

AI-Ready Data at Scale



Austin Gerig

Scan the QR Code to Connect



Question: How ready is your data for AI?

Disclaimer

The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This presentation expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners or other members of the staff.

Part I: Data Preparation

To scale AI, prepare your data

"Over 93% of Chief Data Officers agree data strategy is crucial for generative AI, however it requires high-quality, relevant data, specific to an organization's business and problems to truly unlock value."^

"Most companies are not ready to deploy generative AI at scale because they lack a strong data infrastructure or the controls needed"~

"72% of the CDOs noted that managing data is one of the top challenges preventing them from scaling AI use case"+

^ **Gartner**

~ **Accenture**

+ **McKinsey**

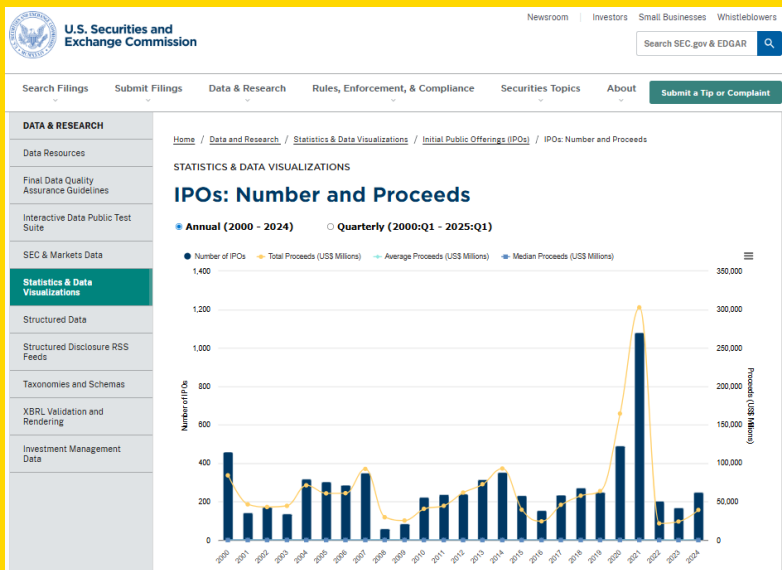
The SEC data landscape

Structured data

- Over 250 datasets
- An exabyte (1000PB) in size!

Unstructured/Semi-structured data

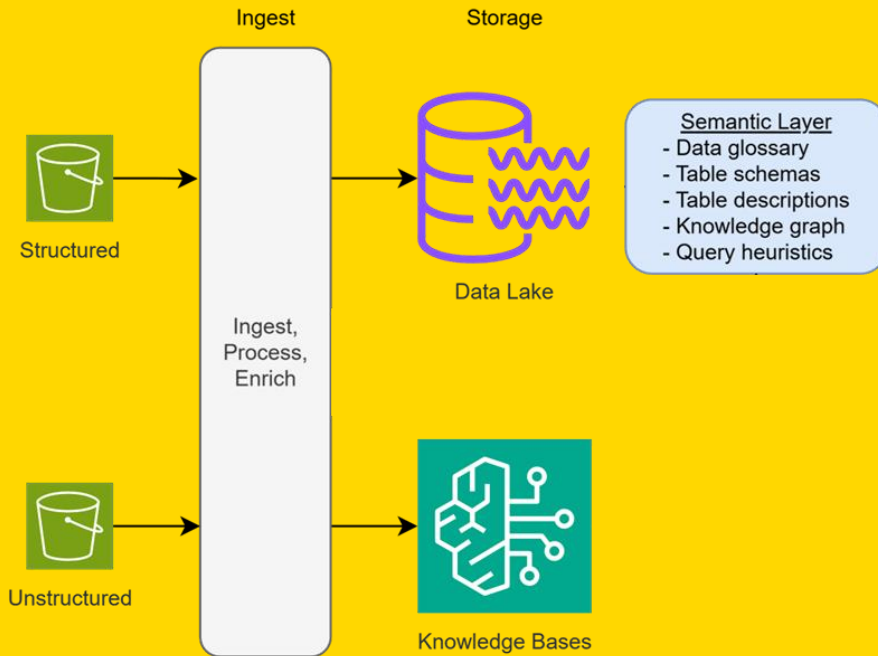
- Over 400 forms
- Over 4500 filings per day



The screenshot shows the 'Investment Adviser Public Disclosure' website. It features a search bar with the text 'Search SEC.gov & EDGAR'. Below the search bar, there are tabs for 'INDIVIDUAL' and 'FIRM'. The 'FIRM' tab is selected. The search criteria include 'Individual Name/CRD#', 'Firm Name or CRD/SEC# (optional)', and 'City, State or ZIP (optional)'. The website also includes a 'Welcome to the Investment Adviser Public Disclosure website' message and a 'For a complete compilation of Investment Adviser Firms' section. A prominent message states: 'The new EDGAR advanced search gives you access to the full text of electronic filings since 2001.' Below this message is a search bar with the text 'Search by keyword, ticker, company name, CIK number or indi...' and a 'SEARCH' button. There is also a '+ more search options' link and a 'Clear all' button.

