Scholarly Artifacts Knowledge Graph: Use Cases for Digital Libraries

Fidan Limani[(✉)] , Atif Latif , and Klaus Tochtermann

Leibniz Information Centre for Economics, Düsternbrooker Weg 120, 24105 Kiel, Germany
{f.limani,a.latif,k.tochtermann}@zbw.eu
https://www.zbw.eu/en/research/science-2-0/fidan-limani

Abstract. Scholarly communication process is constantly evolving to include new artifacts that present a part of the research. The traditional practice of conference and journal publications being the key artifacts to publish have changed to include new or complementary artifacts that help the audience grasp the research outcomes as much as possible as a single unit or whole. We also witnessed this trend with the library community, where different stakeholders – users, library patrons, or management – require that new research artifacts, such as scientific blogs, datasets, or citation links, be included in the library catalogs alongside existing artifacts. Knowledge Graphs (KG) have been applied as means to bring together data from heterogeneous sources, within or across multiple domains, and this has shown to be an effective approach. In this work, we adopt them to bring research artifacts of different types in a more compact research unit. This includes the approach to specify it and its components considering the library environment, and the accompanying methodology used to implement it. Finally, we explore the value and potential use cases from this integration for the library community in the domain of economics.

Keywords: Knowledge graph · Scholarly artifacts · Digital libraries · Semantic Web

1 Introduction

Opening up the research process represents an important development for scholarly communication. Regardless of the extent (include data, source code, scientific workflow implementation, and more), there is evidence of benefits from its adoption by researchers [17]. Moreover, initiatives like the FAIR principles [19] also go in this direction as they provide guidelines for research artifacts (research data, in this case) to increase their re-usability for both humans and machines. Thus, as the Open Science movement and relevant practices build momentum, more research artifacts will be managed (created, described, stored,

etc.), become part of the scholarly communication process, and become valuable assets for research communities.

Research communities employ different, often dedicated, dissemination platforms for their research deliverables. This creates a distributed infrastructure of artifacts which often requires the same effort to publish and access artifacts like dataset, citation link, open notebook, etc., in the corresponding platforms. Libraries are important scholarly infrastructure hubs and in a good position to start to address this opportunity. Thus, as new research artifacts become available, they have an opportunity to offer a more comprehensive research picture, i.e., the ability to provide complementary aspects of a research work as well.

We faced a similar requirement at our library institution – requirements from both authors and library users – for more research artifacts as part of the library collections or catalogs. In this paper we conceptualize and conduct a preliminary exploration of the role of Knowledge Graphs (KG) as means to bring different research artifacts – publications, datasets, blogs, and citations – in a more centralized, one-stop fashion for a library environment.

The paper is structured as follows: In Sect. 2 we provide the research motivation for this work, and the related work in Sect. 3. We then present the datasets selected for the work (Sect. 4), and our technical approach to the KG components in Sect. 5. In Sect. 6 we present few use case scenarios from our data, and make our conclusions in Sect. 7.

2   Research Motivation

In the library context, artifact collections are the single most important asset. While the Open Science is gaining ground, library users are not only generating more artifacts themselves, but they also expect richer research artifact collection from libraries. Libraries, then, should be capable to bring a range of research artifacts together that may consist of different type, metadata standards, research practices, etc., and be able to cope with new artifacts in the future.

Two use case categories we treat in this work revolve around accessing resources that (1) originate from the same research work, or (2) are relevant to a topic of interest:

1. Search for related artifacts: The user is interested to find all the available artifacts that stem from a single research work. For example, given a research publication from a library catalog, she might want to explore the dataset(s), implementation scripts/code, cited publications, etc., used. Moreover, comments on social channels from the scientific community about the paper, such as blogs, could also be of use.
2. Cross-artifacts search: Here the user is interested in a single type of artifact and wants to find relevant artifacts of any type, based on certain criteria (author, topic, publication venue, etc.), thus a cross-artifacts scenario. For example, she might want to know the most cited dataset (on a topic, or any criterion) in a given year; the dataset(s) cited in a research publication;

the number of citations for a publication; the datasets or blog posts on a certain topic or authored by certain authors; most commented publications (by analyzing blog posts from the domain); and so on.

