Towards Cyber Mapping the German Financial System with Knowledge Graphs

Markus Schröder[1], Jacqueline Krüger[2], Neda Foroutan[1], Philipp Horn[2], Christoph Fricke[2], Ezgi Delikanli[2], Heiko Maus[1], and Andreas Dengel[1]

[1] Smart Data & Knowledge Services Dept., DFKI GmbH, Kaiserslautern, Germany
{markus.schroeder, neda.foroutan, heiko.maus, andreas.dengel}@dfki.de
[2] Deutsche Bundesbank, Frankfurt am Main, Germany
{jacqueline.krueger, philipp.horn, christoph.fricke, ezgi.delikanli}@bundesbank.de

Abstract. The increasing outsourcing of financial intermediaries intensifies the interconnection of the financial system with third-party providers. Accompanying concentration risks make cyber incidents at such providers a potential risk for financial stability. With the goal of preserving financial stability, regulators are interested in tracing cyber incidents efficiently. One method to achieve this is cyber mapping, which allows to analyze the connections between the financial network and the cyber network. In this paper, a provenance-aware knowledge graph is constructed to model such a mapping for investment funds which are part of the German financial system. As a first application, we provide a front-end for analyzing their outsourcing behaviors. In a user study with ten experts, we evaluate and show the application's usability and usefulness. Time estimations for certain scenarios indicate our application's potential to reduce the time needed to solve them. Especially for complex analysis tasks, it could provide benefits for cyber risk monitoring.

Keywords: Knowledge Graph Construction · RDF · Ontology · German Financial System · Cyber Mapping · Cyber Incidents · Fund Prospectus

1 Introduction

Financial intermediaries are increasingly outsourcing processes and services to third parties, which manifests in a raising interconnectedness of the financial system with entities outside this network [18]. The trend is further intensified by the increasing digitization, as information and communication technology become the core infrastructure for all financial processes [12]. Concentration risks may arise if outsourcing activities rely on only a few large service providers (e.g., cloud service providers) [21,4]. With the rise of third-party dependencies and increasing concentration risks, the question arises as to possible transmission channels between third-party service providers and financial intermediaries. Cyber incidents at both, financial intermediaries and along their supply chains, pose a risk to financial stability if they significantly impair the provision of key economic

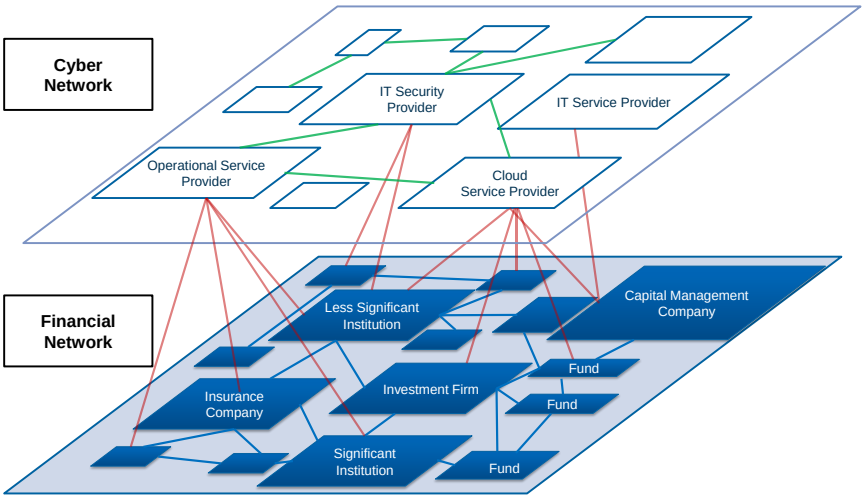

Fig. 1. Conceptual presentation of the cyber mapping methodology showing the cyber network (top) and financial network (bottom). Red arrows indicate the actual *mapping* between them. Source: Deutsche Bundesbank.

functions by the financial system [35,32]. Impacts of recent cyber attacks have already proven to be not restricted to the attacked company itself but of wider range [13] and are also able to induce turmoils at the financial market [24]. Thus, an understanding of the supply chains is inevitable in order to assess any risks to financial stability or intermediaries and address them properly.

Cyber mapping enables the network analysis of the financial system and its third-party service providers (see Figure 1). Mapping, in this context, refers to the connection of nodes from two distinct networks, in our case, from the financial network and the cyber network. Hereby, it allows for the identification of vulnerabilities related to outsourcing activities. Due to limited data availability [19], cyber mapping has been mostly a rather theoretically discussed concept so far [8]. However, recent regulatory changes [38,9] enlarge the data basis and enable a cyber mapping realization.

Cyber mapping has been recognized as essential in cyber risk monitoring of the financial sector and also provides benefits for banking supervision and financial market infrastructure oversight [2,19,34]. Its key feature is the immediate provision of information on potentially affected financial entities in case of a cyber incident and thus enables to determine necessary ad hoc measures. It also supports supervisors and overseers to quickly get an overview on outsourcing risks and third-party dependencies of selected financial entities. By identifying relevant third-party providers bearing concentration risks for the financial system, cyber mapping can also help to determine suitable longer-term regulatory measures or contribute to the development of the regulatory framework.

However, to the best of our knowledge, a cyber mapping of the German financial system with semantic technologies does not exist yet. This work takes first steps filling the gap. Our contribution is an outcome of the transfer lab "Cybermapping"[3] which was set up by Deutsche Bundesbank[4] (the central bank of the Federal Republic of Germany) and the German Research Center for Artificial Intelligence[5] (DFKI) to foster research on this topic. The paper reports on the development of a cyber mapping method by using Knowledge Graphs (KGs) [20] and the Resource Description Framework (RDF) [36]. Especially KGs enable us to model the networks' interconnections, their mappings and metadata by integrating various data silos. The graph is built using a dedicated cyber mapping ontology and linking it to structured financial data as well as extracted information from unstructured texts. Traceability is ensured by considering additional provenance information.

The remainder of the paper is organized as follows: Section 2 covers related work on building Knowledge Graphs in the cybersecurity and financial domain as well as relevant projects. Afterwards, our own construction approach is described in Section 3. Next, initial numbers on our KG and a first user application is provided (Section 4). By conducting a user study, discussed in Section 5, we present first results on our application's usefulness. Finally, Section 6 provides a conclusion and an outlook to further research.

