



## ARTICLE

## The thematic representation of NASA digital images on Flickr the contributions of knowledge organization systems

Danilo Camargo Dias <sup>1</sup>  <https://orcid.org/0000-0002-9838-8861>

Walter Moreira <sup>2</sup>  <https://orcid.org/0000-0001-9454-441X>

Rachel Cristina Vesú Alves <sup>3</sup>  <https://orcid.org/0000-0002-1649-3722>

<sup>1</sup> State University of São Paulo, Marília, SP, Brazil / e-mail: [danilo.dias0512@gmail.com](mailto:danilo.dias0512@gmail.com)

<sup>2</sup> State University of São Paulo, Marília, SP, Brazil / e-mail: [walter.moreira@marilia.unesp.br](mailto:walter.moreira@marilia.unesp.br)

<sup>3</sup> State University of São Paulo, Marília, SP, Brazil / e-mail: [rachel.vesu@unesp.br](mailto:rachel.vesu@unesp.br)

### ABSTRACT

The growth of social networks and the way of representing their resources has become the subject of several studies in Information Science (IS). The digital image carries a difficulty in relation to its thematic representation in these social networks. Thus, the problem with this study is to investigate the retrieval of information on social networks, especially on networks that make available digital images. The objective is to analyze the contributions of the Knowledge Organization System - KOS - of the National Aeronautics and Space Administration - NASA, in the thematic representation of digital images made available on Flickr's social networks. The exploratory and descriptive approach was used as methodology, which enabled the bibliographical review of theoretical aspects and the analysis of terminology in digital images at NASA's Flickr. Therefore, the terms of non-controlled thematic representation described by folksonomies, and the controlled terms from the NASA thesaurus were analyzed. As results of the analysis, the challenges generated by the use of folksonomy in the thematic representation of digital images, and the advantages of using KOS in the processes of thematic representation and information retrieval in these environments are highlighted. It is concluded that folksonomies freely represent terms and reflect the terminological reality of a community, while KOS' contribute to a control of vocabulary and more accurate information retrieval.

### KEYWORDS

Knowledge organization systems. Digital images. Folksonomy. Information retrieval.

## A representação temática de imagens digitais da NASA no Flickr as contribuições dos sistemas de organização do conhecimento

### RESUMO

O crescimento das redes sociais e a forma de representar os seus recursos tornou-se objeto de diversos estudos na Ciência da Informação (CI). A imagem digital carrega uma dificuldade em relação a sua representação temática nessas redes sociais. Assim, o problema deste estudo consiste em investigar a recuperação da informação nas redes sociais, principalmente em redes que disponibilizam imagens digitais. Objetiva-se analisar as contribuições do Sistema de Organização do Conhecimento – SOC – da *National Aeronautics and Space Administration* - NASA, na representação temática de imagens digitais disponibilizadas em redes sociais do Flickr. Utilizou-se como metodologia a abordagem exploratória e descritiva, o que oportunizou o levantamento bibliográfico dos aspectos

teóricos e a análise da terminologia em imagens digitais na NASA do Flickr. Por isso, foram analisados os termos de representação temática não controlados descritos por folksonomias e os termos controlados advindos do tesouro da NASA. Como resultados da análise destaca-se os desafios gerados no uso da folksonomia na representação temática de imagens digitais e as vantagens do uso do SOC nos processos de representação temática e recuperação da informação nesses ambientes. Conclui-se que as folksonomias representam livremente os termos e refletem a realidade terminológica de uma comunidade; enquanto que os SOC contribuem para um controle de vocabulário e recuperação da informação mais precisa.

**PALAVRAS-CHAVE**

Sistemas de organização do conhecimento. Imagens digitais. Folksonomia. Recuperação da informação.



