

A Teoria da Classificação Facetada na Construção de Taxonomias Facetadas

Benildes Coura M. S. Maculan¹, Elisângela Cristina Aganette²

¹Grupo de Pesquisa Representação de Conhecimento e Recuperação da Informação (RECRI), Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC), Escola de Ciência da Informação (ECI), Universidade Federal de Minas Gerais (UFMG), Professora Adjunto, Doutora em Ciência da Informação, ORCID 0000-0003-4303-9071.

²RECRI, PPGGOC, ECI, UFMG, Professora Adjunto, Doutora em Ciência da Informação, ORCID 0000-0003-4357-8016.

benildes@gmail.com, elisangelaaganette@gmail.com

Abstract. *A taxonomy is a type of classificatory structure with terms organized in hierarchical form. It emerged as a dichotomous model in the Biology area (Taxonomy of Lineu), and later in the area of education, with the Bloom Taxonomy, which presented an educational goal classification, built in 1956. In the area of Information Science, Vickery (1960) was the first to define a taxonomy as an instrument to organize the set of entities of a domain. In the construction of ontologies and thesauri, taxonomies (hierarchical structures) are a backbone of these instruments, since they represent a primary data structure of a domain. It is assumed that the bases of Classification Theories, especially the Faceted Classification Theory developed by Ranganathan (1967) can contribute to the determination of classes and facets of subjects and entities in the construction of faceted taxonomies. In this perspective, this article presents subsidies for actions in construction of a facet taxonomy, mainly for a representation of classes and facets of the subject based on the Faceted Classification Theory principles.*

Keywords: *Faceted Classification Theory. Faceted taxonomy. Faceted taxonomy construction.*

Resumo. *Uma taxonomia é um tipo de estrutura classificatória, com termos organizados na forma hierárquica. Ela surgiu como sistema dicotômico, na área da Biologia (Taxonomia de Lineu), e, mais tarde, em 1956, na área da educação, com a Taxonomia de Bloom (1956), que apresentou uma classificação para objetivos educacionais. Na área da Ciência da Informação, Vickery (1960) foi o primeiro a definir a taxonomia como instrumento para organizar o conjunto de entidades de um domínio. Na construção de ontologias e tesouros, as taxonomias (estruturas hierárquicas) são consideradas a “espinha dorsal” desses instrumentos, pois representam a estrutura primária dos dados de um domínio. Considera-se que os fundamentos das Teorias da Classificação, sobretudo da Teoria da Classificação Facetada, desenvolvida por Ranganathan (1967) pode contribuir na determinação de classes e facetas de assuntos e entidades na construção de taxonomias facetadas. Nessa perspectiva, este artigo apresenta subsídios para os procedimentos envolvidos na construção de uma taxonomia facetada, sobretudo para a representação de classes e facetas de assunto, com base nos princípios da Teoria da Classificação Facetada.*

Palavras-chaves: *Teoria da classificação facetada. Taxonomia facetada. Construção de taxonomia facetada.*

1. Introdução

Uma das principais competências da Ciência da Informação (CI) se refere à organização do conhecimento e da informação. Com a realidade da *web*, a produção, armazenamento e distribuição de informação precisa considerar a representação da informação de recursos digitais. Nesse contexto, deve haver tratamento do recurso informacional de maneira que a heterogeneidade de dados possa ser estruturada. Para tanto, os assuntos, tópicos, termos de índice e palavras-chave, por exemplo, podem ser estruturados a partir de diferentes maneiras, seja com a arquitetura da informação, a organização e gerenciamento de metadados ou classificações.

Fazendo uma distinção entre sistema de categorização e sistema de classificação, Silva (2010) estabelece que as classificações têm por características um arranjo sistemático de entidades, em uma estrutura hierárquica de classes fixas, formadas para atender a propósitos e critérios pré-determinados. A autora acresce que a estrutura possui limites rígidos (exclusão mútua), pois uma subclasse pertence ou não a uma superclasse, e, caso positivo, todos os itens dessa classe possuem as mesmas características essenciais, sendo todos igualmente representativos da classe (JACOB, 2004). Por sua vez, a categorização é flexível e dinâmica, sendo que a associação entre os itens da subclasse não ocorre de forma rígida, podendo ser dependente de um contexto em particular.

Uma taxonomia é uma estrutura classificatória, com termos organizados hierarquicamente. De acordo com Novo (2007), o conceito de taxonomia é resultado de um longo processo histórico construído por meio de estudos que culminaram em uma evolução não linear, e nem tampouco ocorreu em um mesmo momento histórico. Um sistema dicotômico originou a taxonomia, na Biologia, com Linnaeus, no século XVIII, que criou um sistema hierárquico dos seres vivos. Mais de dois séculos após, o termo foi utilizado com objetivos pedagógicos, a Taxonomia de Bloom, que apresentou a classificação dos objetivos educacionais em seis níveis: avaliação, síntese, análise, aplicação, compreensão e conhecimento (AGANETTE; ALVARENGA; ROCHA, 2010). No âmbito da CI, Vickery (1960) foi o primeiro a definir a taxonomia como instrumento para organizar o conjunto de entidades de um domínio em uma simples hierarquia.

