



Ambiguidade e o contexto na representação de informações em domínios de especialidade

Benildes Coura Moreira dos Santos Maculan^I

<http://orcid.org/0000-0003-0735-3856>

^I Universidade Federal de Minas Gerais, MG, Brasil.

Doutora em Ciência da Informação.

Professora do PPG-GOC/Escola de Ciência da Informação.

<http://dx.doi.org/10.1590/1981-5344/4300>

Apresenta uma visão geral das questões que permeiam as investigações desenvolvidas sob minha orientação na linha de pesquisa Arquitetura e Organização do Conhecimento, do Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPG-GOC), que tem a Ciência da Informação como área de concentração. Retrata alguns elementos da ambiguidade da linguagem natural que levam a problemas na recuperação da informação. Discute sobre o contexto como componente para a minimização desses problemas, que deve ser considerado na construção de abstrações e modelos da realidade, sobretudo no desenvolvimento de Sistemas de Organização do Conhecimento (SOC) para uso computacional. Trata os SOC como uma linguagem de comunicação composta de um sistema de conceitos que sintetizam e estruturam o conhecimento de um dado domínio. Traz alguns estudos sobre desenvolvimento de sistemas inteligentes que aplicam o contexto como mecanismo mediador que aprimora a comunicação homem-computador. Discorre sobre a complexidade na representação do conhecimento. Como exemplo, descreve a complexidade que envolve a representação no campo do Direito, da informação jurídica. Estabelece que todo e qualquer ato comunicativo não pode ser analisado de modo isolado, sendo um processo envolto em

complexidade, o que exige a consideração de distintos aspectos.

Palavras-chave: *Representação do conhecimento. Sistema de Organização do Conhecimento. Ambiguidade e contexto. Informação jurídica.*

Ambiguity and the Context on Representation of Information in Speciality Areas

This paper presents an overview of the questions that underlie the research carried out under my guidance in the Architecture and Organization of Knowledge research line, of the Graduate Program in Knowledge Management & Organization (PPG-GOC), in the Information Science research area. It shows some elements of natural language ambiguity that lead to problems in information retrieval. It also discusses context as a component for the minimization of these problems, which should be considered in the construction of abstractions and models of reality, especially in the development of knowledge organization systems (SOC) for computational use. It treats SOC as a communication language composed of a system of concepts that synthesize and structure the knowledge of a given domain. It brings some studies about intelligent systems development that apply the context as a mediating mechanism that enhances human-computer communication. It discusses the complexity of knowledge representation. As an example, it describes the complexity of representation of Law documentation – legal information. It establishes that any and every communicative act cannot be analyzed in isolation, being a process involved in complexity, which requires the consideration of different aspects.

Key-words: *Knowledge representation. Knowledge Organization System. Ambiguity and context. Legal information.*

Recebido em 27.02.2020 Aceito em 27.02.2020

1 Introdução

A humanidade sempre foi dependente de alguma forma de comunicação como um meio de sobrevivência e de desenvolvimento da espécie. Passamos pela comunicação por gestos, grunhidos, repetição de processos, fala e escrita, com os primeiros desenhos e hieróglifos. Vieram, então, a tipografia e a prensa de tipos móveis, trazendo a multiplicação das impressões, rumo à ampliação da comunicação social e científica. E, no processo de comunicação, a ideia do contexto é elemento que tem papel importante, impactando positivamente o compartilhamento de recurso informacional.

A linguagem natural é ambígua, pois uma mesma frase pode ter vários significados, dependendo se usamos, por exemplo, a forma denotativa ou conotativa do sentido das palavras. Em geral, quando um certo indivíduo comunica uma frase a outro, o primeiro faz presunções sobre o contexto, as crenças e o conhecimento do segundo sobre aquilo que vai comunicar. A comunicação verbal, assim, implica no uso de uma quantidade mínima de informações preexistentes, que sejam essenciais para o entendimento do receptor de tal mensagem. É como se as pessoas utilizassem um algoritmo-filtro, para que o interlocutor possa verbalizar o mínimo possível, dada a circunstância e o contexto do receptor.

A comunicação é um processo em que, a princípio, dois indivíduos participam de uma troca física ou verbal de duas vias – interlocutor e receptor – tendo em vista alcançar algum entendimento mútuo. Nessa troca, os participantes permutam informações, sentimentos, valores, notícias ou ideias, a partir de escolhas feitas pelo emissor de uma linguagem que possa ser corretamente interpretada pelo receptor. Com isso, evita-se a ambiguidade na comunicação.

Podemos afirmar que a ambiguidade é um tipo de incerteza de significado na forma de um atributo semântico, cujo significado não pode ser resolvido de acordo com uma regra ou processo com um único sentido possível. Cintra *et al.* (2002, p. 55) afirmam que a ambiguidade pode ser “entendida como a possibilidade de uma comunicação linguística se prestar a mais de uma interpretação”. É o contexto que desempenha papel crucial nessa comunicação, limitando os múltiplos sentidos possíveis.

De um modo geral, toda forma de ambiguidade tem algum grau de questão semântica. Existem tipos básicos de ambiguidades com as quais a comunicação tem que lidar: fonológica, lexical, sintática e semântica (LEANDRO FERREIRA, 2000). Para a autora, a primeira caracteriza-se pela semelhança sonora entre duas palavras ou grupos de palavras, ocorrendo o fenômeno das palavras homófonas, que possuem o mesmo som de pronúncia, mas que podem ter grafias ou significados diferentes. Por exemplo, temos duas palavras com o mesmo som – conserto (arrumar

algo) e concerto (show musical) – que possuem grafias e significados diferentes.

Na segunda, ambiguidade lexical, uma palavra pode assumir dois ou mais significados, porque está na forma denotativa ou conotativa (LEANDRO FERREIRA, 2000), como, por exemplo: “*O prefeito pediu um prato ao garçom*”, quando a palavra “prato” pode ser interpretada como o objeto (no qual se serve a comida) ou como a própria comida (um tipo de refeição).

A terceira, a ambiguidade sintática, ou estrutural, não se refere à diversidade de significados das palavras (idem). A autora afirma que, em geral, caracteriza-se pelo duplo sentido causado pela má organização estrutural de uma frase. Basicamente, ocorre devido a três distintas situações: uso incorreto do pronome relativo, por exemplo, em “*Maria pegou o estojo da gargantilha que estava em cima da cômoda*” (o que estava em cima da cômoda? O estojo ou a gargantilha?); má colocação de termos, pronomes ou frases, por exemplo, em “*a mulher reconheceu a menina quando entrou em seu quarto*” (o quarto é da mulher ou da menina?); colocação inadequada do adjunto adverbial, por exemplo, em “*alunos que estudam são em geral mais inteligentes*” (os alunos são em geral mais inteligentes porque estudam ou estudam porque em geral são mais inteligentes?).

Por fim, ambiguidade semântica refere-se aos significados de uma palavra, pelo fenômeno da polissemia, quando uma palavra varia o sentido conforme determinado contexto (LEANDRO FERREIRA, 2000). Por exemplo, se falo “*Roberto está no banco*” não fica explícito se a palavra “banco” se refere a um tipo de assento ou a uma instituição financeira.

Quando trabalhamos na construção de sistemas conceituais para domínios de conhecimento específicos, Tennis (2012, p. 12) estabelece que os modeladores devem empregar a análise de domínio instrumental, elaborando definições e determinando o escopo e alcance, período e propósito, uma vez que “a nossa visão de um domínio muda e os domínios mudam”. Nos domínios de especialidade, um conceito é considerado uma unidade semântica¹ que terá uma designação verbal (o termo) que o representa. Esse termo manterá relação de equivalência com outros termos que carregam o mesmo significado ou com algum grau de semelhança de significado. Esse conjunto de termos equivalentes poderá estar representado por um anel de sinônimos². Poderíamos juntar, por exemplo, os termos morto e falecido ou represa e barragem como

¹ Dahlberg (1978) denomina de unidade de conhecimento e Ranganathan (1967), de unidade de pensamento, mas aqui designo como unidade semântica por se referir a um dado contexto ou domínio.

