

Organization knowledge on the digital network

Mary Eugenia Luna González *

*Paper submitted:
December 5, 2013.*

*Accepted:
October 9, 2014.*

ABSTRACT

This paper presents a documental analysis for the purpose of understanding how knowledge in digital document repertories is organized, specifically those hosted on the digital network. The study is carried out on the basis of an analysis of the knowledge within these information systems, in which the sociocultural context of the creator and the subset of values and beliefs shared by the personnel representing this creator necessarily intervene in conjunction with the informational needs of the end user of the information product offered by the knowledge organization. The paper also discusses the relationship existing between the user and the digital network knowledge retrieval system.

* Ministerio de Petróleo y Minería, Venezuela. malunag2000@yahoo.com

Keywords: Organization of the Knowledge; Context; Digital Network; Recovery of the knowledge; Representation.

RESUMEN

Organización del conocimiento en la red digital

Mary Eugenia Luna-González

Se realiza un análisis documental que permite conocer el estado de la organización del conocimiento en la documentación encontrada en la red digital, partiendo del análisis del conocimiento dentro de dichos sistemas informativos. Intervienen en ello el contexto sociocultural en el que se desenvuelve el creador del conocimiento y el conjunto de valores y creencias que comparte el personal que lo representa, así como las necesidades informativas de los usuarios finales de la información que es resultado de la organización del conocimiento. De igual manera se aborda la relación existente entre el usuario y los sistemas de recuperación del conocimiento en la red digital.

Palabras clave: Organización del Conocimiento; Contexto; Red Digital; Recuperación del Conocimiento; Representación.

INTRODUCTION

Over recent years, technological transformation and advancements have determined how information and knowledge are created, recorded and accessed. Using documental languages based on concepts and controlled or uncontrolled terminologies, the organization of knowledge serves as a representation of such knowledge and information. These languages depend on the reality of professionals in charge of representing said knowledge and on the socio-cultural milieu in which they work. The diverse contexts are determined by socio-economic, political and social-cultural circumstances of the professional, which means there will be diverse approaches to analysis, interpretation and representation of knowledge. The task of representing knowledge is a process performed by the document professional, who stands astride knowledge as an information product and the end user.

On this basis, the objective of this paper is to achieve an understanding of how knowledge is organized in the documental repertoires found online. For the purpose of establishing a conceptual framework for the topic under study, the methodology consists of documental examination and analysis of specialized literature, including both online and printed sources. This approach provides terminological mastery, allowing conclusions to be drawn about the organization of knowledge on the internet.

ORGANIZATION OF KNOWLEDGE

Within the information sciences, the discipline of knowledge organization takes on the task of studying the treatment of knowledge, and the construction and control of the language and instruments used in the process of representing documents that are the products of human knowledge, thereby allowing retrieval by users of these documents. This organization of knowledge is a process in which the document or information professional serves as an intermediary by creating the representation of the knowledge product for the benefit of the user.

Specifically, librarians, archivists and other specialists that organize knowledge create document descriptions, indexations, bibliographic data bases, archives and other types of “Institutional Memories” (Hjørland, 2008: 86). With regard to documents, information and knowledge, Vizcalla Alonso has stated that:

[The] organization of information has been directly associated with the stage known as information processing and has been framed within the traditional aspects of the same. Generically, these activities may be summed up as anything associated with the description of the form and content of the documents included in the system. (1997: 166)

Organization of knowledge establishes systems for representing the reality of the author of such knowledge for the purpose or allowing the end user to access it and understand the context in which the specific knowledge product was created. These systems contemplate documental languages, as cited by Civallero (2005: 1): “[documental languages] undergo transformation from the intermediary to the user, the former codifying both the content of a text the user’s search prompts.” Gil Urdiciain (1996: 1) defines documental language as: “any artificial, standardized system of signs used to facilitate the formal representation of content of documents and thereby

allowing such information to be retrieved either manually or automatically by the user.”

Since documental language is the intermediating factor between the information and the user, it appears inherently subjective to the user, because the representations made in the documental language are immersed in a context and reflect the ideologies of the knowledge organizers.

