Cambridge Semantics: Webinar

July 2023

Greg West - Principal Presales Engineer

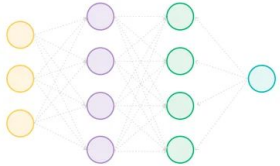
CAMBRIDGE SEMANTICS

Product Demo

CAMBRIDGE SEMANTICS

How does Knowledge Guru compare to a Large Language Model?

KNOWLEDGE GURU



	Knowledge Guru	Large Language Model
Scope	Constrained to specifically an organization's proprietary data	Open World - trained on a huge corpus of public data
Accuracy	Accurate and precise answers - will notify user if query is “out of bounds”	Answers are probabilistic, prone to Hallucinations
Provenance	Can provide query logic used to provide response	Responses are non-deterministic and opaque
Data Security	Role Based Access Control	Custom development

How: The Architecture of Knowledge Guru

CAMBRIDGE SEMANTICS

- **Multi-user web app**

  The “chatboard” is implemented as an Anzo Hi-Res dashboard, leveraging platform infrastructure for access control and data management.

- **Query execution**

  User questions are sent to the Anzo server, where a service manages communication with the LLM. Generated queries are executed on AnzoGraph with the user's identity and results are displayed in the chat session.

- **LLM integration**

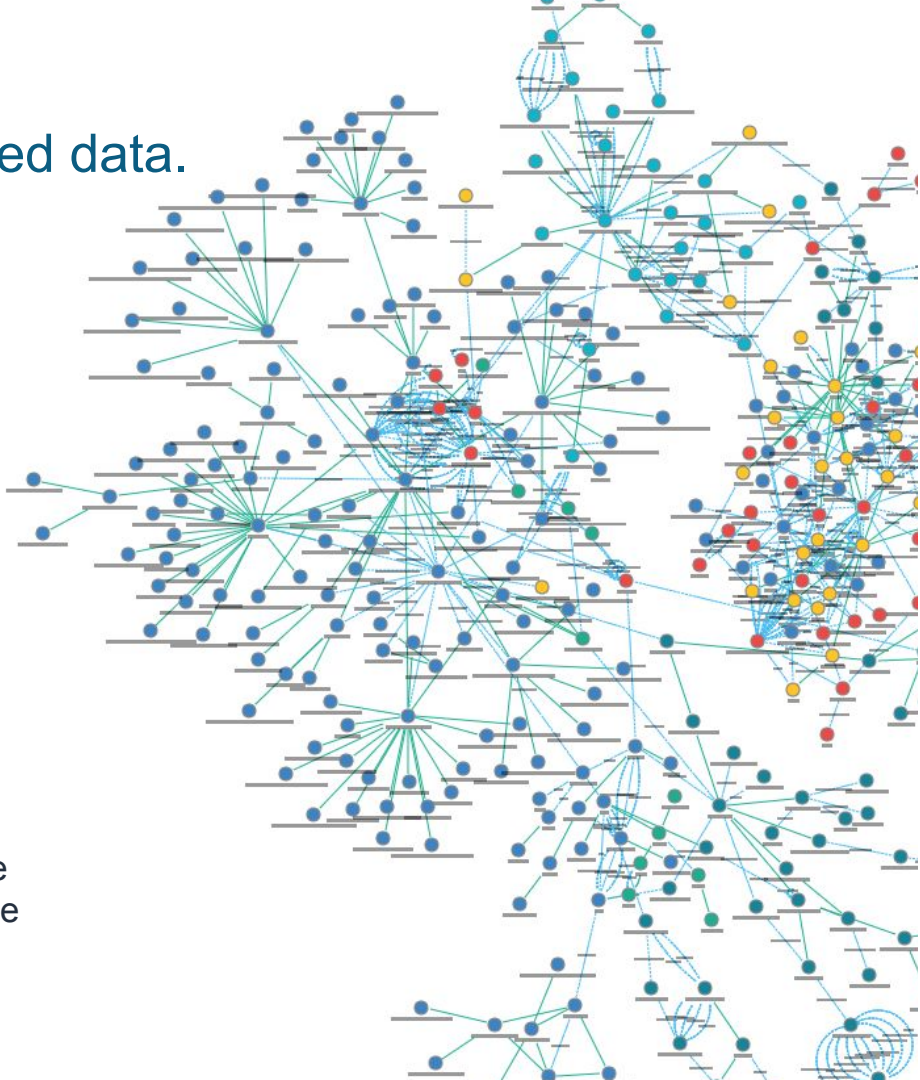
  The app is supported by OpenAI and MS Azure's GPT-4/3.5-turbo models, with configurable endpoints for future LLM availability.

