


RAGs to Data Riches

Talk to your Enterprise Data




Hitachi Solutions
2024

Who Are Our Presenters?



Dr. Michael Green

- **Senior Product Manager** at Hitachi Solutions America (HSAL)
- Ph.D. Artificial Intelligence, New York University, Tandon School of Engineering
- More than 25 published AI research papers



Dr. Micah Groh

- **Senior Data Scientist** at Hitachi Solutions America (HSAL)
- Ph.D. Particle Physics, Indiana University Bloomington
- "Spark-itech," specializing in cluster computing and data engineering.



Topher Partimann

- **Senior Machine Learning Engineer** at Hitachi Solutions America (HSAL)
- M.S. Biomedical Engineering, University of Memphis and UTHSC Joint Program
- Gen AI and MLOps Engineering



Introducing **Hitachi, Ltd.** and **Hitachi Solutions**

HITACHI
Inspire the Next

 **Hitachi Solutions**

Since 1910, **Hitachi, Ltd.** has been a leader in manufacturing innovative products and solutions that support industry and social infrastructure around the globe.



368,300+ Employees Globally



853 Companies (696 outside Japan)



100+ Countries

We are **Hitachi Solutions**.

Hitachi Solutions is a **global systems integrator** with leading capabilities in *Microsoft applications and technologies*. Through **advisory services, industry and technology expertise, and implementation excellence** our *skilled* professionals support and accelerate **our customers'** data and business system modernization initiatives, **end-to-end**.

With our **global reach**, we are well-positioned to support customers at **scale** and *forge strategic relationships for life*.

Agenda

1 What is RAG? Why is it valuable?

2 Centralization & Vectorization

3 Security & Governance

4 Looking Forward

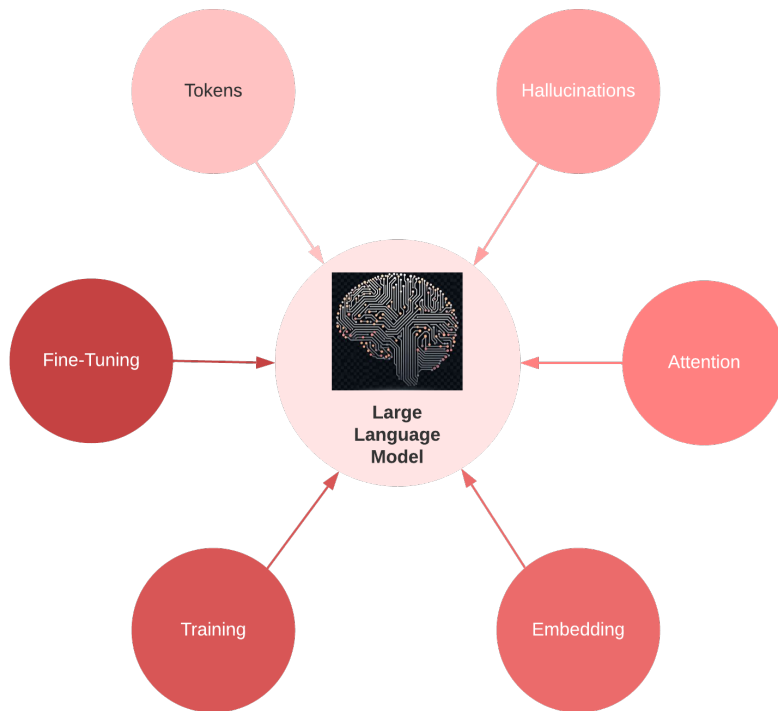
What is RAG?

Why is it valuable?