**JITA:** ID. Knowledge representation.

## 1 INTRODUCTION

The evolution of technologies, whether in relation to hardware, software or means of data transmission, has brought about a social and cultural change. People, companies, public agencies and non-governmental organizations started to share texts, images, audios and videos on social networks. Although this reality has given rise to a new form of social representation, folksonomy and Social Indexing, in which users or consumers of resources are responsible for making thematic representation of resources (HASSAN-MONTERO, 2006), the field of studies of Information Science - IS, is faced with new informational environments that can make use of knowledge representation instruments.

Social networks are highly dynamic objects, they grow and change rapidly over time through the addition of new relations, signifying the emergence of new interactions in the underlying social structure (LIBEN-NOWELL; KLEINBERG, 2007). Flickr, for example, is focused on making digital images available, the representation process can entail some inconveniences in the process of searching and retrieving these resources. The inclusion of many terms in the thematic representation of a digital image, attributed by the user of the social network, does not guarantee a good search and retrieval. The problem is the lack of criteria when selecting terms for representations, since in a given resource there may be many indexing terms while in others these terms are minimal or do not really describe the resource, that in this case is a digital image.

Hassan-Montero (2006) highlights the following point of view in relation to the use of natural language in thematic representations such as folksonomy and Social Indexing:

[...] is done in a non-controlled natural language, and in most cases with selfish motives, there will be a significant proportion of empty terms or tags of collective significance. In other words, the user will often assign tags that only make sense to himself, to which we must add the known problems of polysemy and synonymy.

| 3

The use of non-controlled and established by users terms was denominated folksonomy. For Wal (2006), folksonomies can be defined as “organic systems based on the free and personal labeling of information or objects aiming at organization and retrieval”.

According to Ferneda (2012, p.13), “the function of retrieval is to represent the content of the corpus documents and present them to the user in a way that allows him to quickly select items that fully or partially satisfy his need for information [...]”.

The retrieval of information in environments whose representation was not controlled by any representation instrument, as occurs in social networks, presents a reduced efficiency of the process.

Thus, it is possible to find problems in relation to the retrieval of information on social networks, especially in networks that provide digital images that, by themselves, carry a certain difficulty in relation to their own representation.

Hence, the objective of this article is to analyze the contributions of the Knowledge Organization System - KOS - of the National Aeronautics and Space Administration - NASA, in the thematic representation of digital images made available on Flickr social networks.

The exploratory and descriptive approach was used as methodology, which enabled the bibliographical review of theoretical aspects and the analysis of terminology in digital images at NASA's Flickr. Thus, the non-controlled thematic representation terms described by folksonomies and the controlled terms from the NASA thesaurus were analyzed.

Representations made by NASA were used as samples, which is a world reference in aerospace images, on Flickr, in which initially non-controlled terms were applied and, subsequently, terms of the domain thesaurus. Guimarães (2014, p.16) presents domain as the categorization inherent to the organization of knowledge, in so far as from the identification of a set of common features, it is possible to bring together similar things and separate different ones.

## 2 KOS AND THEMATIC REPRESENTATION

When it comes to the term knowledge, it is impossible to ignore the fact that it is constantly evolving, increasing and undergoing adaptations, as necessary. The acquired knowledge is related to something previously stored in the individual's memory, something that is capable of understanding and drawing conclusions. After that, it can be transformed into information, conveyed by different sources, and again in knowledge, the subject can then classify it.

Registering knowledge is essential for scientific and societal advances. However, as knowledge is incremental and changeable, it is very important that there is some way of organizing such knowledge. One of IS areas of study is how to register knowledge in a structured manner, with the aim of making the information retrieval process more efficient. According to Hjørland (2003, p.87), organizing and retrieving records is the basis of IS:

[...] Knowledge Organization means especially the organization of information in bibliographic records, including citation indexes, full-text records and the Internet. Information Science is basically about the best way to build such bibliographic records, as well as the ideal way to use the records provided.

| 4

Knowledge Representation Schemes are tools created to improve the process of organizing and representing the contents of information resources, with the aim of facilitating the information retrieval process. Carlan and Medeiros (2011, p. 55) state that:

[...] knowledge representation schemes are instruments that translate the contents of original and complete documents, into a systematically structured scheme [...] with the main purpose of organizing information and knowledge and [...] facilitate the retrieval of the information contained in the documents.