How: Where is the real magic?

- Flexibility of Knowledge Graphs for simplifying complex data representation
- OWL Ontologies are the “Rosetta Stone”: OWL makes data both machine readable and human readable and elevates data to a conceptual level without source representation/optimization “clutter” & removes ambiguities
- Ontology provided as context to constrain the LLM response
- LLM transforms the human question into a graph (SPARQL) query drawing its predicates only from the ontology
- Anzo’s support for ad hoc queries can handle any valid query thrown at it

Knowledge Graphs:
Connect siloed data. Simplify complicated data.

- Simplifies access to complex data to address unanticipated questions
- Quickly profiles, connects and harmonizes data from multiple sources, including unstructured
- Presents tailored views, services and experiences to different personas with conceptual models
- Flexibly accommodates new data sources and use cases on the fly, with minimal impact
- CSI enables scalable knowledge graphs capable of accomodating enterprise data sources and use cases



ANZO

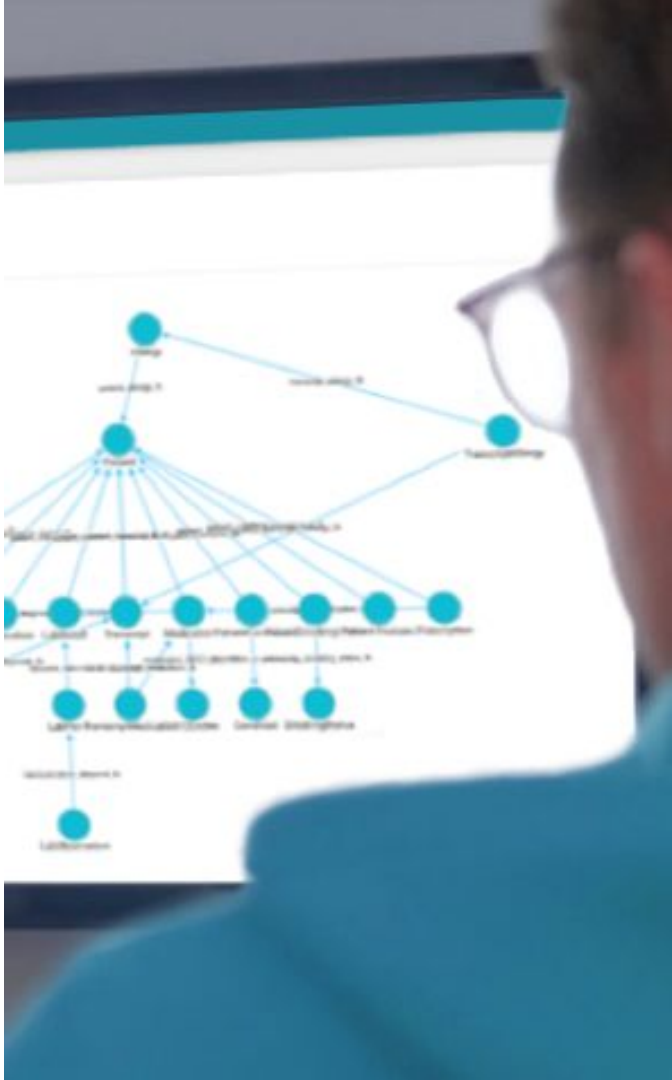
A **scalable knowledge graph platform** for modern data integration and analytics

Anzo connects and models related data in a real-world representation of data at scale, surfacing new insights and fueling pervasive analytics.