Desde então, o conceito de taxonomia sofreu transformações, e, como um sistema de organização do conhecimento (SOC), as taxonomias desempenham um papel importante no esclarecimento de tópicos complexos e para facilitar a compreensão de um domínio (HEDDEN, 2010). Ela é usada para estruturação de informações e considerada uma importante ferramenta para o entendimento de como uma área de conhecimento é organizada e, principalmente, como essa área se relaciona e interage com outras (AGANETTE; ALVARENGA; ROCHA, 2010). São características de uma taxonomia: a) lista estruturada de termos; b) termos organizados hierarquicamente; c) permite a navegação através de seus termos estruturados; d) aceita agregação de dados; e) explicita o modelo conceitual do domínio que representa; f) mecanismo de busca e recuperação de informações (CAMPOS; GOMES, 2008).

As taxonomias podem ser, segundo Conway et al. (2002): descritivas, para navegação e para gerenciamento de dados: i) taxonomia descritiva: consiste em vocabulários controlados construídos a partir de tesouros, e adiciona diversos tipos de palavras, ortografias, formas e dialetos variantes, para que o usuário tenha maior liberdade na hora de buscar um assunto; ii) taxonomia para navegação: envolve agrupamentos apropriados das informações, e pretende descobrir informações por meio do comportamento do usuário mediante a utilização de navegadores (*browsing*), baseando-se nos modelos mentais dos trabalhadores; iii) taxonomia para gerenciamento de dados: contém um pequeno conjunto de termos controlados, com significância particular e específica. As taxonomias construídas para agregar dados, permitem organizar itens de informação em banco de dados. Os metadados são considerados pontos de referência que descrevem e estruturam dados de informação, de maneira estável e uniforme, sob distintas formas (suportes e tipologias) (TAYLOR, 2003). Nessa perspectiva, os metadados são resumos da informação disponibilizada e facilitam o acesso à informação, sua extração e compreensão, e, na

atualidade, são empregados em serviços de informação *online* para armazenamento, autenticação, direitos de autor e busca de informação (CRUZ; DORNELES; GALANTE, 2010). Dessa maneira, elas funcionam como mapas conceituais para os recursos de informação de um domínio, utilizadas como orientação permitida pela estrutura de termos que dão acesso aos tópicos, para exploração (navegação) em um serviço de recuperação (CAMPOS; GOMES, 2008).

Nesses serviços, por vezes encontramos uma taxonomia facetada, na qual um conteúdo não está restrito a uma única dimensão, oferecendo diferentes opções de busca ao usuário, na qual cada termo está em uma ou mais relação tipo pai/filho (geral/específico) em relação a outro termo, conectados em hierarquia ou poli-hierarquia. Com essa característica, ela é uma estrutura composta por um conjunto de taxonomias, e cada taxonomia descreve o conteúdo do banco de dados sob um aspecto (dimensão/perspectiva) diferente, podendo ser usada para a navegação. Assim, em vez de usar os rótulos de uma taxonomia tradicional (constituída por uma única faceta ou filtro) como ponto de partida para encontrar o conteúdo desejado, os rótulos da taxonomia facetada são usados, em especial, para restringir ou filtrar os resultados de uma busca.

O uso de diferentes atributos para limitar os resultados de busca é uma maneira já bastante comum em distintas implementações, tais como em sistemas de pesquisa corporativa (documentos corporativos internos), de sites de comércio eletrônico (seleção de produtos) ou de bancos de dados de bibliotecas (seleção de artigos). Nesses últimos, é essencial haver uma classe para representar os assuntos, visando a recuperação do conteúdo temático de uma coleção. Essa atividade específica, de representação de classes de assuntos, não é corriqueira.

Considera-se que os fundamentos advindos das Teorias da Classificação, sobretudo da Teoria da Classificação Facetada (TCF), desenvolvida por Ranganathan (1967), que dá subsídios para a classificação sistemática e ordenação de termos e assuntos, tem uma importante contribuição na determinação de classes e subclasses de assunto na construção de taxonomias, tornando-as facetadas. Acredita-se que na construção de tesouros, em geral, e de ontologias que determinam relações do tipo <Is-a>, as taxonomias (estruturas hierárquicas) são consideradas a sua “espinha dorsal”, pois representam a estrutura primária dos dados de um domínio. Sendo assim, as taxonomias facetadas poderiam oferecer uma estrutura ainda mais formalizada para reuso na construção de tesouros e ontologias. Discorrendo sobre o reuso de ontologias biomédicas, Campos, Campos e Campos (2010, n.p.) afirmam que a sua precisão pode ser aumentada por meio “da adoção de diretrizes para a definição de recorte de domínio” e que a TCF pode auxiliar nesse sentido, pois, a “perspectiva classificatória estabelece uma forma sistemática de alto nível para se definir e classificar o conhecimento de um domínio” (idem, idem).

Nessa perspectiva, este artigo problematiza a construção de uma taxonomia facetada para a representação de classes e subclasses de assuntos/conceitos/entidades, com base nos princípios da TCF, como subsídios restritivos para a estruturação de entidades em um modelo da realidade. Para isso, depois da introdução, este artigo está organizado como segue: a seção 2 apresenta os princípios da TCF; a seção 3 discorre sobre o uso da TCF na construção de taxonomias facetadas e algumas de suas aplicações; e, finalmente, a seção 4 apresenta as considerações finais.

