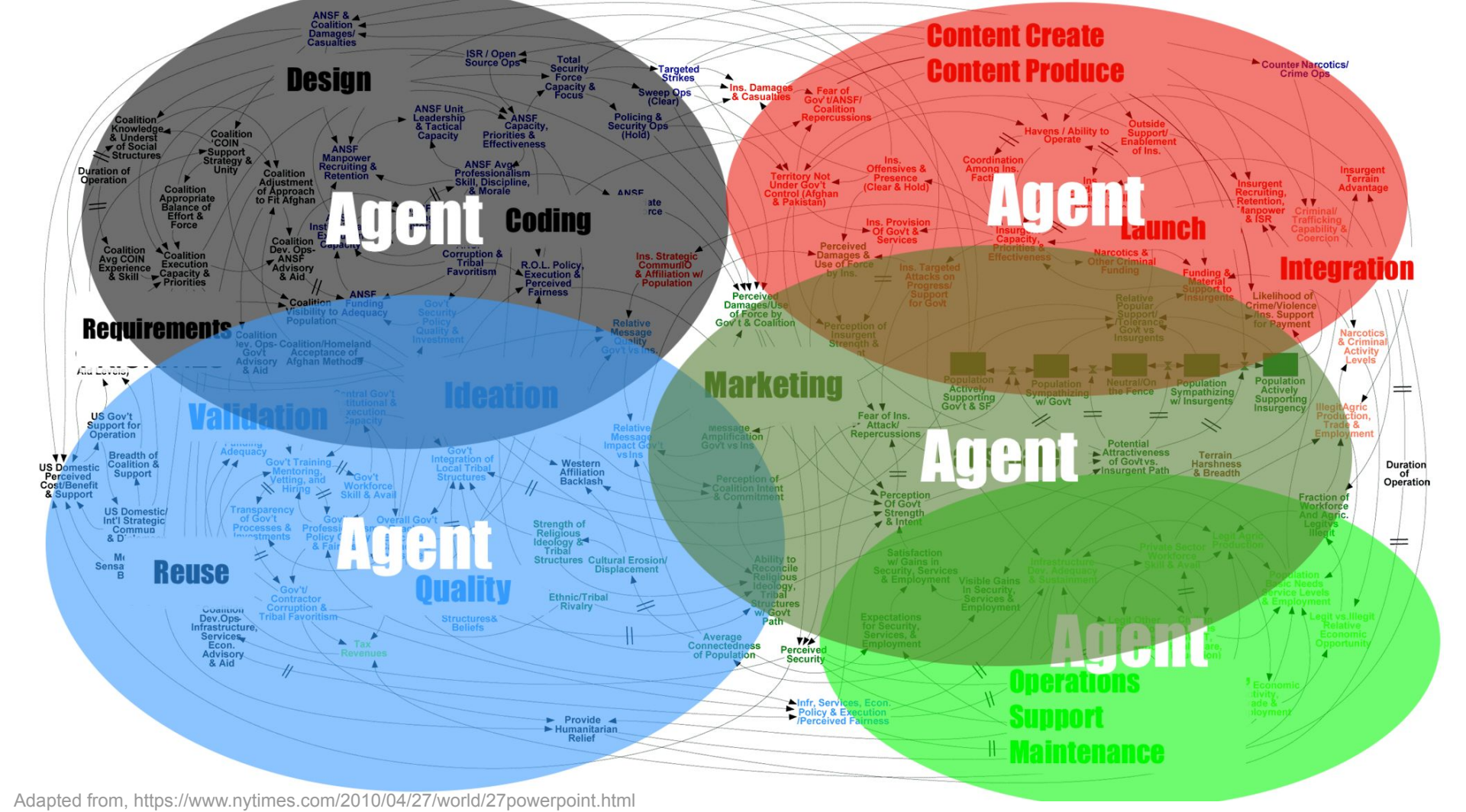
Transform Processes & Business Model

Machine: knowledge base, auto grading

Autonomous

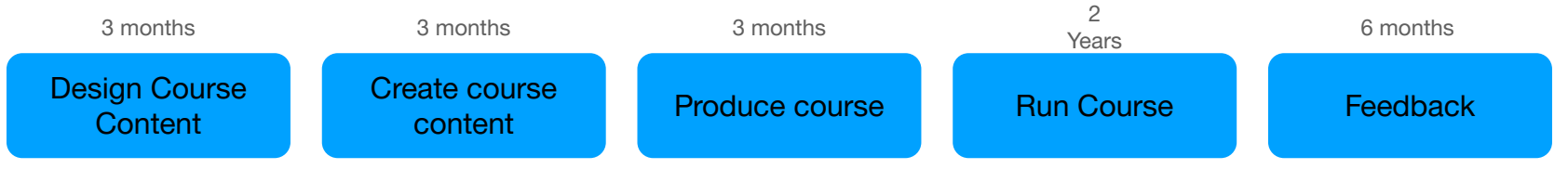
Business/Product Fully Automated with AI.

Fully autonomous



Adapted from, https://www.nytimes.com/2010/04/27/world/27powerpoint.html

Timeline



Collapse to a few hours?

- Automated market research
- Automated course design
- Automated course syllabus

- Automated lecture scripts
- Automated exercise creation
- Automated knowledge base
- Automated reference material
- Train avatars
- Give lectures, record

Fully automated

- Automated marketing
- Automated video promotion
- Automated course administration
- Automated course updates
- Automated content refresh

- Automated Operations
- Automated Delivery
- Automated Support
- Automated Q&A

Abel Corp?

The self-directed company

Enabling the one person company?

Enterprise Strategy I

When will AI be adopted by the enterprise?

- High data privacy concerns
- High IP retention concerns
- Argument: my data is special
- Parallels to cloud in 2010
- Education needed - understand the limitations
- Is cloud storage safe?
    - Nothing has had more high profile leaks than S3

Turf wars

- Data is power
- Classic debate of public cloud vs internal data center
- Stakeholders protecting turf
- Classic journey of data maturity
  - Multi-year internal build
  - Months in the public cloud
- The cloud has proven itself
  - Biggest IT spend globally

Generative-AI, and technology, can drive growth, enhance productivity and improve customer and employee experiences.

CIO ➜ CEO

- 10 years ago the CIO reached out
- Today is the CEO
  - Discuss strategy
    - how can I reinvent what I do with tech
  - Discuss data
    - Mindset: maybe with Gen-AI and my data I can beat the competition
  - I want to do it myself
    - I do not want to give it to OpenAI, Anthropic, etc
    - I want to own the IP
    - I have line out my office of my people saying they can do it

Can you do it yourself?

