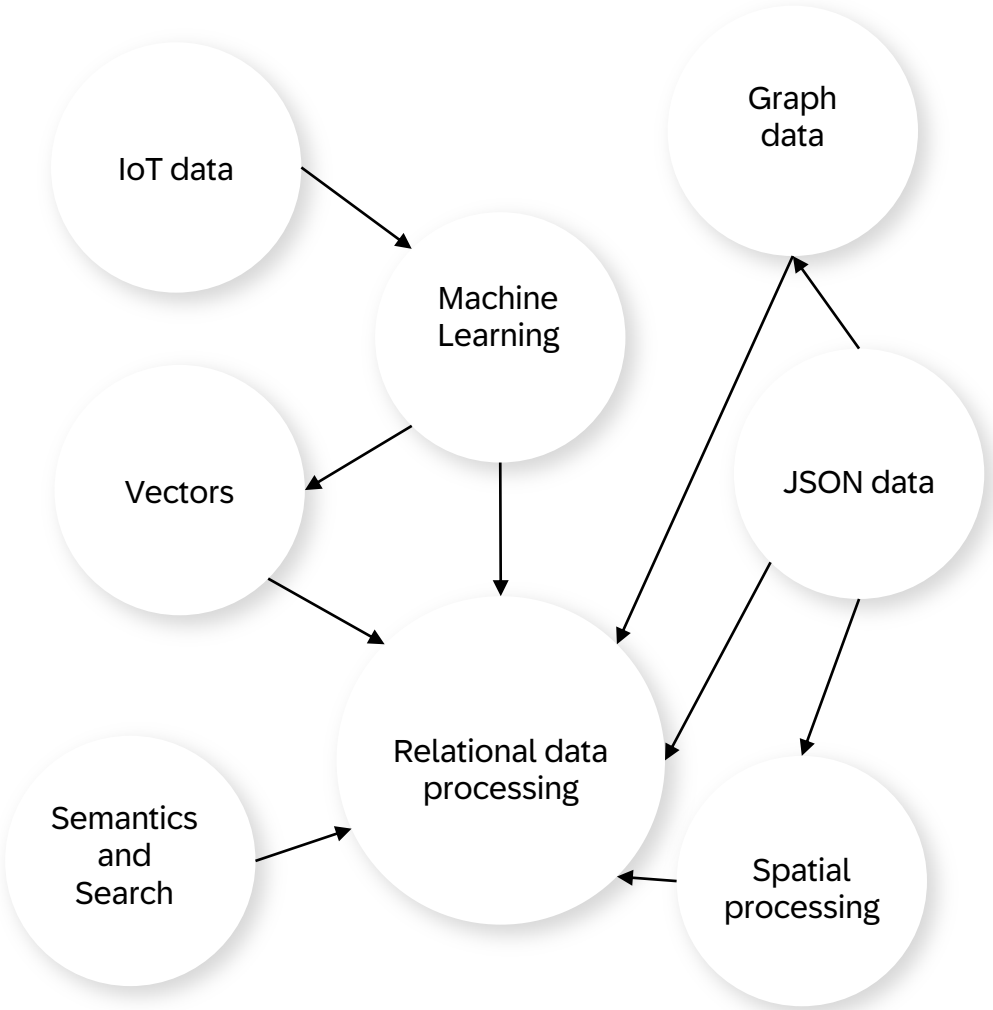
SAP HANA Cloud
Multi-model and Machine Learning

Public

The Modern Data Challenge

Innovate despite data silos



It's complex to manage and leverage many specialized databases.

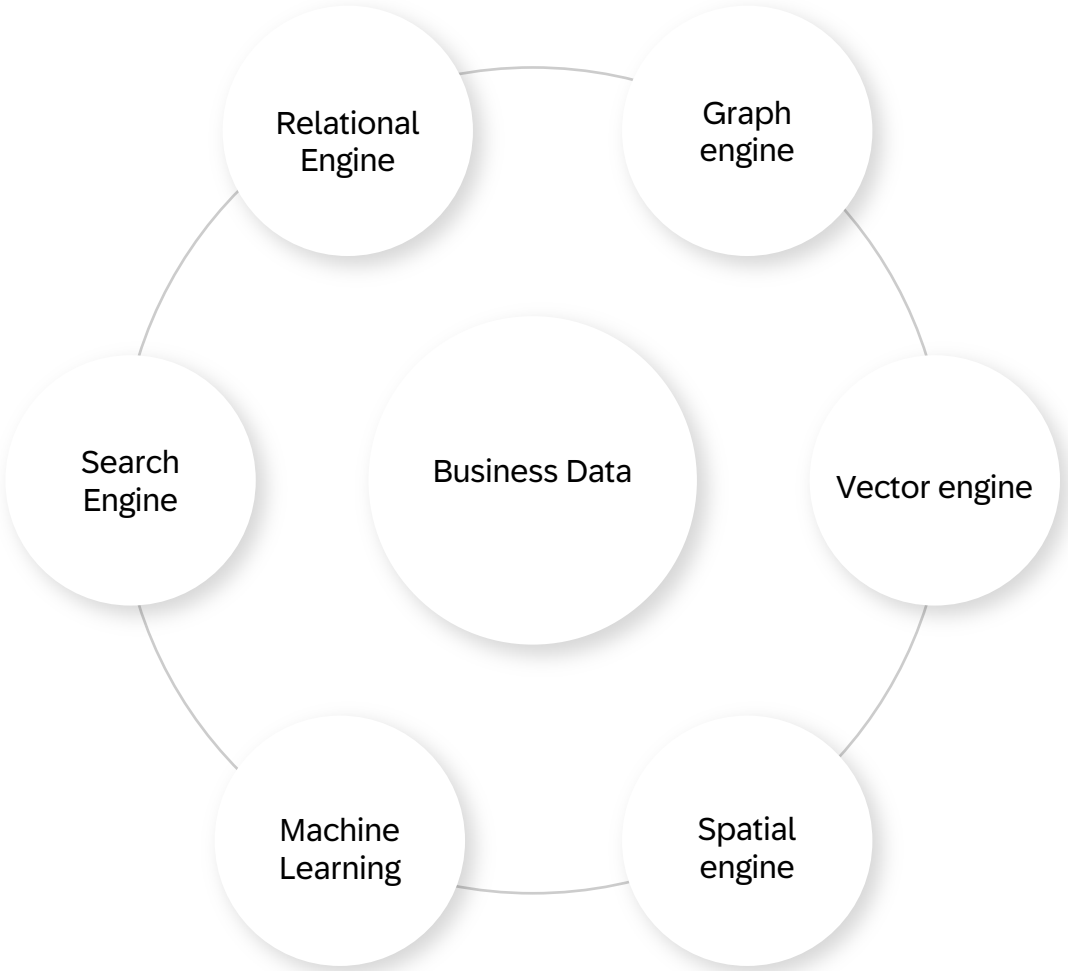
It's frustrating to make untimely decisions.

It's risky to secure data across many separate systems.

It's discrediting to deliver information with consistency issues.

The Modern Data Solution

Consolidate to reduce complexity and enable innovation



The simplicity of a single copy of the data

The speed and scale of a multi-tiered data architecture

The security of a holistic enterprise-grade security framework

The surety of all types of data

SAP HANA Cloud

Power the next generation of Intelligent Data Applications

Virtually blend data from remote sources with locally tiered data

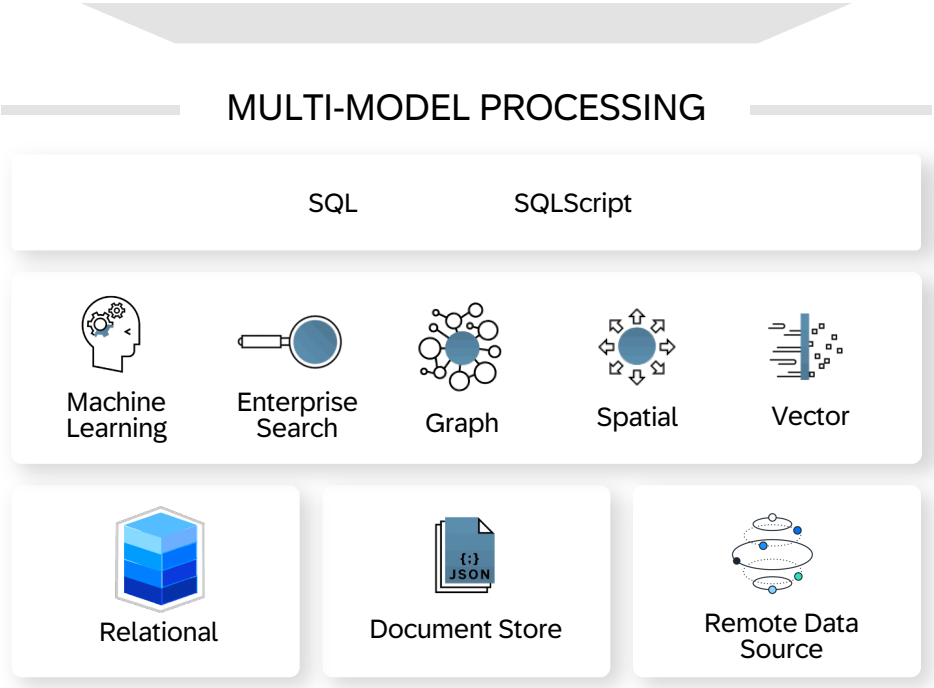
Process all types of business data regardless of data model, type, or volume

