MOTIUS
WE R&D.

Kardex GPT

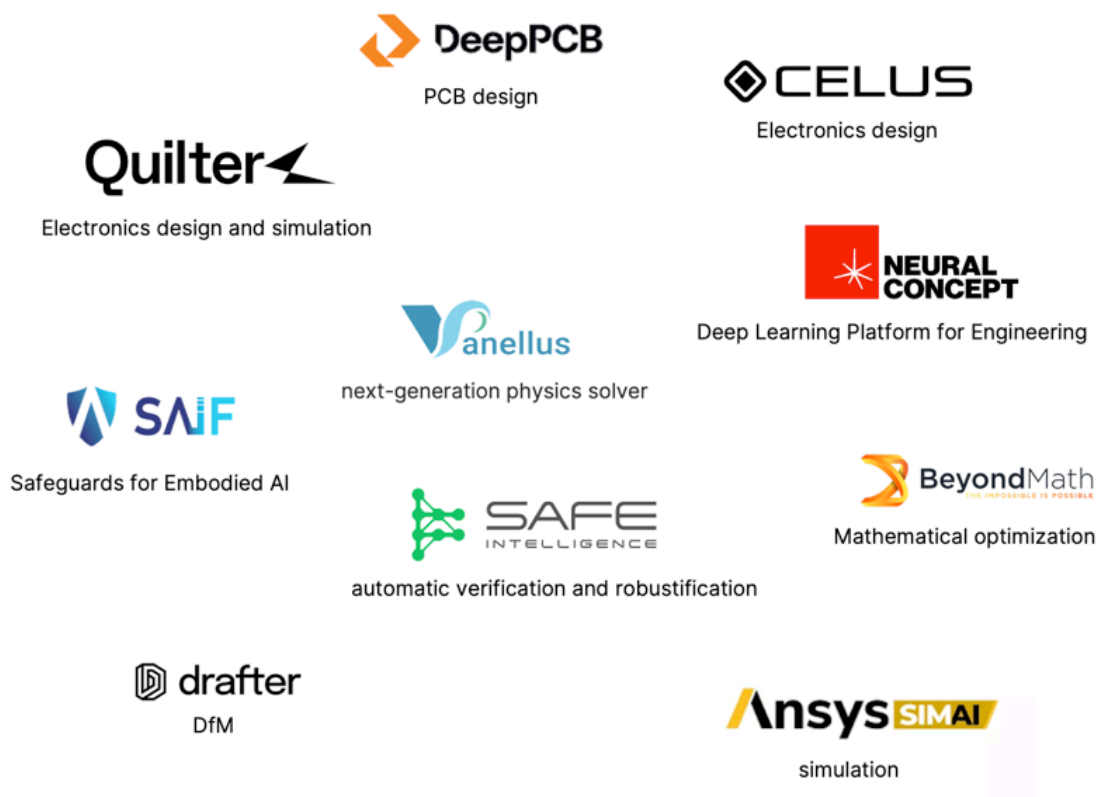# Kardex GPT

For companies like BMW, Allianz, and Sennebogen, Motius has developed an approach to building AI assistants and agents that can work both on-premise and in the cloud, with data ranging from classified internal documents to public data ingested through web search.

A consolidated data layer ensures governance and reusability, while a flexible platform enables the development of **use case-specific AI agents**, including third-party solutions. Users can interact via chat app, custom web interface, or IT system integrations.

DeepPCB — PCB design

CELUS — Electronics design

Quilter — Electronics design and simulation

Vanellus — next-generation physics solver

NEURAL CONCEPT — Deep Learning Platform for Engineering

SAIF — Safeguards for Embodied AI

SAFE INTELLIGENCE — automatic verification and robustification

BeyondMath — Mathematical optimization

drafter — DfM

Ansys SIMAI — simulation

The platform approach provides **enterprise grade scalability and security** while still providing flexibility, in concrete technical implementation. The solution will empower CustomerName with AI-driven insights and automation to enhance workflows and decision-making.

> **ℹ Typical Challenges for other Customers**
>
> - There is a lot of unstructured knowledge in the company, such as documents, Wikis, emails, and chat messages
>
> - Existing systems have limited search capabilities. Usually there is no global search
>
> - Incomplete datasets spread over multiple systems, such as ERP, MES, CRM, and file storage
>
> - Knowledge distributed across multiple locations, in different languages

Demo

Gandalv is an internal product we built to showcase our approach, and help us during the requirements engineering phase of our projects.

