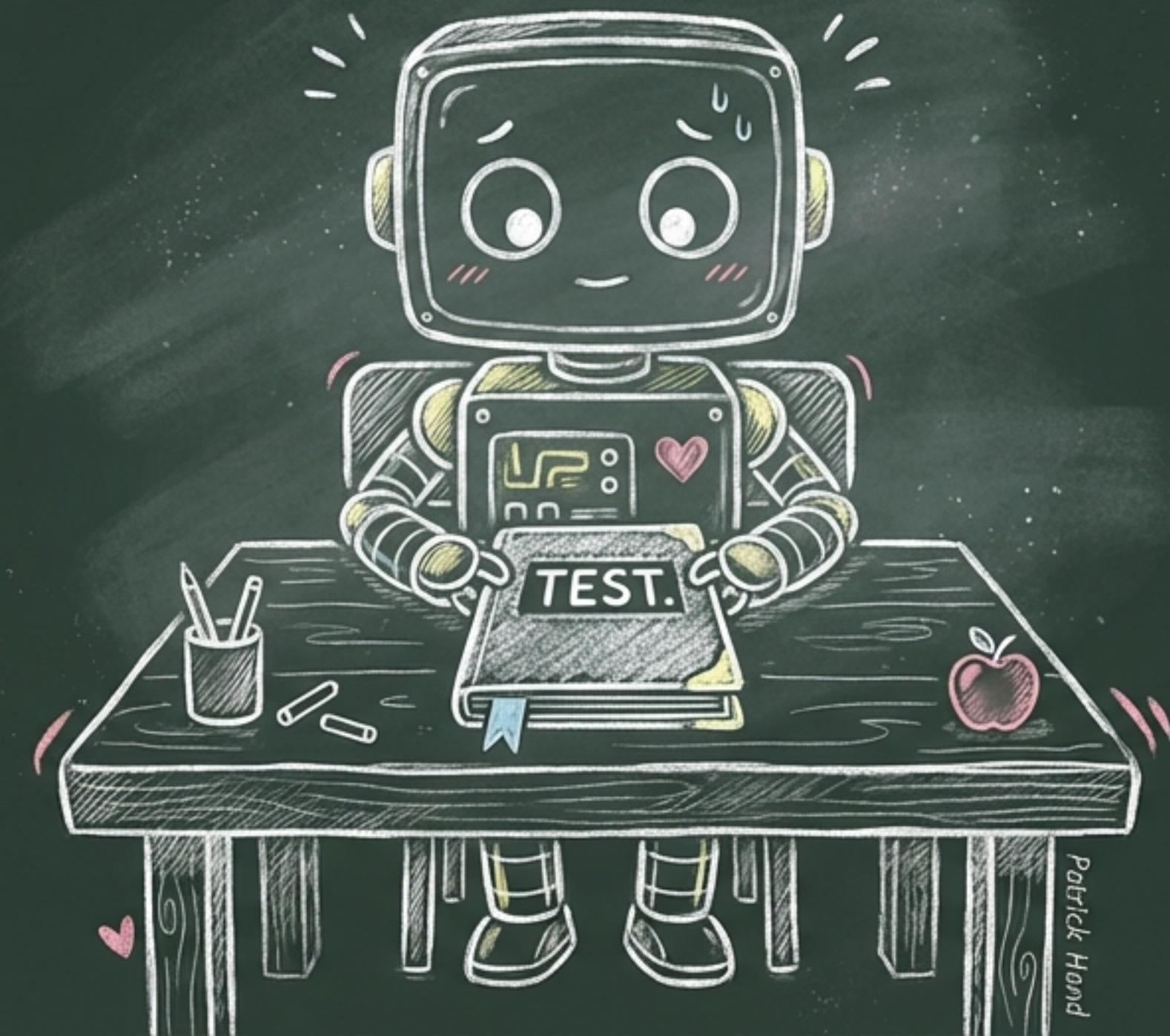


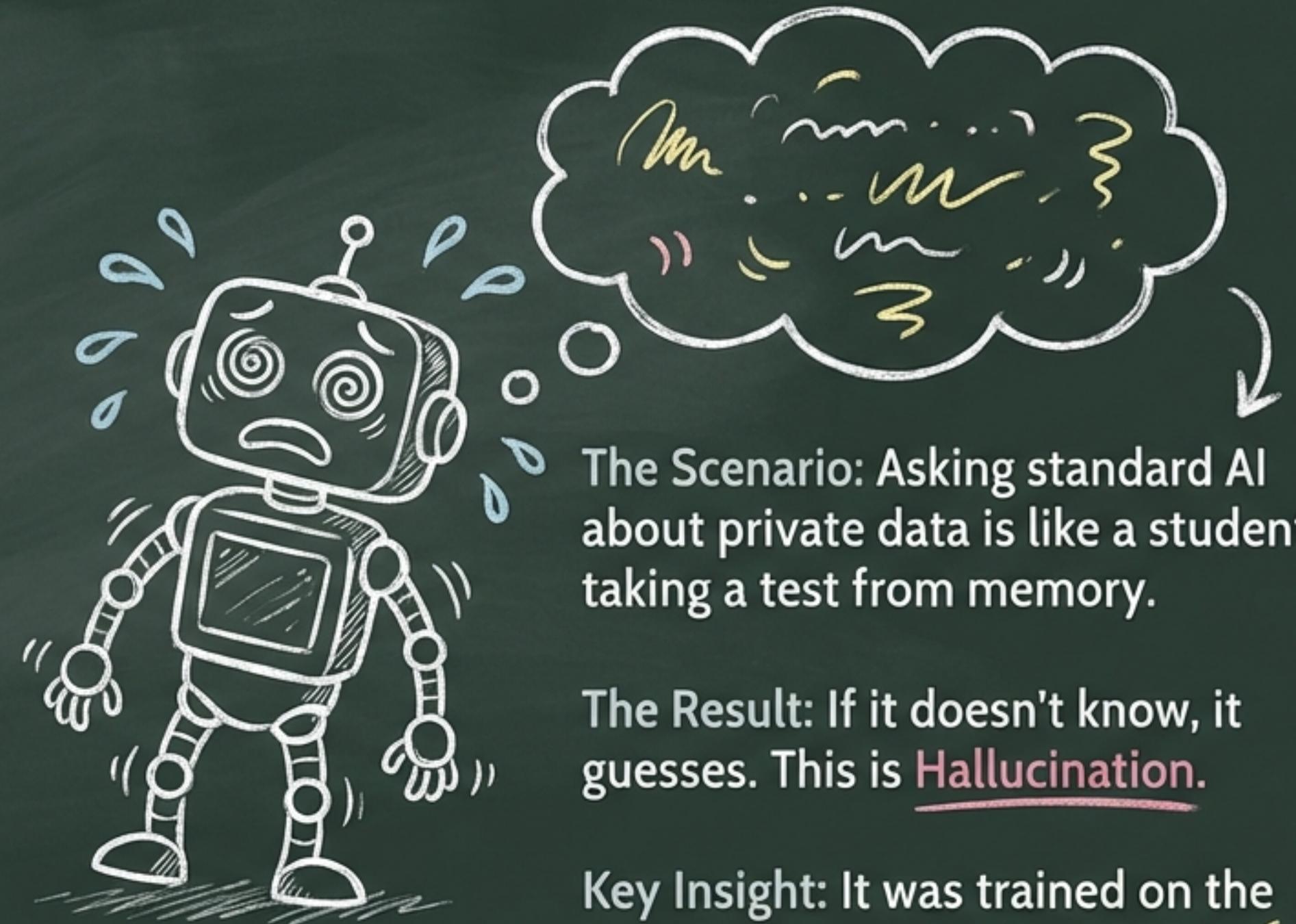
RAG: The Open-Book Test for AI

A simple guide to Retrieval Augmented Generation (without the headaches).



Class is in session.