Two principles for AI-Ready data

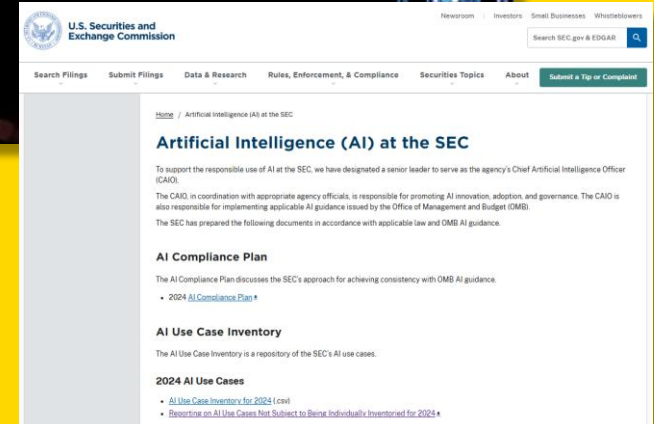
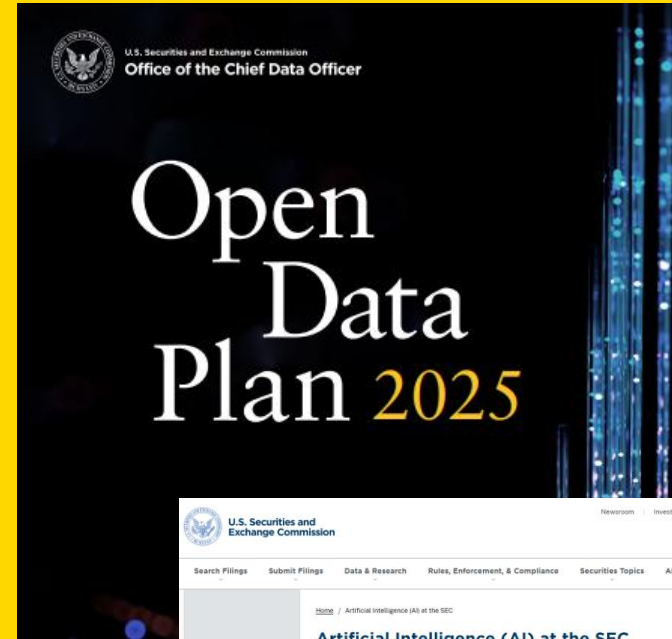
- Make structured data “machine understandable” through a semantic layer
- Curate unstructured data in vector databases that serve AI throughout your organization




SEC Use Cases

- Commission staff are building a semantic layer on top of the raw data assets in the Commission's enterprise data platform, transforming multiple complex data assets to enhance access and usability by staff using AI solutions¹
- Comment Letter Review and Analysis²
 - Comment letter summaries
 - Enhances the review of public comment letters

1. "SEC Open Data Plan" available at https://www.sec.gov/files/2025_sec_opendataplan.pdf
2. SEC AI Use Case Inventory available at <https://www.sec.gov/ai>



