

# Ground Rules

# HOUSE RULES



Mic is on mute by default



Cameras on, if you can!

PS. This session is being recorded.  
Copy of the recording will be shared to you right after the session.

REC

DECODING  
DATA SCIENCE

Ask questions by.



Raising your virtual hands

click Reactions , then click Raise Hand 

Writing them on the chat



# Agenda

---

- Zoom & Streaming Ground Rules
- RAG
- Slido Q & A
- Challenge
- Slido Q & A



EXPERIENCE BUILDING AND LECTURING AT



accenture >

DELL Technologies

مكتبة  
MAJID AL FUTTAIM

Microsoft



**Software Engineer**

2005 - 2010

absolutdata

-an lifogain company



**Data Scientist**

2012 - 2012



**Data Scientist**

2012 - 2014



**Lead Data Scientist**

2014 - 2021



MAJID AL  
FUTTAIM

**Data Science Manager**

2021 - 2022



DECODING  
DATA SCIENCE

**Founder**

2022 - Present



Yale GAP



# 70%

## The Quiet Revolution Powering Modern AI.

A staggering **70% of AI projects** being built today use a technique called Retrieval-Augmented Generation.

It's the secret ingredient making AI smarter, more reliable, and more useful in the real world.

# First, Meet the Brilliant Genius.

A Large Language Model (LLM) is like a genius. It has read billions of documents—practically the entire internet.



Writes fluently



Summarizes complex topics



Answers incredible questions



# But This Genius is Stuck in the Past.

An LLM's knowledge is **frozen in time**.  
It's trained on a massive dataset, but that  
training has a cutoff date.



It knows everything about last June, but nothing about last July.

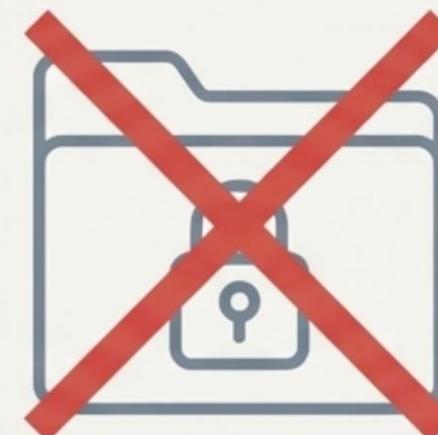


# This Creates Two Critical Knowledge Gaps.



**Blind to Recent Events.**

The Genius can't see your company's latest sales numbers, your internal HR policies, or what happened in the news last week.



**Zero Access to Private Data.**

# So, When It Doesn't Know, It Makes Things Up.



This is called hallucination. The AI is designed to be helpful, so instead of admitting “I don’t know,” it tries to piece together an answer that can be completely wrong.

\*The LLM hallucinates because it does not want to look like a fool.\*

# The Genius Needs a Partner. Meet the Super Librarian.



**\*\*RAG\*\*** (Retrieval-Augmented Generation) is the perfect partner for our Genius. It's a librarian with access to a special, constantly updated library.

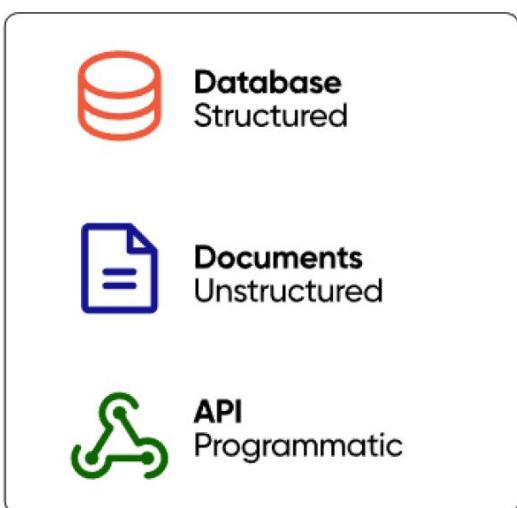
# The Librarian Finds the Facts. The Genius Crafts the Answer.



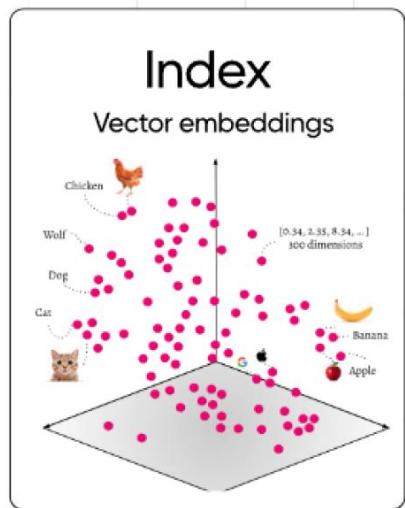
The Librarian finds the exact right page—the perfect document—and hands it to the Genius *before* it answers.

**Official Definition:** RAG optimizes an LLM's output by letting it look at an external, up-to-date knowledge base before it answers a question.

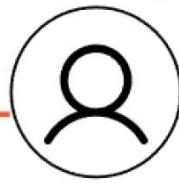
# Your Data



Load →



← Query 💬



User

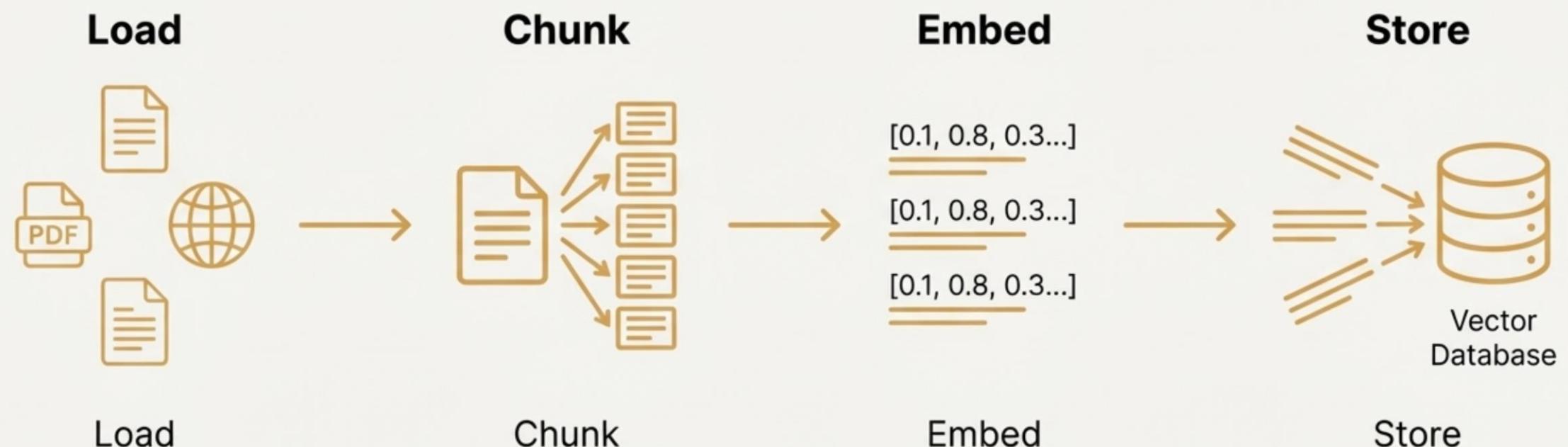
↑ response



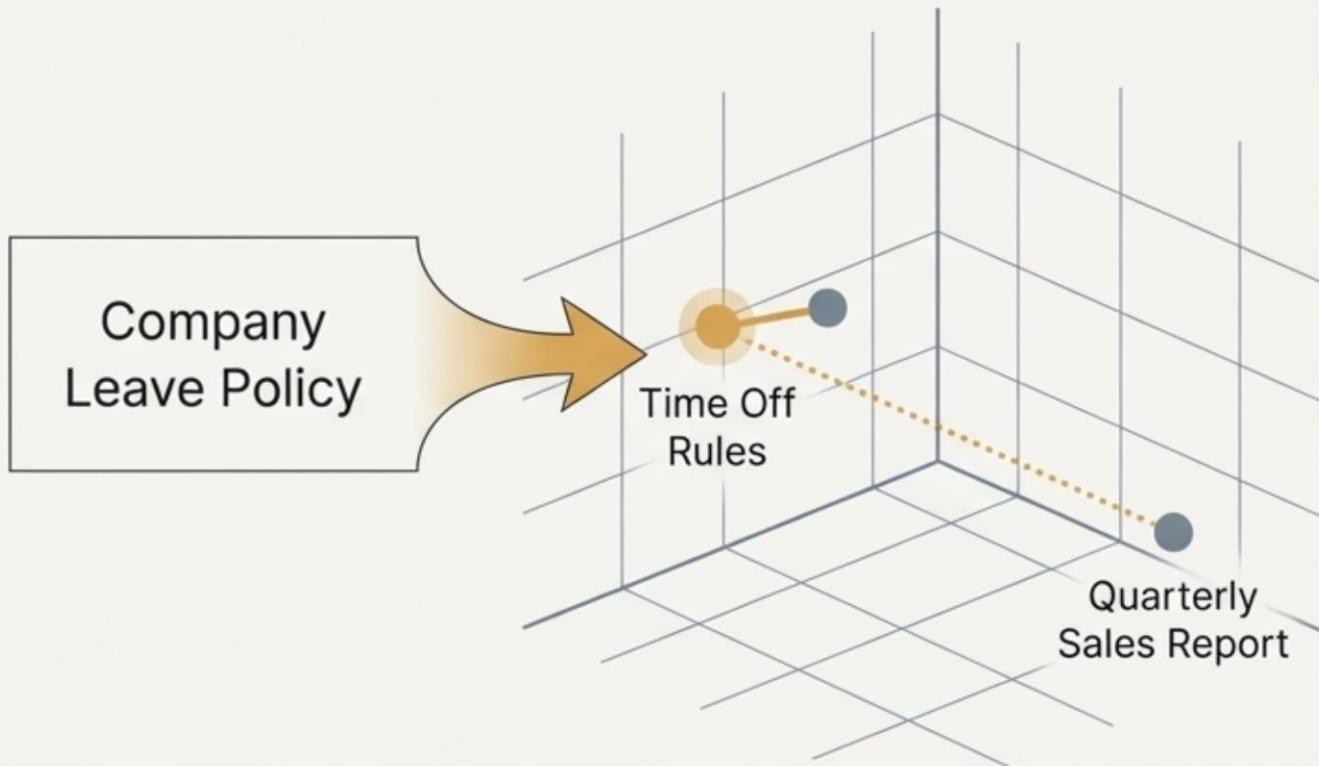
Prompt + query  
+ relevant data

# Step 1: Organizing the Library

Think of it as creating a super-fast digital card catalog.

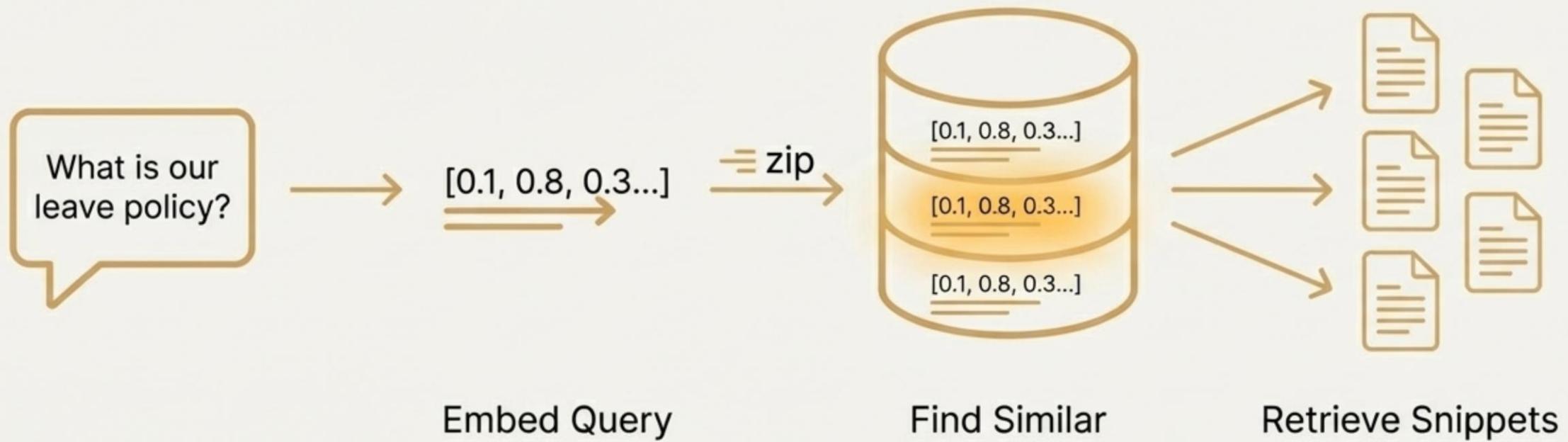


# The Secret Code: Turning Words into Meaning.

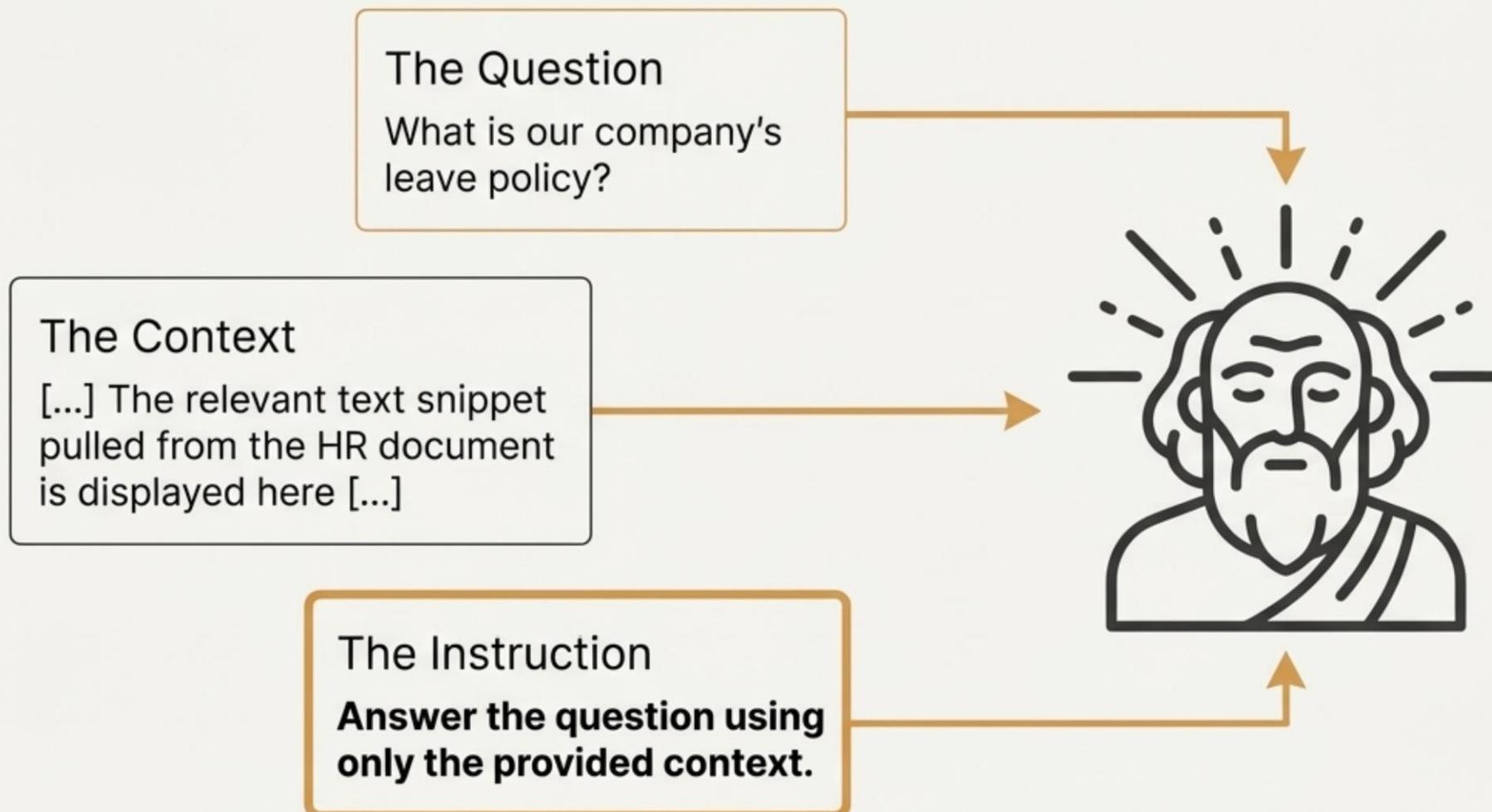


"Embedding" converts text into a series of numbers (vectors). This allows the system to search for **related concepts and meaning**, not just matching keywords. It's incredibly fast and powerful.

# Step 2: Finding the Right Book, Instantly.

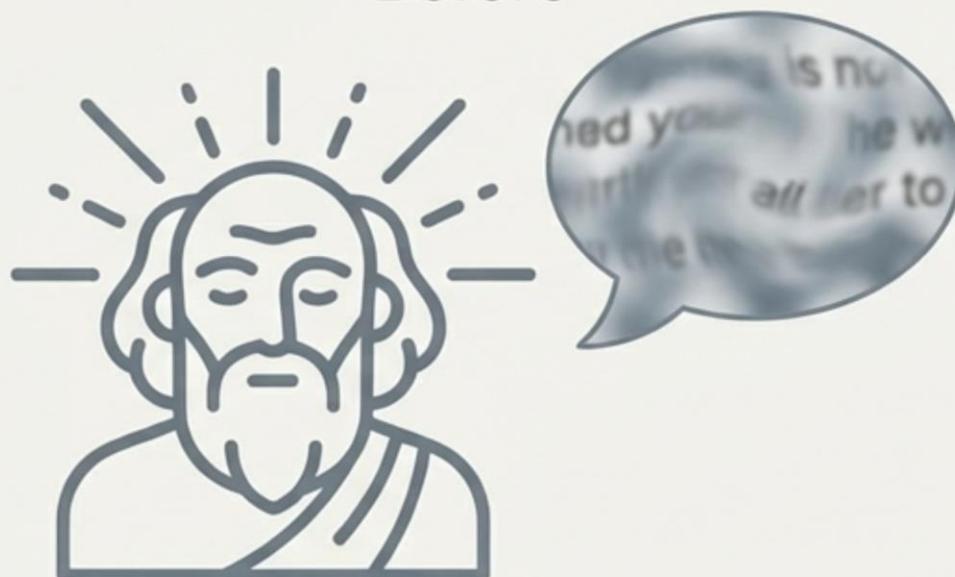


# The Genius is Handed a “Cheat Sheet.”



# The Payoff: Accurate, Trustworthy, and Timely Answers.

Before



Based on old, general knowledge.

After



Based on your specific, up-to-the-minute documents.

- Answers are based on **verifiable facts**.
- **Information is always current**.

# The Payoff: Secure Access and Smart Economics.



**Securely access private data**  
without it ever being part of the  
model's training.



**Highly cost-effective** compared to  
constantly retraining a giant LLM every  
time a new document is created.

# RAG Gives AI a Perfect, Searchable Memory.

This is the leap forward. We can create AI systems that have instant recall of any information we choose, updated in real time.