Understandably, a lot of the use case scenarios depend on the available artifact metadata (for e.g., if bibliographic citations are missing, one could not conduct citation analysis). However, we see this only as a current shortcoming for at least 2 reasons: (1) as the Open Science picks up momentum, there will be more artifacts across domains, thus potential use cases will increase; and (2) libraries have a selection process in place to include research artifacts of interest as part of their catalogs, which matches well with the previous trend.

3   Related Work

Going beyond the publications, many initiatives already strive to capture and express as more complete (with as many artifacts) a research picture as possible. The term "enhanced publication" is often referred to as an umbrella term for this idea. It focuses on "compound digital objects" [18] that capture the different facets of research via its constituting artifacts.

The library domain has seen examples of KG adoption for the similar requirements and goals that other domains have adopted KGs. Haslhofer & Simon [12] point out the continuation and "readiness" of libraries for the adoption of KGs. In terms of the potential benefits, they state that "knowledge graphs can be vehicles for connecting and exchanging findings as well as factual knowledge", which is in line with the benefits generally observed with the KG adoption across domains. In addition, Zhang [20] notes KGs will enable libraries to move from "knowledge warehouses" to "acquisition tools". In the same work he presents the typical structure of such a KG, as well as the (automatic) means to create it. Hienert et al. [13] provide an integrated approach of scholarly resources from social sciences, typically found on multiple platforms, for a digital library. They focus on publications, data, and provide a finer granularity of the data, considering the domain requirements, such as survey details. Angioni et al. [4] apply the KG to unify research deliverables such as articles, research topics, organizations, and types of industry, all from different sources, in order to measure the impact that academic research has on the industry. A final library example, the Open Research KG proposes a model and an architecture to represent research outputs, with an initial focus on survey articles [11,14]. The way the scholarly output is modeled is through the (classes of) *research problem*, *research method*, and *research result*, all part of the *research contribution* under consideration. This allows a more structured and granular comparison between survey articles, such as based on the hypothesis tested, research methods, and so on.

In the context of infrastructure-like providers, Atzori et al. [6] report on a research infrastructure, materialized via its Information Space Graph – a graph representation of scholarly collection that constitutes of different artifacts, such as articles, datasets, people (authors), funders, and grants. In both KG examples, the Semantic Web technologies play an important part of the technical

infrastructure and implementation. Another publishing entity – Springer Nature – offers an enhanced access to its aggregation of scholarly resources, including publications, conferences, funders, research projects, etc., via its SciGraph [1]. Moreover, one of the largest infrastructure institutions in the domain of social sciences is embarking on an infrastructure project, the core of which will be a social sciences KG, bringing together all the collections of this institution, as well as establish links to external collections [2].

In the more academic umbrella of projects, the Microsoft Academic Graph models entities from the scholarly communication (authors, publications, datasets, citations, and other aspects), and represents them as a single graph, whereas the PID Graph connects persistent identifiers (PID) of different research artifacts, across PID schemes, in a single graph for new insights of the research ecosystem [8]. Finally, Aryani et al. [5] report on the graph of datasets linked with other relevant scholarly deliverables, such as publications, authors, and grants, and a corresponding model to represent these artifacts.

Despite the upwards trend of research on KG applications, often even for overlapping scholarly artifacts, the domain of interest, the scope of artifacts, or specific requirements often drive the need for new KG adoptions. We explore such a case to bring (the metadata of) scholarly artifacts in a machine-readable representation for the domain of economics.

4   Dataset Selection

We include several research artifact types in our KG, such as (scientific) blog posts, open access research publications, research data, and citation links. The artifacts were selected based partially on the complementarity they bring, as well as the user interest in a library environment. Next, we present some key descriptions of these 4 collections.

a) **(Open Access) Publications** remain one of the primary means of scholarly communication. They often represent the starting point from where researchers (including library users) search for relevant information. For this artifact type we rely on an Open Access collection of publications from Econ-Stor[1], a publishing platform for scholarly publications from the domain of economics and business administration at the Leibniz Information Centre for Economics[2] (ZBW). The types of publications in this collection include journal articles, conference proceedings, draft papers, and so on. The collection contains more than 108 K publications, and is provided as an RDF data dump, which suits our technology of choice for materializing the Artifacts Graph, as the Semantic Web technology provides a key element in it (see Sect. 5 for more).

