

**XIX** encontro nacional  
de pesquisa em  
ENANCIB ciência da informação

// SUJEITO INFORMACIONAL E AS  
PERSPECTIVAS ATUAIS EM CIÊNCIA  
DA INFORMAÇÃO. //

**22-26**  
**OUTUBRO**  
**2018**  
LONDRINA/PR



## **XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018**

### **GT-2 – Organização e Representação do Conhecimento**

#### **DESENVOLVIMENTO DE SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: PARTINDO DOS SIMPLES PARA OS COMPLEXOS**

**Webert Júnio Araújo (Universidade Federal de Minas Gerais)**

**Gercina Ângela de Lima (Universidade Federal de Minas Gerais)**

**Rainner Finelli Gomes (Universidade Federal de Minas Gerais)**

**Lucas Gomes Fernandes (Universidade Federal de Minas Gerais)**

#### ***DEVELOPMENT OF KNOWLEDGE ORGANIZATION SYSTEMS: FROM THE SIMPLE TO THE COMPLEXES***

##### **Modalidade da Apresentação: Pôster**

**Resumo:** Sistemas de Organização do Conhecimento desempenham papel fundamental no contexto da Biblioteconomia e Ciência da Informação, visto que são insumos básicos para a organização de domínios do conhecimento, bem como do conhecimento produzido (informação registrada). Esta pesquisa se baseia no pressuposto de que o desenvolvimento de Sistemas de Organização do Conhecimento (SOC) de estrutura simples pode auxiliar a construção dos instrumentos de estrutura mais complexa. O objetivo deste trabalho é estudar as etapas de desenvolvimento dos principais tipos de Sistemas partindo daqueles de estruturas simples até aqueles de estrutura mais complexa. A pesquisa se caracteriza como qualitativa, quanto à abordagem; e aplicada, quanto à natureza. Os procedimentos metodológicos foram descritos à medida que cada instrumento foi construído. Com base na experiência de desenvolvimento de cada instrumento, constatou-se através de resultados iniciais que a tarefa de construção dos SOCs de estrutura mais elementar, como um glossário, contribui de forma significativa para a construção daqueles de estrutura mais complexas, como tesouros e ontologias.

**Palavras-Chave:** Sistema de Organização do Conhecimento; Glossário; Taxonomia; Tesouro; Ontologia.

**Abstract:** Knowledge Organization Systems play a fundamental role in the context of Librarianship and Information Science, since they are basic inputs for the organization of knowledge domains, as well as knowledge produced (registered information). This research is based on the assumption that the development of Simple Structure Knowledge Organization Systems can help the construction of

instruments with more complex structure. In view of this, this work is based on the following question: can the development of a simple structure of Knowledge Organization contribute to the development of those with a more complex structure? Thus, the objective of this work is the development of the main types of Systems starting from those of simple structures to those of more complex structure. The research is characterized as qualitative, as to the approach and applied, as to nature. The methodological procedures were described as each instrument was constructed. Based on the development experience of each instrument, it was verified through initial results that the task of constructing the most elementary Structure Systems as a glossary contributes significantly to the construction of those of more complex structure such as thesauri and ontologies.

**Keywords:** Knowledge Organization System; Glossary; Taxonomy; Thesaurus; Ontology.

## 1 INTRODUÇÃO

Representar o conhecimento em domínios é relevante por várias razões: controle terminológico e consenso com relação ao uso de determinado conceito; compreensão do domínio, para facilitar a comunicação entre os especialistas através da delimitação de significados e a consequente redução de ambiguidades. São os Sistemas de Organização do Conhecimento (SOCs) os instrumentos utilizados para esse fim.

SOCs são ferramentas que apresentam a interpretação organizada de estruturas do conhecimento (HJORLAND, 2007). Desempenham papel fundamental no contexto da Biblioteconomia e Ciência da Informação (BCI), visto que são insumos básicos para a organização de domínios do conhecimento, bem como do conhecimento produzido (informação registrada).

A literatura da BCI apresenta técnicas, metodologias e propostas diversas para o desenvolvimento de SOC, que guiam o procedimento de construção desses instrumentos. Esta pesquisa se baseia no pressuposto de que o desenvolvimento de SOC de estrutura simples<sup>1</sup> pode auxiliar a construção de SOC de estrutura mais complexa. Diante disso, este trabalho parte da seguinte questão: o desenvolvimento de um SOC de estrutura simples pode contribuir para o desenvolvimento de SOC de estrutura mais complexa?