Large Language Models 101

Really large AI networks that generate language – Generative AI

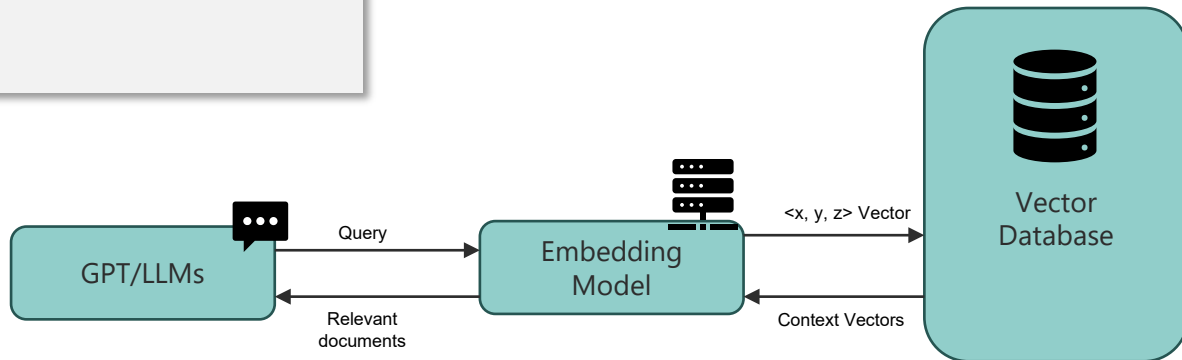
- LLMs = **L**arge **L**anguage **M**odels.
 - LLMs like GPT and DBRX generate words.
 - Embedding LLMs like ada or bge generate vectors
- Models are trained to “predict the next token”.
 - *Sometimes this results in hallucinations!*
- Embeddings are vector representations of words
 - Use of separate models are often used to aid in search
- Token = the basic unit of data processed by LLMs.
 - Models deal in tokens to limit unique words and facilitate calculations



Retrieval Augmented Generation = RAG

Using a Vector Database to enhance Generative AI

- LLM Agents can query a pre-processed Vector DB for **context**.
- Helps “ground” the model.
- Reduces hallucinations.
- Provides citations.



For this to work efficiently, enterprise data must be cataloged and centralized.

Build a strong data lakehouse foundation!
After that, **the rest becomes easier.**

Value of RAG.

Case Study | Q&A Research Assistant

Technical Investment Reports – expert business reviews made accessible through Q&A chat experience

Business Problem

Private Equity firm: Expert investment fundamentals analyses are hundreds of pages per business. Finding and summarizing the most important information to make decisions is a manual burden full of opportunities for mistakes.

Solution

- Implement Lakehouse Architecture
- End-to-end, fully automated generative AI pipeline delivers first draft for SME review.
- **Primary corpus is historical expert reviews.**
- Cost effective model deployment.
- Fully managed AI development environment.

Results

Value



- Faster response time – any investment manager can query the expert reviews with a natural language conversation interface.
- Decision making – data driven, reduce months to hours of effort
- Systematically control sensitive material

Deliverables



- MLOps best practices
- Q&A research assistant user interface
- AI environment that easily flexes to solve other business problems

Case Study | Q&A Research Assistant

Chat with Policies and Regulations

Regulatory Q&A

Question & Answer

Document Search

Question

What is the Type A limit for Co-60? **I**

Response

The Type A limit for Cobalt-60 (Co-60) is 1.2 MBq (Mega-Becquerels).



Case Study | Proposal Generation

Technical Marketing – quickly respond to a Request for Proposal

Business Problem

Design-Build firm: 30 years of proposals to sift through when responding to new job prospects with tight response dead-lines.

Solution

- Implement Lakehouse Architecture
- End-to-end, fully automated generative AI pipeline delivers first draft for SME review.
- **The primary inputs were historical proposals that match the type of work in the new request.**
- Cost effective model deployment.
- Fully managed AI development environment.

Results

Value



- Faster response time
- Insight into new business approaches the during RFP process
- Predictable production schedule
- Systematically control sensitive material

Deliverables



- MLOps best practices
- Automated first draft proposal generation
- AI environment that easily flexes to solve other business problems

Results Summary

Proposal Sections

Section	Final Document Output (Pages)	Average AI Cost (USD)	Average Processing Time (min)
Resume	2 pages	\$1-\$3	5 min
Company Introduction (short)	1 page	\$0.30-\$0.80	2 min
Company Introduction (long)	2 pages	\$0.50-\$1.50	2-5 min
Proposer Experience (Intro)	1-2 pages	\$1-\$2	5 min
Proposer Experience (Project form)	10 pages	\$4-\$6	15-20 min
Project Understanding and Approach	18 pages	\$5-\$8	20+ min

Based on roughly **10-15 past proposal documents**, **~2-5 resume documents per person**, and **10-20+ additional supplementary documentation**.

Costs depend generally on how many different types of searches the AI agent needed to make to compile the information. This is why an Experience section with less pages but with multiple parts might cost just as much to generate as a Project Understanding section.

Case Study | Document Processing

Original Document

Specification Output

005.3739.000 Issued for Construction

SECTION 07 42 47 – FIBER-CEMENT FACADE PANELS

PART 1 - GENERAL

1.1 SUMMARY

A. Section includes fiber-cement composite facade panels which are hereinafter referred to as "fiber-cement facade panel assemblies".

