1-1-2019

Classification as Catachresis: Double Binds of Representing Difference with Semiotic Infrastructure

Lindsay Poirier
University of California, Davis, lpoirier@smith.edu

Classification as Catachresis: Double Binds of Representing Difference with Semiotic Infrastructure

Lindsay Poirier
University of California, Davis

ABSTRACT

Background *This article explores the results of a three-year ethnographic study of how semiotic infrastructures—or digital standards and frameworks such as taxonomies, schemas, and ontologies that encode the meaning of data—are designed.*

Analysis *It examines debates over best practices in semiotic infrastructure design, such as how much complexity adopted languages should characterize versus how restrictive they should be. It also discusses political and pragmatic considerations that impact what and how information is represented in an information system.*

Conclusion and implications *This article suggests that all databased representations are forms of data power, and that examining semiotic infrastructure design provides insight into how culturally informed conceptions of difference structure how we access knowledge about our social and material worlds.*

Keywords *Ethnography; Metadata; Science and technology studies; Knowledge representation; Critical data studies; Semantic Web*

RÉSUMÉ

Contexte *Cet article explore les résultats d'une étude ethnographique ayant duré trois ans sur la manière de concevoir les infrastructures sémiotiques, c'est-à-dire les normes et cadres numériques tels les taxonomies, schémas et ontologies qui donnent un sens aux données.*

Analyse *L'article examine les débats sur les meilleures pratiques dans la conception des infrastructures sémiotiques, tels que le niveau de complexité qu'un langage adopté devrait démontrer par rapport à son caractère restrictif. Il rend compte aussi de considérations politiques et pragmatiques ayant un impact sur le choix d'informations représentées dans un système d'information et la manière de les représenter.*

Conclusion et implications *Cet article suggère que toute représentation dans une base de données est une utilisation de données à des fins de pouvoir, et que l'examen de la manière dont les infrastructures sémiotiques sont conçues peut nous aider à mieux comprendre comment les notions de différence informées culturellement structurent la façon dont nous appréhendons les connaissances de nos univers sociaux et matériaux.*

Lindsay Poirier is an Assistant Professor in the Science and Technology Studies Program at the University of California. Email: lnpoirier@ucdavis.edu .

Canadian Journal of Communication Vol 44 (2019) 361–371 ©2019 Canadian Journal of Communication Corporation http://doi.org/10.22230/cjc.2019v44n3a3455

Mots clés *Ethnographie; Métadonnées; Études en science et technique; Représentation des connaissances; Études sur les données critiques; Web sémantique*

Introduction

Throughout the world, myriad databases are deployed to structure and organize knowledge. Databases are often structured to manifest formal data models, or what many computer scientists refer to as ontologies. Ontologies are abstract models that represent the relationships between a diverse array of objects and concepts in a particular knowledge domain. Within many databases, individual data points are formatted and described with schemas or controlled vocabularies, which are a set of standardized words or phrases used to define, describe, or index digital content. In this sense, ontologies, schemas, and controlled vocabularies are all forms of data infrastructure used to represent the meaning of data. Throughout this article, this set of data infrastructures is referred to as semiotic infrastructures.

Today, there are many standard semiotic infrastructures at a data modeller's disposal. For example, there are standard ontology languages, such as the Web Ontology Language (OWL), which can be leveraged for modelling the relationships between concepts in a database, and there are standard controlled vocabularies, such as Schema.org, which can be leveraged to define data points in a database. Knowledge representation specialists, ontologists, or data scientists, all of which are referred to in this article as semiotic infrastructure designers, design these standard semiotic infrastructures. Semiotic infrastructure designers develop standards and frameworks that data modellers and database engineers can leverage to name, describe, and order concepts within databases. The designers aim to develop codified modes of representation that modellers can use to preserve (across settings, disciplines, and organizations) a shared understanding of the worlds they structure in their databases, and to codify the objects of those worlds in such a way that computers can consistently recognize and compute them. As data modellers leverage semiotic infrastructures to model and define concepts within databases, they produce databased representations of the very social and material phenomena they are abstracting.

