CATEGORIZAÇÃO CONCEITUAL DA INFORMAÇÃO JURÍDICA

CONCEPTUAL CATEGORIZATION OF LEGAL INFORMATION

Fabiana Gulin Longhi Palacio

Vânia Mara Alves Lima

Resumo: organizar a informação e estabelecer procedimentos para a categorização conceitual, na elaboração de linguagens documentárias utilizadas para a sua representação, possibilitam ao usuário o acesso a respostas para solucionar questões pertinentes à sua pesquisa. Delimitamos a área do Direito como domínio a ser trabalhado com o objetivo de identificar princípios para a categorização conceitual da informação jurídica que contribuam para elaborar linguagens documentárias mais eficientes para a indexação da sua produção documental e que preservem o contexto em que o termo jurídico é empregado. O referencial teórico utilizado para a definição do termo "categorização", o qual inclui autores como Aristóteles, Kant, Wittgenstein e Rosch, aliado à Teoria do Conceito de Dahlberg e à Teoria Comunicativa da Terminologia de Cabré, assim como a análise de linguagens documentárias utilizadas para a indexação no âmbito do Direito, permitiram a identificação de alguns princípios que devem ser considerados na elaboração desses instrumentos de representação e controle terminológico para a recuperação da informação. Verifica-se que a categorização conceitual pode ser considerada como pedra fundamental na estruturação de uma linguagem documentária, pois possibilita a relação entre os conceitos, representados por termos, dentro de um domínio específico do conhecimento, tornando mais eficaz a recuperação da informação.

Palavras-chave: Categorização conceitual. Informação jurídica. Linguagem documentária. Recuperação da informação. Terminologia.

Abstract: organizing information and establishing procedures for the conceptual categorization in the development of indexing languages used for such representation allows the user to access the answers for solving issues relevant to his or her search. The area of Law was delimited as the subject area to be worked purposing to identify principles for the conceptual categorization of legal information which will contribute to the elaboration of instruments of representation and terminology control, that means, to produce indexing languages that are more efficient for the indexing of document production and which preserve the context in which the legal term is employed. The theoretical framework used for the definition of the term "categorization", which includes authors such as Aristotle, Kant, Wittgenstein and Rosch, allied to the Theory of Concept by Dahlberg and to the Communicative Theory of Terminology by Cabré, as well as the analysis of indexing languages used for indexation in the scope of Law, allowed the identification of some principles which should be considered in the preparation of these instruments of representation and terminology control for the retrieval of information. It can be seen that the conceptual categorization can be considered a cornerstone in the structuring of an indexing language, since it enables the relation between the concepts represented by terms within a specific area of knowledge making the information retrieval more effective.

Keywords: Conceptual categorization. Legal information. Indexing language. Information retrieval. Terminology.

1. INTRODUÇÃO

O estabelecimento de critérios para a categorização conceitual da informação, objetivo deste trabalho, é utilizado pela Ciência da Informação no armazenamento e na recuperação dos documentos que compõem os acervos dos sistemas de informação. Os procedimentos para o tratamento, organização e recuperação da informação envolvem etapas como a análise, síntese e representação dos conteúdos informacionais do documento, de forma a elaborar uma informação documentária que possibilite ao usuário o acesso às respostas para solucionar questões pertinentes à sua pesquisa.

Nesta pesquisa, delimitamos a área do Direito como domínio a ser trabalhado na busca de princípios para categorização da informação jurídica de maneira a permitir a relação entre as diferentes subáreas do Direito, preservando a particularidade do contexto em que o termo jurídico é empregado, a fim de contribuir com a qualidade da representação desta informação e, consequentemente, facilitar a sua recuperação.

Justifica-se a importância do estabelecimento de princípios claros para a categorização da informação jurídica por possibilitar a elaboração de instrumentos para representação e controle terminológico, ou seja, a construção de linguagens documentárias mais eficientes para a representação e recuperação da informação que também preservem o contexto em que o termo jurídico é empregado.

2. PROCEDIMENTOS METODOLÓGICOS

A pesquisa bibliográfica foi realizada utilizando o Banco de Dados Bibliográficos da Universidade de São Paulo – Dedalus, a Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo e a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação – BRAPCI,¹ utilizando as palavras-chave Categorização, Categoria, Categorias, Conceito, Conceitos, Categorização Conceitual, Terminologia, Linguagem Documentária, Análise Documentária, Filosofia Jurídica, Filosofia do Direito, Teoria Jurídica e Teoria do Direito. Foi constatado, durante a pesquisa, que alguns autores utilizam Classificação quando se referem a Categorização, por isso, a palavra-chave Classificação

¹ Disponível em: <http://www.brapci.ufpr.br>.

também foi utilizada na busca. Com base nas citações e referências dos documentos obtidos foi possível ampliar o número de autores consultados. e acessar o texto completo dos artigos citados na base de dados nos periódicos eletrônico.

O quadro referencial teórico é fundamentado nas definições de categorização sob o ponto de vista da Filosofia com estudos de Aristóteles, Kant, Wittgenstein e Rosch e sob o ponto de vista da Ciência da Informação com estudos de Ranganathan, Vickery, Iyer e os estudos realizados pelo Classification Research Group (CRG). Procurando caracterizar o discurso jurídico a partir de autores consagrados na área, como Alves, Reale e Ferraz Junior, os quais relacionam as principais categorias existentes no Direito.

Como exemplo de como os instrumentos de representação e controle terminológico categorizam a informação jurídica, utilizamos o Vocabulário Controlado da Universidade de São Paulo (VOCAUSP) e o Vocabulário Controlado Básico do Senado Federal (VCB). Para a coleta de termos no âmbito do Direito utilizamos como exemplo o *Iusdata* — banco de dados de artigos de periódicos da Biblioteca da Faculdade de Direito da USP.

Os princípios para a categorização conceitual da informação jurídica se baseiam na Teoria do Conceito de Dahlberg e na Teoria Comunicativa da Terminologia de Cabré.

