

IFC-Bank of Italy Workshop on "Data science in central banking"

18-20 February 2025

SDMX chatbot: a trustworthy assistant to navigate the SDMX documentation¹

Samuel Melm and Brian Buffett,
Bank for International Settlements

¹ This contribution was prepared for the workshop. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.



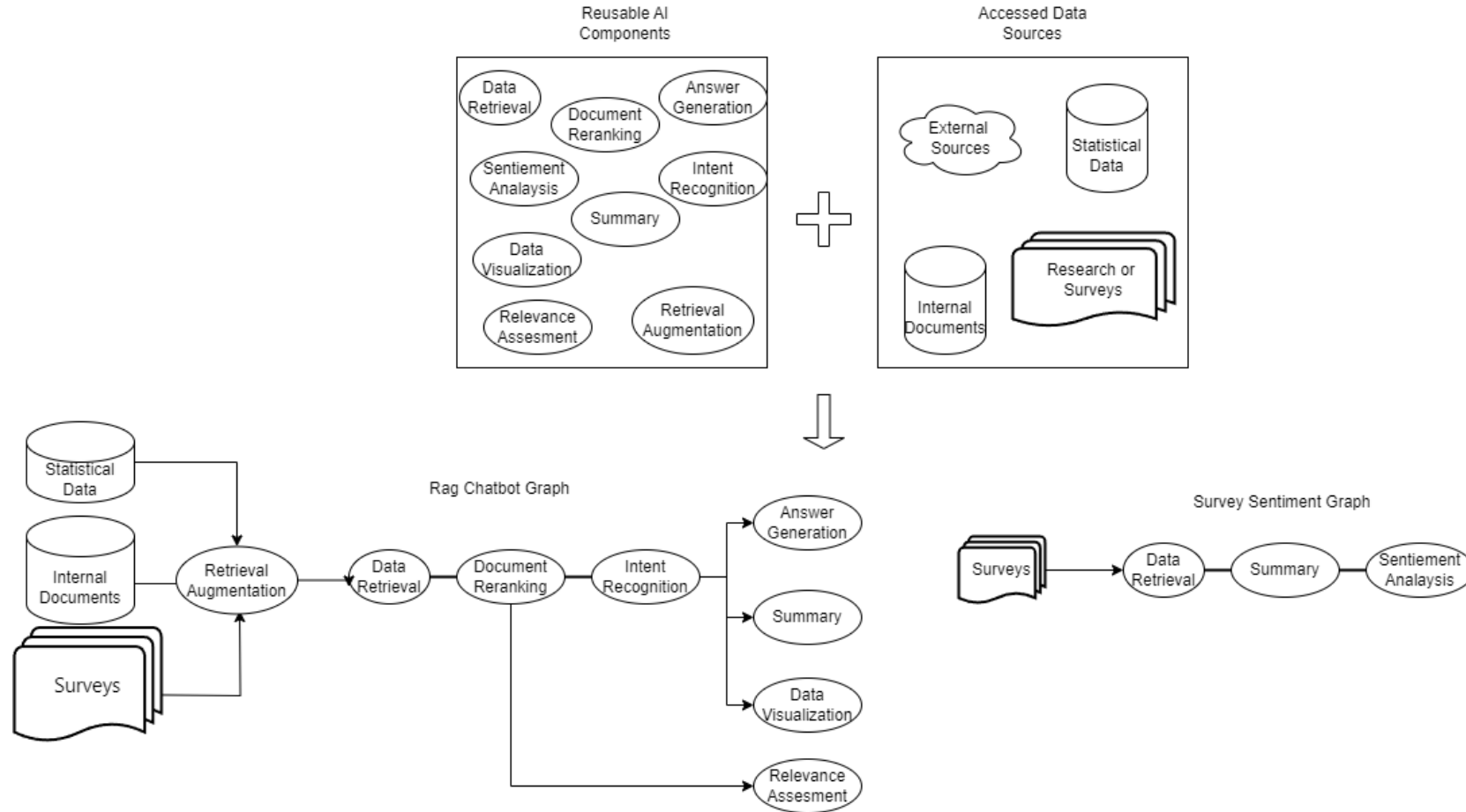
SDMX Chatbot

A Trustworthy Assistant to Navigate the SDMX Documentation

AI at BIS

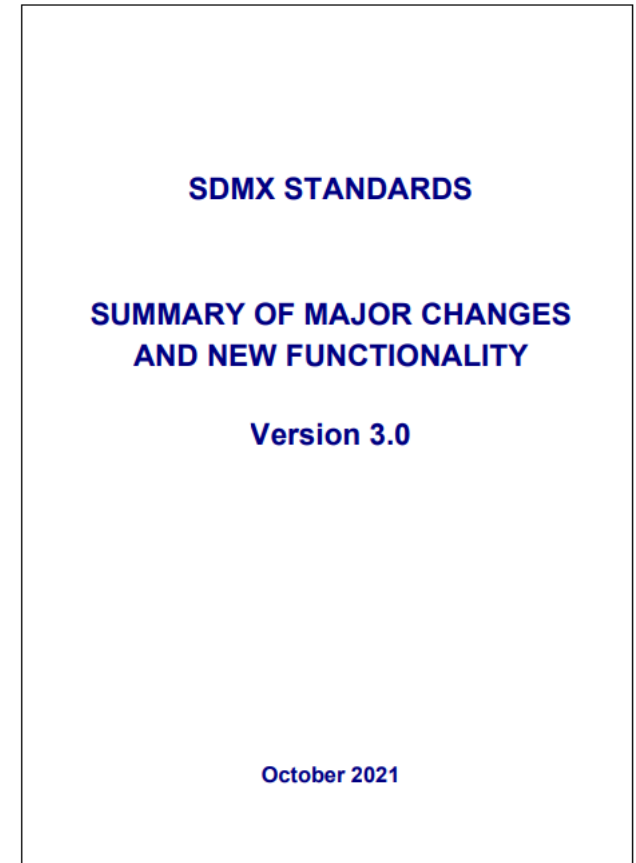