The Problem: The Pop Quiz



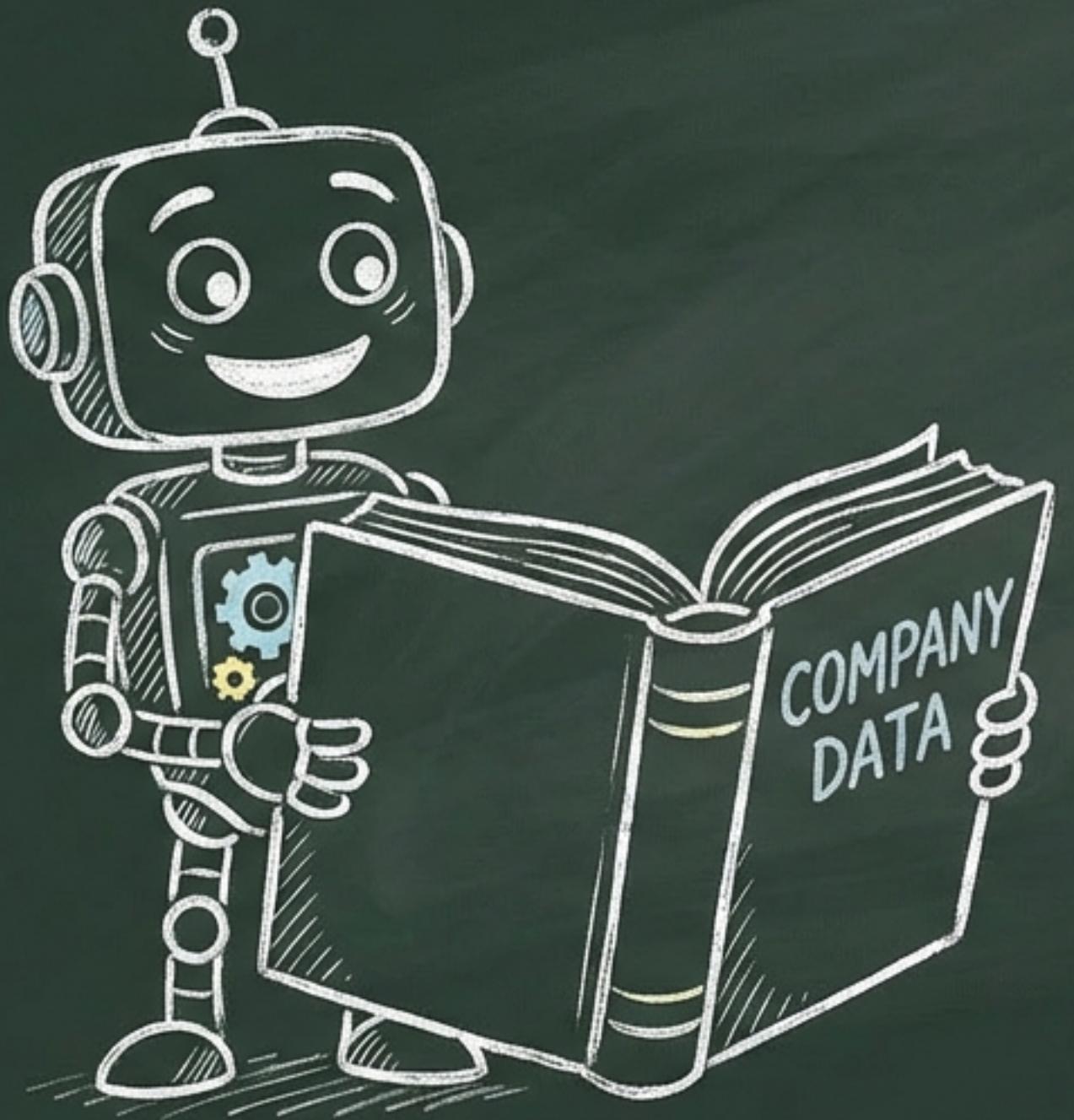
The Scenario: Asking standard AI about private data is like a student taking a test from memory.

The Result: If it doesn't know, it guesses. This is Hallucination.

Key Insight: It was trained on the internet, not your internal files.

----- ANALOGY ALERT -----
Standard AI = Student guessing without studying.

The Solution: Let the Student Cheat (Legally)



- The Shift: What if we let the AI look up the answer before speaking?
- The Concept: This is RAG (Retrieval Augmented Generation).
- The Definition: Giving the AI access to a library of your specific info so it doesn't have to guess.