When looking for a way to organize the record of the content of information resources, implicitly, what is intended is to create a logical structure that facilitates the thematic retrieval of what is desired, in an efficient and univocal way, among other aspects. In the following topics, thesaurus, ontology and folksonomy tools are discussed, their meanings and how they can contribute to the process of thematic representation of digital images.

### 2.1 Thesaurus

The thesaurus is one of the KOS tools that provides IS professionals, and related areas, with a set of standardized terms, by vocabulary control, to improve the process of registering knowledge and retrieving information. Regarding the definition of thesauri, Cervantes (2009) reports that the thesaurus can be seen as a mediation tool capable of allowing index librarians,

reference and information service librarians and users of a specialized information system to share the same vocabulary.

However, the terms contained in a given thesaurus can vary in meaning depending on the domain in which they are used, but they can be understood from the area of knowledge to which they belong. An example is the term “mango”, which in the thesaurus of the Brazilian Agricultural Research Corporation (EMBRAPA), refers to the fruit of the mango tree. On the other hand, if applied to the clothing thesaurus, it refers to the part of clothing that covers the arms.

Several thesauri can be found on the Web, including the Ministry of Infrastructure, Folklore and Popular Brazilian Culture, the Supreme Federal Court, the Superior Electoral Court, the Thesaurus for Cultural Heritage Objects in Brazilian Museums, among others. In addition to nationals, there are several other thesauri such as that of the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations Bibliographic Information System (UNBIS) and NASA. The latter selected to be used as a reference during the developed study.

## 2.2 Ontology

An ontology represents in a structured and logical way the concepts, relationships and restrictions that exist between them. According to Ferneda and Dias (2017, p. 175) “[...] ontology is a conceptual structure that aims to formally represent the concepts and their relations, rules and logical restrictions of a given domain”.

According to Jacob (2003, p.20), ontologies are not new on the Web. Any metadata scheme is an ontology that specifies the set of physical and/or conceptual features of resources that have been considered relevant for a given community of users.

For Silva, Santos and Ferneda (2013, p. 28), “the importance of ontologies in information retrieval is to provide and expand the user's search for a wider range of possible relevant options”.

As social networks do not impose any type of vocabulary restriction during the resource description process, the use of domain ontologies could provide a range of meanings to the terms applied in different situations.

Once the domain is defined, this tool can be used as a reference for the development of automated computer systems, with the objective of being the guiding model for understanding, comprehension, of the concepts and terms applied in the representation, regardless of the domain.

## 2.3 Folksonomy

In the social scenario in which many people, companies and public institutions have access to computer equipment and media, the possibility of creating new forms of data records on the Web appears as a consequence of these changes. Slang, symbols, neologisms, abbreviations and various forms of expression can be used to represent something.

The word folksonomy was created by Thomas Vander Wal in 2004, when he mixed the words folk (people) with taxonomy (taxonomy), that is, taxonomy created by the people. Wal (2006) explains that:

folksonomy is the result of personal tags, free of information and objects (anything with a URL) for your own retrieval. The labelling is made in a social environment

(shared and open to others). The act is that the labelling is made by the person to consume the information.

According to Santarém Segundo and Vidotti (2011, p. 286):

The main purpose of Folksonomy in this context is to allow ordinary users to create labels/tags that can describe or point to the content they are inserting during the self-archiving of digital documents [...] so that the resources can be retrieved later by the user himself user or by others [...].

Having elucidated the issue of the lack of standardization of terms for the representation inherent to the concept of folksonomy, it is possible to understand the complexity in the automatic indexing process in social networks and in digital environments that allow the use of non-controlled vocabularies. This is the case with Flickr, which allows all users to register information regarding the posted images, without any type of tool to assist in this process.

