

Taxonomia e Classificação: o princípio de categorização

Taxonomy and Classification; the principle of categorization

por [Maria Luiza de Almeida Campos](#) e [Hagar Espanha Gomes](#)

Resumo: Taxonomia é, por definição, classificação, sistemática e está sendo conceituadas no âmbito da Ciência da Informação como ferramenta de organização intelectual. É empregada em portais institucionais e bibliotecas digitais como um novo mecanismo de consulta, ao lado de ferramentas de busca. Além destas aplicações, a taxonomia é um dos componentes em Ontologias. A organização das informações através do conceito de Taxonomia permite alocar, recuperar e comunicar informações dentro de um sistema de maneira lógica através de navegação. Este estudo apresenta o método de categorização bem como os princípios e cânones para elaboração de taxonomias como meios de representação e de acesso, já testado em serviços de informação de diferentes naturezas e propósitos.

Palavras-chave: Taxonomia; Teoria da classificação; Categorização.

Abstract: Taxonomy, according to its definition, is classification, systematics, and now is being considered a tool for intellectual organization in the realm of Information Science. It is used in institutional portals and digital libraries, as a new means for consultation together with search engines. Besides these uses, taxonomy is one of the components of ontologies. Organization of information through taxonomy allows filing, retrieving and communicating information resources in an information system as well as accessing them in a logical way, through navigation. This paper presents the method of categorization as well as principles and canons to build taxonomies as means of representation and access already tested in information services of different nature and purposes.

Key words: Taxonomy; Classification theory ; Categorization.

Introdução

Tradicionalmente, o acesso às informações em uma base de documentos se dá via busca por palavras chave, através de navegação por uma lista alfabética de tais palavras, ou ainda por busca sobre outros campos, com valores previamente associados aos documentos. No primeiro caso, o uso de operadores booleanos tem permitido recuperação mais precisa, embora, para

o usuário final, a seleção de palavras-chaves seja sempre um desafio, pois ele não tem acesso ao vocabulário utilizado na entrada do sistema.

A lista alfabética, por seu turno, por ser prática, mas não lógica, deixa de reunir idéias afins, impedindo, por vezes, que o usuário selecione a palavra mais adequada; além disso, restringe a busca a uma única palavra de cada vez. O terceiro caso, a busca por campo, é certamente útil para aquelas situações em que outros campos são de fato o alvo da seleção pretendida pelo usuário e não a busca por assunto ou conteúdo propriamente dita.

Para ampliar o resultado de uma busca, alguns sistemas permitem que documentos sejam recuperados a partir de palavras significativas presentes no texto. Apesar de sua utilidade em situações específicas, isto apenas produz mais ruído na recuperação, a menos que o usuário tenha a opção de selecionar outros elementos cobertos pela busca (*por exemplo, título, palavras-chaves, resumos.*)

Recentemente, o uso de taxonomias tem sido adotado por permitir acesso através de uma navegação em que os termos se apresentam de forma lógica, ou seja, em classes, sub-classes, sub-sub-classes, e assim por diante, em quantos níveis de especificidade sejam necessários, cada um deles agregando informação sobre os documentos existentes na base. Uma vantagem desta forma de acesso é a garantia, para o usuário, da melhor seleção do termo de busca, uma vez que as classes contêm tópicos mutuamente exclusivos. No entanto, como nas listas alfabéticas, a recuperação se restringe a um único aspecto, não permitindo o uso de operadores booleanos. ([Bruno & Richmond](#), 2003; [Gilchrist](#), 2001; [Wyllie](#), 2005)

Entretanto, em qualquer dos modelos de acesso, o sucesso da recuperação depende, basicamente, de alguns fatores:

- *É preciso haver um vocabulário-padrão;*
- *Os termos precisam ser atribuídos segundo critérios previamente estabelecidos, para assegurar consistência no tratamento dos dados;*
- *O software deve ser adequado às características de tal serviço de informação;*

Para contornar a limitação do acesso a um único aspecto, alguns sistemas desenvolvem programas que permitem ao usuário selecionar os vários aspectos disponíveis nas taxonomias e, em alguns casos, utilizar o operador booleano “E” como default.

As taxonomias têm sido bastante empregadas em portais corporativos e em bibliotecas digitais. Além dessas aplicações, o seu uso tem sido também bastante difundido no contexto da Web Semântica. Aqui, a utilização de taxonomias permite que se estabeleçam padrões de alto nível para a ordenação e a classificação de informação através do uso de mecanismos de herança.

O conceito de herança é um dos conceitos mais poderosos no desenvolvimento de software. As máquinas podem compreender corretamente relacionamentos de generalização e especialização entre as entidades atribuindo propriedades às classes gerais e então assumindo que as subclasses herdam estas propriedades. ([Campos, M. L. M; Campos, M. L. A.; Campos, L.M., 2005](#)).

Para que a Web semântica venha a funcionar de forma efetiva, computadores têm que ter acesso às coleções estruturadas de informações e a conjuntos de regras de inferência que se consolidam através de mecanismos como as ontologias. Estas são meios poderosos de inter-relacionar sistemas e neste contexto elas possuem papel de destaque, como podemos observar através dos componentes que integram uma ontologia, ou seja: Termos e Definições; Classes e subclasses - que podem estar organizadas em uma taxonomia; Relações (*também chamadas de propriedades*), que devem representar os tipos de interação entre as classes de um domínio; Axiomas que são regras para determinar a verdade das sentenças; e Instâncias que são utilizadas para representar elementos específicos, ou seja, os próprios dados.

Apesar de ainda existirem muitas restrições computacionais na aplicação de taxonomias em sistemas de informação, seu uso permite que se estabeleçam padrões de alto nível para a ordenação e classificação de informação, além de contribuir para que as organizações possam reconhecer e relacionar atividades agregadoras de valor, diminuindo esforços na produção e utilização do conhecimento.

O papel da taxonomia, neste sentido, possibilita também que os usuários possam “*aprender*” com essas estruturas de conceitos. Um usuário que não seja conhecedor da cultura popular brasileira, por exemplo, usando a hierarquia apresentada na respectiva taxonomia pode aprender os tipos de folgedos¹ existentes em determinada região do Brasil.

Em resumo, a organização das informações através do conceito de Taxonomia permite alocar, recuperar e comunicar informações dentro de um sistema de

maneira lógica através de navegação.

No âmbito da Ciência da Informação as taxonomias podem ser comparadas a estruturas classificatórias como as Tabelas de Classificação, que têm como objetivo reunir documentos de forma lógica e classificada. Atualmente, as taxonomias reúnem todo tipo de documento digital e permitem, diferentemente das estratégias de busca, um acesso imediato à informação. Ao contrário das Tabelas, que oferecem um endereço (notação) que localiza os documentos nas estantes, a taxonomia prescinde de notação.

Este estudo propõe apresentar princípios classificatórios que possam auxiliar na elaboração de taxonomias como um meio de representação e de acesso às informações.

Taxonomia: uma síntese da sua tipologia

A taxonomia ou taxionomia surgiu como Ciência das leis da classificação de formas vivas e, por extensão, ciência das leis da classificação. No ambiente dos sistemas de classificação, das ontologias, da inteligência artificial, é entendida como classificação de elementos de variada natureza.

Tradicionalmente, as taxonomias tiveram por função a classificação das espécies em botânica, e zoologia, adotando uma nomenclatura binária. Taxonomia é, por definição, classificação sistemática.

O resgate da taxonomia nos sistemas de informação considera a unidade sistemática (*taxon*) não mais família, gênero, espécie, mas conceitos. Aqui, as classes se apresentam segundo uma ordem lógica, apoiada igualmente em princípios classificatórios.

As taxonomias atualmente são estruturas classificatórias que têm por finalidade servir de instrumento para a organização e recuperação de informação em empresas e instituições. Estão sendo vistas como meios de acesso atuando como mapas conceituais dos tópicos explorados em um serviço de recuperação. O desenvolvimento de taxonomias para o negócio da empresa tem sido um dos pilares da gestão da informação e do conhecimento. (*volume de informação requer padronização*). ([Bayley, K.](#) 2007; [Gilchrist, A.](#), 2003; [Opdahl, A. L. e Sindre, G.](#), 1994)

Estas taxonomias se caracterizam por:

- *Conter uma lista estruturada de conceitos/termos de um domínio.*

- *Incluir termos organizados hierarquicamente*
- *Possibilitar a organização e recuperação de informação através de navegação.*
- *Permitir agregação de dados, diferentemente das taxonomias seminais, além de evidenciar um modelo conceitual do domínio.*
- *Ser um instrumento de organização intelectual, atuando como um mapa conceitual dos tópicos explorados em um Sistema de Recuperação de Informação.*
- *Ser um novo mecanismo de consulta em Portais institucionais, através de navegação.*

Por outro lado, as taxonomias são restritas em suas possibilidades de exploração por conterem apenas relações hierárquicas e partitivas. Daí a necessidade de serem complementadas com mecanismo de busca, para acesso a outras relações.

Entretanto, a taxonomia como um mecanismo para a organização de dados e informações, pode também conter um vocabulário agregado que permita a recuperação através de busca e não só por navegação. Esta possibilidade de somar outras funcionalidades a uma taxonomia têm criado uma certa discussão entre os pesquisadores no que diz respeito a questões como: Por que o investimento em elaboração de taxonomias tendo em vista que os tesouros são estruturas terminológicas que possuem relações paradigmáticas e sintagmáticas? Não estaríamos retrocedendo adotando um instrumento somente hierárquico?

A resposta a essas questões é muito simples. A função principal de uma taxonomia não é de padronização terminológica para a recuperação da informação, como os tesouros, mas de ordenação/organização de informação e dados. E desta forma, é devido a esta função que as relações paradigmáticas são adotadas especificamente.

Além disso, tanto os [Tesauros](#) quanto as [Tabelas de Classificação Bibliográfica](#) são instrumentos que estão fora do sistema informatizado, ou seja, são utilizados como documentos referenciais, não possuindo assim mecanismos que permitam a agregação de dados. Entretanto, deve-se considerar que todos esses instrumentos que possibilitam a formação de redes de conceitos estruturados possuem como base uma estrutura classificatória que deve se apoiar em princípios da lógica.

As taxonomias como estruturas classificatórias representam os propósitos de organização intelectual de um dado contexto. Neste sentido, são diferentes dependendo do tipo de organização e de informações que pretendem representar. Os estudos que vêm sendo desenvolvidos em nosso grupo de pesquisa ² apontam para a definição de três tipos de taxonomias:

- *Taxonomia canônica, classificação binária (dicotômica), de unidades sistemáticas (família, gênero, espécie)*
- *Taxonomia de domínio*
- *Taxonomia de processos e tarefas gerenciais*

As duas últimas caracterizam-se por serem policotômicas, ou seja, a partir de um domínio ou tarefa várias divisões são possíveis. Por sua complexidade, requerem um primeiro recorte por categorias e no interior destas várias divisões e subdivisões são possíveis em cada passo de divisão. Estas duas são objetos das pesquisas atuais.

As taxonomias não são neutras. Assim, tanto as taxonomias de domínio como as taxonomias para representação de processos e tarefas gerenciais têm seu recorte determinado pelas características da organização a que servem, bem como os propósitos do serviço.

A taxonomia com princípio classificatório policotômico, independente da área de conhecimento ser disciplinar ou multidisciplinar, possibilita uma organização que representa classes de conceitos com um princípio de divisão (*coisas e seus tipos, processos e seus tipos*), não priorizando nenhum dos aspectos, sendo os níveis subseqüentes uma simples especificação do primeiro.

Esta forma classificatória se baseia no princípio de categorias conceituais. O que se representa são os diversos aspectos (fenômenos, objetos etc) que ocorrem naquele campo de saber. Como o documento pode tratar de diversos aspectos, este modelo permite agregar e também acessar os documentos sob estes diversos aspectos.

Entretanto, requer um sistema informatizado que possua possibilidades de agregação de informação em mais de uma entrada, além de um sistema de busca avançada. Alguns sistemas internacionais vêm adotando esta modalidade, que pode ser considerada a mais hospitaleira e adequada à dinâmica do conhecimento.

Alguns princípios básicos ³ de classificação adotados nas taxonomias

Ao contrário do princípio dicotômico adotado na concepção de taxonomia original, pode-se, atualmente, construir taxonomias policotômicas, ou seja, onde um elemento é associado a tantas classes, e subclasses quantas necessárias, dentro de um domínio especializado ou uma tarefa.. Fica evidente a magnitude do problema de mapeamento multidimensional de qualquer área especializada.

Alguns princípios básicos de classificação adotados nas taxonomias:

- *Categorização, que fornece as bases para a apresentação sistemática*
- *Cânones, para o trabalho no plano das idéias (princípios para a construção das classes);*
- *Princípios, para a ordenação das classes e de seus elementos*

Categorização

A Categorização é um processo que requer pensar o domínio de forma dedutiva, ou seja, determinar as classes de maior abrangência dentro da temática escolhida. Na verdade, aplicar a categorização é analisar o domínio a partir de recortes conceituais que permitem determinar a identidade dos conceitos (categorias) que fazem parte deste domínio.

Quem primeiro introduziu a noção de Categoria nos sistemas de classificação foi Ranganathan (1967), com suas cinco Categorias Fundamentais:

*Personalidade,
Matéria,
Energia,
Espaço,
Tempo.*

"Categorias fundamentais" poderiam ser entendidas como "categorias as mais genéricas possíveis e passíveis de se manifestarem de diversas formas, capazes de hospedar todos os objetos da natureza até então conhecidos pelo Homem, e de classificá-los de acordo com sua natureza conceitual, cada um numa e somente numa categoria." Ranganathan pergunta: por que 5 e não 6? "É possível", responde ele; "qualquer pessoa pode explorar isso, pode gostar de seis. O postulado destas Categorias ampliadas poderia ser aceito se produzisse resultados satisfatórios nos

arranjos dos assuntos dos artigos ... [na notação] linear".

Vickery (1960,1980) , do Classification Research Group, desenvolve estas categorias propondo o seguinte desdobramento: Coisas, substâncias, entidades que ocorrem naturalmente; produtos; instrumentos; constructos mentais. Suas partes constituintes, órgãos. Sistemas de coisas. Atributos de coisas, qualidades, propriedades, incluindo estrutura, medidas; processo, comportamento. Objeto da ação (paciente). Relações entre coisas, interações, efeitos, reações. Operações sobre coisas; experimentos, ensaios, operações mentais. (Vickery, 1980, Wilson, T. D. 1972).

Kandelaki (1985), teórico russo da Terminologia, propõe nove categorias para a Ciência e a Técnica: categoria dos objetos, dos processos, dos estados, dos regimes, das propriedades, das grandezas, das unidades de medida, das ciências e ramos das ciências,. dos profissionais e suas ocupações. Como se pode observar, as categorias são meta-níveis.

O método de Categorização é, sobretudo, um método para organizar o pensamento, o raciocínio.(GIL, F. 2000) Serve como princípio para organizar, para reunir classes e estas é que são nomeadas, as facetas. Entretanto, além de se constituírem em princípio para organização do raciocínio, as categorias fornecem uma ordem para a disposição dos tópicos numa taxonomia.(Martínez, A. 2004).

Estes tópicos dependendo do contexto, ou seja, da finalidade de aplicação da taxonomia (*tipo de instituição, tipo de documento agregado, tipo de usuário*) podem ser escolhidos a partir de uma ou de várias categorias. Por exemplo, em uma empresa de óleo Diesel que produz vários Padrões para os processos de produção e distribuição, ela pode utilizar uma única categoria, ou seja, a dos processos. Estes é que norteiam as atividades agregadoras de valor.

No interior de cada categoria, as classes de conceitos são dispostas através de uma organização que deve ser apoiada também por princípios diretivos. Estes princípios foram desenvolvidos e explicitados por Ranganthan em pelo menos três de seus trabalhos, ou seja, *Philosophy of library classification*. (Ranganthan, 1951); *Colon Classification*. (Ranganthan, 1963); *Prolegomena to Library Classification* (Ranganthan, 1967), e constituíram o que ele denominou de *Canônes* para o trabalho no plano das idéias.

O plano das idéias é o espaço onde os conceitos de um dado domínio são organizados formando um sistema de conceitos, cuja síntese apresentamos a

seguir.

Cânones para o trabalho no plano das idéias

No interior de cada categoria os conceitos devem ser organizados em classes. As classes de conceitos são de dois tipos, a saber: cadeias e renques. Cadeias são séries verticais de conceitos que podem ser genéricas e partitivas. As cadeias genéricas formam uma seqüência de conceitos que respondem à seguinte pergunta: é tipo de? As cadeias partitivas por sua vez respondem à pergunta: é parte de? ou é o todo de?. A cadeia pode assim, ser crescente ou decrescente.

Renques são séries horizontais de conceitos e podem também ser genéricos ou partitivos. Os renques agregam conceitos de mesmo nível organizados a partir de um conceito que respondem à seguinte pergunta: São elementos partitivos ou elementos específicos da classe maior?

Para a organização das cadeias e renques alguns cânones, princípios gerais devem ser seguidos. A seguir apresentaremos estes princípios que conduzem o classificacionista na organização de uma taxonomia, minimizando a subjetividade inerente a qualquer processo classificatório.

Separamos assim em quatro momentos, a saber:

Cânones para Cadeias;

Cânones para Renques;

Cânones para Características de Divisão;

Princípios para ordenação das classes e de seus elementos.

Cânones para Cadeia:

Como as cadeias são séries verticais de conceitos é necessário estabelecer a ordem em que cada conceito deverá está em relação aos outros conceitos. Desta forma, dois princípios são fundamentais:

- *Cânone da Extensão Decrescente onde os conceitos de uma série descendente devem crescer em intensão .*

- *Cânone da Modulação onde a ordem dos conceitos em uma classe de conceitos devem ter uma seqüência que respeite cada elo da cadeia. Não se deve quebrar nenhum elo da cadeia.*

Cânones para Renque

Como os renques são formados a partir da reunião de elementos em uma classe, é necessário estabelecer a forma como estes elementos devem ser agregados para formar classes de conceitos.

- *Cânone da Exclusividade – o conjunto de elementos de uma classe não deve constituir uma outra classe, ou seja, os elementos de um renque devem ser mutuamente exclusivos.*
- *Cânone da seqüência Útil – a seqüência dos elementos em uma classe, deve ser útil ao propósito daqueles a quem ela se destina.*
- *Cânone da Seqüência Consistente – sempre que existirem classes semelhantes à seqüência de seus elementos deve ser paralela em todas aquelas classes, aonde a insistência em tal paralelismo não venha a contrariar outros requisitos mais importantes.*

Cânones para Características de Divisão

Alguns princípios são importantes para a ordenação das cadeias. Estes princípios se consolidam no que se denomina de características de divisão, ou seja, os princípios pelos quais as classes podem ser divididas. Cada característica de um esquema associado de características deve satisfazer os quatro cânones seguintes:

- *Cânone de Diferenciação – uma característica usada como base para a classificação de um universo deve fazer a diferenciação de algumas de suas entidades, ou seja, deve dar origem a pelo menos duas classes;*

- *Cânone de Relevância – uma característica usada como base para a classificação de um universo deve ser relevante para o propósito da classificação;*
- *Cânone da Verificabilidade – uma característica usada como base de classificação em um universo deve ser definitiva e verificável. Um universo de entidades pode ter muitas características relevantes. Mas nem todas elas são passíveis de verificação. Este cânone enfatiza que apenas as características verificáveis devem ser escolhidas para divisão do universo de entidades;*
- *Cânone da Permanência – uma característica usada como base para a classificação de um universo continua a ser mantida enquanto não houver mudança no propósito da classificação*

A sucessão de características no esquema associado de características deve satisfazer os três cânones seguintes:

- *Cânone de Concomitância – em caso algum duas características podem ser concomitantes num esquema associado de características, ou seja, elas não devem dar origem ao mesmo renque de conceitos;*
- *Cânone da Sucessão Relevante – a sucessão de características no esquema associado de características deve ser relevante para o propósito da classificação;*
- *Cânone da Sucessão Consistente – a sucessão das características no esquema associado de características deve ser seguida de forma consistente, enquanto não houver mudança no propósito da classificação.*

Princípios para ordenação das classes e de seus elementos

A ordenação dos vários elementos nas classes e sub-classes requer alguma ordem em sua seqüência. Vários são os princípios passíveis de adoção:

- *Princípio do Posterior-no-Tempo – este princípio ocorre principalmente em fenômenos, processos, atividades. Podem ser citadas como exemplo, as escolas de pensamento, as religiões, os movimentos culturais e artísticos, os planos econômicos de desenvolvimento.*
- *Princípio do Posterior-na-Evolução - este princípio, como diz o nome, serve para orientar a organização de conceitos ligados a processos evolutivos. É bastante sutil a diferença entre o princípio anterior e este, pois ambos se referem à contigüidade temporal, mesmo porque o conceito de "evolução" embute a idéia de progressão, o que corresponde à transformação de um processo ou de uma idéia no tempo. Taxonomias consideram, por vezes, que existe uma relação partitiva com o todo que está sendo classificado.*
- *Princípios da Contigüidade Espacial - A seqüência defendida neste Princípio se refere, em geral, a um objeto no todo, quando se necessita ordenar os elementos que o compõem, os quais se apresentam numa disposição espacial, que pode ser numa linha direcional, radial ou circular. Nestes casos, sempre que possível, os assuntos devem ser "arranjados numa seqüência correspondente", exceto quando houver um forte impedimento para arranjos deste tipo. Em geral podem se apresentar em pares antitéticos: a) Entidades numa Linha Vertical - Aqui se incluem os Princípios "de cima para baixo" e "de baixo para cima"; b) Entidades numa Linha Horizontal - o arranjo pode atender os Princípios de Esquerda para Direita, ou de Direita para Esquerda; c) Entidades numa Linha Circular - Dois princípios são propostos: na Direção horária e na Direção anti-horária; d) Entidades numa Linha Radial - Dois Princípios são invocados: da Periferia para o Centro e do Centro para a Periferia;*

e) Contigüidade Geográfica "...continentes, países, províncias, distritos ou outras divisões administrativas estão numa superfície e não numa linha. Portanto, a contigüidade não pode ser determinada de uma única maneira.

- *Princípios para Medida Quantitativa - aqui, também, um par de Princípios é proposto: o de Quantidade Crescente e o de Quantidade Decrescente. A área da Geometria é o exemplo para quantidade crescente, no que se refere ao número de dimensões: linha, plano, três dimensões, cinco dimensões, até n-dimensões.*

- *Princípio da Complexidade Crescente - Se os elementos de uma série horizontal "mostrarem diferentes graus de complexidade, devem ser arranjados de modo correspondente à seqüência de complexidade crescente, exceto quando qualquer outro forte impedimento o exigir." Em tecnologia industrial, poderíamos pensar em: matéria-prima, produto semi-manufaturado, produto manufaturado."*

- *Princípio da Seqüência Canônica - este princípio consagra a tradição. Se existe uma ordem tradicional para citar um conjunto de assuntos, então ela deve ser adotada, se for conveniente.*

- *Princípio da Garantia Literária - este princípio estabelece uma ordem para os assuntos de acordo com a quantidade decrescente de documentos publicados ou a serem publicados. A adoção deste princípio requer cuidado e bom senso*

- *Princípio da Ordem Alfabética.- não por acaso, a ordem alfabética fica por último. Ela deve ser adotada quando nenhuma das outras seqüências for mais útil.*

Um fato a destacar é que a existência de vários princípios não significa que todos tenham que ser adotados. São fatores decisivos para adoção dos princípios: os propósitos da taxonomia, os documentos a serem agregados, a comunidade a ser atendida, o

software disponível.

