

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

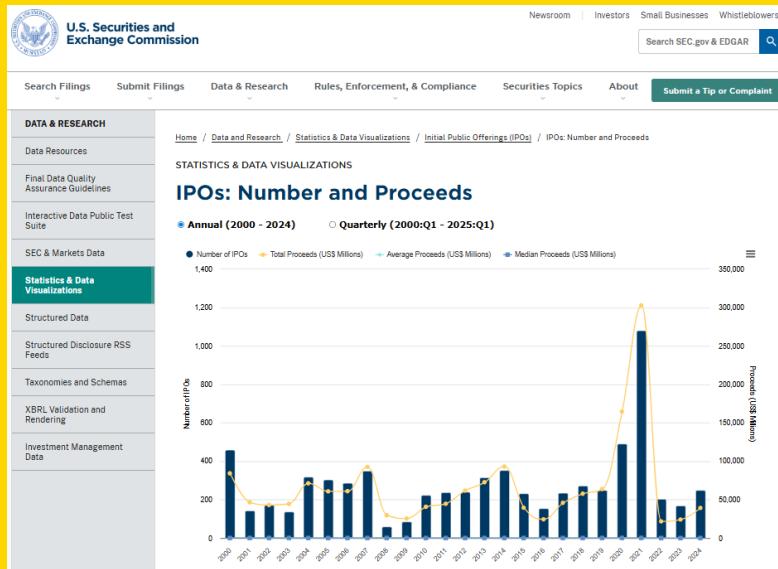
~Accenture

+McKinsey

The SEC data landscape

Structured data

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



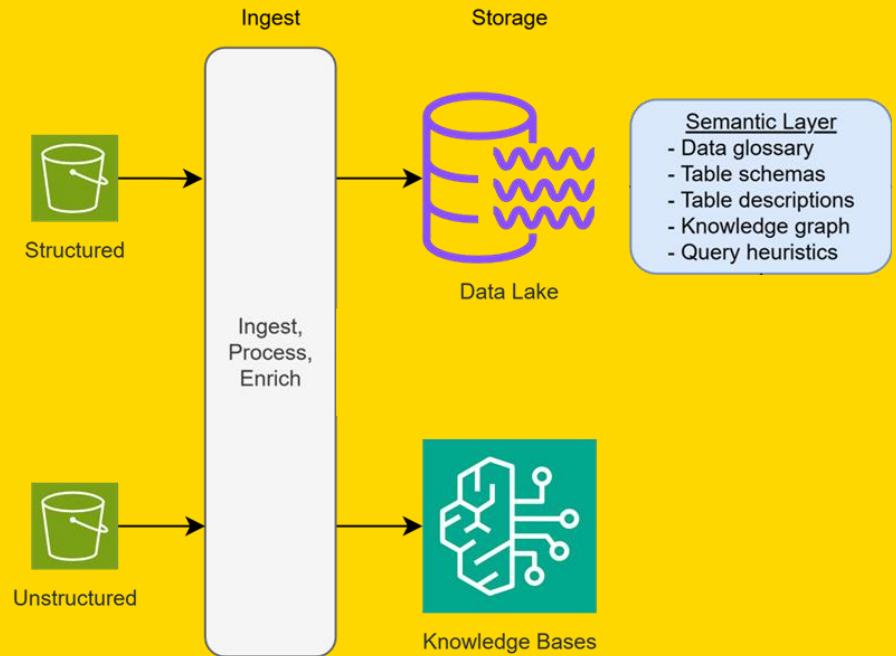
Unstructured/Semi-structured data

- Over 400 forms
- Over 4500 filings per day

The image shows two side-by-side screenshots of SEC websites. On the left is the "Investment Adviser Public Disclosure" website, featuring a search bar for "INDIVIDUAL" and "FIRM" names, and links for "Individual Name/CRD#", "Firm Name or CRD/SEC# (optional)", and "City, State or ZIP (optional)". Below the search bar is a welcome message and a link to "Investment Adviser Data". On the right is the "EDGAR" search interface, featuring a large search bar with placeholder text "Search by keyword, ticker, company name, CIK number or industry", a yellow "SEARCH" button, a "Clear all" button, and links for "FAQ" and "Other search tools". A banner at the bottom of the EDGAR page states: "The new EDGAR advanced search gives you access to the full text of electronic filings since 2001."

