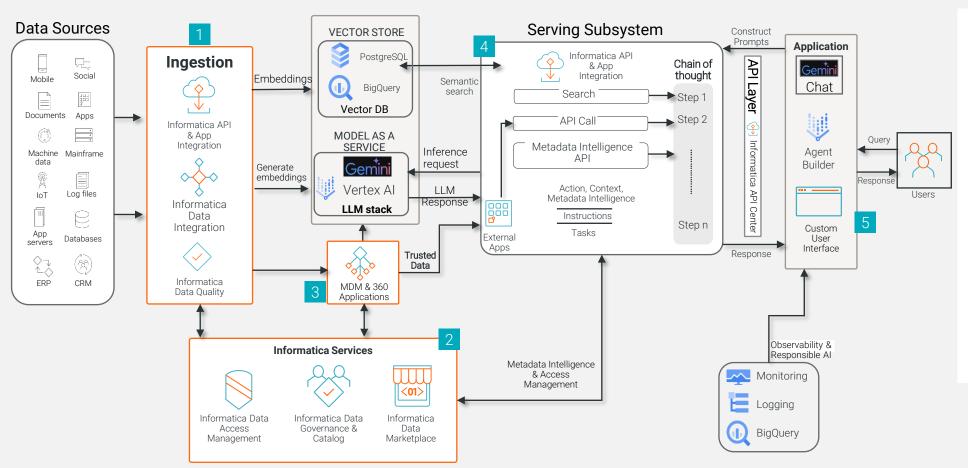
Reference Architecture

GenAl Blueprint for Google Cloud

Retrieval Augmented Generation Framework for GenAl with Informatica on Google Cloud



- Ingestion—Data is ingested from various sources into IDMC. Once ingested, Vertex Al generates embeddings, which are then stored in the Google BigQuery vector database.
- Metadata Intelligence & Access
 Management Data governance, catalog,
 access management, and marketplace
 enhance Generative AI by ensuring highquality, well-managed data with secure, rolespecific access and rich metadata, boosting
 both the precision and innovation of
 retrieval-augmented generation models.
- 3 Master Data Trusted data is provided using Informatica's Master Data Management (MDM), which assists organizations in maintaining a consolidated, accurate record.
- 4 Serving Subsystem —The serving subsystem has been built using Informatica API and Application Integration (CAI). Here, the chain of thought is constructed by CAI through API calls to external systems for metadata or Master data.
- Frontend— This can be a custom User Interface or existing applications. When the user inputs a query, it initiates the RAG chain.

