

A análise facetada na modelagem conceitual de sistemas de hipertexto: uma revisão de literatura

Gercina Ângela Borém Lima

A partir da Teoria da Classificação Facetada, busca resgatar um embasamento para a organização conceitual em sistemas de hipertexto. São apresentados os princípios dessa teoria e revistos trabalhos de autores da área da Ciência da Informação que descrevem seus estudos teóricos ou suas experiências.

Palavras-chave: *Análise facetada, Modelagem conceitual, Hipertexto*

Recebido em 28/06/02 - Aceito em 6/08/02

189

Introdução

Desde o final da década de 1980, diversos pesquisadores começaram a estudar a possibilidade de se utilizarem as teorias da classificação na organização conceitual de sistemas de hipertexto, principalmente a classificação facetada, criada pelo indiano Ranganathan. Como a classificação facetada, os sistemas de hipertexto também objetivam a estruturação do conhecimento, através da organização dos conceitos e das relações entre eles, permitindo o mapeamento de uma área de assunto e a inclusão de novos conceitos, sem que isto signifique uma mudança estrutural do sistema.

Entre os mais recentes avanços que têm surgido para ajudar os autores de hipertextos numa organização semântica, estão o modelo *Resource Description Framework (RDF)* e a linguagem *eXtensible Markup Language (XML)*, propostas pelo grupo *World Wide Web Consortium (W3C)*. A arquitetura RDF é um modelo que permite a representação de dados com um vocabulário distinto para a modelagem da informação. O XML é uma linguagem que fornece um conjunto extensível de marcações que podem ser utilizadas para capturar a estrutura semântica do documento.

Embora se saiba que os processos de autoria de hiperdocumentos têm sido estudados por vários profissionais da área da ciência da computação, como aqueles do grupo *World Wide Web Consortium (W3C)* que, mais recentemente, têm tido a preocupação de inserir o conteúdo semântico nas páginas Web, este artigo não contemplará os trabalhos dessa área. A intenção desta revisão concentra-se na relevância da teoria da análise facetada em relação à modelagem hipertextual, seja ela *offline* (CD-ROM, redes locais) ou *online* (WWW, redes externas)

¹ Professora da Escola de Ciência da Informação da UFMG. Doutoranda do Programa de Pós-Graduação em Ciência da Informação da Escola de Ciência da Informação da UFMG. e-mail: glima@eci.ufmg.br - Site: www.eci.ufmg.br/glima

A teoria da análise facetada de Ranganathan

Shiyali Ramamrita RANGANATHAN (1892-1972), matemático indiano que tornou-se bibliotecário, foi um dos estudiosos que mais contribuiu para a teoria da biblioteconomia no século XX, especialmente na área da classificação de assunto. Sua insatisfação com os sistemas de Classificação Decimal de Dewey (CDD) e Classificação Decimal Universal (CDU) o levou a buscar um sistema mais eficiente. Depois de constatar, em cinco periódicos diferentes, que a maioria dos assuntos tratados eram assuntos compostos, Ranganathan organizou, em 1933, sua Classificação de Dois Pontos (*Colon Classification*), também conhecida como Classificação em Facetas ou Classificação Analítico-Sintética. Sob influência de sua área inicial, a matemática, estruturou o conhecimento de maneira que os assuntos compostos, sinteticamente, surgiam a partir de conceitos elementares. Publicou seis edições desse sistema, vindo a falecer em 1972, quando sua sétima versão estava para ser publicada. Além desse sistema de classificação, RANGANATHAN (1945; 1962) publicou *Elements of Library Classifications* e *Prolegomena to Library Classification* (1967), obras consideradas referenciais pelos teóricos da classificação. Entre os princípios introduzidos por Ranganathan, o mais conhecido é o princípio de análise de faceta (subdivisões de assuntos em suas partes componentes) e a síntese (recombinação destas partes para classificar um documento adequadamente). Na Classificação Facetada, identifica-se cinco categorias: Personalidade (*P, de personality*), Material (*M, de material*), Energia (*E, de energy*), Espaço (*S, de space*) e Tempo (*T, de time*), mnemonicamente conhecido como PMEST. Esta ordem de citação é baseada na idéia da concretividade decrescente. Embora o sistema de Classificação de Dois Pontos não tenha sido aceito mundialmente, a teoria de análise de faceta e síntese, proposta por ele, transformou-se em uma importante base teórica para a área de análise de assunto no século XX (RANGANATHAN, 1985, p.86).