----- ANALOGY ALERT -----
RAG = Student taking an Open-Book Test. -----

Breaking Down the Acronym

R

Retrieval



A

Augmented



G

Generation



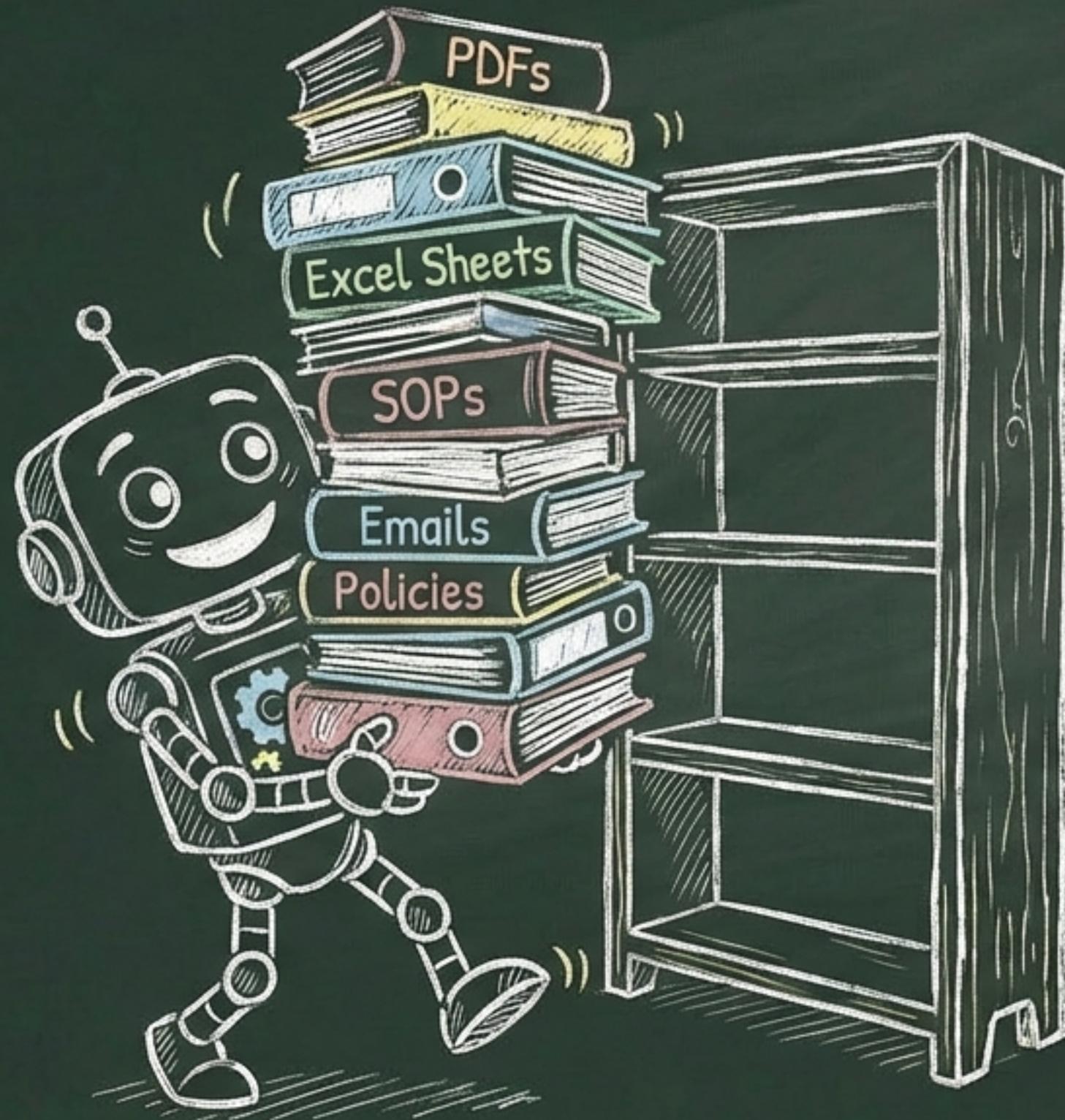
Searching for the
correct info first.

Adding that info
to the question.

Writing a clear
answer based on facts.

Find the info. Add it to the prompt. Write the answer.

Step 1: Building the School Library (Collection)



- Before the AI can answer, we need to give it a knowledge base.
- We collect all your messy documents and store them in one place.
- This becomes your Source Truth.

Step 2: Breaking it Down (Chunking)



- The Issue: **Searching a 200-page PDF** takes too long.
- The Fix: We break large documents into **small “chunks”** (chapters or paragraphs).
- Why? It's easier to **find one specific paragraph** than to **re-read an entire textbook**.

ANALOGY ALERT

Chunking = Breaking the textbook into bite-sized chapters.

Step 3: Translating to Robot Language (Embeddings)



Revenue increased.

