

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

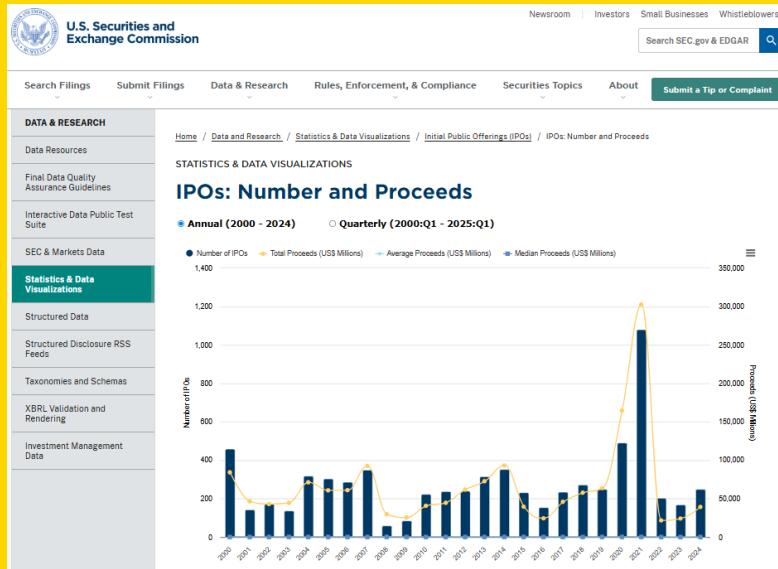
~Accenture

+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 figure consists of two screenshots of SEC websites. The top screenshot shows the 'Investment Adviser Public Disclosure' website with a search bar for 'Individual' and 'Firm' information. The bottom screenshot shows the 'EDGAR' search interface with a search bar and buttons for 'SEARCH', '+ more search options', and 'Clear all'.

Investment Adviser Public Disclosure

Welcome to the Investment Adviser Public Disclosure website

INDIVIDUAL FIRM

Individual Name/CRD# Firm Name or CRD/SEC# (optional) City, State or ZIP (optional)

For a complete compilation of Investment Adviser Firms or individuals, download the [Investment Adviser Data](#).