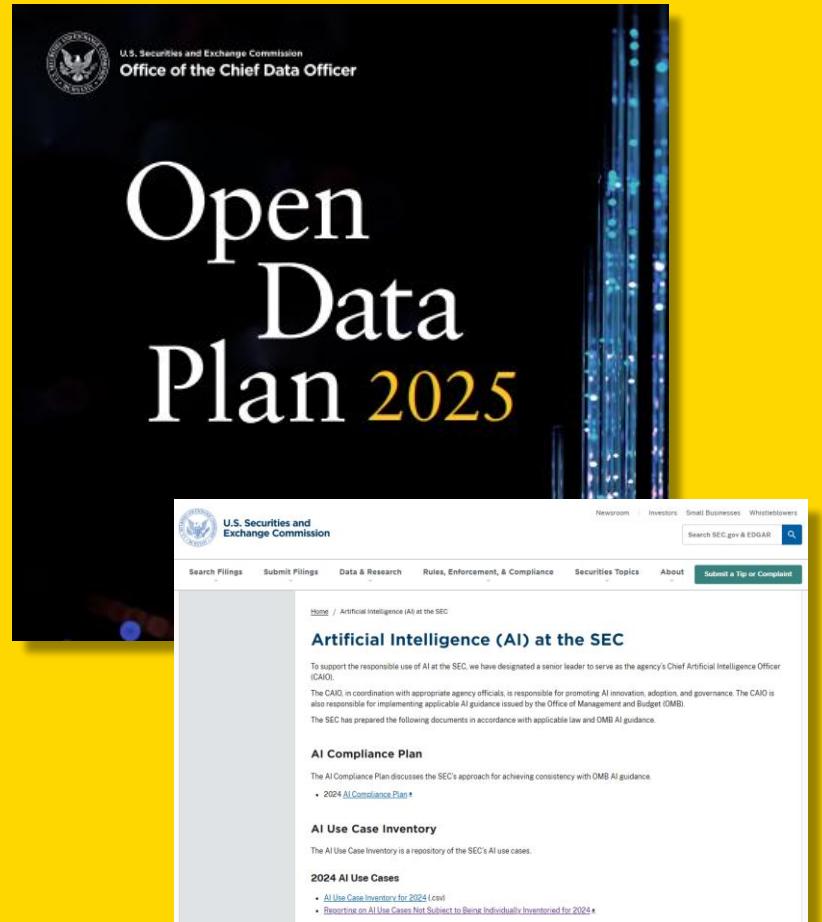
Two principles for AI-Ready data

- 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>



The image shows two screenshots of the U.S. Securities and Exchange Commission (SEC) website. The top screenshot displays the "Open Data Plan 2025" page, featuring large white and yellow text on a dark background. The bottom screenshot shows the "Artificial Intelligence (AI) at the SEC" page, which includes sections for the AI Compliance Plan and the AI Use Case Inventory.

**U.S. Securities and Exchange Commission
Office of the Chief Data Officer**

Open Data Plan 2025

[Search SEC.gov & EDGAR](#)

U.S. Securities and Exchange Commission

[Search Filings](#) [Submit Filings](#) [Data & Research](#) [Rules, Enforcement, & Compliance](#) [Securities Topics](#) [About](#) [Submit a Tip or Complaint](#)

[Home / Artificial Intelligence \(AI\) at the SEC](#)

Artificial Intelligence (AI) at the SEC

To support the responsible use of AI at the SEC, we have designated a senior leader to serve as the agency's Chief Artificial Intelligence Officer (CAIO). The CAIO, in coordination with appropriate agency officials, is responsible for promoting AI innovation, adoption, and governance. The CAIO is also responsible for implementing applicable AI guidance issued by the Office of Management and Budget (OMB). The SEC has prepared the following documents in accordance with applicable law and OMB AI guidance.