Elaboração da taxonomia

A elaboração de taxonomias é uma atividade que requer do classificacionista um planejamento que inicia com o seu dimensionamento até as etapas de construção propriamente. Os estudos e as experiências desenvolvidas por nosso grupo de pesquisa têm mostrado que alguns princípios são fundamentais para garantir a consistência de sua estrutura classificatória e atender a uma boa recuperabilidade.

No que tange ao dimensionamento das taxonomias, de acordo com Holgate (2004), algumas questões são fundamentais e devem ser observadas pelo profissional que irá se envolver com tal tarefa, a saber:

- *Qual o problema que a taxonomia está tendo que responder?*
- *Qual o tipo e o alcance da informação corporativa?*
- *Qual o volume do conteúdo de informação agregada?*
- *Qual a disponibilidade dos especialistas da área para estarem desenvolvendo a taxonomia?*
- *Qual a arquitetura de informação e informática da organização para suportar uma taxonomia?*

Consideramos também, de forma geral, que alguns princípios gerais devem ser adotados na elaboração de taxonomias relativos ao termo, que é o elemento de comunicação e acesso as informações, ou seja:

Comunicabilidade: o termo empregado deve representar a linguagem utilizada pelo usuário;

Utilidade: o nível de especificidade dos termos deve expressar um agrupamento de documentos e não um único documento, ou seja, o termo só é útil quando representativo para um conjunto de documentos;

Estimulação: utilizar termos que induzem o usuário a continuar a navegação pelo sistema;

Compatibilidade: os termos empregados representem o

campo que se está ordenando, fazendo parte das atividades e funções da organização. (Terra, 2005)

A partir das considerações acima, o classificacionista inicia o que denominamos de desenvolvimento da taxonomia, que se configura das etapas apresentadas a seguir.

Captura do conhecimento

A captura do conhecimento é uma etapa da elaboração da taxonomia em que o objetivo é o levantamento dos assuntos que deverão ser acomodados numa estrutura classificatória, servindo como ponto de acesso à informação.

O levantamento dos assuntos pode-se dar através de entrevistas com especialistas da área, de documentos existentes na instituição, de outros instrumentos classificatórios ou terminológicos.

No caso das entrevistas com especialistas é importante definir a metodologia utilizada para a explicitação do conhecimento tácito do especialista para um conhecimento explícito. Neste sentido, diversas técnicas têm sido empregadas, como por exemplo, a observação participativa, onde o observador (*classificacionista*) e o observado (especialista), a partir de um contexto apresentado (*no caso, exemplos de propostas classificatórias existentes*), expressam suas impressões através de discussões consolidadas posteriormente pelo observador, que transforma estas observações em novas propostas, que serão novamente discutidas até se chegar a um modelo adequado que atenda os propósitos da organização.

Análise dos documentos e informações que serão agregados à taxonomia

As Taxonomias têm por finalidade servir de mapa navegacional para uma dada tipologia de documentos/informação, e necessitam de uma estrutura classificatória que expresse a natureza dos documentos agregados. Aqui cabe ressaltar, a diferença apresentada por [Ranganthan](#) em seu Prolegomena (1967) quanto aos princípios da classificação de documentos, entre Universo de Conhecimento e Universo de documentos.

O que se pretende representar em taxonomias são os conhecimentos existentes e explicitados por aquela comunidade de especialistas, ou seja, analogamente o Universo de Documentos, e não um pseudoconhecimento que não expressa o contexto e as visões daquela comunidade. Desta forma, as taxonomias agregam documentos, considerando documentos em sua acepção mais ampla,

e estas informações/documentos devem encontrar hospitalidade no sistema. Isto evidencia a diferença entre a construção de taxonomias tradicionais para representar o conhecimento de uma área e a construção de taxonomias para organizar e recuperar documentos.

Elaboração da Estrutura Classificatória da taxonomia

Primeiramente nesta etapa é importante definir a tipologia de taxonomia que será representada, de domínio ou voltada para representação de processos e tarefas gerenciais, como apresentada nos itens anteriores. A partir de então consideramos importante estabelecer os princípios utilizados na elaboração de estruturas classificatórias apresentados nos itens anteriores.

Além dessas etapas, a taxonomia é também fruto do modelo de representação adotado; assim, definir a forma representacional e o software utilizado para esta representação é fundamental. Desta forma, é importante investigar formas gráficas de representação, porque em alguns casos estas formas "*inibem*" as possibilidades representacionais.

Cabe ressaltar que toda taxonomia é fruto de um processo representacional e classificatório e como todo processo desta natureza é um produto de uma construção que representa o estado e visão do conhecimento de seus elaboradores. Entretanto, consideramos fundamental explicitar os princípios pelos quais tais representações foram construídas, pois isto facilita sua alimentação constante.