Structured Data Example: AI-Powered Queries/Charts


 **Structured Data Query**

Ask a Question

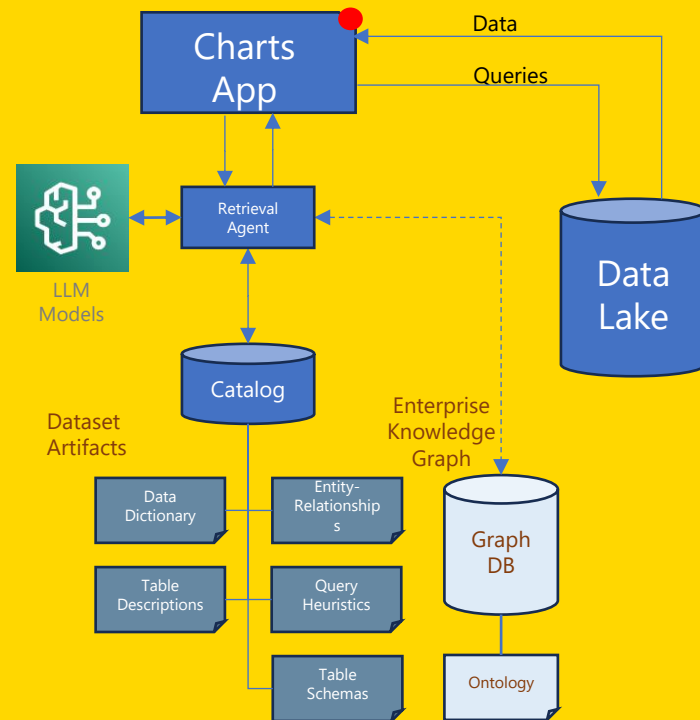
Manual Search

Data Filters

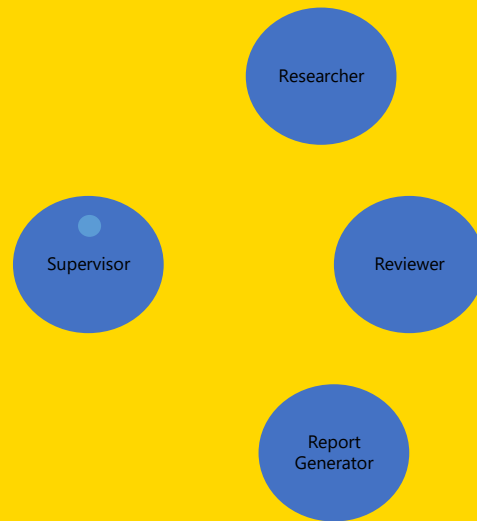
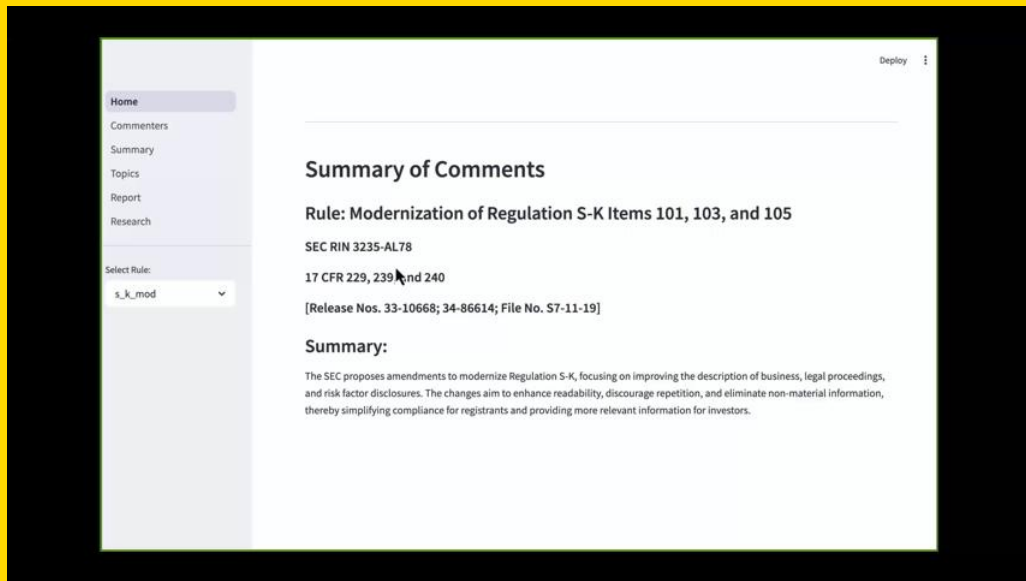
Try: "What is Apple's revenue?" or "Show me Microsoft's assets"



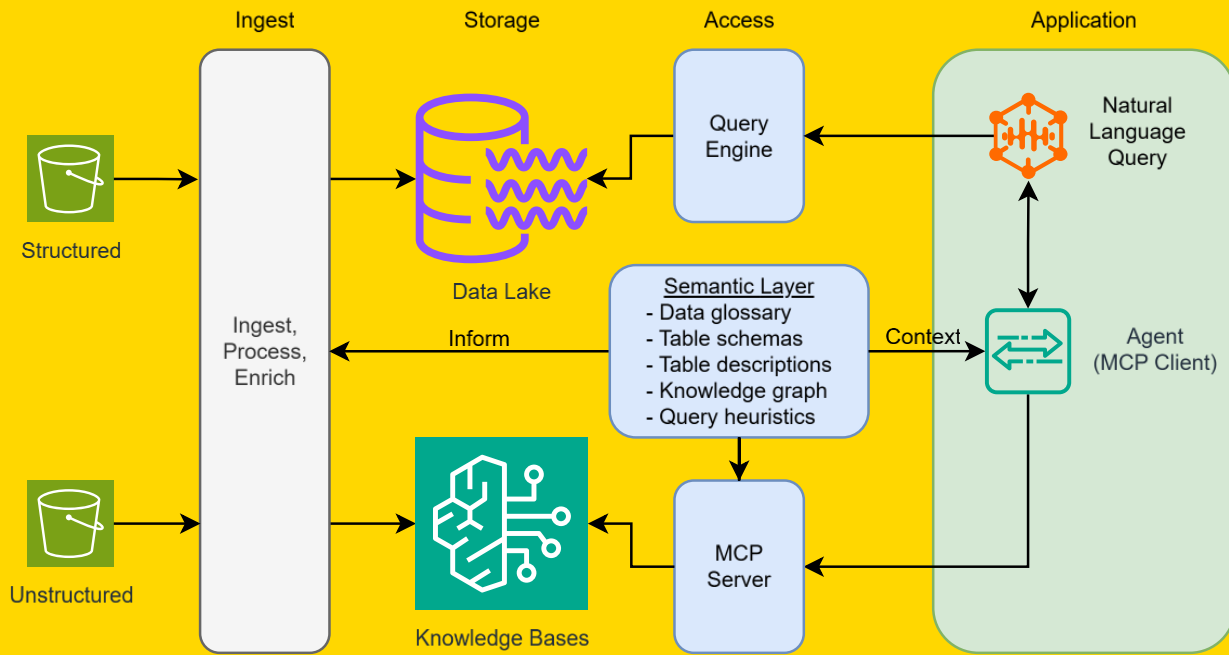
Examples: "Apple revenue", "Microsoft's assets", "Tesla profits"