2. Princípios da Teoria da Classificação Facetada (TCF)

Considera-se que Ranganathan (1967) contribuiu com o maior avanço na área de classificação no séc. XX, desenvolvendo os princípios analítico-sintéticos da TCF, possibilitando uma visão multidimensional e ilimitada do conhecimento. Nesse sentido, a classificação facetada (ou uma taxonomia facetada) deixa a subdivisão unidimensional para apresentar o conhecimento em ramificações, em características que obedecem a postulados pré-determinados.

O preceito multidimensional contesta o princípio dicotômico-binário de Porfírio, denominado Árvore de Porfírio, que reduziu as relações a apenas pai/filho, característicos das taxonomias tradicionais, sistematizando a noção de ideias gerais até alcançar as mais específicas.

Examinando o modelo da dicotomia, Ranganathan (1967) concluiu que o conhecimento é prolífico e multidimensional e sugeriu um padrão para mapear o conhecimento com uma policotomia ilimitada (ilimitadas subdivisões da ciência). Ranganathan (1967) usou a ideia da Árvore Baniana, um tipo de figueira indiana, que se espalha por uma grande área, enviando galhos para o solo. Os galhos criam raízes e formam vários troncos. Essas raízes são aéreas e crescem do tronco principal, tornando-se novos troncos adicionais. Nesse tipo de representação, não há apenas relacionamentos hierárquicos, pois, essa árvore sugere a ideia de que os assuntos (classes mais abstratas) podem ser relacionados uns aos outros de diferentes, complexas e imprevisíveis maneiras, gerando novas classes, em um sentido multidimensional. Desse modo, esse modelo configura uma representação simbólica para o domínio, como um organismo vivo, no qual o conhecimento está em constante transformação, sofrendo ramificações e desenvolvimento ao longo do tempo. No contexto dos estudos sobre a TCF, há distintos conceitos que devem ser entendidos, tais como:

QUADRO 1 – Conceitos relacionados à TCF

Conceito	Definição
Categorias fundamentais	São as grandes classes, as classes com conceitos mais gerais ou mais abstratos, que podem ser utilizadas para reunir outros conceitos. Ex.: PMEST
Faceta/Classe	Manifestações das categorias fundamentais, em aspectos exaustivos, reunindo conceitos que têm determinada característica em comum; a soma dos isolados resultantes da divisão de um assunto por uma característica; também entendida como um ponto de vista ou atributo usado para agrupar conceitos em uma área de assunto (domínio). Ex.: Coisa, Processo, Instrumento, Período, entre outros.
Subfacetas/ Subclasses (ou arrays)	Grupos de termos coordenados, obtidos com a divisão de um assunto por meio de um mesmo princípio, mutuamente exclusivos. Ex.: Instrumento (bateria, baixo, guitarra, entre outros).
Isolado	Cada componente ou indivíduo, obtido a partir da divisão de uma faceta, antes de serem reunidos em facetas e subfacetas. Ex.: guitarra.
Foco	É um isolado já acomodado na estrutura facetada, porém, sem preocupação sobre a relação que tem com outros componentes (termo/indivíduo) da estrutura.
Divisão	O processo pelo qual uma faceta se decompõe em diferentes focos. Ex.: pelo estilo (barroco, pós-moderno); pela forma (redondo, triangular).
Renques	Divisão feita a partir de apenas uma característica, em uma série horizontal de conceitos. Ex.: genéricos (itens específicos da classe maior); partitivos (itens que representam partes específicas da classe maior).
Cadeias	Divisões sucessivas de um mesmo assunto, em uma série vertical de conceitos. Ex.: genéricas (tipo de); partitivas (parte de); Árvore-Árvore frutífera-Macieira

Fonte: adaptado de Ranganathan (1967).

Para Vickery (1960), a essência da TCF é a ordenação de conceitos em um dado domínio, em facetas homogêneas, mutuamente exclusivas, todas derivadas do universo pai por uma única característica de divisão. Segundo o autor, o método da teoria é análogo às regras de divisão lógica, pois as facetas são grupos de conceitos derivados a partir de suas definições reais (gênero, granulação, relações com outros elementos, entre outros). Assim, a TCF respalda a formação de conjuntos de informação que vão desde ideias ou conceitos mais abrangentes (classes básicas) até aos conceitos mais específicos (focos). Em sua origem, esses conjuntos representam a classificação dos assuntos dos documentos, pois são classes de termos agrupados em conformidade com o conteúdo temático dos documentos tratados em uma unidade de informação.

Ranganathan (1967) reelaborou a ideia inicial de Kaiser, criando cinco categorias fundamentais para pensar o domínio, o PMEST: Personalidade (assunto; entidade), Matéria (composição do assunto ou objeto), Energia (ação ou processo em relação ao assunto ou objeto), Espaço (localização geográfica do assunto ou objeto) e Tempo (período). Esse conjunto de

categorias vem sendo empregado no mapeamento de entidades de um domínio e, conforme Campos, Campos e Campos (2010, n.p.), esse procedimento se assemelha ao empregado na construção de ontologias, pois, as “ontologias de topo como a BFO [continuanes e ocorrentes] e a DOLCE [endurantes e perdurantes] distinguem a natureza das entidades do mundo através de categorias que são independentes de domínio”. Para as autoras, no PMEST os “endurantes corresponderiam à categoria de Personalidade, enquanto que perdurantes corresponderiam à categoria de Energia” (idem, idem).