A Classificação Facetada compreende princípios e técnicas para os sistemas de organização e recuperação da informação. Um sistema facetado reconhece muitos aspectos em um único assunto, e tenta sintetizar estes aspectos de maneira que os descreva de forma mais adequada. A relação entre assuntos, feita através do símbolo de dois pontos, seria infinita, provando que o conhecimento pode ser multidimensional e que essas relações podem tomar rumos diferentes, dependendo da síntese entre vários conceitos múltiplos (VICKERY, 1980, p.209).

Para elaborar uma classificação facetada, examina-se a literatura do assunto com a finalidade de identificar seus conceitos e termos e estabelecer suas características e facetas. Após levantar e definir a terminologia própria do assunto, os termos são analisados e distribuídos em facetas. Como já foi definido, faceta é a coleção de termos que apresenta um relacionamento hierarquicamente igual com o assunto global, refletindo a aplicação de um princípio básico de divisão. As facetas obtidas são inerentes ao assunto e, dentro de cada faceta, os termos que as constituem são suscetíveis a novos agrupamentos, pela aplicação de outras características divisionais, dando origem às subfacetas. Os termos, nas subfacetas, são mutuamente exclusivos, isto é, não devem se sobrepor na formação de assuntos compostos. Estabelecidas as facetas e subfacetas, é importante determinar a ordem de citação em que serão apresentadas no sistema de classificação. A seguir, ordenam-

se todos os elementos em ordem de arquivamento, o que permite colocar o assunto geral antes do específico. Após estas etapas, o sistema está pronto para receber uma notação, que deverá ser flexível para permitir a inclusão de novas classes. Finalmente, compila-se um índice com todos os termos e suas respectivas notações (PIEDADE, 1983, p. 80; BARBOSA, 1972, p.76).

A análise facetada reconhece muitos aspectos em um único assunto, e tenta sintetizá-los de maneira que os descreva de forma mais adequada. Esse fato é relevante para a abordagem não linear do hipertexto, porque permite ao usuário ver um assunto sob diferentes perspectivas, possibilitando um reagrupamento de um mesmo conhecimento como parte do todo, ao invés de estar subordinado dentro de uma hierarquia. Essa modelagem conceitual é uma etapa importante na organização da informação para sistemas de hipertexto.

Utilização da análise facetada na modelagem conceitual de sistemas de hipertexto

O objetivo de toda classificação é estabelecer uma ordem ou organização das coisas e dos pensamentos. A classificação facetada pode mostrar um mapeamento do conhecimento científico de um determinado período, ou permitir a descoberta de conexões e analogias entre diferentes campos do conhecimento, facilitando a recuperação da informação (SPEZIALI, 1973, p.462).

Em 1945, Vannevar Bush previu mudanças na organização da informação e nos meios de recuperá-la. Segundo seu raciocínio, a mente humana trabalha por associações e, quando se depara com um item, instantaneamente o relaciona com um próximo, que é sugerido pela associação de pensamentos. Essa associação se dá pelo conceito que cada pessoa tem sobre determinado coisa ou idéia. Se a mente humana pensa rapidamente com itens familiares é, por outro lado, ponderada e muito inconstante com objetos desconhecidos. No processo de recuperação da informação de uma grande base de dados, BUSH (1945, p.29) via, como o maior desafio, não a extração, mas sim a seleção de dados. Para ele, o sistema tradicional de indexação, com sua abordagem alfabética e numérica para organizar a informação, não refletia o modo de funcionamento seletivo da mente humana. Para incorporar essa característica associativa na seleção da informação, Bush sugeriu a indexação associativa na qual, a partir de um item selecionado, pode-se chegar, automaticamente, a outro. Nesse caso, o autor de um documento pode criar *links* associativos ou caminhos entre os itens de informação.