² Um tipo de SOC, que agrupa um conjunto de sinônimos ou equivalentes.

equivalentes semânticos. Contudo, devemos entender que não há sinônimos absolutos e, em alguma medida, esses pares de palavras poderiam ter sentidos diferentes, a depender do contexto de uso dessa linguagem. Ducrot (1977, p. 116) afirma que “não existe enunciado que ocorra fora de um contexto, por mais reduzido que possa ser”, sendo a descrição semântica de um vocabulário o conjunto de maneiras que podem expressar este conhecimento em relação aos seus significados.

Devido a todas as possibilidades de ocorrência de ambiguidade na linguagem natural, a consideração do contexto é essencial para minimizar tais acontecimentos. Considero que nenhuma teoria semântica (estudo da comunicação; do significado) para o processamento de informação pode ignorar o papel do contexto na minimização de problemas de ambiguidade expressos na linguagem natural, pois é o contexto que evidencia o ambiente discursivo, nos quais as expressões são usadas e interpretadas. Nessa perspectiva, este trabalho apresenta uma visão geral das questões que permeiam e perpassam as investigações que desenvolvo na linha de pesquisa Arquitetura e Organização do Conhecimento (AOC), do Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPG-GOC), que tem a Ciência da Informação como área de concentração. É um estudo exploratório e descritivo, e, após esta breve introdução, este artigo foi organizado como segue: na seção 2 descrevo diferentes abordagens para o conceito de “contexto” e apresento algumas aplicações do contexto na modelagem de sistemas inteligentes; na seção 3 discorro sobre a complexidade das representações, trazendo, sobretudo, insumos sobre a formação de conceitos de Dahlberg (2006), Hjørland (2008; 2009) e resultados de minha pesquisa doutoral (MACULAN, 2015); também apresento, nesta seção, como exemplo, a complexidade na representação de informações jurídicas; por fim, na seção 4 trago as considerações finais.

2 O contexto no entendimento do sentido semântico da informação

Para a nossa compreensão de um texto, precisamos nos basear na gramática do idioma textual em questão, nas regras do discurso e no nosso conhecimento do mundo em geral, mas isso ainda será apenas parte do que necessitamos para atingir a compreensão do texto ou da informação comunicada em um documento³. É essencial também

³ Neste artigo, documento é entendido em sentido mais amplo, como o conjunto de objetos portadores de “informações selecionadas e organizadas materialmente em um sistema, cuja significação proposta objetiva orientar processos de significação pelo público” (ORTEGA, 2016, p. 58), que não se limitam a textos, pois podem ser imagens, por exemplo.

entendermos o contexto comunicativo do conteúdo daquele texto ou documento.

No dicionário, encontramos a informação de que o termo “contexto” tem origem no latim – *contextus* – significando “união de vários elementos” (FERREIRA, 2009). No processo comunicacional, o contexto pode ser analisado a partir de diferentes abordagens, entre as quais destaque: física, funcional, social, cultural, situacional e semântica.

O contexto físico da comunicação, segundo Santos García (2012), tem relação com os elementos tangíveis e perceptíveis, por meio dos sentidos, presentes durante um evento comunicativo. O autor complementa dizendo que esses elementos podem ser as formas orais e não orais empregadas na comunicação da mensagem, ou elementos estáticos, que também possuem carga comunicativa, tais como cartazes, arranjos de mesas ou letreiros.

O contexto funcional, tal como a teoria do Aprendizado Situado, destaca a importância do contexto durante o aprendizado, e tem origem em programas militares nos Estados Unidos realizados para o ensino de habilidades basilares e leitura (STICHT, 1975; 1976). Segundo o autor, essa abordagem pressupõe um modelo cognitivo que foca em três elementos básicos: o conhecimento (de longo prazo) do aprendiz, as diferentes habilidades de processamento de informações (linguagem, resolução de problemas), e a capacidade de exposição de argumentos que apresentam informações.