Validação

A validação pelos especialistas tem, pelo menos, duas funções: uma delas é confirmar o trabalho realizado, a outra, transferir o conhecimento do processo de realização. Desta forma, consideramos imprescindível a validação em todas as etapas de definição da taxonomia, que vão dos estudos dos documentos/informações agregadas até a definição das formas gráficas de representação. Isto permitirá que se possa atingir critérios de comunicabilidade, estimulação e compatibilidade.

Resultados e Discussão

A proposta do estudo teve por objetivo traçar um caminho teórico e metodológico visando auxiliar na elaboração de taxonomias consistentes. Os critérios aqui apresentados se pautaram em estudos teóricos desenvolvidos pelos autores, além de um grande envolvimento em ensino e consultorias em diversas instituições e em variados contextos de conhecimento.

Estas experiências incluem, dentre outros, empresas de tecnologia, instituições que atuam em áreas científicas e tecnológicas, órgãos de cultura e organizações ligadas ao entretenimento. Os diferentes recursos objetos das taxonomias incluem além de textos (*de natureza bibliográfica e arquivística*), imagens em movimento, pranchas para cenários, plantas baixas, esquemas para produção de objetos, moda, mapas e cartas de diferentes tipos, fotografias originais e manipuladas, para citar os mais características das organizações.

O sucesso dos empreendimentos se deve, em especial, a dois aspectos a serem considerados: primeiro, a adoção de princípios teóricos e metodológicos de classificação consistentes, para que as taxonomias possam ser elaboradas de forma a serem expandidas e, segundo, estar atento para soluções relativas à arquitetura de informação existente para acomodar as taxonomias.

Sobre o primeiro aspecto, consideramos importante a adoção de princípios classificatórios que estejam explicitados para a elaboração de classes de conceitos que possam evidenciar o conhecimento existente em dado domínio de saber ou atividade e possibilitem sempre a inclusão de novos conceitos. Neste sentido, considera-se que etapas de elaboração, como apresentado, devam envolver as seguintes ações:

1. Captura do conhecimento - que pode se dá através de entrevistas com especialistas da área, de documentos existentes na instituição, de outros instrumentos classificatórios ou terminológicos;

2. Análise dos documentos/informações que serão agregados à taxonomia – adequação das informações existentes nos acervos a terminologia apresentada na taxonomia;

3. Elaboração da estrutura classificatória da taxonomia – que deve se pautar nos princípios teóricos de categorização, de formação e de ordenação das classes de conceitos;

4. Validação - se apóia como base para a certificação da proposta classificatória, visando atender as necessidades da comunidade para quem se destina. Esta

ação permitirá que possamos atingir critérios de comunicabilidade, estimulação e compatibilidade.

Sobre o segundo aspecto, consideramos que os projetos de soluções de arquiteturas de informação precisam atentar para múltiplas formas de disponibilizar uma taxonomia e de integrá-la inclusive com funcionalidades de busca.

Assim, múltiplos ambientes de busca podem ser oferecidos ao usuário. Sobre este último aspecto, nossa experiência revelou ser necessário trabalho conjunto entre profissionais de informação e computação para a adoção de soluções que possam permitir a elaboração de taxonomias dinâmicas e expressivas.

Considerações finais

A disponibilidade de uma taxonomia fornece ao usuário um guia ao assunto a ser investigado. Esta disponibilidade elimina a necessidade do usuário possuir um entendimento completo do assunto antes de submeter uma pergunta. Ela serve como um guia ao processo de pesquisa, até mesmo de forma educativa, progressivamente revelando áreas de interesse ao usuário. A navegação por categorias, técnica utilizado de acesso à taxonomia, encoraja o pensamento associativo e pode guiar o usuário através de processos de descoberta de informação.

Apesar de considerarmos que as técnicas de navegação por taxonomias não são superiores às outras formas de busca, como por exemplo, as buscas por palavras-chave sendo uma complementar à outra, com o advento dos sítios e portais, a utilidade das taxonomias tem-se tornado mais importante e difundida, pois a navegação é uma interface bem projetada sobre a informação desses sítios ou portais. Então, faz-se necessário que o classificacionista se paute em princípios classificatórios explícitos e consistentes.

Notas

[1]Folguedo - Atividade ritual que se expressa como manifestação coletiva composta de elementos dramático, musical e coreográfico. Em geral,

organiza-se ao longo de reuniões periódicas para os ensaios dos integrantes, que são mais ou menos constantes. A divisão de trabalho e a hierarquia interna dos grupos exigem certa permanência, contribuindo para a manutenção de um padrão básico. O folguedo integra dimensões festivas, musicais, estéticas e dramáticas. O componente dramático, nem sempre explicitamente encenado, é sempre identificável nos trajes especiais, na organização de danças, cantorias, embaixadas, cortejos e na existência de personagens. As apresentações ocorrem em ruas e praças públicas, ou em terreiros e estádios, especialmente nos dias de festas do calendário litúrgico ou profano. (*Tesouro de folclore e cultura Popular Brasileira*).