Unstructured Data Example: AI-Agent Comment Letter Analysis



Architecture & Implementation



Part II: Working with AI

Learning by Doing

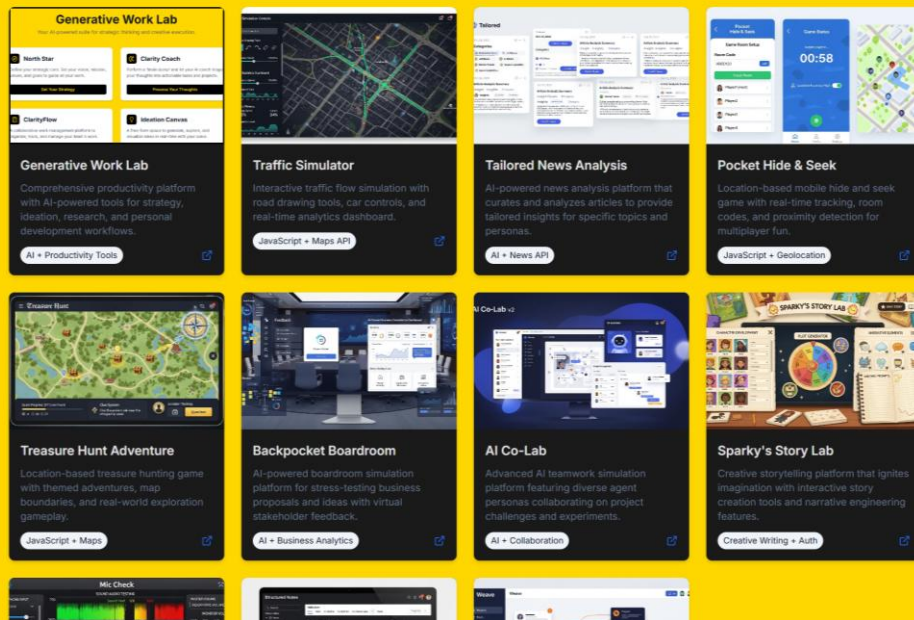
30 Apps in 30 Days

An Experiment in Rapid Prototyping and Creative Exploration

To explore what was possible with LLMs, **Ted Kaouk** and I began with the widest canvas of ideas we could think of: idea capture, design thinking, philosophical reflection, multi-agent environments to model human-like behaviors, decision-making, data modeling, digital design, games, storytelling, and more.

30 Day Apps

A collection of applications created during the 30 day app-a-day challenge.



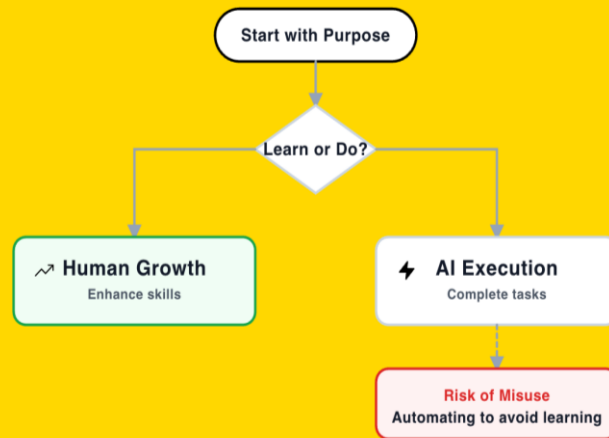
Three Core Principles

We discovered three principles when working with AI:

- 1. Use AI Thoughtfully**
- 2. Learn the Art of Guidance**
- 3. Rethink What is Possible**

1. Use AI Thoughtfully

AI can either foster or hinder your development. Be deliberate with your intent.



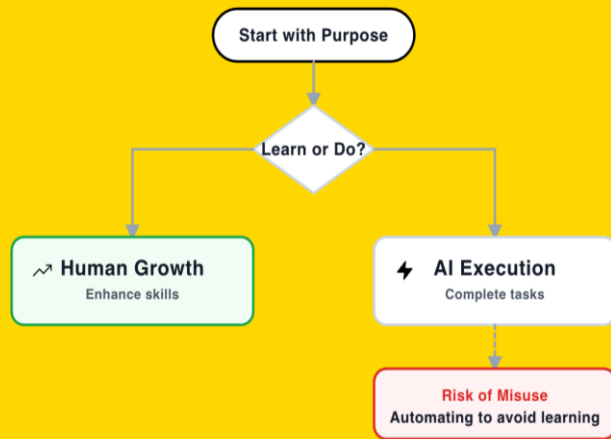
Start with Purpose:

Before you reach for AI, ask yourself a fundamental question. Are you here to explore ideas and learn, or just to get things done?

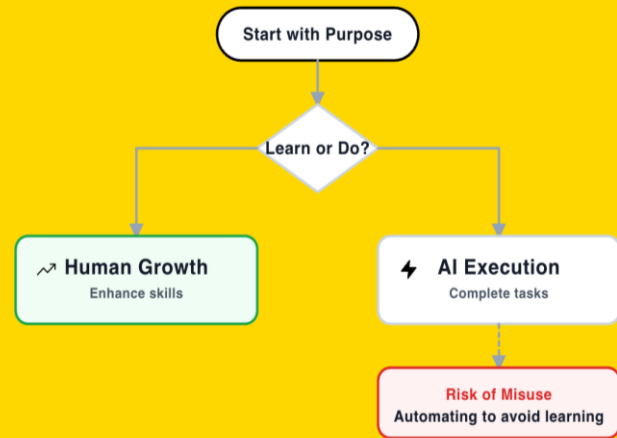
The Risks:

Automating when you should learn hinders growth.