2 Related Work

In literature, Knowledge Graphs (KGs) have been built for various domains (for a survey see [1]). In the field of Information and Communication Technology (ICT), there are efforts to build cybersecurity KGs. To achieve this, works discover relevant concepts and relationships in (un)structured sources with various methods such as Named Entity Recognition (NER) [29], Relation Extraction [33], word embeddings [14] and Extraction-Transformation-Loading (ETL) [30]. Similarly, there are approaches which aim to cover parts of the financial system with KGs. To construct these financial KGs, several sources are considered like annual financial reports [40], financial news articles [16], financial research reports [39] and data on equity [27].

The referenced papers give a comprehensive overview of a variety of KG construction approaches. Although cybersecurity and financial domains are covered, these works lack in the actual mapping between them which is a fundamental aspect of cyber mapping. We therefore have investigated related research projects on this topic, too.

The Financial Supervisory Authority of Norway (Finanstilsynet[6]) together with the Central Bank of Norway (Norges Bank[7]) drafted a first solution on

[3] `https://www.dfki.uni-kl.de/cybermapping`
[4] `https://www.bundesbank.de/en`
[5] `https://www.dfki.de/en/web`
[6] `https://www.finanstilsynet.no/en/`
[7] `https://www.norges-bank.no/en/`

financial sector mapping [28]. Their mapping approach is based on a poll among ministries to identify fundamental national functions, e.g. the execution of financial transactions. In a second step, the relevant organisations and critical service providers related to these functions are identified and mapped to the fundamental national functions. Thus, their top-down concept of a financial sector map differs from our data-driven approach. Moreover, there is no publicly available report on the applied technology or its implementation.

For Germany, the Bundesanstalt für Finanzdienstleistungsaufsicht[8] [German Federal Financial Supervisory Authority] (BaFin) engaged a research team of Innsbruck University on possible scenarios regarding the further development of the financial industry [12,11]. Based on these scenarios, the research team recommended to identify the relevant ICT service provider by mapping the financial system. As a consequence, BaFin produced a first map of the financial system [10], yet it is still restricted to structured data of one financial sector.

In conclusion, a cyber mapping approach with semantic technologies seems not to be available yet. Therefore, we developed an approach to construct a KG for cyber mapping the financial system.

3 Knowledge Graph Construction

Our Knowledge Graph (KG) uses the Resource Description Framework (RDF) to make formal statements about cyber mapping. Figure 2 presents an overview of our KG construction. With the help of a dedicated cyber mapping ontology (covered in Section 3.1), structured data about the German financial system in form of Excel, CSV and XML is mapped to RDF statements (Section 3.2). Similarly, unstructured data in fund prospectus PDFs is extracted using Natural Language Processing (NLP) techniques such as Named Entity Recognition (NER) and Relation Extraction (RE) which is described in Section 3.3. Either way, origins of created RDF statements are recorded in a special provenance box (Section 3.4).

3.1 Cyber Mapping Ontology

To express facts about cyber mapping, we are in need of an appropriate ontology [23]. In literature, some ontologies about the domain of finance and cybersecurity can be found. A prominent one is the Financial Industry Business Ontology[9] (FIBO) [7], which focuses on the business of finance. Another one is the Unified Cybersecurity Ontology (UCO) [37], covering the cybersecurity domain. However, none of them fully satisfy our requirements for cyber mapping. We therefore defined our own Cyber Mapping Ontology (CMO, prefixed `cmo`) which is still work in progress. It has been defined with the well-known Protégé ontology editor[10]

[8] `https://www.bafin.de/EN/`
[9] `https://spec.edmcouncil.org/fibo/`
[10] `https://protege.stanford.edu/`

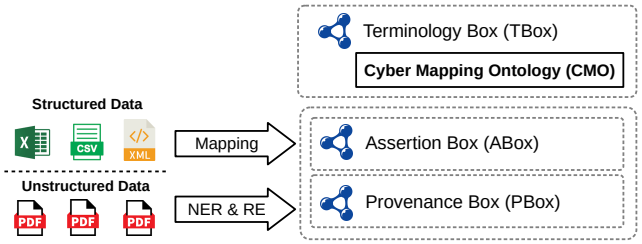

Fig. 2. Knowledge graph construction overview: a cyber mapping ontology (terminology box) provides necessary vocabulary to import (un)structured data into a KG (assertion box). The provenance box keeps track of each RDF statement's origin.

and is published with WIDOCO[11]. A first draft of the ontology's specification is available online[12] in English and German.

Since we are interested in mapping the German financial system, we consulted data provided by the BaFin. Its database about companies lists 44 company types, which we reviewed and generalized to a class tree with 12 nodes under a branch node `cmo:Company`. Similarly, its investment funds database helped us to distil 7 fund classes from 15 types with a generalizing `cmo:Fund` class. All classes in our ontology are generalized to an upper class named `cmo:Entity`.

During the examination of sources, we came across various ways to identify entities which is reflected in our ontology with individual properties. The BaFin uses internal references (`cmo:baFinRef`) and external IDs (`cmo:baFinId`), while the European Central Bank[13] (ECB) manages its own IDs (`cmo:ecbId`). Funds are usually identified by the International Security Identification Number[14] (`cmo:isin`), while companies are commonly recognized by the Legal Entity Identifier (`cmo:lei`). The Register of Institutions and Affiliates Database (RIAD) [17] proposes its own identifier (`cmo:riadId`).

To perform cyber mapping, we are interested in certain relationships between two companies. One such relation is the outsourcing of a service from one company to another (`cmo:outsourcesTo`). Since this typically involves a source entity (`cmo:source`), target entity (`cmo:target`) and the subject matter of the outsourcing (`cmo:Outsourcing`), we decided to model this declaration as an n-ary relation[15] (`cmo:OutsourcingStatement`). In order to keep track where this statement was made, we make use of PROV-O's `prov:wasDerivedFrom` property [6] and the NLP Interchange Format[16] (NIF) [26]. This way, we are able to record, for instance, in which sentence of a PDF file an outsourcing statement occurred.