SEC.gov | EDGAR

FAQ Other search tools

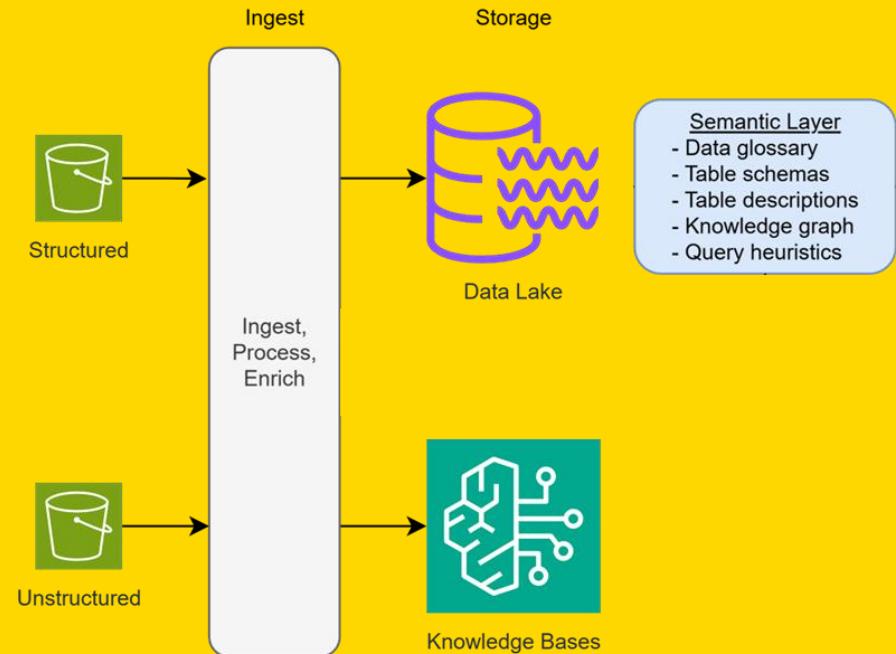
EDGAR

Search by keyword, ticker, company name, CIK number or industry

SEARCH + more search options Clear all

Two principles for AI-Ready data

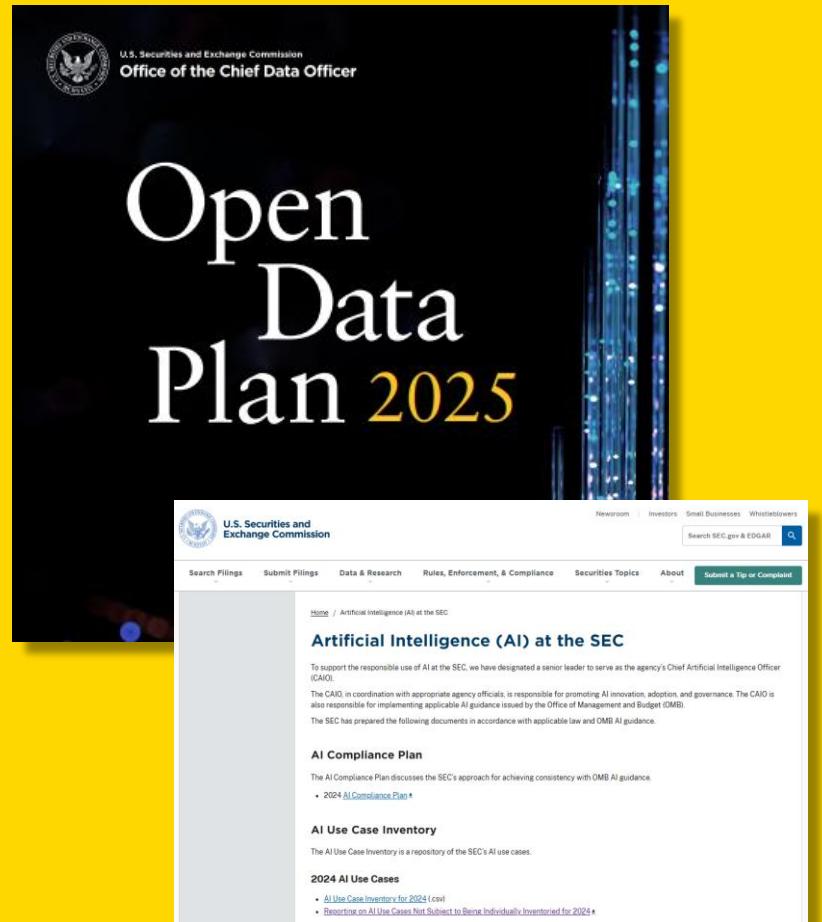
- Make structured data “machine understandable” through a semantic layer
- Curate unstructured data in vector databases that serve AI throughout your organization



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 displays two screenshots of the SEC's AI-related initiatives. The top screenshot is the "Open Data Plan 2025" page, featuring a dark background with the title in large white and gold letters. The bottom screenshot is the "AI Use Case Inventory" page, showing a list of documents and a search bar.

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

Open Data Plan 2025

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](#)