- BIS explores AI in various areas
- State of the art is expanding at a rapid pace
- Self-hosted language and code models
- Development of multiple RAG applications for variety of users
- Modular approach with high software engineering standards

Reusable Graph Architecture



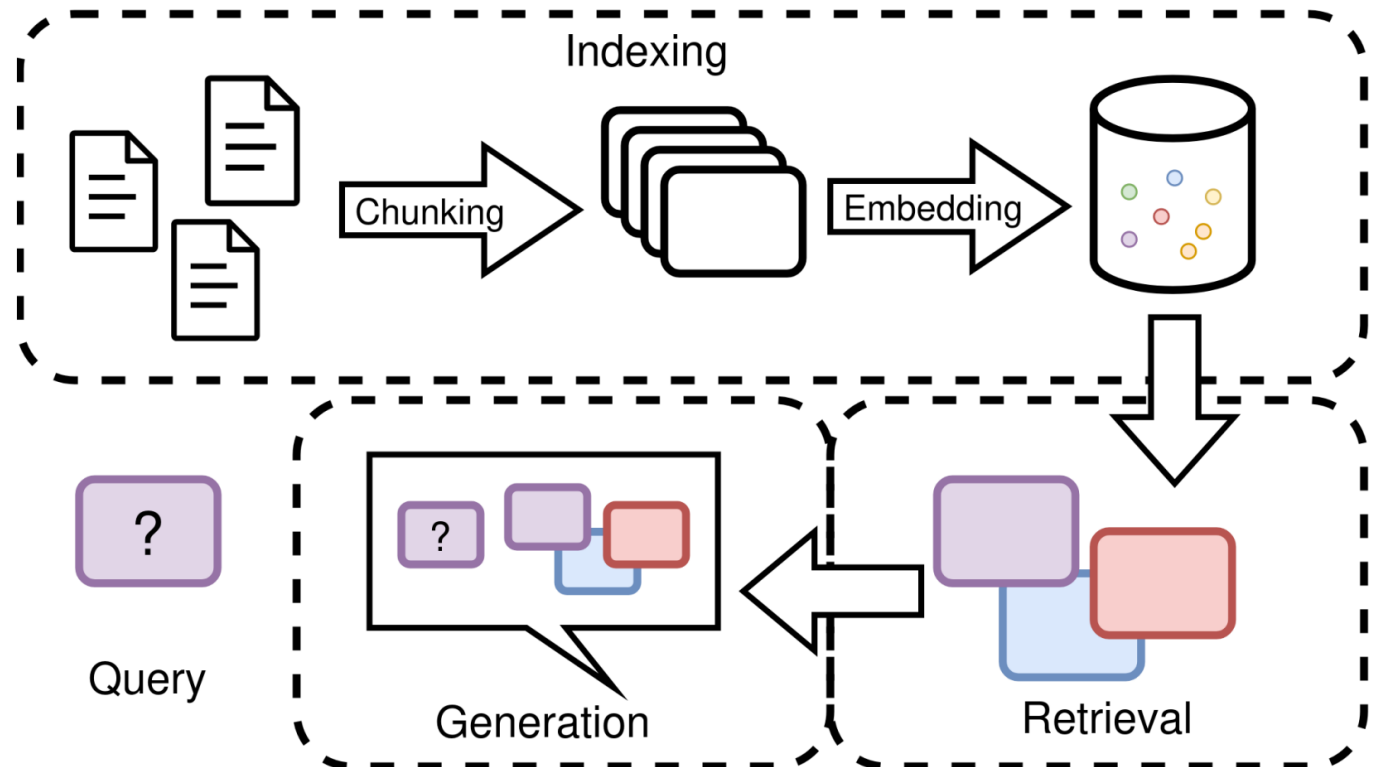
Goals

- Improve engagement with the SDMX standard
- RAG-based question answering system
 - Ingest all SDMX documentation into knowledge base
 - High-quality answers
 - Transparency
 - Verifiability