3. CATEGORIZAÇÃO: DA FILOSOFIA À CIÊNCIA DA INFORMAÇÃO

Na perspectiva filosófica, a categorização foi estudada por Aristóteles (384-322 a.C.) (2010, p. 41), que divide as categorias em *substância* (o quê), *quantidade* (quão grande, quanto), *qualidade* (tipo de coisa), *relação* (com o que se relaciona), *lugar* (onde), *tempo* (quando), *posição* (qual a postura), *estado ou condição* (em quais circunstâncias), *ação* (quão ativo, qual o fazer) e *paixão* (quão passivo, qual o sofrer).

Seguindo a linha aristotélica, mas sob outra perspectiva, Kant (2001, p. 108) nos apresenta as categorias como conceitos puros do entendimento. Ele defende que o conhecimento ocorre de forma empírica por meio das categorias que os seres humanos possuem *a priori*, dentro das quais os fenômenos são entendidos. São elas: a) *Da quantidade* (unidade, pluralidade e totalidade); b) *Da qualidade* (realidade, negação e limitação); c) *Da Relação* (Inerência e subsistência – *substantia et accidens*, causalidade e dependência – causa e efeito e comunidade (ação recíproca entre o agente e o paciente); e d) *Da modalidade* (possibilidade – impossibilidade, existência – não-existência, necessidade – contingência). (KANT, 2001, p. 110-111).

A visão clássica de categorização pode ser bem aplicada às coisas concretas, onde uma palavra representa o conceito de algo que existe na realidade. Wittgenstein, estudioso da linguagem, trata da dificuldade do estabelecimento de fronteiras rígidas entre as categorias.

Wittgenstein (1995, p. 226-229), usa os *jogos* para exemplificar sua teoria das "semelhanças de família", na qual os elementos não apresentam características comuns à classe em que estão inseridos, mas sim características de parentesco entre si. Existem *jogos de tabuleiros*, *jogos de cartas*, *jogos de combate*, *jogos de bola* etc. e, observando as características destes *jogos*, não é possível identificar as mesmas características em todos eles, mas algumas características de parentesco entre eles. Por exemplo, os *jogos de tabuleiros* e os *jogos de bola* apresentam algumas características comuns e outras características que os diferenciam. Tanto o *jogo de xadrez* quanto o *jogo de tênis* são *jogos* em que há vencedores e perdedores, sendo esta uma característica comum a ambos; porém, as habilidades necessárias aos jogadores de *xadrez* são diferentes das habilidades necessárias aos jogadores de *tênis*.

Na área da Ciência da Informação, um dos mais importantes pesquisadores do tema foi Ranganathan, com sua análise e organização de assuntos por facetas (1959, p. 67-68), pela qual cada faceta de qualquer assunto pode ser considerada uma manifestação de uma das cinco categorias fundamentais: a) *Personalidade*, categoria utilizada como faceta residual, depois que todos os outros aspectos foram separados; b) *Matéria*, categoria referente ao material; c) *Energia*, categoria que conota ação, mas também compreende estrutura (morfologia), função (fisiologia), mau funcionamento ou doença, a ação ambiental ou ecologia, filogenia, ontogenia, entre outras ideias semelhantes; d) *Espaço*, categoria referente à área geográfica; e e) *Tempo*, categoria referente ao tempo cronológico, as quais seguem a sequência: [P] *Personalidade*, [M] *Matéria*, [E] *Energia*, [S] *Espaço* e [T] *Tempo*.

Baseado na Classificação de Dois Pontos de Ranganathan e em sua análise de facetas, e com o objetivo de estudar a elaboração de sistemas de classificação mais flexíveis e para áreas especializadas, surgiu, em 1952, em Londres, o Classification Research Group (CRG), o qual determinou as seguintes facetas do objeto/coisa: *Todo/Tipo, Partes, Constituinte/Material, Propriedade/Processos, Operações, Agentes, Lugar, Tempo* e *Formas de Apresentação*. A ordem estabelecida pelo CRG parte do princípio que coloca o fim antes do meio de obtê-lo, o todo antes das partes (BARBOSA, 1972, p. 76-78).

Atualmente a categorização para fins de recuperação em sistemas de informação se apresenta na forma de linguagens documentárias como os tesauros e os vocabulários controlados, que são listas de termos (descritores) que denominam conceitos e se relacionam

entre si dentro de categorias gerais. Essa lista que é consultada tanto no momento da indexação quanto na busca de informações deve prever a categorização conceitual no momento de sua elaboração. Dahlberg (1978) ao elaborar a Teoria do Conceito, na qual discute a categorização, contribuiu consideravelmente neste aspecto.

Partindo do princípio de que o conhecimento fixa-se através dos elementos da linguagem e de que a linguagem constitui-se "na capacidade do homem de designar os objetos que o circundam, assim como comunicar-se com seus semelhantes" (DAHLBERG, 1978, p. 101), a autora divide esses objetos, com os quais o homem se relaciona, em objetos individuais e objetos gerais. O significado de um conceito é entendido formulando-se enunciados sobre os atributos necessários ou possíveis dos objetos que se obtém as características dos respectivos conceitos (DAHLBERG, 1978, p. 102).

Para Dahlberg (1978, p. 102), um atributo pode corresponder a uma hierarquia de características, uma vez que "o predicado de um enunciado pode tornar-se sujeito de novo enunciado e assim sucessivamente até atingirmos uma característica tão geral que possa ser considerada uma categoria. (Entende-se aqui por categoria o conceito na sua mais ampla extensão)."

Baseada nas categorias aristotélicas, Dahlberg (1978, p. 102-103) todas as características possíveis: *Matéria (substância)*, *Qualidade*, *Quantidade (extensão)*, *Relação*, *Processo (atividade)*, *Modo de ser*, *Passividade*, *Posição*, *Localização (lugar)* e *Tempo*.

Para Dahlberg (1978, p. 105), a categorização conceitual combina os conceitos de acordo com: a) *objetos:* plantas, produtos, papel; b) *fenômenos:* crescimento, chuva, tráfego; c) *processos:* imprimir, sintetizar etc.; d) *propriedades:* cego/cegueira, suave/suavidade etc.; e) *relações:* causalidade, necessidade; e f) *dimensão:* espaço, tempo, posição.