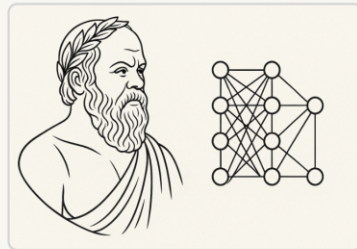
Working on things that should be automated wastes time.



Most work is a blend of learning and doing:
The challenge is to master this blend, moment by moment, deciding what to own and what to offload.



SocraGPTes: We built an AI tool that doesn't give answers, but questions yours to deepen your own understanding.



Socra GPTes

Socrates is not an answer machine; he is a mirror into the quality of your own thinking.

In the age of AI, what should I study in college?

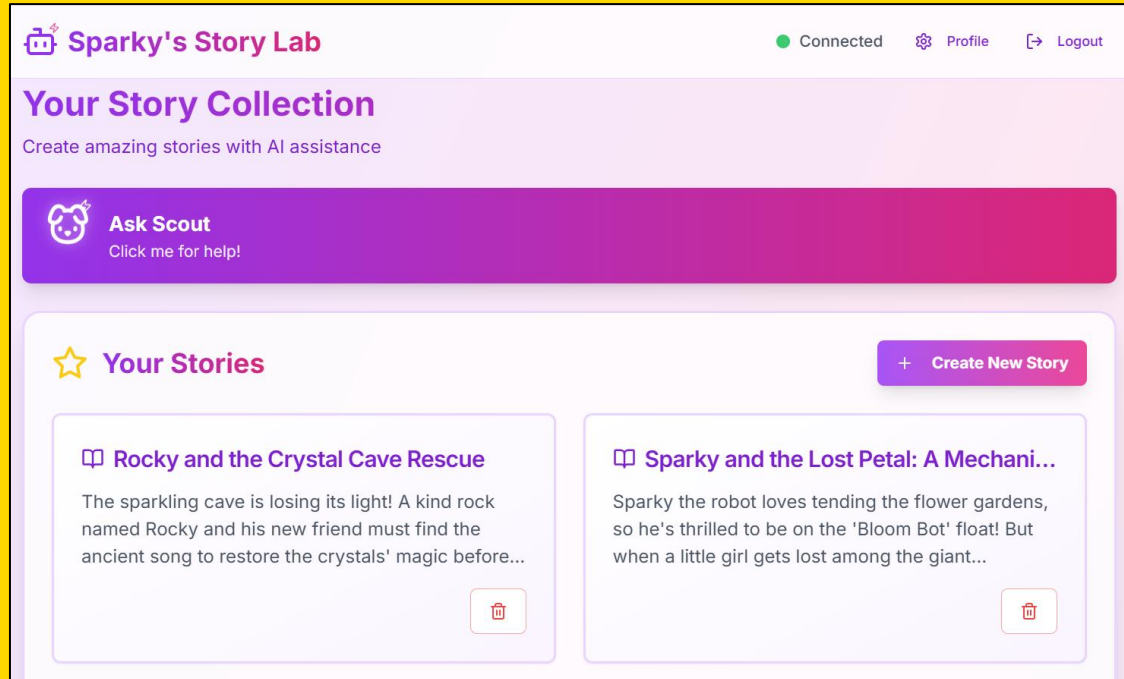
Ah, my most excellent friend, you raise a question that is worthy of long inquiry. But before we hasten to prescribe what you *should* study, I find myself puzzled—for such advice surely depends on what you believe to be the *purpose* of studying, does it not?

Tell me then: what do you believe is the true aim of a college education? Is it to gain wealth? To acquire power? To cultivate the soul? Or is it something else altogether? Let us not discuss what to study until we first understand *why* one studies.

What say you?

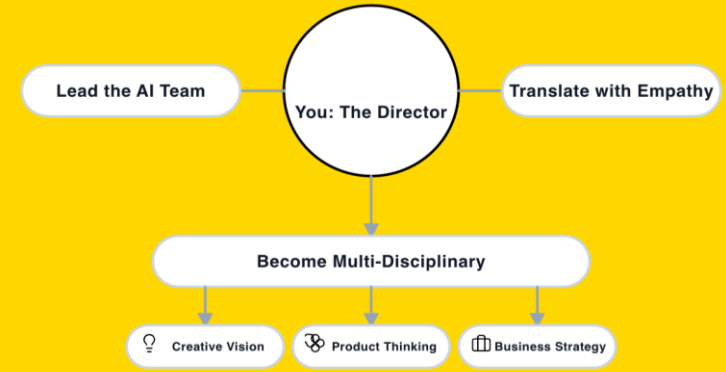


Sparky's Story Lab: An AI storytelling application that doesn't write the story, only assists children when they're stuck.



2. Learn the Art of Guidance

The most critical new skill is shifting from an executor to a director.