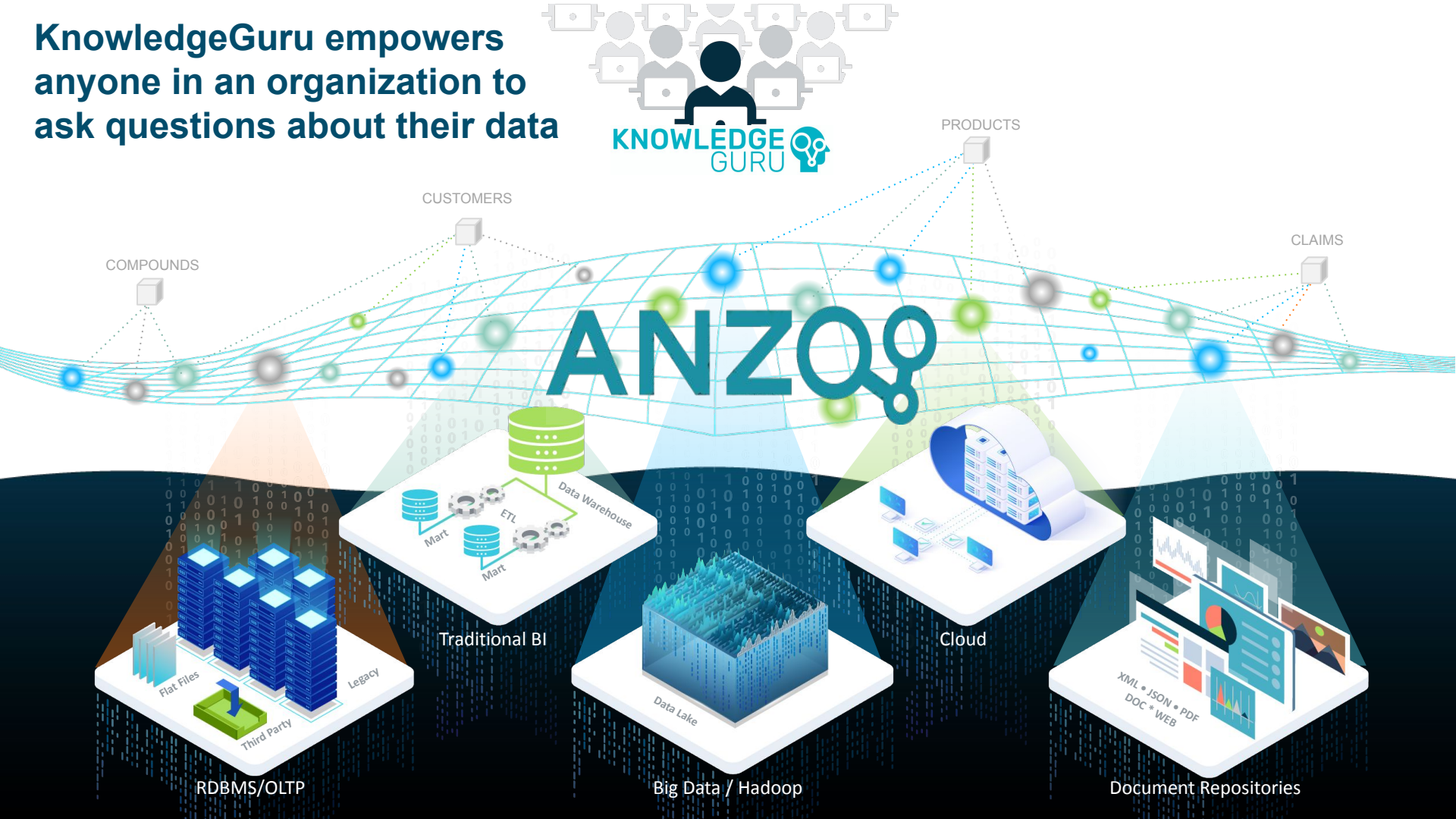
Knowledge Graph Management and Metadata Cataloging