AI Compliance Plan

The AI Compliance Plan discusses the SEC's approach for achieving consistency with OMB AI guidance.

- [2024 AI Compliance Plan*](#)

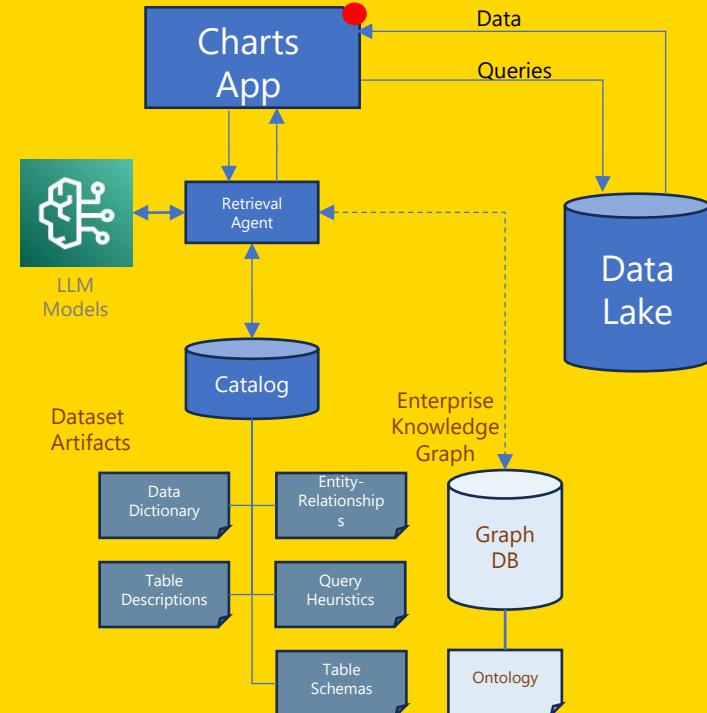
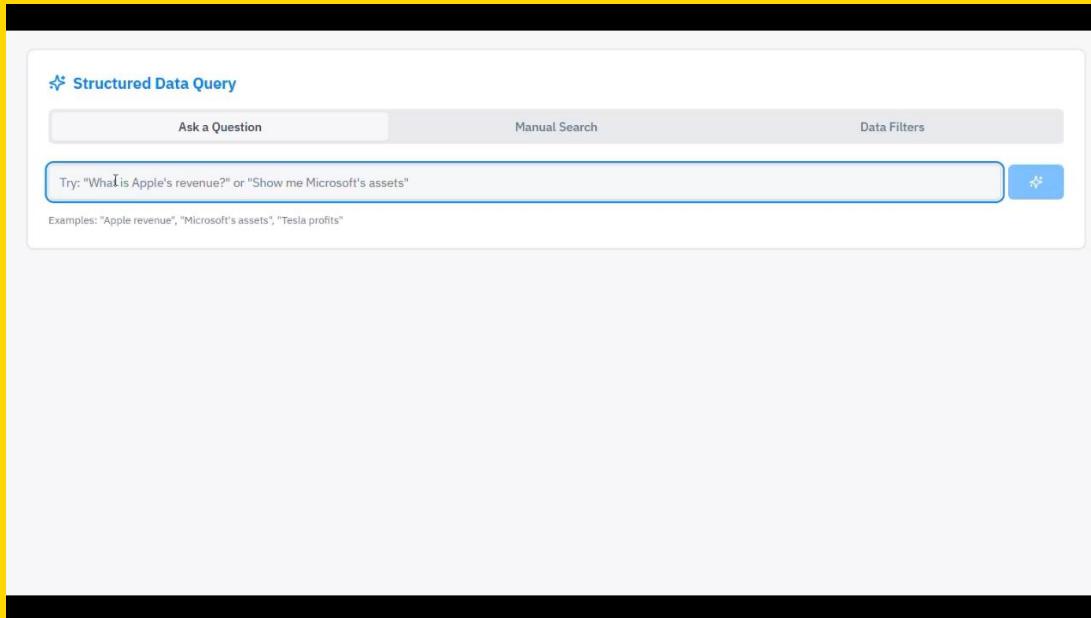
AI Use Case Inventory

The AI Use Case Inventory is a repository of the SEC's AI use cases.