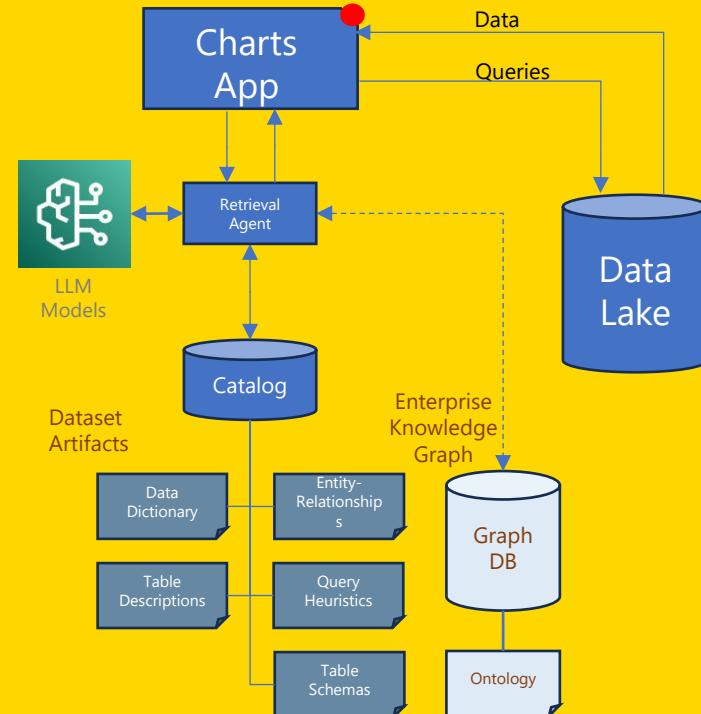
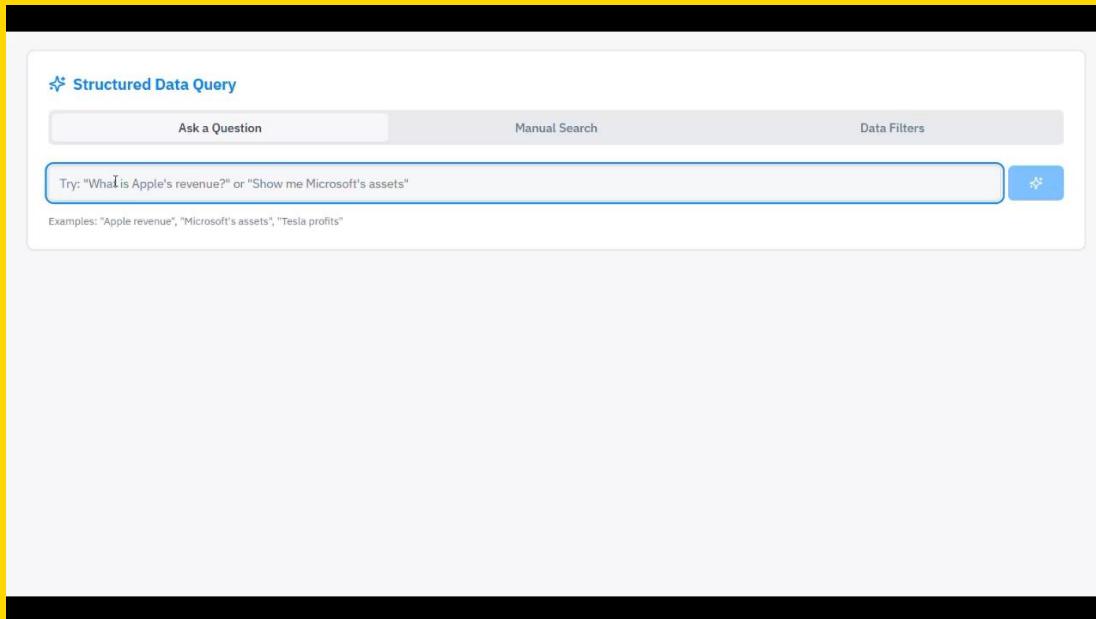
AI Use Case Inventory

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

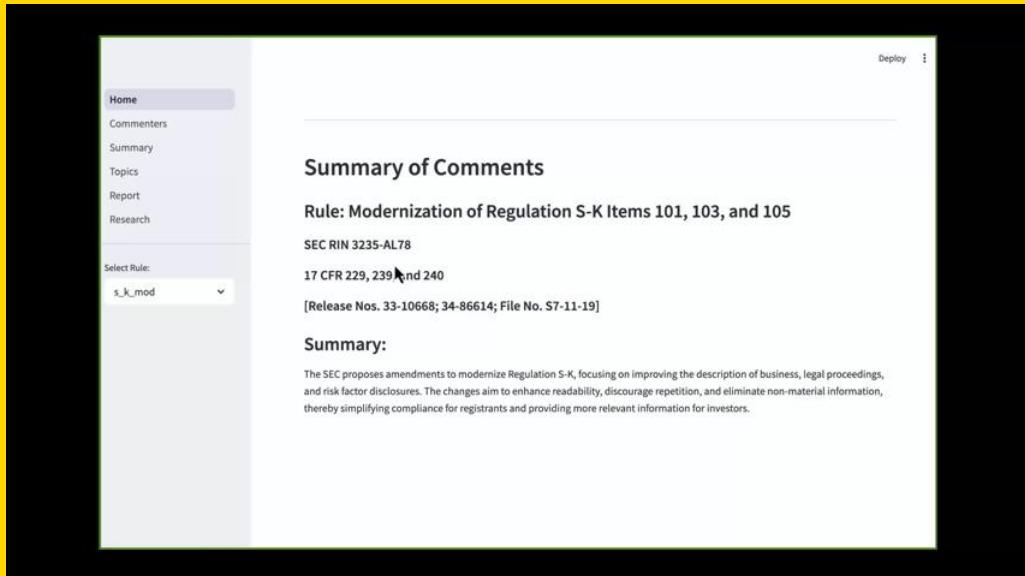
2024 AI Use Cases

- [AI Use Case Inventory for 2024 \(cav\)](#)
- [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