B. Related Requirements:

- Section 06 10 00, "Rough Carpentry," for wall framing.
- Section 06 16 05, "Gypsum Sheathing," for wall sheathing.
- Section 07 21 00, "Thermal Insulation", for wall insulation.
- Section 07 27 26 "Fluid-Applied Membrane Air Barriers" for air and moisture barrier.

1.2 ACTION SUBMITTALS

A. **Product Data:** Submit product data for each type of product indicated. Include construction details, material descriptions, dimensions of individual components and profiles, and finishes for each type of fiber-cement facade panel and accessory. Include fiber-cement facade panel installation instructions.

B. **Deferred Submittal Documentation (Delegated Design):** Submit copies of structural calculations for the fiber-cement facade panel work, indicating complete compliance with the specified performance requirements. Include calculations to show that maximum deflections do not exceed specified performance requirements under full design loading, and calculations for panels, and anchor connections. Calculations and shop drawings shall be prepared, signed and sealed by a Professional Engineer registered in the state of California and who is experienced in providing engineering services of the kind indicated.

- Architect and consultants' review of the Deferred Submittal Shop Drawings and calculations is limited to confirmation that the work described in them is in general conformance to the design of the building, in accordance with CBC Section 107.3.4.1.

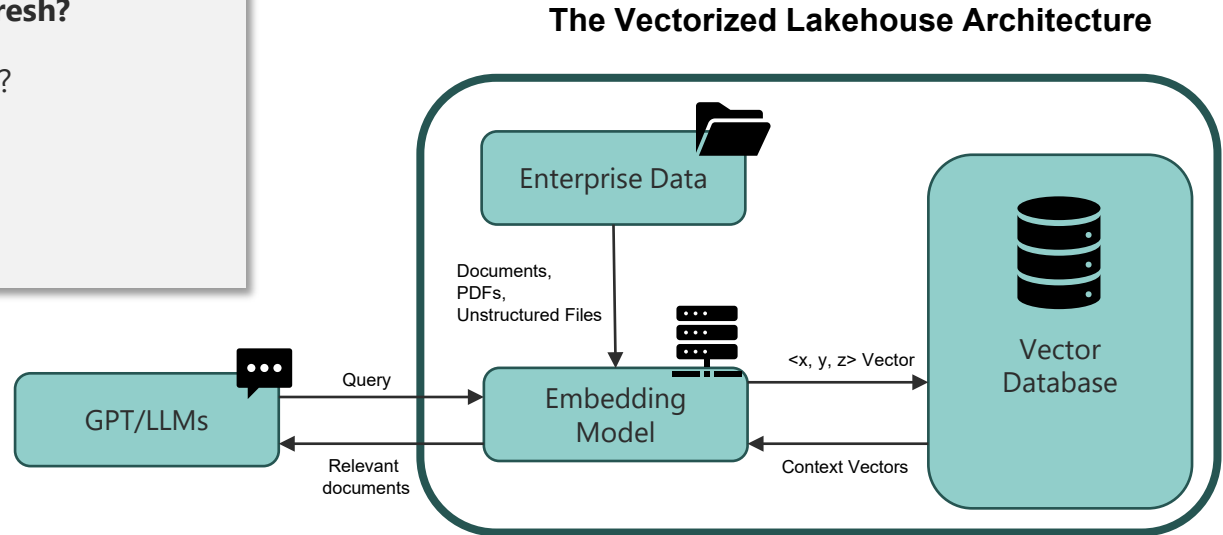
C. **Shop Drawings:** Submit shop drawings to the Architect. The shop drawings shall show scaled elevations, plans, and sections of the fiber-cement facade panel work. Full scale sections shall be prepared and submitted for details of the assemblies that cannot be shown in the elevations or sections. Include with shop drawings thickness of all panel components, panel finish, metal finishes, and all other pertinent information as necessary or requested by the Architect to