2024 AI Use Cases

- [AI Use Case Inventory for 2024 \(.csv\)](#)
- [Reporting on AI Use Cases Not Subject to Being Individually Inventory for 2024*](#)

Structured Data Example: AI-Powered Queries/Charts

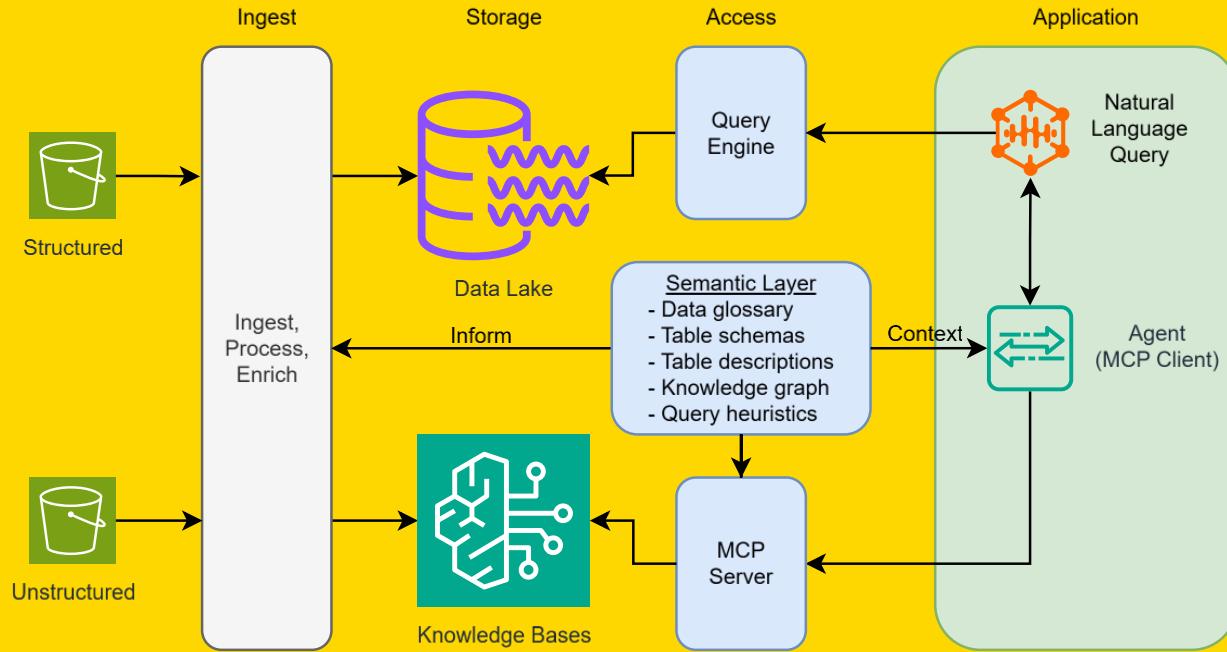


Unstructured Data Example: AI-Agent Comment Letter Analysis

The screenshot shows a web-based application for analyzing comment letters. On the left, a sidebar menu includes Home, Commenters, Summary, Topics, Report, Research, and a dropdown for 'Select Rule' with 's_k_mod' selected. The main content area has a title 'Summary of Comments' and a subtitle 'Rule: Modernization of Regulation S-K Items 101, 103, and 105'. It also lists 'SEC RIN 3235-AL78', '17 CFR 229, 239, and 240', and '[Release Nos. 33-10668; 34-86614; File No. S7-11-19]'. A 'Summary:' section contains a paragraph about modernizing Regulation S-K to improve disclosure readability and eliminate non-material information.



