RESEARCH ARTICLE

Knowledge Graphs, Metadata Practices, and Badiou's Mathematical Ontology

John Huck

University of Alberta

Metadata practices in libraries have been shifting towards a graph–centric data model for a number of years due to the influence of the Semantic Web on metadata standards as well as the ongoing engagement of libraries with linked data. This trend is likely to be sustained by the growth of the knowledge graph domain, which is animated by the interests of large technology companies and which represents a continuation of earlier programmes such as expert systems and the Semantic Web. Given the role of Semantic Web ontologies in knowledge graph development and the relevance of philosophical questions of ontology to cataloguing theory, metadata practitioners require theoretical frameworks suitable for conceptualizing the knowledge graph data model's mixture of data and ontology. To that end, this paper considers the mathematical ontology of philosopher Alain Badiou, which employs set theory to schematize a theory of the multiple. It outlines how Badiou's ontology is compatible with the graph data model and what it offers to metadata practitioners seeking to critically engage the knowledge graph paradigm.

Keywords: knowledge graphs; Semantic Web; philosophical ontology; metadata; critical cataloguing practice; Alain Badiou

Introduction

Due to the influence of the Semantic Web project and its successor, linked data, library metadata practices have been slowly shifting towards a graph-centric data model over the past decade, a model in which individual metadata statements are encoded and exchanged as Resource Data Framework (RDF) triples.[1] The significance of the RDF data model is that it enables schema-less, cross-domain data integration. For libraries, which have long shared cataloguing metadata with each other, this model represents an opportunity to link library data with data from external creators for mutual benefit. In turn, linked data activity in the library sector has helped sustain and even revive waning interest in the Semantic Web project's vision of data sharing and interoperability. Because the Semantic Web's roots lie in the history of expert systems (applications designed to answer questions by analyzing a body of knowledge), its goal is not only to facilitate the exchange of data, but also to enable new data to be derived from existing data by means of logical reasoning. This second aim has resulted in the introduction of Semantic Web ontologies into the landscape of metadata standards, even as the necessity of reasoning over library metadata remains unproven. In the last ten years, the Semantic Web and linked data have been recast as a part of a broader domain called knowledge graphs that incorporates aspects of machine learning and artificial intelligence, and which is animated, in part, by the interests of large technology companies like Google and Apple. These developments hold significant implications for the ways that metadata practitioners conceptualize their data.

Since subject classification plays a central role in metadata practice and metadata statements serve as assertions about reality, philosophical ontology (which engages with questions of being and existence) has long been relevant to cataloguing theory; however, growing library engagement with linked data, which has increased the likelihood that library metadata may be reused outside of libraries, amplifies the impact of choices made in the creation of metadata and gives greater weight to the ontological status of metadata

---

[1] A triple consists of a subject, predicate, and object and asserts that "some relationship, indicated by the predicate, holds between the resources [or entities] denoted by the subject and object" (W3C 2014). A set of triples is called an RDF graph.

statements, as well as to issues of knowledge representation. Using RDF to express metadata statements does not entail an obligation to employ the inferencing tools of Semantic Web ontology that RDF was designed to support. Nevertheless, metadata in this form can easily be incorporated into knowledge graphs that use reasoning tools of one kind or another. Therefore, metadata practitioners need theoretical frameworks suitable for conceptualizing the knowledge graph data model and supporting a critical approach to metadata practice.

To that end, this paper will consider one of the more innovative systems of philosophical ontology of recent times, the ontology of Alain Badiou. In his major work *Being and Event*, published in 1988, Badiou adapts mathematical set theory to elaborate an ontological system that privileges the multiple over the one. This paper argues that Badiou's mathematical ontology is compatible with the graph data model and offers metadata practitioners the benefit of a theoretical framework that accounts for the processes by which entities are recognized and brought into being within the terms of an existing discourse—such as may be represented in a structure like a knowledge graph—without prejudice to the existence of what may lie outside the terms of that discourse.

The first part of this paper considers the implications of the knowledge graph paradigm for metadata practice. It introduces the knowledge graph domain; outlines the parallel development of metadata practice and Semantic Web standards; explains the RDF data model; reviews current activity in the library community related to the Bibliographic Framework Initiative (BIBFRAME) standard; examines a debate within cataloguing theory related to the ontological nature of aboutness; and concludes by offering implications of the graph paradigm for metadata practice. The second part of the paper introduces Badiou's system of ontology; discusses Badiou's theory of the multiple and his account of knowledge; assesses his ontology's compatibility with the graph data model; and considers the implications of it for metadata practice. By providing a framework for the processes by which beings are recognized and added to the terms of an existing discourse, Badiou's ontology encourages metadata practitioners to think about the specific processes of metadata creation through which assertions about people, places, and ideas are put into circulation; and by introducing the distinctive figure of a subject who, acting in "fidelity" to an event, recognizes and enumerates the elements or parts of that event to reveal its meaning, Badiou's ontology invites metadata practitioners to imagine a broader range of participants in this activity.

The term *ontology* has different meanings in different domains. In philosophy, ontology refers to a tradition of thought that theorizes being and existence, while in information science it refers to a method of knowledge organization (Herre 2013). In computer science it is associated with specific technologies and languages related to the Semantic Web, such as RDF Schema (RDFS) and Web Ontology Language (OWL) (Feilmayr and Wöß 2016). These meanings are distinct, though not unrelated. Where the context is not obvious, this paper will refer to these respectively as philosophical ontology, knowledge organization ontology, and Semantic Web ontology.

Knowledge Graphs, Metadata Practices, and the Semantic Web

Knowledge Graphs

Defining Knowledge Graphs

The term *knowledge graph* represents a relatively new label for what was previously referred to as the Semantic Web, an area of research and activity with a long history. Ehrlinger and Wöß (2016) conclude that the term, in its current sense, was introduced in Google's announcement in 2012 that the company was enhancing its search product with information from a knowledge base that it called a knowledge graph.[2] In fact, the term itself has a prior history quite separate from the Semantic Web (detailed in the following section), but Google reintroduced the term, and—in spite of the confusion caused by the company's lack of a formal definition (Ehrlinger and Wöß 2016) and the corresponding lack of consensus over the term's exact meaning—it has gained traction in computer science and the technology sector.[3]

The Semantic Web was itself a development in the history of knowledge bases and expert systems. Ji et al. (2021) place knowledge graphs within that broader history of knowledge bases, which stretches from the 1950s through the development of expert systems like MYCIN in the 1970s and 1980s to RDF, OWL, and the

---

[2] Chen, Jia, and Xiang (2020) and Jie et al. (2021) also cite Google's announcement as a key moment in the history of the term.

[3] Enthusiasm for the term's adoption is likely due, at least in part, to its cachet as a buzzword. However, we may also speculate that, with the emergence of alternatives to RDF like the labelled property graph framework (Barrasa 2017) and a lack of publicly available information about the large commercial proprietary graph projects, the term *knowledge graph* has proven to be useful as a generic term that encompasses all of this activity. It is also free of any association with stale or failed technology, which the term *Semantic Web* may carry as accumulated baggage for some.

Semantic Web in the 1990s. In a survey of methods for reasoning over graphs, Chen, Jia, and Xiang (2020) cite semantic networks like Cyc as precursors to knowledge graphs. Ji et al. (2021) and Chen, Jia, and Xiang (2020) both count WordNet, DBpedia, YAGO, Freebase, NELL, and Wikidata among the major open data sources for knowledge graphs available today.[4]

A key point of difference amongst various definitions of *knowledge graph* is whether processes and meta-structures required to reason over the data are considered part of the graph or not. Ehrlinger and Wöß (2016), in seeking to distinguish knowledge graphs from ontologies (which may include classes and individuals together), offer the following definition: "a knowledge graph acquires and integrates information into an ontology and applies a reasoner to derive new knowledge" (2016, 3). Ji et al. cite this definition but elect to focus on structure, adopting a definition where knowledge graphs, as a variety of knowledge bases, "are sets of entities, relations and facts" (2021, 2). In surveying methods for reasoning over graphs, Chen, Jia, and Xiang (2020) are less concerned with defining knowledge graphs per se, but emphasize that the triple structure of graphs accommodates a wide range of reasoning methods. Oldman and Tanase (2018, 331) propose a conceptualization of knowledge graphs that emphasizes the application of "both automatic reasoning and crucially, collaborative human thinking and creativity." In a handbook on OWL, Uschold indicates, with reference to a figure in the text that shows both OWL classes and named individuals (i.e., class members), that "graphs representing knowledge and data as depicted in [the figure] are commonly referred to as knowledge graphs" (2018, 17). This is essentially the definition of ontology that Ehrlinger and Wöß wish to avoid, but no doubt represents a common understanding of the term amongst practitioners.