Spec #	Spec Name	Page	Statement	Submit #	Type	Relevant
07 42 47	CEMENT FACADE PANELS	pg 337	Product Data: Submit product data for each type of product indicated. Include construction details, material descriptions, dimensions of individual components and profiles, and finishes for each type of fiber-cement facade panel and accessory. ...	1.2-A	Product Data	Yes
07 42 47	CEMENT FACADE PANELS	pg 337	Deferred Submittal Documentation (Delegated Design): Submit copies of structural calculations for the fiber-cement facade panel work, indicating complete compliance with the specified performance requirements. Include calculations to show that maximum deflections ...	1.2-B	Calculations	Yes
07 42 47	CEMENT FACADE PANELS	pg 337	Shop Drawings: Submit shop drawings to the Architect. The shop drawings shall show scaled elevations, plans, and sections of the fiber-cement facade panel work. Full scale sections shall be prepared and submitted for details of the assemblies ...	1.2-C	Shop Drawing	Yes



Enterprise RAG Components

- How do you keep the data **hot and fresh**?
- How do you keep the data processed?
- How do you keep the data **secure**?
- How to do you manage data access?



For this to work efficiently, enterprise data must be cataloged and centralized.

Build a strong data lakehouse foundation!
After that, **the rest becomes easier.**

Centralization & Vectorization

What's Stopping you from just **“doing RAG right now”**?



You're here because you, like many industry leaders, are struggling with data.
It's messy, complicated, and overwhelming to get started on a path to modernize.



Governance and control, who has access?



Fragmented data across disparate data sources



Keeping data “fresh” is complicated!

Empower Data Platform | Key Features



Empower **enables you to unify data** to provide faster **business value**.

1

Acquires Data

Move data from over 200 sources into your own Lakehouse.

2

Transforms Data

Unify this data within your Lakehouse into one model.

3

Publishes Data

Analyze your data in Power BI or other apps with writeback capability.