ELLIS (2000, p.109) aponta o problema básico na criação de hipertextos, onde geralmente

"não há nenhuma proximidade entre o designer ou o criador (que poderia ser qualquer pessoa) e o usuário em potencial (que poderia ser qualquer um ou todos). . . 'quanto mais' distantes as características e necessidades de informação dos usuários estão do tipo de usuário concebido ou servido pelos que criam ou indexam uma base de dados, 'mais provável' será a existência de problemas no acesso de informações relevantes pelos usuários daquela bases de dados."

Vários autores citam a classificação facetada como um instrumento que pode auxiliar na representação do conteúdo intelectual em sistemas de hipertextos. Um dos primeiros autores a tratar deste assunto foi Elizabeth B. Duncan. No seu artigo

Perspect. cienc. inf., Belo Horizonte, v. 7, n. 2, p. 189-196, jul./dez. 2002

Structuring knowledge bases for designers of learning materials DUNCAN (1989, p.20), apresenta um modelo de aplicação direta da análise facetada no *design* de hipertexto, onde sugere que a análise de facetas tem pontos comuns com a teoria lingüística, entendendo que uma combinação entre facetas e análise semântica pode servir de base para o desenvolvimento da tecnologia do hipertexto. As categorias determinadas pela análise das facetas constituem-se nos diferentes pontos de vista do assunto². A autora enfoca três aspectos para a organização e representação do conhecimento: **a) o mapa conceitual**, que representa os conceitos de um domínio do conhecimento, **b) a análise facetada**, que representa os diferentes pontos de vista que o assunto pode ser tratado e **c) links** e seus tipos e a relação entre eles. Em outro trabalho, *A faceted approach to hypertext?*, DUNCAN (1989, p.157) relata uma abordagem da teoria facetada em sistemas de hipertexto, onde considera possibilidades do hipertexto como uma tecnologia a ser explorada dentro de alguns princípios básicos da ciência da informação, como por exemplo, a análise facetada.

INGWERSEN e WORMELL (1992, p.185) também destacam o trabalho de Ranganathan sobre análise facetada, para o desenvolvimento dos sistemas de recuperação da informação. Os autores examinam essa abordagem do conhecimento e sua relevância para o acesso intelectual nas bibliotecas. Concluem que os princípios da organização do conhecimento e a metodologia facetada apresentada por Ranganathan, por ter uma abordagem estruturada, poderiam ser utilizadas como ponto de partida na modelagem de uma estrutura cognitiva e no trabalho de análise de uma determinada área do conhecimento, antes do processo de *design* de um sistema de recuperação da informação.

ELLIS (1996, p.155), comentando os trabalhos de Duncan citados acima, concorda que estes são modelos de aplicação da teoria da ciência da informação no *design* de hipertexto. Discute o uso da análise de facetas na simplificação do processo de reconhecimento de conceitos e suas relações, tanto para a máquina quanto para o usuário, uma vez que a exibição pode se restringir a um tipo de faceta somente.

POLLITT (1997, p.2) sugere que a abordagem facetada para classificação, enquanto alternativa aos sistemas enumerativos da classificação de bibliotecas, seja equivalente à abordagem relacional dos sistemas de bases de dados via computador. O autor descreve o desenvolvimento do sistema HIBROWSE (*a High resolution Interface for BROWsing and SEarching databases*) que trabalha com os princípios da classificação facetada para melhorar a interação através da pesquisa visual (*view-based*). Esse estudo foi realizado na base de dados EMBASE, publicada pela *Elsevier Science*, base que, na época, possuía 7 milhões de referências da literatura biomédica.