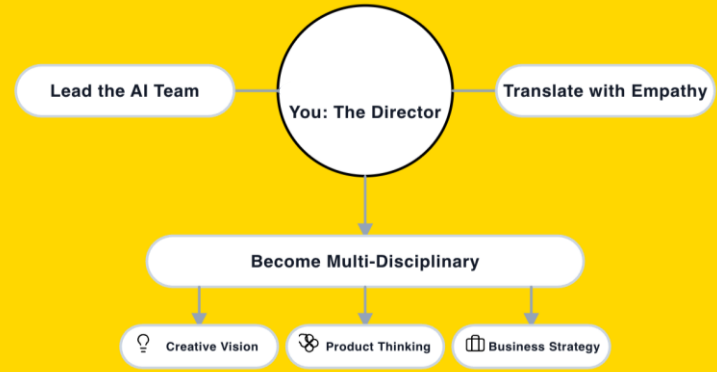


You Must Lead Your AI Team:

To get quality results, you must guide AI by effectively clarifying and communicating the work to be performed.

Become Multi-disciplinary:

To do great work with AI, you need to learn how to provide creative vision, product thinking, business strategy, and enough technical fluency to lead—not just code.



The Experiment: The Bachelorette: a chaotic but perfect stage to test new theories of human behavior.

The First Failure: Classical economics approach (rational utility functions) produced “perfect” but inhuman silence.

The Pivot: Shift to complex economics and agent-based modeling: mimic messy, imperfect real decision-making.




Introducing Psychology: Built with a layered psychological profile (MBTI, schemas, attachment style, conflict style).

Personality, not just goals, guides behavior: unrelenting standards, guarded attachment, conflict triggers, values—producing authentic reactions.

Emergent Social Dynamics: Suitors develop unique motivations; empathy and connection emerge beyond goal-seeking.

Ben "Benny" Kozlowski ✕



Age: 30
Occupation: Carpenter
Hometown: Philadelphia, PA

Bio

Benny is a man who believes in what he can build with his own two hands. He's a third-generation carpenter from South Philly with a heart of gold and a deep suspicion of anyone who seems too slick. He isn't here to play games; he's a straight-shooter who values family, hard work, and honesty above all else.

What you see is what you get with Benny. He comes from a tight-knit family where love was never in question, giving him a quiet, unshakeable confidence. He doesn't need to be the loudest guy in the room because he's comfortable in his own skin. He shows his affection through actions, not empty words, and is looking for a woman who understands that.

💡 Watch For...

- His thick, unapologetic Philly slang ("jawn," "you guys") that sets him apart.
- Expressing his interest by doing things—offering to fix something, giving a small handmade gift.
- His tendency to be a quiet observer, taking everything in before he speaks.

Emergent Social Dynamics:

Suitors develop unique motivations; empathy and connection emerge beyond goal-seeking.

"It's good to get a feel for the room... Kevin seems a bit quiet, perhaps I should go introduce myself to him..."

Ben, Internal Monologue

Become an AI-Anthropologist

To animate a hidden skull in a painting, direct commands failed. The solution was providing a better 'map'—the context the AI was missing—to guide it.



a



b

Become an AI-Anthropologist

To animate a hidden skull in a painting, direct commands failed. The solution was providing a better 'map'—the context the AI was missing—to guide it.



Become an AI-Anthropologist

To animate a hidden skull in a painting, direct commands failed. The solution was providing a better 'map'—the context the AI was missing—to guide it.



Design Thinking: An app with an AI facilitator that listens and summarizes.

[← Back to Dashboard](#)

Discover Phase

Focus Area: How can we make data AI ready?

Design Thinking Sessions

+ New

Session name

Save

Load a session...

Mic: OFF

Suggest Refined Problem

Next: Solution Phase →

```
graph TD; Center((Making Data AI Ready)) --- S[Structured Data Sets]; Center --- U[Unstructured Data]; S --- V[Very Large]; S --- AI[AI Interaction Challenges]; S --- P[Problems]; AI --- UO[Uploading All Not Optimal]; AI --- TI[Requires Tooling Interface]; U --- M[Messy]; U --- D[Disorganized]; U --- L[Don't Shove All into LLM]; P --- L;
```

Transcript

Research

so I think there's quite a few different problems with making data AI ready first there's your structured data sets which are often very very large and AI doesn't interact with them directly at least uploading all of the structured data into AI and asking questions is not really optimal usually you want to tooling type interface

and then for unstructured data it's oftentimes really messy disorganized you don't want to shove it all in at the same time to the llm so figuring out how to deal with that issue as well as a I think two of the biggest problems

3. Rethink What is Possible

Don't just optimize the old. Create the new.

Mindset: Optimization

Optimizing Old Workflows



No Competitive Edge

Mindset: Innovation

Rethink What is Possible



Human + AI Synergy



Breakthrough
Outcomes

**As AI adoption approaches 100%,
simply using it won't set you apart—
only how you use it will.**

**Aim for outcomes neither humans
nor machines could achieve alone.**

Use AI to push against boundaries and
solve problems that were once out of
reach.

Mindset: Optimization

Optimizing Old Workflows



No Competitive Edge

Mindset: Innovation

Rethink What is Possible



Human + AI Synergy



Breakthrough
Outcomes

Case in point: We built [Clarity Coach](#) to manage our chaotic inner monologue—a new class of tool that solves a previously intractable human problem.