***AnzoGraph* MPP OLAP Knowledge Graph Engine**

Enterprise-grade cloud deployment and security

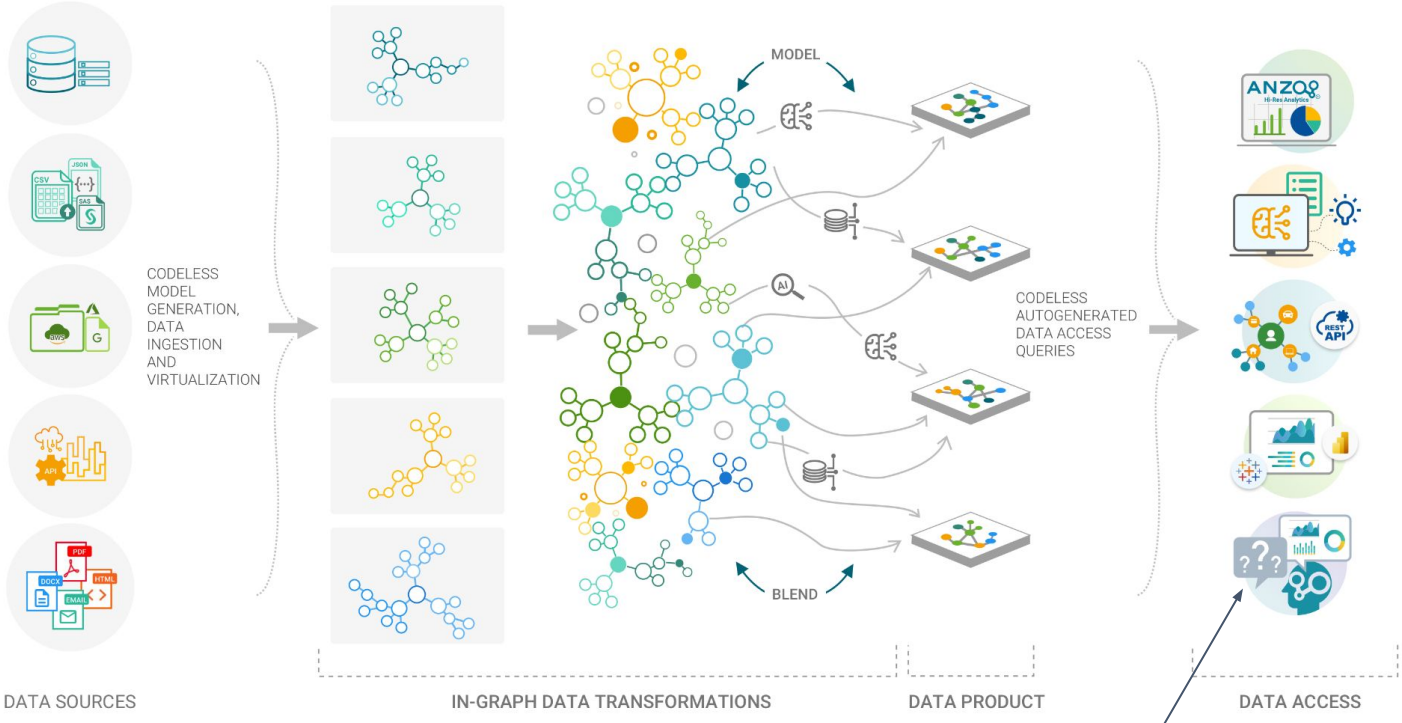


KnowledgeGuru empowers anyone in an organization to ask questions about their data



Anzo - The Knowledge Graph Platform

Turn your data into actionable knowledge.



KNOWLEDGE GURU

Anzo Interoperates with the Data ecosystem

Graphical Application Interface

Knowledge Graph Management

Cataloging, Metadata Management, Governance, and Lineage

Enterprise Data Sources

METADATA

ONBOARD
Register & Capture

- Direct Data Onboarding via GDI
- Metadata Capture

MODEL
Graph Data Model

- Automatic Ontology Generation
- Design Ontologies
- Connect Data Models

BLEND
GraphMarts

- Find Connections - automatically detect relationships
- Data Layers - Combine and Align Related Data Sets

ACCESS
Hi-Res Analytics

- Analyze All Data Together
- Fast, Iterative Queries Ad Hoc, What if
- Code-Free or API

DATA ONBOARDING VIA GDI

AnzoGraph MPP In-Memory Knowledge Graph Engine

Automated Deployment and Operations with Kubernetes

OPTIONAL PERSISTENCE

Data Storage Layer

aws · Windows Azure · Cloud or On-Prem Data Storage Infrastructure · hadoop · Google Cloud

Machine Learning and AI

DataRobot

"Last Mile" Analytics Tools

tableau · sas · TIBCO Spotfire

Enterprise Search

elasticsearch

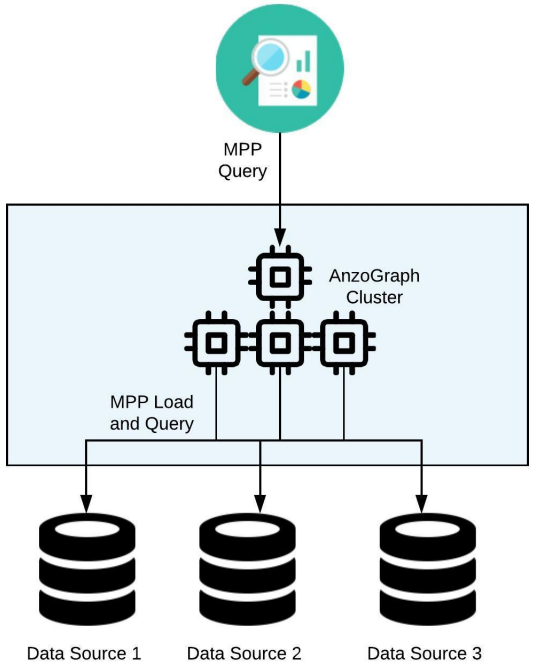
Scaling Virtual Knowledge Graphs

The AnzoGraph engine offers three ways to load data into memory for ***flexible virtualization***

1. Load onboarded RDF graph data from disk

2. Load data into memory directly from sources, APIs, streams (can persist if desired)

3. Load data at query-time through virtualized views



Product Demo

CAMBRIDGE SEMANTICS

Q & A

CAMBRIDGE SEMANTICS