Aimée GLASSEL (1998, p.1), catalogadora da *Internet Scout Project*, provocativamente, escolheu como título do seu artigo "*Was Ranganathan a Yahoo!?*", comparando o trabalho do pioneiro da classificação facetada com a conhecida ferramenta de busca na Internet. Para a autora, Ranganathan reconheceu que o mundo do conhecimento estava crescendo rapidamente, com o aparecimento de novas áreas do conhecimento e novas maneiras de combinar os assuntos existentes. Além disso, nenhuma classificação que tentasse enumerar este enorme e crescente número de assuntos, sem uma capacidade de expansão para permitir novas áreas do

² Um mesmo assunto pode ser tratado sob diferentes aspectos, como por exemplo, um documento sobre o açúcar poderá ser tratado dentro de *agricultura* (planta), dentro de *química orgânica* (análise química) ou dentro de *nutrição* (alimento).

conhecimento, teria um futuro promissor. Pelo fato de a classificação facetada ocupar menos espaço na tabela do que os sistemas enumerativos, e, mesmo assim, permitir maior especificidade, fornecendo o acesso a cada faceta de uma notação, é que a autora afirma que a teoria da análise facetada se parece com a ferramenta de busca *Yahoo!*. Esta ferramenta é conhecida por ser um diretório de assuntos dentro das ferramentas da *Web*, acessível através da navegação hierárquica de categorias, subcategorias e sub-subcategorias, caracterizando um processo de navegação ramificada, que pode chegar até ao quinto nível categórico ou mais. Cada termo, dentro das categorias mais abrangentes dos sites do *Yahoo!*, tem uma palavra com seu próprio significado. Porém, quando combinado com outras palavras dentro de uma classe, um novo contexto é criado, gerando um novo significado. Daí a semelhança com a classificação facetada. Por aplicar análise facetada em uma hierarquia *online*, o *Yahoo!* tem a possibilidade de ter uma classe (*string*) de categorias e subcategorias (facetadas). Pela rotação ou permutação destes termos, essa ferramenta pode fornecer acesso a um único recurso, por meio de uma variedade de subcategorias em uma hierarquia maior.

PRISS e JACOB (1999, p.205) descrevem uma outra experiência que utiliza a estrutura facetada no *design* de sistemas de informação, em que foi feita a análise estrutural de três web sites de instituições escolares. As autoras argumentam que a aplicação da abordagem facetada na organização do conhecimento pode assegurar um processo menos ocasional e mais maleável. Para isto, sugerem o uso de um tesouro facetado como base para uma abordagem sistemática na estruturação dos *sites*, onde os elementos mínimos são as facetadas constituintes de cada conjunto de termos ou conjunto de conceitos. Nesse trabalho são descritos três tipos de construção de facetadas: **a)** por agregação, **b)** por composição de termos e **c)** por composição de termos com limitação.

ELLIS e VASCONCELOS (1999, p.3; 2000, p.97) sugerem a análise facetada como forma de transpor algumas dificuldades associadas com o processo de criação de hipertextos para usuários desconhecidos, e, possivelmente, minimizar o problema da distância entre o *designer* e o usuário em potencial. Os autores descrevem e discutem, detalhadamente, o trabalho de Duncan, considerando-o como um bom modelo para estruturação da Internet e da *WWW*, mas concluem que a análise facetada não resolve o problema de indexação para o usuário genérico. Porém, a adoção de uma abordagem classificatória *a posteriori* e não *a priori*, ou seja, quando a classificação é derivada indutivamente dos conceitos e dos termos usados em uma área de assunto, com base na garantia literária, pode aliviar alguns problemas de pesquisa na *WWW*.

KWASNIK (1992, p.63), citada na primeira página do *site Web Document Management* (<http://www.ou.edu/cas/slis/courses/LIS5990A/slis5990/Catalog/innovation.htm>), diz que

"Classificações são muito parecidas com as teorias. De fato, como as teorias, os sistemas de classificação podem prover um arcabouço que explique a abordagem da palavra a partir de uma perspectiva contextualmente determinada. Os sistemas de classificação não apenas refletem o conhecimento baseando-se na teoria e exibindo-o de uma maneira prática [...] mas as próprias classificações também funcionam como as teorias [...] e desempenham um papel semelhante na investigação".