In the specific case of NASA, which uses Flickr, this happens in a more controlled way, for the purpose of more accurate retrieval, since it has a tradition in representing informational objects and has a domain thesaurus. However, there are some situations in which images that appear in the agency's official repository do not appear on her official Flickr. Nonetheless, some of these are registered by ordinary users, without any professional training, which generates ambiguities and inconsistencies.

#### *2.4 Information Treatment: Thematic Representation and Retrieval*

Basically, the treatment of information addresses issues that concern the process of representing documents, whether physical or digital, either in order to describe them physically or the matters contained therein. All forms of representation are intended to allow the user to retrieve information. According to Dias (2001, p. 3), information treatment is defined as:

[...] the function of describing the documents, both physically and thematically, results in the production of representations [...] that not only constitute easier units to manipulate in an information retrieval system, but also represent syntheses that make it easier for the user to assess the relevance that the full document may have [...].

One of the ways of representing a resource is by thematic representation or by subject, which aims to present the content present in the document so that the researcher user can analyze its relevance to it. For Souza (2013, p. 139) it is:

[...] traversed by the cataloguer, through the subject analysis process, which aims to provide an understanding of the content, [...] through a process that really contributes to the consistent identification of the subject addressed in the document.

According to Kuramoto (2002), one of the great challenges found in information retrieval is how to meet the user's information needs quickly and accurately. For Ferneda (2003, p. 14), information retrieval is:

[...] for some, the operation by which documents are selected, from the collection, according to the user's demand. For others, it consists of providing, based on a demand defined by the user, the elements of documentary information [...] can also be used to designate the operation that provides a more or less elaborate response to a demand [...].

Therefore, representation and retrieval are strongly related, if there are failures in representation, consequently, the retrieval is unsatisfactory. Thematic representation of digital images, especially on social networks, is a challenge, as the quality of the record is directly linked to the indexer, which in this case can be any user, with or without knowledge of terms control to facilitate retrieval.

### 3 TERMINOLOGY ANALYSIS IN DIGITAL IMAGES AT NASA'S FLICKR

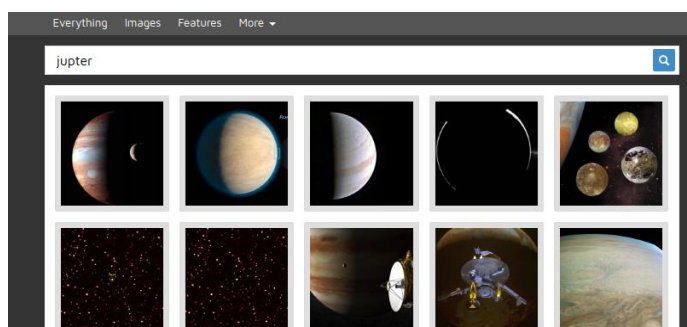
Flickr is a social network that allows its users to digitally store and share photos, images, illustrations, drawings and videos, in addition to enabling the creation of albums and galleries. Like any social network, users can follow other users, like posts, get in touch, bookmark and leave comments. It was developed in 2004 by the Canadian company Ludicorp, which in 2005 was acquired by Yahoo!, and in April 2018, it was sold to the company SmugMug.

Several institutions use Flickr to make digital images available, amongst them the Library of Congress, George Eastman Museum, Amsterdam Museum, NASA, among others.

#### 3.1 Official NASA Repository and Flickr

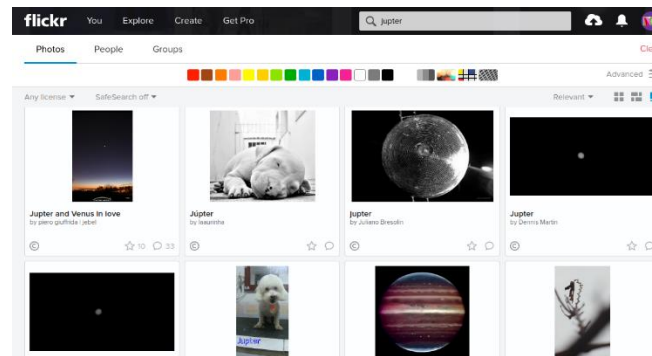
The official NASA repository, which stores digital images, both of the earth and of other planets and space objects, captured by satellites. On the home page, it presents a simple interface. Several blocks with digital images and informational texts about the missions and news that happen at the agency are displayed. Figure 1 shows the repository after a search with the English term “jupiter”.

**Figure 1.** NASA digital repository when searching for the term Jupiter



Source: the author.

The same search was carried out on Flickr and presented different results. Figure 2 shows the results obtained with the term Jupiter on the social network. It was possible to observe the presence of official NASA records as well as ordinary users.

**Figure 2.** Flickr when searching for the term Jupiter

Source: the author.

A repository in a specific domain, such as the NASA official, outperforms a general-purpose repository, such as Flickr. For this reason, no comparison is made between the two tools, as this would render the results invalid.

The searches presented in the following topics were carried out on Flickr, for this purpose, the following steps were defined:

- In the first search, the terms “Big Planet Neptune” were used, common in the aerospace domain, terms that are not included in the NASA thesaurus;
- In the second and third searches, the agency's official thesaurus, NASA THESAURUS VOLUME 1 Hierarchical Listing With Definitions, was used;
- And then, having assembled the following strategies: “Giant Gas Planet Neptune” and “Celestial Giant Gas Planet Neptune”, all the words appear in the agency's thesaurus.

| 8

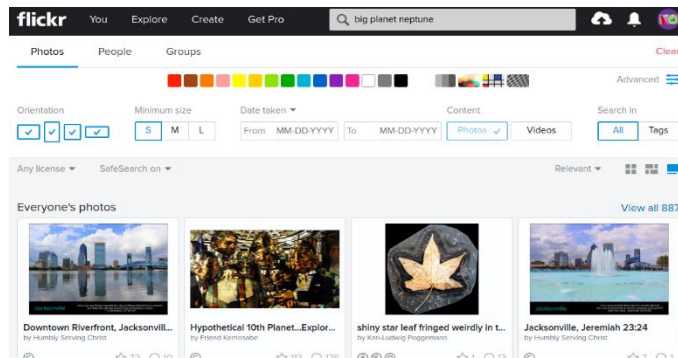
To observe the results, the first 105 records displayed were selected, analyzing the subject or term with which the digital image was indexed.

### 3.2 Image retrieval based on Folksonomic Records

To perform the first search, it was considered that the user was interested in digital image of the planet Neptune, so the expression used was “Big Planet Neptune”, in which only the word “big” is not related to “planet Neptune” in the NASA thesaurus. No types of restrictions or advanced filters have been defined. Figure 3 shows part of the displayed results.



**Figure 3.** Research with Big Planet Neptune



Source: the author.

At the end of the search, the system returned 887 items, however, the first 105 were analyzed, as it considered a sufficient number for this study. Of the 105 objects retrieved, it was possible to observe that only one of the digital images displayed contained the photographic capture of the planet Neptune, although in the search expression the terms “planet” and “Neptune” were used. This means that the relevance was less than 1%.

All digital images retrieved contained in their descriptions all the words used in the search. The inconsistency in the results obtained is due to the variations in which the terms were used in the representation. For example, Neptune, in a digital image referred to the planet, in another to the name of a research institute, in another to a monument and in another to the name of a person.

3.3 Retrieval on Flickr Using NASA's Thesaurus

Before carrying out the second search, a new digital image of the planet Neptune was registered, taken from the NASA website, but edited for easy identification in the retrieval.

During the process of representing the digital image, the terms recommended in NASA THESAURUS were used. Chart 1 shows the final result of registering the image on the website.

**Table 1.** Data and values used in the record of the digital image

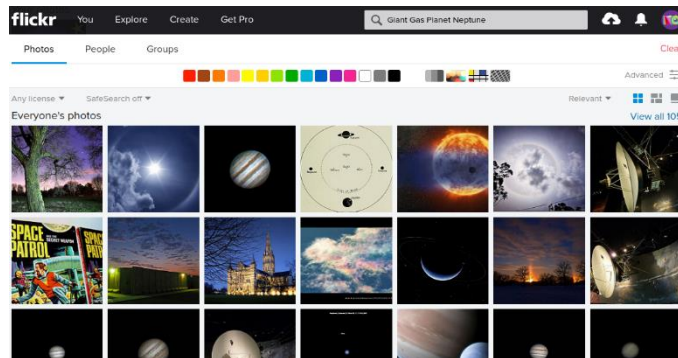
| Atributo        | Valor   |
|-----------------|---|
| Nome do arquivo | neptune_full.png  |
| Descrição       | This is a photo of planet Neptune                       |
| Tags            | Planet, neptune, planets, gas, giant, celestial, bodies |

Source: the author.

3.4 Record-Based Image Retrieval Using Thesaurus

For the second search, the expression used was “Giant Gas Planet Neptune”, with all the terms described in the NASA thesaurus. No types of restrictions or advanced filters have been defined. Figure 4 shows part of the displayed results.

**Figure 4.** Research with Giant Gas Planet Neptune

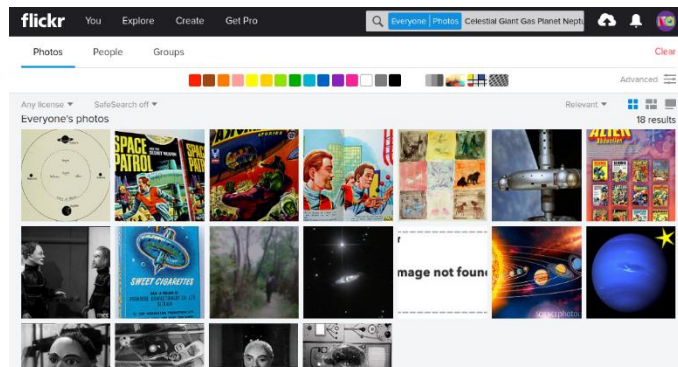


Source: the author.

At the end of the search, the system returned 105 results that met the search, a universe 88% smaller than the first one, considering the total retrieved previously. When analyzing the 105 items returned, it was observed that 12 of the displayed objects contained the capture of the planet Neptune, a relevance just over 11%.

In the third search, an additional term was used, consulted in the NASA thesaurus, the term Celestial. Regarding the second search, significant results were obtained, as shown in figure 5.

**Figure 5.** Research with Celestial Giant Gas Planet Neptune



Source: the author.

At the end of the search, the system returned a total of 18 items, a universe 82% smaller than the second and 97% smaller than the first. When analyzing the retrieved objects, it was observed that 3 digital images displayed were from the planet Neptune, a relevance just over 16%.

The differences between the second and third searches showed a significant difference, especially when looking at the retrieved results, this with the use of just one more term in the search expression.

Thus, results of the analysis present the challenges generated in the use of folksonomy in the thematic representation of digital images and the advantages of using the KOS in the processes of thematic representation and information retrieval in these environments.

## 4 FINAL CONSIDERATIONS

Understanding how to build and use the available KOS tools, both theoretical concepts and practical applications, can make a significant difference in the quality of thematic representation and image retrieval in the digital environment.

Freedom when representing digital images, videos, audios, among others, on social networks must be questioned. These are complex information resources to be represented and that also include the subjectivity issue in their thematic representations. Furthermore, the lack of controlled vocabularies and ontologies, in addition to the use of folksonomies in indexing, reflect a non-specialized vocabulary, which may be devoid of conceptual meaning, but at the same time may represent the terms most used by a community.

The lack of commitment to the thematic representation of these resources, especially the digital image as addressed in this study, may also reflect the lack of interest in the quality of the representation, which will directly impact the retrieval results. It is a fact that, in an open environment such as a social network, little or no imposition, or restriction, is expected on its users due to the interests of companies.

Another factor to be considered is the difficulty in creating KOS tools, such as thesauri and ontologies, in totally open environments. Knowledge representation instruments are built based on an area of knowledge, so that it is possible to standardize the terms used. It is a paradoxical situation to have a free environment in which a control of vocabulary is imposed when the resource is represented.

However, the results obtained in the analysis of this study demonstrate that, even in a social network environment, it is possible, in some cases, to use KOS instruments to improve indexing and retrieval. Professionals involved with representation should pay attention to the process of thematic representation of information, and use, when possible, thesauri from specific domains to consult and use the established controlled terms. The results of the analysis carried out in this study demonstrated that in a search for free terms, the results obtained are not as relevant and accurate as those obtained in searches performed with the terms of the thesaurus.

Therefore, it is considered in this study that the thematic representation of digital images from the use of thesauri, an instrument of the KOS, can significantly contribute to a retrieval with greater efficiency, precision and quality. While folksonomies, typical of social networks, bring a free representation of a digital image and, at the same time, reflect the terminological reality of a community.

## REFERENCES

CARLAN, Eliana. **Sistemas de organização do conhecimento**: uma reflexão no contexto da Ciência da Informação. Brasília/DF: UNB, 2010. Available on:

[https://repositorio.unb.br/bitstream/10482/7465/1/2010\\_ElianaCarlan.pdf](https://repositorio.unb.br/bitstream/10482/7465/1/2010_ElianaCarlan.pdf). Access at: 15 nov. 2018.

CARLAN, Eliana; MEDEIROS, Marisa Bräscher Basílio. Sistemas de organização do conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, Brasília, DF, v. 4, n. 2, p. 53-73, ago./dez. 2011. Available on:

[https://repositorio.unb.br/bitstream/10482/12867/1/ARTGO\\_SistemasOrganizacaoConhecimento.pdf](https://repositorio.unb.br/bitstream/10482/12867/1/ARTGO_SistemasOrganizacaoConhecimento.pdf). Access at: 22 nov. 2018.

CERVANTES, Brígida Maria Nogueira. **A construção de tesouros com a integração de procedimentos terminográficos**. Marília/SP: UNESP, 2009. Available on:

[https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/cervantes\\_bmn\\_do\\_mar.pdf](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/cervantes_bmn_do_mar.pdf). Access at: 22 nov. 2018.

DIAS, Eduardo Wense. Contexto digital e tratamento da informação. Rio de Janeiro: **DataGramZero** - Revista de Ciência da Informação, 2001. Available on:

[http://www.brapci.inf.br/\\_repositorio/2010/01/pdf\\_8df58fec78\\_0007466.pdf](http://www.brapci.inf.br/_repositorio/2010/01/pdf_8df58fec78_0007466.pdf). Access at: 20 jan. 2019.

FERNEDA, Edberto. **Introdução aos modelos computacionais de recuperação de informação**. Rio de Janeiro: Editora Ciência Moderna, 2012.

FERNEDA, Edberto. **Recuperação de informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. Tese (Doutorado em Ciência da Informação e Documentação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003. doi:10.11606/T.27.2003.tde-15032004-130230. Available on:

<https://teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/publico/Tese.pdf>. Access at: 23 jan. 2019.

FERNEDA, Edberto; DIAS, Guilherme Ataíde. OntoSmart: um modelo de recuperação de informação baseado em ontologia. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, n. 2, p. 170-187, jun. 2017. ISSN 19815344. doi:<http://dx.doi.org/10.1590/1981-5344/2081>. Available on: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2081>. Access at: 18 jan. 2019.

| 12

GUIMARÃES, José Augusto Chaves. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, v. 43, n. 1, June 2015. ISSN 1518-8353. Available on: <http://revista.ibict.br/ciinf/article/view/1415>. Access at: 15 mai. 2020.