Across the range of different activities and concerns that these definitions reflect, the graph structure itself remains a consistent feature. Consisting of nodes and edges and expressible as triples, it easily accommodates both data aggregation and various kinds of reasoning methods, and so is suitable for projects that wish to do one or both of these things. Acknowledging the range of definitions and lack of consensus on the issue, this paper defines *knowledge graph* as an extensible graph structure containing data or a mixture of data and elements of ontology.

Earlier Knowledge Graph Research

Even as it differs in certain ways from the contemporary knowledge graph project, the Dutch research programme that coined the term *knowledge graph* in the 1980s is worth revisiting. It is interesting to note that the work originated with mathematicians rather than computer scientists. Nurdiati and Hoede (2008, 1) indicate that knowledge graph theory "was initiated by C. Hoede, a discrete mathematician at the University of Twente, and F.N. Stokman, a mathematical sociologist at the University of Groningen." Beginning as early as 1982, "the initial idea was to use graphs, a discrete mathematical concept . . . as a representation of the contents of medical and sociological texts" (Nurdiati and Hoede 2008, 1). Apparently, the researchers working in this area did not become aware of complementary work on conceptual graphs in the field of semantic networks until after 1988 (Hoede 1995). The conversion of texts into graphs was a central preoccupation. Bakker's Knowledge Integration and Structuring System (KISS) defined three steps for this process: text analysis, construct analysis, and link integration (Nurdiati and Hoede 2008, 2). Stokman and de Vries (1988, 187) characterize this as "structuring knowledge" rather than acquiring knowledge.

It is interesting to note that the emphasis in this work, as a direct application of mathematical graph theory, is not on entities per se but on the causal links between them: knowledge-how rather than knowledge-of. Hoede (1995) defines an ontology for this system that consists of only one type of entity and fourteen types of links, the core of which are derived from set theoretical relationships. Underpinning this technical ontology is a philosophical framework that joins a set theoretical structure with a subjectivist epistemology:

> A mind is able to distinguish somethings, a word we use for perceptions and other awarenesses. We suppose one, outer world and many sets of somethings, as many as there are minds (and computers). . . . The granularity of the world has led both to the awareness of what is called 'something' and to the idea of 'set', the awareness of composite something. (Hoede 1995, 310)

Within this framework, Hoede then links the infinite structure of a set with the infinite structure of a graph: "we consider a mind to have a mind graph reflecting the world as perceived by the mind. Certain substructures are carrying names. Each word in the vocabulary of a mind has a word graph corresponding to it" (1995, 314). Finally, Hoede places the subjective activity of the mind within the context of the communal

---

[4] Ji et al. (2021) call these datasets, while Chen, Jia, and Xiang (2020) call them knowledge graphs, illustrating the terminological confusion mentioned above.

project of knowledge sharing, which he intends his ontology to support: "the conclusion is that minds subjectively give meaning to words. Theory making is therefore always subjective. Objective theory making is the goal of a community of minds, say those of scientists" (1995, 318). Hoede's conceptualization of concepts as labelled subgraphs, or labelled sets, is striking in its simplicity, and reflects an intuitive sense of the associative, subjective nature of thinking, which is nicely summed up in the formula: "thinking is linking somethings" (Hoede 1995, 310).

Current State of Knowledge Graph Activity

The current state of the knowledge graph domain can be characterized as one that combines active research with the proprietary activities of large enterprises like Google and Apple. Enhancing Google Search, as mentioned above, was the original motivation for Google's knowledge graph. Chen, Jia, and Xiang (2020, 16) give "intelligent question-answering systems" as another significant application of graph reasoning, citing products by Apple, Microsoft, and Amazon as examples. There is much that remains publicly unknown about the knowledge bases behind these types of products. At a minimum, though, we do know that they leverage well-known public data sources to some extent. Wikipedia and DBpedia are sources for Siri (Ling 2020). Google says it "draw[s] from hundreds of sources from across the web" (Sullivan 2020), Wikipedia being just one. The Schema.org vocabulary represents the key to another public data source, namely semantic information embedded in websites. The vocabulary, it is claimed, is used by more than ten million websites to encode information in semantic triples, which, when embedded in a website's code, can be harvested as part of regular web crawling activities. According to Schema.org (2021), "many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences."

Google's knowledge graph and knowledge panels have given the notions of the Semantic Web a level of everyday visibility that they never had before. What are presented as benignly useful enhancements, however, do raise important questions about identity and representation in knowledge systems. Monea delivers an incisive critique of Google's knowledge graph, and of graph databases in general, through the interpretive lens of Deleuze, paying attention to "the potentiality of difference" (2016, 456). The concern he raises is this: "what escapes any graph database is difference itself. . . . The new is not enumerated" (2016, 458). Proprietary, commercial graphs, like Google's knowledge graph, are not directly accessible for scrutiny or revision, unlike the public datasets they may draw from, and yet they are likely to occupy central positions in the digital space for some time. Metadata practitioners must therefore be sensitive to questions and concerns about representation, misrepresentation, the unrepresented, and the unrepresentable when contributing to public datasets.

Metadata Practices and the Semantic Web

Common Origins

As an area of descriptive practice distinct from traditional cataloguing, metadata became established in libraries with the emergence of the World Wide Web in the 1990s, which brought with it the new problem of cataloguing online resources, on the one hand, and the rise of digital libraries on the other (Zeng and Qin 2008, 6; Borgman 2001). Emerging practice focused on developing standardized element sets and encoding formats that were suitable for the web, for exchanging metadata, and for cross-walking metadata between standards. A famous workshop held in Dublin, Ohio, in 1995 by library vendor OCLC eventually led to the development of a family of standards under the aegis of the Dublin Core Metadata Initiative (DCMI). In more recent years, as metadata work has become integrated with standard library operations, the term *metadata* has also come to be used to refer to library cataloguing or to cataloguing and non-traditional descriptive work together, depending on the context.

The development of metadata practices happened to coincide with the development of the Semantic Web. In fact, the two projects could be considered parallel responses to the same moment of burgeoning web activity at the end of the 1990s. The precursor to the W3C Semantic Web Activity (W3C 2013), which began in 2001, was the W3C Metadata Activity, which began in 1997 (W3C 2000).[5] At the time, Berners-Lee (1997) posited that the goal of metadata on the web was to support "self-describing information." These W3C Activities organized various working groups and interest groups, whose efforts resulted in the development of key Semantic Web standards: Resource Description Framework (RDF), RDF Schema (RDFS), and the Web Ontology Language (OWL). The first working draft of RDF was produced in 1997 and it became an official W3C Recommendation in 1999. Work on RDF Schema began in 1998 and it was accepted as a Recommendation in 2004. Work on OWL began in 2001 and it reached Recommendation status in 2004 as well. Subsequent versions of all these standards were produced over the following decade, with OWL 2

---

[5] Dates in this section are taken from the various working group pages and outputs linked from the cited W3C activity pages.

released in 2009, followed by a second edition in 2012, and RDF 1.1 and RDFS 1.1 released in 2014. The W3C Semantic Web Activity (W3C 2021) was itself superseded in 2013 by the W3C Data Activity (W3C 2021), which marked the completion of a shift from the original Semantic Web project to a linked data paradigm. The original vision for the Semantic Web imagined computers reasoning over distributed data on the web, which drove the development of RDFS and OWL. By 2006, Berners-Lee (2006) had recast this vision more narrowly, setting a more modest goal of encouraging greater data sharing in the RDF format, which he called linked data.

RDF Data Model and Technologies

RDF, RDFS, and OWL were designed to work together. RDF defines the basic data model, wherein a graph structure is composed of triples, each of which consists of a subject, predicate, and object. Subjects and predicates must be URIs, while objects can either be URIs or string values. URIs represent entities, which are called resources. Each triple represents an assertion about its subject. Because predicates are URIs, they can be the subjects of triples, and this plays a key role in defining semantic structures. RDFS introduces a small number of predicates and classes that can be used to define RDF resources—for instance, asserting something to be a class—and declare basic relationships between resources—for instance, asserting one class to be a subclass of another. Illustrating the symbiotic nature of these two standards, RDF is the source of the predicate rdf:type, which is used to declare that something belongs to a class, while RDFS is the source of the predicate rdfs:label, which is the standard way of associating a string value, like a name, with the URI representing a resource. OWL extends RDFS by introducing additional classes and predicates to allow more sophisticated relationships to be expressed. These relationships, whether declared with RDFS or OWL, serve as the logical framework that enables inferencing across a graph dataset. Allemang and Hendler (2011) provide a practical introduction to these languages, while Uschold (2018) provides a comprehensive guide to OWL. In practice, RDFS and OWL have not seen wide adoption. Graph query performance is affected by complex ontological constructs, and OWL is difficult to employ correctly. The labelled property graph model was developed as an alternative to RDF to address performance issues like these, among others (Barrasa 2017), but at the cost of moving away from a data format optimized for data exchange, as RDF is. Baker (2013, 64) observes that "just as English is useful without being the best of all possible grammars, RDF happens to be what we currently have – the only general-purpose language for data with any traction." Because of their shared lineage, RDFS and OWL remain an implicit aspect of the RDF data model, even if they are only used in limited ways.