Interact in real-time using an industry standard language

Spatial
Location-based intelligence

Graph
Storage and traversal of highly connected data and associated properties

Document Store
ACID-compliant, flexible management of JSON documents



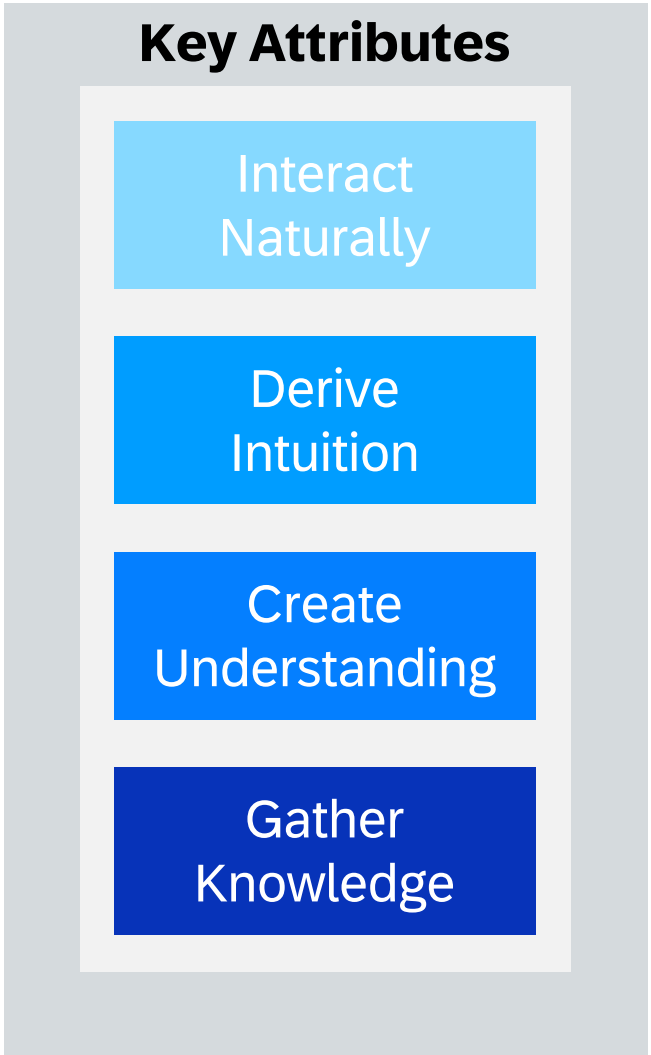
Enterprise Search
Find business entities within the database

Machine Learning
Expert library designed and optimized for fastest in-memory processing

Vector
Embedding storage and similarity search functions

SAP HANA Cloud

An **intelligent data application** is an **independent component** within a **business process** that operates with **autonomy** and/or provides the **user** with the tools **to perform** with **expertise**.



Open Grid Europe

Unifying data to enable field and asset performance

Data silos prevented a holistic view of asset information

SAP HANA Cloud enabled data from Esri and SAP to integrate in a single database

Field workers receive **reliable guidance** in locating pipelines and identifying repair needs

Stakeholders have **immediate transparency** into pipeline operations

OGE

SAP HANA Cloud Data Processing Engines

SAP HANA Cloud Multi-model

Overview

SAP HANA Cloud Spatial

- Spatial data types and reference systems
- Spatial functions, predicates, and algorithms
- Esri and open source GIS integration

SAP HANA Cloud Graph

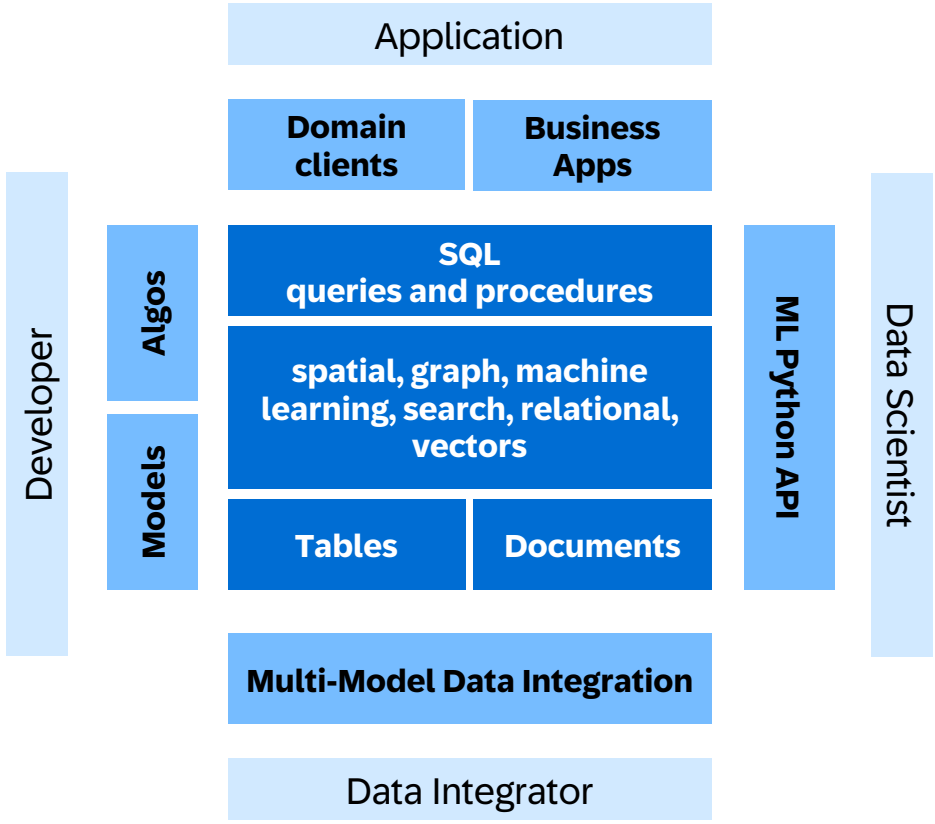
- Relational data as graph/network
- Pattern matching queries
- Built-in graph algorithms and custom network processing

SAP HANA Cloud JSON Document Store

- Native store for JSON documents
- JSON operations via SQL

SAP HANA Cloud Vector Engine (QRC1 2024)

- Vector data type
- Vector distance functions



SAP HANA Cloud Multi-model

Spatial

Spatial data types

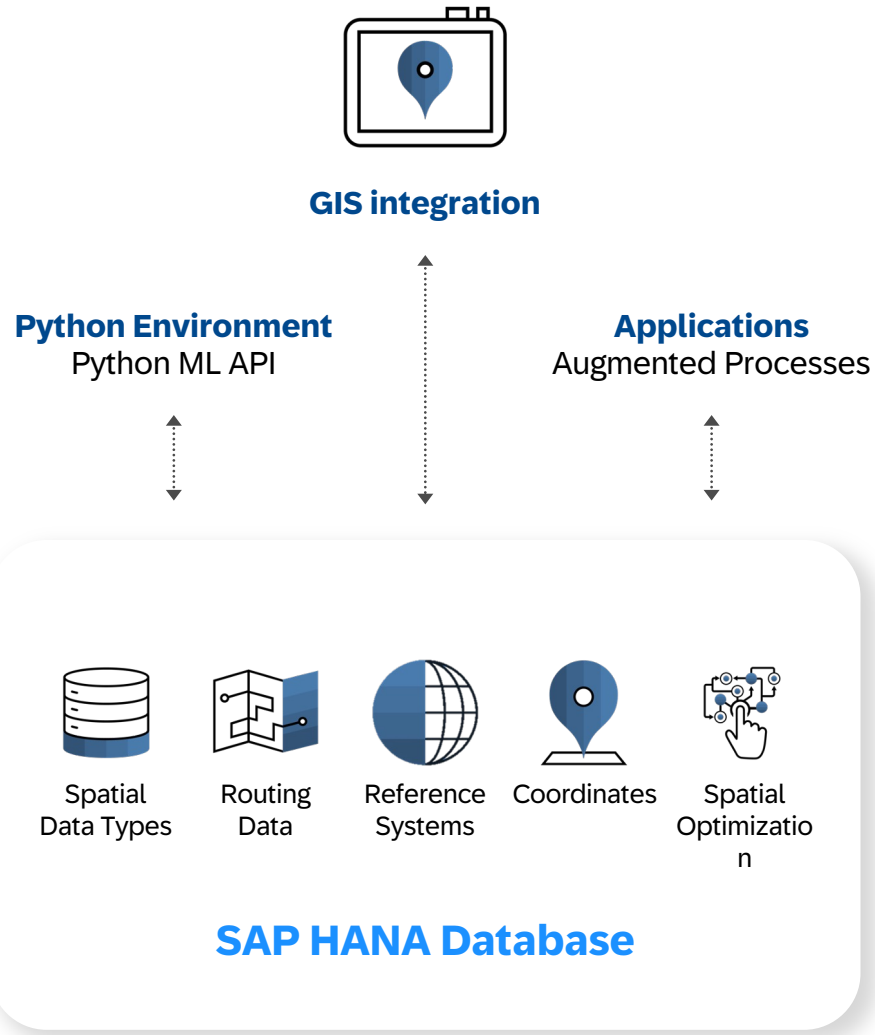
- Native storage for vector data types such as points, lines, polygons...
- 2D and 3D objects supported
- Choice to use a specific spatial reference system

Spatial function and predicates

- Boolean operations, e.g. union, intersect
- Relationship determination, e.g. contains, touches
- Property computation, e.g. length, area
- Transformation and inspection, e.g. SRS, lin. referencing

Consumption and ecosystem

- SQL, ABAP
- Esri ArcGIS, Esri geodatabase, GeoTools/GeoServer, QGIS
- Python client for machine learning (hana-ml)