Iyer (1995, p. 40-41) baseada no triângulo conceitual de Dahlberg, em que o conceito é composto por (A) Referente, (B) Características e (C) Expressão Verbal, afirma que os conceitos inferem informações às palavras e a categorização é baseada nas semelhanças dos conceitos, que, dentro de uma categoria, estabelecem relações uns com os outros. Estas relações podem ser baseadas em características ou atributos físicos, ambientais, ou por funções. Por exemplo, um animal que possua bico e asas pertencerá à categoria de pássaros.

No paradigma clássico, um elemento pertence ou não a uma categoria de acordo com suas características, sendo que, para pertencer a determinada categoria, ele precisa apresentar as características que definem tal categoria; caso contrário, não pertencerá a ela.

Iyer (1995, p. 42-45) considera que dois autores mudaram a visão clássica da categorização: Ludwig Wittgenstein por sugerir que a categorização considerasse a semelhança de família, não definindo um limite rígido para a categoria e Eleanor Rosch por desenvolver o nível básico de categorias e pela teoria dos protótipos, medindo a probabilidade de sugestões válidas para as categorias por meio de um estudo empírico.

Segundo Rosch² (apud IYER, 1995), níveis básicos de categorias são aqueles em que as sugestões de características válidas são maiores. Normalmente são categorias localizadas nos níveis médios das hierarquias dos sistemas taxonômicos. Por exemplo, no estudo Eleanor Rosch foi solicitado a algumas pessoas que listassem propriedades características dos seguintes elementos da categoria Mobiliário:

Mobiliário

Cadeira

Cadeira de Cozinha

O elemento *cadeira* foi o que recebeu mais sugestões de características válidas; o que o torna o objeto de nível básico da categoria. Dentro das categorias, estes objetos são os mais fáceis de serem conceitualizados, reconhecidos e desenhados mentalmente. Neste nível, o conceito é compreendido como um todo, quando o conceito superordenado é muito geral e o conceito subordinado é muito específico (IYER, 1995, p. 45-47).

Na teoria dos protótipos, proposta por Rosch, a categoria é composta por uma estrutura gradual, em que um elemento representa melhor a categoria do que os demais, e este elemento é o protótipo da categoria. Os protótipos podem variar de acordo com a cultura ou o ambiente — por exemplo, na América, a *rosa* pode ser o protótipo da categoria *flores*, mas não o será no Deserto do Kalahari. Nesta graduação, os elementos variam do mais ao menos típico da categoria (IYER, 1995, p. 51).

Pensando na organização do conhecimento, visando sua recuperação e disponibilidade de acesso, G. Lima (2004, p. 64) afirma que, sofrendo forte influência dos processos cognitivos, "a informação passa pelos processos de categorização, indexação e interação homem-máquina até chegar à recuperação da informação".

A categorização conceitual não é simples, ela exige um trabalho cognitivo complexo e só pode ser realizada depois de um estudo reflexivo da linguagem de especialidade.

² ROSCH, Eleanor et al. Basic Objects in Natural Categories. **Cognitive Psychology**, n. 8, p. 382-439, 1976.

Tendências da Pesquisa Brasileira em Ciência da Informação, v.8, n.1, jan./jun. 2015.

[...] Sabe-se que, para organizar o conhecimento de qualquer área, da sua representação à sua recuperação, estudam-se primeiramente os conceitos que compõem esse campo do conhecimento e a relação entre eles. A influência do contexto é tão importante na categorização que, qualquer descrição individual na classificação que o desconsidere, pode mostrar-se muito limitada. Categorias e hierarquias de categorias são a melhor maneira de organizar o conhecimento para recuperação, pelo óbvio motivo de que a informação estruturada é mais fácil de ser recuperada do que uma informação desorganizada [...] (LIMA, 2010, p. 120).

4. O DISCURSO JURÍDICO

Assim como toda linguagem de especialidade, a linguagem jurídica é é originária da linguagem natural. Para Bittar (2001, p. 168), "o universo do discurso jurídico pode ser dito um universo oriundo da linguagem natural, encontrando-se atualmente autônomo em relação a ela, apesar de preservar sua relação dialética com os sistemas que o circundam". O Direito, ao longo do tempo, tão específico que muitas vezes é incompreensível para o leigo.

Alves (2005, p. 210-211), estudioso e professor de lógica e teoria jurídica, discute a categorização conceitual afirmando que os conceitos são ordenados de acordo com sua intensão e sua extensão, premissa adotada por Dahlberg quando aborda a elaboração da definição do conceito (1978). Em um sistema conceitual, hierarquicamente ordenado, os conceitos de maior conteúdo extensivo (extensão) e menor conteúdo intensivo (conceito de ser) encontram-se no extremo superior do sistema — são as denominadas *categorias gerais*. No outro extremo, no limite inferior do sistema conceitual, encontram-se os conceitos de maior conteúdo intensivo e menor conteúdo extensivo, são os *objetos individuais*.

Outro especialista do Direito, Reale (1980, p. 335-337) afirma que, para ser estudada, toda ciência precisa ser dividida em partes claramente discriminadas, e apresenta a primeira divisão do Direito, estabelecida pelos romanos, em *Direito Público*, que se ocupa das coisas do Estado, e em *Direito Privado*, que se ocupa dos interesses particulares dos cidadãos.

Para Ferraz Junior (2007, p. 140-144), a distinção entre *Direito Público* e *Direito Privado* e a classificação dos diferentes ramos dogmáticos do Direito dentro deles é importante para garantir a certeza e segurança na decisão de conflitos.

Seguem, no quadro 1, as propostas de categorização do Direito elaboradas por Reale (1980, p.338-366) e Ferraz Junior (2007, p. 140-144). Ambas podem ser úteis na categorização conceitual da informação jurídica para fins documentários e podem contribuir para a melhoria das linguagens documentárias utilizadas atualmente na organização,

representação e recuperação deste tipo de informação, contribuindo para o desenvolvimento do domínio.