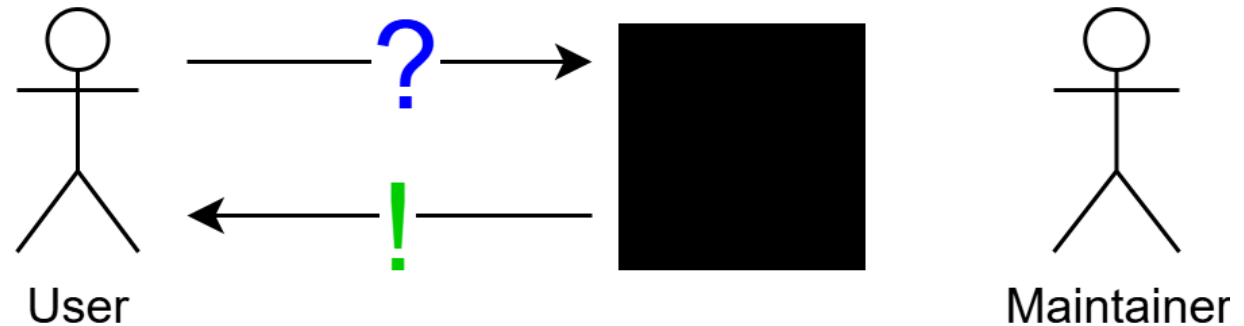
Retrieval-Augmented Generation (RAG)

- Split documents into chunks
- Create vector embeddings
- Retrieve relevant chunks
- Generate answer with LLM