---

[1] https://www.econstor.eu/.
[2] https://www.zbw.eu/.

b) **Research data** are seeing a surge in importance in the scholarly communication. As a result, the ZBW is also engaging in supporting it in its data holdings, such as its engagement with the Journal Data Archive[3] or Project GeRDI. Project GeRDI [3], a research data infrastructure, focused on providing research data management support for long tail research data. It targeted many disciplines, such as social sciences and economics, life sciences and humanities, marine sciences, and environmental sciences. During its 3-year run, it harvested more than 1.1 M dataset metadata, and had 9 pilot research communities to help specify the project requirements for its infrastructure services. Having a multidisciplinary research scope, albeit at different extents (life sciences contributions dominate the collection), provides means to potentially conduct cross-disciplinary use cases, and this is one of the reasons we included it as part of the dataset. We use the RDF version of the dataset, which is publicly available from Zenodo [16].

c) **Links between scholarly artifacts** can be quite complementary in a KG that contains publications and datasets because this enables one to check if a publication or dataset has been cited or not. For the purpose of this work we use the link collection from OpenAIRE's Data Literature Interlinking service, Scholexplorer[4], originally containing more than 126 M citation links – both literature-to-dataset and dataset-to-dataset, and 17 providers that contribute to the collection. For our use cases section, we rely only on a subset of this large collection, also available as an RDF data dump [15].

d) **Blog post collection** Social scientific collections, such as blogs and wikis, are another type of artifact that we have explored in the past as they have become an interesting development of scholarly communication. Blog authors often contact the ZBW to offer their collection to any of its publishing platforms. Moreover, as with many of the emerging research artifacts, different blogs on the topic of economics have been considered for integration. For this artifact type we chose the blog post collection from VoxEU[5], as a portal that provides analysis and articles on more than 30 economic topics. We harvested 8.5 K blog posts, including the ones published as late as April 11 of this year.

5 Artifacts Knowledge Graph: A Technical Perspective

Although the number of projects and research on the topic is increasing, there is still not a commonly accepted definition what a KG is [9]. We adopt the definition from Färber et al., who "use the term knowledge graph for any RDF graph" [10]. In this section we provide our KG adoption approach, starting with its architecture, semantic modeling of the datasets, and the KG instantiation.

---

[3] https://journaldata.zbw.eu/.
[4] http://scholexplorer.openaire.eu/.
[5] https://voxeu.org/.

5.1   KG Architecture

The debate about KGs does not end with their definition, and different approaches exist for their architecture design as well. We adopt the so called "Enterprise KG" architecture from Blumauer and Nagy [7, pp.146], which specifies 3 key "layers" of an KG:

1. Data sources: This layer contains the dataset, and the datasets can be of different representation, metadata description, and so on. In our case, we deal with structured (RDF) and unstructured (information retrieved from Web pages) representations.
2. Enterprise KG Infrastructure: Represents the core of the a KG and typically includes the graph database used to store (and query) datasets; AI/ML activities to populate, maintain, enrich, etc., the data layer; any KG management tasks that might be required, and so on. In our case, we rely on an RDF triple store for the database operations, and apply enrichment with one artifact collection.
3. Data consumers. This is the layer used by end users – developers, data scientists, etc., which offers different services for this purpose. In our case, we rely on the SPARQL service to retrieve artifacts of interest (as defined in the use cases).

Based on these suggestions, we adopted the architecture in Fig. 1. Starting from the data sources, since we are dealing with a variety of data provisions, including RDF and JSON data dumps, and HTML pages, we implement dedicated adapters to access the sources. Data ingestion, then, provides the resulting (meta)data. The Extract Transform Load-like process allows us to conduct different task on this metadata, such as pre-processing or enriching it (when applicable), before finally converting it to RDF – our model of choice – and a single graph representation. We organize, store, and maintain the resulting RDF in a triplestore, which approximately matches layer 2 above. In this layer we also plan for other graph-based technologies, such as the Property Graph, especially when it comes to graph analysis, hence its depiction (although not in use yet). Finally, in layer 3, different set of services can be developed (so far, we only rely on retrieval via SPARQL for our use cases).