[11] `https://github.com/dgarijo/Widoco`
[12] `https://www.dfki.uni-kl.de/cybermapping/ontology`
[13] `https://www.ecb.europa.eu/`
[14] `https://www.isin.org/`
[15] `https://www.w3.org/TR/swbp-n-aryRelations/`
[16] `https://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core`

Section 3.3 shows how these concepts are applied in practice, while more on the provenance topic is covered in Section 3.4.

To model addresses, concepts from DBpedia's ontology[17] and the ontology for vCard[18] are used: the `vcard:hasAddress` property with a `vcard:Address` blank node store `dbo:address` (street name and house number), `dbo:postalCode`, `dbo:city` and `dbo:country`. The property `dbo:subsidiary` models typical company structures.

With our defined ontology, we are able to map structured data about the financial system to our KG, which is described in the next section.

3.2 Structured Data: Financial Domain

The Deutsche Bundesbank and the BaFin legally collect regulatory data from German financial entities (e.g., banks). However, at the beginning of our transfer lab useful data for our project was either still in the process of being collected or subject to strict confidentiality. We therefore decided to initially use public data for building the KG.

Table 1 lists six publicly available data sources about financial intermediaries and funds provided by the BaFin, the ECB and the European Securities and Markets Authority[19] (ESMA). It provides an overview about the number of records (#Rec.), columns (#Col.) and types (#Typ.) as well as the source's data format (XML, Excel or CSV). While entities are always named (mapped to `skos:prefLabel`), different sets of identifiers are provided by each data source which is stated with a check mark (✓). In some cases, data is rather incomplete, which is indicated in the table by a tilde sign (∼). Regarding the *Address* column, this means that only the country is mentioned (mapped to `dbo:country`). The selection of additional properties also greatly varies per dataset.

With the CMO (Section 3.1) and information from Table 1, structured data can be lifted to RDF with an appropriate technique, for instance, KG generation with the RDF Mapping Language[20] (RML) [15]. To give an illustrative example of a resource in our graph, Listing 1.1 depicts in Turtle syntax [5] a fictional German stock company with type, label, identifier, address and subsidiary information. The resource is identified with a Universally Unique Identifier (UUID) in our cyber mapping resource namespace (prefixed `cmr`). Since such entities can be named differently (e.g. abbreviations), `skos:altLabel` records alternative labels.

While the discussed data lists the majority of entities in the German financial system, it lacks in stating specific relationships for the purpose of performing cyber mapping. Helpfully, such information can be found, at least partly, in unstructured texts which is covered in the following section.

[17] `https://dbpedia.org/ontology`
[18] `https://www.w3.org/2006/vcard/ns#`
[19] `https://www.esma.europa.eu/`
[20] `https://rml.io/`

Table 1. Six publicly available data sources about companies and funds in the financial system. For each entry the number of records (#Rec.), columns (#Col.) and types is given (#Typ.) as well as its format. A check mark (✓) states that this property can be found in the dataset, while tilde (∼) indicates incomplete data. The last column lists additional properties available in the dataset, for instance, some funds refer to their Capital Management Company (CMC).

Data Source	#Rec.	#Col.	#Typ.	Format	`skos:prefLabel`	`baFinId`	`ecbId`	`lei`	`isin`	`riadId`	Address	Additional Properties
BaFin Company Database[a]	7,151	8	44	XML	✓						✓	arbitration board
BaFin Investment Funds Database[b]	14,427	10	15	XML	✓	✓			∼		∼	structure; name ref. to CMC
ECB Supervised Entities[c]	900	5	4	Excel	✓			✓			∼	subsidiary; grounds for significance
ECB Investment Funds[d]	78,932	19	3	Excel	✓		✓	✓	✓		✓	capital variability; investment policy; net asset value size class; ref. to CMC
ECB Monetary Financial Institutions[e]	5,720	14	4	CSV	✓				✓	✓	✓	country of registration; metadata about head
ESMA Money Market Funds[f]	472	15	4	CSV	✓			✓			∼	legal framework; ref. to CMC

[a] `https://portal.mvp.bafin.de/database/InstInfo/?locale=en_US`
[b] `https://portal.mvp.bafin.de/database/FondsInfo/?locale=en_US`
[c] `https://www.bankingsupervision.europa.eu/banking/list/html/index.en.html`
[d] `https://www.ecb.europa.eu/stats/financial_corporations/list_of_financial_institutions/html/index.en.html#if`
[e] `https://www.ecb.europa.eu/stats/financial_corporations/list_of_financial_institutions/html/daily_list-MID.en.html`
[f] `https://registers.esma.europa.eu/publication/searchRegister?core=esma_registers_mmf04`

Listing 1.1. A fictional example of a mapped resource in the Cyber Mapping KG expressed in Turtle syntax.

```
cmr:bc57a47d-d990-486f-9b7f-4af78aded30a
  rdf:type          cmo:SignificantInstitution ;
  skos:prefLabel    "Mercurtainment Bank Aktiengesellschaft" ;
  skos:altLabel     "Mercurtainment Bank AG" ;
  cmo:baFinRef      "303846" ;
  cmo:lei           "G9QIEQ1BITM5RF3YCDRQ" ;
  cmo:riadId        "DE70255" ;
  vcard:hasAddress [
    dbo:address     "Musterstr. 42" ;
    dbo:postalCode "60312" ;
    dbo:city        dbr:Frankfurt_am_Main ;
    dbo:country     dbr:Germany
  ] ;
  dbo:subsidiary cmr:c613bf97-07da-46cd-ab5c-eba0454679a9 .
```

3.3 Unstructured Data: Fund Prospectus

The German Capital Investment Code (Kapitalanlagegesetzbuch, KAGB) governs in § 164[21] the creation of sales prospectuses, or more specifically fund prospectuses. Capital Management Companies (CMCs) are obligated to describe which activities are outsourced to specific companies (see § 165, Sec. 2, No. 33 KAGB). By interpreting these texts, we are able to map funds to companies with our `cmo:outsourcesTo` property.