As businesses, governments, and scientists increasingly rely on big data analysis to drive automated decision-making, data modellers increasingly rely upon semiotic infrastructure to ensure that the data in the systems they design can move through various "data assemblages" and consistently maintain what they represent (Kitchin & Lauriault, 2018). It is important to study these data representations as they exert a form of infrastructural data power. This article will discuss the results of a three-year ethnographic research project that examined the work of semiotic infrastructure designers and the infrastructures they build. It considers how semiotic infrastructure designers—often trained to see identity and difference through the lens of formal logic—are learning to confront and at times challenge the politics of representation that emerges in their work. Before explaining the nature of the study, this article will first provide some background information about the domain of inquiry and how the analysis is framed.

Background: Semiotic infrastructure design

The researchers and practitioners interviewed for this study and referred to here as semiotic infrastructure designers would not refer to themselves this way. They would call themselves knowledge representation specialists, ontologists, data scientists, or database engineers. Most of the interviewees are trained in analytical philosophy (a field of philosophy that applies logical techniques to elucidate problems and draw conclusions), computer science, and/or information science. Many are also well versed in linguistic theory, often citing Alfred Tarski (1944), Noam Chomsky (1957), Charles Pierce (1960), and Ludwig Wittgenstein (1953) in philosophical debates about how to best approach the representation of phenomena. Most share a common goal of building data models to preserve the meaning of data as they move through diverse settings. Theorists and practitioners in this community are referred to in this article as semiotic infrastructure designers, not so much to characterize the history of scholarship that has informed their thinking and work but more to characterize the processes of signification that they are embroiled within. Semiotic infrastructure designers enact meaning and mediate difference as they formalize data standards.[1]

Semiotic infrastructure designers often are trained in and draw inspiration from the field of knowledge representation. Knowledge representation is a subfield of artificial intelligence (AI) concerned with developing strategies and tools to formally encode the meaning of concepts so that computers can solve complex tasks. Since the early 1970s, knowledge representation experts have sought to engineer naming practices, information models, and logical frameworks in databases (see, for example, Hayes, 1979; Lenat, 1995; Minsky, 1974; Sowa, 2000).

The work of these pioneers can be seen today in collaborative efforts such as Schema.org, where infrastructure designers are building structured data markup schemas that Web designers use to ensure that search engines can read the data and information on their pages. Another commonly used standard is Dublin Core, which can be characterized as a library cataloguing system for data and information on the Web. At its most basic, Dublin Core is a metadata standard that helps describe digital or physical objects and indexes them with a set of structured vocabularies also recognized by search engines or cataloguing systems. OWL outlines primitives, which are encoded properties or attributes for modelling the relationships between data points in such a way that computers can make logical assertions about the data. These are but three of the many kinds of semiotic infrastructures being built today.[2] In formalizing each of these standards, designers negotiated what makes one schema concept distinct from another concept or what types of primitives were needed to mark a diversity of differences. As data modellers leverage these standards to describe and order their data, they produce systems that structure how we know and find information.

Framework: The politics and pragmatics of knowledge representation

There are pragmatic trade-offs to formalizing meaning into semiotic infrastructure (Bowker & Star, 1999). To ensure that diverse communities adopt standard data languages in a consistent way and to ensure that automated systems can reliably interpret these languages, semiotic infrastructure designers strive to harden the meaning of words and how things are described by structuring the relationship between concepts in "formal" ways. In doing so, they reduce the complexity of what is represented. The more they restrict the expressivity[3] of the models they develop, the less opportunity there is for diverse communities to use the terms in these models in ways the designers did not intend or in ways that computers cannot understand. In other words, while languages and the meanings of things evolve culturally, fluidity and flexibility are often not characteristic features of formal semiotic infrastructures as their objective is to constrain interpretation. Thus, the boundary between formal and informal is a source of much deliberation among semiotic infrastructure designers.

In the 1990s, ethnographers studying knowledge representation communities emphasized the practice of "formalizing" knowledge as a mode of power (Adam, 1998; Forsythe, 1993; Star, 1995). For example, Diana Forsythe (1993), in her ethnographic study of five U.S.-based artificial intelligence laboratories from 1986 to 1993, observed that, in trying to articulate standard rules for knowledge that could be codified into reasoning systems, knowledge representation experts tended to employ what she characterized as "I am the world thinking," (p. 458) whereby the experts' worldviews were assumed to be representative of all thinking. They decided which terms would be used to describe things in a database, how those things were described and defined, and how they related to other things in the system or model of the phenomenon they aimed to represent. In doing so, they determined what could and could not be represented, therefore foreclosing interpretation and thus delimiting what becomes knowable in those particular information systems.

Some scholars now investigate the social and cultural processes that impact how knowledge gets ordered within research environments, governments, and on the Web. Some examine the design of ontologies (Illiadis, 2018; Leonelli, 2010; Ribes & Bowker, 2009), others semantic search engines (Ford & Graham, 2016; Waller, 2016), and still others the Semantic Web (Halford, Pope, & Weal, 2013). They study the processes by which complex ideas are reduced into machine-processable data and how designers negotiate contested facts to fit into prescribed information infrastructures. The literature echoes decades of scholarship in critical semiotics and linguistic anthropology that discusses how language is a reductive and constantly negotiated semiotic infrastructure (Gal & Irvine, 1995), one that is ordered and enforced by those in power but also constantly negotiated and evolving (Milroy & Milroy, 1985).

Methodology

The case study presented in this article is the result of three years of ethnographic fieldwork for a research project examining how the community of researchers designing standards for encoding the meaning of data confront the politics of representation in their work. It examines semiotic infrastructure design discourse and work within three communities: first, the Semantic Web community, which comprises practitioners designing frameworks for structuring and encoding the meaning of data on the World Wide Web; second, technologists and researchers designing standards and protocols for data sharing as part of the Research Data Alliance (RDA); and third, a community of practitioners designing languages and models for describing and ordering data about the accessibility of human services and the criteria by which individuals become eligible for them.

Eighteen semi-structured interviews were conducted with lead semiotic infrastructure designers in each of these three communities; designers were encouraged to elaborate on the history of ideas that influence their design approaches, the debates that emerge in their design communities, and the trade-offs they confront in their work. I conducted several years of participant observation research at Semantic Web conferences and Research Data Alliance (RDA) plenaries, examining how designers in these communities addressed critiques of their work. Finally, the article references several decades of digitally archived email correspondence between semiotic infrastructure designers, where the design of semantic frameworks, languages, and schemas such as the Resource Description Framework (RDF), OWL, and Schema.org was deliberated.

Approaches to semiotic formality

In the communities that were studied, formality as a design approach was consistently debated. In the late 1990s, as data infrastructure designers worked to establish Dublin Core as an international metadata standard, there were many heated arguments between two camps that were often referred to as the minimalists and the structuralists (Weibel, Iannella, & Cathro, 1997). Minimalists argued that the schema needed to be restricted to terms for which the greatest number of people could agree upon a single, consistent definition. In achieving this broad semantic agreement, minimalists could ensure the greatest interoperability between dispersed datasets and minimize the possibility that the meaning of things would drift. Structuralists, on the other hand, argued for designing a more complex structure into the metadata schema so that diverse communities could leverage the schema more flexibly. This would enable those describing their data with the terms used in their domains to do so with the precision and specificity they needed to characterize the knowledge unique to their domain.

Similarly, in the design of Semantic Web languages, there were many arguments between what had historically been known in the AI community as the neats versus the scruffies (Poirier, 2018). On the one hand, neats argued for clean, consistent, and restrictive data languages; they often prioritized the ability for automated systems to reliably produce correct answers about data above other design considerations, such as modelling the complexity of relationships between data. Their design approach was to restrict the number of primitives available to classify the relationships between concepts, formalize definitions for how and when these primitives should be used, and model relationships between concepts based on the rules of first-order logic.