RDF data can be serialized in a number of formats, including turtle, RDFa, and JSON-LD, and it can be published in a number of ways as well. For instance, it can be embedded in a website, compiled into a down-loadable dataset, or made accessible through a SPARQL endpoint. However, in order to be queried and manipulated as a graph, RDF data must be ingested into a specialized application, called a triplestore or graph database, that combines storage for triples with support for a query language, usually SPARQL, and a reasoner to perform logical inferencing. A single triplestore may contain multiple distinct graphs or data-sets. Since RDFS and OWL expressions are RDF triples, a triplestore stores and treats them like any other triple, and when they are added to a given dataset, they become a part of that graph. Graphs may grow over time as new data is added and, in the same way, semantic structure may be added to a graph at any time simply by adding RDFS or OWL triples. Even a small number of these triples can create semantic structures that span a large dataset. A single triple that asserts the equivalence between two predicates—for instance, sdo:creator and dcterms:creator—will cause a query processor to treat them as equivalent in a graph that contains millions of triples. Adding small bits of ontology like this can allow data from different sources to be combined and integrated.

Whether a knowledge graph is comprised of RDF triples stored in an RDF triplestore or non-RDF data stored in a graph store like neo4j, the data and whatever bits of ontology accompany it constitute a specific, localized, and instantiated dataset, albeit one that can be endlessly extended with new data. In this way, knowledge graphs have a dimension that simple linked data does not: namely, the ontological presence or absence of a given entity or assertion within a given graph.

Influence of the Semantic Web on Metadata Standards

Over the past two decades, Semantic Web standards have influenced the development of metadata standards as the latter were adapted to be compatible with RDF expression and as the conceptual implications of the RDF data model came to be better understood. The specifications of the DCMI are a good example of this phenomenon (DCMI 2021). The original set of Dublin Core metadata elements was introduced in 1999, and a specification for expressing Qualified Dublin Core in RDF was introduced in 2002. In 2003, a unified specification called DCMI Metadata Terms was released, which brought together for the first time several complementary specifications, including the Dublin Core elements, element refinements, encoding

schemes, and the DCMI Type vocabulary. In 2008, several specifications to support RDF expression of Dublin Core metadata were released, and these were gradually integrated into the main DCMI Metadata Terms specification (DCMI 2020). By adapting to the RDF data model, Dublin Core metadata standards have remained useful in a wide range of contexts since they can be easily integrated with other linked data predicates. The Data Catalog Vocabulary (DCAT), for instance, reuses certain Dublin Core terms for basic properties (W3C 2020). DCAT is an example of a recent metadata standard that has been designed as an RDF vocabulary from the ground up.

The adaptation of the Metadata Object Description Schema (MODS) standard to RDF represents a story with a different outcome. Developed by the Library of Congress in 2002 as an XML-based standard broadly compatible with bibliographic records (Library of Congress 2016) and still actively maintained in 2021, MODS has seen wide adoption by digital libraries and repositories. Work to express the MODS element set in RDF began around 2013 and resulted in an OWL ontology that modelled top-level XML elements as RDF classes and created custom properties for most of the lower-level elements. No subsequent version of this ontology has appeared and it is not thought to have been widely adopted, even as maintenance work on the XML standard has continued. Having found that the official ontology did not meet operational needs, a working group within the Samvera digital repository community set out to devise an alternate mapping from MODS to RDF, one that made minimal use of classes and reused properties from existing vocabularies (Samvera 2019). Hardesty and Young (2017) provide an account of the shift from XML to RDF within the Samvera community. The example of MODS serves to underline the fact that RDF represents an entirely different data model from the hierarchically nested structure of XML. It also serves to illustrate that, while there is an assumption that RDF should be used with an OWL ontology, there are significant challenges to pursuing this path.

A conservative approach to semantics informed the development of the Simple Knowledge Organization System (SKOS) standard, which allows thesauri and classification schemes to be expressed in RDF. The SKOS designers followed a principle of *minimal ontological commitment* and ensured that SKOS concepts would not be represented by OWL classes (Baker et al. 2013). Similarly, the Shape Expressions (ShEx) and Shapes Constraint Language (SHACL) validate RDF graph structures on a syntactic level rather than a semantic one (Gayo et al. 2018), a shift that reflects an emerging consensus that "overspecified ontologies can create unwanted entailments and complications when used in new and unanticipated contexts" (Coyle and Baker 2013, 3).

Linked Data and the Library Cataloguing Community

One of the most significant, or at least most visible, examples of the influence the graph data paradigm has exerted on metadata practices is the development of the BIBFRAME framework for bibliographic data, which is intended as a replacement for the MARC record format (Library of Congress 2021). BIBFRAME 1.0 was released in 2011, with 2.0 following in 2016. Schreur (2018) recounts some of the history of its development. BIBFRAME is part of an ongoing, sector-wide shift within the library cataloguing community towards creating and managing bibliographic metadata as RDF data, also referred to in this context as linked data. Whether cataloguing practices will fully transition to linked data remains an open question given the associated technical challenges, not to mention the as yet unproven value proposition of linked data for library cataloguing. Folsom sums up this state of uncertainty, tempering optimism with realism, thus:

> Many have assumed—and I might have once believed—that linked data would completely replace MARC. I'm not sure if MARC will ever 'die' as Roy Tennant famously suggested, but if I had to guess, it will be a combination of different types of data that replace MARC, not just linked data. . . . Libraries have economic incentives to limit the types of data we create and have to process in order to account for our collections. (Folsom and Jones 2021, 9)

Even so, Folsom continues to see a place for linked data in cataloguing practice: "I hope we can use scalable and flexible metadata practices to go beyond inventorying needs to highlight the strengths and unique collections through local discovery environments" (Folsom and Jones 2021, 9).

Alongside development of the BIBFRAME specification itself, libraries have collaborated to advance linked data practices. The Linked Data for Production (LD4P) initiative (Branan and Futornick 2018) began in 2016 as a Mellon-funded collaboration between the Library of Congress and five major university libraries seeking "to begin the transition of technical services production workflows to ones based in Linked Open Data (LOD)" (Schreur 2018, 7). The initiative has been extended with subsequent phases in 2018 (Futornick 2020) and 2020 (Branan and Futornick 2021). Along the way, an LD4 Community group was formed, which sponsors a number of affinity or interest groups as well as an annual conference. The group's vision statement proposes a relationship of mutual exchange between libraries and the wider linked data domain:

"the world enriched with library data; libraries enriched with the world's data" (LD4 Community 2021). This two-way exchange of linked data is a common rationale for its adoption by libraries (Yoose and Perkins 2013). Yet, engagement with linked data also represents an opportunity for libraries and librarians to bring professional ethics to the linked data conversation, one of the motivations behind the 2020 online LD4 conference, which drew over 1,500 participants from twenty-two countries: "by bringing together a broad range of perspectives, and centering diversity, equity, inclusion, and ethics in our discussions, we will create a community of practice for linked data in libraries" (LD4 2020). These two concerns—data sharing and ethical considerations—were represented at the conference by thematic tracks for Wikidata and ethics in linked data, and both topics also have dedicated affinity groups within the LD4 Community. This activity signals a recognition that public datasets like Wikidata represent a key point of connection between libraries and the broader domain of linked data and knowledge graphs. It also shows that library professionals hold a range of concerns and interests in linked data beyond the technical.

Ontological Nature of Aboutness in Subject Authority Data

Philosophical ontology is relevant to cataloguing theory, especially with regard to the nature of aboutness. When the IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR) came to devise a new model for subject authority data—Functional Requirements for Subject Authority Data (FRSAD)—it noted a lack of consensus about the ontological nature of aboutness and characterized two opposing views that it called nominalist and realist:

> For the thoroughgoing nominalist . . . aboutness should be conceived not as a property of works but rather as a relation, constructed by a particular person at a particular time, between a particular set of works and a particular linguistic expression (i.e., a name or label). The realist, on the other hand, is content to proceed on the assumption that subjects are real things that exist separately from the linguistic expressions that we use to name them, and that it is possible to determine "the" subject(s) of any given work. (IFLA Working Group on the FRSAR 2011, 10)

While acknowledging that most knowledge organization work is likely to be carried out within a realist framework, the FRSAD authors decline to take a position on the philosophical question, choosing instead to be guided by user expectations, which include the ability to search by subject. A feature of the resulting FRSAD model is the distinction between a topic or thing and its label, which are termed *thema* and *nomen* and modelled as separate entities.