0:00
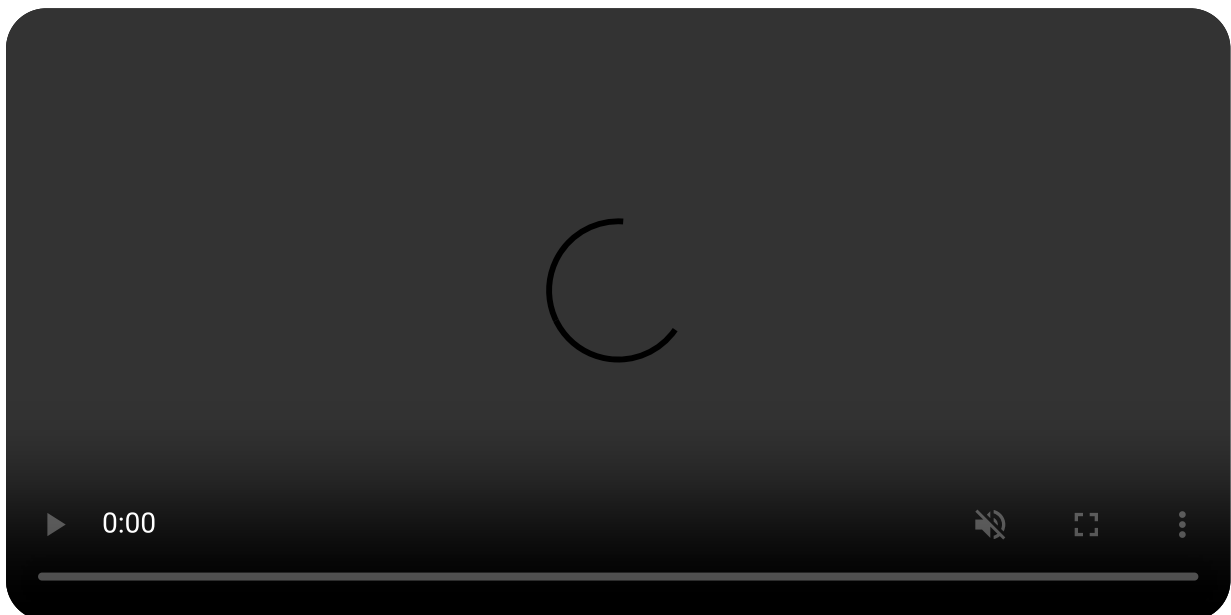
Key Features

While the demo shows a specific use case, the platform is designed to be flexible and extensible:

🛡️ On-Premise or Cloud

Our approach works on-premise or using a cloud-based AI model in Azure or AWS. We use open-source components to stay flexible and avoid vendor lock-in, while using the state of the art in AI and LLMs.

🤖 AI-Powered Assistant

The AI assistant is capable of understanding and processing natural language queries, providing intelligent responses and insights based on the data it has access to. It seamlessly works with different input languages in CustomerName's data.

🗄️ Data Integration

Using MCP and RAG we ensure that the AI model has access to the most relevant data, without spending months preparing and cleaning it.

⏱️ Speed & Quick Iterations

With powerful open-source components and state of te art models, Motius only needs a few weeks to set up the - infrastructure and first use cases for CustomerName. From there, we can quickly iterate on data sources and new use cases.

Approach

❶ Knowledge Structuring and Extraction

A team from Motius first analyzes existing content and develops a strategy for providing a structured knowledge base. The strategy may include the following technologies and approaches:

Technology	Data Source	Description

RAG	Wikis, internal documentation, operating manuals	AI-powered search and extraction of relevant information from various sources. Very efficient for searching, but requires dedicated infrastructure
MCP	Confluence, CAD systems, ERP, MES	Standardized interface for querying information. Ad hoc, requires very little infrastructure, but is slower than RAG

❷ Rapid Prototyping and Integration

Once the data sources have been reviewed and connected, a prototype is developed usually using on-premise LLMs and open source components.

Employees at CustomerName can start working with the system and provide feedback after just 1-2 sprints (each 2 weeks).

> ✏️ **Note**
>
> If possible, we deploy this system to a cloud environment, such as Azure or AWS, to ensure scalability and performance.
>
> Frequently, internal security policies require that the system runs on-premise, in which case we use open source LLMs and components.