Ranganathan (1967) também indicou três níveis distintos de trabalho: 1) plano das ideias: análise conceitual; 2) plano verbal: expressão verbal (termo; signo) dos conceitos; 3) plano notacional: códigos para a localização dos termos e sinais que podem sugerir as relações entre os conceitos tratados nos documentos. Dahlberg (1979) salienta que os processos de Ranganathan distinguiu, em níveis, o princípio de escolha das facetas e a sua ordem de citação, facilitando a compreensão daquilo que está sendo descrito e representado. Analisando o *Prolegomena* de Ranganathan (1967), percebe-se a seguinte síntese sobre os três planos de trabalho (Quadro 2):

QUADRO 2 – Síntese dos planos e princípios de trabalho com a TCF

Planos/Princípios	Descrição
Plano das ideias	a) Diferenciação: distinção de critério de divisão. Ex.: Tinta: (pelo brilho) Tinta fosca-Tinta brilhosa; (pelo componente) Tinta a óleo; Tinta à base de resina. b) Relevância: características adequadas ao objetivo proposto. Ex.: no <i>Colon Classification</i> , Ranganathan determinou as características: Língua, Gênero, Autor e Obra. c) Verificação: uso de facetas definitivas e que possam ter seu uso validado na literatura ou pelos especialistas do domínio (compromisso ontológico); d) Permanência: a escolha das facetas deve obedecer ao critério de ter um sentido estável dentro do domínio; e) Homogeneidade: ser coeso, com facetas divididas apenas sob um critério (Ranganathan estabelece a “exclusão mútua”, que parece reafirmar o caráter de homogeneidade).
Plano verbal	1) Contexto: o significado de todo termo na classificação deve ser determinado em razão da classe à qual pertence, pois é preciso que todos os termos da classe denotem um só sentido em completude. 2) Aceitabilidade: o termo usado para a faceta ou isolado deve ser aceito no contexto de uso. Caso o termo escolhido se torne obsoleto, o mesmo deve ser trocado por outro mais adequado. 3) Restrição ou Reticência: o termo utilizado não deve ser escolhido na acepção pessoal do classificador.
Plano notacional	1) Sinônimo: havendo mais de um termo representando um mesmo conceito, ainda assim haverá, apenas, um número para ilustrar esse conceito, sejam quantos forem os sinônimos; 2) Homônimo: nenhum número poderá representar mais de um conceito, ainda que sejam homônimos (mesma grafia ou pronúncia e significados diferentes); 3) Hospitalidade: ser flexível, sempre permitindo a inclusão de novos termos, pois o sistema notacional deve acomodar mudanças, atualizações e adições de novos termos; 4) Arquivamento: respeitar a ordem estabelecida pela política utilizada pelo sistema.

<p>Princípios para a escolha da sequência das classes e subclasses</p>	<p>1) Sucessão relevante ou útil: a sequência dos assuntos deve ser útil e adequado ao objetivo proposto. Ex.: do complexo ao simples, do simples ao complexo, contiguidade espacial ou geométrica, ordem convencional ou canônica, garantia literária, ordem evolucionária, posterior no tempo, cronológica, ordem crescente ou decrescente, quantidade crescente ou decrescente, alfabética.</p> <p>2) Sucessão consistente: para classes semelhantes, de preferência, usar uma mesma sequência de assuntos.</p> <p>3) Concomitância: duas características não devem dar origem ao mesmo renque de conceitos. Ex.: as características idade e data de nascimento não devem ser usadas para classificar um mesmo conjunto de indivíduos, pois formariam dois renques com mesmo conjunto de conceitos.</p>
<p>Princípios para a formação de renques e cadeias</p>	<p>1) Exaustividade: todos os elementos (conceitos, entidades, objetos, assuntos) possíveis de uma classe devem fazer parte dela. Ex.: Planeta-Mercúrio-Vênus-Terra-Marte-Júpiter-Saturno-Urano-Netuno</p> <p>2) Mútua exclusão: os itens de uma classe não devem constituir outra classe (renques mutuamente exclusivos). Ex.: no campo da Economia, o termo Exportação: 1) processo de venda (Processo) e 2) quantidade ou valor de produtos e serviços vendidos (Entidade), não pode ser utilizado com as duas definições; a solução poderia ser usar o termo Exportação para Processo e o termo Exportações para Entidade.</p> <p>3) Extensão decrescente: na cadeia a hierarquia deve respeitar a ordem do mais geral para os mais específicos. Ex.: Vertebrado-Mamífero-Felino-Tigre.</p> <p>4) Modulação: termos organizados em módulos, formando uma cadeia (uma classe para cada natureza de subdivisão). Ex.: Instituição Cultural-Biblioteca-Biblioteca Pública-Biblioteca Pública Estadual.</p>
<p>Princípios para a formação de assuntos</p>	<p>1) Dissecção: divisão do domínio de entidades em partes coordenadas entre si, ou seja, de mesmo nível;</p> <p>2) Laminação: construção de camadas de assuntos básicos e ideias isoladas, com a superposição de faceta sobre faceta;</p> <p>3) Desnudação: segmentação dos recursos informacionais do domínio para obter maior especificidade nos assuntos;</p> <p>4) Reunião: combinação de assuntos básicos ou compostos, originando a representação de um assunto complexo;</p> <p>5) Superposição: ligação de assuntos isolados com características, atributos e propriedades distintos, formando isolados de universos diferentes.</p>

Fonte: Ranganathan (1967); Spiteri (1998), Campos (2001), Lima (2004) e Maculan (2011).

Ranganathan também utiliza a noção de classe básica para a análise de assuntos dos conteúdos de documentos e para a organização desses assuntos (isolados) em uma estrutura conceitual, quando cria as facetas. Isto se dá em dois momentos: na Formação de Assuntos e na Formação de Categorias. Para a formação de assuntos, propõe cinco métodos: Dissecção, Laminação, Desnudação, Reunião/Agregação e Superposição. Já na formação de classes básicas, são avaliadas as características comuns de cada conceito, analisados e definidos os termos isolados, conforme suas características comuns, para o procedimento de categorização e a definição das facetas. A partir disso, a análise dos conceitos e a relação entre eles servem como base para a estruturação do sistema facetado.

3. Procedimentos da TCF na construção de taxonomias facetadas

Quando se trata de construir taxonomias facetadas, os rótulos das facetas que representam o conteúdo do recurso informacional podem parecer óbvios: Tipo de Documento, Tipo de Conteúdo, Local, Público, Propósito, Produto, entre outros. O problema aparece quando se vai determinar uma faceta para Assunto. Em geral, as taxonomias facetadas para a representação de produto, como em sites de comércio eletrônico, por exemplo, não têm assuntos, mas, sim, classes de produtos. Também nas taxonomias corporativas, utilizadas para o gerenciamento de documentos, em geral, não há uma faceta para Assunto, podendo ter facetas para Tipo de Documento, Atividades de Negócios, Linhas de Negócio e/ou Departamentos.

Contudo, para recursos de informação de biblioteca, base dados ou repositório de artigos, nos quais haverá relatórios de pesquisa, manuais, apresentações, materiais instrucionais e de treinamento, imagens, vídeos, entre outros, é importante haver uma faceta para Assunto, visando a recuperação do conteúdo de uma coleção. Quando a abrangência de assuntos é ampla, é necessário criar um número de facetas e subfacetas suficientes para cobrir adequadamente a amplitude e a profundidade do conteúdo.

Uma taxonomia facetada é construída com uma combinação de mapeamento *top-down* (identificando facetas e principais termos) e *botton-up* (identificando os termos individuais necessários para representar o conteúdo informacional). A abordagem *top-down* de determinação dos assuntos começa com a identificação do domínio e do escopo do assunto e, em seguida, de quaisquer divisões principais nesse domínio, com base nos princípios do PMEST. A abordagem *botton-up* envolve a pesquisa de vários itens de conteúdo individuais para determinar os principais tópicos sobre os quais os recursos informacionais tratam, visando identificar termos para representar esses tópicos.

Determinar quais são os itens de conteúdo e quais termos descrevem esses conteúdos se assemelha à atividade de indexação, pois é uma tarefa analítico-sintética, cujos princípios estão estabelecidos na TCF. Para manter o número de termos mais gerenciável para análise, é necessário revisar e editar os termos, testando-os e validando-os continuamente.

Para aplicar a TCF na construção de taxonomias facetadas sugerem-se os seguintes procedimentos, adaptados dos princípios de Ranganathan (1967) e do detalhamento elaborado por Oliveira (2018):

- (a) examinar recurso informacional para identificar o domínio – seus conceitos, termos e atributos –, o propósito e o público-alvo;
- (b) definir a linguagem que será utilizada, assim como o nível de especificidade dos assuntos apresentados e o tipo de navegação que será disponibilizado;
- (c) estabelecer as características e/ou atributos do conteúdo – dos recursos de informação –, para a formação de facetas;
- (d) analisar e distribuir os rótulos específicos nas diferentes facetas, sem que eles se sobreponham (mutuamente exclusivos);
- (e) utilizar o método da análise facetada para a formação de assuntos e determinação de rótulos, atendo-se aos princípios de Dissecção, Desnudação e Laminação;
- (f) estabelecer relacionamentos de hierarquia (instância, todo-parte, gênero-espécie), a partir da representação <isA> (é um), incluindo os poli-hierárquicos, conforme norma ISO 25964-1 (2011): aplicação de outras características divisionais do 1) Plano das Ideias: Cânone das Características: Diferenciação, Relevância, Verificação, Permanência, Homogeneidade, Mútua Exclusão; e do 2) Plano Verbal: Cânones do Contexto, da Enumeração (ordem sequencial do assunto), da Atualidade (terminologia atual) e da Restrição (conceito/assunto/atributo aceito no domínio);
- (g) ordenar e agrupar os assuntos e levantamento das facetas dando origem às subfacetas,

observando os princípios de Ranganathan: utilizar as Categorias Fundamentais do PMEST e os Cânones do Plano das Ideias: Formação de *Arrays*: Exaustividade; Formação de Cadeias; Modulação;

- (h) determinar a ordem de apresentação no sistema de classificação facetado, para a organização da estrutura hierárquica (das facetas e subfacetadas), observar os princípios de Sucessão relevante, Sucessão consistente e Sequência útil;
- (i) validar a estrutura e fazer a readequação quando e onde necessário.

O desenvolvimento de uma taxonomia facetada é um processo tanto interativo quanto contínuo e, provavelmente, serão observadas tendências e padrões. Ademais, classes e subclasses podem emergir à medida que as características, atributos e propriedades vão sendo mapeados. Destaca-se que uma taxonomia deve continuar a ser atualizada mesmo após ter sido implementada, sobretudo para as facetadas de assunto, pois um novo conteúdo poderá introduzir novos tópicos que ainda não haviam sido incluídos na estrutura. Para tanto, é necessário a criação de uma política para inserção de rótulos e/ou assuntos na estrutura da taxonomia. As ações de manutenção de uma taxonomia devem prever ações que possam gerar modificações nos referidos instrumentos de gestão terminológica, tais como: exclusão ou inclusão de termos; surgimento de novos processos de negócio; ineficiência na recuperação de informações; tratamento de exceções.

3.1 Aplicações da TCF na construção de taxonomias facetadas

Na contemporaneidade o conceito de taxonomia tem sentido abrangente, e é utilizada em meio digital para criar estruturas de um domínio nos quais são aplicadas, como, por exemplo, para: definir conflitos semânticos em bancos de dados corporativos (KASHYAP; SHETH, 1996); classificar doenças (BURGUN; BODENREIDER, 2001); gerenciar construção (EL-DIRABY; LIMA; FEIS, 2005); classificar os erros humanos em acidentes de trem (REINACH; VIALE, 2006); cobrir causas de mudança e efeitos de mudança em projetos de construção (SUN; MENG, 2008); facilitar a interoperabilidade e a recuperação da informação (CHENG et al., 2010), entre outros. É nos ambientes digitais da atualidade que o aparecimento e uso das taxonomias se relacionam com as formas automatizadas de criação da informação, tornando-as um foco de estudos, principalmente na CI (EDOLS, 2001). Assim, na CI, alguns estudos elaboram estruturas facetadas para representar uma área do conhecimento (ARAÚJO, 2005; TRISTÃO, 2005; NOVO, 2007; COSTA; RAMOS, 2008) ou para aplicações práticas de organização do conhecimento em sistemas de recuperação de informações – SRI (CAFÉ; BRATFISCH, 2007; MACULAN, 2011). Nesta seção serão apresentados alguns trabalhos que desenvolveram sistemas de classificação e/ou taxonomias facetadas para aplicações práticas de organização e acesso a informações em SRI especializados, a saber:

1) *FilmFinder* (<<http://www.film-finder.com/>>): protótipo criado por Ahlberg e Shneiderman (1994), da área da Ciência da Computação e Informática, e aprimorado por Doan, Plaisant e Shneiderman (1995); possui uma modelagem em facetadas, implementado no portal da *National Aeronautics and Space Administration* (NASA).

2) *HIBROWSE* (*High resolution Interface for BROWsing and SEarching*): projeto criado por Pollitt et al. (1996) e desenvolvido a partir de um tesouro facetado que foi construído com base nas recomendações de Ranganathan e do *Classification Research Group* (CRG); apesar de algumas limitações no mecanismo de busca, mostrou o potencial das classificações facetadas para a estruturação do conteúdo de documentos.

3) *Flexible information Access using Metadata in Novel Combinations* – FLAMENCO (<<http://sourceforge.net/projects/flamenco>>): desenvolvido de 1990 a 2000 por uma equipe interdisciplinar (HEARST, 2000), é o projeto de navegação e busca facetada em coleções de documentos mais recorrentemente citado; possui seu recurso informacional

estruturado em diferentes classes básicas e multi-hierarquias, que pode ser navegado por uma interface baseada em metadados facetados, denominada *MatrixView*.

4) *Relation Browser* ou RAVE (<<http://idl89.ils.unc.edu/rb07>>): desenvolvido em uma Escola de Biblioteconomia e Ciência da Informação, iniciado em 1998-1999 e em constante evolução; parte do portal governamental que oferece informações estatísticas de agências americanas; em 2008 estava na versão RB07, quando foi adicionado um mecanismo que exhibe múltiplas facetadas.

5) Mapa hipertextual (MHTX): protótipo criado por Lima (2004), da área da CI, para a organização hipertextual de teses e dissertações, empregando os princípios da TCF para a modelagem conceitual dos documentos, desde a identificação dos termos relevantes até a formação das classes básicas, para uma Biblioteca Digital de Teses e Dissertações.

6) Portal Semântico Educacional (POSEDU): mecanismo de busca e navegação facetada para portais semânticos, criado por Lachtim (2008), da área de Sistemas e Ciência da Computação, com princípios baseados na TCF, oferecendo informações estruturadas em facetadas, por meio de regras formais estipuladas em uma ontologia.

7) TAFNAVEGA: taxonomia facetada navegacional criada por Maculan (2011), da área da CI, para a organização da informação e padronização dos conteúdos informacionais de teses e dissertações para uma Biblioteca Digital de Teses e Dissertações que adotou a estrutura textual do documento e os princípios da TCF como subsídio à sua construção.

Destaca-se o tesauro facetado da Arte e Arquitetura (AAT), desenvolvido pelo *Getty Institute*, em 1994, disponibilizado online, acoplado a um SRI, que foi construído tendo por base os princípios da TAF e a partir de uma classificação facetada, onde as hierarquias são compostas por facetadas e subfacetadas, sendo que cada facetada contém uma classe homogênea de conceitos, cujos itens compartilham características que os distinguem dos itens de outras classes. Como exemplo, cita-se o conceito <mármore> (substância usada na criação de arte e arquitetura) que pertence à classe <Materiais>; e o conceito <impressionista> (estilo de arte) que pertence à classe <Estilos e Períodos> (BROUGHTON, 2005). Segundo Broughton (2005), a classificação facetada é ideal para a construção de tesouros, uma vez que a formação de classes (facetadas) no esquema facetado significa que todos os itens em uma facetada devem ter o mesmo “valor” (ou seja, todos devem ser processos ou todos devem ser agentes, por exemplo). Dessa forma, a autora afirma que, por padrão e essência, as únicas relações exibidas dentro de cada facetada são as hierárquicas, conforme mostrado no Quadro 3.

QUADRO 3 – Formas de organização dentro dos *arrays*

Mostrar múltiplos princípios de divisão dentro de um array	Esclarecer as relações dentro de um array
Sapatos (por temporada) Inverno Primavera (por função) Caminhada Corrida (por estilo) Chuteiras Sandálias (por recurso) Tirinhas Mary Janes (boneca) (por tipo de salto)	Mobília (por material) Móveis de madeira Móveis de plástico (por estilo) Móveis rococo Mobília moderna (por quarto) Mobília de quarto Móveis de escritório (por função) Mobília de armazenamento (por forma) Estantes

Anabelas Meia Pata (pela altura do salto) Salto alto Salto baixo	Mesas Roupeiros Agências Móvelia de dormir
--	---

Fonte: elaborado pelas autoras (2018)

Assim, há duas formas de organização dentro dos *arrays*: a) mostrar a relação entre um conceito superordenado com um subordinado, em especial, quando é necessário usar múltiplos princípios de divisão; b) esclarecer a relação de coordenação para estabelecer relacionamentos entre conceitos de mesmo nível dentro de um array.

Em relação à construção de ontologias, Campos, Campos e Campos (2010, n.p.) afirmam que os princípios da TCF podem “ajudar na formação mais sistemática de renques e cadeias, com o foco da sua compreensão e manipulação por pessoas, sem o prejuízo do seu tratamento pela máquina”, sobretudo no que se refere aos Cânones da: Modulação, Concomitância, Diferenciação, Exclusividade e Sucessão de Classes e Subclasses.

4. Considerações finais

No contexto da CI, a taxonomia facetada constitui um importante instrumento no campo da Organização da Informação e do Conhecimento. Assim, acredita-se que estudos teóricos e conceituais sobre os princípios, métodos, aplicações e a concepção de sistemas de classificação facetada são pertinentes e atuais, considerando o contexto informacional ao qual se está inserido e a utilização de soluções *web*, aplicativos, entre outros, que demandam uma estrutura de conteúdos que possibilite a representação da informação de modo sistematizado. Considera-se que este artigo evidencia que a TCF é um importante insumo para o mapeamento de um domínio para a construção de estruturas hierárquicas ou taxonomias facetadas.

Nota-se a habilidade da TCF em estabelecer uma ordem ou organização de coisas (objetos; entidades), mostrando um mapeamento do conhecimento ou dos recursos informacionais de um dado domínio. Considera-se que essa teoria é um aporte teórico importante na contemporaneidade, pois fornece uma abordagem estruturada, que pode ser utilizada como ponto de partida na modelagem de uma estrutura cognitiva adequada ao usuário, para navegação em sistema de recuperação da informação, assim como para reuso na construção de tesouros e ontologias.

5. Referências

- AGANETTE, E.; ALVARENGA, L.; SOUZA, R. R. Elementos constitutivos do conceito de Taxonomia. *Inf. & Soc.: Est.*, João Pessoa, v. 20, n. 3, p. 77-93, set./dez. 2010.
- AHLBERG, C.; SHNEIDERMAN, B. Visual information seeking: tight coupling of dynamic query filters with starfield displays. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS: Celebrating Interdependence, Boston, MA, USA, Apr. 24-28, 1994. *Proceedings...* New York: ACM, 1994. p. 313-317.
- ARAÚJO, C. A. A. *Análise temática da produção científica em comunicação no Brasil baseada em um sistema classificatório facetado*. 2005. 427 f. Tese (Doutorado em Ciência da Informação), Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.
- BROUGHTON, V. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings: New Information Perspectives*, v. 58, n. 1/2, p. 49-72, 2006.
- BURGUN, A.; BODENREIDER, A. Aspects of the taxonomic relation in the biomedical domain. In: INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY IN

- INFORMATION SYSTEMS, 2001 (FOIS '01), ACM, New York, NY, USA. *Proceedings...* New York, NY: FOIS, 2001. p. 222-233.
- CAFÉ, L. M. A.; BRATFISCH, A. Classificação analítico-sintética: reflexões teóricas e aplicações. *TransInformação*, Campinas, v. 19, n. 3, p. 237-250, set./dez. 2007.
- CAMPOS, L. M.; CAMPOS, M. L. A.; CAMPOS, M. L. M. Diretrizes para definição de domínio no reuso de ontologias biomédicas: uma abordagem baseada na análise do compromisso ontológico. In: XI ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, Inovação e inclusão social: questões contemporâneas da informação, 11., Rio de Janeiro, 25 a 28 de outubro de 2010. *Anais...* Rio de Janeiro, ENANCIB, 2010.
- CAMPOS, M. L. A. *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. 2001. 190 f. Tese (Doutorado em Ciência da Informação) – CNPq/ IBICT/ UFRJ/ECA: Rio de Janeiro, 2001.
- CHENG, C. P. et al. Albert Jones regulation retrieval using industry specific taxonomies. *Artificial Intelligence and Law*, v. 16, n. 3, p. 277-303, Sept. 2008.
- CONWAY, S.; SLIGAR, C. Building taxonomies, en su *Unlocking knowledge assets*. Redmont: Microsoft Press, pp. 105-124, 2003.
- COSTA, F. C.; RAMOS, L. B. Análise facetada: em busca de uma classificação para o teatro. *Ponto de Acesso*, Salvador, v. 2, n. 3, p. 30-42, dez. 2008.
- CRUZ, R.; DORNELES, C. F.; GALANTE, R. Especificação de uma Taxonomia para Metadados Multifacetados em Ambientes de Integração de Dados na Web. In: VI ESCOLA REGIONAL DE BANCO DE DADOS, 6., 2010, Joinville. *Anais...* Joinville: ERBD, 2010.
- DAHLBERG, I. *Teoria da classificação, ontem e hoje*. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA, Rio de Janeiro, 12-17 de setembro de 1972, Brasília. *Anais...* Brasília: IBICT/ABDF, v. 1, 1979.
- DOAN, K.; PLAISANT, C.; SHNEIDERMAN, B. Query previous in networked information systems. In: INTERNATIONAL FORUM ON RESEARCH AND TECHNOLOGY ADVANCES IN DIGITAL LIBRARIES, 3., 1995, Washington. *Proceedings...* Washington, DC: IEEE Computer Society, 1995.
- EDOLS, L. Taxonomies are what?. *Free Pint*, n. 97, Oct. 2001.
- EL-DIRABY, T. A.; LIMA, A.; FEIS, B. Domain taxonomy for construction concepts: toward a formal ontology for construction knowledge. *Journal of Computing in Civil Engineering ASCE*, v. 19, n. 4, p. 395-406, 2005.
- HEARST, M. A. Next generation web search: setting our sites. *IEEE Data Engineering Bulletin*, v. 23, n. 3, p. 38-48, 2000.
- KASHYAP, V.; SHETH, A. Schematic and semantic similarities between database objects: a context-based approach. *VLDB Journal*, v. 5, n. 4, 1996.
- LACHTIM, F. A. *Organização e instanciação automática de conteúdos em portais semânticos*. 2008. 192 f. Dissertação (Mestrado em Ciência da Computação), Instituto Militar de Engenharia do Rio de Janeiro, Rio de Janeiro, 2008.
- LIMA, G. A. B. O. *Mapa hipertextual (MHTX): um modelo para a organização hipertextual de documentos*. 2004. 199 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência

- da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.
- LIMA, G. Â. B. O. *Mapa hipertextual (MHTX): um modelo para organização hipertextual de documentos*. 2004. 204 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.
- MACULAN, B. C. M. S. *Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos*. 2011. 191 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
- NOVO, H. F. *A elaboração de taxonomia: princípios classificatórios para domínios interdisciplinares*. 2007. Dissertação (Mestrado em Ciência da Informação) - UFF/IBICT.
- NOVO, H. F. *A elaboração de taxonomia: princípios classificatórios para domínios interdisciplinares*. 2007. 172 f. Dissertação (Mestre em Ciência da Informação), Universidade Federal Fluminense, Instituto de Arte e Comunicação Social, Niterói, Rio de Janeiro, 2007.
- OLIVEIRA, E. D. *Proposta de um modelo de hipertexto com abordagem semântica para a representação do conhecimento no domínio temático da Intensificação Agropecuária*. 2018. 231 f. Dissertação (Mestrado em Ciência da Informação), Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Belo Horizonte: PPGOC, 2018.
- POLLITT, A. S. et al. View-based searching systems: progress towards effective disintermediation. In: ONLINE INFORMATION'96, INTERNATIONAL ONLINE INFORMATION MEETING, 20.,1996. London, England. *Proceedings...* London, England: [s.n], 1996. p. 433-441.
- RANGANATHAN, S. R. *Prolegomena to library classification*. 3. ed. London: Asia Publishing House, 1967.
- REINACH, S.; VIALE, A. Application of the human error framework to conduct train accident/incident investigations. *Accident Analysis and Prevention*, v. 38, p. 396- 406, 2006.
- SILVA, A. R. *Estudo dos princípios de categorização na Biblioteconomia e Ciência da Informação: Ranganathan - entre a teoria clássica e a abordagem cognitiva contemporânea*. 2010. 193f. Dissertação (Mestrado em Ciência da Informação), Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.
- SUN, M.; MENG, X. Taxonomy for change causes and effects in construction projects. *International Journal of Project Management*, v. 27, n. 6, p. 560-572, 2009.
- TAYLOR, Chris. *An introduction to metadata*. Brisbane, Australia: University of Queensland Library, 2003.
- TRISTÃO, A. M. D. *Classificação da informação na indústria da construção civil: uma aplicação em placas cerâmicas para revestimento*. 2005. 269 f. Tese (Doutorado em Engenharia Civil), Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Santa Catarina. Florianópolis, PPGECC, 2005.
- VICKERY, B. C. *Faceted classification: a guide to the construction and use of special schemes*. London: Aslib, 1960.