5.2   Semantic Modeling of Artifacts

We were directly involved with the conversion of the last three collections, but we are familiar with that of the first one. Next we present the selection of vocabularies/ontologies for the datasets:

1. **Publications** This collection is already provided in RDF. Being that its metadata are mainly of descriptive nature, the Dublin Core Metadata Initiative (DCMI) is used to model the larger part of the collection, with more than 30 properties – both using the `DC Elements` and the `DC Terms` specifications. In few of the cases, there are more vocabularies that cover the same artifact
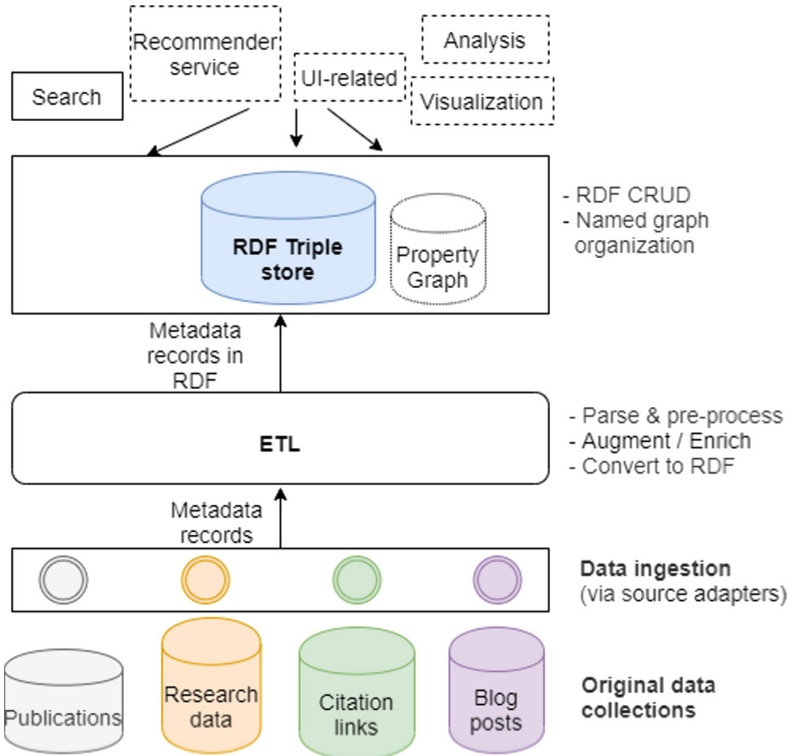
Fig. 1. The knowledge graph components

attribute. For example, to denote the `author` of a publication, in addition to DCMI, the `maker` from Friend of a Friend (FOAF) and `author` from the Semantic Web Conference Ontology (SWRC) have been used, as have the classes for `Document` (FOAF), `Paper` (SWRC), and `Item` from Semantically-Interlinked Online Communities (SIOC), for every literature item. Finally, the Bibliographic Ontology (BIBO) is used to capture additional bibliographic aspects, such as DOI and handle identifiers, ISSN, and so on.

2. **Research data** The type of an artifact is important during retrieval. We use the BibFrame Initiative (its `Dataset` class) for this purpose; targeting bibliographic descriptions, it supports a wide range of types (13 such types in its latest version), able to accommodate new scholarly artifacts in the future. To represent dataset identifiers, we relied on the DataCite Ontology: `PrimaryResourceIdentifier` class for identifiers of type DOI, and `AlternateResourceIdentifier` class for the rest. Due to the fact that the datasets come from different institutions or providers, we used the Europeana Data Model (EDM) to specify the dataset provider, specifically the class `Agent` to denote the provider. For the descriptive aspects of the datasets, we used the DCMI specification (`creator`, `date`, `description`, `format`,