Scruffies, on the other hand, argued that because diverse communities might disagree on the meanings of terms, and because the relationships between concepts could be so messy and at times paradoxical, the languages for contextualizing the meaning of things needed to be able to represent complexity. They advocated for designing semiotic infrastructures that maintained flexibility, where the meaning and order of concepts could evolve as diverse users leveraged the infrastructure to represent diverse datasets. This contrasted with neats, where meaning was hard-coded into the schemas and ontologies ahead of time. To achieve this, scruffies had to sacrifice the guarantee that automated systems could consistently reason with concepts. During an October 2016 interview, Ora Lassila, one of the original co-authors of an article introducing the Semantic Web[4] (Berners-Lee, Hendler, & Lassila, 2001), explained that the success of the Semantic Web could be attributed to this design approach:

> Standardization as a means of achieving interoperability is that we decide in advance what are the possible things to say and what do they mean—so that two systems can then talk to one another. … The Semantic Web takes a different approach. We don't actually try to standardize what are the possible things that you can say. We only standardize how to say them. And we give a framework that allows me to give you some clues as to how to interpret the things that I'm saying that you have not heard before, which I refer to as delayed semantic commitment. So standardization has this disadvantage in that it tends to be a limiting thing for technology because you cannot anticipate everything in advance, and if you haven't anticipated those things then maybe they cannot be done because the standard precludes them. And Semantic Web, particularly RDF was constructed to be free of this limitation.

Standardization stabilizes meaning by restricting what data modellers can represent as they describe data points and model the relationships between them. For Lassila, however, this was not the goal of the Semantic Web. Knowledge representation always delimits what becomes knowable, but the strategy that Lassila describes here pushes against this limit by opening standards up to diverse viewpoints and iterations in meaning. Lassila's attention to the limits of standards signifies a significant departure from "I am the world" (Forsythe, p. 485) thinking.

Semiotic infrastructure designers have attempted to design standards and frameworks for categorizing and organizing concepts within datasets and databases so that they can more readily be retrieved and for modelling the meaning of individual data points so that they can more readily be shared. However, developing a "shared language" is not only challenging because diverse communities apply such diverse definitions to their words or because their words are polysemous; it is also challenging because the more semiotic infrastructure designers standardize and formalize definitions, the more they restrict the complexity and diversity of what can be represented and the more they foreclose the possibility for meaning to evolve. The following section describes how semiotic infrastructure designers are confronting the limits of formal representation in the face of these challenges.

Double binds of representing difference

There are important reasons to contextualize concepts and to preserve their meaning across geographic, disciplinary, and cultural boundaries. In many social science fields that collect and interpret empirical data, researchers need to take great care to describe their datasets (such as interviews, images, and historical documents) because interpreting these data out of context can harm the communities represented within them. Government agencies and human service organizations that have traditionally worked in silos and often have diverse definitions for the same words need semiotic infrastructure to structure and share data in ways that enable them to more effectively serve their constituents and clients, and to communicate with each other. For instance, specialists at human service referral centres need to be able to quickly reference their service databases to direct individuals to programs they are eligible for. However, eligibility definitions for words such as homelessness, food insecurity, poverty, and gender vary by region, government agency, and culture. Without a standard way to interpret eligibility data, specialists may direct individuals with complex and sensitive issues to services or programs they are not eligible for. This could lead them to needlessly take a day out of work, potentially expose sensitive information (such as their immigration status), or it could foil an attempt to escape an abusive relationship. In such domains, restricting how words and meanings should be interpreted can advance ethical research practices, good governance, and ensure the safety of individuals. However, keeping words standard enough to advance common understanding and interpretation among dispersed communities requires devising formal languages—languages that, in turn, restrict how diversity can be represented and foreclose the possibility of representing what is "Other" to the language.

Semiotic infrastructure designers confront this challenge as they are called upon to design semiotic infrastructures that can standardize meaning across heterogeneous domains, including the World Wide Web, the social sciences, or human services. In these domains, words and definitions are not standardized and often highly politicized. In the process of modelling sameness and difference with these diverse domains in mind, many semiotic infrastructure designers are forced to rethink what it means to formalize meaning and are learning to design more flexibility and complexity into their neat, consistent, and restricted data languages.