Architecture & Implementation



Part II: Working with AI

Learning by Doing

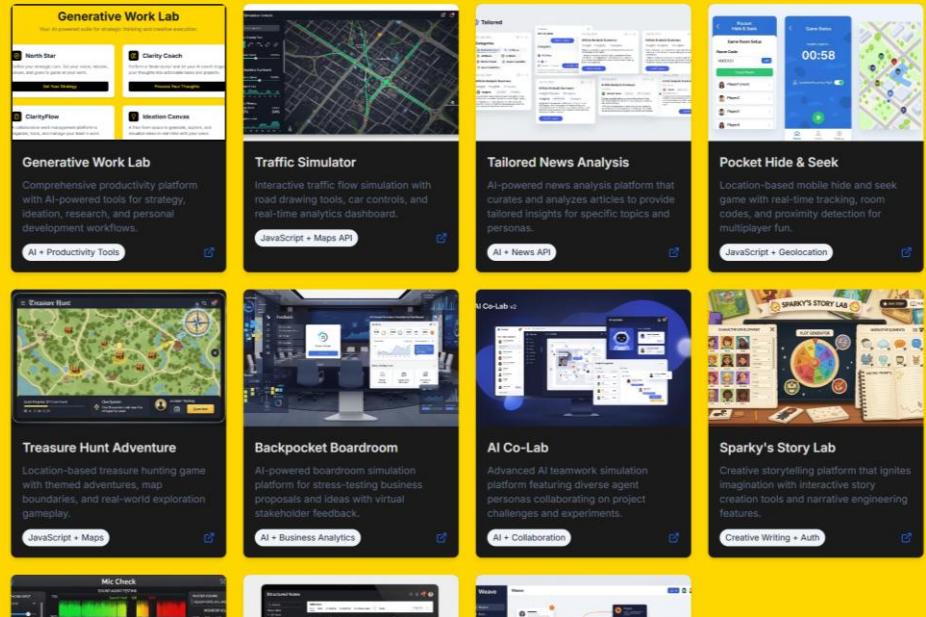
30 Apps in 30 Days

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.



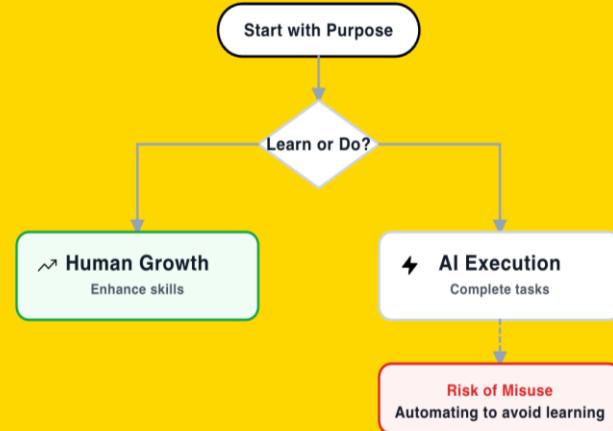
Three Core Principles

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.



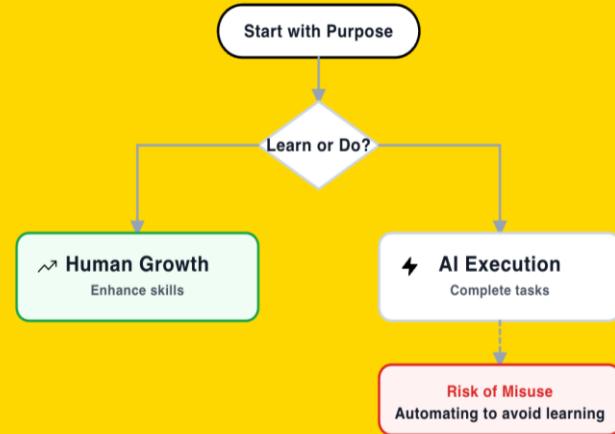
Start with Purpose:

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:

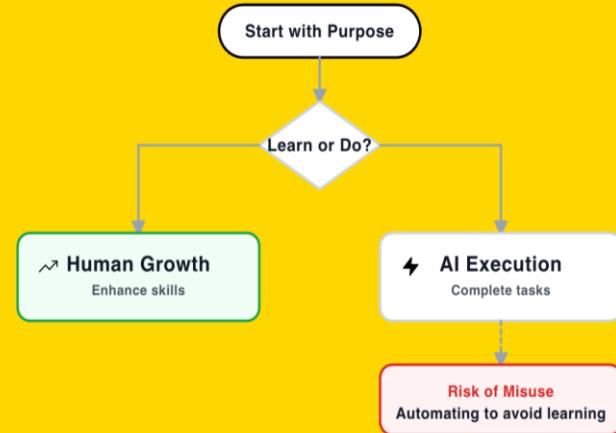
Automating when you should learn hinders growth.

Working on things that should be automated wastes time.

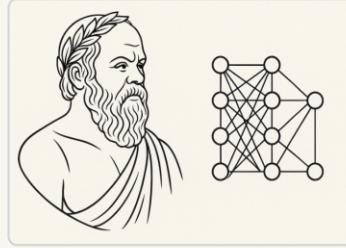


Most work is a blend of learning and doing:

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.

The screenshot shows the Sparky's Story Lab application interface. At the top, there is a purple header bar with the title "Sparky's Story Lab". On the right side of the header are three buttons: "Connected" (green dot), "Profile" (gear icon), and "Logout" (sign-out icon). Below the header, the main content area has a light blue background. The first section is titled "Your Story Collection" in bold purple text, followed by the sub-instruction "Create amazing stories with AI assistance". Below this is a purple rectangular button labeled "Ask Scout" with a small dog icon and the text "Click me for help!". The main content area is divided into two columns. The left column is titled "Your Stories" with a yellow star icon. It contains a card for a story titled "Rocky and the Crystal Cave Rescue". The story description reads: "The sparkling cave is losing its light! A kind rock named Rocky and his new friend must find the ancient song to restore the crystals' magic before...". To the right of this card is a small trash bin icon. The right column also has a "Your Stories" section and a card for a story titled "Sparky and the Lost Petal: A Mechanic...". The story description reads: "Sparky the robot loves tending the flower gardens, so he's thrilled to be on the 'Bloom Bot' float! But when a little girl gets lost among the giant...". To the right of this card is another small trash bin icon. At the bottom right of the main content area is a purple button with a plus sign and the text "Create New Story".

Sparky's Story Lab

Your Story Collection

Create amazing stories with AI assistance

Ask Scout

Click me for help!

Your Stories

Rocky and the Crystal Cave Rescue

The sparkling cave is losing its light! A kind rock named Rocky and his new friend must find the ancient song to restore the crystals' magic before...

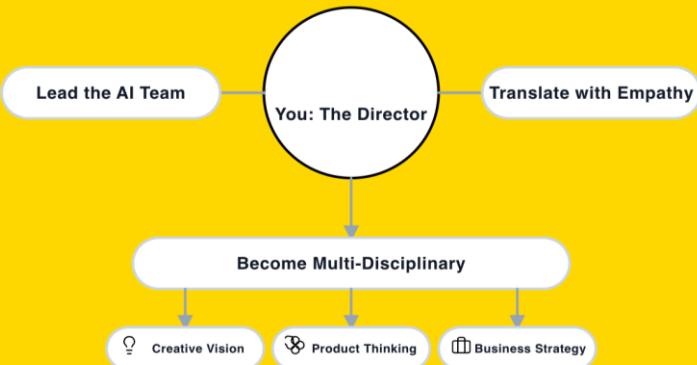
Sparky and the Lost Petal: A Mechanic...

Sparky the robot loves tending the flower gardens, so he's thrilled to be on the 'Bloom Bot' float! But when a little girl gets lost among the giant...