In prior work [22], we compiled a corpus of 1,054 fund prospectuses (PDFs). From these documents, 948 extracted sentences were manually annotated with 5,969 named entity annotations and 2,573 Outsourcing–Company relationship annotations. The resulting German dataset on Company Outsourcing in Fund prospectuses (CO-Fun) is used in this paper to acquire structured RDF statements. An example of such a construct is presented in Listing 1.2. While the `cmo:outsourcesTo` property simply relates a fund to a company, the corresponding statement (typed `cmo:OutsourcingStatement`) additionally states the outsourcing and provenance information (`prov:wasDerivedFrom`). This way, we are able to reconstruct the sentence in a certain fund prospectus that leads to an outsourcing statement. Regarding the prospectus itself, we can state its CMC (`cmo:managementCompany`) and the fund (`cmo:fund`) it is about.

Having such additional information is very useful to trace the origins of statements. We therefore considered provenance for all statements in the ABox which is discussed next.

3.4 Provenance Information

A special feature in our cyber mapping KG is the existence of a Provenance Box (PBox in analogy to TBox and ABox, see Figure 2). Its purpose is the storage

[21] `https://www.gesetze-im-internet.de/kagb/__164.html`

Listing 1.2. Illustration of an outsourcing statement lifted from a sentence in a fund prospectus. For readability, some UUIDs are shortened and some literals are formatted.

```
cmr:1c956834
  cmo:outsourcesTo cmr:2d56a950 .

cmr:8ea294fd-9f0e-4158-a2e3-c14f93c2b4b2
  rdf:type               cmo:OutsourcingStatement ;
  cmo:source             cmr:1c956834 ;
  cmo:target             cmr:2d56a950 ;
  cmo:outsourcing        cmr:DataCenterService ;
  prov:wasDerivedFrom    cmr:s96048cb .

cmr:s96048cb
  rdf:type               nif:String, nif:Sentence ;
  nif:anchorOf           "The company has outsourced data center
                         services to Mercurtainment & CO KGaA." ;
  nif:referenceContext   <file://fund-prospectus.pdf> .

<file://fund-prospectus.pdf>
  rdf:type               nfo:FileDataObject ;
  nfo:fileName           "fund-prospectus.pdf" ;
  dct:hasPart            cmr:s96048cb , cmr:s5d53b88 , cmr:sd4219e7 ;
  cmo:managementCompany  cmr:cmcf78d3 ;
  cmo:fund               cmr:1c956834 .
```

of additional statements for *every* statement asserted in the ABox to enable a comprehensive traceability. To implement this, we make use of RDF-star[22] [25], which allows us to annotate statements in RDF with metadata. Listing 1.3 demonstrates an example how the asserted `skos:altLabel` statement from Listing 1.1 is annotated. Using RDF-star, the triple is quoted (`<<...>>`) on subject position. With the provenance ontology (PROV-O) [6], several aspects about the `skos:altLabel`-statement are recorded: one is the agent who is responsible for creating the statement by using the `prov:wasAttributedTo` property. Usually, this involves a certain activity, for instance, an importing procedure or interface usage, which is stated with a `prov:wasGeneratedBy` statement. To note the origin of the quoted statement, a `prov:wasDerivedFrom` property refers to the data source (e.g., a CSV file). Further RDF statements are made about agent, activity and source to provide useful metadata about them. Dublin Core's[23] `dct:date` attribute is used to be able to reconstruct a chronological order.

By performing all steps discussed in Section 3, our construction approach yields to a first version of a cyber mapping KG. In the next section, we discuss the graph's content and a first utilization of it.

[22] `https://www.w3.org/2021/12/rdf-star.html`
[23] `http://purl.org/dc/terms/`

Listing 1.3. An example how provenance information is annotated with RDF-star and PROV-O. URIs and UUIDs are shortened for readability.

```
<< cmr:bc57a47d skos:altLabel "Mercurtainment Bank AG" >>
  rdf:type              prov:Entity ;
  prov:wasAttributedTo  <https://.../agent/smith> ;
  prov:wasGeneratedBy   <https://.../activity/1868ccf> ;
  prov:wasDerivedFrom   <https://.../entity/some.csv> ;
  dct:date              "2023-07-19T14:32:54.812Z"^^xsd:dateTime .
```

4 Knowledge Graph Application

Graph. An initial version of our cyber mapping KG contains 1,725,383 RDF statements about 108,030 entities, including 93,253 funds and 14,777 companies in the financial system. The latter are separated into 8,184 (financial) institutions, 5,307 capital management companies and 1,286 insurance companies. However, duplicates likely still exist, particularly in the case of `cmo:Fund` instances, as these were imported from several independently managed data sources (as shown in Section 3.2).

Regarding documents, metadata about 917 fund prospectuses with 686 sentences are available, as shown in Listing 1.2. We acquired 7,239 outsourcing statements, which refer to 375 outsourcing entities. 4,033 `cmo:outsourcesTo` relationships between funds and companies could be identified.

Our Provenance Box (PBox) graph contains 10,338,786 triples attributed to two users, one of whom was responsible for invoking the main data import steps. Statements were generated by 17 activities, primarily RDF mapping procedures and a few user interactions via interfaces. Resulting triples were derived from 15 sources: besides the main data sources (see Table 1), this also includes auxiliary data and users. The total number of provenance triples is higher than expected due to multiple iterations of import steps and kept provenance information.

Gathering such an RDF dataset about parts of the German financial system naturally raises concerns of potential misuse. We therefore abstain from making our KG publicly available and only provide fictional examples in this paper.

Application. Having such a KG successfully constructed enables us to provide useful applications for end users such as overseers and supervisors. One use case is the analysis of outsourcing relations in the context of cyber risks. With a focus on outsourcing companies, we would like to allow users to perform data exploration using an initial application, which is presented in Figure 3. The application provides several features (F) for inspection and filtering the KG.