- Yes, but it’s hard, you need elite talent, access to hardware, and lots of cash
- Requires lots of GPUs
- Vendors have already figured out how to do it at scale
  - OpenAI & Scale, Microsoft
  - Anthropic, Amazon & Google
  - AWS, GCP, Azure
  - Databricks, Mosaic

Strategy - Parallels to the internet

- Cisco $224M in 1990 ➜ by 2000 $500B, bypassing Microsoft
- Today Microsoft $2.7T ➜ Cisco $193B
- Hypercompetitive: foundational models, developer platform & tools
- Big opportunity: Apps on top of LLM platforms

VISUALIZED

Forbes SEQUOIA MERITECH

Apps

CONSUMER USES

ENTERTAINMENT
character.ai
Midjourney
Pika

PRODUCTIVITY
ChatGPT
Claude
DeepL
perplexity
Notion
tome

ENTERPRISE STACK

GENERAL PRODUCTIVITY
ChatGPT
DeepL
glean
tome
ADEPT
Claude
Notion
WRITER

LEARNING & DEVELOPMENT
Sana
synthesia

CUSTOMER EXPERIENCE
CRESTA
SIERRA

DEVELOPER & DATA TEAMS
AssemblyAI
kumo
codeium
GitHub*

INDUSTRY VERTICALS

CREATIVE
Midjourney
Photoroom
runway
Rosebud AI
Pika
IIElevenLabs
Leonardo.Ai

DEFENSE
ANDURIL
VANNEVAR Labs

HEALTH & BIO
ABRIDGE
Cradle
insitro
OWKIN

INDUSTRIAL
FIGURE
TRACTIAN
waabi

PROFESSIONAL SERVICES
Harvey.
Hebbia

Infrastructure

INFERENCE PROVIDERS
anyscale
databricks
Replicate
together.ai
baseten

APP DEV FRAMEWORKS
LangChain

MODEL HUBS
Hugging Face
Replicate

FOUNDATION MODEL PROVIDERS
ANTHROPIC
cohere
MISTRAL AI_
OpenAI

STORE & COMPUTE

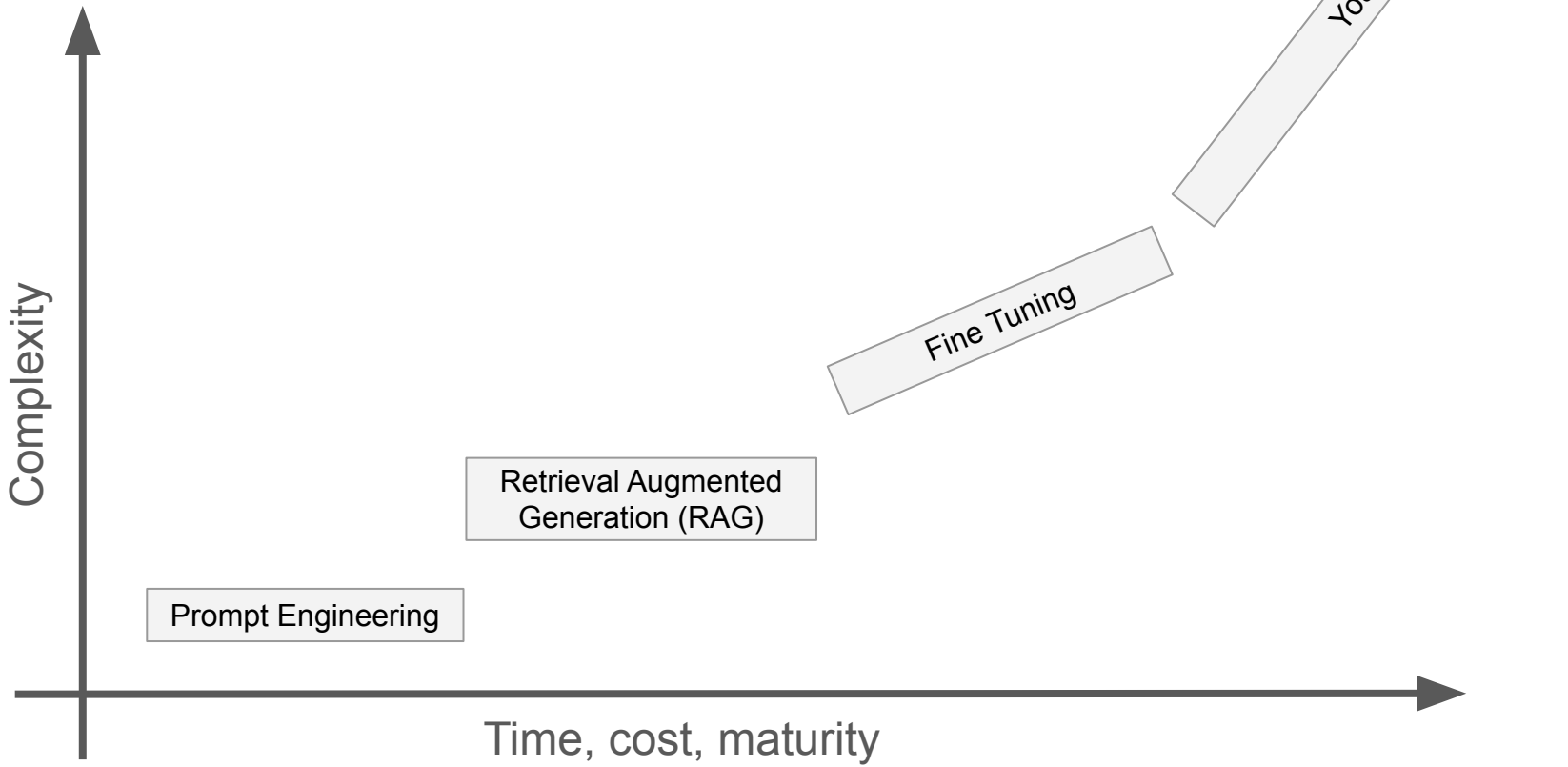
LABEL / PROCESS DATA
Cleanlab
scale
UNSTRUCTURED

CLOUD DATA PROVIDERS
databricks
Pinecone
Weaviate
MongoDB*
snowflake*

CLOUD SERVICE PROVIDERS
Google Cloud*
aws*
Azure*

HARDWARE
cerebras
NVIDIA*
AMD*
intel*
(etc.)

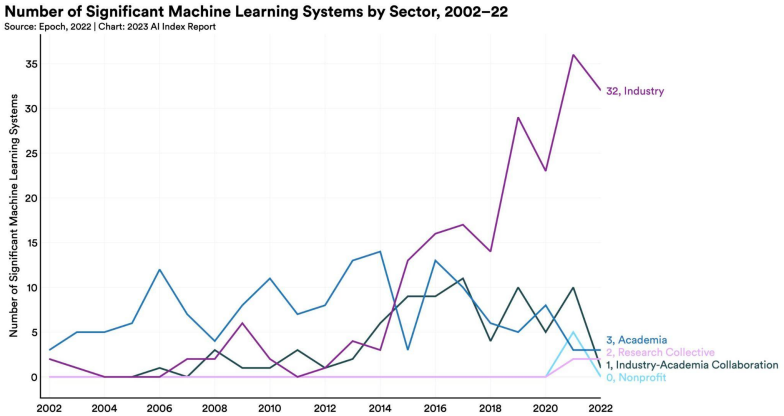
Maturity evolution

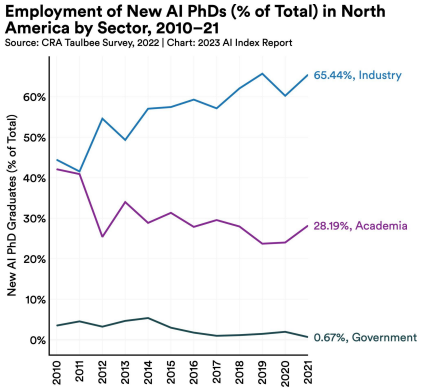


Enterprise Strategy II

Why is university struggling?