For example, I interviewed Christian Bizer, the co-founder of DBpedia, a project to extract information from Wikipedia and structure it with semantics, at the 25th World Wide Web Conference in October 2016. When asked if he saw limits to achieving global data integration with Semantic Web technologies, Bizer responded, "Usually as a human, if I [ask]: Is a village and a tunnel the same? Or is a populated place and a tunnel the same? You would say no." After I nodded, he went on to describe this as class disjointness, which is a logic rule in many ontology languages (including OWL) that is used to indicate that two categories are disjointed, or never overlap. Data modellers use the class disjointness rule to restrict their ontologies, making it easier for automated systems to figure out which information to consider and exclude when asked questions. Bizer went on:

> A tunnel is not a populated place. [But] if you look at reality, or even if you look into Wikipedia, you find that there's a tunnel in India that contains a slum, so a tunnel is a populated place. It violates your logical assumption, but still the logical assumption is quite useful. So if you want to cleanse web data, even though it's only 98% or 99% true, the class disjointness helps you. But there are cases, which are true which violate the axiom. So basically, I think the Semantic Web community thought for a long time that things would be easy, but now as we look at reality, as this example nicely illustrated, it turns out that things are not as easy as we hoped.

Semiotic infrastructure designers constantly encounter exceptions to the rules their languages encode. Even when they explicitly aim to encode a diversity of concepts and ideas, they are constantly confronted with more concepts and ideas that do not fit into their identified categories. They are finding that applying the formal approaches to encoding sameness and difference theorized in analytical philosophy cannot help them escape this bind.

Grounded in other philosophical traditions, critical feminist and post-colonial scholars, such as Edward Said (1979), Gayatri Spivak (1993), and Teresa de Lauretis (1984), have drawn attention to the double binds[5] of representing differences since the 1970s. In order to acknowledge diversity, exceptions, or what is typically "Other" to dominant discourse, analysts need to identify, name, and define that difference; they must stabilize it as a common category. To be able to talk about what is "Other" they must have a way to refer to it. However, as soon as they stabilize a category by encoding it with a sign, they represent the category as homogenous, eclipsing difference within it or difference that may emerge from it.

Spivak (1990) acknowledged this as she described the challenge of representing "woman," "worker," and the "proletariat" in dominant discourse. She argued that while using these words to describe a group of people can be violent (because no one can point to a "true" woman, worker, or proletariat), we certainly do not want to exclude them from representation. She argued that all naming, defining, and categorizing can be categorized as "catachresis"[6] (p. 104)—or the (often forced) misuse of a signifier to represent a concept. In using this term, she drew inspiration from Jacques Derrida (1982), who wrote that "catachresis" "concerns first the violent and forced, abusive inscription of a sign, the imposition of a sign upon a meaning which did not yet have its own proper sign in language" (p. 255).

In designing standard languages for naming, describing, and ordering data, semiotic infrastructure design can be categorized as catachresis. As semiotic infrastructure designers work to formalize definitions for those things that do not fit into their existing definitions, they persistently eclipse more things that do not fit into the formal definitions. This is inevitable; as Drucilla Cornell (1992) argues, there is always an "Other" to a system, and it would be unethical to assume that this can ever be fully overcome. While we need semiotic infrastructure to communicate (for instance, you could not read this article if we had not shared some exposure to infrastructure that defines the words on this page), the words and primitives semiotic infrastructure designers define will never cover the full scope of knowledge and will always mean more than what the designers intend.

For those approaching semiotic infrastructure design with the explicit goal of representing difference in fair and just ways, figuring out how much complexity to represent versus how much complexity to restrict is not straightforward. Often naming a difference can do just as much violence as restricting a language to only represent certain differences. The challenge of designing just semiotic infrastructure demands considering when it makes a political, ethical, and practical difference to encode difference. In pursuit of this aim, semiotic infrastructure designers in certain domains are beginning to approach their work with a different cognitive lens than the analytical lens that has historically guided knowledge representation work; they are learning to attune themselves to the ways in which their knowledge representation systems are embroiled in a politics of representation and in doing so, bringing new mindsets about identity and difference to their work.