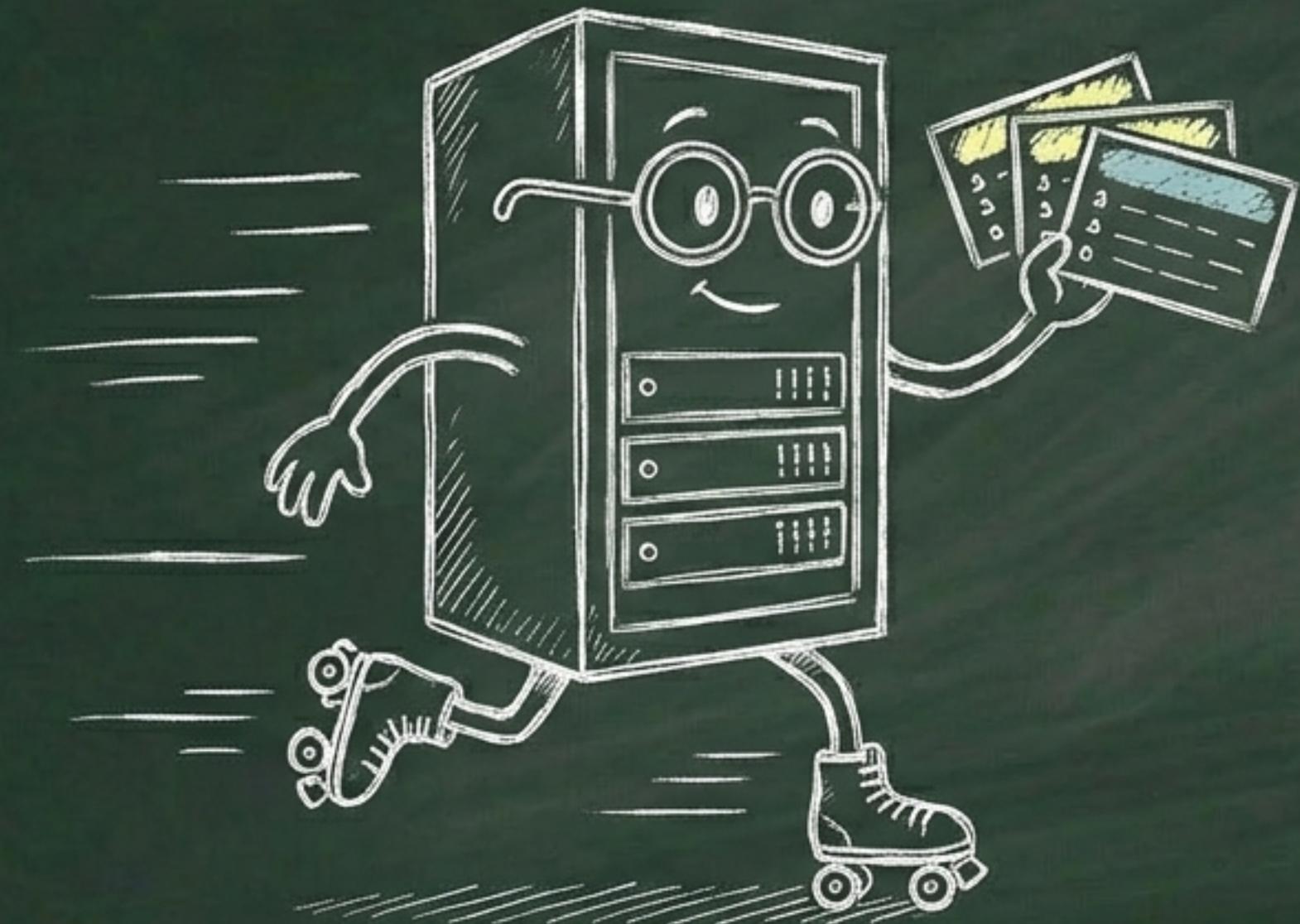
- **The Tech:** Computers don't understand English; they understand numbers.
- **The Process:** We convert every text chunk into a unique number pattern called an **Embedding**.
- **The Magic:** This lets the AI compare the meaning of ideas, not just match keywords.

ANALOGY ALERT

Embeddings = Giving every paragraph a unique barcode.