- Funding: AI research is too expensive
- Too hard to get GPUs
- Talent drain, big money in industry
- Industry is innovating faster
- Talent seeks other talent
- The same is true for you!





Figures from https://aiindex.stanford.edu/report/

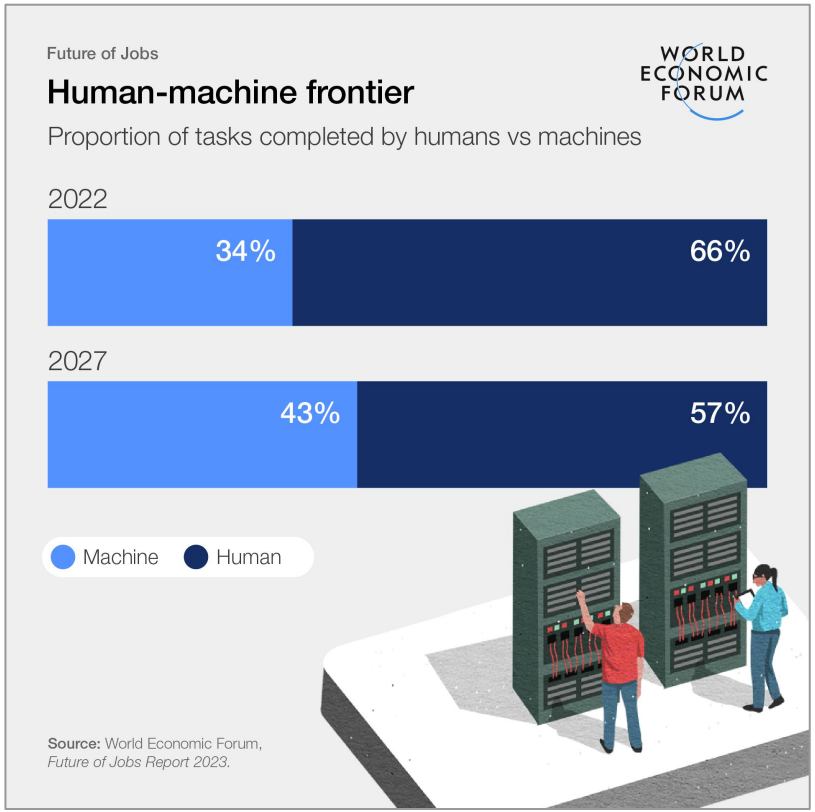
The AI-enabled economy

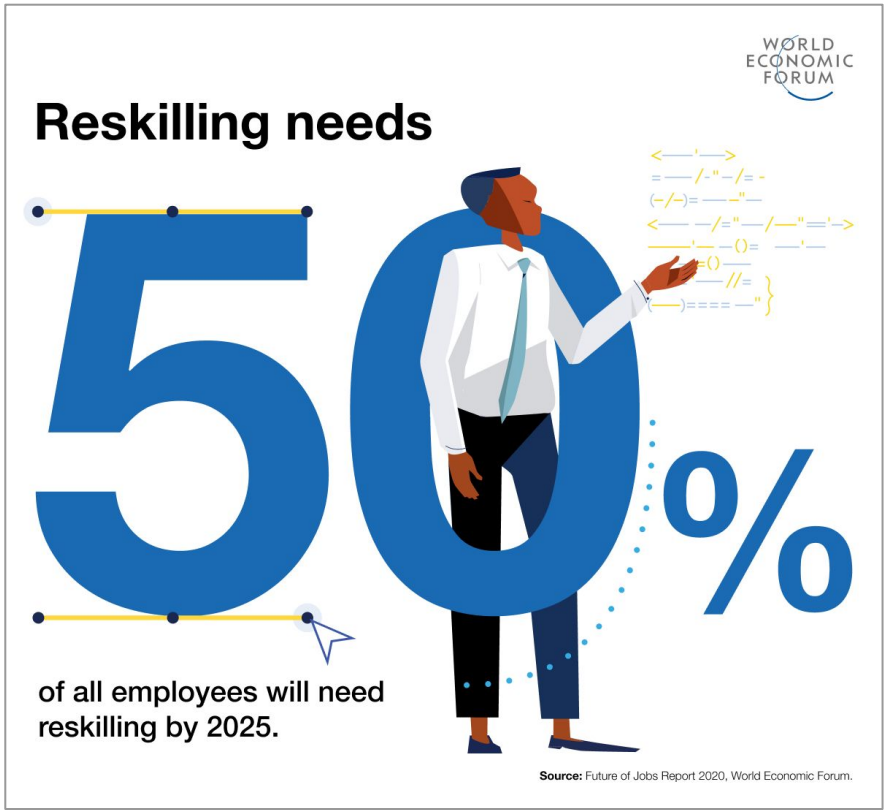
The key drivers of progress, discovery and strategic advantage in AI are access to people, data, compute and finance – all of which face huge global competition.