Responding to the FRSAD report, Furner (2012) argues that philosophical considerations are entirely appropriate for the task of subject modelling and asserts that the FRSAD model betrays an implicit realist viewpoint in spite of its authors' stated intentions. Gemberling (2016), responding to Furner's critique, disputes the need for an ontological approach and argues that Peircean semiotics provides a framework that, for the purposes of cataloguing theory, sufficiently accounts for the contextually situated functioning of signs and their meaning, which vary over time. In Peirce's framework, the functioning of a sign is understood to rely on the relationship between the sign, its object, and the interpretant, a term that Gemberling takes to mean "our habitual way of seeing things as we consider the sign" (2016, 139).

Radio shares Furner's concerns about the implications of a realist orientation to aboutness for knowledge work and proposes a third orientation, critical realism, which recognizes that, "while subjects exist as the output of social conditions, that particular environment is susceptible to change in a way that concrete objects are not" (2018, 39). However, Radio is perhaps more concerned about the effect of interpretive or abstract metadata, like subjects or classifications, in general. Given the inevitable incompleteness of such systems and their tendency to propagate what Adorno calls *identity thinking*, which means a "way of knowing an object through classification" (Radio 2018, 39), Radio contends that "by incorporating aspects of identity thinking in bibliographic practices, models of the world and knowledge of it are reified in ways that obscure the possibility of alternatives" (2018, 42). Here we may remember Monea's (2016) critique of Google's knowledge graph as a project that purports to represent reality, even as a certain difference will always escape it. Considering both works and subjects together within the framework of Peircean semiotics, Radio concludes that "concrete [metadata] elements have an ontological primacy or immediacy over their counterparts that stem from the interpretive act" (2018, 36). He argues in favour of exploring a non-identarian cataloguing practice and proposes several ideas, including, intriguingly, Adorno's concept of the *constellation*, where objects are known in the context of their relations (Radio 2018).

In considering the limitations of library subjects and classification, Olson (2002) does not go so far as to reject them entirely, but her suggestions for remedies are not dissimilar. She proposes that practitioners seek out *eccentric techniques* that "breach the limit to create space for the voice of the Other" (2002, 227). This

strategy builds on the theory of Drucilla Cornell, which posits, Olson explains, that "since the limit is the location of marginalization and exclusion, it is at the limit that a relationship between mainstream, margins and the excluded can be developed" (2002, 226). Several of the practical projects Olson cites as examples involve creating networks or new relational paths, including an application "to allow searchers to save the items they select in a search as a 'user-defined-collection,' which is then available to others" (2002, 234) and a cataloguing project that supplements LCSH terms with terms from a second subject vocabulary to make "a collection documenting the Luiseño people of southern California . . . available from two cultural perspectives" (2002, 236). In the first case, a user collection forms a constellation or set of relations among materials reflecting an individual perspective. In the second case, relationships between the terms in separate vocabularies are constructed through the materials they are applied to. Cataloguing theorists have identified the ontological status of subjects as a problem within cataloguing theory that may carry implications for equitable representation within subject and classification systems. For librarians and metadata practitioners, critical evaluation of such information systems often leads to a search for practical ways to address their shortcomings.

Implications of the Knowledge Graph Paradigm for Metadata Practice

The knowledge graph paradigm has several implications for metadata practice. The first is to bring into focus the latent status of metadata statements as assertions about reality by introducing a data model wherein each piece of metadata is separate, addressable, and actionable within a structure that supports axiomatic reasoning. This does not represent a change in status for metadata statements, but the stakes have arguably been raised, and questions about accuracy in metadata, perspectives reflected in metadata, or even metadata's ontological basis carry more weight, particularly when the metadata is shared publicly. Using RDF predicates for metadata does not necessarily entail storing the metadata in a graph database, just as storing metadata in a graph does not necessarily entail applying Semantic Web ontologies to it for inferencing. Nevertheless, encoding descriptions in metadata standards designed for RDF puts that metadata in a form where there is a greater potential for its reuse in knowledge graphs, which may use reasoning of one kind or another.

The second implication is to further normalize a realist viewpoint on abstract categories within metadata practice. The RDF data model draws a distinction not between concrete and abstract objects, but between "web documents and concepts from the real world—people, organisations, topics, things" (W3C 2008). The FRSAD model, in separating the concepts of topic (thema) and label (nomen), aligns with this model, which its authors see as a pragmatic advantage: "subject authority data that are modelled based on FRSAD and encoded in SKOS and OWL will be able to become part of linked open data and contribute to the further development of the Semantic Web" (IFLA Working Group on the FRSAR 2011, 50). Furner is therefore not wrong to claim that the FRSAD model reflects a realist bias, although this does not prevent us from also accepting Gemberling's counter-argument that there is "probably no reason a self-professed nominalist could not accept the distinction between Thema and Nomen if it is useful" (2016, 143). While there is no obligation to include abstract entities in one's data, RDF easily accommodates such entities and accords them the same ontological status within the context of a graph as concrete entities. As the realist viewpoint is further normalized in information systems, it may paradoxically become harder to ignore questions about the ontological status of abstract subjects since stronger claims about reality are bound to elicit stronger critical reactions.

The third implication of the knowledge graph paradigm for metadata could be said to counterbalance the second, which is that the data model provides the ready means to pursue the kind of non-identarian descriptive practice that Radio advocates for, where links between resources become the basis for establishing meaning. In this way, it is not inhospitable to the nominalist orientation. Whether in the RDF data model or otherwise, stable identifiers for the nodes are necessary to support the graph structure, and these identifiers clearly provide a basis for linking between resources. Linking to create meaning is at the heart of the rationale that Oldman and Tanase offer for ResearchSpace, a semantic web-oriented platform for collaborative research, which allows for "better contextual engagement by placing things within historical and theoretical settings, not provided by raw Linked Data" (2018, 339). It is also reflected in Hoede's observation that "thinking is linking somethings" (1995, 310).

As outlined in this section, metadata practice has seen a gradual but steady shift towards RDF and the graph data model due to the ongoing influence of the Semantic Web, linked data, and knowledge graphs. In addition, libraries have come to recognize the value in public linked data datasets like Wikidata and see value in contributing to them. Taken together, the implications of a knowledge graph paradigm, including the three implications given above, shape conditions for metadata practice today and therefore provide the context for the research questions that will be addressed in the second part of this paper: is Alain Badiou's mathematical ontology compatible with the knowledge graph model? And, if it is, what might it offer to metadata practitioners?

Badiou's Mathematical Ontology[6]

Introduction to Badiou's Ontology

Badiou proposes that ontology is "science of being-qua-being. Presentation of presentation. Realized as thought of the pure multiple, thus as Cantorian mathematics or set theory. . . . Obliged to think the pure multiple without recourse to the One, ontology is necessarily axiomatic" (2005a, 517).[7] It is out of the scope of this paper to provide a detailed explanation of Badiou's ontology in full, but this section provides a brief overview of Badiou's use of set theory to establish an ontological system that—in contrast to ontological theories that espouse the idea that a singular, unified totality can exist a priori and that it holds ontological primacy over the many—considers multiplicity as the basis for being. Badiou distinguishes between pure, or inconsistent, multiplicity—by which he means "pure presentation understood as non-one, since being-one is solely the result of an operation" (2005a, 511)—and consistent multiplicity, which is "multiplicity composed of 'many-ones', themselves counted by the action of structure" (2005a, 503). Consistent multiplicity is what is recognized, or grasped, within a given context. Badiou names the operation that produces a consistent multiplicity the "count-as-one," and calls the combination of a multiple and its count-as-one framework a "situation." He names the framework of the count-as-one operation "structure," and he calls the framework by which the structure of a situation is counted-as-one "the state of the situation," "metastructure," or the "count-of-the-count." Structure is "what prescribes, for a presentation, the regime of the count-as-one. A structured presentation is a situation" (2005a, 522). It is worth noting that none of the structural entities in Badiou's ontology sit outside the ontological framework, and so they are sets as well.

In set theory, there are only sets. There is no intrinsic difference between a set and a member of a set. The members of a set are decomposable into sets, and those subsets are composed of members that are sets, and so on. This framework introduces an aspect of vertical structure, or a scale on which sets sit, in terms of being decomposable to a very small extent or composable to a very large extent. In choosing the multiple (inconsistent multiplicity) as the basis for his ontology, Badiou is saying there is no a priori totality of everything, no transcendental unity, and any such set would be the result of a count-as-one operation.