SAP HANA Cloud Multi-model

Graph

Graph model

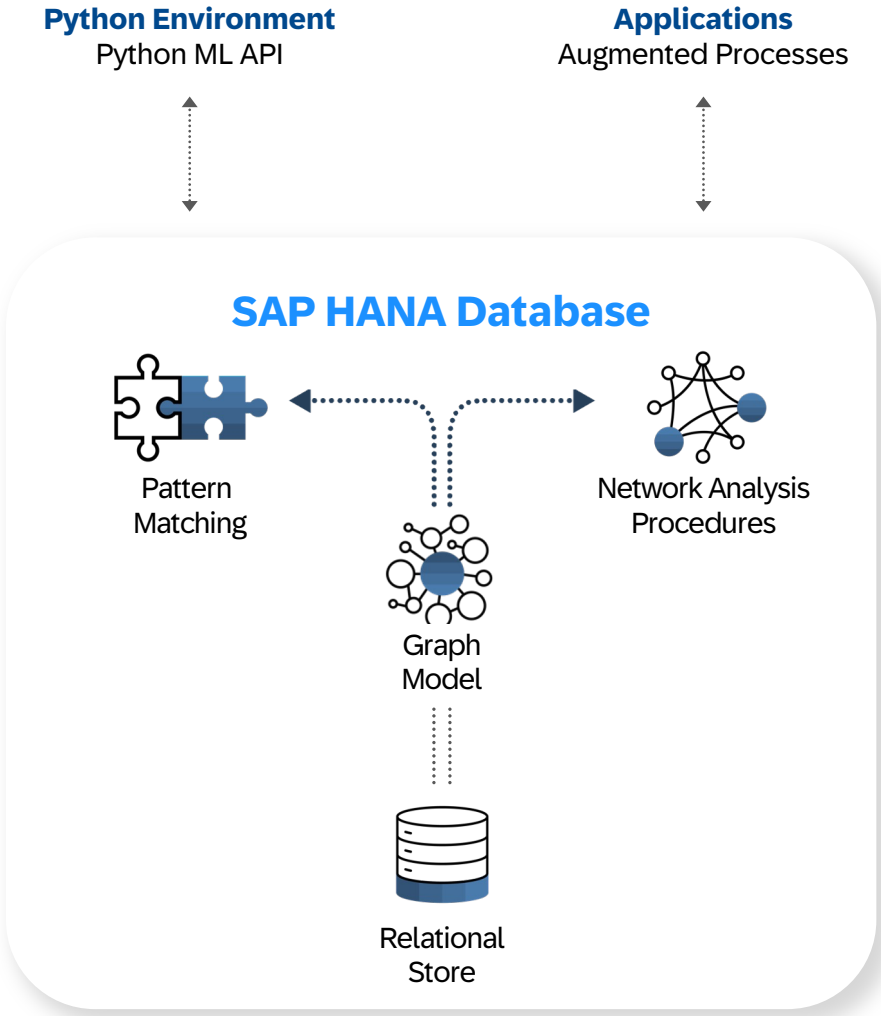
- Labeled property graph embedded into SQL/relational

Graph queries and algorithms

- Cypher for pattern matching
- Built-in algorithms, e.g. shortest path, page rank, link prediction
- In-database Graph procedures (“GraphScript”)

Consumption

- SQL, ABAP (via AMDP)
- Python client for machine learning (hana-ml)
- HANA Database Explorer, Cytoscape (preview)



SAP HANA Cloud Multi-model

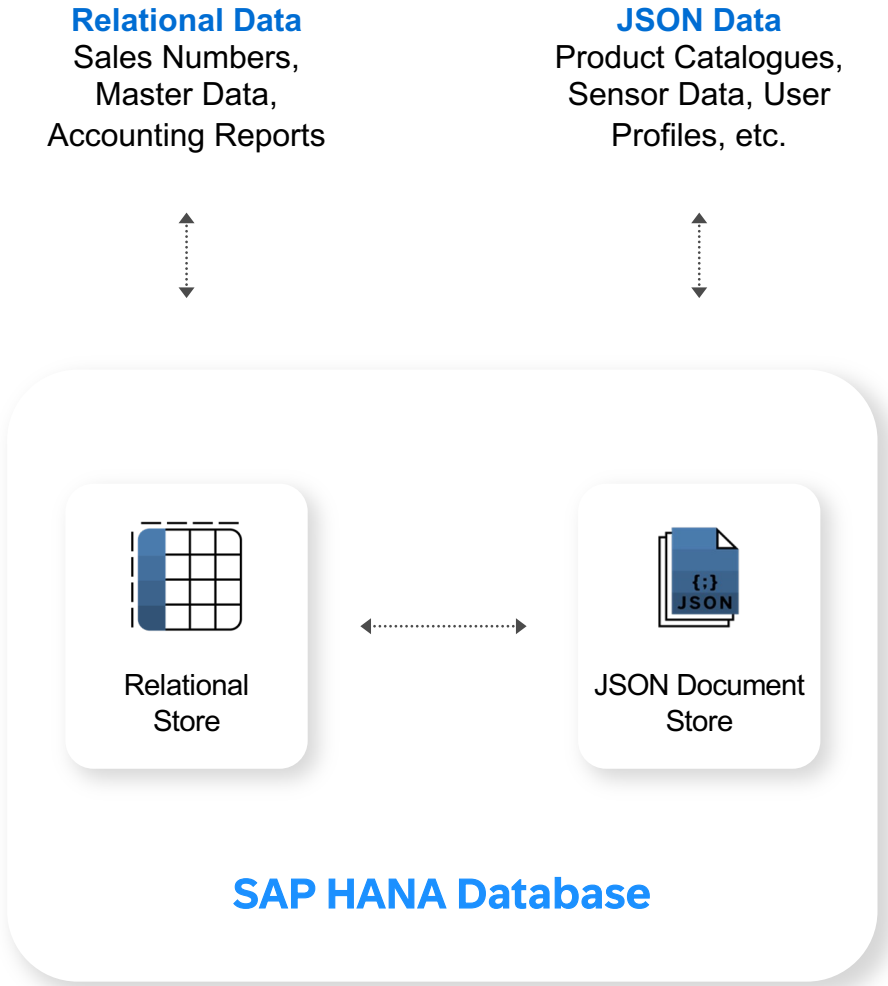
JSON Document Store

Native store for JSON documents

- Schema-less, hierarchical, semi-structured data
- Arrays, objects, key-value pairs
- Complementary, fully integrated store
- ACID across all stores (row, column, document)
- Indexing, SQL parameters, paging to disk, import/export
- Backup/restore, encryption, failover

Native JSON operations via SQL

- Projection, filtering, aggregation
- Unnesting arrays
- Joins with relational data



SAP HANA Cloud Multi-model

Full-text Search

In-database search

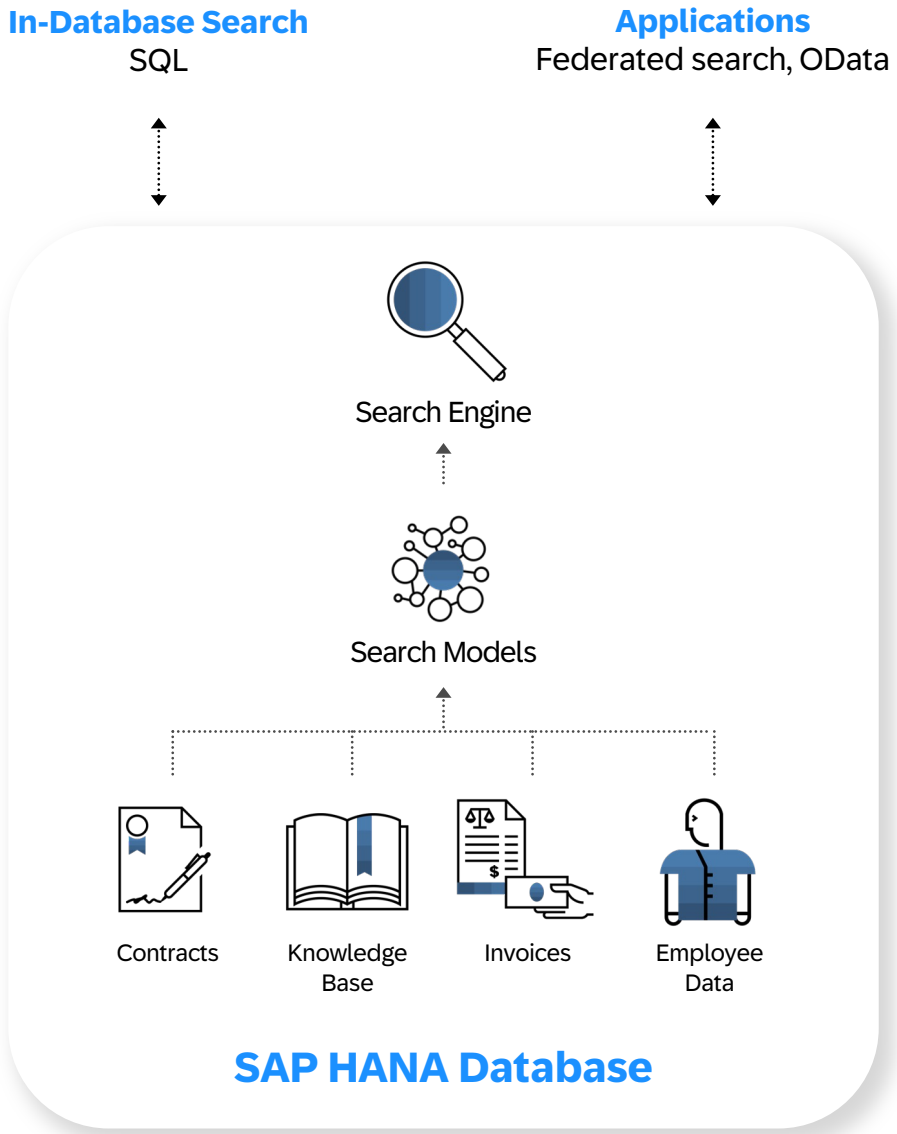
- Column store data types: string, date, numeric
- Fuzzy search index
- Speed up fuzzy searches and enable text search