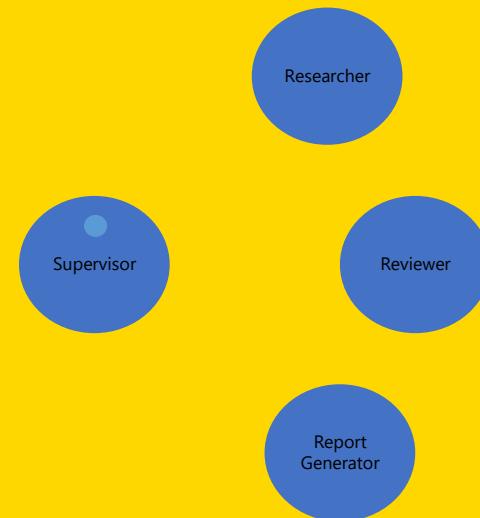


The screenshot shows a web-based application for analyzing comment letters. The interface includes a sidebar with navigation links: Home, Commenters, Summary, Topics, Report, and Research. A dropdown menu for 'Select Rule' is open, showing 's_k_mod'. The main content area is titled 'Summary of Comments' and displays the following information:

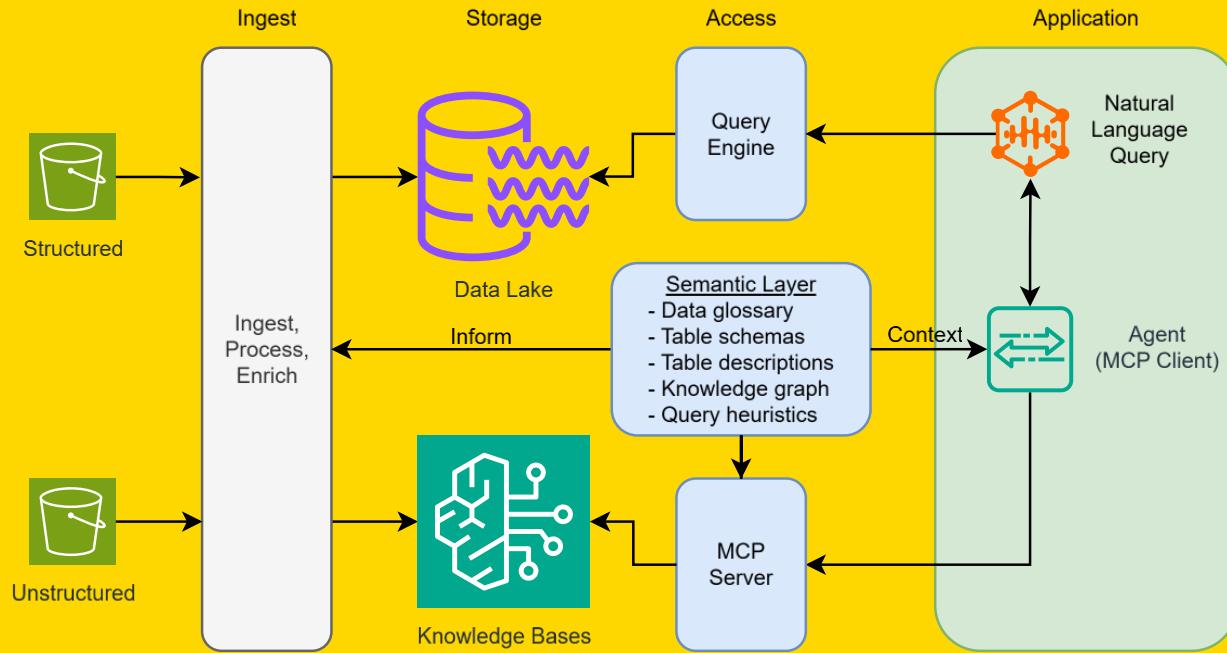
- Rule:** Modernization of Regulation S-K Items 101, 103, and 105
- SEC RIN:** 3235-AL78
- 17 CFR:** 229, 239, and 240
- [Release Nos.:** 33-10668; 34-86614; File No. S7-11-19]

Summary:

The SEC proposes amendments to modernize Regulation S-K, focusing on improving the description of business, legal proceedings, and risk factor disclosures. The changes aim to enhance readability, discourage repetition, and eliminate non-material information, thereby simplifying compliance for registrants and providing more relevant information for investors.