With only one kind of entity—the set—set theory employs one type of relation, belonging, to create structure between entities. Elements are said to *belong* to a set. However, a given set may be said to *include* another set if all of the elements of the other set are also elements of the given set. Badiou takes this pair of set-theoretical terms, belonging and inclusion, and associates them with the paired notions of presentation and representation as well as the two levels of the count—namely, the situation and the state of the situation. He contends that "philosophically it would be said that a term (an element) belongs to a situation if it is presented and counted as one by that *situation*. Belonging refers to presentation, whilst inclusion refers to representation" (Badiou 2005a, 501, emphasis added). He also states that "a term will be said to be included in a situation if it is a sub-multiple or a part of the latter. It is thus counted as one by the *state of the situation*" (Badiou 2005a, 511, emphasis added). The state of the situation is thus concerned with ensuring that there are no gaps in a given situation with regard to belonging: that everything is accounted for. An element, or subset, that is included in the situation but not recognized by it—something implicit or latent, for instance—does not belong to the situation, but the state of the situation will ensure that it is counted. Indeed, this theory corresponds to the axiom of the power set in Zermelo-Fraenkel set theory, which Badiou calls the *axiom of subsets* or *of parts*. The power set contains all possible subsets of a given set.[8] With this construct, Badiou allows for terms in a situation that are "normal," meaning they are both presented by the situation and represented by the state of the situation, and "not-normal," meaning they are presented but not represented. The possibility of latent or unrecognized elements in a situation is relevant for Badiou's conception of the event, discussed later in the paper.

As the next two sections outline, two aspects of Badiou's ontology share a particular resonance with the knowledge graph model: his theory of the multiple and his conception of knowledge.

---

[6] There are two main currents of criticism against Badiou. The first is that, in appropriating methods and technical innovations wholesale from another discipline, he is using them inappropriately or relying on them to show necessity where there is none. Nirenberg and Nirenberg (2011, 2012) and Rahman (2018) represent this view. The second criticism is a charge of muscular or militant thinking, an argument advanced by Laruelle in his book *Anti-Badiou* (Smith 2013). A sympathetic reading of Badiou's project is offered by Zalamea, who concludes that Badiou understands mathematics "as a sophisticated sheaf of methods and constructions for the systematic exploration of the transitory" (2012, 90).

[7] Zermelo-Fraenkel set theory is defined by a set of ten axioms (when the axiom of choice is included), so when Badiou says that ontology is axiomatic, it means that he has chosen to build his ontology on these axioms. Badiou clarifies that he is proposing mathematics as a language or schema for talking about being, not being itself: "the thesis I support does not in any way declare that being is mathematical, which is to say mathematical objectivities. It is not a thesis about the world, but about discourse" (2005a, 8). In *Briefings on Existence*, Badiou says that "being is only exposed to thought as a local site of its untotalizable unfolding," and that a *situation*, a term that will be explained below, is "this localization of the site of ontological thought" (2006, 161).

[8] In designating the power set the "state of the situation," Badiou's choice of words is not accidental. He is also thinking about the political state, which naturally has an interest in regulating what happens within it. This is the type of thinking-by-analogy that Nirenberg and Nirenberg (2011) criticize Badiou for, but the political is a significant aspect of Badiou's philosophical project.

Badiou's Theory of the Multiple

The development of an ontology of the multiple is perhaps the most compelling part of Badiou's ontology. The term *multiplicity*, which Badiou uses interchangeably with *multiple,* is from Gregor Cantor, the mathematician whose work initiated the discipline of set theory, as are the varieties of consistent and inconsistent multiplicities. As mentioned above, Badiou lands on the multiple as a way of rejecting the One of theology, but he also makes clear his desire to break with the linguistic turn in philosophy (Badiou 2006, 121). Badiou calls pure multiplicity "the manifold unfolding the unlimited reserve of Being as a subtraction from the power of the One" (2006, 35). Badiou's interest lies in the infinite rather than enumeration—in potential, not determination—and the multiple, which can be aggregated or decomposed, represents both that promise of inexhaustibility and a common schema for thinking both the given and the new.[9]

It would be almost impossible not to mention Gilles Deleuze in connection with the notion of the multiple, given that the notion of multiplicity figures prominently in the philosophies of both Badiou and Deleuze, and the two were engaged in debate with each other about their respective systems (Badiou 1999).[10] Deleuze's multiplicity, broadly speaking, represents an avoidance of categories and finds expression in such distinctive elements of his philosophy as *vectors of becoming* and *bodies without organs.* Badiou is looking for something different: an escape from totality. The nesting of sets and the possibility of infinite sets, as provided in Cantorian set theory, are thus of particular use for his project.[11]

Set theory forms the basis of applied reasoning languages like OWL and, for this reason, there is an obvious correspondence between Semantic Web ontologies and the Cantorian multiplicity. But there is also a more basic structural correspondence between the graph and the set, for which Badiou provides a demonstration. We might recall that Hoede (1995) made this association as well. Badiou, concerned to demonstrate that relations within or between sets do not represent a species of structure different than sets, shows that a simple relation between two entities can be expressed as an ordered set (Badiou 2005, 443). A pair can be ordered by expressing the second term as a set with two elements: the first and the second: [{a},{a,b}]. A relation, then, is "a set such that all of its elements have the form of ordered pairs" (Badiou 2005a, 445). The compatibility of Badiou's ontology with the graph data model will be taken up in more detail below.

Badiou's Account of Knowledge

In the course of preparing the ground for his introduction of Cohen's forcing technique in *Being and Event*, Badiou offers a set of what he terms *orientations of thought*, which include constructivist, generic, and transcendent (2005a, 282–85). Each of these orientations represents a different approach to the problem of being exceeding the capacity of thought to think it. Badiou associates these orientations with historical projects and figures in mathematics, philosophy, and political theory, and claims that mathematical ontology is able to accommodate each of them. These orientations also figure in *Briefings on Existence*, where Badiou explains that "it is when you decide upon what exists that you bind your thought to Being. That is precisely when, unconscious of it all, you are under the imperative of an orientation" (2006, 57).

Badiou designates the constructivist orientation as the province of knowledge and says that, within this orientation, "that which is not susceptible to being classified within a knowledge is not. 'Knowledge' designates here the capacity to inscribe controllable nominations in legitimate liaisons" (2005a, 293). The constructivist orientation attempts to make the excess that escapes knowledge as small as possible by cataloguing as much as possible. The impetus to discern, distinguish, and name lures the constructivist orientation into thinking that being can be known, but Badiou contends the gap will never be bridged in this way:

> Rather than being a distinct and aggressive agenda, constructivist thought is the latent philosophy of all human sedimentation; the cumulative strata into which the forgetting of being is poured to the profit of language and the consensus of recognition it supports. Knowledge calms the passion of being: measure taken of excess, it tames the state, and unfolds the infinity of the situation within the horizon of a constructive procedure shored up on the already-known. (2005a, 294)

---

[9] For Badiou, multiplicity is closely associated with the political: "of course, every situation is ontologically infinite. But only politics summons this infinity immediately, as subjective universality. . . . the situation is open, never closed, and the possible affects its immanent subjective infinity" (2005b, 143).

[10] It is interesting to note that they derive their respective concepts of multiplicity from different sources. Deleuze draws on Henri Bergson as the source for his multiplicity, reintroducing Bergson's thought, which had fallen out of fashion, in his 1996 book *Bergsonism*. In Bergson, quantitative multiplicity is spatial while qualitative multiplicity is temporal. Qualitative multiplicity is associated with the quintessential Bergsonian concept of duration and reflects a state of simultaneity or coexistence (Lawlor and Moulard-Leonard 2020).

[11] Laruelle (2013, 185) observes that the two philosophers' concepts of multiplicity represent "decisions that are symmetrically opposed," even as Laruelle rejects both of them.

One could say that Badiou recognizes in knowledge a Sisyphean inevitability: all the activities of human meaning, understanding, and communication are pursued without end, yet destined never to be complete. Knowledge will always fall short of capturing the difference of being, but this comes as no surprise. This is the charge Monea (2016) makes against knowledge graphs. Monea locates the limit in the act of enumeration itself and explicitly rejects Badiou's ontology on this basis in favour of Deleuze's framework of difference. As we have seen, Badiou is concerned with what falls outside the terms of the situation and looks to the technique of forcing and the generic to resolve this impasse. The provocative thesis in Badiou's conception of knowledge is not merely that an accumulation of knowledge is insufficient for change but that it actually discourages change, an inversion of the conventional view that positions knowledge as an element of empowerment. Badiou is not saying that knowledge has no effect, but rather that it cannot be a catalyst for action. Even the "infinity of each language . . . and the heterogeneity of languages" (2005a, 291), as well as the assiduously pursued endeavors to say more with them, are insufficient to overcome the structural deficit that forms the other half of the bargain offered by the *controllable nominations* by which we exchange knowledge.

Compatibility of Badiou's Ontology with the Graph Data Model

On a very basic level, Badiou's ontology is compatible with the knowledge graph data model. The set (or multiple) and the graph are both extensible structures. One can also think about a graph as a set of different kinds of sets: sets of nodes, sets of relationship types, and sets of specific relationships between nodes. However, there is a key difference between these basic structures in the context of Semantic Web ontology languages, which draw a distinction between classes and individuals. Individuals are members of a class, and they all share the property of belonging to that class. When designing data models, ontologists make decisions about which entities will be classes and which will be individuals, and the same entities can be modelled in different ways by different modellers. Limitations in the modelling language may play a role in these decisions. In OWL, some restrictions represent trade-offs that are necessary to allow the language to function at all (Uschold 2018). For instance, while a class may be a subclass of another classes, it is prohibited from belonging, as an individual, to another class. These restrictions create a distinction between classes and individuals that does not exist in set theory.