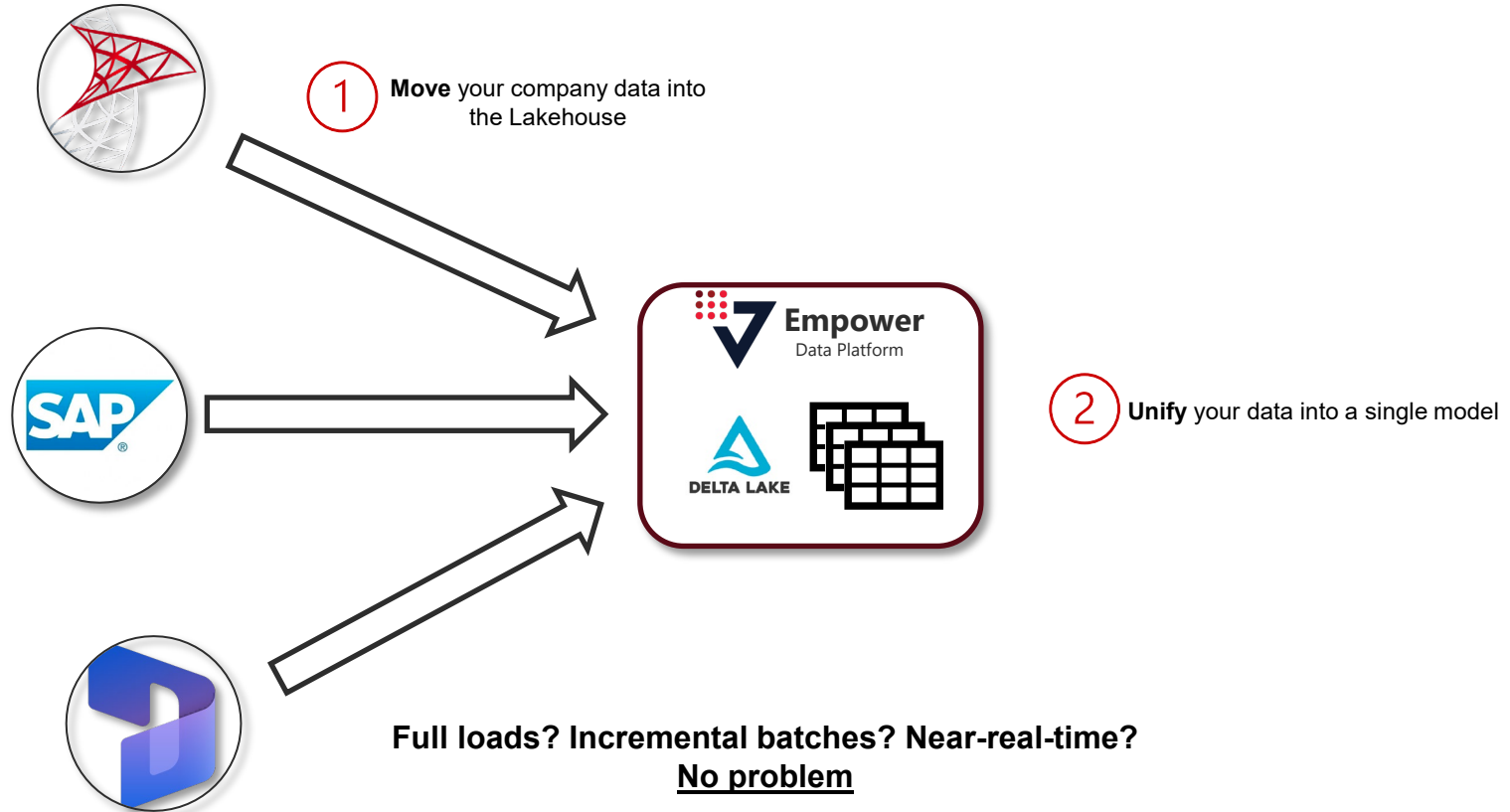
4

Unlocks Value

Users unlock their data estate, enabling every citizen developer to achieve more.

Built on the Lakehouse

Empower Data Platform | How does it work?



Demo

Automating Away Data Acquisition w/ Empower

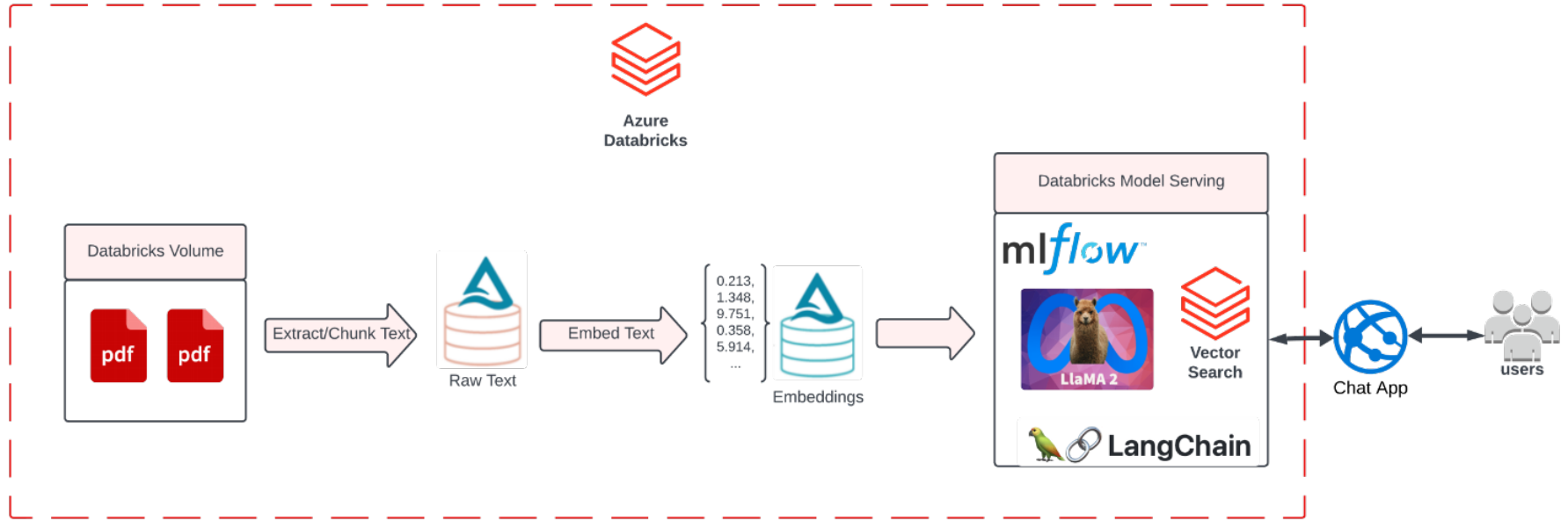
A brief disclaimer...

Focus today is on unstructured data.

Later on, we propose an architecture w/ unstructured & structured data.

Now your data is centralized and fresh...

..time to chunk and vectorize!