Architecture & Implementation



Part II: Working with AI

Learning by Doing

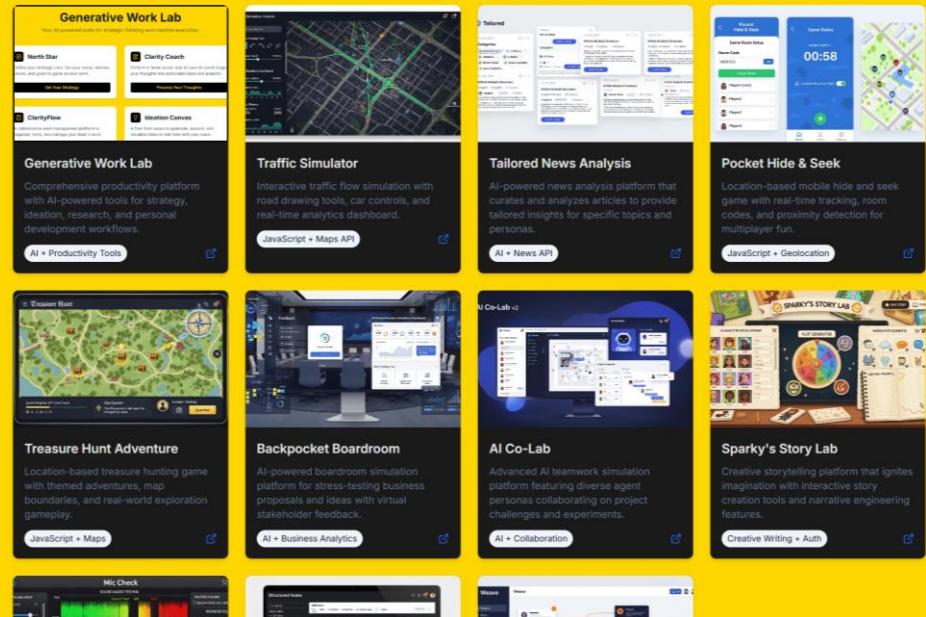
30 Apps in 30 Days

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.



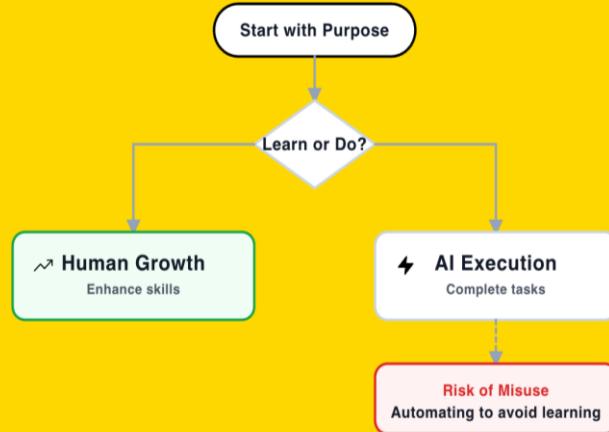
Three Core Principles

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.



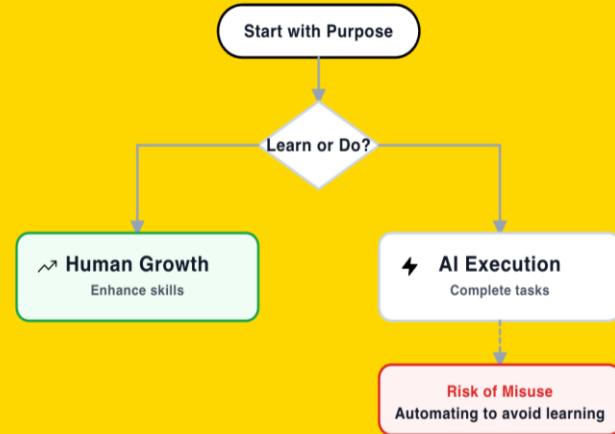
Start with Purpose:

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:

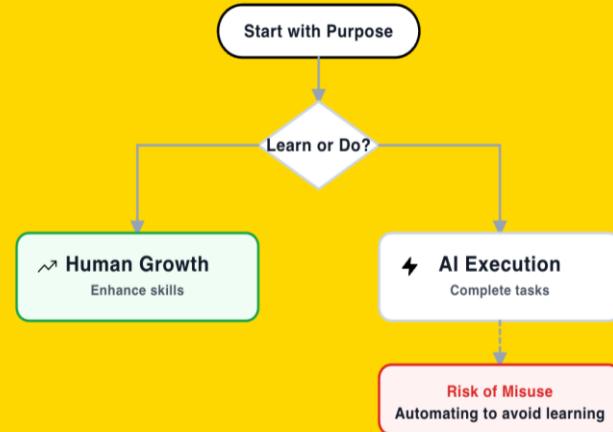
Automating when you should learn hinders growth.

