

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-2 – Organização e Representação do Conhecimento

VISUALIZAÇÃO DA INFORMAÇÃO: REPRESENTAÇÃO DE ESTRUTURAS DE CLASSIFICAÇÃO

INFORMATION VISUALIZATION: REPRESENTATION OF CLASSIFICATION STRUCTURES

Tainá Regly – Universidade Federal do Rio de Janeiro
Rosali Fernandez Souza – Universidade Federal do Rio de Janeiro

Modalidade: Resumo Expandido

Resumo: A classificação é um processo fundamental da natureza humana e principal atividade de organização e representação de informação. O trabalho trata de duas técnicas para representar estruturas de classificação: a estrutura de árvore e a de grafos. A estrutura de árvore representando ligações entre conceitos e a estrutura de grafos representando relações entre conceitos. Conclui relatando que essas duas técnicas, apesar de diferentes, são aplicáveis ao objetivo comum de representar o conhecimento em diferentes níveis de complexidade.

Palavras-Chave: Classificação; Representação do Conhecimento; Visualização da Informação.

Abstract: Classification is a fundamental process of human nature and the main activity of organizing and representing information. This communication deals with two techniques for representing classification structures: the tree structure and the graph structure. The tree structure representing links between concepts and the graph structure representing relationships between concepts. It concludes by reporting that these two techniques, although different, are applicable to the common goal of representing knowledge at different levels of complexity.

Keywords: Classification; Knowledge representation; Information visualization.

1 INTRODUÇÃO

No livro “As palavras e as coisas”, Foucault (2002) afirma que a classificação é espelho da cultura da qual se origina. Para ele, o ato de classificar é natural na existência dos seres humanos e está arraigado em sua cognição já que nossa maneira de lidar com o mundo real consiste em ordená-lo baseando-nos em esquemas de classificação e independe da cultura na qual estamos inseridos (MAZZOCCHI, 2013).

A necessidade de representar o conhecimento e a estrutura conceitual de um domínio sempre ocuparam destaque nas atividades de classificar o conteúdo de assunto e a guarda dos documentos físicos. Com a transição para o mundo digital, essa necessidade se tornou ainda mais substancial uma vez que, num mundo conectado pela Web, são inúmeras as relações entre os conceitos a serem representadas. A dificuldade não é mais, principalmente, pensar em formas de organizar documentos para uma guarda efetiva. A necessidade de uma linguagem que torne possível a representação dos mais diversos tipos de informação se tornou um dos maiores desafios a serem enfrentados pelos profissionais envolvidos. Conforme Simões (2016, p. 46),

[...] o desenvolvimento de uma linguagem com maior flexibilidade em suas relações de termos vai proporcionar a organização e representação da informação necessária para a formulação de estratégias de busca com maior dinâmica e confiabilidade das informações com vistas a sua disseminação.

O presente trabalho trata de aspectos epistemológicos da classificação como uma atividade inerente ao pensamento e existência do homem, bem como de duas formas de representar as estruturas classificadas: as técnicas de árvore e de grafos.

2 A CLASSIFICAÇÃO

O ser humano elabora modelos para compreender o mundo, estabelecer padrões de comunicação e representar de forma simplificada um objeto ou uma situação no mundo. Os modelos conceituais, aqui considerados como modelos conceituais ou de domínio como sinônimo de objetos classificados, são construídos a partir de abstrações que evidenciam o relacionamento entre conceitos, trabalhando semelhanças, diferenças e outras associações de significado (SALES; SAYÃO; MOTTA, 2012). “Conhecer aquilo que pertence propriamente a um indivíduo é ter diante de si a classificação ou a possibilidade de classificar o conjunto

dos outros. A identidade é aquilo que a marca se define pelo resíduo das diferenças” (FOUCAULT, 2002, p. 199). O autor afirma que não há nada mais empírico ou que exija tanto da cognição humana do que a ordenação das coisas. O filósofo conta que essa ação “não se trata de ligar consequências, mas sim de aproximar e isolar, de analisar, ajustar e encaixar conteúdos concretos” (FOUCAULT, 2002, p. XV).