[← Back to Dashboard](#)

Clarity Coach

Perform a "brain dump" to triage your thoughts into actionable items.

Brain Dump Sessions

+ New

Session name

Save

Load a session...

Brain Dump

Throw anything at it: tasks, projects, worries, ideas. The more, the better.

so I am a little bit worried that the presentation is taking too long and it's not that interesting I'm also worried about a work project that is pretty complex and requires us to Think Through the size of different data sets and I have quite a few different things around the house that I need to take care of including mowing the yard and cleaning up the bathrooms

Mic: OFF

Go to Ideation Canvas

Process My Thoughts

Messy Problems

Complex challenges needing a structured approach.

work project that is pretty complex and requires us to Think Through the size of different data sets

Start Design Thinking

Simple Tasks

Clear, concrete, single-step actions.

mowing the yard

Inbox

cleaning up the bathrooms

Inbox

Rumination

Mental chatter to acknowledge and release.

worried that the presentation is taking too long and it's not that interesting

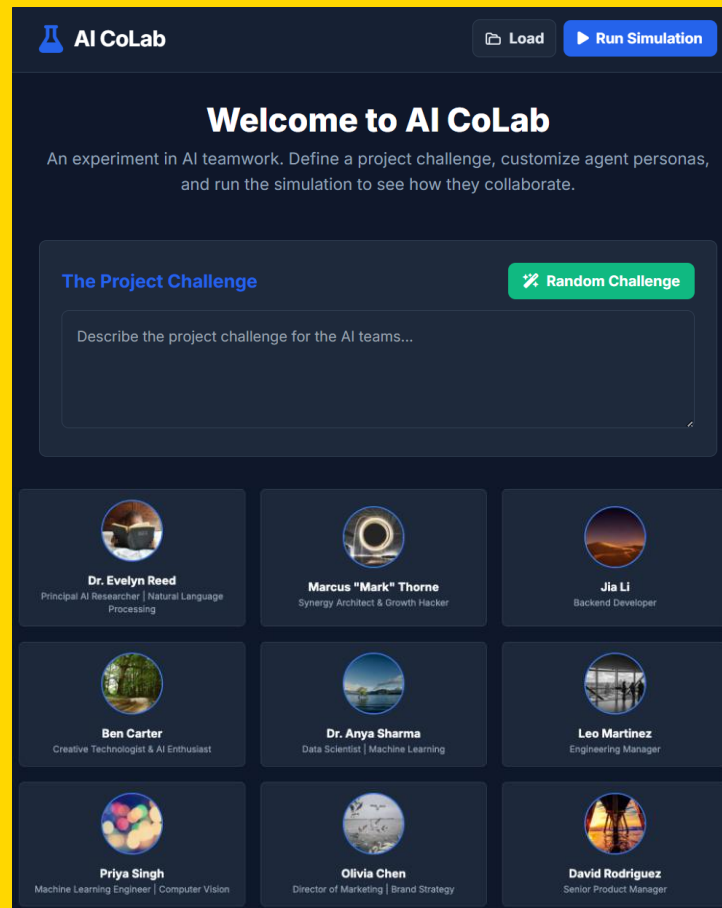
Someday / Maybe

Ideas to park for later.

None identified.

[AI CoLab](#) simulates **complex project teams** composed of autonomous AI agents, each with distinct skills, experience, and working styles.

These agents are tasked with solving real-world business challenges — designing strategies, writing proposals, allocating resources, and delivering outcomes — and their behaviors, decisions, and interactions are all analyzed to generate insights.



Key features include:

Dynamic Peer Assessment: At the outset, each team member rates the others based on perceived capabilities and expected contributions. At the conclusion, they do so again based on actual performance, collaboration, leadership, and follow-through — revealing how perceptions evolve over the life of a project.

Project-Level Evaluation: Each team also scores the projects themselves, providing a structured assessment of quality, feasibility, creativity, and execution.

Emergent Insights: The resulting data uncovers patterns in team dynamics — how trust develops, how leadership emerges, where friction arises, and what conditions lead to the best outcomes.