Regarding inspection, a graph visualization (F1) derived from the KG is presented in the center view showing funds and companies as connected nodes. Larger red nodes suggest a high number of incoming edges indicating the third-party providers where funds mostly outsource their services to. Below, a table (F2) lists outsourcing relationships together with the relevant text passage from the linked fund prospectus (F3). In the top right corner, a table shows for the

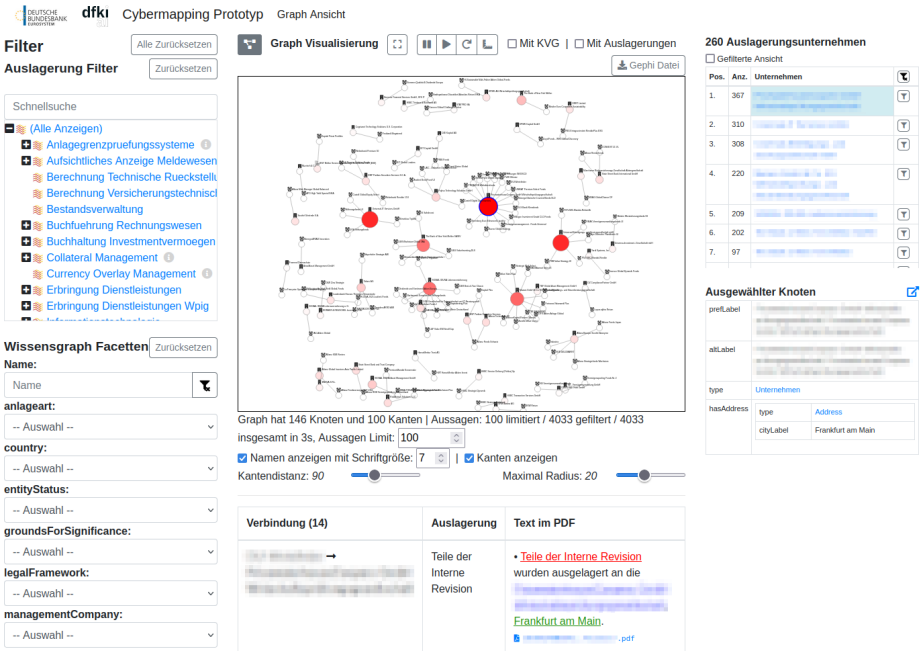

Fig. 3. Web application (in German) to interactively analyze outsourcing relationships between funds and companies. Left: Graph filter options; Center: Graph visualization and connection list; Right: Top outsourcing companies and selected node.

current view outsourcing companies ordered by their incoming edges. Additional properties about a selected node (F4) are presented on the bottom right.

In case of filters (left), a taxonomy of outsourcing categories (top left, F5) lets users restrict the graph to a certain outsourcing type such as Information Technology (IT). Outsourcing categories have expandable definitions attached for a better understanding (F6). Properties in our KG (bottom left, F7) can be used to further filter nodes, for instance, by name (`skos:prefLabel`), location (`dbo:country`) or other metadata (e.g., `cmo:groundsForSignificance`).

In the next section, a study is presented where the application's usefulness for potential users is evaluated.

5 User Study

A user study was conducted in order to evaluate our application regarding its perceived user experience, features and potential time saving for given scenarios. Its setup is described in Section 5.1, followed by a description and interpretation of the results in Section 5.2.

5.1 Setup

We conducted a study with ten selected experts (E1–E10, 9 male, 1 female) from three different departments of the Deutsche Bundesbank with a work experience of at least three years. In particular, the experts' work experience is on average 16.1 years (s.d. 8.8, min. 4, max. 33). The distribution among age groups is almost balanced starting from 25 to older than 55 years (ten year spanned).

Sessions were conducted in one-on-one interviews or small group meetings with up to three experts. Each expert was provided with individual access to the application described in Section 4. At the beginning of each session, a short introduction was provided by a conductor (an author of this paper) which took about 15 minutes. The introduction consisted of the provision of basic information, such as the elaboration of the cyber mapping concept and its data basis. After that, the conductor provided a practical orientation in the application by presenting its key features. In a subsequent testing phase, the experts were asked to use our application for at least ten minutes in order to familiarize themselves with the application and to explore its features. Questions could be asked anytime, followed by further clarifications provided by the conductor. In the end, experts spent on average 20.2 minutes testing our application (s.d. 4.35, min. 14, max. 25).

After the testing phase, the experts were provided with a structured questionnaire, consisting of four parts. In the first part, 26 questions were asked by employing a standardized User Experience Questionnaire (UEQ) [31] to measure the experts' experience regarding the following six factors: attractiveness, perspicuity, efficiency, dependability, stimulation and novelty. In the second part, questions were asked with respect to the perceived usefulness of the seven features of our application (F1–F7) using a 7-point Likert scale from useless (1) to useful (7). Furthermore, feedback from the experts was collected with regard to further aspired features the current application should be extended with.

Last, participants had to estimate the time needed to fulfil tasks in given scenarios (S1–S3) with and without our application. For the definition of the scenarios, we considered essential cyber risk monitoring features of cyber mapping. The scenarios were formulated as follows:

Affected Intermediaries (S1) "You learn of a cyber incident at an IT company that is no longer able to perform its tasks. Suppose you would like to find out which financial intermediaries relevant to you are potentially affected."

Outsourcing Relations (S2) "Suppose you would like to know to whom a specific financial intermediary has outsourced its processes to."

Outsourcing Relevance (S3) "Suppose you would like to identify the top ten outsourcing companies for the financial system regarding outsourcing of accounting services."

Notably, the identification of possible transmission channels between a third-party provider and financial intermediaries in case of a cyber incident were included in S1. S2 covers exemplary the understanding of the supply chain of a financial intermediary. S3 represents a first step in the analysis of potential concentration

Table 2. Questionnaire results stating for each Expert (E) their feature ratings and estimated times. F1 to F7 encompass the questions posed regarding the features' perceived usefulness. S1 to S3 cover the supposed times needed in minutes for solving the three scenarios with (w) and without (w/o) our application ('n' denotes unsolvable; minutes rounded for better readability). Difference (diff) between the estimations are provided. Below: mean and standard deviation (s.d.) values.