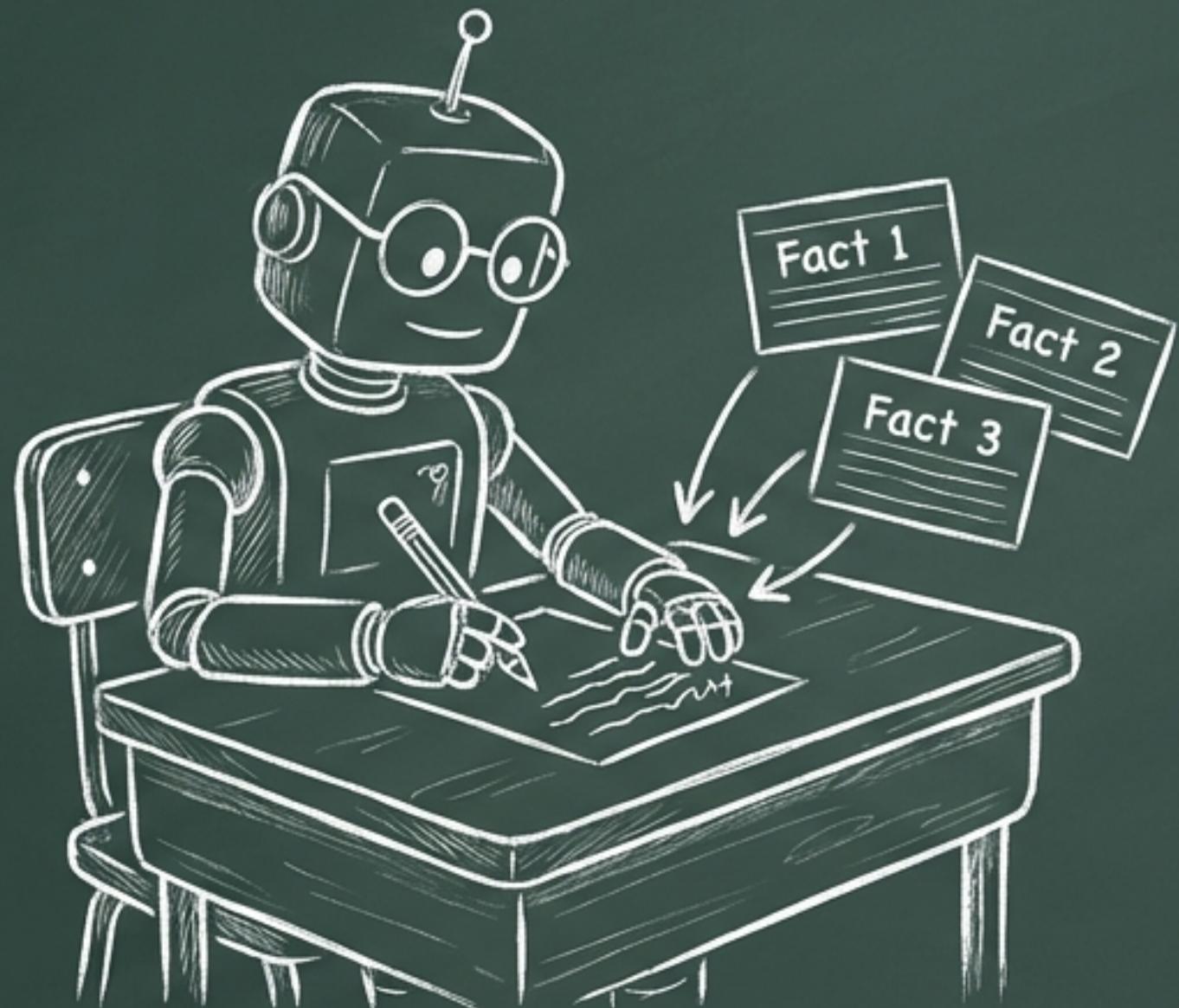
Step 4: The Super Librarian (Vector Database)



- We store all those “barcodes” in a Vector Database.
- The Search: When you ask a question (“Why did revenue drop?”), the system finds the matching number patterns.
- Result: It retrieves only the most relevant paragraphs.

ANALOGY ALERT ----- ★
Vector DB = A Google Search inside your company files. |

Step 5: Writing the Essay (Generation)



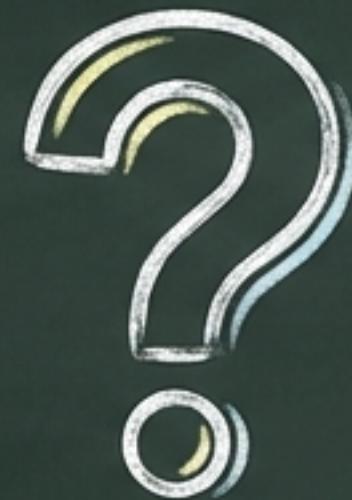
- The Hand-off: The system gives the AI two things: 1) Your Question and 2) The Relevant Facts it found.
- The Output: The AI reads the facts and generates a structured, clean answer in plain English.
- Crucial Difference: It isn't guessing. It is summarizing the facts provided.

| ANALOGY ALERT ⭐

| Generation = Explaining the answer after reading the page.

Report Card: Standard AI vs. RAG

The Guesser (Standard AI)



- ✓ - Answers from memory (Internet training).
- ✓ - Can hallucinate (make things up).
- ✓ - General knowledge only.