Table 1. Artifacts collection features: Size, Vocabularies & Ontologies, and organization

	Collection 1	Collection 2	Collection 3	Collection 4
Artifact type	OA Publications	Research data	Citation links	Blog posts
Source	EconStor portal	Project GeRDI	ScholeXplorer service	VoxEU
Total items	>108 K	>1.1 M	>126 M citation links	8,752
Total RDF triples	3.6 M	24.2 M	3.9 B	396 K
Vocabularies/ collection	DCMI, FOAF, SIOC, SWRC, and BIBO	BiBFrame, DCMI, DataCite, EDM, Schema.org, PROV-O	CiTO, BibFrame, DataCite, DCMI, EDM, FRBR, PROV-O	Schema.org, DCMI, SIOC, PROV-O
Named graphs/collection	1	11	30	1

subject, and `title`), and Schema.org (its `keywords` property) to represent the dataset keywords. Finally, we used the PROV Ontology to add provenance information to the RDF dataset, including classes such as Generation, `Collection`, `Activity`, and `SoftwareAgent`, or properties like `generated`, `used`, `startedAt`, `endedAt`, `wasGeneratedBy`. Every artifacts collection has its own provenance metadata, which should help wither when reusing the (KG) collection, or using individual collections.

3. **Citation links** The Citation Type Ontology (CiTO) models the citation links. `Citation` class is used for that, whereas properties such as `hasCitationCharacterization`, `hasCitationDate`, `hasCitingEntity`, `hasCitedEntity`, capture the type of the link (references, relates to, supplements, etc.) between source and target. BibFrame and DataCite Ontology are used as before: to define the type of the resources being linked (publications and datasets) and the representation of identifiers; the same goes with the DCMI and PROV Ontology, too. EDM is also used as before, with the addition of the `isRelatedTo` property, used to model a citation type in the collection. The last vocabulary, Functional Requirements for Bibliographic Records (FRBR), is used to provide few link types that the previous ontologies did not support – `supplement` and `supplementOf`.

4. **Blog posts**: There are no new elements to model for this collection – classes or properties – as its items resemble the artifacts we already modeled previously, especially research publications and datasets. SIOC's `BlogPost` class denotes a blog post, whereas its `content` property denote the blog post content. The major metadata of every item are covered by DCMI, with Schema.org covering the blog post keywords, and the PROV Ontology providing support for the collection's provenance information.

Table 1 contains some information about the source collections, such as their size, number of RDF triples after the conversion, vocabularies/ontologies used, as well as the number of named graphs used to organize them in the triplestore. Deciding on the number of named graphs typically depended on the source data.

Namely, the research data are organized based on the data providers, whereas the citation link collections is organized based on its source files (30 in total).

5.3   KG Instantiation

The Semantic Web and Linked Data (LD) provide a great conceptual and technological fit for this research undertaking. Among other things, they are well suited for bringing heterogeneous collections in a common representation model (RDF, in this case), which fits well with the KG definition we referred to earlier.

We rely on the data ingestion and Extract, Transform, Load (ETL) components of the KG to harvest, enrich or linked up artifacts with external collections. In our case, we applied this for the blog post collection, as it provided enough text to engage in tasks such as automatic term assignment; link citations and research data, due to the short (textual) values for their metadata, did not lend themselves to such activities. Given that the ZBW has adopted the Thesaurus for Economics[6] to describe its data collections (EconStor dataset, for example), we decided to apply automatic term assignment (using Maui indexer[7]) to blog posts based on it, and add up to 3 terms to every blog post.

In this way, regardless of the term vocabulary used to describe the blog posts, we assign terms from a vocabulary that the library already uses. This bridges (to some extent) the terminology gap between heterogeneous collections (in this case EconStor publications and blog posts). In addition, this task enables us to explore the different components of instantiating KGs based on the domain practices – STW thesaurus and the domain of economics, in our case. Finally, we provide access to the resulting RDF data of the KG via SPARQL or as data dumps.