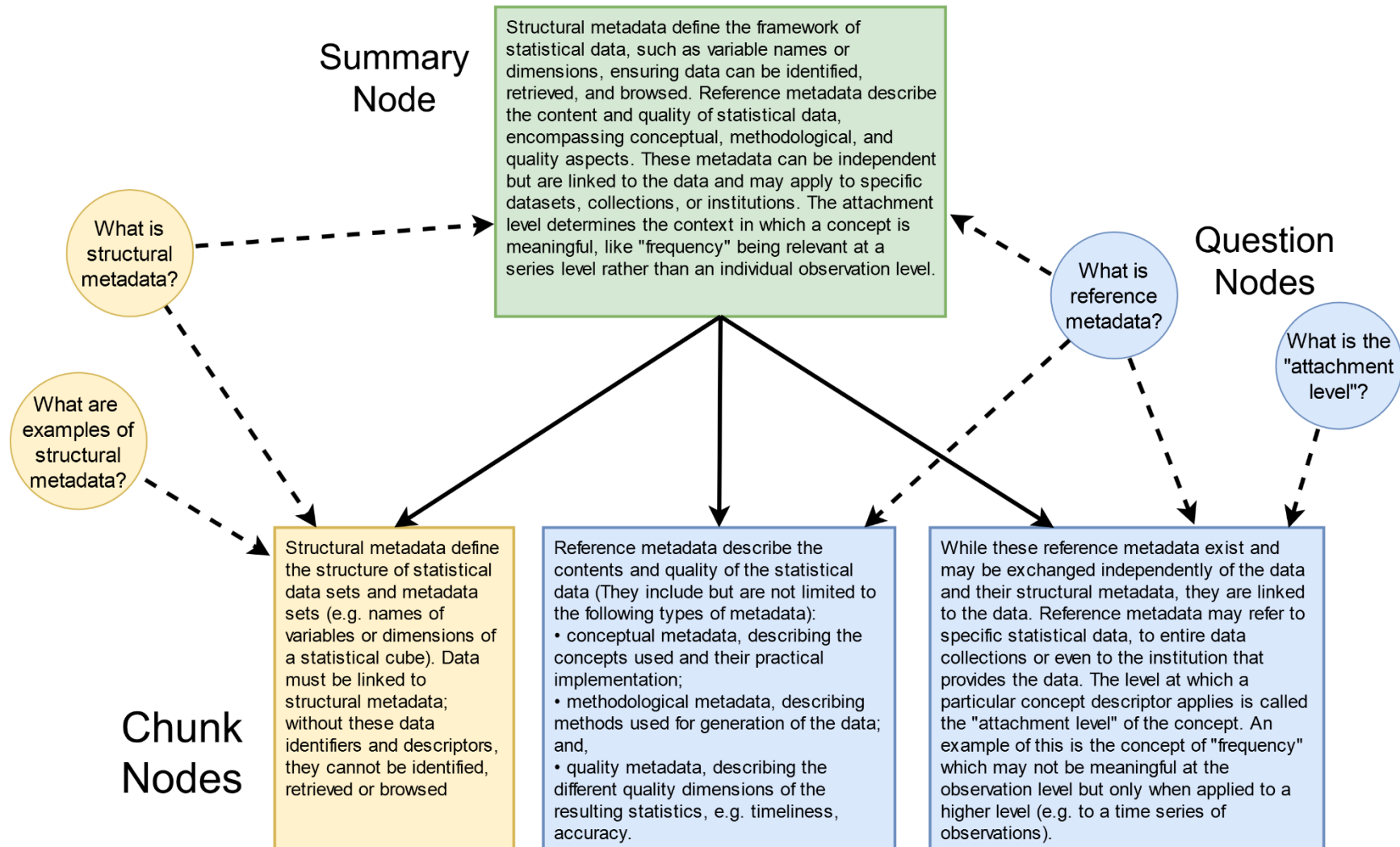


Transparency

- RAG systems are black boxes
- Users see only question and answer
 - Citations
- Maintainers must understand...
 - ...capabilities of the system
 - ...coverage of the documentation

