

Da análise documentária à terminologia: percurso teórico e metodológico



Vânia Mara Alves Lima
Universidade de São Paulo
vamal@usp.br

O objetivo desta comunicação é apresentar o percurso teórico e metodológico percorrido em nossa pesquisa sobre a utilização da terminologia de um domínio do conhecimento como referente para a construção e gerenciamento de linguagens documentárias, mais especificamente os vocabulários controlados e os tesouros.

Esse problema de pesquisa começa a se delinear na década de 90 durante nossa prática profissional no processamento técnico de um dos maiores acervos bibliográficos da Universidade de São Paulo, o Serviço de Biblioteca e Documentação da Faculdade de Filosofia, Letras e Ciências Humanas, que nos apresentava diariamente questões sobre a organização, representação e recuperação de informações produzidas sobre um mesmo assunto por diferentes domínios do conhecimento, e que conseqüentemente abordavam diferentes pontos de vista.

Quando nos transferimos para o Serviço de Biblioteca e Informação da Faculdade de Arquitetura e Urbanismo da USP, nos deparamos com outro problema relacionado à representação da informação, ali o conhecimento produzido em um mesmo domínio era representado por diferentes designações em função, muitas vezes, do suporte em que se encontrava registrado, isto é, livros, mapas, projetos de arquitetura, slides, fotos, artigos de revista.

A busca de fundamentação teórica para a solução de problemas práticos nos levou aos trabalhos de Smit, Tálamo, Cunha, Lara, Kobashi, Guimarães e Fujita, fundadores do Grupo Temma que tem na obra de Jean Claude Gardin seus alicerces teóricos, mas que ampliaram consideravelmente os horizontes da denominada análise

documentária, seus processos e seus produtos, principalmente no que se refere à construção das linguagens documentárias.

Os estudos do Grupo Temma demonstraram que a Análise Documentária definida por Gardin (1981) como “um conjunto de procedimentos para expressar o conteúdo dos documentos científicos sob formas destinadas a facilitar a consulta” é uma disciplina metodológica que sugere procedimentos para a análise de textos com o objetivo de selecionar conteúdos informativos os quais poderão ser representados, recuperados e disseminados (TÁLAMO, LARA, KOBASHI, LIMA, 1992 ; LARA, 2011) independente do suporte em que são registrados. Ao conceito de linguagem documentária como “um conjunto de termos utilizados para representar certo conteúdo de documentos científicos com fins de classificação ou de pesquisa retrospectiva de informações (CROS, GARDIN, LEVY 1968)”, o Grupo Temma acrescenta outras características que enfatizam seu caráter de “linguagem”, como uma estrutura, cujos termos devem necessariamente estar relacionados para que possam significar de modo determinado, funcionando como veículo de comunicação que represente os campos conceituais respeitando a cultura da comunidade a qual serve (VOGEL, 2004).

Neste quadro é que se insere nossa abordagem linguística e terminológica tanto sobre a elaboração da informação documentária, entendida como o resultado da representação do conteúdo de um documento, quanto da utilização das terminologias dos domínios como referente para a construção de linguagens documentárias.

No que se refere à informação documentária, procuramos esquematizá-la a partir das aproximações com os referenciais linguísticos, identificando-a como um signo documentário, analogamente ao signo linguístico, definido por Peirce (1977) como aquilo que, sob certo aspecto ou modo, representa algo para alguém e está no lugar de alguma outra coisa sob algum aspecto ou capacidade, ou seja, a informação documentária se propõe no lugar do conhecimento registrado e, portanto da mesma maneira que o signo linguístico é passível do processo de semiose, denominada por Lara (2006) de semiose documentária..

Entendendo-se a semiose como a construção do significado pelo intérprete, no interior de um contexto determinado, pode-se inferir que a semiose documentária é a construção de significados fundamentados em referenciais terminológicos os quais remetem às estruturas conceituais dos domínios. Portanto, as referências contextuais

de produção da informação documentária são essenciais para que ela efetivamente represente um conjunto de enunciados verdadeiros sobre o conhecimento registrado, os quais são compilados nas definições de cada termo presente na terminologia de um domínio.

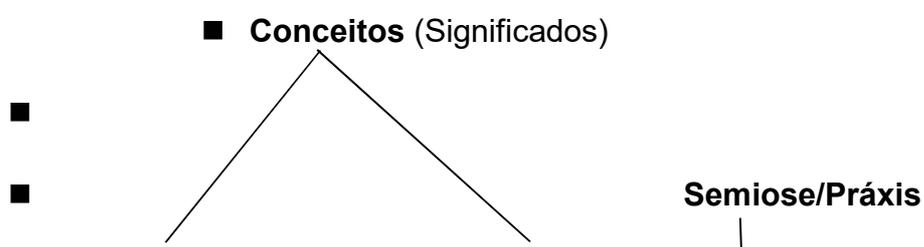
Devemos esclarecer que quando falamos em terminologias estamos nos referindo à terminologia enquanto produto, isto é, ao conjunto de termos de uma especialidade, o qual é elaborado, através da utilização de normas terminológicas, propostas pela Terminologia enquanto disciplina que se ocupa dos termos especializados (CABRÉ, 1995).

Enquanto a terminologia do domínio é o reflexo formal da organização conceitual de uma especialidade e meio inevitável de expressão e comunicação profissional, o qual garante a transmissão do conhecimento, a linguagem documentária tem por função normalizar a busca e garantir a recuperação desse conhecimento registrado através da elaboração da informação documentária. Assim, entende-se que esses dois instrumentos são complementares.

Segundo Gardin (1981), informação documentária é produto da análise documentária, resultado de uma operação semântica formulada dentro de uma linguagem documentária, que transforma um texto original em uma ou mais palavras chaves, e que mesmo apresentando, na linguagem documentária, a mesma forma que na linguagem natural, não tem necessariamente todas as acepções presentes em um dicionário geral da língua

Nesse contexto inferimos a informação documentária enquanto uma representação do conteúdo do documento, a partir dos conceitos do domínio a que pertence, os quais designados pelos termos desse domínio servem como referente para os descritores da linguagem documentária, pois contextualizam a sua significação a partir de uma práxis nesse domínio (LIMA, 1998).

Na figura a seguir esquematizamos a informação documentária representada pela tríade conceito/termo/descritor resultado de processo de semiose documentária aliada à práxis em um domínio do conhecimento.



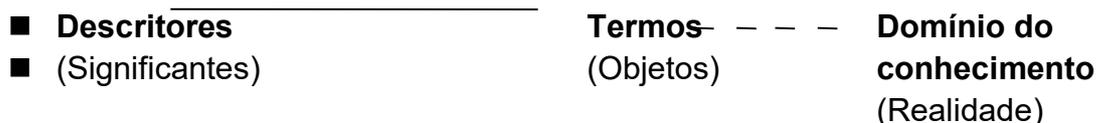


Figura 1 – A informação documentária

Esse esquema da informação documentária atende à colocação de Smit (2000, p.28) de que para que a informação documentária funcione efetivamente e corretamente como mediadora, entre o usuário e o estoque de informação, o usuário deve ter condições para contextualizá-la e decodificá-la, pois ao utilizar a terminologia do domínio como referente para a linguagem documentária, possibilita ao usuário o reconhecimento dos termos que fazem parte de sua práxis. Dito de outro modo, codificamos a informação documentária a partir de referências que serão reconhecida, decodificadas e interpretadas pelo usuário no processo de semiose documentária.

A garantia de qualidade na representação e da recuperação da informação depende da efetiva a codificação e a decodificação da informação documentária, mas não está restrita a compreensão de sua estrutura, ela deve considerar também algumas tensões dialéticas entre forças contrárias presentes quando do gerenciamento e manutenção da linguagem documentária, como por exemplo, a **conservação x mudança e consenso x especificidade** (LIMA, 2004).

A **conservação** assegura o entendimento entre os sujeitos, e a **mutação**, satisfaz as necessidades de mudança da sociedade, ou seja, no processo de gerenciamento e manutenção de uma linguagem documentária deve-se manter pontos de acesso de sistemas anteriores ao mesmo tempo em que deve possibilitar a inclusão de novos pontos de acesso (LIMA, 2004).

A linguagem documentária deve atender também tanto ao **consenso** quanto a **especificidade**, isto é, de um lado atender a sua característica de ser institucional, pois sempre é construída para ser utilizada em função dos objetivos de determinada instituição assegurando a intercompreensão dos sujeitos e, de outro lado, oferecer elementos de uma experiência específica a cada usuário (LIMA, 2004).

A necessidade da linguagem documentária em garantir o consenso ao mesmo tempo em que atende a especificidade exige de seus descritores uma precisão e coerência que só é alcançada a partir da compilação de enunciados verdadeiros que vão explicitar seu significado e tornar explícitas as relações semânticas que articulam sua rede conceitual. Em outras palavras, é necessário identificar as características que formam o conceito, cuja designação (termo) é referente para o descritor que

atribuído a um documento constitui a informação documentária. Enfim, o sentido da informação documentária é expresso pela definição do conceito que ela denomina.

Por sua vez, a definição do conceito reúne atributos, ou características, que permitem estabelecer as categorias de uma linguagem documentária. Essas categorias se definem pelo traço comum presente a toda uma classe de conceitos/termos/descriptores que por isso mesmo estão associados. O limite da cada categoria se estabelece pelos traços específicos que permitem individualizar cada conceito/termo/descriptor estabelecendo a disjunção entre as informações documentárias elaboradas.

Aqui incluímos no percurso a noção de classe semântica que também foi abordada por Gardin (1966) quando na descrição do léxico documentário ao se referir à necessidade de organizar os seus termos de modo a explicitar as relações hierárquicas existentes entre eles, seja por afinidade ou diferença semântica.

Determinar o campo semântico em linguística é, segundo os pressupostos epistemológicos, procurar descobrir a estrutura de um domínio dado, ou propor a ele uma estruturação (Dubois, et al, 2011). Trier, citado por Lopes (1987), observa que as unidades léxicas de uma língua encontram-se organizadas em grupos estruturados de tal modo que, cada unidade fica ali definida pelo lugar que ocupa respectivamente à posição das demais. Assim, inferimos que o sentido de um conceito/termo/descriptor é especificado por sua semelhança e sua diferença em relação aos outros elementos relevantes do campo semântico, pois uma palavra só adquire seu significado por oposição as outras unidades do mesmo campo (Germain, 1981).

Genouvrier e Peytard (1974) no que se refere ao campo semântico, o definem como o conjunto de empregos de uma palavra (ou sintagma, ou lexia) onde e pelos quais a palavra adquire uma carga semântica específica e a delimitação destes empregos se daria pelo levantamento de todos os contextos imediatos que a palavra recebe num texto dado.

Segundo Hernando Cuadrado (1995) a condição mínima para que as palavras pertençam a um mesmo campo é que tenham um traço significativo (sema) comum (quanto maior o número de semas tanto mais coerente será o campo semântico e no geral um menor número de palavras o integra); uma palavra pode tomar parte de todos os campos semânticos que se construam sobre qualquer dos traços significativos que se descubram nela; quando uma palavra possui vários significados, cada um deles pertence a um campo semântico distinto.

Por exemplo, ao mesmo tempo em que podemos identificar com pertencentes ao campo semântico da palavra *mesa*, devido à característica comum “*objeto que permite reunir em torno de*” as seguintes palavras: *mesa de refeições*, *mesa redonda*, *mesa da Assembléia*, *mesa de operação*, cada uma poderá fazer parte de uma categoria diferente, em uma linguagem documentária, devido à características delimitadoras como: “para refeição”; “para debate”; pra cirurgia”, etc. Em outro caso, a palavra *lodo* por ser incluída em diferentes categorias de uma linguagem documentária a partir de características delimitadoras, a saber, *ser uma matéria prima*; *ser um produto*, *ser um reagente*.

No momento em que se discute a interoperabilidade entre as linguagens documentárias na web semântica acreditamos ser necessário aprofundar as discussões sobre o mapeamento dos campos semânticos que constituem um domínio, pois só a partir das características do conceito elencadas nas definições dos termos, que vão servir de referente e permitir a contextualização dos descritores das linguagens documentárias, será possível tornar mais eficaz a representação e a recuperação da informação.

Referências

CABRÉ, M. T. La terminologia hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, Brasília, v.24, n.3, p.289-298, set./dez. 1995.

CROS, R.C.; GARDIN, J-C; LEVY, F. **L'automatisation des recherches documentaires: un modele general: le Syntol**. Paris: Gauthiers-Villars, 1968.

DUBOIS, J. ET AL. **Dicionário de linguística**. São Paulo: Cultrix, 2011.

GARDIN, J. C. Éléments d'un modele pour La description des lexiques documentaires. **Bulletin des Bibliothèques de France (en ligne)**. n.5, 1966. Disponível em <http://bbf.enssib.fr/consulter/bbf-1966-05-071-001>, Acessado em 20 jun. 2015.

GARDIN, J. C. **La logique du plausible**. Paris: Editions de la Maison des Sciences de l'homme, 1981.

GENOUVRIER, E., & PEYTARD, J. **Linguística e ensino do português**. Coimbra: Almedina, 1974.

GERMAIN, C. (1981). **La sémantique fonctionnelle**. Paris: PUF, 1981.
Hernando Cuadrado, L. A. **Introducción a la teoría y estructura del lenguaje**. Madrid : Editorial Verbum, 1995.

LARA, M. L. G. Conceitos de organização e representação do conhecimento na ótica das reflexões do Grupo Tema. **Inf. Inf.**, Londrina, v. 16. n. 3. p. 92 – 121, jan./ jun. 2011.

LARA, M. L. G. É possível falar em signo e semiose documentária? **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, 2º número esp., 2º sem 2006

LARA, M. L. G. **Representação e linguagens documentárias: bases teórico-metodológicas**. (doutorado), Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1999.

LIMA, V.M.A. **Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de codificação e enunciação de decodificação da informação documentária**. (doutorado), Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo

LIMA, V.M.A. **Terminologia, comunicação e representação documentária**. 1998. Dissertação (Mestrado) Escola de Comunicações e Artes, Universidade de São Paulo, 1998

LOPES, E. **Fundamentos da linguística contemporânea**. São Paulo: Cultrix, 1987. SMIT, J. W. Informação. In: LIMA, Y.D.; SMIT, J.W. (Org.) **Organização de arquivos**. São Paulo: IEB/ECA, 2000. p.19-31

PEIRCE, C.S. **Semiótica**. São Paulo: Perspectiva, 1977.

TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N.Y. & LIMA, V.M.A. Instrumentos de controle terminológico: limites e funções. **Actas do II Simpósio Latino-americano de Terminologia**, 1990. Disponível em: <<http://www.riterm.net/actes/2simposio/indice90.htm>>. Acesso em 20 de junho de 2015.

VOGEL, M.J.M. A influência de Jean-Claude Gardin e a linha francesa na evolução do conceito de linguagem documentária. **Perspectivas em Ciência da Informação**, v.14 , número especial, p.80-92, 2009.