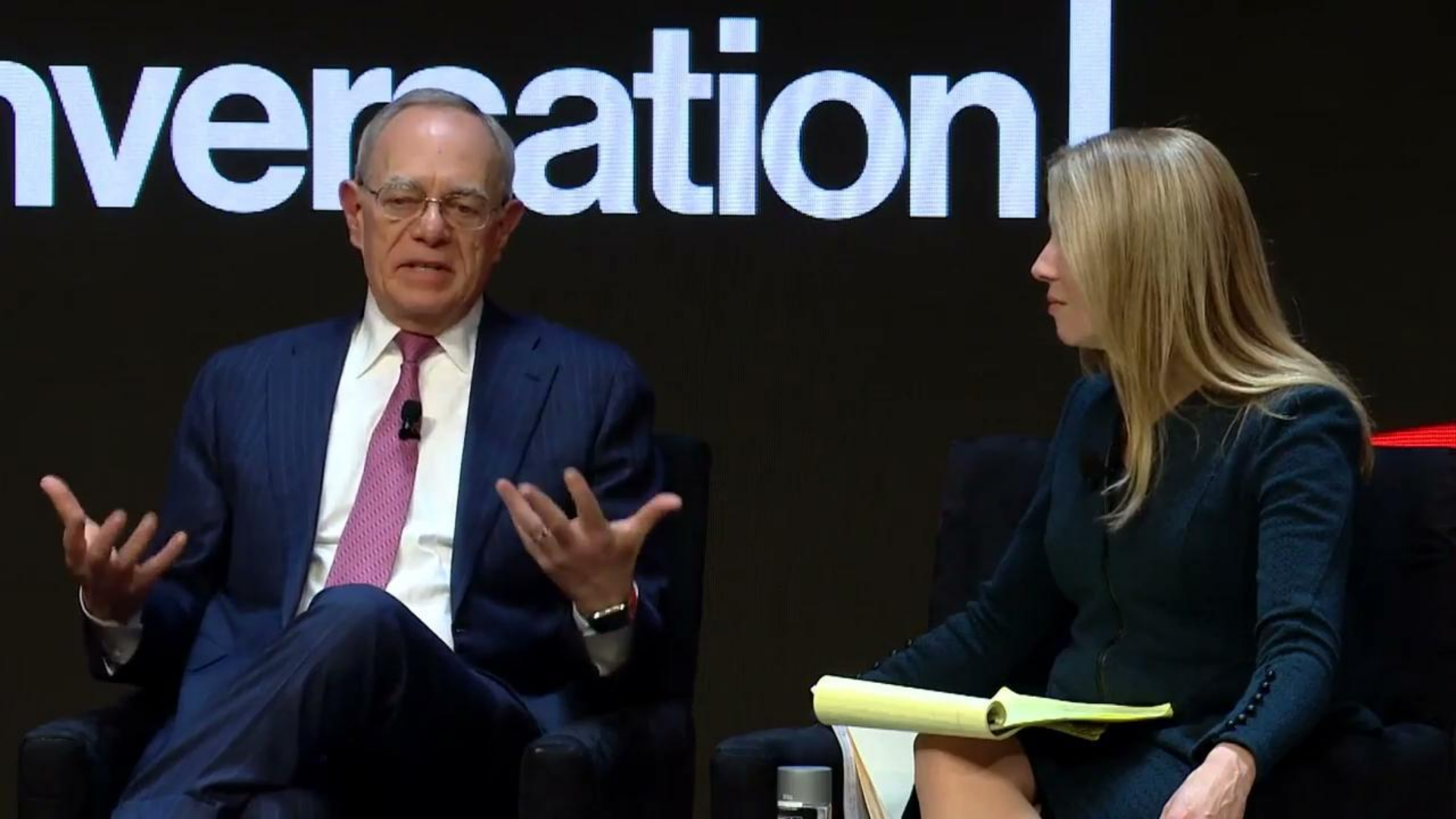
Millions need to be retrained for the AI-enabled economy

- Industry
- Education
- Government
- Defense
- Finance
- Healthcare
- Retail
- Transportation
- Infrastructure

Reskilling

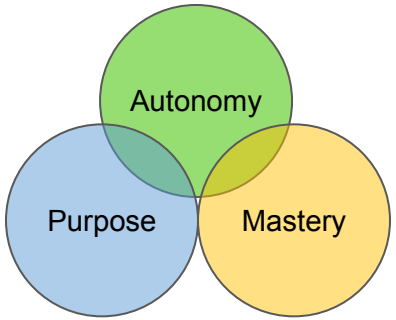

Future of Jobs
Human-machine frontier
WORLD ECONOMIC FORUM
Proportion of tasks completed by humans vs machines
2022
34%
66%
2027
43%
57%
Machine
Human
Source: World Economic Forum, Future of Jobs Report 2023.


WORLD ECONOMIC FORUM
Reskilling needs
50%
of all employees will need reskilling by 2025.
Source: Future of Jobs Report 2020, World Economic Forum.

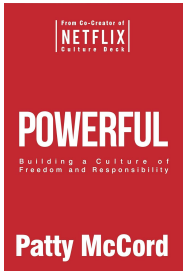

nversation

How to attract talent?

- You already have some
- Others are highly motivated to learn
- 10-to-1 stars - magnets
- Autonomy - get out of the way
- Career advancement without being a manager
- small teams - two pizza teams
- project to product - DevDataOps
- Salaries - MIT data



Drive, Daniel Pink



Developer marketing does not exist

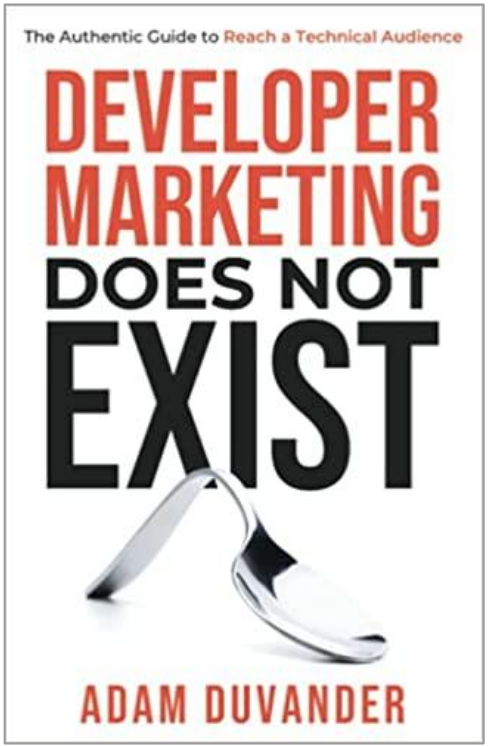


TABLE OF CONTENTS

Organizational Change

There is a lot to do

Gen-AI Digital Transformation: Amplify, Reimagine, Autonomous

Customer Experience Transformation

- Content Generation
- Predictive Personalization
- Interactive AI Solutions

Operational Process Transformation

- Intelligent Automation
- Advanced Analytics
- Enhanced Resource Management

Business Model Transformation

- AI-Driven Products/Services
- Dynamic Business Strategies
- Innovative Ecosystems

Cross-Cutting Factors

- Robust Ethical Frameworks
- Advanced AI Talent Development
- Scalable and Future-Proof Infrastructure
- Regulatory Foresight

KPIs

1. Innovation Rate
2. Automation Rate
3. Accuracy Improvement
4. Adoption Rate
5. Customer Satisfaction
6. Operational Efficiency
7. Innovation ROI
8. Data Utilization Effectiveness
9. Compliance and Security
10. Employee Skill Advancement

Maturity Models

1. Awareness and Education
2. Experimentation and Exploration
3. Integration and Adoption
4. Optimization and Expansion
5. Innovation and Leadership

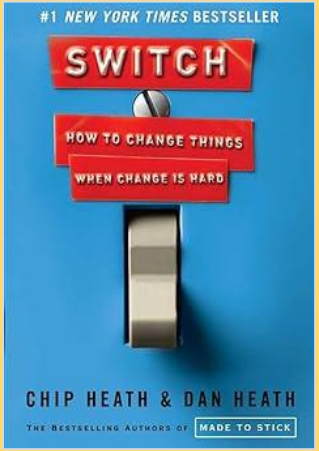
Key Considerations:
- Ethics and governance
- Talent development
- Stakeholder engagement
- Infrastructure for scalability

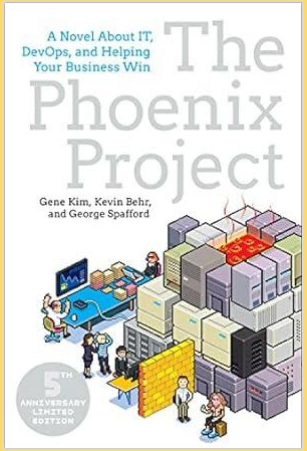
Ethics - AI policy

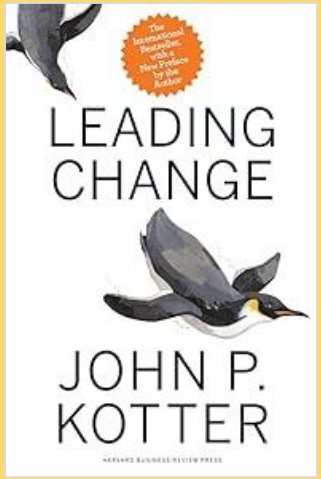
1. Stakeholder Engagement
2. Define Objectives and Scope
3. Ethical and Legal Considerations
4. Governance and Oversight
5. Risk Management
6. Development and Deployment
7. Training and Awareness
8. Monitoring and Evaluation