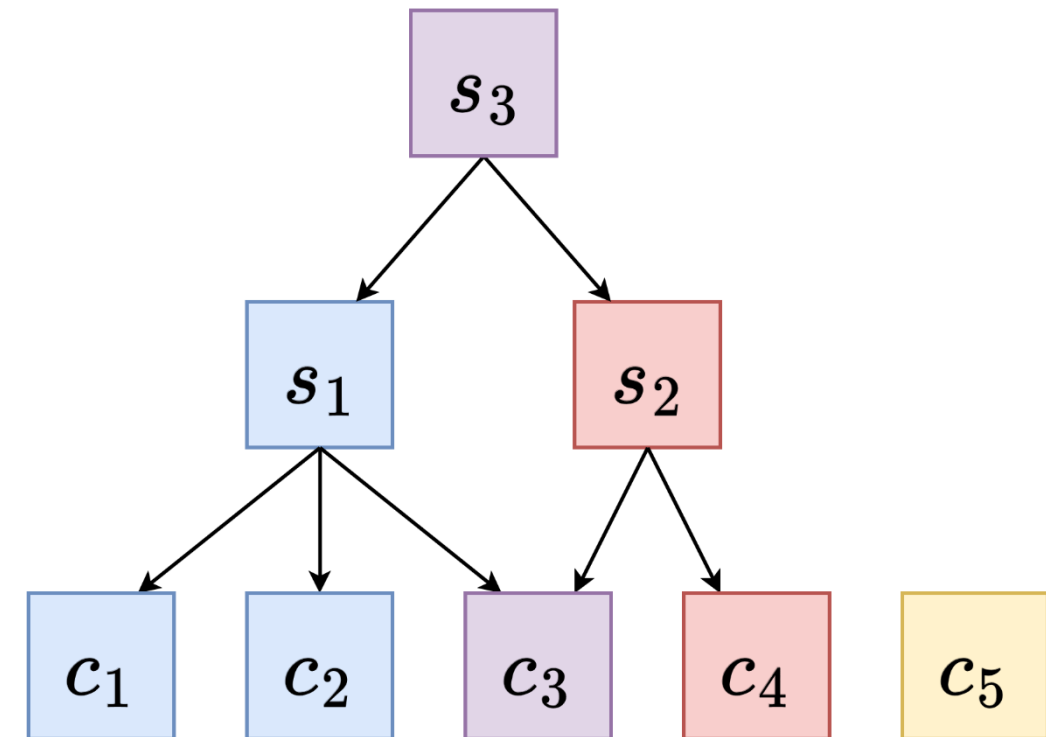


Document-Question Network

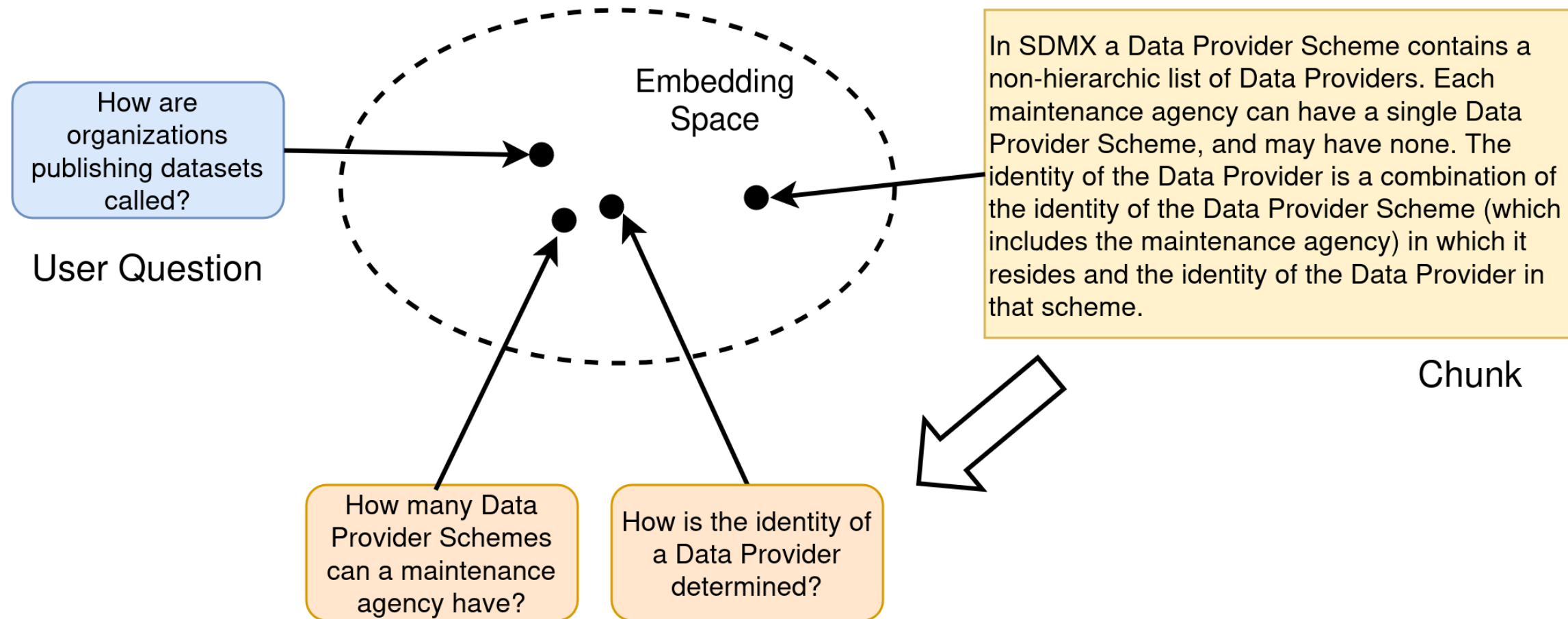


Summarization

- Goal: Improve answer quality
- Summaries are retrieved alongside chunks
- Based on RAPTOR
 - Clustering with UMAP + GMMs
 - Recursive summarization
- Reveals topics in corpus
- Reduces details and introduces abstraction
- Compress the information into fewer elements

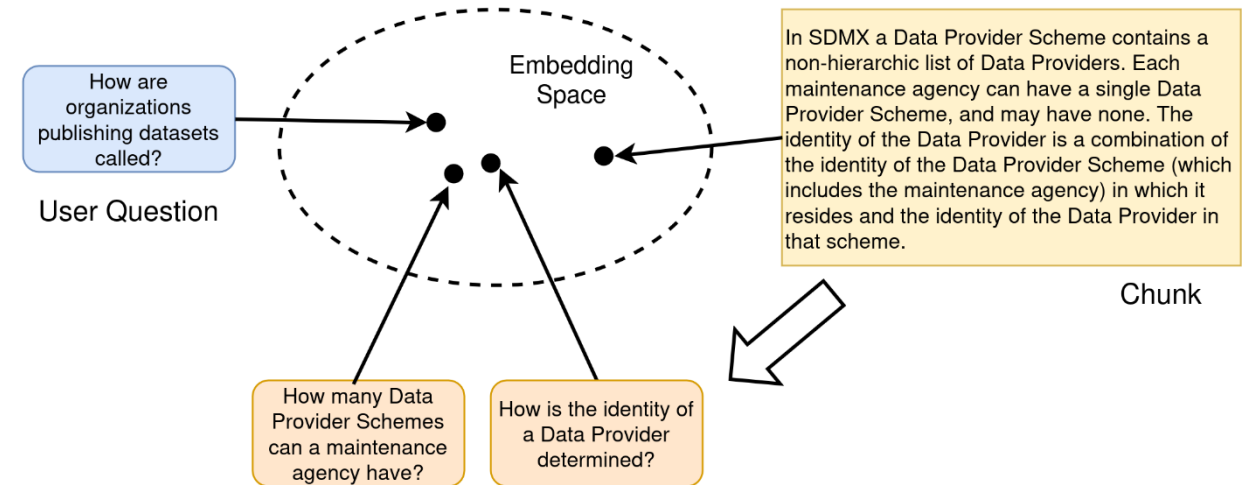


Hypothetical Questions



Hypothetical Questions

- Commonly used to improve retrieval
- Generate questions answered by chunk
 - Closer to user query
- What does chunk answer?
- What would a user ask?
- How would the system answer?
- Caching of answers