A categoria de *Direito Público* elaborada por Ferraz Júnior contém mais ramos do Direito do que a mesma categoria elaborada por Reale. Já na categoria de *Direito Privado*, a principal diferença entre eles está no fato de Reale ter apresentado os ramos jurídicos subordinados ao *Direito Civil*, na qual o *Direito Comercial* se encontra, enquanto Ferraz Junior colocou o *Direito Civil* e o *Direito Comercial* coordenados entre si e subordinados ao *Direito Privado* e não apresentou os ramos jurídicos subordinados ao *Direito Civil*.

Quadro 1 - Categorização sistematizada do Direito

Área do Direito	REALE (1980)	FERRAZ JUNIOR (2007)
Direito Público	Direito Público	Direito Público
Direito Privado	Direito Privado Direito Civil Direitos Pessoais Direitos Obrigacionais Direitos Associativos Direitos Reais Direito de Família Direito das Sucessões Direito Comercial Direito Agrário	Direito Privado

5. LINGUAGENS DOCUMENTÁRIAS PARA REPRESENTAÇÃO DA INFORMAÇÃO JURÍDICA

Um domínio ou área de especialidade é representado pela linguagem documentária por meio de descritores baseados na garantia literária, isto é, na terminologia do domínio retirada da literatura da área de especialidade e na garantia de uso, ou seja, na linguagem dos especialistas. Quando uma linguagem documentária é elaborada obedecendo a critérios de padronização e categorização bem definidos, a informação é estruturada de forma que possibilite a sua adequada recuperação. Conforme Kobashi (1996, p. 9), a linguagem documentária vai "dar forma às informações selecionadas, de modo a torná-las documentariamente manipuláveis" visando à sua recuperação pelo usuário.

As linguagens documentárias mais representativas na área jurídica devido aos acervos que representam são o Vocabulário Controlado Básico do Senado Federal (VCB) e o Vocabulário Controlado da Universidade de São Paulo (VOCAUSP).

O VCB é a linguagem documentária utilizada pela Rede Virtual de Bibliotecas – Congresso Nacional – RVBI desenvolvido na década de 1980 pelo então Subsistema de Administração de Bibliotecas (SABI) ou Rede SABI. (GRUPO de Estudo do Tesauro da Rede Virtual de Bibliotecas do Congresso Nacional (RVBI), 2005, p. 2). O VCB foi definido por Loddo (1991, p. 3) como "uma relação de termos de indexação com vistas à uniformidade de armazenagem de informações, bem como à facilidade de recuperação de documentos". A organização e as relações dos descritores desta linguagem documentária são apresentadas por meio de relações de equivalência (USE / NÃO USE), relações hierárquicas (Termo Geral – TG e Termo Específico – TE) e relações partitivas e associativas (Termo Associado – TA) (GRUPO de Estudo do Tesauro da Rede Virtual de Bibliotecas do Congresso Nacional (RVBI), 2005,p.10). Quando acessamos os registros dos termos, o sistema nos apresenta o termo geral, os termos específicos, os termos relacionados e as notas referentes aos termos, o que nos permite conhecer o contexto dos termos a que eles se referem, suas definições etc.

É importante destacar que a sua estrutura hierárquica não está visível na interface do usuário, porém, observa-se que se trata de uma linguagem categorizada devido ao resultado da pesquisa pela Base de Dados Vocabulário Controlado Básico – VCB e pela Autorias Padronizadas – AUTR. Por exemplo, o resultado da busca por *Fraude Eleitoral* nos apresenta os campos Termo geral e Sub-base, onde verificamos a seguinte categorização:

Direito Eleitoral (Sub-base)

Crime Eleitoral (Termo geral em relação a Fraude Eleitoral)

Fraude Eleitoral

O VOCAUSP³ foi elaborado com o objetivo de ser usado para indexação de assuntos no Banco de Dados Bibliográficos da Universidade de São Paulo – Dedalus. Composto por termos autorizados para indexação, não autorizados (remissivas) e os falsos elos, que servem para agrupar termos mais específicos, o Vocabulário Controlado USP "abrange as áreas do conhecimento inerentes às atividades de ensino, pesquisa e extensão da Universidade de São Paulo." (LIMA; KOBASHI; IMPERATRIZ, 2002, p. 225). Entre estas áreas do conhecimento, encontra-se a área jurídica.

A categorização do Vocabulário Controlado USP foi definida quando do estabelecimento dos subgrupos responsáveis pela coleta e estruturação dos termos dentro de domínios afins e isso se reflete na sua macroestrutura (grandes categorias) dentro de cada domínio. No caso do *Direito*, a categorização foi determinada pelos indexadores da própria área, a partir da análise conjunta com os demais indexadores dos domínios participantes no mesmo subgrupo, a saber: *Administração*, *Economia* e *Ciências Sociais*.

A área jurídica encontra-se estruturada com o termo *Direito* como classe geral e todos os ramos jurídicos subordinados a ele, causando divergência com a estrutura clássica proposta por especialistas consagrados da área jurídica, que dividem o Direito em *Direito Público* e *Direito Privado*, e colocam os demais ramos do Direito subordinados a um ou a outro (NARUKAWA, 2012, p. 4).

Em 1986, a Biblioteca da Faculdade de Direito da Universidade de São Paulo iniciou o Serviço de Disseminação Seletiva da Informação (SDI) para divulgar aos docentes da Faculdade de Direito os artigos acadêmicos dos periódicos nacionais e internacionais disponíveis no acervo da Biblioteca. Além de divulgar os últimos artigos publicados na área jurídica, este serviço também exerce a função de avaliação do conteúdo destas publicações de acordo com os interesses dos usuários desta área. Em 1998 este banco de dados foi disponibilizado online com o nome de *IusData*. A pesquisa é realizada por assunto, título e/ou autor do artigo, o que possibilita uma expressiva economia de tempo para o usuário e leva a uma grande aceitação e utilização pelos docentes da FDUSP.

_

³ Disponível em: http://143.107.73.99/Vocab/Sibix652.dll.