HASSAN-MONTERO, Yusef. **Indización social y recuperación de información**. No Solo Usabilidad Journal, 2006. Available on:

[http://www.nosolousabilidad.com/articulos/indizacion\\_social.htm](http://www.nosolousabilidad.com/articulos/indizacion_social.htm). Access at: 16 jan. 2019.

HJØRLAND, Birger. Fundamentals of knowledge organization. **Knowledge Organization**, v. 30, n. 2, p. 87-111. Available on:

<http://ppggoc.eci.ufmg.br/downloads/bibliografia/Hjorland2003.pdf>. Access at: 17 nov. 2018.

JACOB, Elin K. Ontologies and the semantic web. **Bulletin of the American Society for Information Science and Technology**, Apr./May, 2003. Available on:

<https://onlinelibrary.wiley.com/doi/pdf/10.1002/bult.283>. Access at: 19 jan. 2019.

KURAMOTO, Hélio. Sintagmas Nominais: uma nova proposta para a recuperação de informação. **DataGramZero** - Revista de Ciência da Informação, Rio de Janeiro, v.3, n.1, 2002. Available on: <https://ridi.ibict.br/bitstream/123456789/150/1/KuraData2002.pdf>. Access at: 16 jan. 2019.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. 2. ed. Brasília/DF: Briquet de Lemos, 2004.

LIBEN-NOWELL, David; KLEINBERG, Jon. The link-prediction problem for social networks. **Journal of the American Society for Information Science and Technology**, v. 58, n. 7, p. 1019-1031, May, 2007.

SANTARÉM SEGUNDO, José Eduardo.; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Representação iterativa e folksonomia assistida para repositórios digitais. Rio de Janeiro: **Liinc em Revista**, v.7, n.1, 2011. Available on: <http://revista.ibict.br/liinc/article/view/3300/2916>. Access at: 19 jan. 2019.

SILVA, Renata Eleuterio da; SANTOS, Renata Eleuterio da; FERNEDA, Edberto. Modelos de Recuperação de Informação e Web Semântica: a questão da relevância. **Informação & Informação**, João Pessoa, v. 18, n. 3, p. 27-44, out. 2013. ISSN 1981-8920. doi:<http://dx.doi.org/10.5433/1981-8920.2013v18n3p27>. Available on: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/12822>. Access at: 17 dez. 2018.

SOUZA, Brisa Pozzi de. Representação Temática da Informação Documentária e sua Contextualização em Biblioteca. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 9, n. 2, p. 132-146, nov. 2013. ISSN 1980-6949. Available on: <https://rbbd.febab.org.br/rbbd/article/view/249>. Access at: 10 dez. 2018.

TRISTAO, Ana Maria Delazari; FACHIN, Gleisy Regina Bóries; ALARCON, Orestes Estevam. Sistemas de classificação facetados e tesauros: instrumentos para organização do conhecimento. **Ci. Inf.**, Brasília, v. 33, n. 2, p. 161-171, agosto 2004. [.https://doi.org/10.1590/S0100-19652004000200017](https://doi.org/10.1590/S0100-19652004000200017). Available on: <http://www.scielo.br/pdf/ci/v33n2/a17v33n2.pdf>. Access at: 13 jan. 2019.

WAL, Thomas Vander. **Folksonomy Definition and Wikipedia**. 2005. Available on: <http://www.vanderwal.net/random/entrysel.php?blog=1750>. Access at: 17 jan. 2019.

WIKIPEDIA. Flickr. Available on <https://pt.wikipedia.org/wiki/Flickr>. Access at: 21 jan. 2019.

Article submitted to the similarity system



Submitted: 01/03/2020 – Accept: 11/05/2020 – Published: 18/05/2020

---