Testing & Evaluation

- SDMX experts are unreplaceable
 - Sustainability: can parts be automatized?
- Initial dataset of 100 evaluation questions
 - Use automatic question generation to find more
- RAGAS
 - LLM-based evaluation framework
 - Multiple metrics
 - Factual correctness
 - Faithfulness
 - Context quality (precision & recall)

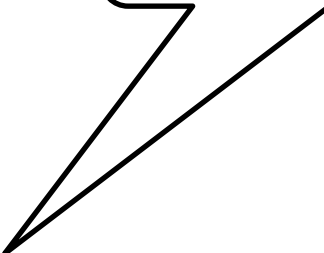
Results & Outlook

- Results compared to basic RAG pipeline
 - Up to 14% increase in correctness
 - Summarization
 - Filtering of context
 - Hypothetical questions reduce correctness
 - A lot more insights into the knowledge base
- Outlook
 - Beta testing with SDMX experts
 - Expansion of source documents
 - Presentation to the SDMX secretariat
 - Integration into sdmx.org search, following SDMX secretariat approval

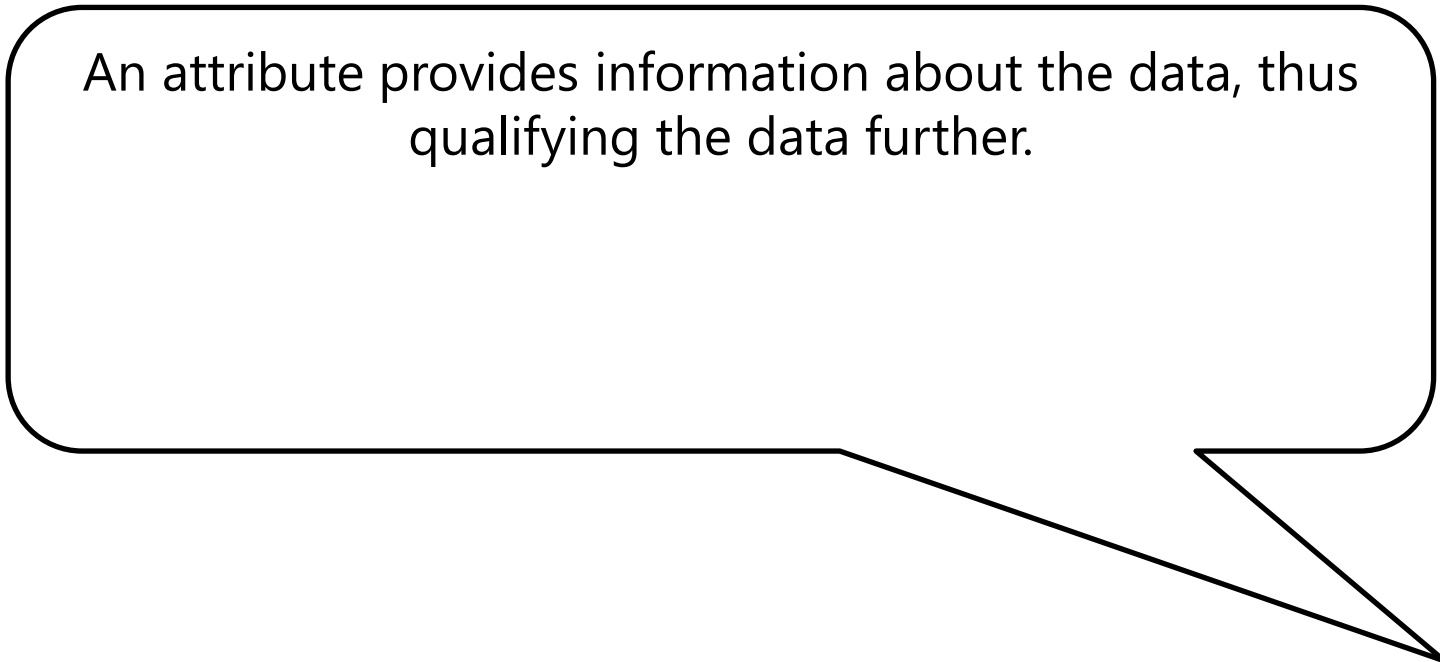


Demo

Example 1: Basic RAG

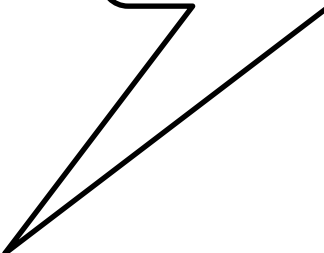


What is an attribute?