Search models

- SQL views (joins, w/ parameters), table functions
- Search configuration via built-in procedure or HDI
- Entity type and property annotations
- Multi-value, sub-objects, multilingual texts

Search

- SQL contains predicate
- Federated search using built-in procedure
- OData, Enterprise Search SDK



SAP HANA Cloud Multi-model

Vector

In preview*
(*subject to change)

Vector data type

- REAL_VECTOR
- Natively store high-dimensional vectors

Vector functions

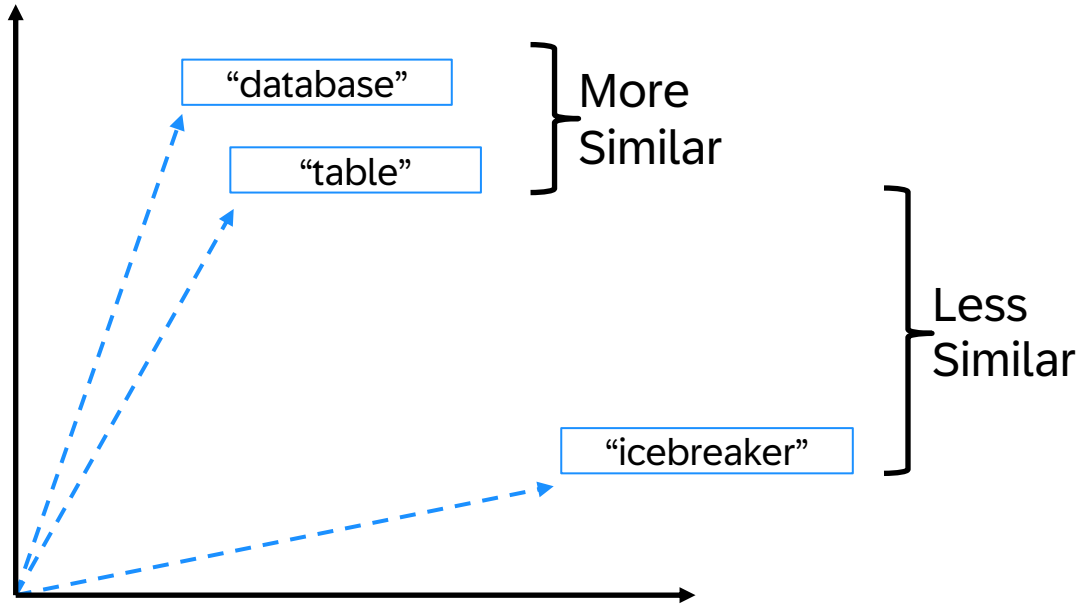
- Combine vector functions with other SQL operations
- L2DISTANCE()
- COSINE_SIMLARITY()

Consumption

- SQL
- python (hana-ml)

Roadmap & Vision

- LangChain plug-in, CAP support, SAP Generative AI Hub integration
- Vector indexes, approximate nearest neighbor (ANN) search
- In-database text embedding



*Planned Q1 2024

SAP HANA Cloud Multi-model
Machine Learning

Embedded Machine Learning Libraries

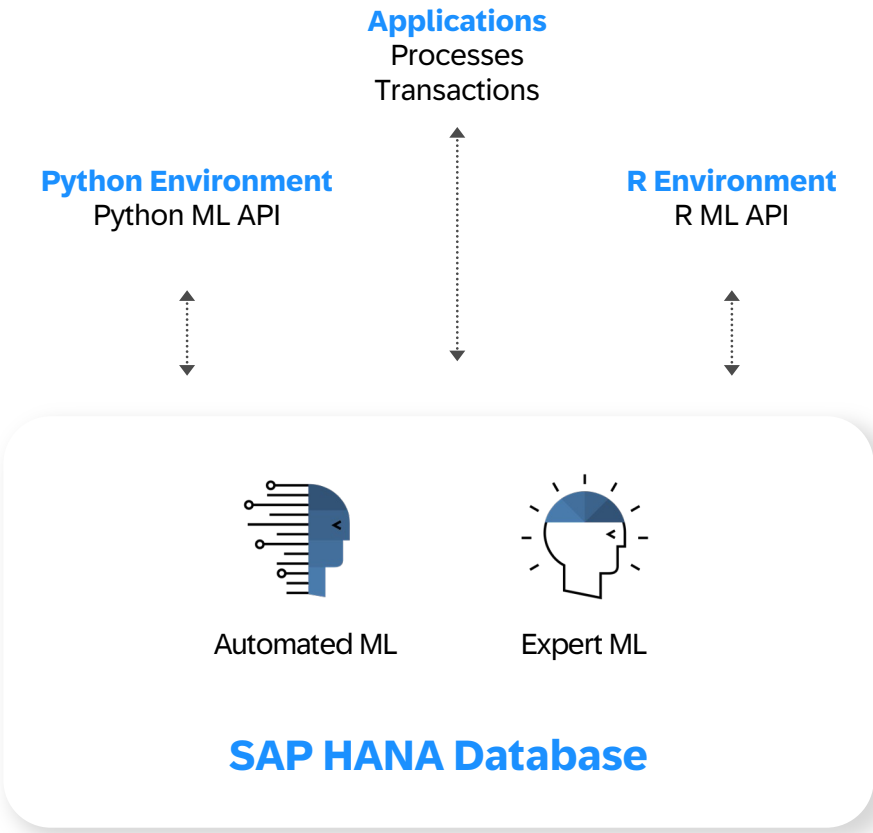
- Predictive Analysis Library (PAL) for experts
- Automated Predictive Library (APL)

Key Machine Learning Scenarios

- 100+ classic and trending ensemble algorithms
- Time series forecasting, classification, regression
- Machine learning functions optimized for massive parallel in-memory processing

Additional Features

- Native interfaces for data scientists in R and Python
- Improved data science in Python to development handshake for building an intelligent data application
- Simple ML scenario generation for SAP Business Application Studio
- New PAL AutoML framework supporting classification, regression, and time-series scenarios



Recent Innovation Highlights

Innovation Highlights

Spatial Functions in the Graph Engine

Spatial Intelligence Meets GraphScript

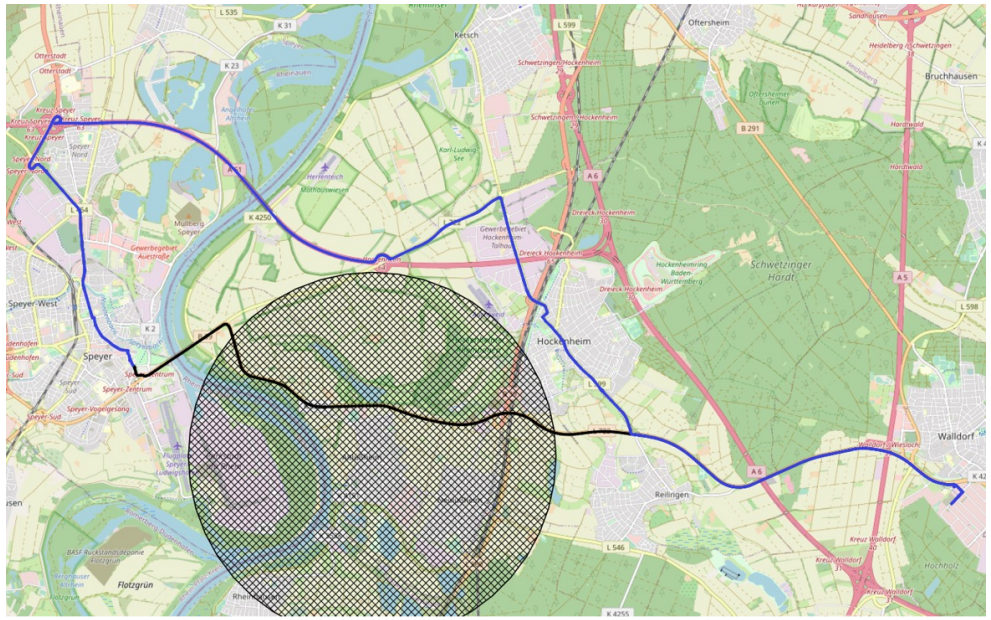
- GraphScript now supports all methods of ST_GEOMETRY which do not return ST_GEOMETRY or VARBINARY data types as outputs
- All the ST_GEOMETRY types listed in the SAP HANA Cloud, SAP HANA Database Spatial Reference are supported

Benefits

- Apply spatial facet to graph networks
- GraphScript algorithms can now evaluate the spatial characteristics of graph vertices and edges

Use Cases

- Identifying optimal routes
- Avoiding restricted areas when determining shortest paths
- Dynamic calculation of edge weights within a graph network



Spatial Pruning: Avoid no-go area

Innovation Highlights

GraphScript on JSON Document Store

Enable Graph Processing on JSON Documents

- Create a graph workspace on document store collections
- Support for the document store data model in graph engine
- Adjacency Index(Persistent Index) built on top of the JSON collection
- No need to rebuild after updates

