

# Improved Analysis of Survey Data using Knowledge Graphs

Anna Gossen<sup>1,2</sup>, Dr. Eike Nicklas<sup>1,2</sup>

<sup>1</sup>Bank for International Settlements

<sup>2</sup>HMS Analytical Software GmbH

## Abstract

In this work, we present an internal knowledge-graph-based application for the storage and analysis of data regarding central banking practice on governance, management and organizational matters. We discuss a custom ontology as well as the high level application architecture and implementation challenges we experienced.

## Keywords

CEUR-WS, Knowledge Graphs, Ontologies

Knowledge Graphs are widely used in financial sector for various purposes, including fraud detection, banking oversight or customer profiling. Traditional database technologies often do not solve expanding analytical needs in the growing markets environment. By introducing semantically meaningful meta data and integrating instance data with the contextual structure, knowledge graphs offer a smart and efficient way to create, store, query, analyze data and convert it into direct value.

In the context of central banking, the analysis of governance and organizational structures and processes involves various challenges due to heterogeneous, but strongly inter-connected organizational structures and often qualitative or textual data. In addition, hierarchical metadata is commonly required for grouping or aggregation of analytical results.

An exemplary use case demonstrating these challenges is the following:

*The data analysts need to find central banks that have a supervisory board, where the governor is not chair, and a dedicated Monetary Policy Committee, where the governor is chair, and at least one member*

The task will require a lot of time, effort, resources, and table join operations when using a standard relational database. This exercise gets even more complicated when taking temporal aspects into account. In addition, the data required for this analysis is often not available in one database, but distributed across multiple systems or files.

In this work, we present an internal knowledge-graph-based application for the storage and analysis of data regarding central banking practice on governance, management and organizational matters. It supports the management and analysis of data on various topics and with

different data structures, some of which is collected via surveys. Its core is a custom Dataset ontology, which provides generic data structures to store information about given entities including revisions, temporal versions, provenance and data access information. The ontology reuses W3C standards and open linked vocabularies (SKOS<sup>1</sup>, PROV-O<sup>2</sup>, DCAT<sup>3</sup>), which makes it easy to understand and to apply. Figure 1 provides a high level overview.

Extensions to this ontology allow for the storage of additional metadata. For example, referenced entities and concepts can be structured in taxonomies for efficient data selection and aggregation, supported by standard ontologies such as the ORG ontology. In addition, related data can be grouped in datasets, e.g. to trace data that was collected in the same survey.

Supported by the use knowledge graphs technologies, the application offers:

- Generic, ontology-driven data analysis
- Advanced, inference-based full text search and contextual filtering
- Data provenance tracking
- Time-based data analysis
- Seamless datasets integration from heterogeneous sources
- Data quality validation

In this presentation, we describe the generic dataset ontology and a domain-specific extension, as well as the high-level application architecture. In addition, we will share experiences regarding implementation challenges and discuss how the application supports the data analysis work in the context of central banking organization and governance.

Published in the Workshop Proceedings of the EDBT/ICDT 2022 Joint Conference (March 29-April 1, 2022), Edinburgh, UK

✉ [anna.gossen@bis.org](mailto:anna.gossen@bis.org) (A. Gossen); [eike.nicklas@bis.org](mailto:eike.nicklas@bis.org) (Dr. E. Nicklas)



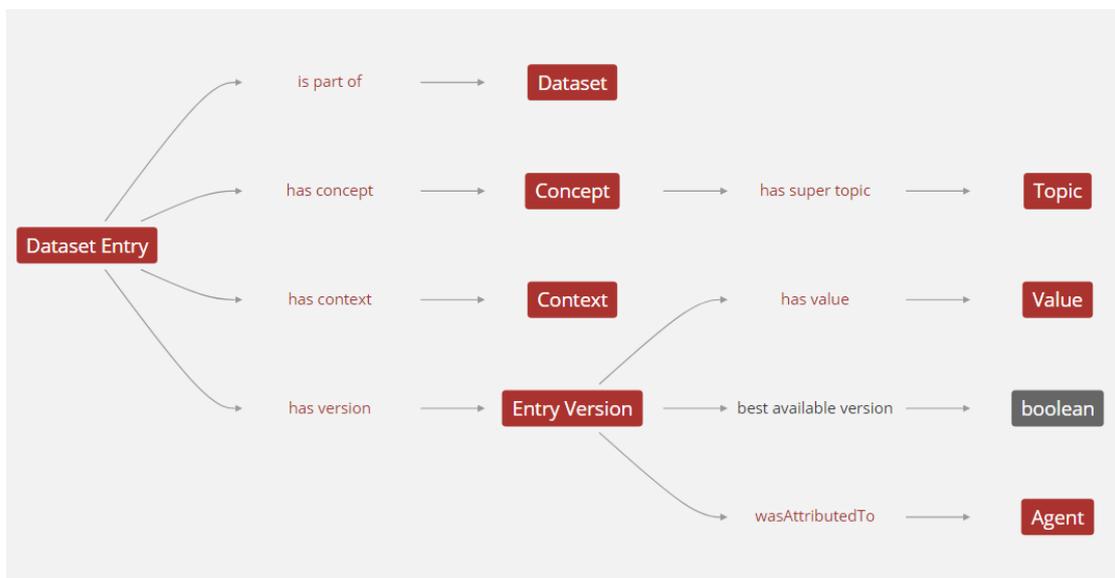
© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

<sup>1</sup><https://www.w3.org/TR/2009/REC-skos-reference-20090818/>

<sup>2</sup><https://www.w3.org/TR/prov-o/>

<sup>3</sup><https://www.w3.org/TR/vocab-dcat-2/>



**Figure 1:** Schematic overview of the main concepts in the dataset ontology.