Reasoning languages and logical classes are not an essential aspect of the graph model, though, and there are alternative ways to create structure in a graph. On a basic level, whole-part relationships between individuals can be expressed. The Open Archives Initiative Object Reuse and Exchange (OAI-ORE) specification was designed for making arbitrary aggregations in this way (Open Archives Initiative n.d.). In fact, most of the primitives in Hoede's (1995) ontology for knowledge graphs were for different types of relationships rather than class structures. Moreover, Semantic Web ontologies are not the only method of reasoning over graphs (Chen, Jia, and Xiang 2020). Syntactical validation of RDF with ShEx or SHACL and modelling with minimal ontological commitment both represent an approach that de-emphasizes Semantic Web ontologies in graph data. For these reasons, the basic compatibility of Badiou's ontology with the knowledge graph paradigm is not invalidated by the differentiation between classes and individuals in Semantic Web ontology. Even so, one must be careful not to draw a simple correspondence between sets and classes.

Another point of alignment concerns the types of entities that Badiou's ontology addresses. As a system of thinking about being, it treats concrete and abstract entities in the same way: both are considered multiples and subject to the count-as-one. Badiou is often considered a Platonist for this reason. This represents an essential point of agreement with the RDF data model, which does not distinguish between concrete and abstract entities either and considers both to be "concepts from the real world" (W3C 2008).

One final comparison worth noting is between the infinite aspect of the multiple and the *open world assumption* that underlies the RDF data model. The open world assumption is technically an aspect of RDFS and OWL, not RDF, although it is commonly associated with the RDF data model. It represents the premise that a given graph is not a closed system, which is an assumption the reasoning engine in a graph database relies on to generate inferences. Practically speaking, this means that the absence of an assertion does not prove that the assertion is false. Uschold explains that open world reasoners "can distinguish between 'no' and 'don't know'" (2018, 57). In theory, a given system could be configured to employ a closed world assumption instead, which might be useful in certain cases. However, the open world assumption reflects the expectation that a graph will grow and receive new data over time. A graph may include defined classes, which add meta-structure to the data, but the graph itself is a fundamentally flat and non-hierarchical structure, and so there is no structural impediment to adding new data. Perhaps more than other aspects, this feature of knowledge graphs bears the greatest affinity with the notion of the multiple or infinite set.

Implications of Badiou's Ontology for Metadata Practice

In the present-day context of the knowledge graph paradigm, metadata practitioners have an interest in theoretical frameworks to conceptualize their data and their work. Because they cannot control how the

linked data they create is used by others once it is shared, practitioners must take what care they can with the metadata they do create, which may include critical evaluation of its theoretical underpinnings. Library authority data represents a particular kind of metadata of high value for knowledge graphs, because it is about people, places, and ideas. Is it also more sensitive than other types of metadata for the same reason. Radio (2018) finds that abstract elements—such as subject, genre, target audience, and classification—present the risk of perpetuating what Adorno calls *identity thinking*, because they involve an element of interpretation in their application, and proposes limiting their use in descriptive practice for this reason. Whether one shares this view or not, it is evident that abstract metadata carries with it all the intrinsic problems of knowledge representation. People and places are not abstract per se, but they may be associated with abstract entities like groups that represent different aspects of their identity. Biographical or historical details that may be included in authority data about them will touch on questions of identity and representation and may involve a measure of conjecture or interpretation. Consequently, the ontological status of these types of entities within knowledge graphs matters a great deal, both because of their sensitive nature and because there is an implied correspondence between the reality of the world (or worlds) in which we live and a knowledge graph, which purports to be a model of it.

The principle value that Badiou's ontology offers to metadata practitioners working within a knowledge graph paradigm is in the provision of a theoretical framework of ontology that accounts for the processes by which beings are recognized and brought into being within the terms of an existing discourse, such as may be represented in knowledge structures like graphs, thesauri, or metadata. Moreover, it is a framework that is not prejudiced against the existence of what lies outside the already recognized. For Badiou, this is the difference between consistent and inconsistent multiplicity. Badiou's model is also hospitable to both the nominalist and realist perspectives on the aboutness problem in cataloguing theory, just as it accommodates multiple orientations of ontological thought. Laruelle (2013) claims that Badiou's ontology does not add much philosophy to the set theory it employs, but this could also be seen as an advantage. Therefore, the aboutness debate need not be definitively resolved in order to employ Badiou's model. Finally, by proposing a procedure for the generic, Badiou attempts to address, in a theoretical manner, the loss of difference that occurs within a constructivist knowledge project, which is Monea's principle concern about Google's knowledge graph. Further study is needed to determine whether or not this technically complex innovation achieves its purpose and what it might mean in practical terms for metadata practice.

By framing ontological recognition as a procedure, either the count-as-one or the procedure of fidelity, Badiou's ontology encourages metadata practitioners to think about the specific processes through which assertions about people, places, and ideas are enumerated and put into circulation in their own areas of knowledge work, and to consider who is performing the enumeration and establishing the *controllable nominations*. Thinking concretely about the entities that regulate the *situations* in which knowledge is codified—Google can easily be counted alongside traditional library authorities here—and considering whether they belong to the *structure* that regulates the count-as-one or to the meta-structure of the *state of the situation* may lead metadata practitioners to recognize or invite the participation of additional parties in metadata activities.

One of the most compelling aspects of Badiou's ontology is the figure of the subject acting in "fidelity" to an event. Badiou's event is a multiple, composed of many elements, large and small, but it is also, we might say, meta-ontological, in that it is an interruption of rather than a consequence of the state of the situation. The elements that are part of an event, what can be said to belong to it, are identified after the fact by a subject or subjects acting in sympathy with or fidelity to the event. Like the count-as-one, fidelity is an operation, defined as "the procedure by means of which one discerns, in a situation, the multiples whose existence is linked to the name of the event that has been put into circulation by an intervention" (2005a, 507). Procedures of fidelity are also closely associated with the process for discerning a truth, which Badiou defines as "the gathering together of all the terms which will have been positively investigated by a generic procedure of fidelity supposed complete (thus infinite). It is thus, in the future, an infinite part of the situation" (2005a, 524). The suggestion is that a procedure of discernment that addresses each and every term to determine whether it belongs to a truth will ultimately produce a result, meaning that truths do exist in the situation. However, because of the infinite nature of the situation, the procedure would also need to be infinite to consider every term. This could be likened to the approaching of a limit.[12]

---

[12] In elaborating the procedure of fidelity as the technical support for truths, Badiou turns to the mathematics of Paul Cohen, and specifically to the technique of forcing, which Cohen famously employed to prove the independence of the continuum hypothesis from the axioms of Zermelo-Fraenkel set theory ("Forcing Method" n.d.). A detailed explanation of this aspect of Badiou's ontology is beyond the scope of this paper. In very broad terms, it involves expanding the terms of a set by adding a generic model to it. Badiou joins the notion of the generic from the forcing method to his procedure of fidelity because he wishes to ensure that there is an element of these operations that falls outside the terms of the situation. In a completely closed or determined system, there could be no true change and no unforeseen events.

There are a number of unanswered questions about the relationship between Badiou's event and this subject (Gratton 2010), but the formulation does bear a certain intuitive sense. Events continue to unfold in the realm of ideas long after their time. Before one can speak of an event and "locate it," an understanding of the parts of that event is required, and this understanding is bound to vary between people and across time. The event is, in some ways, the quintessential multiple; we speak about events all the time, and they are understood to have effects, but when we try to define an event, we find ourselves considering a set of smaller events and facts that sit within its frame, asking, "which of these elements are relevant?" More to the point, we might also ask, "relevant to whom?" Badiou says that "the identification of multiples connected or unconnected to the supernumerary name (circulated by the intervention) is a task which cannot be based on the encyclopedia. A fidelity is not a matter of knowledge. It is not the work of an expert: it is the work of a militant" (2005a, 329). We may detect here some resonance with the charge against Badiou of authoritarianism. On the other hand, "militant" could describe the figure of an active, engaged, passionate human being.

The procedure of fidelity that Badiou describes imagines that the subject acting in fidelity to an event undertakes the task of identifying what belongs to it by considering elements of the situation one by one, making a *report* on each. Badiou says that a set of reports make up a finite enquiry, and he proposes that "a truth groups together all the terms of the situation which are positively connected to the event" (205, 335). This procedure of fidelity bears a striking resemblance to the theme of subjective engagement with and traversal of knowledge graphs and knowledge systems that runs through many of the works cited in this paper. Hoede (1995) conceptualizes a mind graph that reflects subjective understanding. Olson (2002) proposes facilitating the sharing of user-curated collections. Oldman and Tanase (2018, 331) include "collaborative human thinking and creativity" as a necessary element of a knowledge graph. Radio (2018) points to Adorno's concept of constellation. Monea argues for "graph databases and information extraction mechanisms that are open to the public" and, "rather than one graph, many graphs, a 'proliferant continuance'" (2016, 460). The inscrutable nature of the major commercial knowledge graph projects remains a barrier to a decentralized knowledge environment, notwithstanding the various open datasets these proprietary graphs are partially built from. Nevertheless, one conclusion we might draw is that a progressive approach to knowledge graphs should include not only diverse data sources, representing heterogeneous points of view, but also an element of creative, subjective engagement with that data.