Demo

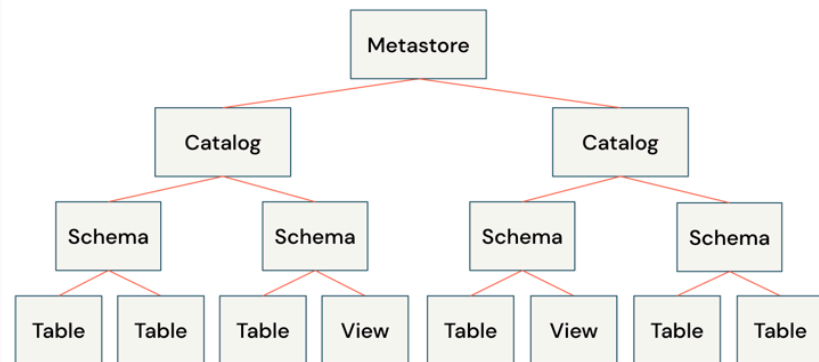
Chunking and Databricks Vector Search

Security & Governance

Unity Catalog for your (Vectorized) Lakehouse

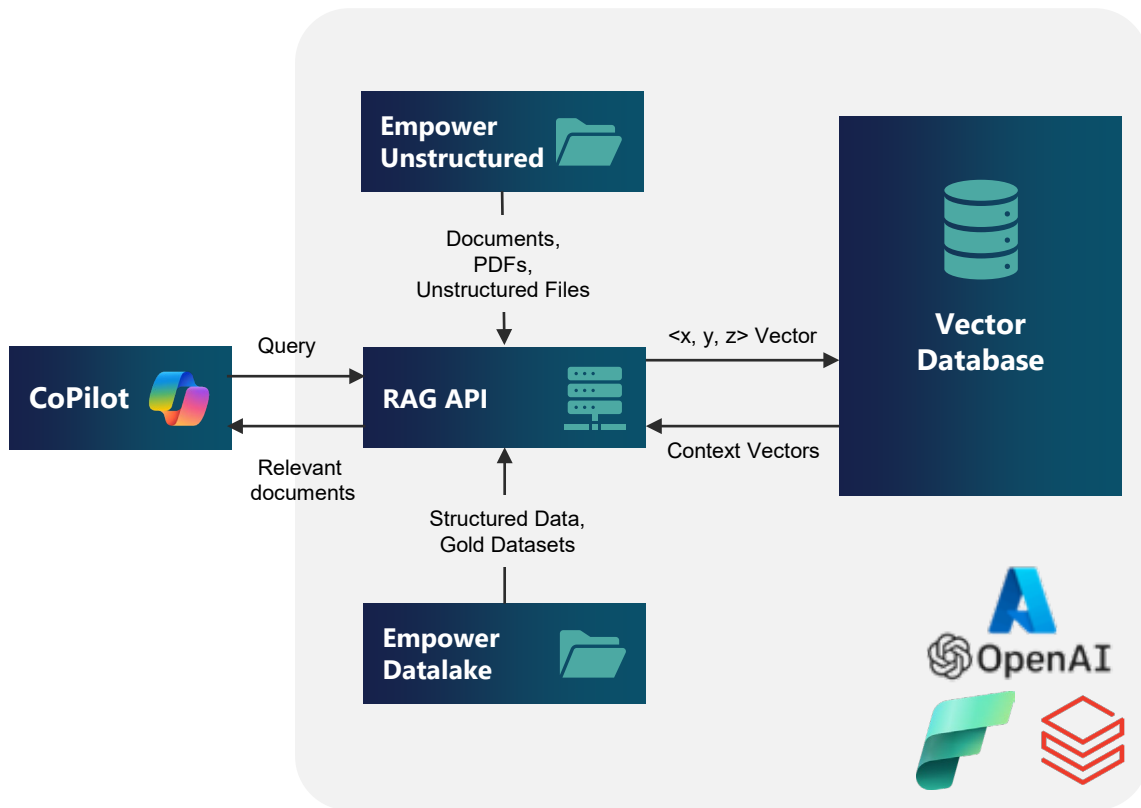
Unified Security and Governance for your lakehouse.

- **Security and Access** – Define access management as low as a row-by-row and column-by-column granularity. Create user groups to simplify data access, **all using SQL statements**.
- **Governance** – Track changes to your data across your lakehouse on an object-by-object basis. Know exactly what changes from table to table, and who/what made it.
- **Internal Data Sharing** – Never copy data within your organization for other Lakehouse's, Empower environments, or products. Now you simply grant permission and go. No cloning, no pipelines, no hassle.
- **Delta Sharing with Partners** – Share, or even buy and sell data between your tenant and others: updates are *instantly available* from the source with no latency, meaning that you can share (or use shared data) the moment it lands in the source table.