KOEHLER (2001, p.1), autor do documento web nesse site, ressalta a importância da classificação facetada na identificação de características individuais de vários conceitos, combinando-os de uma maneira sintética, para criar uma estrutura

fluída e interativa. Citando STAR (1996), argumenta que a classificação facetada é particularmente importante para a definição da informação e sua recuperação, e levanta as seguintes questões: **(a)** Os sistemas de classificação têm sido desenvolvidos para a *WWW*?; **(b)** O que eles têm em comum?; **(c)** O que seria mais útil neste caso: os sistemas de classificação facetados ou os universais?; **(d)** O bibliotecário classificador tradicional seria útil para *Web*? Embora não responda a essas questões, o autor traz à tona importantes pontos a serem refletidos pela comunidade de profissionais e acadêmicos da área de ciência da informação, especialmente o bibliotecário, cuja atualização e adaptação a um perfil mais moderno é urgente. Pode-se constatar que os sistemas de classificação não têm sido desenvolvidos para a *WWW*, embora haja um consenso de que eles poderão ser instrumentos relevantes para a organização e pesquisa nesse ambiente. Os sistemas de classificação facetados são mais úteis para a *Web* por que sua característica analítico-sintética permite ao usuário ver um assunto sob diferentes pontos de vista, do mesmo modo que participam da não linearidade do hipertexto.

No meio acadêmico brasileiro, destacam-se, até o momento, uma tese e uma dissertação dedicadas ao estudo da classificação facetada como possível contribuição para a estruturação de hiperdocumentos. Primeiro, SANTOS (1996) em sua dissertação de mestrado *Engenharia da informação para sistemas hipertexto*, apresenta os princípios teóricos que regem a construção de sistemas de conceitos na Teoria da Classificação Facetada e a Teoria Geral da Terminologia, como opção na autoria estruturada de hiperdocumentos complexos, em especial na execução da modelagem conceitual dos sistemas de hipertextos. A autora faz um paralelo entre as duas teorias e suas aplicabilidades nos sistemas de hipertexto, apontando para a necessidade de métodos que tracem princípios consistentes para a autoria de hiperdocumentos. Ela conclui que, através de princípios que orientam a modelagem conceitual, essas teorias podem promover a construção de sistemas hipertextos que expressem cada vez mais o processo cognitivo humano.

Segundo CAMPOS (2001), em sua tese de doutorado *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para a realização da autoria*, apresenta uma preocupação com a autoria no processo de produção de hiperdocumentos. Para ela, a falta de metodologias apropriadas afasta o autor (que desenvolve o documento temático) do analista de sistema (que implementa o hipertexto), ressaltando que os modelos conceituais atualmente constituídos e aceitos são deficientes. Através de uma perspectiva interdisciplinar, entre a ciência da informação, a ciência da computação e a área da terminologia, a autora propõe estudar critérios que venham a auxiliar a elaboração de metodologias para modelos conceituais voltados para organização de unidades de conhecimento, como os hiperdocumentos. Além de discorrer sobre o objeto do hipertexto, apresenta e compara a Teoria da Classificação Facetada, a Teoria da Terminologia, a Teoria do Conceito, e métodos de construção de modelos segundo a Teoria de Orientação a Objetos e a Ontologia. Seu estudo culmina na proposição de uma série de requisitos para a elaboração de metodologias para modelos conceituais de hiperdocumentos.

Considerações finais

O objetivo deste trabalho foi mostrar, na literatura, a relevância da análise facetada em relação à organização e à recuperação da informação em sistemas hipertextuais, seja ele *online* ou *offline*. Ainda não se pode afirmar se a análise facetada resolveria o problema de classificação para o usuário desconhecido. Porém, por adotar uma abordagem classificatória *à posteriori*, adotando conceitos ou termos usados no campo específico do assunto e baseando-se na garantia literária, pode-se dizer que ela alivia alguns problemas na busca de informações em meios eletrônicos. Além disso, vários estudos têm sido feitos na implementação da Web semântica, onde o conteúdo poderá ser agregado ao símbolo. É um assunto emergente em que se busca um modelo que permita uma melhor organização e recuperação da informação. Verifica-se, na literatura, muitos esforços no sentido de relacionar procedimentos sobre a teoria da classificação facetada com a autoria de hipertexto. Entretanto, aguarda-se ainda, soluções viáveis.