O objetivo deste trabalho é estudar as etapas de desenvolvimento dos principais tipos de SOC partindo daqueles de estruturas simples (glossário) até aqueles de estrutura mais complexa (ontologia). Propõe-se o desenvolvimento preliminar dos seguintes instrumentos: glossário, taxonomia, sistema de classificação, tesouro e ontologia.

---

<sup>1</sup> Essa nomenclatura se respalda em Zeng (2008).

Este estudo se justifica, pois se trata de um trabalho empírico no desenvolvimento dos SOCs mais comumente usados no contexto da BCI, que utilizam como base estrutural a hierarquização do conhecimento, partindo da representação conceitual por meio da definição terminológica, com o glossário e a representação desses conceitos de forma estruturada. Portanto, espera-se que o percurso metodológico utilizado para o desenvolvimento deste estudo possa contribuir para as discussões na perspectiva de construção desses instrumentos.

Para realizar este estudo, foram utilizados os conceitos do texto “Sistema de busca e recuperação de informação” (SISTEMAS, [20--]) como amostra, para o desenvolvimento de cada SOC, conforme descrito nas próximas seções deste trabalho.

## **2 NATUREZA DA PESQUISA**

Esta pesquisa se caracteriza como qualitativa, quanto à abordagem; e aplicada, quanto à natureza. Os procedimentos metodológicos usados para a construção dos SOCs serão descritos na seção 3. Optou-se por essa forma de apresentação dos procedimentos por acreditar-se que essa estrutura facilitará a leitura e a compreensão do caminho realizado para se alcançar os resultados da pesquisa.

## **3 PROCEDIMENTOS METODOLÓGICOS**

Nas próximas subseções, apresentam-se breves definições de cada SOC, com base em autores da Biblioteconomia e Ciência da Informação, e as etapas utilizadas para o desenvolvimento dos sistemas. O Glossário é o primeiro SOC desenvolvido, por ser o mais elementar, e servirá de base terminológica para a construção dos outros instrumentos. Depois constrói-se a Taxonomia e o Sistema de Classificação com base na categorização dos termos do Glossário e a hierarquização das classes. Em seguida, desenvolve-se o SOC Tesouro, que se trata de um instrumento mais completo do que os anteriores, com as relações hierárquica, de equivalência e associativa. Por fim, desenvolve-se a Ontologia, considerada neste trabalho o SOC mais complexo, pois envolve conceitos, relacionamentos entre eles, restrições aos relacionamentos e axiomas, que possibilitam uma representação mais completa da realidade.

### **3.1 Glossário**

De acordo com o verbete do Dicionário de Biblioteconomia e Arquivologia (CUNHA, CAVALCANTI, 2008), glossário é a obra que explica o significado de vocábulos pouco comuns,

bem como termos técnicos e científicos. Aparece, às vezes, como apêndice de livros técnicos e científicos ou mesmo de obras literárias.

O desenvolvimento do Glossário foi a primeira atividade realizada, que, por consistir de uma lista de termos com suas definições, serviu como base para construção dos outros instrumentos, uma vez que, definidos os termos de um domínio, fica mais fácil categorizá-los e classificá-los.

A construção do Glossário consistiu nas seguintes etapas: a) leitura do documento para identificar a sua temática; b) extração dos termos identificados como relevantes, totalizando 70 termos; c) definição dos termos: os termos foram definidos com base no Dicionário de Biblioteconomia e Arquivologia e em artigos científicos da BCI.

### **3.2 Taxonomia**

Segundo o Dicionário de Biblioteconomia e Arquivologia (CUNHA, CAVALCANTI, 2008), uma taxonomia é o estudo teórico das bases, leis, regras e princípios de uma classificação e, em adição, Lima e Maculan (2017) afirmam que ela é composta de um conjunto de termos arranjados em hierarquias representando o domínio modelado, que organiza conceitos segundo as suas semelhanças e diferenças. As autoras também exemplificam o mapeamento do conhecimento de um domínio e o estabelecimento de um rótulo para as informações disponibilizadas, como algumas de suas funções.