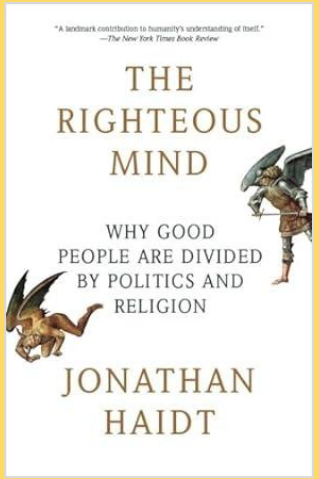
The Change Management

1. The elephant
2. The rider
3. The environment (situation)





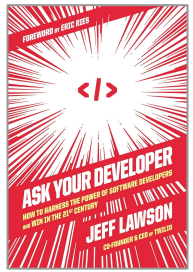




Learn & Change

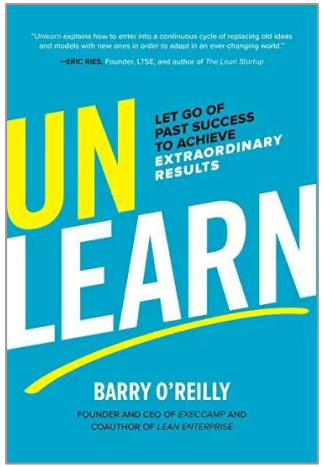
Transformation is hard

70% of transformation projects fail
80% of digital transformation fail
85% of big data projects fail
GE, $10 billion failure

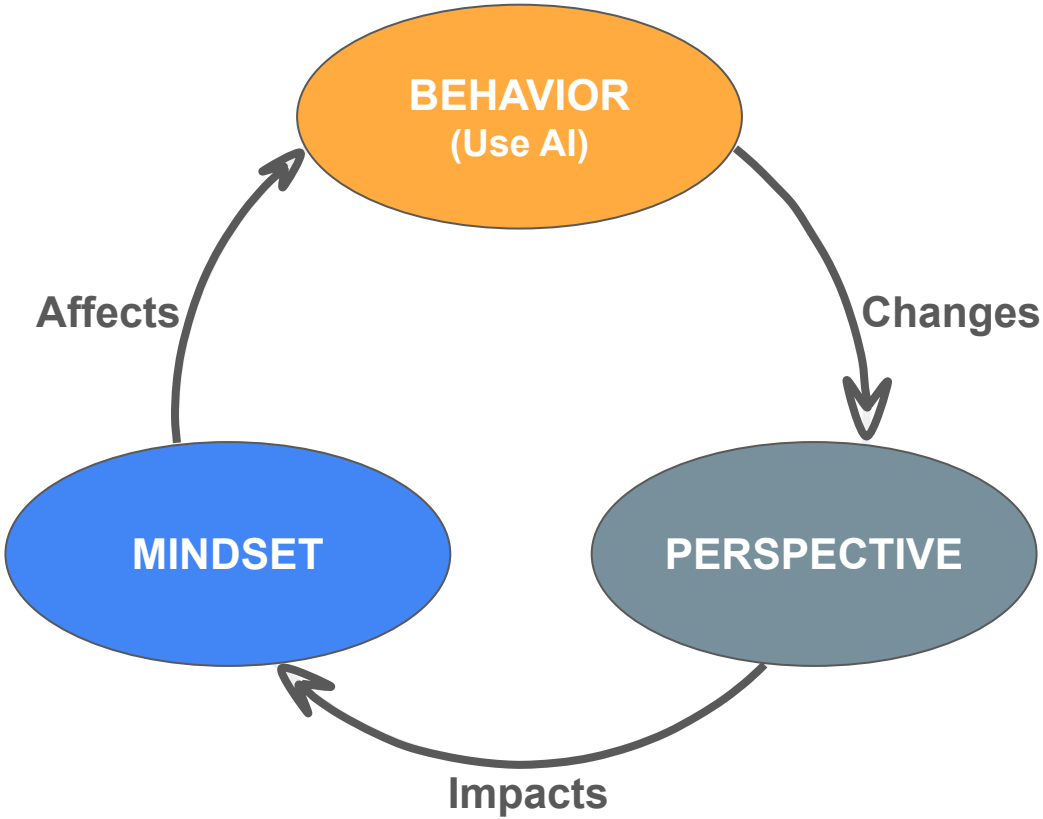


Disruption

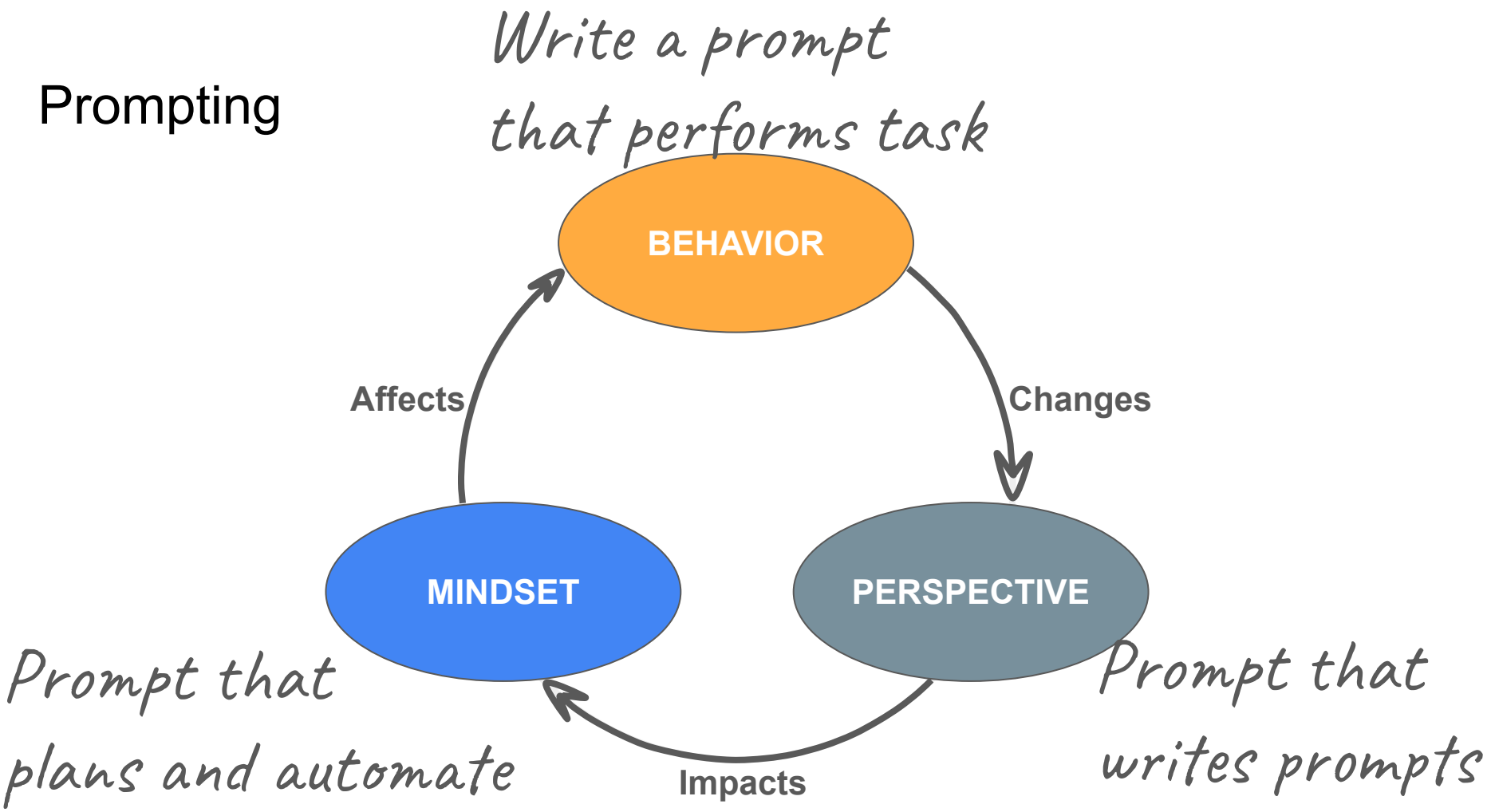
Disruption does not actually apply to organizations. The truth is it applies to individuals.