Conclusion

As sharing and analyzing data become increasingly pervasive, it has never been more important to ensure that concepts represented in data are robustly contextualized. However, even in the design of data infrastructure, there are notable politics and pragmatics to representing differences through standard languages for naming, describing, and ordering the world. Semiotic infrastructure designers have no choice but to consider trade-offs when it comes to deciding upon the complexity/restrictiveness, flexibility/hardness, scruffiness/neatness of their standards and frameworks. To make ethical and just decisions when evaluating these options, semiotic infrastructure designers need skill, not just in finding the edges of terms or logically modelling meaning but also in discerning when and why it is important to represent complexity and when and why it is important to foster shared or restrictive language. Scholarship examining and critiquing the design of such information infrastructures has admirably characterized the political consequences of formalizing definitions and reducing knowledge to fit them; however, there are ripe opportunities to examine how designers are confronting the problematic politics of representation their data systems produce and how they are devising creative solutions to enduring the binds of representing difference. All representations, including databased representations, are forms of power, so it is critical that scholars examining the social and cultural consequences of data systems pay attention to the work that shapes how difference gets demarcated within them.

Notes

1. For several decades, work in science and technology studies has shown how signs can become intertwined in technologies. See, for example, Keith Grint and Steve Woolgar (1997) on how technologies can be treated as texts.

2. For a more comprehensive list of technologies that could be categorized as semiotic infrastructures, see Archana Patel and Sarika Jain (2019).

3. In knowledge representation, expressivity refers to the breadth and complexity of ideas that can be represented in a knowledge representation language.

4. While work on the Semantic Web began in the mid-1990s, one of the first publications about the Semantic Web was published in 2001 (Berners-Lee, Hendler, & Lassila, 2001).

5. For Gregory Bateson (1972), double binds emerge when an individual is exposed to irreconcilable competing injunctions.

6. Spivak (1993) acknowledges that defining catachresis is itself catachresis.

Acknowledgement

The authors would like to thank Rensselaer Polytechnic Institute's Humanities, Arts, and Social Sciences Fellowship for the funding of this project.

Websites

DBpedia, wiki.dbpedia.org
Dublin Core Metadata Initiative, dublincore.org
Research Data Alliance, rd-alliance.org
Schema.org, schema.org
Web Ontology Language, w3.org/OWL/

References

Adam, Alison. (1998). *Artificial knowing: Gender and the thinking machine.* London, UK & New York, NY: Routledge.

Bateson, Gregory. (1972). *Steps to an ecology of mind.* Chicago, IL: University of Chicago Press.

Berners-Lee, Tim, Hendler, James, & Lassila, Ora. (2001). The Semantic Web. *Scientific American, 284*(5), 29–37.

Bowker, Geoffrey, & Star, Susan Leigh. (1999). *Sorting things out: Classification and its consequences.* Cambridge, MA: MIT Press.

Chomsky, Noam. (1957). *Syntactic structures.* The Hague, NL: Mouton.

Cornell, Drucilla. (1992). *The philosophy of the limit.* New York, NY: Routledge, Chapman & Hall, Incorporated.

de Lauretis, Teresa. (1984). *Alice doesn't: Feminism, semiotics, cinema.* Bloomington, IN: Indiana University Press.

Derrida, Jacques. (1982). *Positions.* Chicago, IL: University of Chicago Press.

Ford, Heather, & Graham, Mark. (2016). Semantic cities: Coded geopolitics and the rise of the semantic web. In R. Kitchin & S. Perng (Eds.), *Code and the city* (pp. 200–214). Abingdon, UK: Routledge.

Forsythe, Diana. (1993). Engineering knowledge: The construction of knowledge in artificial intelligence. *Social Studies of Science, 23*(3), 445–477. URL: https://doi.org/10.1177/030631279302 3003002

Gal, Susan, & Irvine, Judith. (1995). The boundaries of languages and disciplines: How ideologies construct difference. *Social Research, 62*(4), 967–1001.