For the RDF conversion, its storage to a triplestore, and querying, we relied on Apache Jena Framework[8], and its TDB storage component[9]. We provide all our datasets based on the Creative Commons BY-NC 4.0 license[10]; due to the large collection (especially that of citation links), we only provide a subset and few exemplary SPARQL queries online[11].

6   Use Cases: Explored Scenarios

Generally, the use cases revolve around (fine-grained) search and involve different metadata elements across artifacts, such as: publication date, resource provider, persistent identifier, resource type of the resource and dataset size, etc.

---

[6] https://zbw.eu/stw/.
[7] https://github.com/zelandiya/maui.
[8] https://jena.apache.org/.
[9] https://bitbucket.org/fidanLimani/workspace/projects/KG.
[10] https://creativecommons.org/licenses/by-nc/4.0/.
[11] https://zbw.eu/beta/sparql-lab/about.

Publications and Data. GeRDI dataset collection contains a provider from social sciences that we can use to demonstrate cross-artifact search. After retrieving the subject terms for the datasets from this provider, we select the *"household composition"* to search the publications collections with. We find 2 publications that are also described with this subject term (*Inputs, Gender Roles or Sharing Norms? Assessing the Gender Performance Gap Among Informal Entrepreneurs in Madagascar*, and *The analytical returns to measuring a detailed household roster*). This is relatively specific subject for this provider, but if we want to check for *earnings*, another subject from the dataset collection, this gives us more than 200 matches. On the other hand, *health and satisfaction indicators* does not have an exact match from publications, although *health indicators* provides 3 matching publications.

In a user study, for example, while a user is checking a dataset on the topic of "household composition", we can show her two publications from EconStor on this topic. For the cases where there are more results, we can apply additional metadata to further filter the results (publication date, text similarity, or resource identifiers, for example).

On another scenario, the user retrieves the RD that directly support (as a primary source of data) the research paper at hand. If she wants to further specify the result, she can refine the query to include the most re-used RD in the collection (based on the number of times it has been cite). On the other hand, if the results are scarce or the user wants to broaden the search, she could also retrieve all the RD by the same author of the paper. In another scenario, a user can retrieve the "trending" RD (RD being cited the most in a more recent time frame) – and their corresponding publications – for a quick impression of what her community is currently working on. In a final scenario for this part, the user can rely on the "subject term" to search for a field of interest across link providers for a more interdisciplinarity search scenario (search for a fish type to see its fishing quotas, market fluctuations in a certain period, as well as the impact from climate conditions on its habitat).

Importantly, the DL collection we are working with primarily supports publications, thus in presented scenarios we assume the user first selects a publication and then proceeds to find RD. This aspect can easily be reversed (start with RD of interest, and find publications and/or RD).

Publications and Blogs. Another search across artifacts could be between publications and blog posts. On the topic of *health*, for example, there are 134 matches from the relatively small blog post collection, one matching dataset from one of the dataset providers (from social sciences), and many more (>700) from the literature collection.

Similar examples, with many more elements, can be devised to search across all the different collections, but due to space limit, we will provide such examples alongside the SPARQL service for the KG (see the link above).

Generic Use Case Scenarios: This set of scenarios provides more general information, which, although not the first use case of choice, could turn useful to the researchers. Few list examples are:

– List resources that are linked by the same publisher, publication date, domain, and other relevant metadata.
– Based on links that cite my research artifacts (publications or RD), who is using my RD? In what scenarios and context (information you get after reading a citing paper or RD, for example)? This question would apply both to individual researchers and institutions.
– Show the potential of relevant resources based on a certain criteria such as: classification terms for the subject of coverage, resources type, number of files that constitute a resources, etc.

While the trends of combining different types of research artifacts are already emerging, finding the right artifact collections to demonstrate the possible reuse scenarios is always challenging. This is so even within the same/similar research domains. More of these use cases will only be possible if sufficient information is present within the KG, specifically the new sources it will harvest.