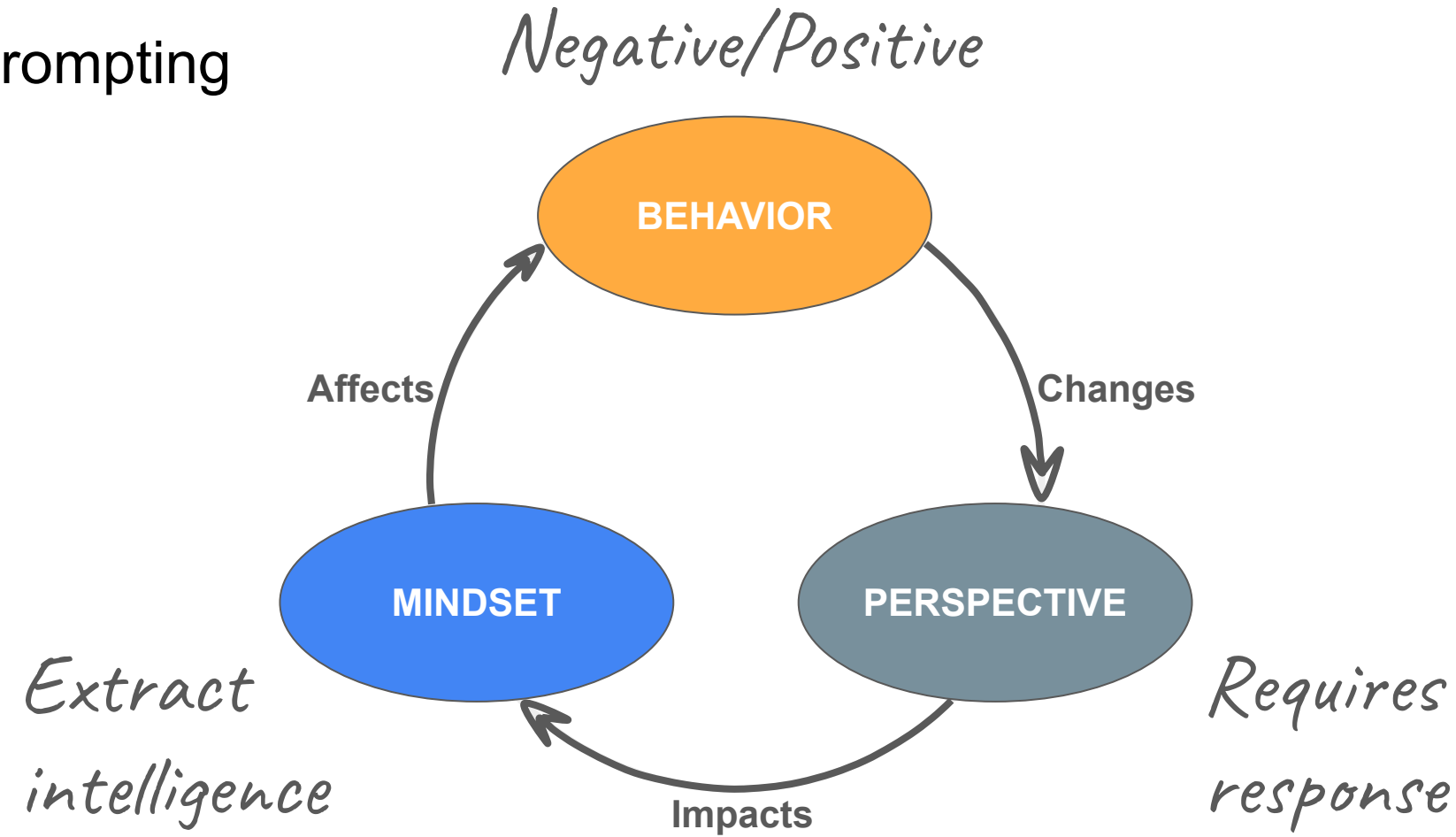
Transformation - live the life, get in the pool