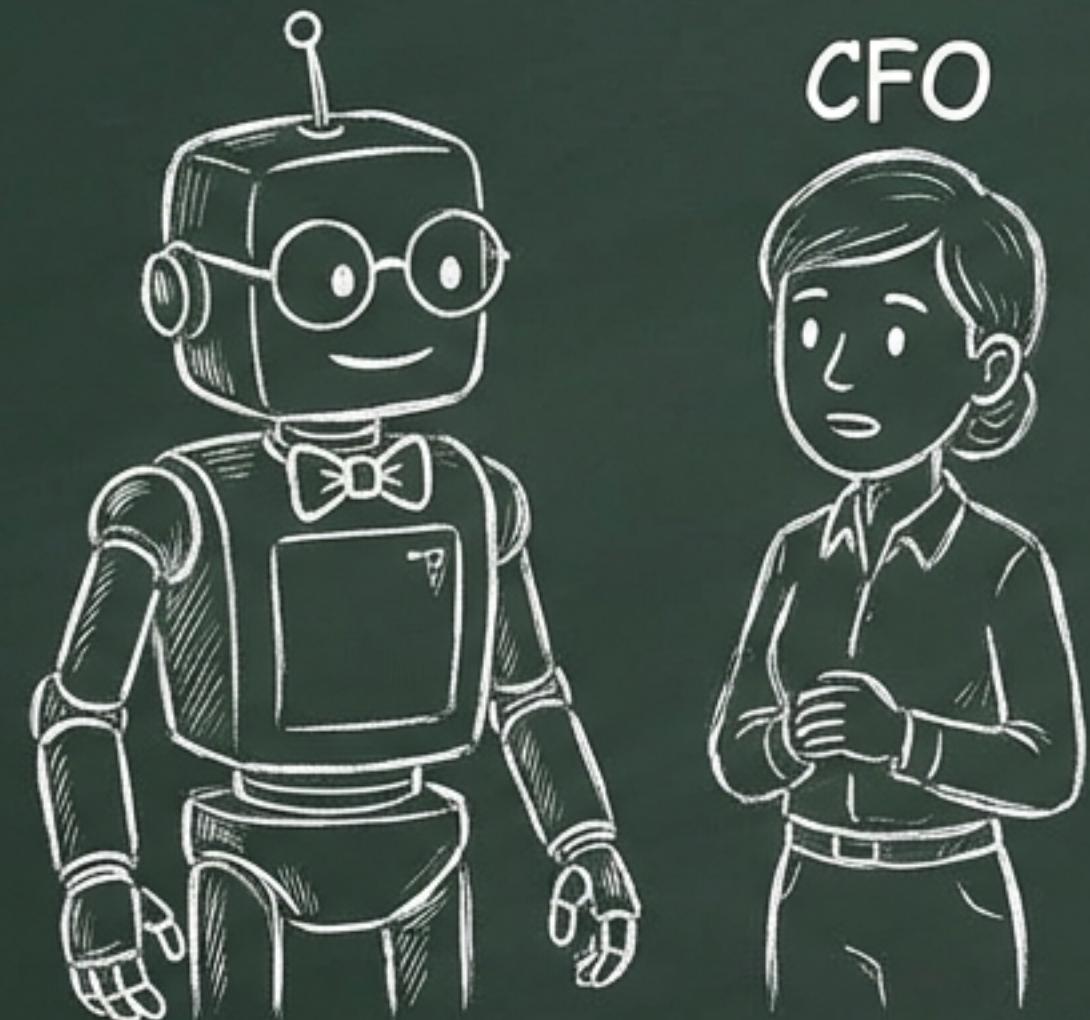
The Researcher (RAG System)



- ✓ - Answers from YOUR documents.
- ✓ - Reduces hallucination (facts only).
- ✓ - Context-specific knowledge.

RAG combines the brains of AI with the facts of your files.

Graduation Day: The Business Application



CFO

Why did engineering costs exceed forecast in July?

Costs were up 12% due to one-time restructuring expenses found in the July Variance Report.

- Without RAG: Vague answers ('Costs vary due to many factors').
- With RAG: The system finds the specific 'Variance Report' and 'Project Excel'.
- Key Benefit: Specificity and citation.