Benefits

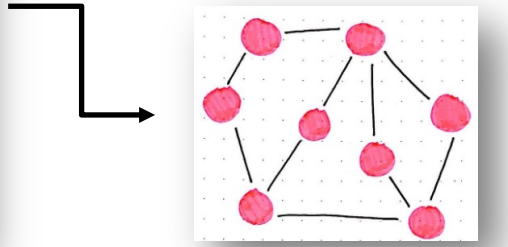
- Ability to analyze graph data stored in JSON document store
- Interoperability with other multi-model engines

Use Cases

- Social Network analysis with user profiles in unstructured JSON format
- E-Commerce Product Recommendation
- Fraud detection in financial transactions

```
CREATE COLLECTION MY_SCH.MY_COLLECTION;

CREATE HASH INDEX MY_HASH_INDEX ON MY_SCH.MY_COLLECTION ("id_path"."id");

-- ID, SOURCE and TARGET: 63 bit positive integer values
-- TYPE: string - either "vertex" or "edge"

CREATE ADJACENCY INDEX MY_ADJ_INDEX ON MY_SCH.MY_COLLECTION (
    ID "id_path"."id",
    SOURCE "src_path"."src",
    TARGET "tgt_path"."tgt",
    TYPE "type_path"."type");


CREATE [OR REPLACE] GRAPH WORKSPACE A_SCH.A_GRAPH ADJACENCY INDEX MY_SCH.MY_ADJ_INDEX;
```

```
[{
    "id": 0,
    "type": "vertex",
    "name": "Mary",
    "role": "senior manager"
},
{
    "id": 1,
    "type": "vertex",
    "name": "John",
    "role": "manager"
},
{
    "id": 2,
    "type": "vertex",
    "name": "Sara",
    "role": "architect"
},
{
    "id": 3,
    "type": "vertex",
    "name": "Paul",
    "role": "senior developer"
},
```



Vector Engine (Planned Q1 2024)

SAP HANA Cloud Vector Engine

Definitions

Vector

A vector is a list of numerical float values that has magnitude and direction.
In the context of Generative AI, it is used to represent ‘something’ (object) like a book, car, customer record, etc. that describes the object itself or its attributes or characteristics, which can further be compared

Embeddings

An embedding is one of the numbers in the vector representation that represents data in a way that captures meaningful information, semantic relationships, or contextual characteristics.

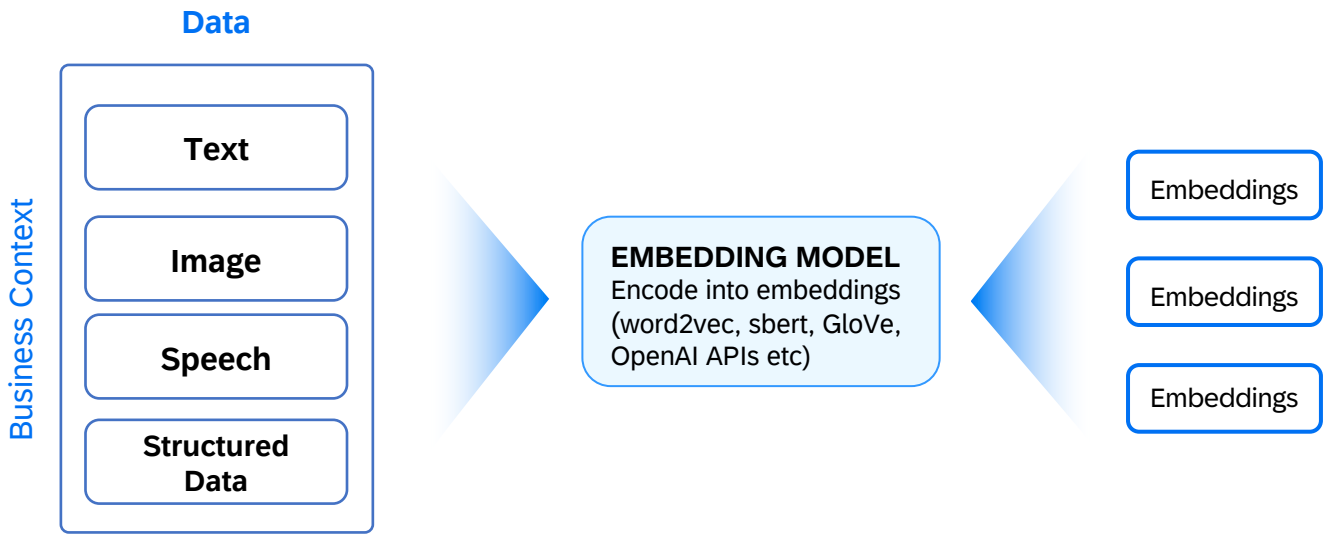
SAP HANA Cloud Vector Engine

Definitions

Embedding Model

A type of machine learning model that is designed to transform high-dimensional categorical data into lower-dimensional, continuous vector representations. This model is particularly used for transforming high-dimensional and complex data into a more concise and meaningful representation.

Example : Word2Vec, GloVe, Sentence Bert, Open AI APIs etc.



SAP HANA Cloud Vector Engine

Planned for Q1-2024 Release

In preview*
(*subject to change)

Vector data type: 'REAL_VECTOR'

- Natively store high-dimensional vectors
- E.g. text embeddings

Vector functions: L2DISTANCE(), COSINE_SIMILARITY()

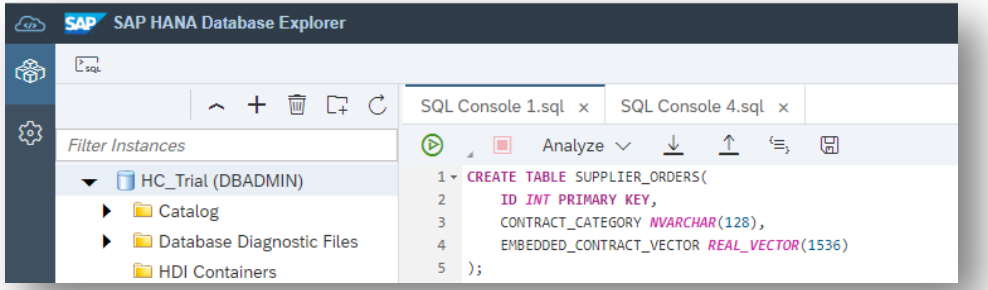
- Vector distance functions
- Combine vector functions with other SQL operations

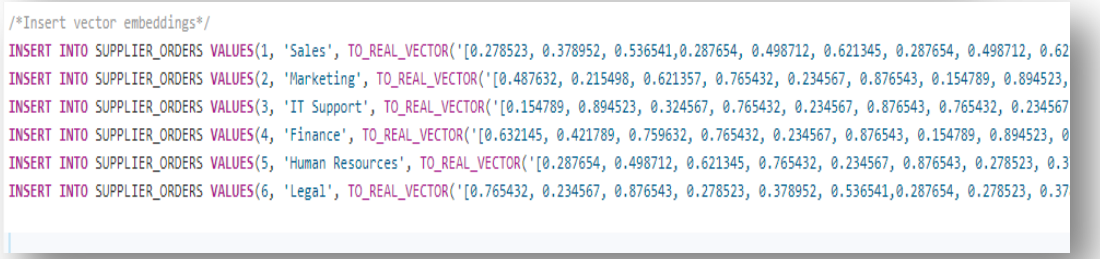
Consumption

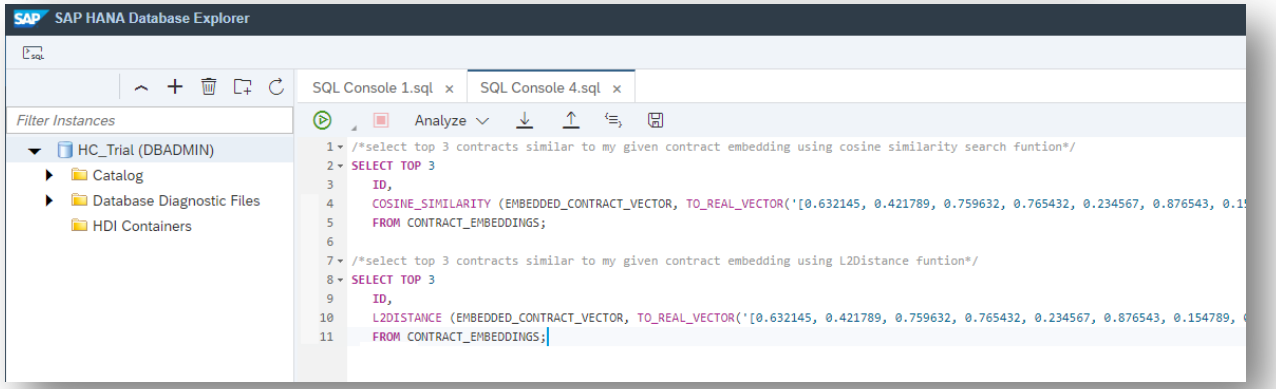
- SQL
- python (hana-ml)

Roadmap & Vision

- LangChain plug-in, CAP support, SAP Generative AI Hub integration
- Vector indexes, approximate nearest neighbor (ANN) search
- In-database text embedding







SAP HANA Cloud Vector Engine

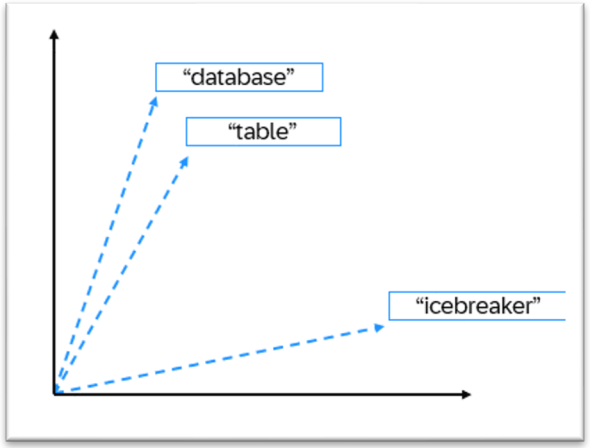
Similarity Search Example

In preview*
(*subject to change)

Search Term: “Database”

Search Term (attributes stored as Vector Embeddings): [0.18, -0.55, 0.77, ...]