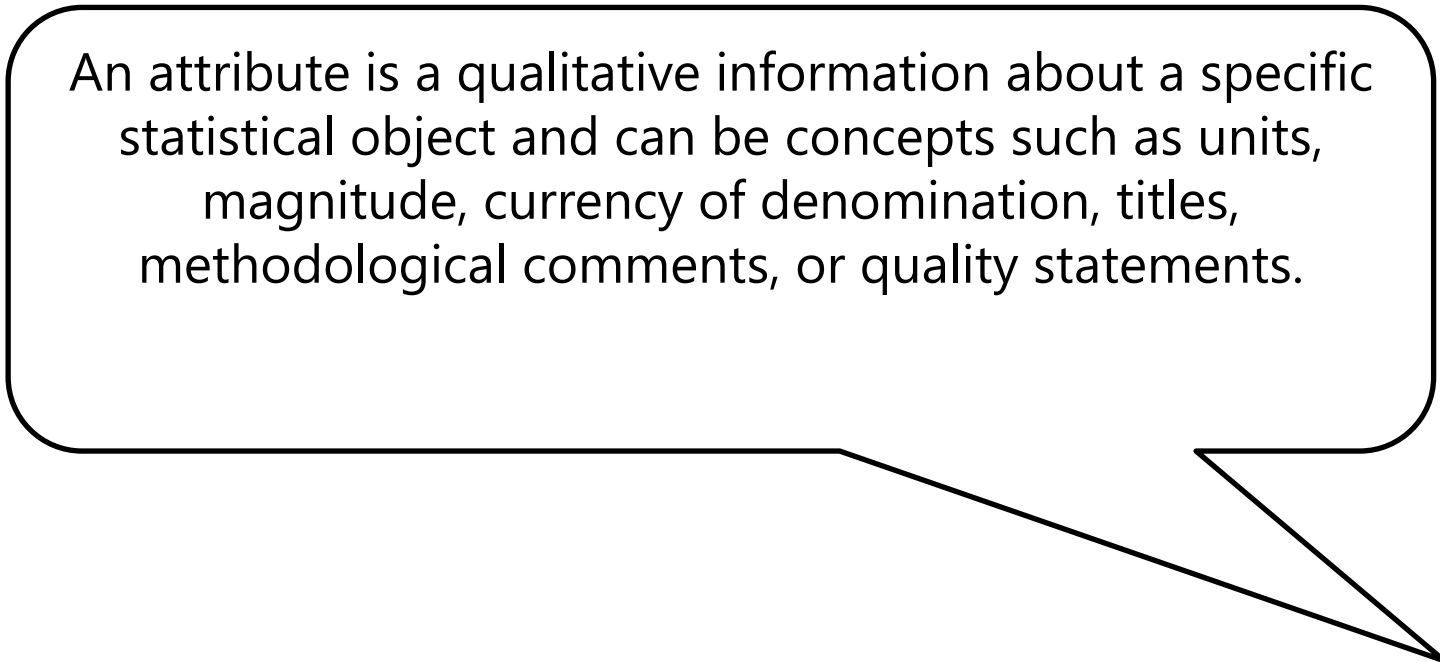


An attribute provides information about the data, thus qualifying the data further.

Example 1: Improved RAG



What is an attribute?



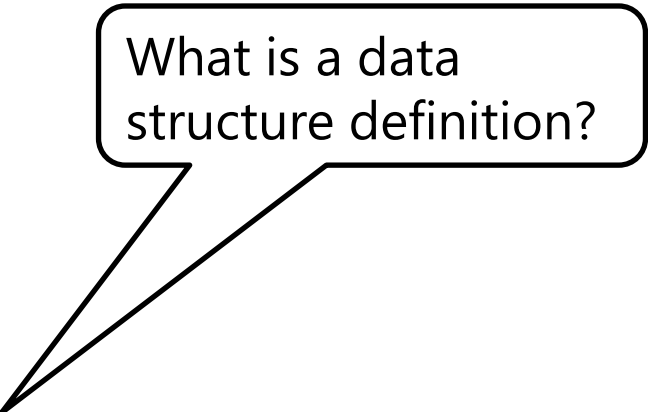
An attribute is a qualitative information about a specific statistical object and can be concepts such as units, magnitude, currency of denomination, titles, methodological comments, or quality statements.