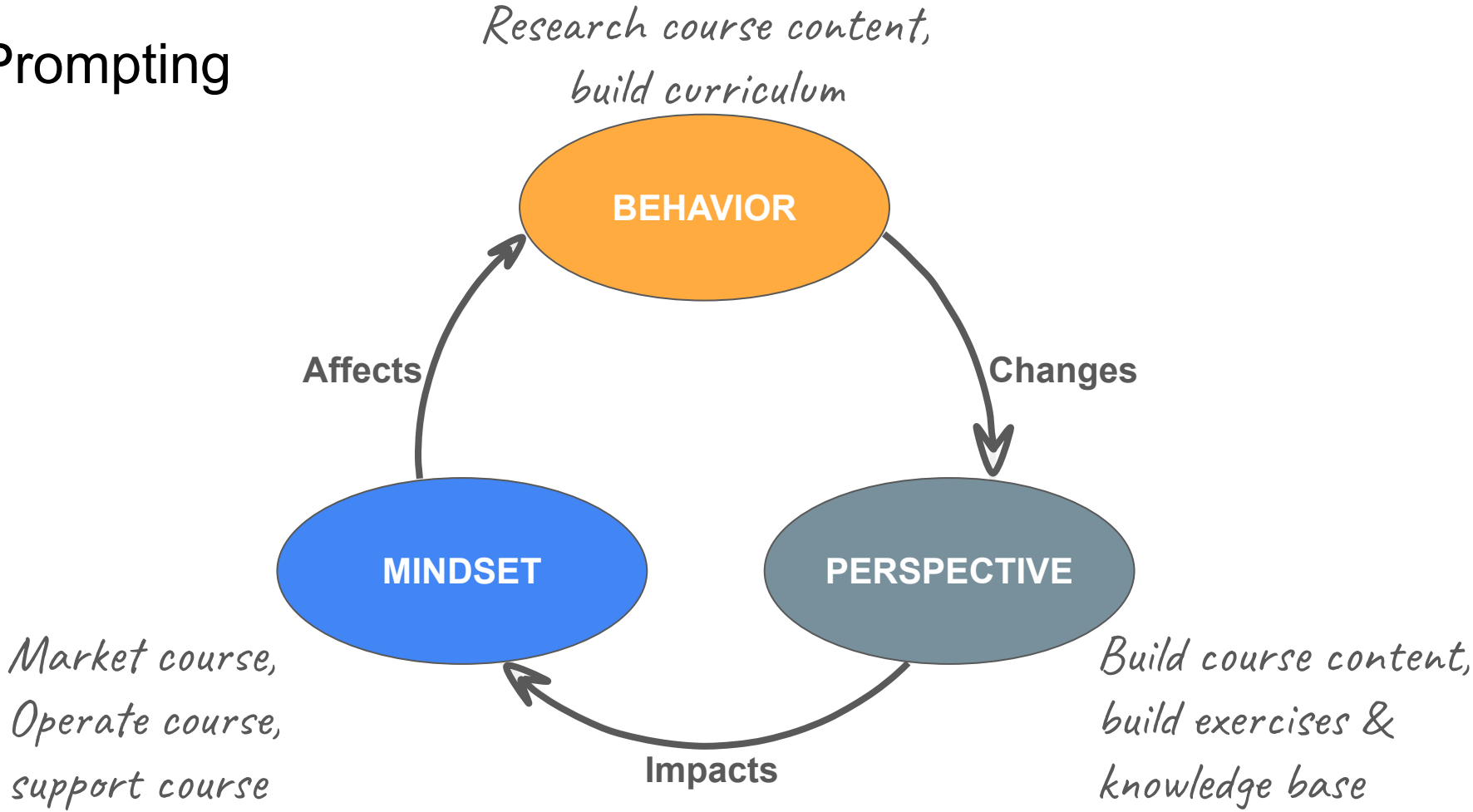
Prompting


Write a prompt that performs task
BEHAVIOR
Changes
PERSPECTIVE
Prompt that writes prompts
Impacts
MINDSET
Prompt that plans and automate
Affects

Prompting


Negative/Positive
BEHAVIOR
Changes
PERSPECTIVE
Requires response
Impacts
MINDSET
Extract intelligence
Affects

Prompting


Research course content,
build curriculum
BEHAVIOR
Changes
PERSPECTIVE
Build course content,
build exercises &
knowledge base
Impacts
MINDSET
Market course,
Operate course,
support course
Affects

I hear and I forget.
I see and I remember.
I do and I understand.

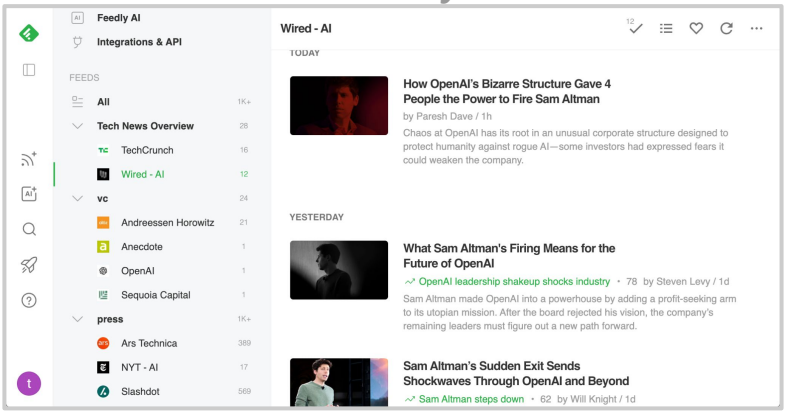
Attributed to Xunzi

Invest an hour per week

- 10 minutes/day - 5 minute scan, 5 minute dive
- See recommended sources
- 30 minutes/week - new service, app, or device, you need to touch it
  - Use it, “Live the life”
- Get a Gen-AI buddy
- Make it fun!

pressreader

Feedly



Sample Sources

Tech News Overview

- TechCrunch
- Wired - AI

vc

- Andreessen Horowitz
- Anecdote
- OpenAI
- Sequoia Capital

press

- Ars Technica
- NYT - AI
- Slashdot
- Stories by Netflix Technology
- The GitHub Blog
- The Official Google AI Blog
- The Official Google Blog

data

- Brent Ozar Unlimited®
- Engineering Blog – Databricks
- KDnuggets
- Statistical Modeling, Causal Infer
- Towards Data Science
- Data Engineering Podcast
- Seattle Data Guy