E	F1	F2	F3	F4	F5	F6	F7	S1			S2			S3		
								w	w/o	diff	w	w/o	diff	w	w/o	diff
E1	6	4	6	5	6	7	7	5	5	0	5	5	0	5	n	-
E2	5	6	6	5	7	7	5	5	120	115	5	60	55	5	240	235
E3	7	7	7	3	7	7	5	8	60	53	8	30	23	8	n	-
E4	7	7	7	5	5	6	5	10	n	-	5	n	-	15	n	-
E5	7	4	7	7	7	7	5	60	n	-	1	60	59	15	n	-
E6	5	7	7	6	6	7	7	5	120	115	1	60	59	10	180	170
E7	7	7	4	6	4	6	5	5	n	-	5	n	-	5	n	-
E8	7	7	7	7	6	6	5	1	1	0	1	n	-	60	n	-
E9	7	7	6	6	7	5	5	15	120	105	10	30	20	15	n	-
E10	7	6	7	5	7	7	7	1	60	59	0	15	15	45	n	-
mean	6.5	6.2	6.4	5.5	6.2	6.5	5.6	5.6	69.4	63.9	4.1	37.2	33	7.9	210	202.1
s.d.	0.8	1.2	1	1.2	1	0.7	1	4.4	52.7	50.4	3.2	23.1	24.3	5.6	42.4	46

risks for a defined scope. Since our application was tested in different departments, the experts were asked to assume the application would already include their relevant data sources.

5.2 Results

Regarding UEQ results (see Figure 4), the scaled results show that all factors were rated higher than the values of the benchmark data set based on 468 studies [3]. The ratings show highest levels of attractiveness and perspicuity close to "excellent" results, while stimulation and efficiency are located in the "good" area, followed by dependability and novelty being "above average". A rather low value for the ergonomic quality aspect dependability (i.e. predictable, secure) might reveal that further explanations of the way the application functions and simplifications towards a more intuitive interface could be helpful. The quality aspect novelty (i.e. innovative, creative) might be low because the application's front-end mainly comprises data visualization and filtering options, which is rather standard and thus expected by the experts. Still, overall results show us that our initial application let participants have a good user experience with potential for improvements.

The results regarding the experts' ratings of the usefulness of the seven specific features (F1–F7) and the scenario results are depicted in Table 2. With an overall mean value of 6.1 close to the max. value of 7 (useful), we derive that the features of our application are well received. The definition of outsourcing categories (F6)

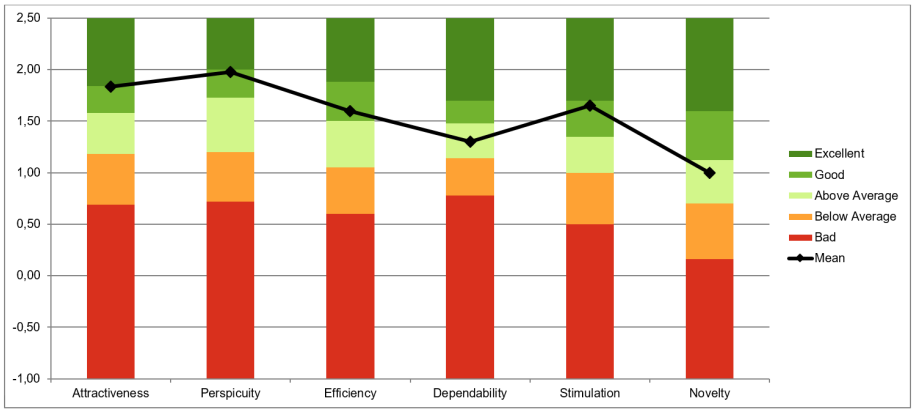

Fig. 4. Mean and distribution of the six factors from the User Experience Questionnaire (UEQ) [31] derived from all expert answers to the 26 questions.

and the visualization of the graph (F1) are rated most useful (mean 6.5), followed closely by the linkage of fund prospectuses (F3 with 6.4). The filter option for outsourcing categories (F5) and the display of the relevant text passages from fund prospectuses (F2) receive both on average 6.2. Lowest values are gained by the filter options for the knowledge graph facets (F7 with 5.6) and the selected company's information (F4 with 5.5). Interpreting the results, visualizations, explanations and links to further information are perceived as rather useful. However, in its current state, the KG's properties for filtering and inspection provide room for improvement.

Our application is perceived to save time in solving the given scenarios (S1–S3). Using our application, participants estimate to complete them on average in under 10 minutes. Conversely, without our application, experts report a completion time of about 30–60 minutes for the same tasks. Considering S1 and S2, the time saving is approx. by a factor of 10. Looking at absolute numbers, the supposed time saving ranges on average from around 30 minutes (S2) and one hour (S1) up to three hours (S3). Especially for S3, which encompasses a broader cross-sector scope, eight out of ten experts stated not to be in a position to solve this task without our cyber mapping application (indicated with 'n'). The remaining experts estimated the time needed to solve this task via a manual workaround taking 3–4 hours.

The study also collected feedback to aspired features in our application. Mostly, experts asked for new features regarding the integration of more data sources and functions. In particular, capabilities such as full-text search, auto completion and fine-grained filter options were suggested. Since our application currently covers only a part of the financial system, participants recommend to extend the KG with further data relevant to their jobs.

In conclusion, results in our user study show that our initial application provides a good user experience, notably regarding perspicuity and attractiveness.

Features about visualization, definitions and references were perceived as most useful, yet further improvements in the KG's content and filter operations need to be implemented. Time estimations indicate that our application has potential to reduce the time needed to investigate on cyber incidents in the financial system. Especially for complex analysis tasks (like S3), our application could provide benefits for cyber risk monitoring.

6 Conclusion and Outlook

After motivating the importance of cyber mapping for ensuring financial stability, we presented a first approach towards this goal by utilizing Knowledge Graphs (KG). Although there were some endeavors in the past, no approach applying semantic technologies has been published so far. Therefore, a KG construction approach was proposed which consists of a dedicated cyber mapping ontology, the integration of (un)structured knowledge and its traceability. Having such a KG at hand, we implemented an application to let users analyze outsourcing relations of funds and third-party providers. A user study with ten experts was conducted to collect feedback about the usability and usefulness and to estimate possible time saving potential of our application. With room for improvements, results have indicated a good user experience and the features' usefulness. Especially, time estimations indicate that our application has potential to reduce efforts. In case of complex tasks, our cyber mapping solution could provide benefits for cyber risk monitoring. With this work, we provide a first step towards cyber mapping the German financial system with knowledge graphs.