Compare against table of records stored in the database:



	Term	Vector Embeddings	Distance	Rank
Most similar to ‘Database’ →	Table	[0.18, -0.65, 0.75, ...]	0.24	1
	Schema	[0.12, 0.75, 0.14, ...]	0.28	2
	Token	[0.67, 0.45, -0.16, ...]	0.57	3
Least similar to ‘Database’ →	Icebreaker	[0.89, 0.77, 0.34, ...]	0.68	4

Considering distance calculated using L2Distance() function

SAP HANA Cloud Vector Engine

Benefits

In preview*
(*subject to change)

Execute fast in-memory vector searches using SQL

Combine vector similarity search results with other business data

Ground generative AI queries for better responses

Implement “human intuition” specific use cases

*Planned Q1 2024

SAP HANA Cloud Vector Engine

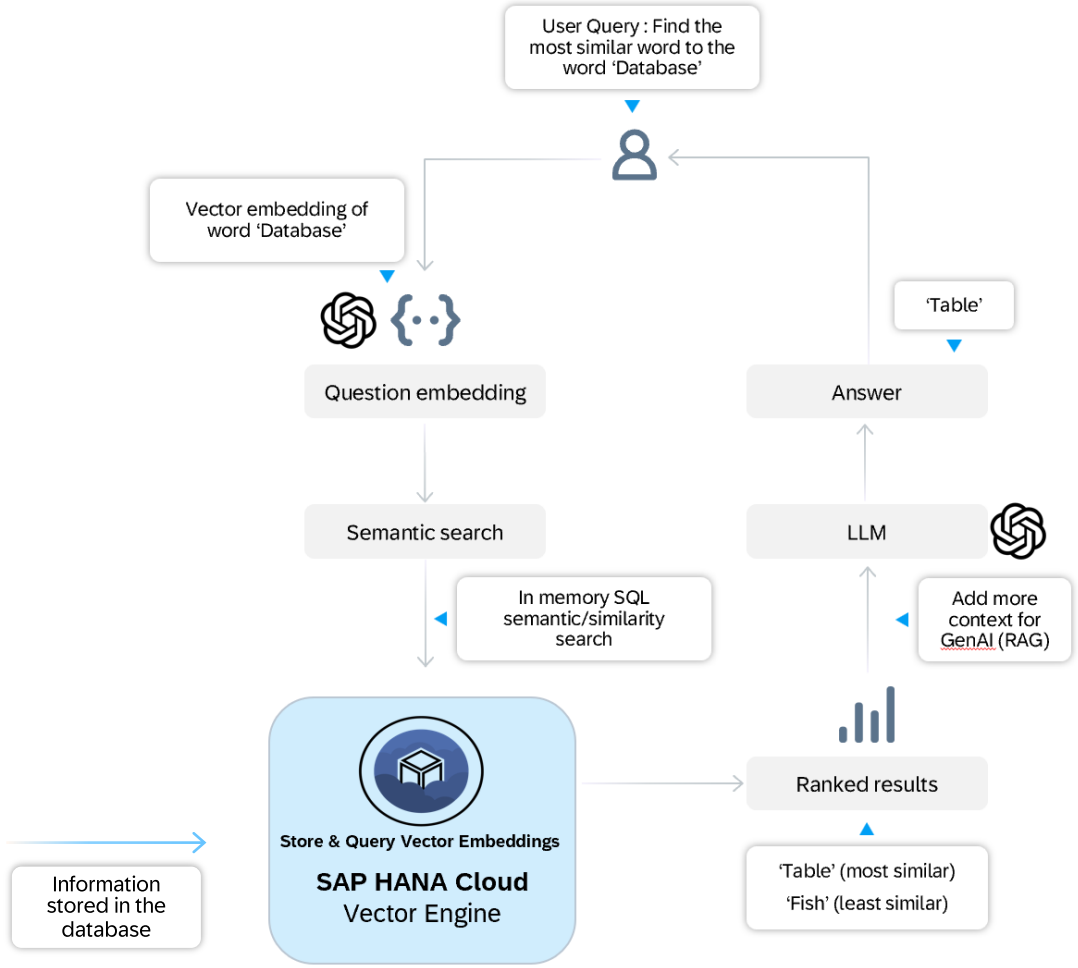
Execute fast in-memory vector searches using SQL

In preview* (*subject to change)

Store and retrieve vectors swiftly using SQL, optimizing data access and retrieval

Leverage the speed of in-memory processing for faster query response times

Term	Vector Embeddings
Table	[0.18, -0.65, 0.75, ...]
Schema	[0.12, 0.75, 0.14, ...]
Token	[0.67, 0.45, -0.16, ...]
Fish	[0.89, 0.77, 0.34, ...]



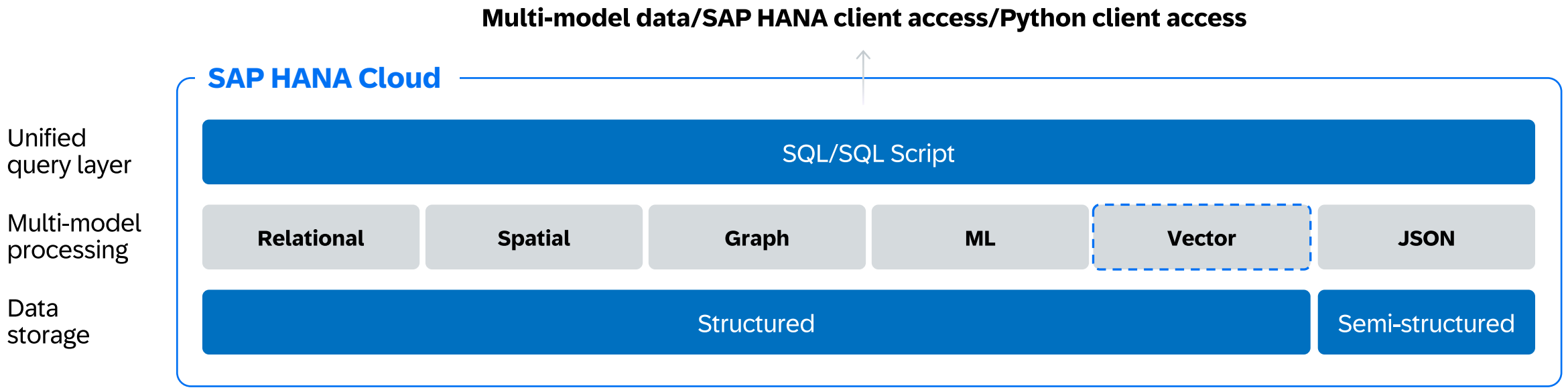
*Planned Q1 2024

In preview*
(*subject to change)

SAP HANA Cloud Vector Engine

Combine vector similarity search results with other business data

- Integrate business data with graph, spatial, document, and vector data in a single platform
- Eliminate the need for numerous specialized database platforms
- Prevent data silos & data duplication efforts



*Planned Q1 2024

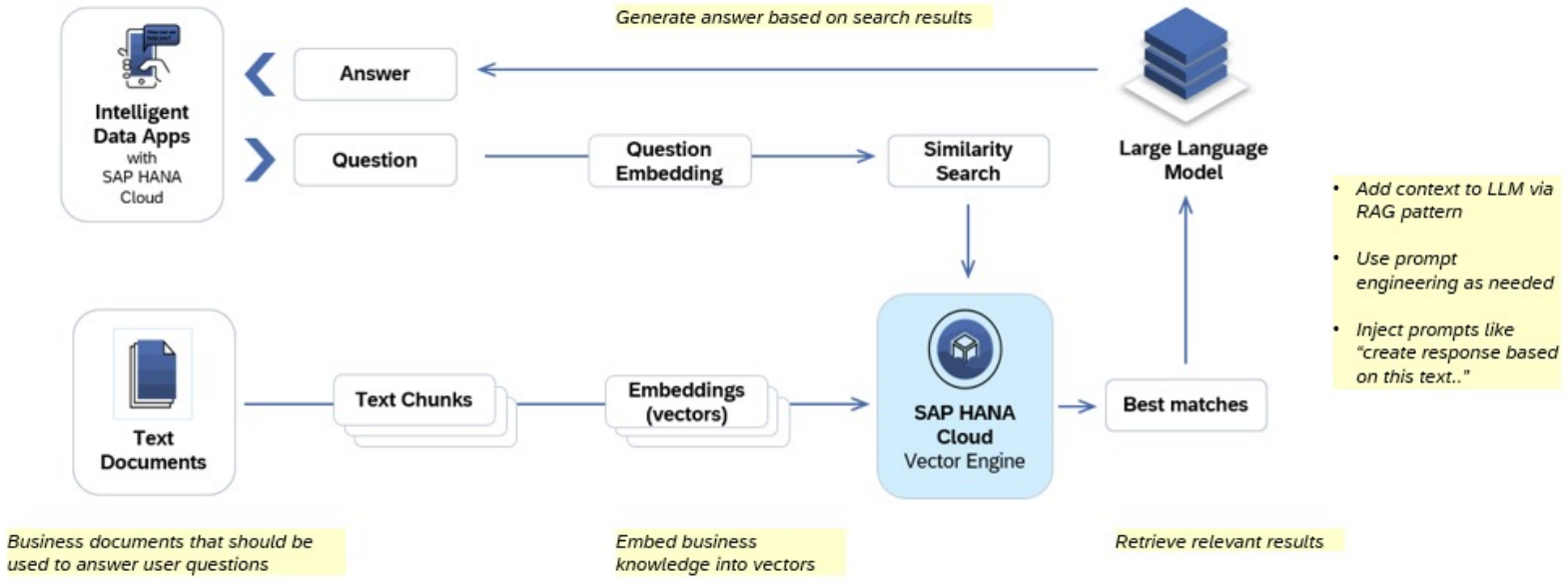
SAP HANA Cloud Vector Engine

Ground generative AI queries for better responses

In preview*
(*subject to change)

Contributes to grounding in Large Language Models by providing semantic and contextual aspects of language

Enable more contextually aware language processing



*Planned Q1 2024

SAP HANA Cloud Vector Engine

Use Case Overview

In preview* (*subject to change)

Retrieval Augmented Generation

Overcome insufficient or incorrect results from Large Language Models (LLM). SAP HANA Cloud's vector engine provides contextual data to add the right supplemental data for better responses.