O contexto social pode ser considerado quando tratamos de circunstâncias ou ocorrências de comunicação interpessoal e individual, pois ele trata da relação que um indivíduo tem com outras pessoas. Estima-se que os sujeitos, por exemplo, modificam o comportamento devido à posição social ou hierárquica da pessoa com a qual estejam em situação comunicativa. E, da mesma maneira, os sujeitos podem desempenhar diferentes papéis na sociedade, a depender das instâncias de socialização (trabalho, lazer, família) nas quais estão circunscritos.

Por sua vez, o contexto cultural é o conjunto de valores, atitudes, crenças e visões, e está integrado dentro dos princípios fundamentais de cada indivíduo. Essas características influenciam a maneira como pensamos, nossas preferências, necessidades, comportamentos, tal qual a forma como enxergamos o mundo à nossa volta, também o compartilhamento dos significados. O contexto oferece ajuda para selecionarmos comportamentos apropriados para cada situação e a determinarmos os conteúdos explícito e implícito de nossos pensamentos. Enfim, o contexto engloba todas as circunstâncias (configuração, posições de valor, adequações comportamentais, público, lugar de referência) que envolvem uma mensagem.

O contexto situacional – ou pragmático – engloba analisar um conjunto de circunstâncias ou fatos que cercam um dado evento ou situação, criando um significado para a mensagem comunicada. Dessa maneira, a identificação do contexto auxilia na avaliação da informação e terá papel crucial no sucesso, ou não, do processo comunicativo, que também envolve o conhecimento prévio do assunto tratado, que deve ser compartilhado com os interlocutores.

O contexto semântico, por fim, refere-se aos significados das palavras e suas correlações. No âmbito da educação, o contexto semântico é considerado como *priming* semântico, que se refere à capacidade do sujeito em identificar palavras, signos, objetos ou imagens, após uma experiência recente com tais elementos (SQUIRE; KANDEL, 2003). Segundo Coan (2000, p. 25), contexto semântico “é o contexto de ordem linguística”, que não deve ser confundido com o contexto situacional (ou pragmático), que está sujeito às circunstâncias. Sobre isso, Van Dijk (1999, n.p.) afirma que a “compreensão pragmática se dá paralelamente ao processo relativo de compreensão semântica, no qual os discursos e conhecimentos prévios do contexto são importantes para a interpretação [...]” de uma dada preposição. O contexto semântico é, assim, definido no componente semântico do discurso, como elemento linguístico.

A partir do exposto, é possível fazer inferências sobre o papel do contexto na construção de representações do conhecimento de um domínio, considerados no sentido de instrumentos terminológicos comunicativos que quero apontar aqui. Levar em conta o contexto significa que, entre outros elementos, consideramos que diferentes sujeitos, em situações comunicativas, podem ter concepções distintas sobre determinados conceitos, fazendo-se necessário controlar a terminologia adotada na linguagem construída.

Considero que a determinação do contexto nos estudos sobre construção de linguagens de representação na Ciência da Informação desempenha papel essencial, como ocorre em disciplinas tais como semântica, pragmática, psicologia cognitiva, e até mesmo na Inteligência Artificial (IA). A próxima seção trata um pouco sobre isso, trazendo uma breve descrição de trabalhos que usaram o contexto na modelagem de sistemas inteligentes de informação.

2.1 Aplicações do contexto na modelagem de sistemas inteligentes

A ideia da consideração do contexto na criação de linguagens de representação no campo da IA contraria a perspectiva da semântica realista, advinda dos princípios aristotélicos, que utiliza uma abordagem que declara haver uma relação direta entre a linguagem e os fenômenos e

objetos do mundo real (BREZOLIN, 2017). Aqui considero que quando se trabalha em domínios de especialidade, sobretudo não científicos, essa abordagem não, necessariamente, precisa ser adotada. Mas continuo ainda acreditando que, em algumas situações comunicativas, a abordagem realista pode ser aplicável e até desejável. No Processamento da Linguagem Natural (PNL), por exemplo, quando queremos estruturar informações declaradas em linguagem natural, o uso da abordagem realista talvez não traga um resultado satisfatório. O contexto, quando considerado e adotado em um sistema de informação inteligente, é um mecanismo mediador entre o vocabulário dos usuários e a linguagem desse sistema (BREZOLIN, 2017). O PNL está estritamente ligado a tecnologias de inteligência artificial como a de aprendizado de máquina (*machine learning*), que se refere à promoção da capacidade de reconhecimento de padrões de dados pelos computadores.