❸ Data Enrichment and Automation

After the first tests, we add more data sources and build agents that can automate tasks and provide insights for different business processes, such as:

- ☑ Accessing knowledge across many systems and regions, asking follow-up questions, and getting relevant documents as sources in the response
- ☑ Automating repetitive tasks, such as generating reports, summarizing documents, or extracting key information
- ☑ Tool calling with MCP, which allows the AI assistant to act on behalf of the user, but using their credentials and permissions

Architecture
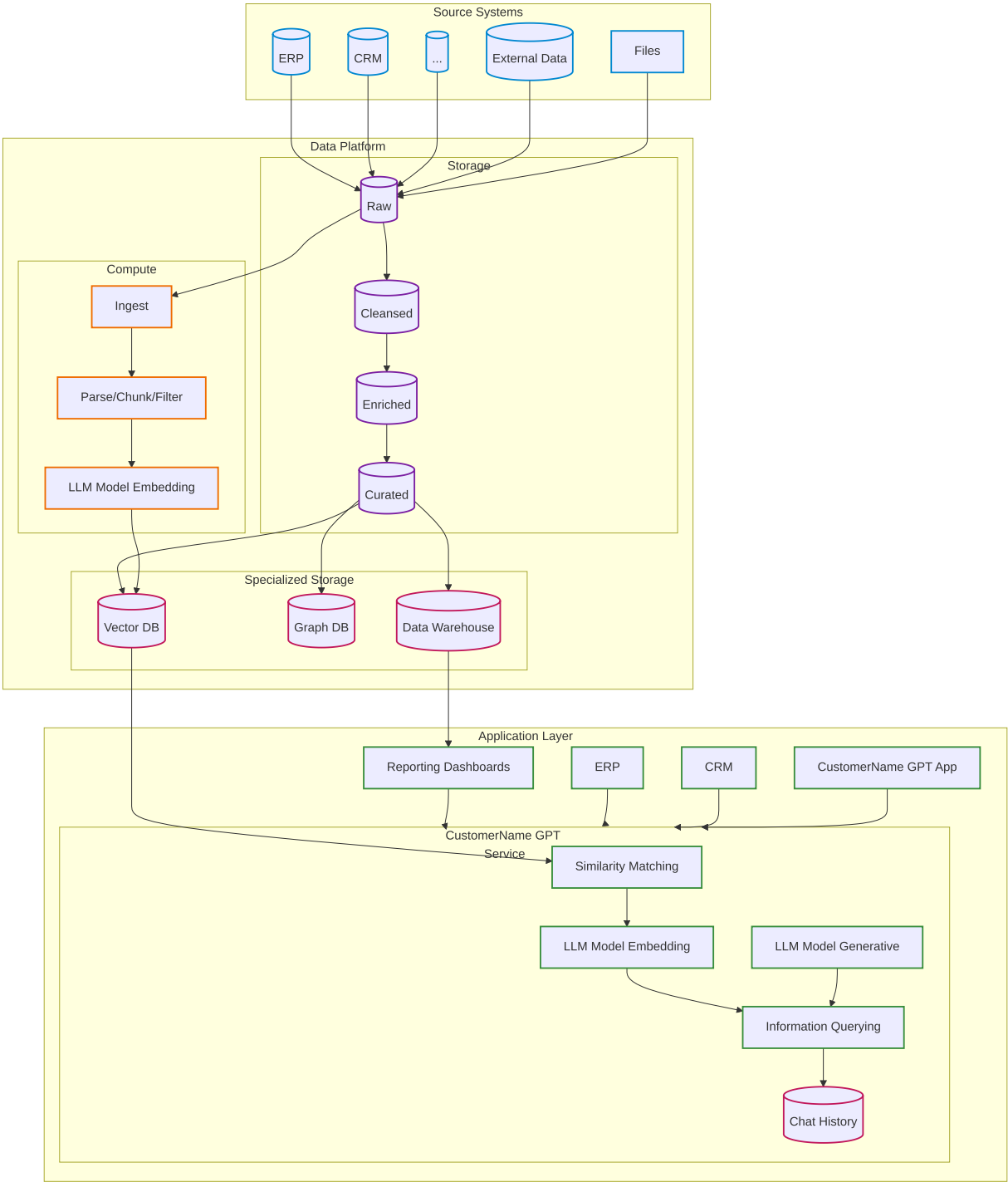
On-Premise

We deploy a mix of open source components and open-weight LLM models:

Application at CustomerName

With this approach, CustomerName can start building AI assistants for specific use cases within weeks, while building on an architecture that allows for future growth:

- ☑ Cloud and AI-agnostic architecture means CustomerName can choose the best providers with no lock-in

- ☑ Immediate benefits of AI assistants for specific use cases, such as customer support, internal knowledge management, and process automation

- ☑ Go further with AI agents that can automate tasks, and even let your engineering teams build their own agents