[2] Grupo de Pesquisa registrado no CNPq Ontologias e Taxonomias: princípios

[3] Estes princípios foram sintetizados a partir do trabalho publicado no site: < <http://www.conexaorio.com/bitl/revisitando/revisitando.htm> > , intitulado Revisitando Ranganthan: a classificação na rede, de autoria de Hagar Espanha Gomes, Dilza Motta e Maria Luiza de Almeida Campos.

Referências Bibliográficas

ADAMS, Katherine. The semantic web: differentiating between taxonomies and ontologies. Online, p.20-24, 2002.

BAILEY, Kenneth D. Typologies and taxonomes: an introduction to classification techniques. Disponível em: <<http://www.sagepub.com/booksProdDesc.nav?contribId=502366&prodId=Book4445>> Acesso em: 04 de abril de 2007.

BRUNO, D & RICHMOND, H. The true about taxonomies. Information Management Journal, v. 37, n.2, p.44-46, mar-apr 2003.

CAMPOS, Maria Luiza Machado; CAMPOS, Maria Luiza de Almeida; CAMPOS, Linair Maria. Web Semântica e a Gestão de Conteúdos Informacionais. In: MARCONDES, Carlos H. e (org). Bibliotecas Digitais: Saberes e Práticas. Salvador/Brasília: EDUFBA/IBICT, p. 55-78, 2005.

GIL, Fernando. Categorizar. In: ENCICLOPÈDIA Einaudi. Conhecimento. [s.l.]: Imprensa Nacional – Casa da Moeda, p.52-89, v. 41, 2000.

GILCHRIST, Alan. Taxonomies for knowledge management. Online information 2001. Proceedings of the International Online Information Meeting. London, 4-6 December 2001. Oxford: Learned Information Europe Ltd, 2001, p. 37-9.

GILCHRIST, Alan. Thesauri, taxonomies and ontologies – an etymological note. Journal of Documentation, v. 59, n. 1., p. 7-18, 2003

HOLGATE, I. Creating and using taxonomies to enhance enterprise search. Information Today, New Jersey, v.7, n.21, jul/ago. 2004.

KANDELAKI, T. I. Les sens des termes et les systèmes de sens des terminologies scientifiques et techniques. In: Rondeau, G. & Felber, F. org. Textes choisis de terminologie. Québec, GISTERM, 1981, p. 135-184

MARTÍNEZ, Ana; RISTUCCIA, Cristina; PISARELLO, Rosa; STUBBS, Edgardo; CAMINOTTI, Laura; BALPARDA, José; VALDEZ, Julia; MANGIATERRA, Norma. Las categorías o facetas fundamentales: una metodología para el diseño de taxonomias corporativas de sitios web argentinos. Ciência da Informação, v. 33, n. 2, maio/ago. 2004, p. 106-111

OPDAHL, Andreas L.; SINDRE, Guttorm. A taxonomy for real-world modelling concepts. Information systems, v. 19, n. 3, 1994, p. 229-241.

RANGANATHAN, S.R. Colon Classification. Bombay: Asia Publishing House, 1963. 126p

RANGANATHAN, S.R. Philosophy of library classification. New Delhi: Ejnar

Munksgaard, 1951

RANGANATHAN, S. R. Prolegomena to Library classification. Bombay: Asia Publishing House, 1967. 640p

TERRA, José C. C. et al. Taxonomia : elemento fundamental para a gestão do conhecimento. Rio de Janeiro: Campus, 1998.

VICKERY, B. C. Classificação e indexação nas ciências. Rio de Janeiro: BNG/BRASILART, 1980

VICKERY, B. C. Faceted Classification. London, Aslib, 1960.

WILSON, T.D.. The Work of the British Classification Research Group. In: WELLISCH, H. ed. Subject retrieval in the seventies. Westport, Greeword Publishing Co, 1972. p. 62-71.

WYLLIE, Jan. Taxonomies: Frameworks for Corporate Knowledge. 2nd. London: Ark Group in association with David Skyrme Associates, 2005.

Sobre os autores / About the Authors:

Maria Luiza de Almeida Campos

mlcampos@nitnet.com.br

Doutora em Ciência da Informação pelo IBICT/UFRJ, Professora Adjunta do Departamento de Ciência da Informação da Universidade Federal Fluminense/Professora do Programa de Pós-Graduação UFF/IBICT.

Hagar Espanha Gomes

hagar.espanha@terra.com.br

Livre-docente pela Universidade Federal Fluminense, Consultora

independente.