Similarity Search

Allows users determine similarities based on the actual objects. Unlocks on-demand access to information crucial for trend analysis, pattern recognition, personalized recommendations, and various decision-making processes.

Semantic Search

Empowers users to compare the attributes or semantics of an object(s) for enhanced decision-making, content discovery, and knowledge extraction.

Information Retrieval

Facilitates interactive and efficient querying, allowing users to retrieve critical data for decision-making, analysis, and reporting with exceptional speed and precision. SAP HANA Cloud empowers users to navigate vast datasets seamlessly, transforming the way information is accessed and utilized.

Content Based Filtering

SAP HANA Cloud enhances user experience and optimizes engagement by delivering targeted content that resonates with each user's unique profile. This use case extends beyond e-commerce, influencing sectors such as streaming services, news platforms, and more.

Intelligent Data Applications

Build applications that redefine decision-making. This involves real-time data processing, integrating advanced analytics and machine learning, and adopting a scalable, flexible architecture. The unified data management approach also ensures comprehensive insights from diverse data sources.

*Planned Q1 2024

SAP HANA Cloud Vector Engine

Retrieval Augmented Generation

In preview*
(*subject to change)

A system first retrieves relevant information from a knowledge source on a given context or user query. This retrieved information is then used to augment or influence the subsequent process of content generation.

Examples:

- **Question Answering Systems:** Enhancing responses with real-time information
- **Content Summarization:** Creating concise and informed summaries
- **Chatbots:** Offering contextually relevant and personalized interactions
- **Content Creation:** Improving the quality of generated articles, reports, or responses
- **Adaptive Learning:** Refine retrieval and generation strategies based on user feedback

SAP HANA Cloud Vector Engine

Similarity Search

In preview*
(*subject to change)

Similarity search focuses on retrieving items that are similar or related to a given query, based on certain features or characteristics.

Examples:

- **Content Recommendation Systems:** Improving recommendations based on similar user preferences
- **Image Retrieval:** Searching for visually similar images in large datasets
- **Anomaly Detection:** Identifying unusual patterns or outliers in data
- **Product Matching:** Matching similar products in e-commerce platforms

SAP HANA Cloud Vector Engine

Semantic Search

In preview*
(*subject to change)

Semantic search involves understanding the meaning behind the user's query and the content being searched. This process goes beyond keyword matching and takes into account the context, intent, and relationships between words.

Examples:

- **Natural Language Processing (NLP):** Enabling more precise and context-aware language understanding
- **Semantic Image Retrieval:** Finding images based on semantic content rather than just visual appearance
- **Document Similarity Analysis:** Identifying documents with similar semantic meaning
- **Customer Support Chatbots:** Enhancing conversational understanding for more accurate responses
- **Content Tagging and Classification:** Improving content organization based on semantic context

SAP HANA Cloud Vector Engine

Information Retrieval

In preview*
(*subject to change)

Information retrieval refers to the process of finding relevant information using similarity versus Boolean-style logic.

Examples:

- In Text-Based Applications for:
  - **Document Retrieval:** Finding relevant documents based on vector representations
  - **Question Answering Systems:** Enhancing accuracy in finding answers to user queries
- In Real-Time Applications for:
  - **Web Search:** Provide users with instant and relevant search results as they type
  - **Sensor Data Retrieval:** Gather and process data from sensors and devices, enabling quick decision-making

SAP HANA Cloud Vector Engine

Content Based Filtering

In preview*
(*subject to change)

Search for items based on intrinsic content and features where each vector contains the inherent characteristics or attributes of the object.

Examples:

- **Topic Matching**
  - News Platforms
  - Recommend content based on the user's interests and previously read articles
- **Feature Matching**
  - E-Commerce Platforms
  - Recommend products based on the user's preferences, historical purchases, and current cart selections

Vector Engine and Generative AI

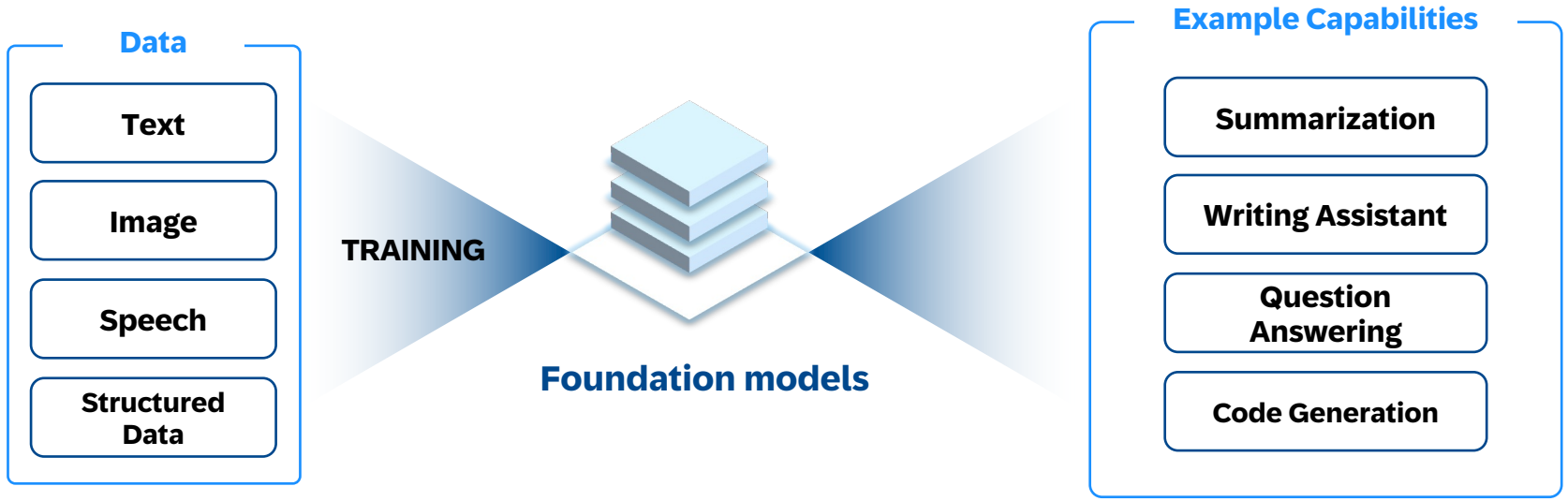
Power of Generative AI

In preview*
(*subject to change)

Execute business tasks using natural language

Overcome inadequate or incomplete pre-trained foundational models

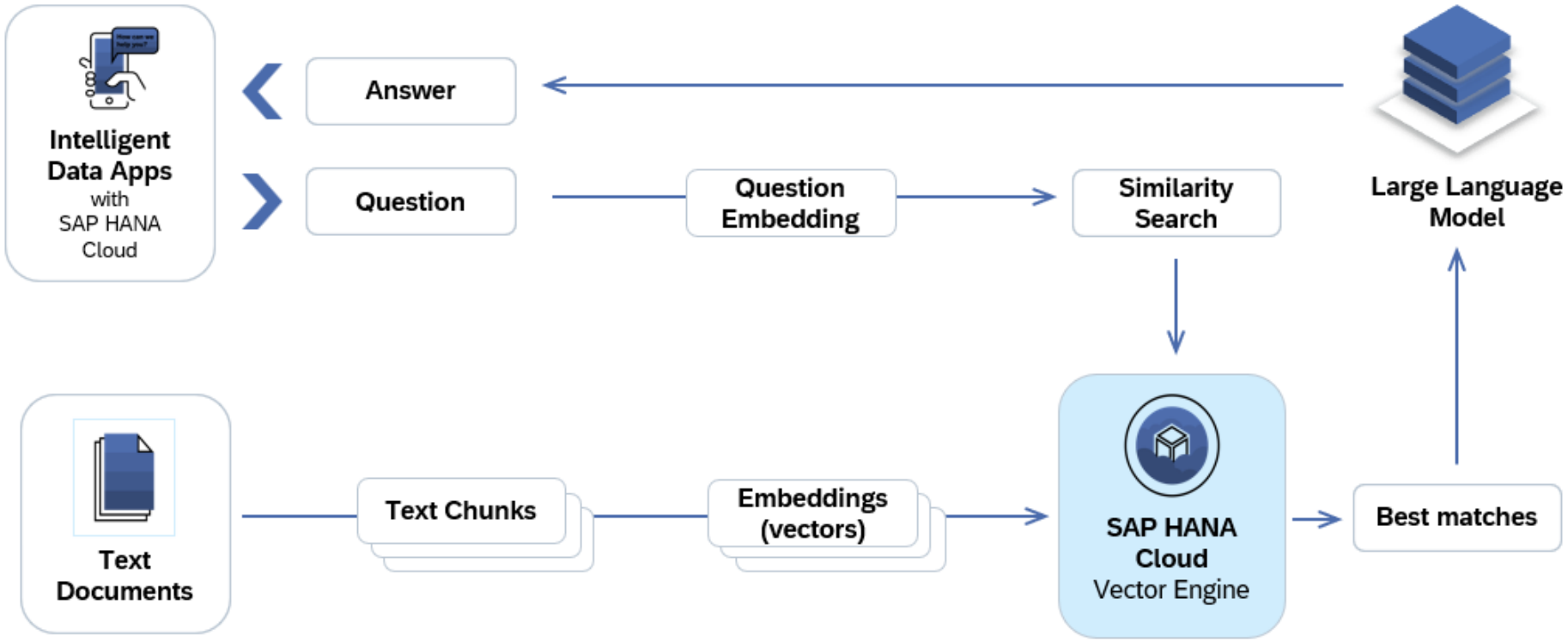
Improve results using prompt engineering, fine tuning, and retrieval augmented generation (RAG).