Acting in fidelity to an event could almost be construed as a kind of moral imperative to act to bring about change except that Badiou's account of knowledge, falling under the so-called constructivist orientation, finds that knowledge is insufficient for creating change and may, in fact, impede it. At the very least, this provocative assertion serves as a rejoinder to the notion that knowledge work, such as that undertaken by metadata practitioners and librarians, is in some way virtuous for its own sake. Practitioners who wish to make a difference in the world (Wenger-Trayner and Wenger-Trayner 2020) need to consider that their ethical commitments may extend beyond their work.

Conclusion

Because of the ongoing influence of the Semantic Web project, metadata practitioners today often create and manage metadata using standards expressed in the RDF data model and make contributions to the store of open linked data on the web. Now that the knowledge graph domain has become established and linked to other areas of computer science like machine learning and artificial intelligence, metadata practitioners have a need for theoretical frameworks that are compatible with the graph data model. Knowledge graph activity external to the library ecosystem challenges metadata practitioners and librarians to reconsider the ontological status of the entities in their vocabularies and metadata. Even if library cataloguing does not ultimately transition to a practice based exclusively in linked data, it is reasonable to expect that authority data will continue to be shared with the public, and it remains true that metadata made available publicly as RDF could be aggregated by another party into a knowledge graph that drives some kind of information service or function.

Badiou's ontology represents a creative reuse of the mathematical framework of set theory. Based on the multiple, it has a natural structural alignment with the graph data model. Badiou's ontology recontextualizes the debate about the ontological status of aboutness within cataloguing theory, providing a common framework for concrete and abstract entities that accommodates both the nominalist and realist perspectives. In it, entities are recognized in a given context either by the normative processes of an authority carrying out the count-as-one or by the intervention of a subject acting in fidelity to an event. As such, Badiou's theory suggests that subjects play a role in determining the meaning of the

events that affect them. Badiou's procedure of fidelity shares an affinity with the kinds of subjective, path-making methods proposed for knowledge graphs and classification work by other writers cited in this paper.

The possibility that mathematics could form the basis for a critical philosophical ontology as well as for ontologies of reasoning offers the prospect of approaching the set-theoretical underpinnings of knowledge graphs from two distinct, yet related, perspectives. However, the technical elaboration of Badiou's ontology presents a barrier for the non-specialist, and it is difficult for the reader who does not already understand the mathematics to evaluate Badiou's use of it. Whether Badiou's re-deployment of set theory is successful, whether it is valid, and whether it is invalidated by the criticism of author-itarianism are all questions those wishing to apply his ontology should consider. At the same time, metadata practitioners are accustomed to employing different models and schemas, reusing and conjugating between them with an eye on flexibility often motivated by practical considera-tions. When evaluating such a model, a pragmatic approach will tend to ask what the model has to offer.

Badiou's ontology reminds us that one of the strengths of graph structures is the ability to aggregate data that contains contradictions. Inferencing is often applied to a graph to resolve contradictions and fill in gaps. A lighter touch that leaves contradictions in place might be appropriate in cases where data representing heterogeneous viewpoints is integrated. Finally, whether or not a philosophical ontology based on set the-ory is advisable, metadata practitioners may wish to look to the field of mathematics as a source for new approaches to knowledge organization problems: on a technical level, category theory has been successfully employed to characterize graph data (Reformat, D'Aniello, and Gaeta 2018), while on a philosophical level, Zalamea's (2012) vision of a transitory ontology for mathematics points to the potential in mathematics for modelling synthetic philosophical approaches.

Competing Interests

The author has no competing interests to declare.

References

Allemang, Dean, and Jim Hendler. 2011. *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*. 2nd ed. Waltham, MA: Morgan Kaufmann.

Badiou, Alain. 1999. *Deleuze: The Clamor of Being*. Translated by Louise Burchill. Minneapolis: University of Minnesota Press.

Badiou, Alain. 2005a. *Being and Event*. Translated by Oliver Feltham. London: Continuum.

Badiou, Alain. 2005b. *Metapolitics*. Translated by Jason Barker. London: Verso.

Badiou, Alain. 2006. *Briefings on Existence: A Short Treatise on Transitory Ontology*. Translated by Norman Madarasz. Albany, NY: State University of New York Press.

Badiou, Alain. 2009. *Logics of Worlds: Being and Event, 2*. Translated by Alberto Toscano. London: Bloomsbury.

Baker, Thomas, Sean Bechhofer, Antoine Isaac, Alistair Miles, Guus Schreiber, and Ed Summers. 2013. "Key Choices in the Design of Simple Knowledge Organization System (SKOS)." *Journal of Web Semantics* 20: 35–49. https://doi.org/10.1016/j.websem.2013.05.001.

Baker, Tom. 2013. "Designing Data for the Open World of the Web." *JLIS: Italian Journal of Library, Archives and Information Science = Rivista italiana di biblioteconomia, archivistica e scienza dell'informazione* 4 (1): 63–66. https://doi.org/10.4403/jlis.it-6308.

Barrasa, Jesús. 2017. "RDF Triple Stores vs. Labeled Property Graphs: What's the Difference?" https://neo4j.com/blog/rdf-triple-store-vs-labeled-property-graph-difference/. Archived at: https://perma.cc/V4JE-5QXL.

Berners-Lee, Tim. 1997. "Metadata Architecture." W3C. https://www.w3.org/DesignIssues/Metadata.html. Archived at: https://perma.cc/M4ZJ-32PX.

Berners-Lee, Tim. 2006. "Linked Data." W3C. http://www.w3.org/DesignIssues/LinkedData.html. Archived at: https://perma.cc/B9CK-9Q2L.

Borgman, Christine L. 2001. *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World*. Cambridge, MA: MIT Press.

Branan, Bill, and Michelle Futornick. 2018. "Linked Data for Production (LD4P)." Last modified January 10, 2020. https://wiki.lyrasis.org/display/LD4P.

Branan, Bill, and Michelle Futornick. 2021. "Linked Data for Production: Closing the Loop (LD4P3)." Last modified March 21, 2022. https://wiki.lyrasis.org/pages/viewpage.action?pageId=187176106.

Chen, Xiaojun, Shengbin Jia, and Yang Xiang. 2020. "A Review: Knowledge Reasoning over Knowledge Graph." *Expert Systems with Applications* 141: 112948. https://doi.org/10.1016/j.eswa.2019.112948.

Coyle, Karen, and Tom Baker. 2013. "Dublin Core Application Profiles: Separating Validation from Semantics." http://www.w3.org/2001/sw/wiki/images/4/4a/RDFVal_Coyle_Baker.pdf. Archived at: https://perma.cc/M2KV-TFF5.

Deleuze, Gilles. 1988. *Bergsonism.* Translated by Hugh Tomlinson and Barbara Habberjam. New York: Zone Books.

Dublin Core Metadata Initiative. 2020. "DCMI Metadata Terms." https://www.dublincore.org/specifications/dublin-core/dcmi-terms/. Archived at: https://perma.cc/H454-7BKK.

Dublin Core Metadata Initiative. 2021. "Specifications: Dublin Core." https://www.dublincore.org/specifications/dublin-core/.

Ehrlinger, Lisa, and Wolfram Wöß. 2016. "Towards a Definition of Knowledge Graphs." In *Posters & Demos @ SEMANTiCS 2016 and SuCCESS'16 Workshop*, edited by Michael Martin, Martí Cuquet, and Erwin Folmer. http://ceur-ws.org/Vol-1695/paper4.pdf. Archived at: https://perma.cc/24CE-W8VD.

Feilmayr, Christina, and Wolfram Wöß. 2016. "An Analysis of Ontologies and Their Success Factors for Application to Business." *Data & Knowledge Engineering* 101: 1–23. https://doi.org/10.1016/j.datak.2015.11.003.

"Forcing Method." n.d. *Encyclopedia of Mathematics.* http://encyclopediaofmath.org/index.php?title=Forcing_method&oldid=51543. Archived at: https://perma.cc/ZQ8S-VGQA.

Folsom, Steven M., and Edgar Jones. 2022. " Interview with Steven Folsom on Linked Data." *Cataloging & Classification Quarterly* 60 (1): 1–12. https://doi.org/10.1080/01639374.2021.2010854.