Segundo Mazzocchi (2013), as classificações existem para dar limites às coisas e está atrelada à adaptação dos seres humanos ao mundo e à sua sobrevivência, pois permite que esses compartilhem percepções e até mesmo, em alguns casos, formas de comunicação. Diferentes formas de ordenar uma realidade nos permitem vislumbrar os limites da nossa compreensão e conhecer outras maneiras possíveis de compreender o que se passa ao nosso redor, visto que para se categorizar, as variáveis dependerão da percepção, experimentação e conceituação da realidade de cada um. Para o autor, ao realizar uma pesquisa, cientistas devem observar múltiplos pontos de vista em detrimento de apenas um e, por isso, a classificação é um tópico importante na ciência.

A organização do conhecimento é baseada em esquemas de classificação (MAZZOCCHI, 2013). Categorizar conceitos permite que um domínio seja representado de modo que reflexões lógicas e epistemológicas se manifestem na organização pragmática do conhecimento. É a partir das estruturas classificadas que nos é permitido perceber referências cruzadas que elucidam relações (links) entre os objetos categorizados. Esses links agrupam características comuns e evidenciam diferenças entre os conceitos ou até mesmo contradições (FEDELI, 2013).

Conforme Le Moigne (apud CAMPOS, 2001, 2004), modelar é conhecer e conceber, para um objeto, um modelo que possibilite o ser humano conhecê-lo, compreendê-lo e interpretá-lo. Essa compreensão ocorre através da construção de modelos de um campo do conhecimento que permite descrever, fornecer explicações, conceituar e representar os fenômenos que observamos no mundo. A forma de representar um domínio serve como princípio norteador para a organização de dados? documentos, informação e conhecimento, pois permite a estruturação de seus conceitos e relações através da busca, identificação, descrição e entendimento de determinado contexto (SOLANO, 2014).

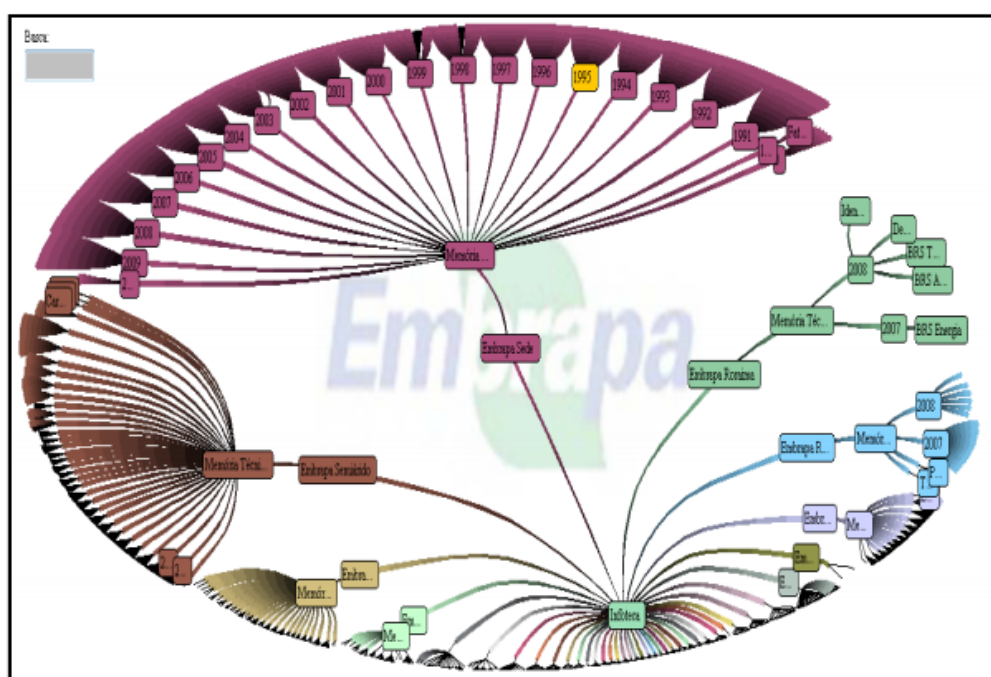
3 REPRESENTAÇÃO DE ESTRUTURAS DE CLASSIFICAÇÃO

De acordo com Fedeli (2013) a representação espacial de dados permite e promove o mapeamento da informação através do uso de categorias linguísticas sintáticas e semânticas. Abordaremos dois modelos de representação: a estrutura de árvore e de grafos, que foram desenvolvidos em diferentes contextos e épocas, porém com o objetivo comum retratar a complexidade do conhecimento.

No ocidente, a estrutura da árvore baseada na lógica aristotélica tem sido o modelo predominante na representação de classificações. Essa estrutura ramificada se tornou icônica na representação da ordem hierárquica obtida por uma série de bifurcações onde, em cada nível, as diferenças essenciais entre os objetos e suas características definidoras são concebidas (MAZZOCCHI, 2013).

Para Fedeli (2013) a árvore deve se ramificar o máximo que for possível, partindo de um objeto geral que seria o seu tronco e subir para as principais classes e, dos ramos maiores, partir para os menores. Na figura 1 vemos o exemplo de uma técnica de visualização originada da estrutura de árvore utilizada para representar conceitos e as relações entre eles. Essa técnica é capaz de revelar subordinação entre os conceitos e de evidenciar relações de gênero-espécie e partição ao representar graficamente as informações ali contidas.

Figura 1 - Browser hiperbólico da Infoteca-e/Embrapa



Fonte: Vieira e Correa (2011)

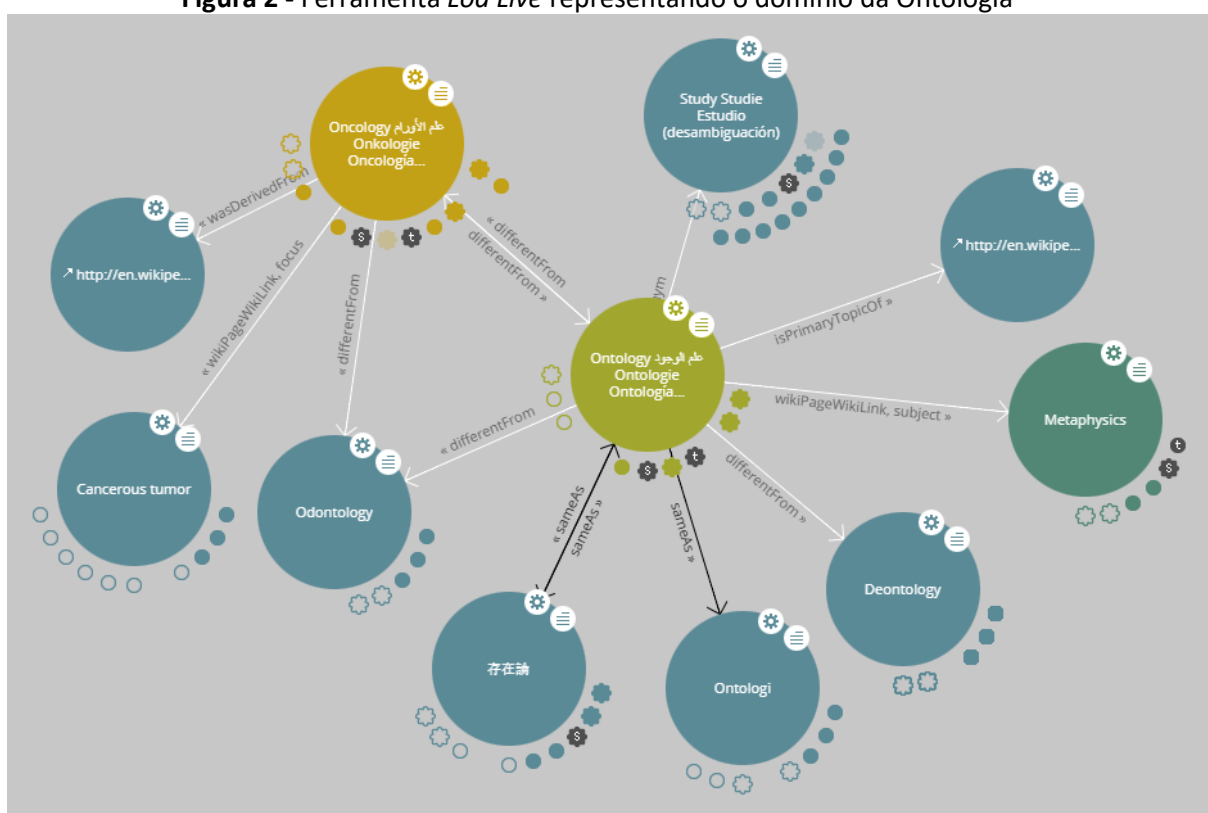
Shneiderman (1996) e Olive (2018) contam que os nós e os links entre os dados representados em estruturas hierárquicas podem apresentar diversos tipos de atributos e que aparecem em taxonomias, nas estruturas de organizações, nas genealogias, em diretórios, na Classificação Decimal de Dewey, entre outros. Dados representados nessas estruturas permitem que o usuário produza inferências a partir de micro-observações dos elementos individuais, de macro-observações de grandes grupos, do encontro de determinado nó, da visualização do contexto de uma hierarquia, do exame da estrutura e das relações encontradas em uma árvore.

Eco (1984 apud MAZZOCCHI, 2013) aponta a insustentabilidade lógica da estrutura da árvore para representar algo tão complexo quanto o conhecimento. Nesse tipo de visualização, diversos tipos de hierarquias podem ser desenvolvidos desde que sejam originárias de um mesmo ponto de vista. O autor destaca a complexidade multidimensional do universo do conhecimento de modo que os conceitos estariam relacionados de muitas mais formas do que apenas com a relação gênero e espécie possibilitada pelas técnicas de árvore. Ou seja, a técnica de árvore não é capaz de representar domínios do conhecimento de maneira complexa apenas com hierarquias que elucidam as relações de gênero e espécie ou partitivas dentro de uma maneira única e universal de classificar.

Fedeli (2013) conta que a teoria dos grafos se desenvolveu e se tornou a principal ferramenta para analisar conceitos e seus relacionamentos. A teoria permite a construção de sistema de conexões como as ontologias que são úteis para examinar os domínios complexos do conhecimento. De acordo Shneiderman (1996), representar itens através de uma estrutura de árvore em determinadas situações pode não ser conveniente, dado que determinados elementos possuem complexidades que estão além da representação permitida por uma hierarquia. Por mais que hierarquias sejam formas especializadas de redes, a representação perde sua estrutura hierárquica ao evidenciar diversos tipos de relacionamentos e, por isso, necessita de um algoritmo diferente para posicionar os nós (HEER; BOSTOCK; OGIEVETSKY, 2010). Estruturas gráficas gerais como links de nós, hipermídia, redes semânticas e teias permitem que pesquisadores encontrem caminhos mais curtos ou baratos para conectar dois itens localizados numa rede através da representação de relações de equivalência contidas nas categorias que relações partitivas e hierárquicas.

Para Dias (2007), as estruturas de grafos representam diversos relacionamentos entre as instâncias de determinada classe. Essa visualização espacial entre instâncias auxilia na busca de informações ao permitir que usuário tenha noção da dimensão do contexto sobre o qual está navegando. Na técnica de Grafos (figura 2), é escolhida uma representação visual onde se define a posição de cada vértice e aresta, onde o primeiro pode representar termos ou documentos e, o segundo, as relações semânticas ou ocorrência dos termos.

Figura 2 - Ferramenta Lod Live representando o domínio da Ontologia



Fonte: <http://en.lodlive.it/?http://dbpedia.org/resource/Ontology>

Na figura 2, destacamos a aplicação da técnica de Grafos na ferramenta Lod Live (2019) para representar o domínio Ontologia. Na imagem, fica visível a aplicação de diferentes tipos de relações entre os conceitos através dos links que são nomeados conforme cada relação é apresentada.

Nessa técnica é escolhida uma representação visual onde se define a posição de cada vértice e aresta, onde o primeiro pode representar termos ou documentos e, o segundo, as relações semânticas ou ocorrência dos termos. A técnica de grafos pode ser utilizada para o desenvolvimento de sistemas, o que permite uma visualização mais

evidenciada da representação de conceitos e seus relacionamentos (VIEIRA, 2014).

Nessas estruturas podemos atingir um alto grau de representação devido à capacidade da técnica de Grafos em manipular os links, seja dando nome às linhas que os formam, alterando sua cor ou o seu formato dependendo da relação que se pretende representar. Podemos aplicar essa técnica de visualização da informação em links de nós hipermídia, redes semânticas e teias que possibilitam pesquisadores encontrar caminhos mais curtos ou baratos para interligar itens localizados numa rede.

3 CONSIDERAÇÕES FINAIS

A modelagem conceitual auxilia na construção de instrumentos voltados para a classificação, a representação e a recuperação da informação. O produto do processo de modelização consiste na representação de uma ideia existente em determinada realidade. Elaborar uma representação consistente, que expresse efetivamente as relações entre os conceitos de um domínio, promove aumento cognitivo no usuário de um sistema, tornando a recuperação da informação satisfatória ao permitir a manifestação de padrões e agrupamentos de dados, bem como de informações camufladas em grandes volumes de dados.

As estruturas de árvore e de grafos são formas úteis para se representar classificações. Ainda que haja ressalvas de alguns autores quanto às técnicas de visualização que utilizam os princípios hierárquicos da árvore, a técnica cumpre sua função de representação em contextos mais simples onde não prevalecem relações mais complexas que as partitivas ou de gênero-espécie enquanto a representação através dos grafos que permitem demonstrar conceitos e suas relações em contextos mais complexos.

REFERÊNCIAS

CAMPOS, M. L. A. **A organização de unidades do conhecimento em hipertextos: o modelo conceitual como um espaço comunicacional para a realização de autoria.** 2001. Tese (Doutorado em Ciência da Informação) - Conselho Nacional de Desenvolvimento Científico e Tecnológico, Instituto Brasileiro de Informação em Ciência e Tecnologia, Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

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

DIAS, M. P. **A contribuição da visualização da informação para a Ciência da Informação**. Dissertação (Mestrado em Ciência da Informação) – Pós-Graduação em Ciência da Informação, Centro de Ciências Sociais Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, 2007.

FREITAS, C. M. D.S. *et al.* Introdução à visualização de informações. **RITA: Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 8, n. 2, p. 143- 158, out. 2001.

HEER, J.; BOSTOCK, M.; OGIEVETSKY, V. A tour through the visualization zoo: a survey of powerful visualization techniques, from the obvious to the obscure. **ACM Queue**, v. 8, n. 5, p. 1-22, maio 2010.

LOD LIVE. Disponível em: <http://en.lodlive.it/?http://dbpedia.org/resource/Ontology>. Acesso em: 19 set. 2019.

OLIVE. *OLIVE*: Oline Library of Information Visualization Environments. Disponível em: <http://otal.umd.edu/Olive/>, Acesso em: 23 abr.2018.

SALES, L. F.; SAYÃO, L. F.; MOTTA, D. F. Modelagem de relações conceituais para a área nuclear. *In*: SEMINAR ON ONTOLOGY RESEARCH IN BRAZIL, 5., INTERNATIONAL WORKSHOP ON METAMODELS, ONTOLOGIES AND SEMANTIC TECHNOLOGIES,7., 2012, Recife. **Proceedings** [...]. Recife, 2012.

SIMÕES, K. O. **Proposta de modelagem conceitual para organização da informação da Biblioteca Virtual em Saúde prevenção e controle de câncer**. 2016. 150 f. Dissertação (Mestrado Profissional em Biblioteconomia) - Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2016.
Disponível em: <http://www.unirio.br/ppgb/arquivo/katia-de-oliveira-simoes>. Acesso em: 10 abr. 2018.

SHNEIDERMAN, B. The eyes have it: a task by data type taxonomy for information visualizations. *In*: SYMPOSIUM ON VISUAL LANGUAGES. 1996, Boulder, **Proceedings** [...]. Boulder: IEEE, 1996.

SOLANO, V. O. **Cientometria e modelização de domínios de conhecimento**: análise da produção técnica da Embrapa Pantanal no período de 2007-2012. 2014. 181 f. Dissertação (Mestrado em Ciência da Informação) –Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

FEDLI, Gian Carlo. **Metaphors of order and disorder**: from the tree to the labyrinth and beyond. *Knowledge Organization*, n. 6, p. 366-374, 2013.

FORUM: the philosophy of classification. **Knowledge Organization**, n. 5, p. 398-404, 2011.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

FOUCAULT, M. **As palavras e as coisas**: uma arqueologia das ciências humanas. São Paulo: Martins Fontes, 2002.

MAZZOCCHI, Fulvio. Images of thought and their relation to classification: the tree and the net. **Knowledge Organization**, n. 6, p. 375-383, 2013.

VIEIRA, Jessica Monique de Lira. **A contribuição da organização e da visualização da informação para os sistemas de recuperação de informação**. 2014. Dissertação (Mestrado em Ciência da Informação)– Universidade Federal de Pernambuco, Recife, 2014.

VIEIRA, J. M. L.; CORREA, R. F. Visualização da Informação na construção de interfaces amigáveis para Sistemas de Recuperação de Informação. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 16, n. 32, 2011. Disponível em:
<http://www.periodicos.ufsc.br/index.php/eb/article/view/15182924.2011v16n32p73>. Acesso em: 12 abr. 2018