Joseph T. Tennis. (2004). "URIs and Intertextuality: Incumbent Philosophical Commitments in the Development of the Semantic Web." In *Knowledge Organization and the Global Information Society: Proceedings of the Eighth International ISKO Conference.* (London, United Kingdom, July 13-16, 2004). Advances in Knowledge Organization vol. 9. Würzburg: Ergon: 103-108.

PREPRINT

Joseph T. Tennis
Information School, University of Washington

URIs and Intertextuality: incumbent philosophical commitments in the development of the semantic web

Abstract: Examines two commitments inherent in Resource Description Framework (RDF): intertextuality and rationalism. After introducing how rationalism has been studied in knowledge organization, this paper then introduces the concept of bracketed-rationalism. This paper closes with a discussion of ramifications of intertextuality and bracketed rationalism on evaluation of RDF.

1. Introduction

The practices and technologies of the semantic web provide a robust landscape of philosophical commitments to knowledge organization. This paper explores rationalism, one of those philosophical commitments, in relationship to traditional discourse in library and information science. First this paper outlines the object of study: Resource Description Framework (RDF). It then introduces the standpoint of bracketed-rationalism. Finally this paper closes with the analysis of RDF from the bracketed-rationalistic viewpoint. Central to this analysis is the intertextual nature of Uniform Resource Identifiers (URIs) – that is the linking of resources and concepts through references. Through this analysis, this paper outlines the nature and problems of the philosophical commitments of URIs as an explicit intertextuality. Such an analysis has ramifications for the construction and evaluation of potentially global semantic web technologies built on RDF.

2. Definitions

The Resource Description Framework (RDF) is a framework for describing metadata. Metadata is used in the networked environment to organize knowledge, to manage digital objects, and to provide other functionality like access or preservation. RDF is expressed as a graph relationship between three things: the Resource, a Property, and a Value. Each of these can be a Uniform Resource Identifier, or URI. The most common URI now is a URL, or Uniform Resource Locator. Intertextuality is "the principle whereby the textuality of any one text arises from a interaction with other texts." (De Beaugrande, 1980, 242). In this instance textuality is the composition of the text through linking other resources and concepts from other texts through hypertext links, citations, or quotations.

3. Resource Description Framework, Uniform Resource Identifiers, and Intertextuality

The first part of the paper introduces the Resource Description Framework as it relates to knowledge organization theory and practice. It then introduces URIs, Uniform Resource

Identifiers, a key technology for RDF. Finally, this section closes with a justification for viewing URIs as intertextuality.

3.1. Resource Description Framework

RDF is a framework that describes relationships between metadata. In this first section RDF is discussed in relationship to the classification theory literature of S. R. Ranganathan. From this context RDF and its intertextual nature surfaces as a logical extension to the work of these library science thinkers, an extension of the implicit assumptions of bracketed-rationality.

RDF is one of the many technologies that are part of the w3c's version of the semantic web. Like Ranganathan's classification theory, RDF outlines the methods necessary for converting an N-dimensional space into points in a knowledge organization system. For Ranganathan the N-dimensional space needs to be converted into a single line. For proponents of RDF, the N-dimensional space must be converted into machine readable relationships. Both of these frameworks, RDF and Ranganathan's classification theory, operate on certain assumptions. First, both frameworks assume that there are concepts that exist that can be represented in a knowledge organization structure, and secondly, that those relationships can be formalized in such a way as to reduce their complexity by establishing a set of logical relationships. RDF and Ranganathan's theory, because they share these two assumptions, are both rationalistic models of knowledge organization systems. Section 4 outlines rationalism in knowledge organization.

3.2. Uniform Resource Identifiers

A key technology for RDF is Uniform Resource Identifiers. These, more commonly seen today as Uniform Resource Locators, or URLs, are the anchors for concepts, documents, and links among these in the networked environment.

3.3. Intertextuality

Intertextuality, being the principles whereby the textuality of one text is established by relationships and interactions with other texts (De Beaugrande, 1980, 242), is present in every type of text. It is present in novels through allusion. It is present in scientific and scholarly texts through citation. It is present in the web through URIs, and it is present in controlled vocabularies through various types of warrant. What is the intertextuality of URI's? URI's link all types of resources together in the networked environment. One straightforward way to see this is to observe hyperlinks. These establish a relationship and an interaction between two internet resources – in this case most likely web pages. In more advanced applications URIs can pull together meaning from two different controlled vocabularies and meaning from the Dublin Core Element Set. Because URIs can link the Dublin Core Element set and these controlled vocabularies to this particular application URI's establish an intertextual relationship between among these resources.

Joseph T. Tennis. (2004). "URIs and Intertextuality: Incumbent Philosophical Commitments in the Development of the Semantic Web." In *Knowledge Organization and the Global Information Society: Proceedings of the Eighth International ISKO Conference.* (London, United Kingdom, July 13-16, 2004). Advances in Knowledge Organization vol. 9. Würzburg: Ergon: 103-108.

4. Rationalism in Knowledge Organization

Birger Hjørland outlines four epistemological approaches to knowledge organization (Hjørland, 1999, 163). To Hjørland, all knowledge organization structures and practices are built on rationalistic, empirical, historic, or pragmatic assumptions. Pragmatic assumptions base knowledge organization structures and practices on use. Historic assumptions base knowledge organization structures and practices on precedence including past practice and established structures. Knowledge organization structures and practices built on perceptible evidence, like text corpora or empirical findings from user studies is an approach based on what Hjørland calls an empirical approach. The fourth approach is a rationalistic approach.

Hjørland outlines at least two aspects of the rationalistic approach to knowledge organization. The first aspect is existence of fundamental categories and the second is the logical process of division and aggregation along prescribed steps. The next section discusses these to aspects in detail.

4.1. Fundamental Categories

Fundamental categories are one aspect of the rationalistic approach to knowledge organization. Fundamental categories are peppered in throughout many theories and implementations of knowledge organization structures. The fundamental categories example *par excellence* is the work of S. R. Ranganathan's PMEST – Personality, Matter, Energy, Space, and Time (Ranganathan, 1967). Another example is D. W. Langridge's subject analysis categories (Langridge, 1989). The major criticism of a rationalistic view of knowledge organization is its assumption that knowledge falls into basic categories a priori. Opponents of rationalism feel that this approach fails to account for use and empirical evidence of texts in knowledge organization. However, there are two problems with this criticism. It appears that these critics fail to see the various levels of knowledge organization practice and theory. First, there is confused relationship between facets and foci in the PMEST theory of Ranganathan. The second is the clouded relationship between rationalism and empiricism, especially in knowledge organization structures and processes. The former argues that even when we find foci for use in controlled vocabularies we are somehow not finding categories (fundamental or not) of things. The second ignores a historical problem between schools of rationalism and schools of empiricism. Simply put the problem between the two schools is the fact that categories (fundamental or not) and perception inhere in both schools of thought. Therefore it is not possible to have rationalism without some empiricism and vice versa.

4.2. Logical Division

Logical division is a method evident in most classification schemes. Yet it is a rationalistic method. To Hjørland it is a defining characteristic for rationalistic knowledge organization: that is to establish rules and then follow those rules in the organization of knowledge is a rationalistic method. There seems from his description of four approaches to knowledge organization another way to organize controlled vocabularies, besides a rational way. That is that an empirical or pragmatic way would in no way be rational. This seems unlikely in a common sense way, and if it is true, needs refining in the theoretical literature.

4.3. Rationalism in the networked environment

In light of Hjørland's four epistemological approaches to knowledge organization, S. R. Ranganathan's work is a rationalistic approach. Intertextuality, as a theory of presence in the network of meaning is rationalistic. If RDF is an intertextual technology and is also rationalistic, what then are the theoretical ramifications of a rationalistic model of a global information system? In what ways does it matter that RDF and Ranganathan's classification theory are rationalistic? Does it influence the design of information systems? Specifically, what influence does this have on the evaluation of semantic web applications? Is some flavor of bracketed-rationalism necessary for any and all work in knowledge organization, ranging from classification schemes and including the semantic web?

5. Bracketed-Rationalism

Bracketed-rationalism is always present in knowledge organization structures. It is an inherent component of knowledge organization systems. The assumptions inherent in bracketed-rationalism have massive repercussions on the critiques of subject analysis, domain analysis, and controlled vocabulary construction. These assumptions also affect the how the semantic web can be critiqued. RDF, as one component technology of the semantic web, makes design commitments that are rationalistic. The critical question for systems design and evaluation is whether there is a significant difference between rationalism and bracketed rationalism in the instantiation of a knowledge organization system. That is, does it matter that RDF or S. R. Ranganathan's work seems rationalistic to a knowledge organization scholar? How might an information system be built that did not resemble a rationalistic or bracketed-rationalistic structure?

The commitments of bracketed-rationalism state that once created, a knowledge organization structure is always and necessarily a work of rationalism. That is, in order to arrive at a working structure for any knowledge organization tool like classification schemes, thesauri, or RDF descriptions, the designer and the user must accept that what is given is what, in this context, constitutes the fundamental categories and the logical division of the universe described. So formally, bracketed-rationalism assumes:

1. knowledge organization structures are built
2. once knowledge organization structures are built they operate, if not solely then mostly, on the assumptions of rationalism
3. the assumptions of rationalism are twofold: a utilization of fundamental categories and a utilization of logical division among parts of the knowledge organization structure
4. these assumptions may be based on studies or beliefs in other philosophical commitments (empiricism, historicism, or pragmatism for example)
5. but that these other assumptions do not reflect the resultant knowledge organization structure
6. which, as the result of a process, is a product that is rationalistic
7. and because a knowledge organization structure may not be totally built on the assumptions of rationalism, but may be built on a variety of assumptions and from variety of approaches, it is important to qualify the resultant structure as being bracketed-rationalistic.

Subject analysis, domain analysis, and controlled vocabulary construction will always have a rationalistic component. And because of this, any evaluation of a knowledge organization structure or process including subject analysis, domain analysis, and controlled vocabulary construction must account for bracketed-rationality as a given. This shifts Hjørland's analysis substantially. There are problems with isolating rationalism in his four-part division of knowledge organization epistemologies. We do not have a sophisticated enough evaluation or critique apparatus in his four-part division unless we incorporate the assumption of bracketed-rationality.

5.1. Example number one of bracketed-rationalism

I will give two examples and then move into a discussion of RDF as it relates to bracketed-rationalism. First is the example of universal classification schemes. In universal classification schemes various methods for maintenance and editing take place. Some are empirical, like using literary warrant. Other techniques are historic – based on past organizing schemes of subjects like history, philosophy, and social sciences. However, the resultant structure, e.g. the Dewey Decimal Classification or the Universal Decimal Classification, operates as a structure that commits to categories (and in some cases they can be called fundamental categories) and logical division. The major classification schemes, then in this analysis are bracketed-rationalistic schemes.

5.2. Example number two of bracketed-rationalism

Another example is the subject analysis provided by Birger Hjørland in his 1997 text (Hjørland, 1997, 93-98). This text provides two examples of subject analysis from an activity-theoretic point of view. This would be a pragmatic approach according to the four-part epistemological analysis above. Yet the resultant structure of what Hjørland provides is a description that assumes known categories (within a domain) and follows a logical ordering of concepts along those categories.

5.3. RDF and bracketed rationalism

RDF follows the same structure. It is a structure that assumes basic categories and a consequent logical division and aggregation of those categories. RDF is based on a three-part model: Resource, Property, and Value. It is assumed in the model that all knowledge falls into these categories. The model also assumes that these are not mutually exclusive categories. Resources, Properties, and Values can be grouped in various ways, provided they follow the rules associated with marking these up in the networked environment. That specifically means placing each of these conceptual entities at a place in the network – located at a URI. Linking URIs with other URIs in the Resource Description Framework is an act of intertextuality. It is also explicitly an act of bracketed-rationality. It assumes that representations of knowledge can fit into these conceptual categories and be linked with one another in a logical manner.

6. Intertextuality, URIs, and RDF

The intertextual nature of URIs is a foundational assumption to RDF. Linking representations of concepts, actions, and an information professional's analysis of these concepts and actions is a basic function of RDF. Concepts and actions, the results of analysis, are texts in the network environment. And each text of RDF is linked to another text of RDF through URIs. Each URI represents a place in the network where a concept is reserved. And RDF places these concepts, represented by URIs, into explicit relationships with one another.

7. Ramifications of Intertextual Bracketed-Rationalistic RDF

Why does it matter that RDF is a bracketed-rationalistic intertextual technology? In investigating the epistemological foundations of knowledge organization structures, and specifically semantic web technologies and practices, classification theory must account for bracketed-rationalism. That is, knowledge organization structures, specifically RDF, represent categories and logical relationships among those categories. And if we are going to advocate a user-based or use-based knowledge organization philosophy, then we must understand what that means for designing structures. There are two basic issues that spring from this analysis: 1) authoritative texts and 2) user-guided structures and categories. I will address the latter first.

7.1. User-guided bracketed-rationalistic structures

If bracketed-rationality is inherent in RDF and other knowledge organization structures, what remains to be included in the discussion is the question of how the categories are chosen and how the logical division should proceed. It is here that empirical, pragmatic, historical, task-based, and other approaches enter into the work of RDF and knowledge organization structures. The source of categories and the source of criteria for division happen from a data, use, or historical perspective, not before. And this should be based on evidence.

7.2. Authoritative texts

The second component here is the authority of the texts involved in RDF intertextuality. Why would we trust an analysis or representation of a concept fixed to a URI? What evidence do we have for working with this concept instead of that concept? Acknowledging the textuality and intertextuality of RDF offers designers a platform to create authoritative representations based on evidence of use, users, institutional or collectivistic authority, all represented in texts, linked to texts through URIs in RDF.

8. Conclusion

To the extent that theory influences system design, it is important to understand implicit methodological commitments in order to evaluate resultant knowledge organization systems. To that end in order to evaluate the structures and processes established by technologies like RDF and theoretical frameworks like Ranganathan’s it is important to study the incumbent philosophical commitments.

The most basic finding from this analysis is that information systems are all, at least in part, rationalistic. In order to evaluate the effectiveness of information systems it is important

for knowledge organization theory to accept this premise as an explicit statement of practice. Further, the ramifications of a rationalistic, or bracketed-rationalistic, semantic web must be understood. Perhaps our critiques should not be on bald and overly general categories of knowledge organization work, but with particular and finely differentiated approaches. Perhaps we should as knowledge organization theorists make explicit the various epistemic components of extant knowledge organization systems and then evaluate to what degree they work together with bracketed-rationalistic structures to aid information organization and access, not just at a local level, but at a global level as well.

References

De Beaugrande, R. (1980). *Text, discourse and process: toward a multi-disciplinary science of texts.* (Norwood, NJ: Ablex).

Hjørland, B. (1997). *Information seeking and subject representation: an activity theoretical approach to information science.* (Westport, CT: Greenwood).

Hjørland, B. (1999). The classification of psychology: a case study in the classification of a knowledge field. In *Knowledge Organization*. 25(4):162-201.

Langridge, D. W. (1989). *Subject analysis: principles and procedures.* (London: Bowker-Saur).

Ranganathan, S. R. (1967). *Prolegomena to library classification.* 3rd ed. (Bombay: Asia Publishing House).