Why Companies Love RAG



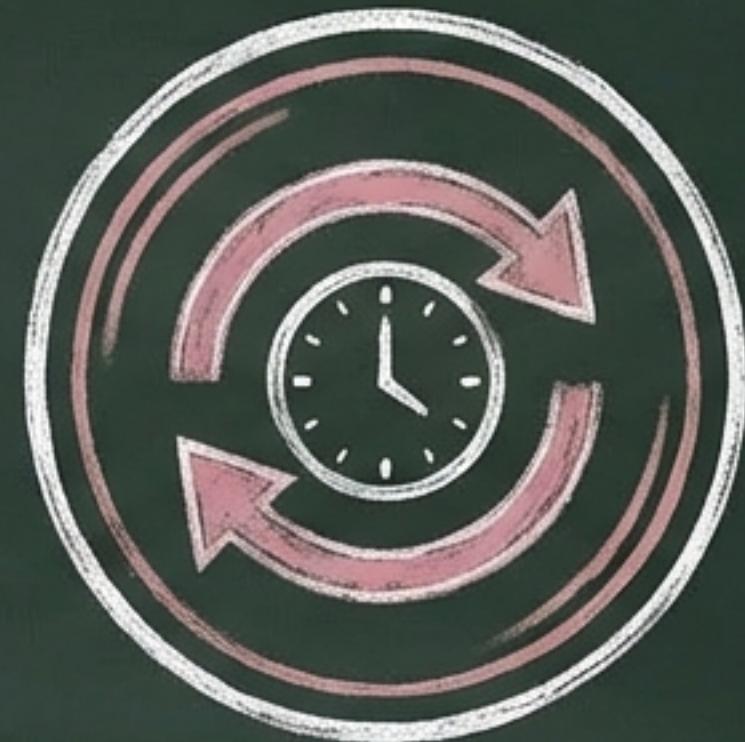
Accuracy

(No guessing).



Security

(Data stays private).



Freshness

(Instant updates).

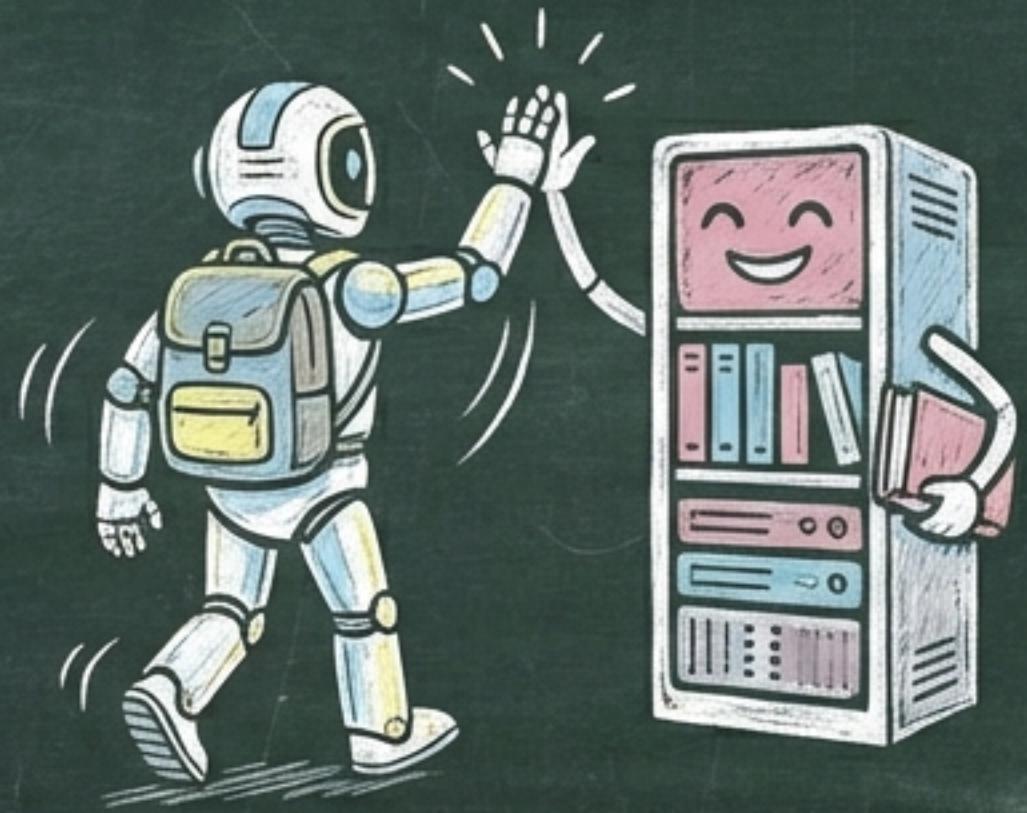
No need to ‘retrain’ the brain. Just add a new document to the folder, and the AI knows it immediately.

The Golden Rule

RAG is like giving AI access to your company's library so it can look up the correct answer before speaking.

Smart Search + Smart Generation = RAG

Class Dismissed!



- Don't let your AI sweat the pop quiz.
- Give it an open book.



Patrick Hand

No students were hallucinated in the making of this deck.