Facet analysis in the conceptual modeling for hypertext: A bibliographic review

Departing from Ranganathan's Facet analysis, this article reviews the literature to find elements which may help the conceptual organization in hypertext systems. It presents the principles of this theory, theoretical studies and experiments.

Key-words: *Faceted analysis, Concept modeling, Hypertext*

Referências

- BARBOSA, Alice Príncipe. Classificações facetadas. *Ci. Inf.* Rio de Janeiro, v.1, n. 2, p. 73-81, 1972.
- BUSH, Vannevar. As we may think. *Atlantic Monthly*, v.176, n.1, p.101-108, July 1945.
- CAMPOS, Maria Luiza Almeida. *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. 2001. 190f. Tese (Doutorado em Ciência da Informação) - CNPq/IBICT-URFJ/ECO, Rio de Janeiro.
- DUNCAN, Elizabeth B. A faceted approach to hypertext? In: McALEESE, Ray. *Hypertext: theory into practice*. Nowood, NJ: ABLEX, 1989. p. 157-163.
- _____. Structuring knowledge bases for designers of learning materials. *Hypermedia*, v.1, n.1, p. 20-33, 1989.
- ELLIS, David. *Progress and problems in information retrieval*. 2nd ed. London: Library Association, 1996. 220p.
- ELLIS, David; VASCONCELOS, Ana. Ranganathan and the Net: using facet analysis to search and organise the World Wide Web. *Aslib Proceedings*, v. 51, n. 1, p.3-10. Jan. 1999.
- _____. The relevance of facet analysis for World Web subject organization and searching. In: THOMAS, Alan R.; SHEARER, James R. *Internet searching and indexing: the subject approach*. New York: Haworth, 2000. p. 97-114.
- GLASSEL, Aimee. *Was Ranganathan a Yahoo!* Disponível em: <http://scout.cs.wisc.edu/addserv/toolkit/enduser/archive/1998/euc-9803.html>. March 1998. Recuperado em 14/05/2001.
- INGWERSEN, Peter; WORMELL, Irene. Ranganathan in the perspective of advance information retrieval. *Libri*, v.42, n.3, p.184-201, July/Sept. 1992.
- KOEHLER, Wallace. Concepts. Disponível em <http://www.ou.edu/cas/slis/courses/LIS5990A/slis5990/catalog/coordination/concepts.htm> em 06/06/2001.
- KWASNIK, B. The role of classification structures in reflecting and building theory. In: AMERICAN SOCIETY FOR INFORMATION SCIENCE, 1992, Pittsburgh, PA. *Advances in Classification Research: proceedings...* Medford: Learned Information, 1992. v.3, p. 63-81.

Perspect. cienc. inf., Belo Horizonte, v. 7, n. 2, p. 189-196, jul./dez. 2002



- PIEADADE, M. Requião. *Introdução à teoria da classificação*. 2.ed. Rio de Janeiro: Interciência, 1983.
- POLLITT, A. Steven. *Interactive information retrieval based on faceted classification using views*. Disponível em <http://www.hud.ac.uk/school/cedar/dorking.htm>. Recuperado em 07/10/1997.
- PRISS, Uta; JACOB, Elin. Utilizing faceted structures for information systems design. In: AMERICAN SOCIETY FOR INFORMATION SCIENCE. ASIS Annual meeting, 62, 1999, Washington, D.C. *Proceedings...* Medford: Learned Information, 1999. p.203-212.
- RANGANATHAN, S. R. Faceted analysis. In: CHAN, L.M. et al. (Ed.) *Theory of subject analysis: a sourcebook...* Littleton, CO: Libraries Unlimited, 1985. p. 86-93.
- SANTOS, Paula Xavier. *Engenharia da informação para sistemas de hipertexto*. 1996. 83f. Dissertação (Mestrado em Ciência da Informação) - CNPq/IBICT-URFJ/ECO, Rio de Janeiro:
- SPEZIALI, P. Classifications of the sciences. In: *DICTIONARY of the history of ideas*. New York: Scribners, 1973. p. 462-467.
- VICKERY, B.C. *Classificação e indexação nas ciências*. Rio de Janeiro: BNG/Brasilart, 1980.