CONTEXT AND IDEOLOGY IN THE ORGANIZATION OF KNOWLEDGE

Documental professionals and others in charge of representing knowledge perform an intermediary task between the documentation and the final users. To this end they employ controlled languages that attempt to ensure neutrality, but in reality they provide records and summaries loaded with subjectivity, if for no other reason than they are human and have developed personally and professionally within a specific environment.

Context plays a fundamental role in the description and representation of written information, because it involves social and cultural facets. The language used in the discourse of authors of knowledge and that of final users varies in accord with ideologies and contextual relationship. Consequently, as Moreiro states:

The tasks of selecting and attributing terms by information professionals is based largely on the cultural context to which they belong and their experience as interlocutors, rather than on the needs of the user. These tasks are performed within the universe of possibilities for representing the selected concepts, ranging from controlled vocabularies through all of the possible circumstances that bring them to the decisions to include a representation term for the original content all the way to free languages. (cited from Bufrem, Silva and Breda (2005: 124. Translated from Spanish)

In this regard, Van Dijk (cited in Silva, 1997) states that all levels of discourse contain contextual fingerprints, such as gender, class, ethnicity, age, origin, position and other group identifiers; and these play a fundamental role in representation. In this way, the physical paradigm demands the professional mediator represent information in a scientific, controlled way, even when the primary source’s use of language is not deployed and structured in this way. Human nature, however, does not lend itself to neutral, objective transcriptions. These professionals, therefore, are prone to include elements from their reality in anything they

attempt to represent. Actual practice seems to contradict the idea of organizing information in terms of the user's needs to retrieve and exploit information.

Similarly, the practice of using controlled languages to confer neutrality on representations of knowledge relies on rigid guidelines of the physical paradigm. The crisis in the paradigm stands in high relief when we look at the "gulf that exists between the opinion of researcher and practitioner, who believe the work of our professionals is neutral and innocuous" (García Gutiérrez, 2001: 3. Translated from Spanish).

Since ideology determines the language used by the author and mediator, one might assert that it is understood as a subset of values and beliefs shared by the members of a given society and it is assimilated by the members unconsciously, thereby becoming their conception of reality, while serving to model the identity and inform the personality (Civallero, 2005).

This ideological *schemata* structures the opinion of persons about a wide range of specific topics, while also motivating the way they act. When negative connotations of ideas are presented, these persons may respond aggressively or discriminatively toward the groups they presume are making such asseverations.

As stated by Civallero (2005: 3. Translated from Spanish), in diverse controlled language tools, we can see that such negative connotations are considered unbiased:

A documental work tool used internationally includes, representative, unambiguous and relevant descriptors controlled by professionals to classify entire peoples and races as "colonial" or "primitive" (in contrast to "developed" or "highly developed," which are also included). These labels imply the establishment of mental and physical differences that widen the gap existing between human beings. In this way the politics of hate, marginalization, dominance and scorn are perpetuated, all of which are and have been the hallmarks of human history. What is perhaps even more worrisome is that these labels, these descriptors, have been deemed "neutral" by the society creating the documental language. What is more, these terms are deemed necessary, pertinent, important and objective for classification and understanding of certain human groups.

Since the information professional will always make decisions regarding the relevance of online information, and objective, neutral representations are not possible; the idea that such organization is actually performed on the basis of the needs users is not credible.

THE ORGANIZATION OF
KNOWLEDGE ONLINE

With the evolution of new IC technologies, representation and organization of knowledge, and storage and access processes have also undergone significant change.

One of these technologies is the digital network, a mechanistic invention developed to control the rapid growth of knowledge. It is understood as an open information system by which upload and update of knowledge to the representation system allows individuals and specialists to interact with each other and publish documents on specific topics. This process, at the end of the day, causes information overload and in fact can overload the digital system.

In this sense, the organization of knowledge on digital networks provides direct access through a single platform. Since the knowledge available in information systems consists of codified mediations of real products, these real products are decontextualized; and it falls to the user to learn how to refine search parameters in order to find relevant information.

The selection and analysis work of documental professionals to represent and disseminate information through material and technological supports is done for the purpose of creating a resources for source for users. Chacón Gutiérrez (1995) states that this process is abstract, in that it uses a series of codes in the storage, search and retrieval system; and it is anonymous, in that the analyst and final user do not know each other and are unaware of one and other's needs. Thus, the user delegates his search to another individual, who is unaware of the reasons motivating the search

The abstract procedure of organizing knowledge on the web, in which codes or metadata are assigned to the document in order to make it more explicit and easy to represent, index and subsequently search, is done without a set of standardized procedures. In terms of products, this can lead to less than complete representations. In this regard, Codina states that “[...] people make mistakes and webpage makers make mistakes: they forget to use metadata, they transcribe these erroneously, they use them on some pages and not on others, and they make spelling errors [...]” (2003: 151. Translated from Spanish).

In sum, indexing of intellectual production and knowledge on said servers is performed by a large number of experts in charge of selecting, simpli-

fyng and representing documentation published by specialists and authors writing on a given topic, which causes said information to be distorted or manipulated by the person performing the mediation task. In this way the original concepts may be changed by a mediating context that is distinct from the original context of the producer. This change can exert an effect on the results a user gets from the search process.

ORGANIZATION OF KNOWLEDGE IN DIGITAL INFORMATIONAL RETRIEVAL SYSTEMS

Results coming from search engines are the product of textual coincidences, rather than of content or the interrelation of said terms. In this sense, no logical retrieval mechanism exists, because the search is divorced from semantics. Consequently, the system links to terms that oversaturate and obstruct useful information retrieval often yielding without any relevance to the topics searched, or with genuine results buried under mountains of superfluous or inappropriate information. The user, then, will often abandon the search in the belief that these electronic systems are without scientific foundation and cannot meet the information demands with a reasonable investment of time and effort.

Indexes developed by documental personnel working for online information retrieval companies do not make distinctions between the diverse users they serve. On the other hand, the user secures information that has not been tailored to his needs because he enjoys free access to the data. According to García Gutiérrez (2001: 6. Translated from Spanish): “there is no innovation in the knowledge organization logic, which conservatively preserves the classic structures and it is the support itself that determines what is changed.” With regard to the documental languages used to index documentation in electronic search engine sites, these can be manipulated for any of the following reasons (Chacón Gutiérrez, 1995).

1. Like the creator of the documental language, the analyst is immersed in a society that imposes ideological and cultural models.
2. In intellectual operations, the documental mediator’s subjectivity influences the selection of document topics and key words or descriptors used in representation.
3. The documental language is codified and, even though the user and documental intermediary share documental codes, such codes are subject to the bias of the person performing the mediation.

4. As a process in which the natural language is converted to documental language, indexing may fail to reflect the implicit connotation of the original language and the socio-cultural context where it was produced. Likewise, the lack of knowledge of the indexer can have a negative influence on the choice of concepts and keywords used in the documentation.

García Gutiérrez (2001: 5-6. Translated from Spanish) states: “for now, the hypertext does not go beyond the old systematic or rotated indexes. It is nothing more than a chaotic transversal index with special effects.”

The user-system interaction in information searches is more productive when the user is capable of articulating his information needs in a way that coincides with the index representation language used. This means, of course, that such searches will be limited by the features of the representation language, because electronic online information retrieval systems use diverse processes to perform searches. One such method of retrieval is performed by means of logical search, in which Boolean operators are used to associate and combine terms in a logical way. These terms are generally *and*, *or* and *not*.

The conjunction or logical intersection *and* is used to associate two or more terms required to appear at the same time in the documents. The logical sum *or* is used to search documents containing one or all of the search terms. The logical negation *not* is used to exclude documents containing a given term (Lancaster and Pinto, 2001).

Likewise, searches carried out with information retrieval systems using free text, with words and phrases in natural language, yields term coincidences in the title and body text of the document. With this procedure, the strategy to achieve the search consists of phrases the user believes are relevant to his needs.

The advantage provided by the use of free, natural language for searches resides in being able to perform searches that are more narrowly targeted than searches using controlled language, while nonetheless reaping results with the user's phrases that coincide with the documental language of the professionals who perform the subjective representations that otherwise might interfere with the neutrality in the organization of knowledge.

NEUTRALITY IN THE ONLINE
ORGANIZATION OF KNOWLEDGE

There is a relationship between the socio-cultural context and the expression of knowledge represented in the documental discourse. This relationship is built by the individual participants in said discourse, with each individual having his or her own interpretation of the context in which he is immersed.

The persons processing the documental content online do not make distinctions between the client users. Since the user enjoys free access to the content of said data bases, he receives information that is not tailored to needs. This is because information professionals cannot make neutral, objective representations that are uninfluenced by the subjectivity inherent in their life and work realities.

In this regard, Chacón Gutiérrez (1995. Translated from Spanish) has made the following observation regarding description and indexing:

Because it is an intellectual activity, the analyst's subjectivity is decisive in the selection of topics contained in a document and in the selection of descriptors or key words to represent them. The documental professional acts as an end reader of the primary document, interprets the document he receives and sends the content (sender-intermediary) translated in the codified language (co-author).

This is why the documental analysis and indexing performed by documentation professionals entails simplified representations of the work of authors on specific topics in order to make them available to web users. Consequently, the documental analysis and representation may distort or manipulate information in the process of aligning terms and concepts with his or her reality.

With regard to manipulation that occurs in the indexing process when natural language is translated to documental language, Chacón Gutiérrez (1995. Translated from Spanish) states:

The secondary document translated from the primary document is an intermediary between the user and the original. If it has been manipulated during the indexing process, the user is also manipulated. The manipulation may arise from unresolved problems inherent in all translations, such as lack of fidelity to the

original; failure to find equivalencies between natural and documental languages; failure to convey implicit connotations of the original or those of the socio-cultural context of the author; or when the indexer's subjectivity or lack of inurnment negatively influence the selection of concepts existing in the document or key works or descriptors to represent these topics.

Influenced by their social cognitions, documental professionals read the primary document, assuming the standards, values, objectives and socially shared interests these contain in order to provide the foundation for drafting the secondary document. In this way they become co-authors of the final information made available to the user. "The data offered by a Documentation Center are often not verified by users. The users are subject to a double manipulation, that of the secondary document and that of the primary document on which the former is based" (Chacón Gutiérrez, 1995. Translated from Spanish). Similarly, Bufrem, Silva and Breda (2005: 123. Translated from Spanish) observe:

[...] when the mediator or interlocutor reader interprets a text for subsequent representation, he must conceive of it as socially situated and targeted at an exact moment, and he must accept that the structure of the enunciation is determined by the most immediate social situation and the larger social medium.

With regard to searches in the digital sphere, documentation is made searchable through representations in the search engine and the use of specific key words chosen from those available and belonging to a social subset. As such, the concept selection process performed by documental specialists or information professionals within a search website is based on their own socio-cultural context and experiences rather than on the needs of the users. As stated by Moreiro (cited in Bufrem, Silva and Breda, 2005), such tasks are performed within a universe of possibilities for the representation of selected concepts, spanning the continuum from controlled languages to free languages, and entailing circumstances that determine whether to include one term or another to represent the knowledge (Moreiro, cited in Bufrem, Silva and Breda, 2005).

Finally, when the user performs an information search, the coincidence of concepts (as distinct from content) in electronic retrieval systems yields irrelevant information, while ignoring content relevant to the search intention. These results interfere with the neutrality of the information represented.

CONCLUSIONS

1. Representation of knowledge is the manifestation of thought that is expressed through language, which allows how individuals reproduce thought as members of a given socio-cultural group to be deciphered.
2. Organization of knowledge is a field that plays a fundamental role in the general access to and exchange of knowledge and information. It rests on the technological networks used to retrieve represented information, though the terminology and content used in these sites must be entrusted to professionals in the field of information science.
3. The participants in the documental discourse build the relationship between the social-cultural context and its expression, while each mediator creates representations in accord with his particular context and individual reality, meaning such representation cannot be completely neutral.
4. Some measures are needed to control the documental representation tools that pretend to be neutral but which are in fact imbued with bias and/or negative or discriminatory connotations with regard to specific socio-cultural groups.
5. Traditional human activities are increasingly being performed using computerized media. Professionals in the area of information are focused on applying information and communication technologies and the tasks of interpretation and semantic analysis of content is often afforded a secondary status.
6. With the advent of the internet and as a function searches performed, the relationship between the knowledge producer and the end user has intensified. Searches in which the indexers employs both standardized and free languages are never free of the personal context, and this will for the most part cause results to be based on word coincidences rather than on semantics that best match the needs of the user.
7. Despite the advantages offered by the use of Boolean operators in searches of knowledge in digital information retrieval systems, users are often not very familiar with their correct use and combination. This causes the system to yield incorrect or unwanted information.
8. When using Boolean logic in interactive digital information systems, users often fail to represent information needs properly, which sometimes leads to the search failing to yield useful, relevant information.

9. Free searches can be more specific than those performed with controlled languages, but they often exert an effect on retrieval neutrality leading to irrelevant results and exclusion of relevant information.
10. The organization of content on the web is threatened when the personnel of the information indexing are unaware of user needs and injects personal bias into their analyses.

WORKS CITED

- Bufrem, L. S.; Silva H. F., and Breda S. M. (2005), "Reformulación de los elementos teóricos de la organización del conocimiento: bases lingüísticas y culturales y estructuras de representación." In *ISKO, capítulo español. Congreso 7º*, 121-127. Barcelona.
- Chacón Gutiérrez, I. (1995), "La mediación documental." *Cuadernos de Documentación Multimedia* 4. Accessed January 10, 2011. <http://www.ucm.es/info/multidoc/multidoc/revista/cuadern4/chacon.html>
- Civallero, E. (2005), "Espinass ocultas: lenguajes documentales, ideologías negativas y revisiones." *Fuentes del Congreso. Boletín de la Biblioteca y Archivo Histórico del H. Congreso Nacional de Bolivia* 4 (20). Accessed January 10, 2011. http://issuu.com/bib.csinfo/docs/civallero__2005
- Codina, L. (2003), "La web semántica: una visión crítica." *El Profesional de la Información* 12 (2) (May-April): 149-152.
- García Gutiérrez, A. (2001), "Redes digitales y exomemoria." Paper presented at *I Congreso Ibérico de Comunicación. La Sociedad de la Comunicación en el siglo XXI*. Universidad de Málaga, Málaga, May, 2001.
- Gil Urdiciain, B. (1996), *Manual de lenguajes documentales*. Madrid: Noesis.
- Hjørland, B. (2008), "What is Knowledge Organization (KO)?" *Know-ledge Organization* 35 (2-3): 86-101.
- Lancaster, F. (1995), *El control del vocabulario en la recuperación de información*. Valencia: Universidad de Valencia.
- Lancaster, F., and Pinto, M. (2001), *Procesamiento de la información científica*. Madrid: ArcoLibros.

- Silva, O. (1997), "El estudio del discurso en el camino de Teun Van Dijk." *Revista Frontera* 16: 97-106. Universidad de la Frontera, Temuco.
- Sojo, V. (2008), *Normas de la American Psychological Association (APA) para las citas y referencias bibliográficas*. Caracas: Escuela de Psicología, U.C.V.
- Van Dijk, T. (2001), "Algunos principios de una teoría del contexto." *Revista Latinoamericana de Estudios del Discurso* 1 (1): 69-81.
- Vizcalla Alonso, D. (1997), *Información: procesamiento de contenido*. Argentina: Parhadigma Ediciones.



Para citar este artículo como revista electrónica:

Luna González, Mary Eugenia. 2015. "Organización del conocimiento en la red digital". *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*. 67: 77-89. Aquí se agrega la dirección electrónica (Consultado el día-mes-año)

Para citar este artículo tomado de un servicio de información:

Luna González, Mary Eugenia. 2015. "Organización del conocimiento en la red digital". *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*. 67: 77-89. En: Aquí se agrega el nombre del servicio de información y la dirección electrónica (Consultado el día-mes-año)