Working on things that should be automated wastes time.

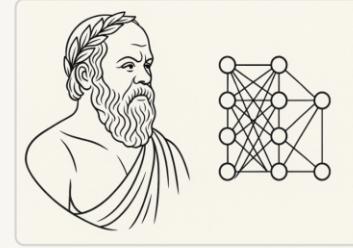


Most work is a blend of learning and doing:

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?

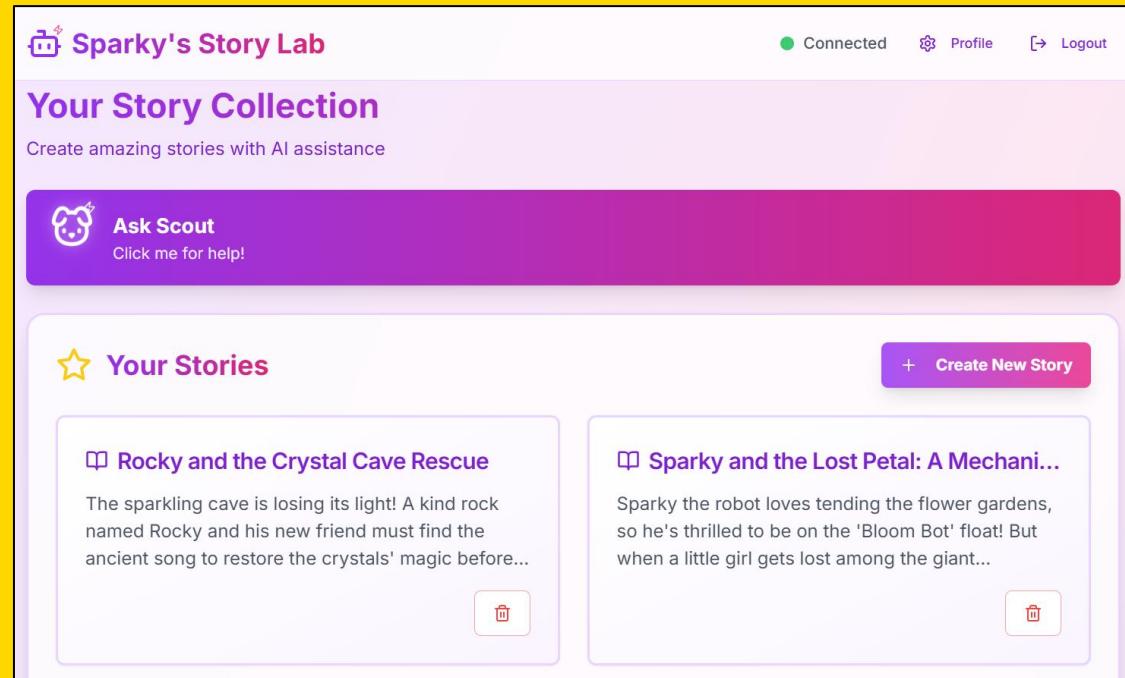
Ah, my most excellent friend, you raise a question that is worthy of long inquiry. But before we hasten to prescribe what you *should* study, I find myself puzzled—for such advice surely depends on what you believe to be the *purpose* of studying, does it not?

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 text 'Sparky's Story Lab' and a small robot icon. To the right of the header are three buttons: 'Connected' (green dot), 'Profile' (gear icon), and 'Logout' (Logout icon). Below the header, the main content area has a light blue background. The title 'Your Story Collection' is displayed in a large, bold, purple font. Below the title, the subtext 'Create amazing stories with AI assistance' is shown in a smaller, purple font. A purple button labeled 'Ask Scout' with a small robot icon and the text 'Click me for help!' is located on the left side. The main content area is divided into two sections: 'Your Stories' and 'Create New Story'. The 'Your Stories' section contains two story cards. The first card is titled 'Rocky and the Crystal Cave Rescue' and features a short description: '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...'. The second card is titled 'Sparky and the Lost Petal: A Mechan...' and features a short description: '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...'. Each story card has a small trash can icon in the bottom right corner.

Sparky's Story Lab

Connected Profile Logout

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...

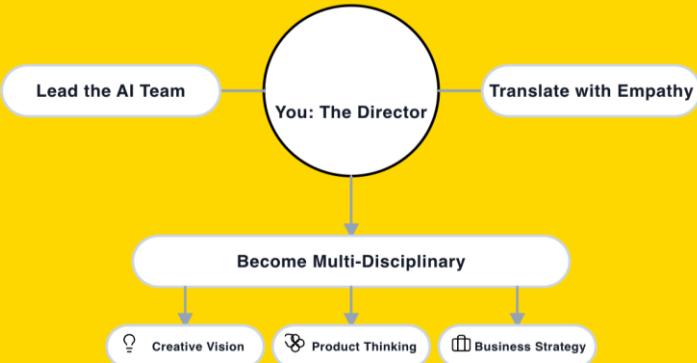
Sparky and the Lost Petal: A Mechani...

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.

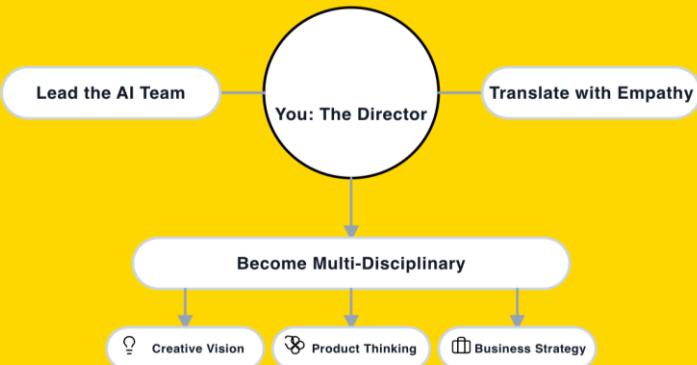


You Must Lead Your AI Team:

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 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 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.



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[Not Direct Interaction]; B --> F[Very Large]; C --> G[Problems]; C --> H[Messy]; C --> I[Disorganized]; J[Requires Tooling Interface] --- D
```

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

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.

The image shows the Clarity Coach application interface. At the top, there is a navigation bar with a "Back to Dashboard" button and a title "Clarity Coach". Below the title, a sub-instruction reads: "Perform a 'brain dump' to triage your thoughts into actionable items."

Brain Dump Sessions: A section for managing sessions. It includes a "+ New" button, a "Session name" input field, and a "Save" button. A dropdown menu is open, showing the option "Load a session...".

Brain Dump: A large text area for inputting thoughts. The placeholder text is: "Throw anything at it: tasks, projects, worries, ideas. The more, the better." Below this, a user's input is shown: "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: A button indicating the microphone is disabled.

Process My Thoughts: A large, prominent black button.

Messy Problems: A section for complex challenges. It includes a sub-instruction: "Complex challenges needing a structured approach." and a note: "work project that is pretty complex and requires us to Think Through the size of different data sets". A "Start Design Thinking" button is present.

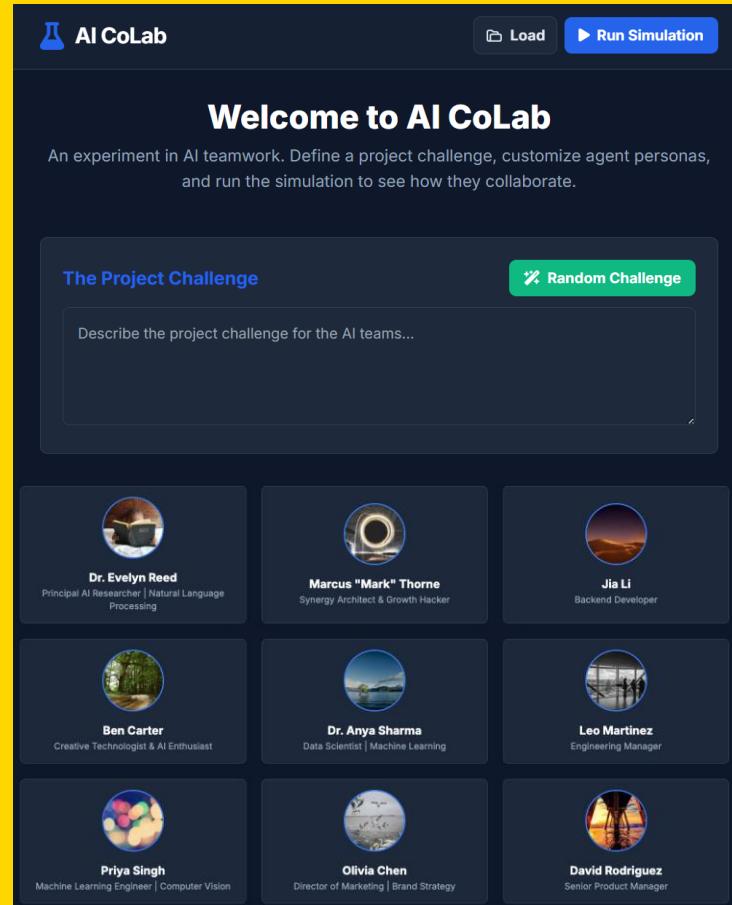
Simple Tasks: A section for clear, single-step actions. It includes a sub-instruction: "Clear, concrete, single-step actions." and a list: "mowing the yard" and "cleaning up the bathrooms". Each item has an "Inbox" dropdown and a "Send" button.

Rumination: A section for mental chatter. It includes a sub-instruction: "Mental chatter to acknowledge and release." and a note: "worried that the presentation is taking too long and it's not that interesting".

Someday / Maybe: A section for ideas to park for later. It includes a sub-instruction: "Ideas to park for later." and a note: "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 interface. At the top, there is a logo with a flask icon, the text "AI CoLab", and 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 "Random Challenge" button is also present. The main area is titled "The Project Challenge" and contains a text input field with the placeholder "Describe the project challenge for the AI teams...". 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)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Pragmatic

Fatima Al-Jamil (Pragmatist)
Collab: Lone Wolf
Ethic: Hard-working

Comm: Rude
Approach: Pragmatic

Dr. Evelyn Reed (Leader)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Pragmatic

Team Delta

Kenji Tanaka (Specialist)
Collab: Lone Wolf
Ethic: Hard-working

Comm: Rude
Approach: Pragmatic

Chloe Davis (Harmonizer)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Pragmatic

Liam O'Connell (Innovator)
Collab: Collaborative
Ethic: Lazy

Comm: Friendly
Approach: Innovative

Team Bravo

Olivia Chen (Communicator)
Collab: Collaborative
Ethic: Lazy

Comm: Friendly
Approach: Innovative

Samuel "Sam" Jones (Pragmatist)
Collab: Lone Wolf
Ethic: Hard-working

Comm: Rude
Approach: Pragmatic

Ben Carter (Innovator)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Innovative

Team Echo

Priya Singh (Specialist)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Innovative

Isabella Rossi (Communicator)
Collab: Collaborative
Ethic: Hard-working

Comm: Friendly
Approach: Pragmatic

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
Dr. Evelyn Reed

Reviewee
David Rodriguez

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
David Rodriguez

Reviewee
Dr. Evelyn Reed

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.

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



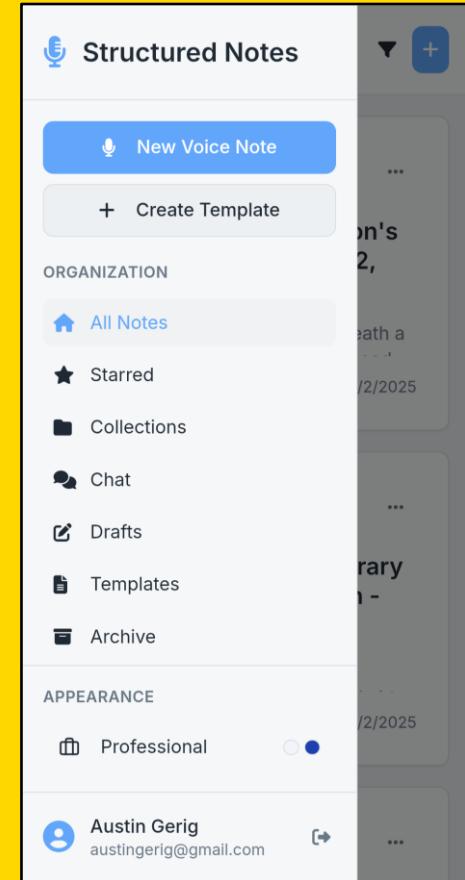
Snippets

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.



Additional Demos

30 Day Apps

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