Our collaborative research lab is still running. For this time, we aspire to use the feedback acquired from the user study to improve our KG and provide further applications. In this context, results showed us that including further specific data sources would be useful. For a comprehensive cyber mapping, incorporating regulatory data sources would be essential. Moreover, our ambition is to improve our data integration pipeline: on the one hand, we would like to keep the data up-to-date, for instance, by introducing virtual knowledge graphs. On the other hand, we intend to process unstructured financial data with state-of-the-art technology in the field of neural networks and large language models.

References

1. Abu-Salih, B.: Domain-specific knowledge graphs: A survey. J. Netw. Comput. Appl. **185**, 103076 (2021). https://doi.org/10.1016/J.JNCA.2021.103076
2. Adelmann, F., Elliott, J.A., Ergen, I., Gaidosch, T., Jenkinson, N., Khiaonarong, T., Morozova, A., Schwarz, N., Wilson, C.J.: Cyber risk and financial stability : It's a small world after all. IMF Staff Discussion Notes (2020), https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2020/12/04/Cyber-Risk-and-Financial-Stability-Its-a-Small-World-After-All-48622

3. Andreas Hinderks and Martin Schrepp and Jörg Thomaschewski: User Experience Questionnaire, Data Analysis Tools. Website (2023), https://www.ueq-online.org/
4. Bank of England: Operational resilience: Critical third parties to the UK financial sector. PRA Discussion Paper 3/22 and FCA Discussion Paper 22/3 (2022), https://www.bankofengland.co.uk/prudential-regulation/publication/2022/july/operational-resilience-critical-third-parties-uk-financial-sector
5. Beckett, D., Berners-Lee, T., Prud'hommeaux, E., Carothers, G.: RDF 1.1 Turtle (2014), https://www.w3.org/TR/turtle/
6. Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S., Zhao, J.: PROV-O: The PROV Ontology (2012), http://www.w3.org/TR/prov-o/
7. Bennett, M.: The financial industry business ontology: Best practice for big data. Journal of Banking Regulation **14**(3), 255–268 (Jul 2013). https://doi.org/10.1057/jbr.2013.13
8. Brauchle, J.P., Göbel, M., Seiler, J., von Busekist, C.: Cyber mapping the financial system. Tech. rep., Carnegie Endowment for International Peace (2020), http://www.jstor.org/stable/resrep24291
9. Bundesamt für Justiz [German Federal Office of Justice]: Gesetz zur Stärkung der Finanzmarktintegrität (Finanzmarktintegritätsstärkungsgesetz – FISG) [Act to Strengthen Financial Market Integrity (Financial Market Integrity Strengthening Act)] (2021), https://www.bgbl.de/xaver/bgbl/start.xav#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl121s1534.pdf%27%5D__1699873137312
10. Bundesanstalt für Finanzdienstleitsungsaufsicht [German Federal Financial Supervisory Authority] (BaFin): Auslagerungen: Landkarten bieten Orientierung [Outsourcing: Maps Provide Orientation] (2022), https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Fachartikel/2022/fa_bj_2208_Auslagerungen_Landkarten.html
11. Bundesanstalt für Finanzdienstleitsungsaufsicht [German Federal Financial Supervisory Authority] (BaFin): Wertschöpfungsketten im Finanzsektor: Empfehlungen zur IT-Aufsichtspraxis [Value chains in the financial sector: recommendations for IT supervisory practice] (2022), https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Fachartikel/2022/fa_bj_2207_uni_innsbruck_wertschoepfungsketten.html
12. Böhme, R., Pesch, P.J., Fritz, V.: Auswirkungen sich verändernder Wertschöpfungsketten im Finanzsektor auf die IT-Sicherheit [Effects of changing value chains in the financial sector on IT security] (2022), https://www.bafin.de/SharedDocs/Downloads/DE/Bericht/dl_abschlussbericht_forschungsprojekt_uni_innsbruck.pdf?__blob=publicationFile
13. Clancy, L., Mourselas, C.: Ion cyber outage continues as banks rely on workarounds (2023), https://www.risk.net/derivatives/7955967/ion-cyber-outage-continues-as-banks-rely-on-workarounds
14. Deng, Y., Lu, D., Huang, D., Chung, C., Lin, F.: Knowledge graph based learning guidance for cybersecurity hands-on labs. In: Proceedings of the ACM Conference on Global Computing Education, CompEd 2019, Chengdu,Sichuan, China, May 17-19, 2019. pp. 194–200. ACM (2019). https://doi.org/10.1145/3300115.3309531
15. Dimou, A., Sande, M.V., Colpaert, P., Verborgh, R., Mannens, E., de Walle, R.V.: RML: A generic language for integrated RDF mappings of heterogeneous data. In: Proceedings of the Workshop on Linked Data on the Web co-located with the 23rd International World Wide Web Conference (WWW 2014), Seoul, Korea,