Furner, Jonathan. 2012. "FRSAD and the Ontology of Subjects of Works." *Cataloging & Classification Quarterly* 50 (5–7): 494–516. https://doi.org/10.1080/01639374.2012.681269.

Futornick, Michelle. 2020. "Linked Data for Production: Pathway to Implementation (LD4P2)." Last modified April 20, 2021. https://wiki.lyrasis.org/display/LD4P2.

Gayo, Jose Emilio Labra, Eric Prud'hommeaux, Iovka Boneva, and Dimitris Kontokostas. 2018. *Validating RDF Data.* Burlington, MA. Morgan & Claypool Publishers. https://doi.org/10.2200/S00786ED1V01Y201707WBE016.

Gemberling, Ted. 2016. "FRSAD, Semiotics, and FRBR-LRM." *Cataloging & Classification Quarterly* 54 (2): 136–44. https://doi.org/10.1080/01639374.2015.1133751.

Gratton, Peter. 2010. "Change We Can't Believe In: Adrian Johnston on Badiou, Žižek, & Political Transformation." *International Journal of Žižek Studies* 4 (3). http://zizekstudies.org/index.php/IJZS/article/view/363/363.

Hardesty, Juliet L., and Jennifer B. Young. 2017. "The Semantics of Metadata: Avalon Media System and the Move to RDF." *Code4lib Journal* 37. https://journal.code4lib.org/articles/12668.

Hoede, Cornelis. 1995. "On the Ontology of Knowledge Graphs." In ICCS 1995: *Conceptual Structures: Applications, Implementation and Theory*, edited by Gerard Ellis, Robert Levinson, William Rich, and John F. Sowa, 308–22. Berlin: Springer. https://doi.org/10.1007/3-540-60161-9_46.

IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR). 2011. *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model.* Edited by Marcia Lei Zeng, Maja Žumer, and Athena Salaba. https://repository.ifla.org/handle/123456789/835.

Ji, Shaoxiong, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S. Yu. 2021. "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications." *IEEE Transactions on Neural Networks and Learning Systems* 33 (2): 494–514. https://doi.org/10.1109/TNNLS.2021.3070843.

Laruelle, François. 2013. *Anti-Badiou: The Introduction of Maoism into Philosophy.* Translated by Robin Mackay. London: Bloomsbury.

Lawlor, Leonard, and Valentine Moulard Leonard. 2020. "Henri Bergson." *The Stanford Encyclopedia of Philosophy* Edited by Edward N. Zalta. https://plato.stanford.edu/archives/fall2020/entries/bergson/. Archived at: https://perma.cc/9RN4-TGDQ.

LD4. 2020. "2020 LD4 Conference on Linked Data in Libraries." https://sites.google.com/stanford.edu/ld42020/.

LD4. 2021. "LD4 Community." https://sites.google.com/stanford.edu/ld4-community-site/home.

Library of Congress. 2013. "MODS RDF Ontology: Primer." Metadata Object Description Schema. http://www.loc.gov/standards/mods/modsrdf/primer.html. Archived at: https://perma.cc/K9FV-A7AH.

Library of Congress. 2016. "MODS: Uses and Features." Metadata Object Description Schema. https://www.loc.gov/standards/mods/mods-overview.html.

Library of Congress. 2021. "Bibliographic Framework Initiative." https://www.loc.gov/bibframe/.

Ling, Xiao. 2020. "Knowledge Base Construction at Apple Siri." Presentation at Stanford CS 520 Knowledge Graphs seminar, April 7, 2020. YouTube video, 1:53:27 (presentation begins at time-mark 1:03:50). https://youtu.be/ZWM-Dlw3VCM.

Monea, Alexander. 2016. "The Graphing of Difference: Numerical Mediation and the Case of Google's Knowledge Graph." *Cultural Studies ↔ Critical Methodologies* 16 (5): 452–61. https://doi.org/10.1177/1532708616655763.

Nirenberg, Ricardo L., and David Nirenberg. 2011. "Badiou's Number: A Critique of Mathematics as Ontology." *Critical Inquiry* 37 (4): 583–614. https://doi.org/10.1086/660983.

Nirenberg, Ricardo L., and David Nirenberg. 2012. "Reply to Badiou, Bartlett, and Clemens." *Critical Inquiry* 38 (2): 381–87. https://doi.org/10.1086/662748.

Oldman, Dominic, and Diana Tanase. 2018. "Reshaping the Knowledge Graph by Connecting Researchers, Data and Practices in ResearchSpace." In *The Semantic Web – ISWC 2018*, edited by Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl, 325–40, Berlin: Springer. https://doi.org/10.1007/978-3-030-00668-6_20.

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries.* Dordrecht: Kluwer Academic Publishers.

Open Archives Initiative. n.d. "Open Archives Initiative Object Exchange and Reuse." https://www.openarchives.org/ore/.

Radio, Erik. 2018. "Abstraction, Concrescence, and Identity in Descriptive Metadata." *Journal of Library Metadata* 18 (1): 31–44. https://doi.org/10.1080/19386389.2018.1461455.

Rahman, Shiva. 2019. "On Why Mathematics Can Not Be Ontology." *Axiomathes* 29: 289–96. https://doi.org/10.1007/s10516-018-9406-2.

Reformat, Marek Z., Giuseppe D'Aniello, and Matteo Gaeta. 2018. "Knowledge Graphs, Category Theory and Signatures." In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 480–87. https://doi.org/10.1109/WI.2018.00-49.

Samvera MODS to RDF Working Group. 2019. *MODS to RDF Mapping Recommendations.* https://repo.samvera.org/concern/generic_works/b922d5e1-abfe-4305-a696-0c6d72c962ff.

Schema.org. 2021. https://schema.org/.

Schreur, Philip. 2018. "The Evolution of BIBFRAME: From MARC Surrogate to Web Conformant Data Model." Paper presented at IFLA WLIC 2018, Kuala Lumpur. http://library.ifla.org/id/eprint/2202.

Smith, Anthony Paul. 2013. Review of *Anti-Badiou: On the Introduction of Maoism into Philosophy*, by François Laruelle. *Notre Dame Philosophical Reviews.* https://ndpr.nd.edu/reviews/anti-badiou-on-the-introduction-of-maoism-into-philosophy/. Archived at: https://perma.cc/224N-HCM7.

Sri Nurdiati, S. N., and Cornelis Hoede. 2008. "25 Years Development of Knowledge Graph Theory: The Results and the Challenge." *Memorandum* 2/1876 : 1–10. https://research.utwente.nl/en/publications/25-years-development-of-knowledge-graph-theory-the-results-and-th.

Stokman, Frans N., and Pieter H. de Vries. 1988. "Structuring Knowledge in a Graph." In *Human-Computer Interaction: Psychonomic Aspects,* edited by Gerrit. C. van der Veer and Gijsbertus Mulder, 186–206. Berlin: Springer. https://doi.org/10.1007/978-3-642-73402-1_12.

Sullivan, Danny. 2020. "A Reintroduction to Our Knowledge Graph and Knowledge Panels." *The Keyword* (blog). Google. https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/. Archived at: https://perma.cc/7FRL-7T6V.

Uschold, Michael. 2018. *Demystifying OWL for the Enterprise.* Burlington, MA: Morgan & Claypool Publishers.

W3C. 2000. "W3C Metadata Activity Statement." https://www.w3.org/Metadata/Activity.html. Archived at: https://perma.cc/42VP-ZYEK.

W3C. 2008. "Cool URIs for the Semantic Web." https://www.w3.org/TR/cooluris/. Archived at: https://perma.cc/HTG8-78GR.

W3C. 2013. "W3C Semantic Web Activity." https://www.w3.org/2001/sw/. Archived at: https://perma.cc/R3PD-W2TZ.

W3C. 2014. " RDF 1.1 Concepts and Abstract Syntax." https://www.w3.org/TR/rdf11-concepts/.

W3C. 2020. "Data Catalog Vocabulary (DCAT) - Version 2." https://www.w3.org/TR/vocab-dcat-2/. Archived at: https://perma.cc/N7UK-TL2T.

W3C. 2021. "W3C Data Activity: Building the Web of Data." https://www.w3.org/2013/data/. Archived at: https://perma.cc/98J2-FD5W.

Wenger-Trayner, Etienne, and Beverly Wenger-Trayner. 2020. *Learning to Make a Difference: Value Creation in Social Learning Spaces.* Cambridge: Cambridge University Press.

Yoose, Becky, and Jody Perkins. 2013. "The Linked Open Data Landscape in Libraries and Beyond." *Journal of Library Metadata* 13 (2–3): 197–211. https://doi.org/10.1080/19386389.2013.826075.

Zalamea, Fernando. 2012. *Synthetic Philosophy of Contemporary Mathematics.* Translated by Zachary Luke Fraser. Falmouth: Urbanomic.

Zeng, Marcia Lei, and Jian Qin. 2008. *Metadata.* New York: Neal-Schuman Publishers.

OPEN ACCESS