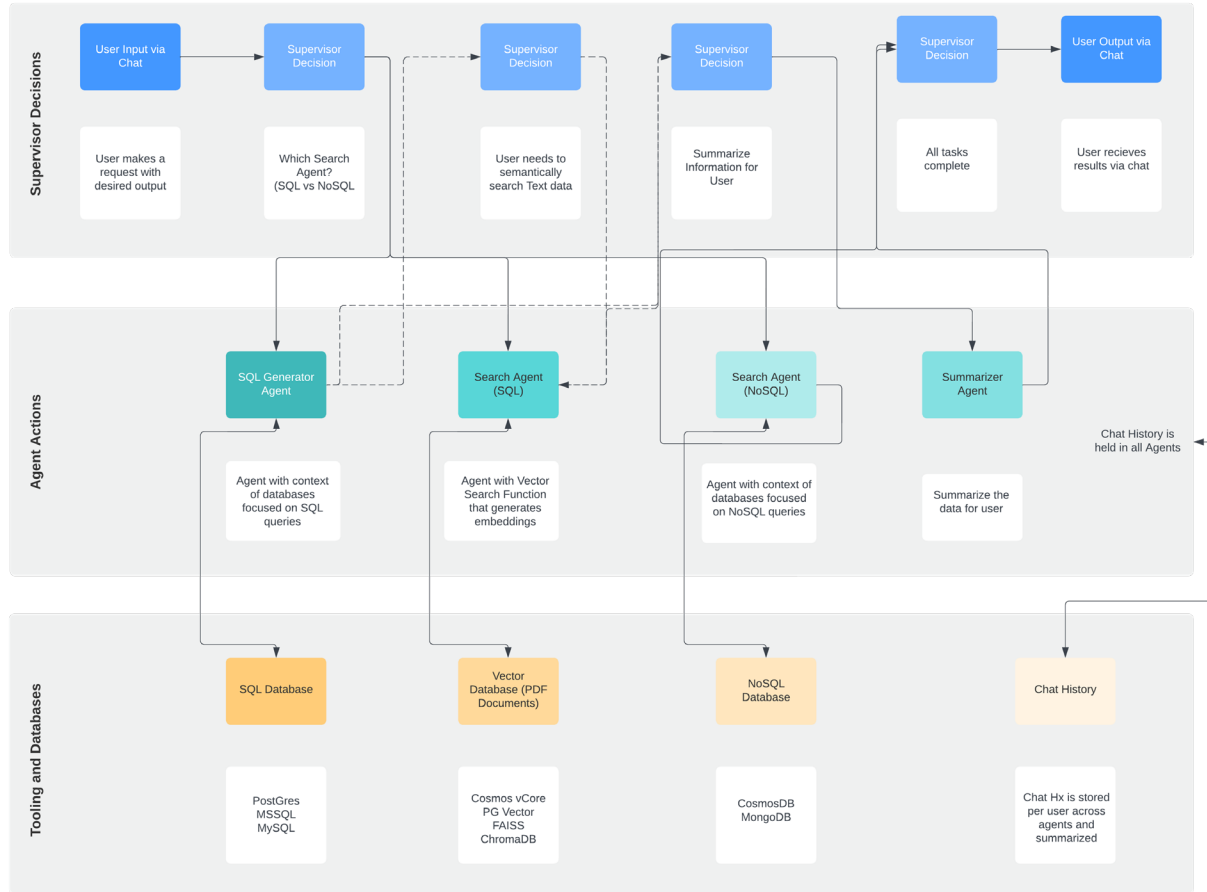


Looking Forward

Looking Forward | MSFT Fabric, OneLake, Unity Catalog & Copilot



Looking Forward | Complex Agents & Structured Data



Next Steps...

Where can we engage with you?

Learn more about
Empower



- **Modernize your Data Estate** with our Industry expertise.
- **Unity Catalog** and **Serverless** in everything that we do.
- **Unique Delta Connectors** for Dynamics 365, SAP, and MSFT Stack.

Contact Info



Looking for Answers?

Reach out to us!



Dr. Michael Green

Senior Product Manager

michaelgreen@hitachisolutions.com

+1 615-558-8035



Don Scott

VP of Product Development

dscott@hitachisolutions.com

267-567-3266