Para a indexação destes artigos, este banco de dados dispõe de uma lista de assuntos, baseada, inicialmente, no VCB do Senado Federal; porém, como se trata de uma lista de assuntos, não há uma categorização e nem está previsto o controle da sinonímia, da polissemia e da homonímia, acarretando problemas na recuperação da informação jurídica.

Embora não seja uma linguagem documentária categorizada, a lista de assuntos utilizada para a representação e recuperação dos artigos de periódicos referenciados no *IusData* foi estudada neste trabalho como fonte de terminologia jurídica atualizada devido à dinâmica deste tipo de publicação, pois os artigos de periódicos estão sempre atualizados com o desenvolvimento científico da área, de modo que termos novos ou candidatos sejam facilmente percebidos e identificados nesta linguagem.

6. CATEGORIZAÇÃO CONCEITUAL DA INFORMAÇÃO JURÍDICA

O processo de categorização deve ser embasado fundamentalmente nos conceitos e no relacionamento entre os conceitos de um domínio, o que poderíamos denominar aqui de *categorização conceitual*. A categorização conceitual permitiria a utilização de descritores corretos, uma vez que facilitaria a identificação e o controle da sinonímia, da polissemia e da homonímia, tanto na representação quanto na recuperação da informação jurídica. Os princípios para uma categorização conceitual podem ser elencados a partir da Teoria do Conceito de Dahlberg e da Teoria da Terminologia.

A terminologia é entendida neste trabalho como o conjunto de termos (designações), de um domínio específico (ISO 1087-1 2000, p. 10), o qual vem sendo indicado como fonte de referência para a construção das linguagens documentárias. Isto porque o termo, ao denominar um conceito do domínio, ou seja, ao trazer em sua definição enunciados que permitem a compreensão do conceito, serve de referência para o descritor de uma linguagem documentária que pretenda representar a informação produzida naquele domínio. No dizer de V. Lima (2007, p. 121) e Lara (1993, p. 75), a terminologia serve de referência e funciona como base legal para a construção da linguagem documentária, pois, ao articular o conceito e o termo, possibilita a elaboração de uma informação documentária que vai efetivar a comunicação entre o usuário e o sistema de informação.

Nesta pesquisa nos embasamos nos trabalhos de Cabré e sua Teoria Comunicativa da Terminologia (TCT), que reconhece o caráter comunicativo da língua através da circulação dos termos em domínios diferentes e entre estes e a língua geral e tem se mostrado a mais

adequada para embasar uma terminologia que funcione como referente de uma linguagem documentária. "Para a TCT, os termos *são usados* em um domínio" (CABRÉ, 1999, p. 141,⁴ apud BARROS, 2004, p. 59). A TCT visa, além da representação da informação especializada, a transferência desta informação por meio da comunicação, mas, de qualquer maneira, a elaboração de uma terminologia ainda hoje é objeto das normas internacionais ISO 1087-1 e ISO 704, que definirão o conceito, o termo e as relações que permitem a sua estruturação.

De acordo com a norma ISO 1087-1 (2000, p. 2-5), o conjunto de conceitos estruturados de acordo com as relações entre si é chamado de "sistema conceitual", tomando como ponto de partida o "campo conceitual", que é o conjunto de conceitos não estruturados relacionados tematicamente. O critério de subdivisão do sistema conceitual é estabelecido pelas "características" que descrevem os conceitos. As definições dos diferentes tipos de conceitos e suas relações permitem a categorização da terminologia.

Comparando a norma ISO 1087-1 (2000), que estabelece um vocabulário básico para teoria e aplicação do trabalho terminológico, com a norma ISO 704 (2000), que estabelece princípios e métodos para a realização do trabalho terminológico, o *termo* é definido como:

Para a norma ISO 1087-1 (2000, p. 6) o *termo* é a "designação verbal de um conceito geral em um campo de assunto específico". E a norma ISO 704 (2000, p. 24) define *termo* como "designação que consiste em uma (termo simples) ou mais palavras (termo composto) para representar um conceito geral em uma linguagem específica".

Percebe-se semelhança nas definições das duas normas terminológicas. O termo é considerado uma designação para um conceito dentro de uma linguagem de especialidade. Tal designação pode ser feita por um termo simples, representado por uma única palavra, ou por um termo composto, representado por mais de uma palavra.

O termo designa um conceito em uma determinada linguagem de especialidade e é representado por uma palavra. Para Maciel (2001, p. 40), "são os termos que revelam a especialização e que caracterizam imediatamente a linguagem de especialidade".

Segundo Cintra e col. (1996, p. 20), a Terminologia, através de procedimentos sistemáticos, seleciona e/ou cria termos para as noções dos domínios, relacionando-os através de definições. Assim, a Terminologia possibilita a relação entre os termos, questão importante para a representação da informação por meio da categorização conceitual.

⁴ CABRÉ, M. T. Hacia una teoría comunicativa de la terminología: aspectos metodológicos. In: CABRÉ, M. T. **La terminologia**: representación y comunicación. Barcelona, IULA, 1999, p. 129-150.

A terminologia de determinado domínio do conhecimento pode ser buscada na literatura da área, na linguagem dos especialistas, na linguagem utilizada pelos usuários na busca pela informação etc. Porém, embora restrita a uma área de especialidade, normalmente há divergências nesta linguagem, pois o usuário e especialista podem utilizar termos diferentes para o mesmo conceito.

Para Cabré (1999, p. 245), a normalização terminológica facilita a representação da informação para o indexador devido à rigidez desta normalização, mas força o usuário a adequar sua linguagem natural à linguagem artificial utilizada pelo sistema de informação.

A própria Cabré (1999, p. 246), com a TCT, apresenta uma solução para a questão. Ela propõe uma normalização terminológica rígida para a representação da informação, mas flexível para a recuperação da informação, prevendo a sinonímia, a polissemia e a homonímia, de acordo com a linguagem do usuário, utilizando-se de recursos tecnológicos.

A terminologia do domínio funciona como referente para a elaboração da linguagem documentária, onde os termos denominados de *descritores* representam os conceitos que constituem esse domínio. Na linguagem documentária, ao operacionalizar as categorias, cada termo deverá ser considerado em relação a cada categoria e em relação aos demais termos que compõem esse instrumento. Para isso, é necessário delimitar o domínio destes documentos e atribuir significado para os termos dentro do sistema, sempre considerando o contexto em que se encontra inserido.

Dentro de um sistema de informação, o domínio é representado pela forma que a linguagem documentária dá a ele por meio da categorização conceitual. É esta categorização que dá consistência ao significado dos descritores da linguagem documentária, pois permite reunir os descritores em classes e relacioná-los com outros descritores.

Especialistas consagrados da área jurídica já se preocuparam em categorizar este tipo de informação, seja para fins estudo da área (REALE, 1980, p. 335), seja para a finalidade prática da decidibilidade de conflitos (FERRAZ JUNIOR, 2007, p. 140). O estudo destes especialistas gerou categorizações que podem ser utilizadas para fins documentários no sentido de organizar, representar e disponibilizar este tipo de informação.

Segue, no quadro 2, uma síntese dos pontos abordados no processo de categorização da Filosofia, da Ciência da Informação e do discurso jurídico, na tentativa de elencar os itens que devem ser considerados na categorização conceitual da informação jurídica.

Quadro 2 - Processo de categorização na Filosofia, na Ciência da Informação e no discurso jurídico

CATEGORIZAÇÃO Filosofia Discurso Jurídico Ciência da Informação ARISTÓTELES RANGANATHAN **ALVES** Divide as Categorias em: Cria a classificação facetada, na A divisão dos termos segue a qual cada faceta pode ser lógica Todo-Parte, que pode Substância (o quê); Quantidade considerada uma manifestação ser: Todo real: as partes são (quão grande, quanto); Qualidade das cinco categorias distintas, não necessariamente (tipo de coisa); Relação (com o fundamentais: Personalidade separáveis; Todo lógico: suas que se relaciona); Lugar (onde); (categoria fundamental); partes são distintas pela razão; Tempo (quando); Posição (qual a Matéria (material); Energia Todo moral: as partes, postura); Estado ou condição (em quais circunstâncias); Ação (quão (ação, função); Espaço (onde); distintas, são unidas por um Tempo (quando). As categorias fim comum, devendo ser: ativo, qual o fazer); Paixão (quão passivo, qual o sofrer). Já também indicam qual Adequada ou completa: a soma característica essencial do das partes deve ser igual ao demonstra a escolha de uma conceito permite a classificação todo: Irredutível: cada parte da característica essencial do dos assuntos de um documento divisão não deve conter outra conceito para que o objeto seja inserido em determinada em determinada categoria. parte: Fundada no mesmo **CLASSIFICATION** princípio de divisão; categoria. Ordenada hierarquicamente; RESEARCH GROUP (CRG) **KANT** O CRG parte da Classificação de Breve. Apresenta as categorias como Dois Pontos de Ranganathan e A divisão também pode ser conceitos puros do entendimento indica as categorias: Todo/Tipo; dicotômica, na qual o termo que os seres humanos possuem a Partes; Material/Constituinte; tem expressão positiva em sua priori. Divide as classes em três Propriedade; Processos; Operações; classe e negativa nas demais categorias, sendo as duas primeiras representadas por uma Agentes; Lugar; Tempo; Formas de classes. Apresentação. dicotomia e a terceira resultado da **REALE VICKERY** Categoriza o Direito ligação da segunda com a Membro do CRG, Vickery primeira de sua classe. Divide as observando como afirma que a categorização está características essenciais o categorias em: Quantidade presente tanto na indexação conteúdo da relação jurídica de (unidade, pluralidade e quanto na recuperação da um lado (o Direito Público totalidade); Qualidade (realidade, informação. visa ao interesse geral e o negação e limitação); Relação **DAHLBERG** Direito Privado visa ao (inerência e subsistência, causa e Na Teoria do Conceito, os interesse particular) e a forma efeito e ação recíproca entre o objetos são expressos por da relação jurídica de outro (de agente e o paciente); Modalidade coordenação, como no caso de conceitos cujas características (possibilidade – impossibilidade, traduzem os atributos desses compra e venda; de existência - não-existência e objetos que podem ser subordinação, como no caso da necessidade -contingência). categorizados como: Matéria convocação para eleição). WITTGENSTEIN (substância); Qualidade; As características presentes nos Ouantidade (extensão): Relação: membros de uma classe não são Processo (atividade); Modo de sempre as mesmas, mas se cruzam ser; Passividade; Posição; e se sobrepõem umas às outras, localização (lugar); resultando em uma complicada

Continua

rede de parentesco.

Continuação

CATEGORIZAÇÃO				
Filosofia	Ciência da Informação	Discurso Jurídico		
ROSCH	IYER	FERRAZ JUNIOR		
Na teoria dos protótipos, o	Afirma que as semelhanças	Considera que a categorização		
elemento que melhor representa a	dos conceitos inferem	da área jurídica deve garantir a		
categoria é considerado o seu	informações às palavras e à	certeza e segurança da		
protótipo; os demais elementos	categorização e estabelecem	decidibilidade de		
formam uma estrutura gradual, na	relações uns com os outros	conflitos.Tempo.		
qual os elementos são mais ou	dentro de uma categoria			
menos típicos da categoria de				
acordo com as características				
compartilhadas com o protótipo				
da categoria.				

Observa-se que o elemento central que permite a inclusão de um determinado objeto em uma determinada categoria é denominado de *característica do conceito*, o qual estabelece as relações entre os demais objetos dessa categoria.

Quando a teoria clássica não dá conta da categorização, pois o mesmo termo é parte de classes diferentes dentro de uma classe maior, muito comum nas ciências humanas, a teoria de semelhança de famílias de Wittgenstein e a teoria dos protótipos de Rosch podem ser utilizadas como alternativa na elaboração de linguagens documentárias das ciências humanas, como é o caso da linguagem jurídica, na qual a sobreposição de características dos conceitos é uma dificuldade encontrada para a divisão da área em categorias rígidas e previamente estabelecidas, já que alguns conceitos são aplicáveis a mais de uma categoria.

A teoria de Wittgenstein não estabelece limites rígidos entre as categorias, mas considera as semelhanças de família compartilhadas entre seus membros, como é o caso do termo *Embargos de Declaração*, termo que compartilha de semelhanças familiares das categorias: *Processo Civil, Processo Penal* e *Processo Trabalhista*.

A teoria de Rosch estabelece que um membro que seja o mais representativo de determinada categoria será considerado o protótipo da categoria. Os demais membros da categoria precisam compartilhar das mesmas características do protótipo da categoria, que é composta por uma estrutura gradual, na qual seus membros são considerados mais ou menos representativos da categoria de acordo com o número de características semelhantes àquelas presentes em seu protótipo. O membro que compartilhar de mais características com o protótipo será mais representativo nesta categoria e o membro que compartilhar de menos características com o protótipo será menos representativo nesta categoria.

Utilizando como exemplo o termo *Fiança*, quando ele estiver relacionado ao *Direito Civil* compartilhará de mais características do protótipo desta categoria e de menos características do protótipo da categoria *Processo Penal*, e acontecerá o inverso quando ele estiver relacionado ao *Processo Penal*. Desta forma o termo será categorizado de acordo com as características que envolvem o contexto no qual ele se insere. O que determina o assunto do documento é o conjunto de descritores da linguagem documentária utilizada na representação da informação, ou seja, a informação documentária.

7. CONSIDERAÇÕES FINAIS

Cabe à Ciência da Informação estabelecer procedimentos para o tratamento, organização e recuperação da informação, mais especificamente pela elaboração de uma informação documentária que possibilite ao usuário o acesso às respostas para solucionar questões pertinentes à sua pesquisa.

A categorização tem sido utilizada pela Ciência da Informação na construção de instrumentos utilizados para a organização da informação nos sistemas de informação, com o objetivo de representar e recuperar as informações produzidas e registradas pela humanidade ao longo do processo de construção do conhecimento. Especificamente, ela é aplicada na estruturação dos descritores das linguagens documentárias, os quais representam os assuntos dos documentos armazenados no sistema para fins de recuperação da informação, conforme demonstram o Vocabulário Controlado Básico do Senado Federal e o Vocabulário Controlado da Universidade de São Paulo.

A categorização conceitual aliada aos procedimentos terminológicos mostra-se uma importante alternativa tanto para a construção quanto para o uso de linguagens documentárias, seja no momento da representação ou no momento da recuperação da informação. Isso porque são os procedimentos terminológicos que possibilitam a definição dos conceitos e o estabelecimento das relações entre eles, representados por termos, dentro de um domínio específico do conhecimento. A terminologia de uma área de especialidade — no caso deste trabalho, a terminologia jurídica — deve ser utilizada como fonte para a elaboração de uma linguagem documentária, e os especialistas da área podem contribuir sobremaneira para a compreensão da estrutura do domínio.

A partir da fundamentação teórica, identificamos como princípios para a categorização conceitual da informação jurídica no interior dos instrumentos de indexação:

- A definição dos conceitos e suas designações, os termos, do domínio do Direito, a partir de obras de referência e estudos de especialistas consagrados na área, como, por exemplo, Alves, Reale e Ferraz Junior conforme a norma ISO 1087, a qual estabelece a forma de organizarmos o "sistema conceitual", partindo do "campo conceitual" e utilizando como critério de subdivisão do sistema conceitual as "características" utilizadas para descrever os conceitos;
- A análise dos enunciados destas definições, de maneira a identificar as características essenciais do conceito, a partir da Teoria do Conceito de Dahlberg;
- A identificação dos relacionamentos entre as características essenciais que estabelecem tanto a "semelhança de famílias", proposta por Wittgenstein, quanto o "protótipo", proposto por Rosch, que vão indicar as categorias conceituais a que pertencem cada termo;
- A estruturação das categorias conceituais, procurando-se também estabelecer as suas definições e relacionamentos enquanto designação de conceitos mais gerais no interior do domínio;
- A delimitação das categorias conceituais enquanto descritores de uma linguagem documentária para representação e recuperação da informação jurídica.

A partir dessas considerações é necessário testar estes princípios, o que pode ser feito utilizando-se a lista de assuntos do *IusData*, pois é um banco de dados que não só abrange a já consolidada terminologia jurídica, mas também pode ser considerado uma fonte para coleta dos novos termos do domínio.

8. REFERÊNCIAS

ALVES, Alaor Caffé. **Lógica**: pensamento formal e argumentação: elementos para o discurso jurídico. 4. ed. São Paulo: Quartier Latin, 2005. 405 p.

ARISTÓTELES. Categorias. In: ARISTÓTELES. **Órganon**. Tradução Edson Bini. 2. ed. Bauru, SP: EDIPRO, 2010. p. 39-80.

BARBOSA, Alice. Classificações facetadas. **Ciência da Informação**, Rio de Janeiro, v. 1, n. 2, p. 73-81, 1972.

BARROS, Lídia Almeida. **Curso básico de terminologia**. São Paulo: EDUSP, 2004. (Acadêmica; 54).

BASE de dados Vocabulário Controlado Básico – VCB e Autorias Padronizadas – AUTR. **Portal O Senado**. Disponível em: http://www.senado.gov.br/senado/biblioteca/pesquisa/pesquisa.asp>. Acesso em: 04 fev. 2012.

BITTAR, Eduardo Carlos Bianca. Linguagem jurídica. São Paulo: Saraiva, 2001.

CABRÉ, M. T. Terminología y documentación. In: CABRÉ, M. T. La terminología: representación y comunicacion: elementos para uma teoria de base comunicativa y otros artículos. Barcelona: IULA/Universitat Pompeu Fabra, 1999. p. 231-247.

CINTRA, Anna Maria Marques et al. Linguagens documentárias e terminologia. In: ALVES, I. M. (Org.). A constituição da normalização terminológica no Brasil. São Paulo: FFLCH / CITRAT, 1996. p. 17-22 (Cadernos de Terminologia).

CINTRA, Anna Maria Marques et al. **Para entender as linguagens documentárias**. 2. ed. rev. e ampl. São Paulo: Polis, 2002. 96 p. (Coleção Palavra-Chave, 4). DAHLBERG, Ingetraut. Teoria do conceito. Tradução para o português do Prof. Astério Tavares Campos. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.

DINIZ, Maria Helena. **Dicionário jurídico**. 3. ed. rev., atual. e aum. São Paulo: Saraiva, 2008. 4 v.

FERRAZ JUNIOR, Tercio Sampaio. **Introdução ao estudo do direito**: técnica, decisão, dominação. 5. ed. rev. e ampl. São Paulo: Atlas, 2007.

GRUPO de Estudo do Tesauro da Rede Virtual de Bibliotecas do Congresso Nacional (RVBI). **Elaboração do tesauro da Rede Virtual de Bibliotecas – Congresso Nacional – RVBI**: um esboço metodológico. Brasília: Senado Federal, Secretaria de Biblioteca, 2005. 15 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 1087-1**: Terminology work: Vocabulary: Part 1: theory and application. 2000.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 704**: Terminology work: Principles and methods. 2000.

IYER, Hemalata. Cognition and categories. In: IYER, Hemalata. **Classificatory structures**: concepts, relations and representation. Frankfurt/Main: Indeks Verlag, 1995. 232 p., p. 40-57. (Textbooks for Knowledge Organization, v. 2).

KANT, Immanuel. **Crítica da razão pura**. Tradução de Manuela Pinto dos Santos e Alexandre Fradique Morujão. Introdução e notas de Alexandre Fradique Morujão. 5. ed. Lisboa: Fundação Calouste Gulbenkian, 2001. 680 p.

KOBASHI, Nair Yumiko. Análise documentária e representação da informação. INFORMARE – Cadernos do Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, v. 2, n. 2, p. 5-27, jul./dez. 1996.

LARA, Marilda Lopes Ginez de. Linguagens documentárias, instrumentos de mediação e comunicação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 26, n. 1/2, jan./jun. 1993.

LIMA, Gercina Ângela Borém de Oliveira. **Mapa hipertextual (MHTX)**: um modelo para organização hipertextual de documentos. 2004. 199 f. Tese (Doutorado)—Escola de Ciência da Informação da Universidade Federal de Minas Gerais, Minas Gerais, 2004.

LIMA, Gercina Ângela Borém de Oliveira. Modelos de categorização: apresentando o modelo clássico e o modelo de protótipos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 2, p. 108-122, maio/ago. 2010.

LIMA, Vânia Mara Alves. A informação documentária: codificação e decodificação. **Transinformação**, Campinas, v. 19, n. 2, p. 119-127, maio/ago. 2007.

LIMA, Vânia Mara Alves; KOBASHI, Nair Yumiko; IMPERATRIZ, Inês Maria de M. Vocabulário Controlado USP: desenvolvimento, implantação e gerenciamento. In: INTEGRAR: CONGRESSO INTERNACIONAL DE ARQUIVOS, BIBLIOTECAS, CENTROS DE DOCUMENTAÇÃO E MUSEUS, 1., 2002, São Paulo. **Textos...** São Paulo: Imprensa Oficial do Estado, 2002. p. 225-235.

LODDO, Maria Eliza Nogueira. **Vocabulário Controlado Básico**: VCB: regras, convenções e instruções para sua utilização. Brasília: Senado Federal, Subsecretaria de Biblioteca, 1991. 29 p.

MACIEL, Anna Maria Becker. Terminologia, linguagem de especialidade e dicionários. In: KRIEGER, Maria da Graça; MACIEL, Anna Maria Becker (Orgs.). **Temas de Terminologia**. Porto Alegre; São Paulo: UFRGS; Humanitas, 2001. p. 39-46.

MACROESTRUTURA com as relações Lógico-Semânticas com as áreas e terminologia em seus diferentes níveis. In: **Vocabulário Controlado USP**. Disponível em: http://143.107.73.99/Vocab/Sibix652.dll/Mac. Acesso em: 22 fev. 2012.

NARUKAWA, Cristina Miyuki et al. Metodologia de trabalho para atualização do Vocabulário Controlado da Universidade de São Paulo (USP) da área jurídica. In: Seminário Internacional de Bibliotecas Universitárias, 17., 2012, Gramado. **Anais...** Porto Alegre: Universidade Federal do Rio Grande do Sul, 2012. Disponível em: http://www.snbu2012.com.br/anais/pdf/4RHH.pdf. Acesso em: 30 mar. 2013.

NEVES, Emerson Douglas. **Avaliação da recuperação da informação jurídica a partir do Vocabulário Controlado Básico – VCB**. 2008. 12 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia)—Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

RANGANATHAN, Shiyali Ramamrita. Facet analysis: fundamental categories. In: RANGANATHAN, Shiyali Ramamrita. **Elements of library classification**. London: Association of Assistant Librarians, 1959. 108 p., p. 67-84.

REALE, Miguel. **Lições preliminares de direito**. 7. ed. rev. São Paulo: Saraiva, 1980. **THES**. Disponível em: http://webthes.senado.gov.br/thes/navthes.htm>. Acesso em: 04 fev. 2012.

VICKERY **Classificação e indexação nas ciências**. Tradução de Maria Christina Girão Pirolla. Rio de Janeiro: BNG/Brasilart, 1980. (Coleção Biblioteconomia, documentação, ciência da informação).

VOCABULÁRIO Controlado USP. Disponível em: http://143.107.73.99/Vocab/Sibix652.dll>. Acesso em: 22 fev. 2012.

WITTIGENSTEIN, Ludwig. **Tratado lógico-filosófico. Investigações filosóficas**. 2. ed. rev. Tradução e prefácio de M. S. Lourenço. Lisboa: Fundação Calouste Gulbenkian, 1995. 611 p.