Create New Story

2. Learn the Art of Guidance

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

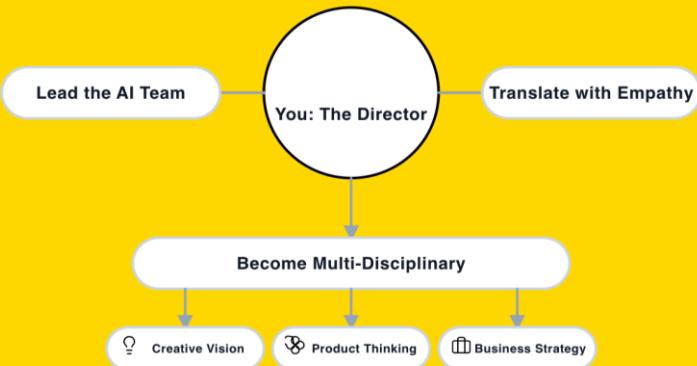


You Must Lead Your AI Team:

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 Bachelor: 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 X



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.

Age: 30
Occupation: Carpenter
Hometown: Philadelphia, PA

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," "yous 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.



Veo

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; A((Making Data AI Ready)) --> B[Structured Data Sets]; A --> C[Unstructured Data]; B --> D[Uploading All Not Optimal]; B --> E[Requires Tooling Interface]; B --> F[Not Direct Interaction]; C --> G[Very Large]; C --> H[Problems]; C --> I[Messy]; H --> J[Don't Shove All into LLM]; H --> K[Disorganized]
```

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

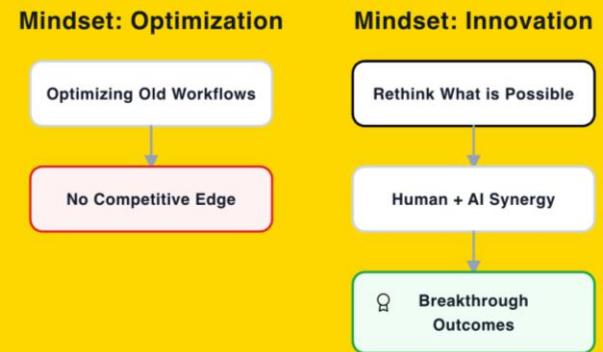
💡 Breakthrough Outcomes

No Competitive Edge

**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.



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.

The image shows the Clarity Coach application interface. At the top left is a "Back to Dashboard" button. The main title is "Clarity Coach" with a subtitle "Perform a 'brain dump' to triage your thoughts into actionable items." Below this is a "Brain Dump Sessions" section with a "+ New" button, a "Session name" input field, and a "Save" button. A dropdown menu is open, showing "Load a session...".

The central area is titled "Brain Dump" with the instruction "Throw anything at it: tasks, projects, worries, ideas. The more, the better." A text box contains the following note:

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

At the bottom of this section are buttons for "Mic: OFF", "Go to Ideation Canvas", and "Process My Thoughts".

To the right, the interface is divided into several sections:

- Messy Problems:** Complex challenges needing a structured approach. One item listed is "work project that is pretty complex and requires us to Think Through the size of different data sets". A yellow button says "Start Design Thinking".
- Simple Tasks:** Clear, concrete, single-step actions. Examples include "mowing the yard" and "cleaning up the bathrooms".
- Rumination:** Mental chatter to acknowledge and release. One entry is "worried that the presentation is taking too long and it's not that interesting".
- Someday / Maybe:** Ideas to park for later. The note states "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.

The screenshot shows the AI CoLab application interface. At the top, there is a logo with a flask icon, followed by the text "AI CoLab". To the right are two buttons: "Load" and "Run Simulation". Below this is a section titled "Welcome to AI CoLab" with the sub-instruction: "An experiment in AI teamwork. Define a project challenge, customize agent personas, and run the simulation to see how they collaborate." A large input field labeled "The Project Challenge" contains the placeholder text "Describe the project challenge for the AI teams...". To the right of this input field is a green button labeled "Random Challenge". The main area below this is a grid of nine cards, each representing an AI agent with a profile picture, name, and job title:

Dr. Evelyn Reed Principal AI Researcher Natural Language Processing	Marcus "Mark" Thorne Synergy Architect & Growth Hacker	Jia Li Backend Developer
Ben Carter Creative Technologist & AI Enthusiast	Dr. Anya Sharma Data Scientist Machine Learning	Leo Martinez Engineering Manager
Priya Singh Machine Learning Engineer Computer Vision	Olivia Chen Director of Marketing Brand Strategy	David Rodriguez Senior Product Manager

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

David Rodriguez (Leader)	Comm: Friendly Approach: Pragmatic
Collab: Collaborative Ethic: Hard-working	

Fatima Al-Jamil (Pragmatist)	Comm: Rude Approach: Pragmatic
Collab: Lone Wolf Ethic: Hard-working	

Dr. Evelyn Reed (Leader)	Comm: Friendly Approach: Pragmatic
Collab: Collaborative Ethic: Hard-working	

Team Delta

Kenji Tanaka (Specialist)	Comm: Rude Approach: Pragmatic
Collab: Lone Wolf Ethic: Hard-working	

Chloe Davis (Harmonizer)	Comm: Friendly Approach: Pragmatic
Collab: Collaborative Ethic: Hard-working	

Liam O'Connell (Innovator)	Comm: Friendly Approach: Innovative
Collab: Collaborative Ethic: Lazy	

Team Bravo

Olivia Chen (Communicator)	Comm: Friendly Approach: Innovative
Collab: Collaborative Ethic: Lazy	

Samuel "Sam" Jones (Pragmatist)	Comm: Rude Approach: Pragmatic
Collab: Lone Wolf Ethic: Hard-working	

Ben Carter (Innovator)	Comm: Friendly Approach: Innovative
Collab: Collaborative Ethic: Hard-working	

Team Echo

Priya Singh (Specialist)	Comm: Friendly Approach: Innovative
Collab: Collaborative Ethic: Hard-working	

Isabella Rossi (Communicator)	Comm: Friendly Approach: Pragmatic
Collab: Collaborative Ethic: Hard-working	

Leo Martinez (Leader)	Comm: Rude Approach: Pragmatic
Collab: Collaborative Ethic: Hard-working	

Reviewer	Reviewee	Reviewer	Reviewee
Reviewer	Reviewee	Reviewer	Reviewee
David Rodriguez	Fatima Al-Jamil	David Rodriguez	Dr. Evelyn Reed
Score	8/10	Score	9/10
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.	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: 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.	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	Reviewee	Reviewer	Reviewee
Fatima Al-Jamil	Dr. Evelyn Reed	Dr. Evelyn Reed	David Rodriguez
Score	9/10	Score	9/10
Strength: Understands the technical requirements and consistently refined her suggestions to be more robust and implementable, aligning with proper engineering practices.	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: She wastes time attempting to collaborate on writing, which is inefficient. She should focus on providing direct technical input when requested.	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.		



Contacts

[View All](#)[Favorites](#)[Recent](#)

Theodore ...
3 weeks ago



Alexandra ...
4 weeks ago

Reminders

[View All](#)[Upcoming](#)[Overdue](#)

Meet on Generative W...
Tue, Sep 9 ✓

[Home](#)[Contacts](#)[Reminders](#)[Snippets](#)

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.

The screenshot shows the Weave application's "Structured Notes" feature. On the left is a sidebar with the following sections and options:

- New Voice Note** (button)
- Create Template** (button)
- ORGANIZATION**
 - All Notes (selected)
 - Starred
 - Collections
 - Chat
 - Drafts
 - Templates
 - Archive
- APPEARANCE**
 - Professional (selected)

At the bottom of the sidebar, there is a contact card for "Austin Gerig" with the email "austingering@gmail.com".

Additional Demos

30 Day Apps

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