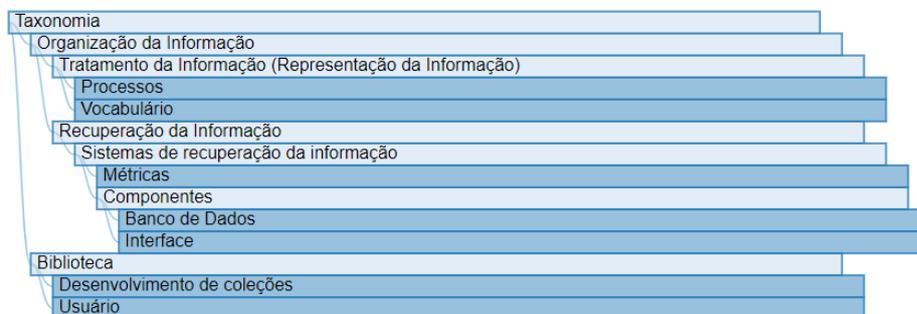
Bocato (2011) indica a ISO 25964 como a norma responsável pelas diretrizes de elaboração de taxonomias, a qual também inclui as redes semânticas e os tesouros, e destaca, também, os sistemas de classificação, os sistemas de classificação de negócios para gerenciamento de registros, as listas de cabeçalhos de assunto. Além disso, inclui, também, os mapas conceituais, as terminologias/banco de termos e as listas de autoridades de nomes.

As classes definidas foram baseadas na lista dos termos extraídos do artigo utilizado neste trabalho e, em seguida, agrupados de acordo com a proximidade de suas definições. O processo em questão é classificado como indutivo, o qual, segundo Campos (2004), é uma abordagem *bottom-up* e possibilita a elaboração de modelos, partindo, desde o início, da representação dos elementos/objetos e relações de um contexto.

Na figura 1 são exibidas as principais classes utilizadas na Taxonomia. A versão completa está implementada com o recurso dinâmico de expandir/contrair e foi

disponibilizada na *web*. O endereço para acessá-la é:  
<<https://dlucasgf.github.io/aoc2/ind.html>>.

**Figura 1: Principais classes da Taxonomia proposta**



Fonte: Elaborada pelos autores.

### 3.3 Sistema de Classificação

De acordo com o Dicionário de Biblioteconomia e Arquivologia (CUNHA, CAVALCANTI, 2008), um Sistema de Classificação é um esquema específico de classificação; plano para o arranjo de documentos de acordo com princípios predeterminados ⇔ Classificação bibliográfica. Para Lima e Maculan (2017), os Sistemas de Classificação são construídos a partir de signos estruturados e normalizados, podendo ser gerais, quando buscam cobrir todos os campos do conhecimento, ou especializados, quando representam um assunto específico. Ademais, eles podem ser analíticos, quando sistematizam fenômenos físicos, viabilizando seu entendimento, ou podem ser documentais (de assuntos ou bibliográficos) e implicam a classificação de documentos, em qualquer tipo de suporte e mídia, visando facilitar, principalmente, a localização do item em um dado acervo.

Dito isso, o sistema de classificação proposto neste trabalho é considerado documental, uma vez que utiliza como base um artigo. Apoiados pela estrutura elencada por Lima e Maculan (2017), a taxonomia elaborada na seção anterior já apresentava os elementos agrupados por semelhança de acordo com o contexto, o que resultou apenas na escolha dos símbolos numéricos (1>1.1>1.1.1) para representar as classes.

### 3.4 Tesouro

O Tesouro teve sua origem devido à necessidade da utilização e organização de um número grande de materiais especializados. De acordo com o verbete do Dicionário de Biblioteconomia e Arquivologia (CUNHA, CAVALCANTI, 2008, p. 362), o Tesouro é um “vocabulário controlado e dinâmico de descritores relacionados semântica e genericamente, que cobrem de forma extensiva um ramo específico de conhecimento”.

Segundo Sales e Café (2009, p. 102), tesouros “são vocabulários formados por termos-descretores semanticamente relacionados, e atuam como instrumentos de controle terminológico”.

A construção do Tesouro consistiu nas seguintes etapas: seleção dos termos já extraídos, identificados e definidos pelo glossário, e uma pesquisa e análise de outros tesouros já elaborados. Os tesouros usados como base foram os Descritores em Ciências da Saúde (Decs), elaborados pela Biblioteca Virtual em Saúde, e o Tesouro Brasileiro de Ciência da Informação, elaborado pelas autoras Lena Vania Ribeiro Pinheiro e Helena Dodd Ferrez, com auxílio do IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia). Dessa forma, a versão preliminar do Tesouro proposto possui cerca de 200 descritores.

### **3.5 Ontologia**

Ontologia é um instrumento que possibilita a representação consensual de determinada abstração da realidade, por meio da estruturação de conceitos, associações entre esses conceitos e restrições sobre a interpretação desses, para evitar ambiguidades. Uma das definições mais citadas na literatura sobre ontologia é a de Gruber (1993), na qual ele diz que ontologia é uma especificação formal e explícita de determinada conceitualização.

Na literatura das áreas de Ciência da Computação e Informação, existem algumas metodologias para construção de ontologias. Entre as principais, podem-se citar: TOVE, Methontology, Neon, Método 101, Sensus. Recentemente, foi criada uma metodologia no âmbito da Escola de Ciência da Informação – UFMG, a Ontoforinfoscience, desenvolvida na tese de doutorado de Fabrício Mendonça, sob orientação do Professor Maurício Almeida. Essas metodologias fornecem as etapas para o desenvolvimento de ontologias, principalmente as ontologias de domínio.

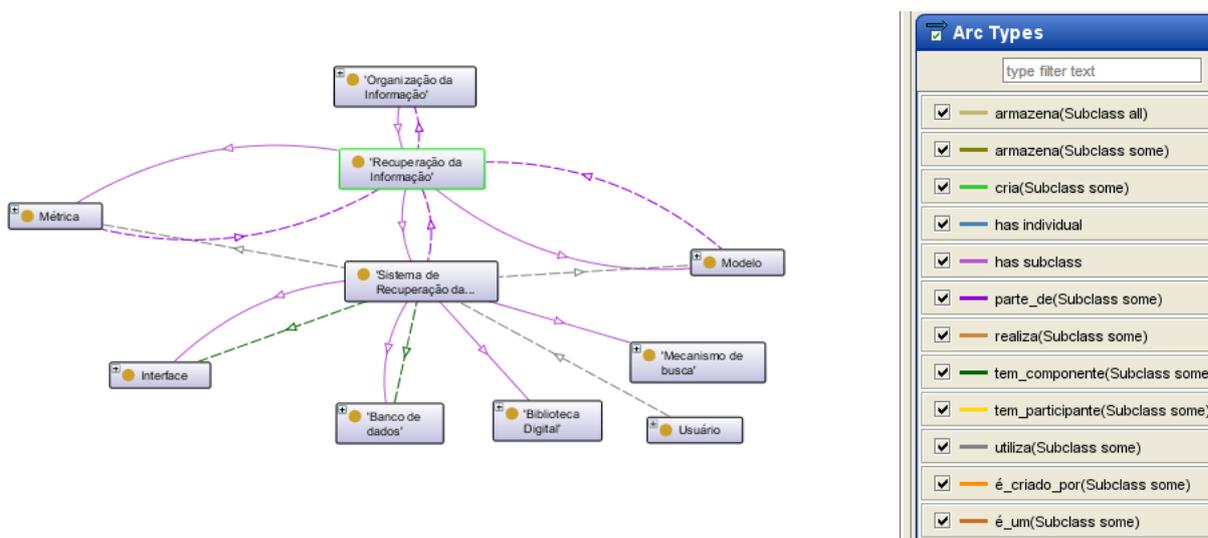
Neste trabalho, seguiu-se, ainda que de forma imprecisa (por essa ser uma pesquisa em andamento, as outras etapas serão desenvolvidas futuramente), as etapas de desenvolvimento da Metodologia Ontoforinfoscience de Mendonça (2015). A razão da escolha dessa metodologia se deve ao fato de ser uma metodologia especialmente criada para pesquisadores da CI interessados na construção de ontologias, e por ela apresentar conceitos técnicos que são do domínio dos pesquisadores e explicar outros que não são tão comuns na CI.

Com base na Metodologia Ontoforinfoscience e nos SOCs desenvolvidos anteriormente, as etapas utilizadas para o desenvolvimento da ontologia neste trabalho estão descritas a seguir:

- a) especificação da ontologia: utilizaram-se 5 Questões de Competência para definição do escopo da ontologia. São elas: i) Quais as técnicas de tratamento da informação? ii) Quais os instrumentos de tratamento da informação? iii) Quais tipos de documentos uma BDTD armazena? iv) Quais estratégias de busca podem ser usadas em um SRI? v) Quais as métricas de um SRI?;
- b) aquisição e extração de conhecimento: foi utilizado o glossário descrito na subseção 3.1, a taxonomia, o tesauro e o texto base;
- c) conceitualização: para esta etapa, utilizou-se mapa conceitual para contribuir na modelagem do domínio. Elencaram-se os conceitos principais e, em seguida, definiram-se os conceitos específicos, ou seja, utilizou-se a abordagem dedutiva;
- d) fundamentação ontológica (utilizar ontologia de alto nível): não foi realizada;
- e) formalização parcial da ontologia: criaram-se alguns axiomas e propriedades lógicas para a ontologia, visto que a formalização completa exige mais insumos (conhecimento do domínio);
- f) avaliação da ontologia: não foi realizada;
- g) documentação da ontologia: o presente documento;
- h) disponibilização da ontologia: disponibilização do arquivo .owl gerado no Software Protégé.

Apresenta-se, por meio de algumas classes e tipos de relações, parte dessa ontologia, proposta como resultado da aplicação da Metodologia Ontoforinfoscience, conforme ilustrado na figura 1.

Figura 1 – Parte da ontologia em forma de grafo



Fonte: Elaborada pelos autores – 2018

#### 4 CONSIDERAÇÕES FINAIS

Após a realização de todos os processos do trabalho, conseguiu-se identificar o quão importante é planejar para conseguir escolher e determinar os termos apropriados para cada Sistema de Organização do Conhecimento. Resultados iniciais mostram que o desenvolvimento de um SOC com estrutura mais elementar, como um glossário, contribui de forma significativa para o desenvolvimento de SOCs com estrutura complexa, como tesouros e ontologias.

Entre os obstáculos identificados na criação dos SOCs, destaca-se o processo de categorização do domínio. O processo de categorização é subjetivo, visto que depende do propósito da aplicação do instrumento e da visão de mundo dos classificacionistas. Portanto, as decisões de quais categorias eram mais pertinentes exigiu análise e estudo da temática.

#### REFERÊNCIAS

BOCCATO, Vera Regina Casari. Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. **InCID: Revista de Ciência da Informação e Documentação**. Ribeirão Preto, v. 2, n. 1, p. 165-192, jan./jun. 2011.

CAFÉ, Lígia; SALES, Rodrigo. Diferenças entre tesouros e ontologias. **Perspectivas em Ciências da Informação**. Belo Horizonte, v. 14, n. 1, p. 99-116, jan./abr. 2009.

CAMPOS, Maria Luiza de Almeida. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**. Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

CUNHA, Murilo Bastos da; CAVALVANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília, DF: Briquet de Lemos/Livros, 2008.

DeCS – Descritores em ciências da saúde. **Pesquisa a consulta de índices permutado**. Disponível em: <<http://decs.bvs.br/>>. Acesso em: 10 jun. 2018.

FERREZ, Helena Dodd; PINHEIRO, Lena Vania Ribeiro. **Tesouro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), 2014.

GRUBER, Thomas R. Toward principles for the Design of Ontologies Used for Knowledge Sharing. In: **FORMAL ONTOLOGY IN CONCEPTUAL ANALYSIS AND KNOWLEDGE REPRESENTATION**, 1993, Padova. Italy. **Proceedings...** Padova. Italy: [s.n], 1993.

HJORLAND, B. Knowledge Organization Systems. 2007. Disponível em: <[https://www.db.dk/bh/lifeboat\\_ko/CONCEPTS/knowledge\\_organization\\_systems.htm](https://www.db.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_systems.htm)>. Acesso em: 20 jun. 2018.

LIMA, Gercina Ângela de; MACULAN, Benildes Coura Moreira dos Santos. Estudo comparativo das estruturas semânticas em diferentes sistemas de organização do conhecimento. **Ciência da Informação**. Brasília, v. 46, n. 1, p. 60-72, jan./abr. 2017.

MENDONÇA, Fabrício. **Ontoforinfoscience**: metodologia para construção de ontologias pelos cientistas da informação: uma aplicação prática no desenvolvimento da ontologia sobre componentes do sangue humano (HEMONTA). 2015. 322 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, UFMG, Belo Horizonte, 2015.

SISTEMAS de busca e recuperação de informação. [S.l.]: PUC Rio certificação Digital, [20--]. Disponível em: <[http://www2.dbd.puc-rio.br/pergamum/tesesabertas/0313143\\_06\\_cap\\_05.pdf](http://www2.dbd.puc-rio.br/pergamum/tesesabertas/0313143_06_cap_05.pdf)>. Acesso em: 20 jun. 2018.

VIEIRA, Jessica Monique de Lira; SANTOS, Monick Trajano dos; LAPA, Remi Correia. Estudo da construção e aplicação do tesouro na recuperação da informação de teses e dissertações do programa de pós-graduação em desenvolvimento urbano. **Biblionline**. João Pessoa, n. esp., p. 71-80, 2010.

ZENG, M. L. Knowledge Organization Systems. **Knowledge Organization**. [S.l.], v. 35, n. 2, p. 160-182. 2008.