7 Conclusion and Future Work

In this work we explored the role of a KG in providing more holistic view of research deliverables. Moreover, we specified the components and an approach to instantiate the KG with data from the social sciences domain. The KG developed in this work represents a nice basis for future work, including the following:

– Showing an approach to bring emerging research artifacts (links, datasets, blog posts, etc.) to a library environment based on KGs;
– The possibility to explore new use cases for this environment, important before adopting any new (KG) strategy for an interested institution;
– Domain-specific operations that aid with the resource (re)usage, such as the automatic term assignment based on a common thesaurus;
– Contribute with a KG with resources from social sciences, and provide new/emerging research artifacts to the research community;

Summing it up, providing a more holistic access to research deliverables is a good research direction. By adopting KGs as means to do it, in addition to the listed outcomes above, one supports extensibility for new artifacts, as well as means to exchange with other (open) KGs. In the future, we would like to tackle the aspects of scalability (especially when introducing new artifacts), as well as UI to enable an exploration from a broader set of users.

References

1. SN SciGraph: A Linked Open Data platform for the scholarly domain. https://www.springernature.com/gp/researchers/scigraph. Accessed 20 Aug 2021

2. Knowledge Graph Infrastructure. https://www.gesis.org/en/research/applied-computer-science/knowledge-graph-infrastructure. Accessed 24 Aug 2021
3. GeRDI: Generic Research Data Infrastructure. https://www.gerdi-project.eu/. Accessed 30 Feb 2020
4. Angioni, S., Osborne, F., Salatino, A., Reforgiato Recupero, D., Motta, E.: Integrating knowledge graphs for comparing the scientific output of academia and industry (2019)
5. Aryani, A., et al.: A research graph dataset for connecting research data repositories using rd-switchboard. Sci. Data **5**(1), 1–9 (2018)
6. Atzori, C., Bardi, A., Manghi, P., Mannocci, A.: The OpenAIRE workflows for data management. In: Italian Research Conference on Digital Libraries (IRCDL), Modena, Italy (2017, September 25). https://doi.org/10.5281/zenodo.996006
7. Andreas, B., Helmut, N.: The Knowledge Graph Cookbook Recipes That Work. Edition mono/monochrom, Vienna (2020)
8. Cousijn, H., et al.: Connected research: the potential of the PID graph. Patterns **2**(1), 100180 (2021)
9. Ehrlinger, L., Wöß, W.: Towards a definition of knowledge graphs. SEMANTiCS (Posters, Demos, SuCCESS) **48**, 1–4 (2016)
10. Färber, M., Bartscherer, F., Menne, C., Rettinger, A.: Linked data quality of dbpedia, freebase, opencyc, wikidata, and yago. Semant. Web **9**(1), 77–129 (2018)
11. Fathalla, S., Vahdati, S., Auer, S., Lange, C.: Towards a knowledge graph representing research findings by semantifying survey articles. In: Kamps, J., Tsakonas, G., Manolopoulos, Y., Iliadis, L., Karydis, I. (eds.) TPDL 2017. LNCS, vol. 10450, pp. 315–327. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67008-9_25
12. Haslhofer, B., Isaac, A., Simon, R.: Knowledge graphs in the libraries and digital humanities domain. arXiv preprint arXiv:1803.03198 (2018)
13. Hienert, D., Kern, D., Boland, K., Zapilko, B., Mutschke, P.: A digital library for research data and related information in the social sciences. In: 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL), pp. 148–157. IEEE (2019)
14. Jaradeh, M.Y., et al.: Open research knowledge graph: next generation infrastructure for semantic scholarly knowledge. In: Proceedings of the 10th International Conference on Knowledge Capture, pp. 243–246 (2019)
15. Limani, F.: OpenAIRE ScholeXplorer data dump in RDF (0.1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3820026
16. Limani, F.: Project GeRDI collection in RDF (0.3.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4104280
17. McKiernan, E.C., et al.: Point of view: how open science helps researchers succeed. elife **5**, e16800 (2016)
18. Vernooy-Gerritsen, M.: Enhanced Publications: Linking Publications and Research Data in Digital Repositories. Amsterdam University Press, Amsterdam (2009)
19. Wilkinson, M.D., et al.: The FAIR guiding principles for scientific data management and stewardship. Sci. Data **3**(1), 1–9 (2016)
20. Zhang, L.: Describe library resources with knowledge graph (2017)