Example 2: Basic RAG

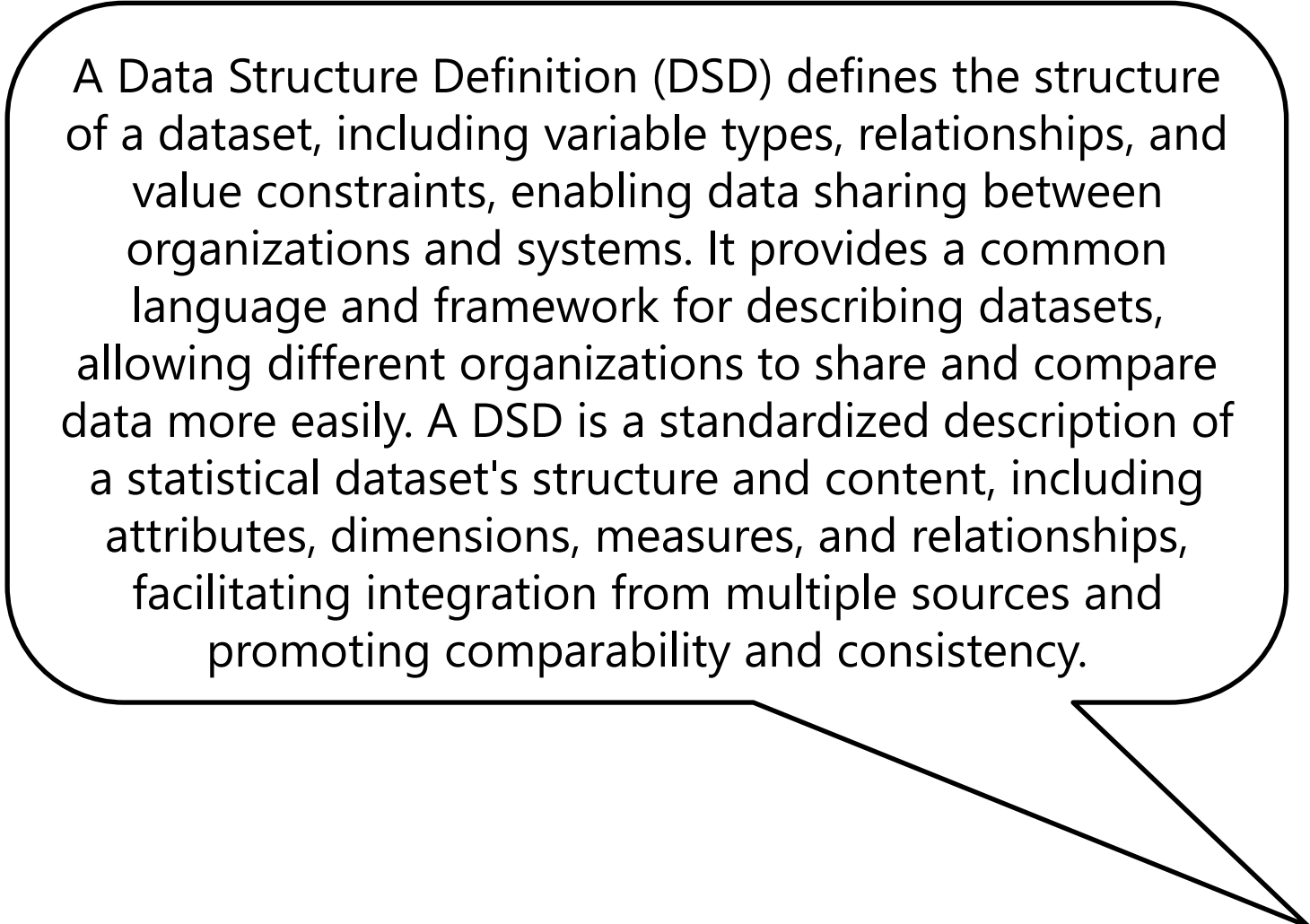
What is a data structure definition?

A Data Structure Definition (DSD) is a definition developed for the specific needs of one organisation only. **No, it's not that.** According to the text, a Data Structure Definition is a specification of statistical concepts necessary for identifying series and used as dimensions, attributes, and measures.

Example 2: Improved RAG



What is a data structure definition?



A Data Structure Definition (DSD) defines the structure of a dataset, including variable types, relationships, and value constraints, enabling data sharing between organizations and systems. It provides a common language and framework for describing datasets, allowing different organizations to share and compare data more easily. A DSD is a standardized description of a statistical dataset's structure and content, including attributes, dimensions, measures, and relationships, facilitating integration from multiple sources and promoting comparability and consistency.