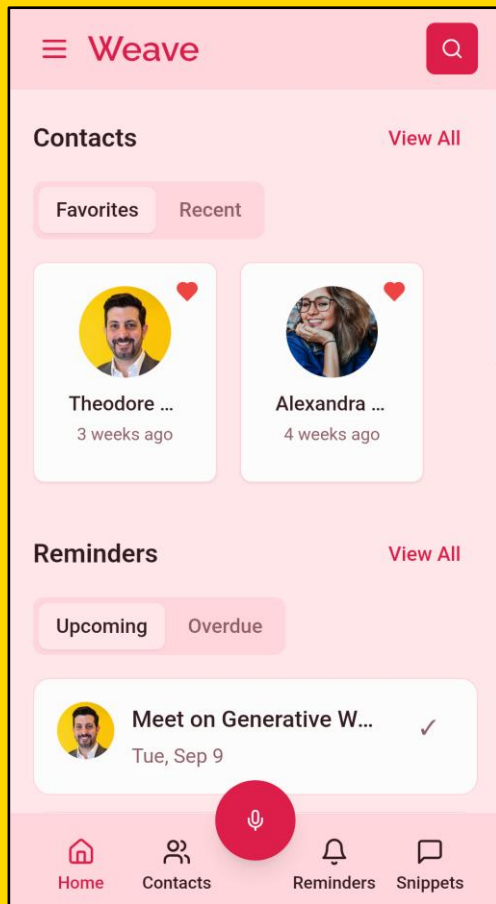
Team Assignments & Personas			
Team Alpha		Team Bravo	
David Rodriguez (Leader) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Pragmatic	Olivia Chen (Communicator) Collab: Collaborative Ethic: Lazy	Comm: Friendly Approach: Innovative
Fatima Al-Jamil (Pragmatist) Collab: Lone Wolf Ethic: Hard-working	Comm: Rude Approach: Pragmatic	Samuel "Sam" Jones (Pragmatist) Collab: Lone Wolf Ethic: Hard-working	Comm: Rude Approach: Pragmatic
Dr. Evelyn Reed (Leader) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Pragmatic	Ben Carter (Innovator) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Innovative
Team Delta		Team Echo	
Kenji Tanaka (Specialist) Collab: Lone Wolf Ethic: Hard-working	Comm: Rude Approach: Pragmatic	Priya Singh (Specialist) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Innovative
Chloe Davis (Harmonizer) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Pragmatic	Isabella Rossi (Communicator) Collab: Collaborative Ethic: Hard-working	Comm: Friendly Approach: Pragmatic
Liam O'Connell (Innovator) Collab: Collaborative Ethic: Lazy	Comm: Friendly Approach: Innovative	Leo Martinez (Leader) Collab: Collaborative Ethic: Hard-working	Comm: Rude Approach: Pragmatic

Reviewer David Rodriguez	Reviewee Fatima Al-Jamil
Score	8/10
Strength: Fatima's biggest strength is her unwavering commitment to technical rigor and practical implementation. She consistently ensured our proposed solutions were robust, reliable, and accounted for critical engineering details like version control and CI/CD. This brought essential depth to our project.	
Improvement: While her individual contributions were excellent, I believe a slightly more collaborative approach during the initial drafting phases could have further enriched our documents and fostered even stronger team cohesion.	

Reviewer David Rodriguez	Reviewee Dr. Evelyn Reed
Score	9/10
Strength: Evelyn's greatest strength lies in her exceptional ability to synthesize diverse ideas and bridge strategic concepts with technical realities. She consistently found ways to integrate everyone's expertise, acting as a crucial unifier and elevating our collective output. Her insights into AI-driven guidance were also invaluable.	
Improvement: Evelyn is a fantastic collaborator. Perhaps, continuing to inspire and gently guide team members towards collaborative document creation, even when there are preferences for individual work, could further enhance team synergy.	

Reviewer Fatima Al-Jamil	Reviewee Dr. Evelyn Reed
Score	9/10
Strength: Understands the technical requirements and consistently refined her suggestions to be more robust and implementable, aligning with proper engineering practices.	
Improvement: She wastes time attempting to collaborate on writing, which is inefficient. She should focus on providing direct technical input when requested.	

Reviewer Dr. Evelyn Reed	Reviewee David Rodriguez
Score	9/10
Strength: David consistently fostered a positive and collaborative environment, expertly integrating diverse ideas and ensuring our strategic vision was well-articulated. His ability to connect our technical solution back to the market analysis and business case was invaluable.	
Improvement: While excellent at facilitating, David could sometimes push a bit more for collaborative writing when the complexity of a section might benefit from combined input, even when there's a strong preference for individual work.	

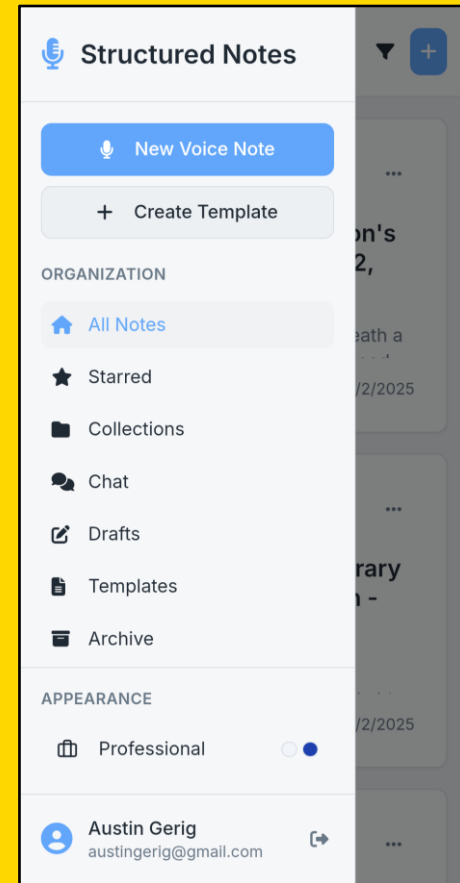


Weave transforms unstructured information about your friends and colleagues into structured contact entries.

It enables **quick and easy capture of information** after meetings and social gatherings.

Structured Notes takes transcripts, spoken notes, and unstructured data, and structures them into user-defined templates.

It is **built for a range of use cases**, from filling out documentation, to travel planning, and tracking medical appointments.



Additional Demos

30 Day Apps

A collection of applications created during the 30 day app-a-day challenge.