crypto

- Bitcoin Magazine - Bitcoin News
- CoinDesk
- Ethereum Blog
- The Coinbase Blog

languages

- JavaScript Weekly
- Planet Python
- Real Python
- RisingStack Engineering

We need AI

Average Person

750 books/life

130 million exist

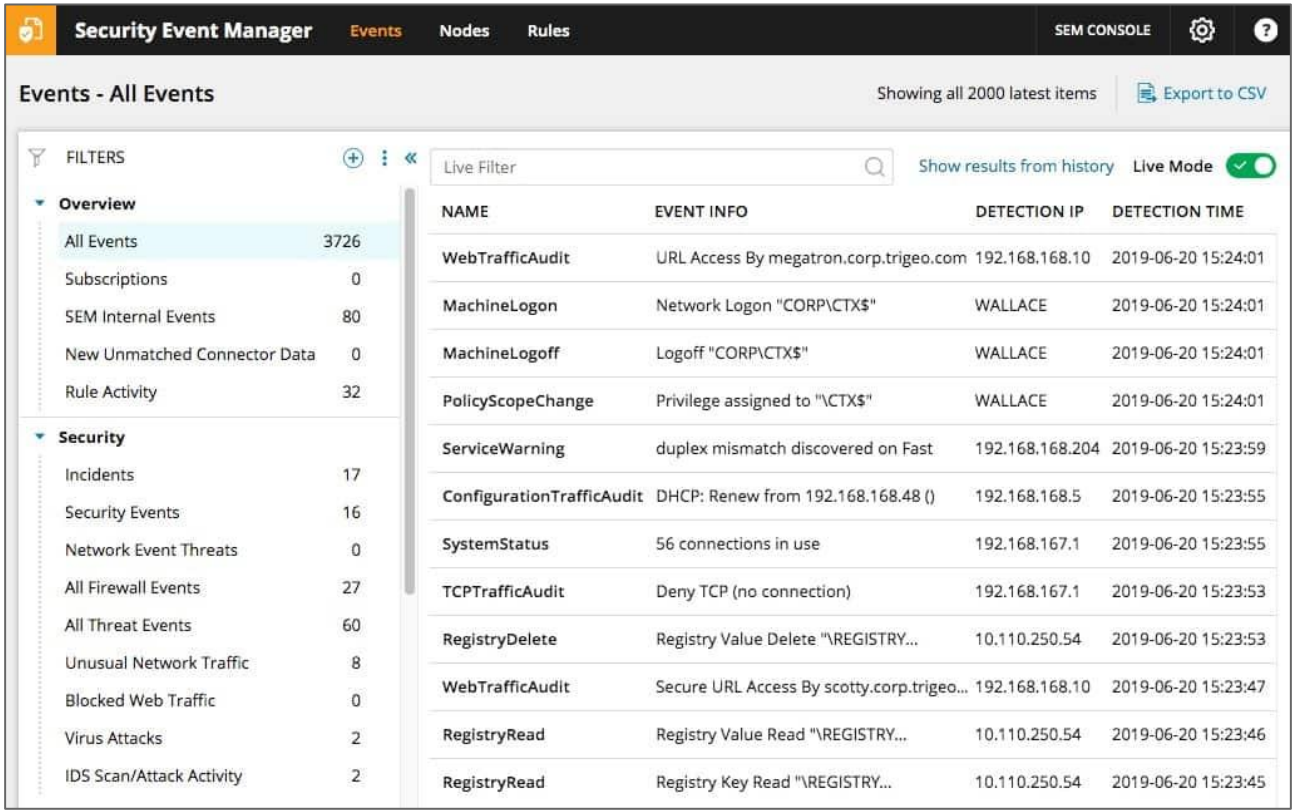
0.0000058

Digital Data Universe

175 zettabytes

1.4 billion x bigger

We need AI: Intrusion detection system events, millions/day




32,000 deaths/year, leading cause of death, ages 4-to-35
CHANDLER CITY HALL
WAYMO
Potential for
80% reduction

UPS - best route

Each UPS driver makes an average of 120 stops per day. There are:

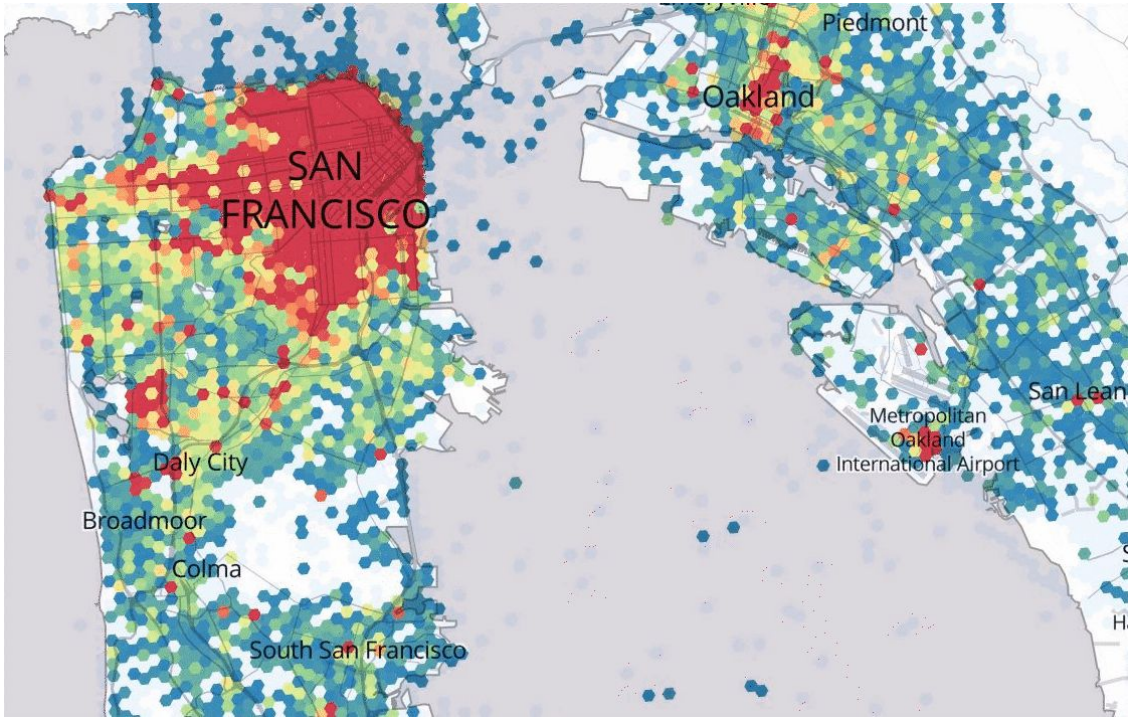
```
6,689,502,913,449,135,000,000,000,000,000,000,000,000,000,000,000,000
,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,
000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,0
00,000,000,000,000,000,000,000,000,000,000,000,000,000,000
```

alternatives for ordering those stops. Which option is the most efficient, after considering variables such as special delivery times, road regulations, and the existence of private roads that don't appear on a map?

Even if an optimal answer exists, the human mind will never figure it out.

UPS best route,
120 stops, 7e^198

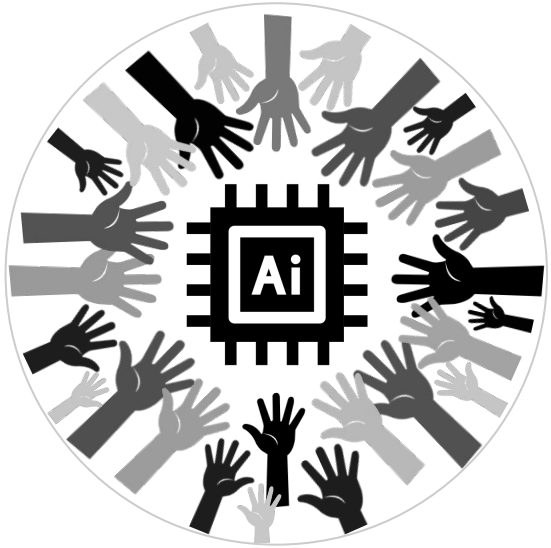
Uber - 1 million predictions per second



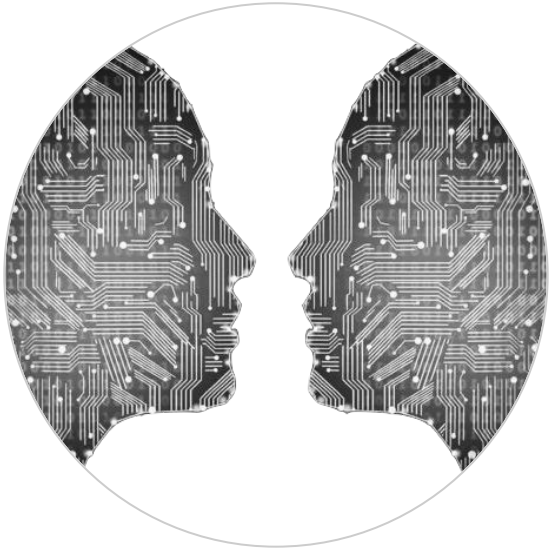
Marketplace forecasting in California's Bay Area allows us to direct drivers to high-demand areas

Democratized access - A new opportunity

Data



AI



Human
Simulation

END