Grint, Keith, & Woolgar, Steve. (1997). *The machine at work: Technology, work and organization.* Hoboken, NJ: John Wiley & Sons.

Halford, Susan, Pope, Catherine, & Weal, Mark. (2013). Digital futures? Sociological challenges and opportunities in the emergent semantic web. *Sociology, 47*(1), 173–89. URL: https://doi.org /10.1177/0038038512453798

Hayes, Patrick. (1979). The naïve physics manifesto. In D. Michie (Ed.), *Expert systems in the microelectronic age* (pp. 242–270). Edinburgh, UK: University Press.

Iliadis, Andrew. (2018). Algorithms, ontology, and social progress. *Global Media and Communication, 14*(2), 219–230. URL: https://doi.org/10.1177/1742766518776688

Kitchin, Rob, & Lauriault, Tracey P. (2018). Towards critical data studies: Charting and unpacking data assemblages and their work. In J. Thatcher, A. Shears, & J. Eckert (Eds.), *Thinking big data in geography* (pp. 3–20). Lincoln, NE: University of Nebraska Press.

Lenat, Douglas. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38*(11), 33–38. URL: https://doi.org/10.1145/219717.219745

Leonelli, Sabina. (2010). Documenting the emergence of bio-ontologies: Or, why researching bioinformatics requires HPSSB. *History and Philosophy of the Life Sciences, 32*(1), 105–125.

Milroy, James, & Milroy, Lesley. (1985). *Authority in language: Investigating language prescription and standardisation.* New York, NY: Routledge & Kegan Paul.

Minsky, Marvin. (1974, June 1). *A framework for representing knowledge.* URL: http://dspace.mit.edu /handle/1721.1/6089 [October 15, 2015].

Patel, Archana, & Jain, Sarika. (2019). Present and future of semantic web technologies: A research statement. *International Journal of Computers and Applications*, 1–10. URL: https:// doi.org/10.1080/1206212X.2019.1570666

Peirce, Charles. (1960). *Collected papers of Ch. S. Peirce* (1931-1935). Cambridge, MA: Harvard University Press.

Poirier, Lindsay. (2018). Making the web meaningful: A history of web semantics. In N. Brugger & I. Milligan (Eds.), *The SAGE handbook of web history* (pp. 256–269). Thousand Oaks, CA: SAGE Publishing.

Ribes, David, & Bowker, Geoffrey. (2009). Between meaning and machine: Learning to represent the knowledge of communities. *Information and Organization, 19*(4), 199–217. URL: https://doi.org/10.1016/j.infoandorg.2009.04.001

Said, Edward. (1979). Orientalism. New York, NY: Vintage.
Sowa, John. (2000). *Knowledge representation: Logical, philosophical and computational foundations.* Pacific Grove, CA: Brooks Cole Publishing Co.
Spivak, Gayatri. (1993). *Outside in the teaching machine.* London, UK: Psychology Press.
Spivak, Gayatri & Harasym, Sarah. (1990). *The post-colonial critic: Interviews, strategies, dialogues.* New York, NY: Routledge.
Star, Susan Leigh. (1995). The politics of formal representations: Wizards, gurus, and organizational complexity. In S.L. Star (Ed.), *Ecologies of knowledge: Work and politics in science and technology* (pp. 88–118). Albany, NY: SUNY Press.
Tarski, Alfred. (1944). The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research, 4*(3), 341–76. URL: https://doi.org/10.2307/2102968
Waller, Vivienne. (2016). Making knowledge machine-processable: Some implications of general semantic search. *Behaviour & Information Technology,* 35(10), 1–12. URL: https://doi.org/10.1080/0144929X.2016.1183710
Weibel, Stuart, Iannella, Renato, & Cathro, Warwick. (1997, June). *The 4th Dublin Core Metadata Workshop Report.* D-Lib Magazine. http://mirror.dlib.org/dlib/june97/metadata/06weibel.html. [March 14, 2019].
Wittgenstein, Ludwig. (1953). *Philosophical investigations.* Oxford, UK: Blackwell.