*Planned Q1 2024

Retrieval Augmented Generation

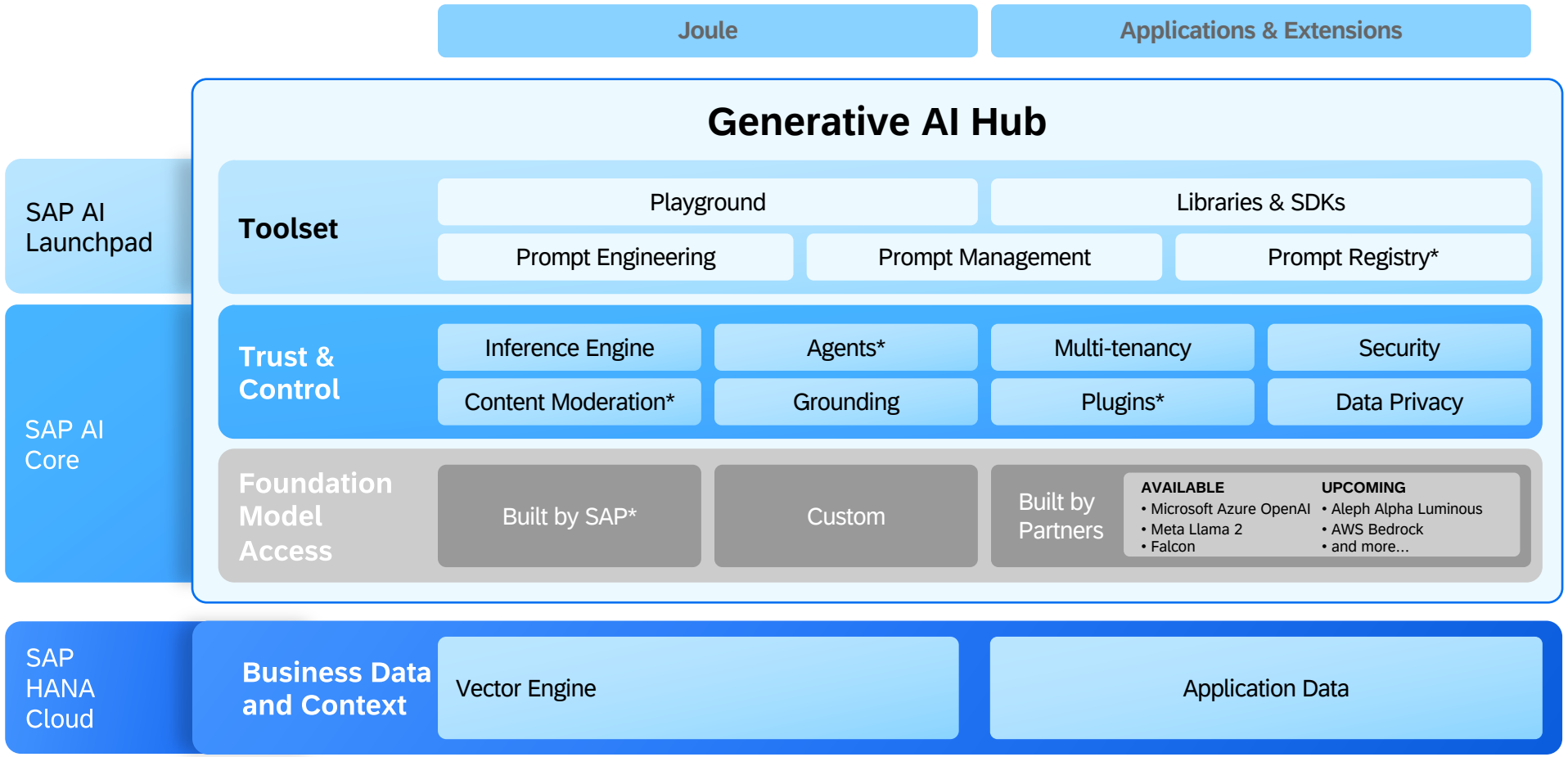
In preview*
(*subject to change)



*Planned Q1 2024

SAP Generative AI Hub

In preview* (*subject to change)



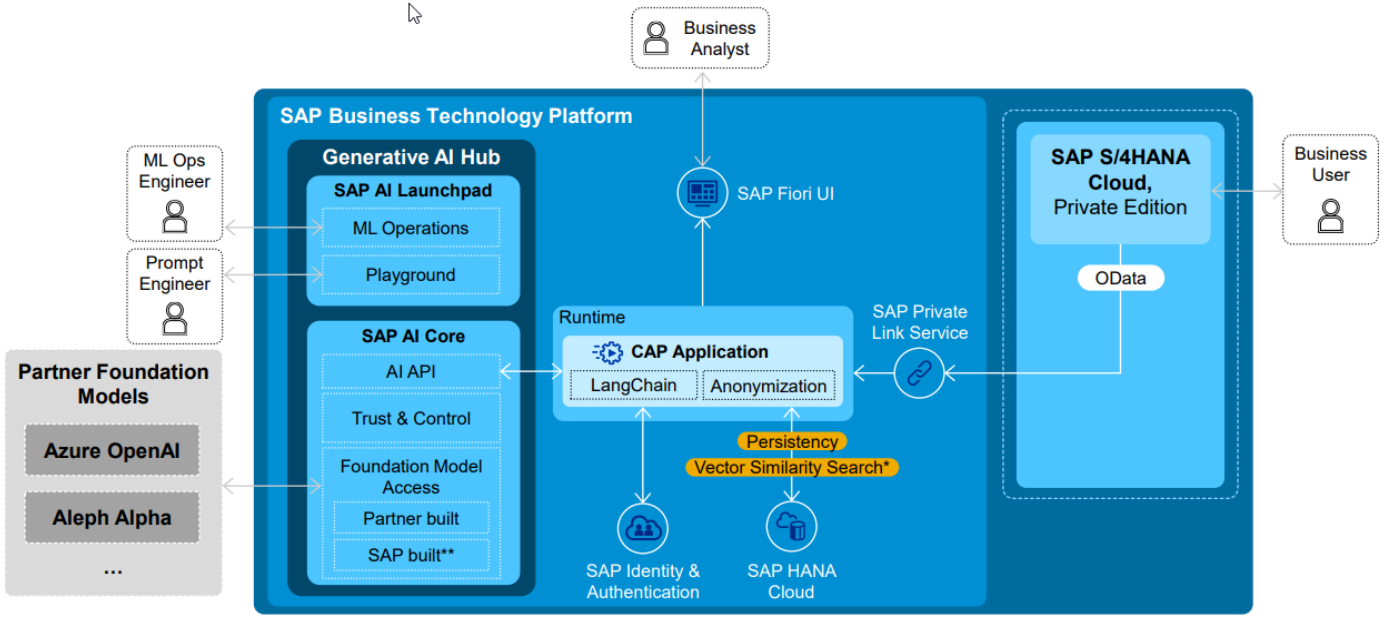
*Planned 2024

[GenAI Hub Roadmap]

SAP Cloud Application Programming (CAP)

Build GenAI-enabled Intelligent Data Applications

In preview* (*subject to change)



- Provide end-to-end support for Vector datatype & built-in function
- Support for semantic search and RAG capabilities
- CAP CDS to support vector data type by Q1-2024

*Planned 2024

Demo

Q&A

Find out more about SAP HANA Cloud

Learn about SAP HANA Cloud

Check out the **sap.com/hanacloud** website, which has valuable resources for fast-tracking your knowledge of SAP HANA and a rich support section designed to help you get the highest quality answers quickly and easily from SAP experts

Read our blogs
- community.sap.com

Get started for Free
- sap.com/hanacloud

Customer stories
- sap.com/hanacloud

Roadmap
- roadmaps.sap.com

Get involved in the discussion

Engage with community experts on the SAP Community program to accelerate the development of SAP HANA Cloud powered solutions

Influence the future
- influence.sap.com

Stay current
- youtube.com/SAPTechnology/SAPHANACloud
- #whatsnewinsaphanacloud

Spread the word
- https://twitter.com/sapBTP

SAP is here to help.

Contact your local SAP representative

sap.com/corporate/en/company/office-locations.html

Thank you

Contact information: