

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-2 – Organização e Representação do Conhecimento

O DESENVOLVIMENTO DE ONTOLOGIAS NA PERSPECTIVA DA SEMIÓTICA DISCURSIVA: ASPECTOS TEÓRICO-CONCEITUAIS

Daniel Libonati Gomes - Universidade Federal do Pará (UFPA)

Thiago Henrique Bragato Barros - Universidade Federal do Pará (UFPA)

THE DEVELOPMENT OF ONTOLOGIES IN THE PERSPECTIVE OF DISCURSIVE SEMIOTICS: THEORETICAL-CONCEPTUAL ASPECTS

Modalidade da Apresentação: Pôster

Resumo: O incremento na quantidade e necessidade de refinamento dos sistemas de representação leva a ciência da informação à busca as melhores maneiras de lidar com seu objeto. Diante disso, visa-se compreender a informação enquanto um discurso, analisando-o pela perspectiva da semiótica discursiva. Para tanto, estuda-se o efeito dessa abordagem no sistema de organização do conhecimento denominado ontologia, procurando entender como essa ótica pode contribuir com o seu aprimoramento e buscando, assim, a melhoria da qualidade da busca pela informação relevante. O trabalho em desenvolvimento divide-se em quatro fases: a primeira, de revisão bibliográfica sobre o uso da semiótica como ferramenta de análise na ciência da informação; a segunda, que visa à compreensão profunda das ontologias; a terceira, que busca verificar de que modo a perspectiva da semiótica discursiva pode afetar a maneira com que se lida com a informação, especialmente nas ontologias; e a quarta, cujo foco é o desenvolvimento de uma metodologia para a criação de ontologias, sendo o objeto de teste algumas ontologias retiradas do *website* Watson Semantic Web Search. A pesquisa encontra-se em fase intermediária e já pôde verificar que as abordagens do objeto da ciência da informação através das teorias semióticas (especialmente de Saussure e Peirce) até o momento encontraram limitações, visto essas teorias não oferecem conceitos que sistematizem a forma como o fenômeno da significação ocorre. Portanto o presente trabalho busca propor uma resolução a essa questão central para ciência da informação.

Palavras-Chave: Informação; Semiótica Discursiva; Discurso; Ontologia; Sistemas de Organização do Conhecimento.

Abstract: The increase in the quantity and need for refinement of representation systems leads the information science to the search of the best ways to deal with its object. The objective is to understand information as a discourse, analyzing it through the perspective of discursive semiotics. In order to do so, we study the effect of this approach in the system of knowledge organization called ontology, trying to understand how this optics can contribute to its improvement and thus seeking to improve the quality of the search for the relevant information. The work under development is divided

into four phases: the first, a bibliographical review on the use of semiotics as an analysis tool in information science; the second, which aims at deep understanding of ontologies; the third, which seeks to verify how the perspective of discursive semiotics can affect the way in which information is dealt with, especially in ontologies; and the fourth, whose focus is the development of a methodology for the creation of ontologies, being the object of testing some ontologies taken from the Watson Semantic Web Search website. The research is at an intermediary stage and has already been able to verify that the approaches of the object of information science through the semiotic theories (especially of Saussure and Peirce) until the moment they found limitations, since these theories do not offer concepts that systematize the way the phenomenon of signification occurs. Therefore the present work seeks to propose a resolution to this central issue for information science.

Keywords: Information; Discursive Semiotics; Discourse; Ontology; Systems of Knowledge Organization.

1 INTRODUÇÃO

Diante da grande produção de dados que é realizada todos os dias, especialmente após a popularização da internet, é necessário levantar a questão de como um usuário de um sistema qualquer pode encontrar a informação que lhe é mais relevante. Essa questão é uma das principais bases para o interesse no tema de pesquisa aqui proposto, assim como é base para toda a Organização e Recuperação da Informação (SARACEVIC, 2009).

Acredita-se que uma possível resposta a ela esteja no tratamento da informação como um discurso, ou melhor, um sistema significativo. Conforme demonstra o trabalho *Estudos da semiótica na Ciência da Informação: relatos de interdisciplinaridades*, de Barros e Café (2012), a maioria dos trabalhos realizados na Ciência da Informação que fazem a ligação entre a CI e a semiótica tomam a informação como um signo, utilizando, para isso, as teorias de Charles Sanders Peirce ou Ferdinand de Saussure. Contudo, considerando aqui tais dados, acredita-se que essas perspectivas sógnicas pouco revelam acerca de como é possível trabalhar o objeto informativo, visto que ambas as teorias, apesar de considerarem a significação um fenômeno sociocultural, não oferecem conceitos que especifiquem de maneira sistemática como realmente ocorre esse fenômeno. Como afirmam Alves, Moraes e Almeida (2014, p. 55):

Dentre as diversas linhas semióticas, a Semiótica Greimasiana parece a que mais elementos fornece aos estudos textuais, bem como a identificação de conceitos. Diferentemente da abordagem de Charles Peirce, com orientação lógica, e da abordagem clássica de Ferdinand de Saussure, cujo objetivo foi lançar as bases estruturais da Semiologia francesa.

Após Saussure, o conceito de signo linguístico foi aprimorado por diversos outros estudiosos, dentre os quais é possível destacar o dinamarquês Louis Hjelmslev. Ele formulou as noções de plano de expressão e plano de conteúdo para distinguir as partes que constituem

o signo linguístico, ampliando assim as noções de significante e significado, formuladas pelo linguista genebrino (FONTANILLE, 2012). Partindo disso, Fontanille propõe adotar, para compreender de fato como a significação ocorre, a perspectiva da linguagem, não do signo, de maneira que os planos de expressão e conteúdo entram em cena e levam à noção da semiótica discursiva de sistemas de significação.

A semiótica discursiva expõe a estrutura de um discurso, seus elementos constituintes, do menos ao mais complexo, através de uma ferramenta chamada Percurso Gerativo de Sentido (PGS), o qual divide o discurso em três níveis: fundamental, narrativo e discursivo. Diante disso, se a informação é um discurso, não apenas um signo, é possível verificar sua estrutura.

Assim, de posse desses conhecimentos, é possível esclarecer o objetivo desta pesquisa e o que a motivou. Conhecendo a estrutura do objeto informativo, suas características de produção e em quais situações ele pode ser requisitado (vistas principalmente nos níveis discursivo e narrativo do PGS), a construção de um sistema de organização pode ser feita de forma mais eficaz.

Para verificar como isso pode ser posto em prática, optou-se por trabalhar com as ontologias, que são sistemas de organização e representação do conhecimento que trabalham com domínios de conhecimento, inserindo as informações que compõem esses domínios em classes e relacionando-as. No que se refere a sua utilidade prática, as ontologias, em geral, são usadas no projeto denominado Web Semântica, que visa a facilitar o contato entre um computador e um usuário. O que ocorre, basicamente, é a utilização de *softwares* que permitem ao computador compreender as ontologias e interpretá-las, levando o usuário a ter acesso a documentos e objetos específicos.

Assim, a opção pela semiótica discursiva fica mais clara: considerando a informação como um discurso e conhecendo sua estrutura, é possível aplicar essa abordagem às ontologias, um sistema de organização de caráter também estrutural. Por exemplo, pode-se realizar uma abordagem parecida àquela aplicada no artigo *Semiótica do discurso científico: um estudo sobre novas perspectivas para a análise documental de conteúdo*, de Alves, Moraes e Almeida (2014), em que os autores investigam através da semiótica discursiva o discurso científico, descrevendo os elementos do nível narrativo do PGS, ou seja, a conjunção de um sujeito, o cientista, com um objeto, o conhecimento acerca da algo. A análise que a presente pesquisa propõe é similar, considerando o sujeito como sendo o usuário de um sistema de

informação e o objeto como sendo um objeto informativo qualquer. Outra aplicação pode ser na relação entre ontologias diferentes, a interoperabilidade, visto que, através do nível fundamental do PGS é possível compreender as bases de um discurso, de modo que é possível relacionar discursos diferentes através delas. Essa mesma lógica pode ser aplicada a ontologias: diferentes domínios, acredita-se, podem ser relacionados através de seus elementos fundamentais, expostos pela análise através do PGS.

Diante disso, é cabível acreditar que o estudo das ontologias tendo como fundamento a semiótica discursiva pode agregar muito ao desenvolvimento desses sistemas de organização da informação, dado o fato de que essa disciplina expõe de maneira bastante clara a relação entre a informação produzida e sua relevância para um determinado usuário, tudo isso visto na própria estrutura da informação (visualizada através do PGS), o que se soma à noção de cooperação entre homem e máquina trazida pela Web Semântica.

2 ASPECTOS TEÓRICOS E METODOLÓGICOS DA PESQUISA

2.1 O Signo, os Sistemas Significativos e as Ontologias

Um dos principais conceitos tratados na pesquisa é o de signo. Para tanto, é necessário destacar duas teorias, a de Ferdinand du Saussure e a de Charles Sanders Peirce. O primeiro, ao tratar sobre signo linguístico, dividiu-o em duas partes interdependentes: o significante e o significado. O significante é a “imagem acústica”, a imagem formada na mente de um sujeito a partir de uma série de sons por ele ouvidos, enquanto que o significado é a ideia, o conceito. Assim, quando um sujeito ouve alguma palavra, em sua mente aqueles sons formam uma imagem, que é atrelada a um conceito. As línguas para Saussure, portanto, são sistemas de signos e, conseqüentemente, as informações também caracterizam signos (SAUSSURE, 2006). No caso do signo de Charles Peirce, o autor definiu que toda a realidade seria apreendida através de três categorias: a primeiridade, que se refere ao sentimento, à consciência de que o ser humano está inserido no tempo; a secundidade, que é a consciência da existência daquilo que é externo; e a terceiridade, que se refere a forma como o ser humano, unindo a primeiridade e a secundidade, reconhece o mundo, apreende-o (BARROS & CAFÉ, 2012).

Contudo a pesquisa, como já explanado na Introdução, trabalhará com a noção de sistemas de significação. Para tanto, é necessário lembrar que foram desenvolvidos por Hjelmslev os conceitos de plano de expressão e plano de conteúdo da linguagem, que dão

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

margem ao desenvolvimento dos chamados sistemas de significação. Assim, é justamente a partir das noções de plano de expressão e plano de conteúdo que Fontanille (2012), baseando-se no trabalho de Algirdas Julien Greimas, propõe investigar o fenômeno da significação, visto que a visão do signo limita o trabalho, não esclarecendo, de fato, como é possível que um sujeito apreenda o sentido de algo. Portanto é essa postura que é proposta aqui: investigar a informação não como um signo, mas como um sistema significativo. Sobre isso, assim explica Baquião (2011, p. 53):

Entende-se significação como a apreensão das relações entre os elementos da linguagem: é o ato por meio do qual o mundo faz sentido, significa. De modo que a significação é um conceito primordial na relação entre homem e mundo. Assim, a semiótica conceitua o texto como um conjunto formal de significação que se manifesta em diversas substâncias da expressão: verbais, visuais, audiovisuais, esculturais, arquitetônicas etc.

Diante disso, quando o signo, a informação, é tomado como sendo apenas uma união entre imagem acústica e conceito, pouco se diz acerca do contexto em que aquela informação é necessária ou produzida, ou seja, os dados são limitados. Agora, se a informação for tomada como um sistema de significação, entra em cena toda a questão contextual que a envolve, o que é importante, visto que um dos principais objetivos da Recuperação da Informação é a busca pelo elemento da relevância. Conforme afirma Saracevic (1996, p. 56): “[...] desde os primórdios da CI, a noção de eficácia (por exemplo, a comunicação eficaz do conhecimento, o acesso eficiente aos recursos informacionais, a relevância e utilidade da informação, a qualidade da informação) tem sido uma preocupação central”.

Esses sistemas de significação são o objeto de estudo da semiótica discursiva (também chamada de semiótica de linha francesa ou greimasiana). Assim, acredita-se que a utilização da metodologia desenvolvida por Greimas na semiótica discursiva pode ter um importante papel na área da Organização da Informação, visto que essa metodologia dá margem para o entendimento da estrutura de qualquer tipo de texto (discurso), desde sua produção até suas unidades fundamentais. Para isso, a semiótica discursiva utiliza em suas análises o método chamado Percurso Gerativo de Sentido (PGS) (BARROS, 2005), o qual se baseia na ideia de que todo texto é constituído por três níveis com graus variáveis de complexidade e profundidade. Baquião resume os três níveis do PGS (p. 53-54):

- *Nível fundamental*: no qual se situam as relações lógicas – representadas por caracteres simbólicos – com investimento semântico mínimo;

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

- *Nível narrativo*: no qual ocorre a organização e o desenvolvimento semântico dos elementos lógicos que estruturam a narrativa;
- *Nível discursivo*: no qual os níveis anteriores se manifestam temática e figurativamente; nele se examinam os investimentos mais concretos, que se manifestam em diversos textos: verbais, visuais, sincréticos etc. Nesse nível, o sujeito da enunciação – produtor do discurso – organiza as estruturas narrativas em categorias discursivas de pessoa, tempo e espaço.

Como é possível notar na citação acima, o nível discursivo é o mais complexo e aquele que está mais próximo da realidade (mais superficial), sendo onde ocorre a relação entre estruturas mais simples sem contexto bem definido e noções como ideologia, sociedade, momento histórico e circunstância; o nível narrativo é intermediário em profundidade e complexidade, sendo o nível em que serão verificados os elementos motivadores de um texto, como seus elementos constitutivos interagem; e o nível fundamental é o mais profundo (longe da realidade) e simples, sendo formado apenas por estruturas semânticas básicas (actantes) que se opõem uma a outra e norteiam o andamento de um texto (por exemplo, *vida e morte*) (BARROS, 2005).

Acerca das ontologias, é importante destacar que elas atuam da forma definida por Vital e Café (2011, p. 118): “A partir de uma determinada área de domínio (campo do conhecimento que se deseja representar), a ontologia se propõe a classificar as coisas em categorias, na perspectiva do sujeito e da linguagem do domínio”. Ainda, segundo Carlan e Medeiros (2011, p. 58):

As ontologias promovem e facilitam a interoperabilidade entre sistemas de informação. Por meio de um processo “inteligente” dos computadores, é possível compartilhar e reutilizar o conhecimento entre os sistemas. Na literatura alguns autores definem ontologias como linguagens documentárias pelos seus elementos de formação: termos, definições e relações. Porém, apesar de possuírem elementos comuns que outros SOC, as ontologias são mais que linguagens documentárias, elas possuem funcionalidades que permitem que máquinas possam processar o “raciocínio” automatizado por meio de regras e inferências.

Assim, percebe-se que o estudo das ontologias possui especial importância, principalmente por sua presença em sistemas de informação, os quais são, atualmente, uma constante presença no mundo.

2.2 Procedimentos Metodológicos

A presente pesquisa tem caráter qualitativo, sendo a investigação realizada em quatro fases, de modo que cada uma seja consequência direta do que foi realizado na fase anterior. A primeira fase (que está atualmente sendo realizada) se volta à busca de informações em trabalhos anteriores que expõem o contato da semiótica (levando em consideração qualquer corrente de pensamento dessa disciplina) com a Ciência da Informação e, mais especificamente, com ontologias, sendo, portanto, uma fase de revisão bibliográfica.

A segunda fase se foca no estudo das ontologias propriamente ditas, sua estrutura, importância e desenvolvimento (com foco nas metodologias e nas linguagens utilizadas na representação). Isso feito, iniciar-se-á terceira fase: de posse dos dados obtidos na primeira e segunda fase, será estudado como a semiótica discursiva pode realizar o tratamento das informações a serem inseridas numa ontologia. Para tanto, será verificado de que forma a estrutura de uma ontologia se comporta a partir de uma abordagem feita através dos níveis do Percurso Gerativo de Sentido, destacando a forma como os dados referentes às informações – chamados de *conceitos*, como propõe Gruber (2005 apud MORAIS & AMBRÓSIO, 2007) – serão expostos. Por exemplo, ao serem analisadas as classes de uma ontologia, sabe-se, pela Semiótica Discursiva, que uma determinada classe sempre possuirá relação com outra – oposta, intermediária ou semelhante – (nível fundamental) e que essas relações são guiadas com base nas motivações que levam um usuário a escolher uma ou outra classe específica (nível narrativo), motivações essas fundamentadas em algum fator sócio-histórico-ideológico (nível discursivo). Com isso, objetiva-se trazer à tona, nas ontologias, a relevância que uma informação pode ter para os usuários e de que modo, a partir dessa ótica, é possível relacionar ontologias de diferentes domínios.

Na quarta fase, será então criada uma metodologia de desenvolvimento de ontologias, a qual procurará evidenciar questões envolvendo contextos de busca por uma determinada informação e que benefícios ela pode trazer à reformulação de ontologias já criadas, tomadas da página Watson Semantic Web Search.

3 CONSIDERAÇÕES FINAIS

O presente trabalho ainda está em pleno desenvolvimento, de maneira que tem se focado em obter mais informações acerca do tratamento do objeto informativo como um sistema significativo. Notou-se que, de fato, a maior parte dos trabalhos (ao menos aqueles que

se preocupam em destacar o modo de tratamento da informação) em ciência da informação abordam seu objeto como um signo, sendo que essa não é uma premissa recente.

O trabalho de Raber e Budd (2003), um artigo intitulado *Information as sign: semiotics and information science*, é um exemplo disso. Nele, os autores fazem uma comparação entre a informação e o sistema linguístico de Saussure, chegando a algumas conclusões interessantes. Contudo, mesmo essa abordagem acaba tendo o mesmo problema descrito na introdução: não descreve como realmente ocorre a apreensão da informação, dando pouca margem a um aprimoramento real da interação com esse objeto.

Portanto acredita-se que a presente pesquisa poderá muito contribuir tanto com a questão epistemológica do objeto da ciência da informação quanto com as áreas da organização e recuperação da informação, visto que busca um real aprimoramento de um dos mais presentes sistemas de organização.

REFERÊNCIAS

ALVES, R. C. V.; MORAES, J. B. E.; ALMEIDA, C. C. Semiótica do discurso científico: um estudo sobre novas perspectivas para a análise documental de conteúdo. **Scire**, v. 20, n. 2, p. 55-59, 2014. Disponível em: <<http://www.iberid.eu/ojs/index.php/scire/article/viewFile/4184/3778>>. Acesso em: 3 set. 2017.

BARROS, C. M.; CAFÉ, L. M. A. Estudos da Semiótica na Ciência da Informação: relatos de interdisciplinaridades. **Perspectivas em Ciência da Informação**, v. 17, n. 3, p. 18-33, 2012. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1501/1050>>. Acesso em: 5 ago. 2017.

BARROS, D. L. P. de. **Teoria semiótica do texto**. São Paulo: Ática, 2005.

BETTENCOURT, M. P. L.; CIANCONI, R. B. Gestão do conhecimento: um olhar sob a perspectiva da Ciência da Informação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, n. 1, 2012. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/13195>>. Acesso em: 4 ago. 2017

CARLAN, E.; MEDEIROS, M. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, 2011. Disponível em: <<http://periodicos.unb.br/index.php/RICI/article/view/6209/5102>>. Acesso em: 5 ago. 2017.

FONTANILLE, J. **Semiótica do discurso**. São Paulo: Contexto, 2012.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens**. Goiás: [S.l], 2007. Disponível em:
<http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-07.pdf>. Acesso em: 7 ago. 2017.

PINHEIRO, L. V. R. Campo interdisciplinar em ciência da informação: fronteiras remotas e recentes. **Investigación Bibliotecológica**, v. 12, n. 25, p.132-163, 1998. Disponível em:
<<http://rev-ib.unam.mx/ib/index.php/ib/article/view/3884/3436>>. Acesso em: 8 ago. 2017.

RABER, D.; BUDD, J. M. Information as a sign: semiotics and information science. **Journal of Documentation**, v. 59, n. 5, p. 507-522. Disponível em:
<<http://www.emeraldinsight.com/doi/pdfplus/10.1108/00220410310499564>>. Acesso em: 15 ago. 2017.

SANTAELLA, Lúcia. **O que é semiótica**. São Paulo: Brasiliense, 1985.

SARACEVIC, T. Information science. BATES, M.; MAACK, M. N. (Eds.) **Encyclopedia of Library and Information Sciences**. Nova York:Taylor & Francis. p. 2570-2586, 2009. Disponível em:
<<https://tefkos.comminfo.rutgers.edu/SaracevicInformationScienceELIS2009.pdf>>. Acesso em: 1 ago. 2017.

_____. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 1, n. 1, 1996, p. 41-62. Disponível em:
<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>>. Acesso em: 6 ago. 2017.

SAUSSURE, F. **Curso de linguística geral**. São Paulo: Cultrix, 2006.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas Contribuições para a Ciência da Informação. **Ciência da Informação**. Brasília, vol. 33, n. 1, p. 132-141, 2004. Disponível em:
<<http://www.scielo.br/pdf/ci/v33n1/v33n1a16.pdf>>. Acesso em: 10 ago. 2017.

VITAL, L. P; CAFÉ, L. M. A. Ontologias e Taxonomias: diferenças. **Perspectivas em Ciência da Informação**, v. 16, n. 2, p. 115-130, 2011. Disponível em:
<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/200/927>>. Acesso em: 7 ago. 2017.