April 8, 2014. CEUR Workshop Proceedings, vol. 1184. CEUR-WS.org (2014), http://ceur-ws.org/Vol-1184/ldow2014_paper_01.pdf
16. Elhammadi, S., Lakshmanan, L.V.S., Ng, R.T., Simpson, M., Huai, B., Wang, Z., Wang, L.: A high precision pipeline for financial knowledge graph construction. In: Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020. pp. 967–977. International Committee on Computational Linguistics (2020). https://doi.org/10.18653/V1/2020.COLING-MAIN.84
17. European Central Bank (ECB): Guideline (EU) 2018/876 of the European Central Bank of 1 June 2018 on the Register of Institutions and Affiliates Data (ECB/2018/16). Official Journal of the European Union pp. 3–21 (06 2018), https://eur-lex.europa.eu/eli/guideline/2018/876
18. European Systemic Risk Board (ESRB): Systemic cyber risk. Tech. rep., European System of Financial Supervision (ESFS) (2020), https://www.esrb.europa.eu/pub/pdf/reports/esrb.report200219_systemiccyberrisk~101a09685e.en.pdf
19. European Systemic Risk Board (ESRB): Mitigating systemic cyber risk. Tech. rep., European System of Financial Supervision (ESFS) (2022), https://www.esrb.europa.eu/pub/pdf/reports/esrb.SystemiCyberRisk.220127~b6655fa027.en.pdf
20. Fensel, D., Simsek, U., Angele, K., Huaman, E., Kärle, E., Panasiuk, O., Toma, I., Umbrich, J., Wahler, A.: Knowledge Graphs - Methodology, Tools and Selected Use Cases. Springer (2020). https://doi.org/10.1007/978-3-030-37439-6
21. Financial Stability Board (FSB): Third-party dependencies in cloud services - considerations on financial stability implications. Tech. rep., FSB (2019), https://www.fsb.org/wp-content/uploads/P091219-2.pdf
22. Foroutan, N., Schröder, M., Dengel, A.: CO-Fun: A german dataset on company outsourcing in fund prospectuses for named entity recognition and relation extraction. In: The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING) (2023), https://www.dfki.uni-kl.de/~foroutan/9bf47629/CO_Fun.pdf, currently under review
23. Gruber, T.: A translation approach to portable ontology specifications. Knowledge acquisition **5**(2), 199–220 (1993)
24. Handelsblatt: Nach Cyberangriff: Evotec verlässt MDax wegen Fristverletzung [After cyber attack: Evotec leaves MDax due to deadline violation] (2023), https://www.handelsblatt.com/finanzen/maerkte/aktien/chart-des-tages-nach-cyberangriff-evotec-verlaesst-mdax-wegen-fristverletzung/29133970.html
25. Hartig, O.: Foundations of rdf⋆ and sparql⋆ (an alternative approach to statement-level metadata in RDF). In: Proceedings of the 11th Alberto Mendelzon International Workshop on Foundations of Data Management and the Web, Montevideo, Uruguay, June 7-9, 2017. CEUR Workshop Proceedings, vol. 1912. CEUR-WS.org (2017), https://ceur-ws.org/Vol-1912/paper12.pdf
26. Hellmann, S., Lehmann, J., Auer, S., Brümmer, M.: Integrating NLP using linked data. In: The Semantic Web - ISWC 2013 - 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II. Lecture Notes in Computer Science, vol. 8219, pp. 98–113. Springer (2013). https://doi.org/10.1007/978-3-642-41338-4_7
27. Huakui, L., Liang, H., Feicheng, M.: Constructing knowledge graph for financial equities. Data Analysis and Knowledge Discovery **4**(5), 27–37 (2020)

28. International Monetary Fund (IMF): Norway: Financial Sector Assessment Program. Technical Note - Cybersecurity Risk Supervision and Oversight. IMF Staff Country Report 2020/262. Tech. rep., IMF (2020), https://www.imf.org/~/media/Files/Publications/CR/2020/English/1NOREA2020004.ash
29. Jia, Y., Qi, Y., Shang, H., Jiang, R., Li, A.: A practical approach to constructing a knowledge graph for cybersecurity. Engineering **4**(1), 53–60 (2018). https://doi.org/https://doi.org/10.1016/j.eng.2018.01.004, cybersecurity
30. Kiesling, E., Ekelhart, A., Kurniawan, K., Ekaputra, F.J.: The SEPSES knowledge graph: An integrated resource for cybersecurity. In: The Semantic Web - ISWC 2019 - 18th International Semantic Web Conference, Auckland, New Zealand, October 26-30, 2019, Proceedings, Part II. Lecture Notes in Computer Science, vol. 11779, pp. 198–214. Springer (2019). https://doi.org/10.1007/978-3-030-30796-7_13
31. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: HCI and Usability for Education and Work, 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings. Lecture Notes in Computer Science, vol. 5298, pp. 63–76. Springer (2008). https://doi.org/10.1007/978-3-540-89350-9_6
32. Panetta, F.: The quick and the dead: building up cyber resilience in the financial sector. Tech. rep., European Central Bank (ECB) (2023), https://www.ecb.europa.eu/press/key/date/2023/html/ecb.sp230308~92211cd1f5.en.html
33. Pingle, A., Piplai, A., Mittal, S., Joshi, A., Holt, J., Zak, R.: Relext: relation extraction using deep learning approaches for cybersecurity knowledge graph improvement. In: ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining, Vancouver, British Columbia, Canada, 27-30 August, 2019. pp. 879–886. ACM (2019). https://doi.org/10.1145/3341161.3343519
34. Resano, J.R.M.: Digital resilience and financial stability. the quest for policy tools in the financial sector. Tech. rep., Banco de España (2022), https://dx.doi.org/10.2139/ssrn.4336381
35. Ros, G.: The making of a cyber crash: A conceptual model for systemic risk in the financial sector. ESRB: Occasional Paper Series No. 2020/16 (2020), https://dx.doi.org/10.2139/ssrn.3723346
36. Schreiber, G., Raimond, Y.: RDF 1.1 Primer (2014), https://www.w3.org/TR/rdf11-primer/
37. Syed, Z., Padia, A., Finin, T., Mathews, M.L., Joshi, A.: UCO: A unified cybersecurity ontology. In: Martinez, D.R., Streilein, W.W., Carter, K.M., Sinha, A. (eds.) Artificial Intelligence for Cyber Security, Papers from the 2016 AAAI Workshop, Phoenix, Arizona, USA, February 12, 2016. AAAI Technical Report, vol. WS-16-03. AAAI Press (2016), http://www.aaai.org/ocs/index.php/WS/AAAIW16/paper/view/12574
38. The European Parliament and the Council of the European Union: Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 (2022), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2554&from=FR
39. Wang, W., Xu, Y., Du, C., Chen, Y., Wang, Y., Wen, H.: Data set and evaluation of automated construction of financial knowledge graph. Data Intell. **3**(3), 418–443 (2021). https://doi.org/10.1162/DINT_A_00108
40. Zehra, S., Mohsin, S.F.M., Wasi, S., Jami, S.I., Siddiqui, M.S., Raazi, S.M.K.: Financial knowledge graph based financial report query system. IEEE Access **9**, 69766–69782 (2021). https://doi.org/10.1109/ACCESS.2021.3077916