Dey e Abowd (2000, n.p.) afirmam que melhorando “o acesso do computador ao contexto, aumentamos a riqueza da comunicação na interação humano-computador”. Para os autores, no âmbito da computação,

Contexto é qualquer informação que caracteriza a situação de uma entidade, onde uma entidade é uma pessoa, lugar ou objeto considerados relevantes para a interação entre um usuário e uma aplicação, incluindo o próprio usuário e a aplicação. O contexto é tipicamente a localização, a identidade, e o estado das pessoas, grupos e objetos físicos e computacionais (DEY; ABOWD, 2000, n.p.).

Os autores também orientam sobre uma categorização de tipos de contexto, tendo em vista que definições prévias de contexto facilitarão evidenciar onde ele será mais útil no aplicativo. Os aplicativos com reconhecimento de contexto analisam quem [who], onde [where], quando [when] e o que é [what]”, sendo o “o que” aquilo que o usuário está fazendo (DEY; ABOWD, 2000, n.p.). Quanto ao elemento “por que”, os autores explicam que é o designer que determina por que uma dada “situação está ocorrendo e usa isso para codificar alguma ação no aplicativo” (idem). Os autores explicam que existem tipos de contexto que são mais relevantes que outros, tais como os que determinam a localização (objetos, pessoas e atividades próximas), identidade, atividade e tempo, que tanto respondem às questões onde, quem, o que (ação) e quando, respectivamente, quanto “atuam como índices em outras fontes de informação contextual” para identificar contextos secundários (idem).

Bouquet *et al.* (2003) discorrem sobre o uso do contexto e destacam diferentes estudos no âmbito da IA: (a) processamento da linguagem natural: resolver ambiguidades e atribuir uma interpretação para afirmação (enunciado predicativo); (b) recuperação da informação: reunir as consultas dos usuários; (c) sistemas multiagentes/agentes autônomos: permitir formalmente flexibilizar o projeto

no sentido de estabelecer comunicação entre os diferentes agentes; (d) interação homem-máquina: projetar aplicações e interfaces personalizáveis.

Bouquet *et al.* (2003) analisaram diversos estudos sobre o uso do contexto numa área da IA denominada representação do conhecimento e raciocínio (em inglês, KRR), que têm como objetivo conceber linguagens para representar agentes inteligentes e processos de raciocínio para a criação de novos conhecimentos (inferências). Segundo os autores, foi a partir do final dos anos 1980 e início dos anos 1990 que o contexto em IA se tornou mais efetivamente empregado, quando “Giunchiglia (1993) e McCarthy (1987; 1993) começaram a trabalhar em uma teoria formal do contexto, cujo objetivo era explicar as propriedades do contexto e do raciocínio contextual de maneira sistemática” (BOUQUET *et al.*, 2003, p. 457). A partir de então, os autores destacam que os estudos se proliferaram, descrevendo uma síntese sobre alguns desses estudos e apresentando a própria proposta de uso de contexto.

O estudo de Bouquet *et al.* (2003) se concentrou em demonstrar duas teorias do contexto para uso em KRR: a teoria de dividir-e-conquistar, com o contexto servindo de norte para criar um modelo de mundo mais simples e subdividido, e a teoria de compor-e-conquistar, quando o contexto é empregado para um modelo local, mas mantendo a interoperabilidade, ou seja, uma rede de relações com outros modelos locais. Para analisar as teorias, os autores usaram os seguintes exemplos: problema da generalidade, formalização de atitudes proposicionais e integração de conhecimento e dados. Os autores concluíram que, para cada tipo de questão tratada no KRR, o contexto será empregado para soluções conceitualmente diferentes, mesmo considerando os mesmos problemas. Sendo assim, vai depender do propósito do sistema construído. Os autores também identificaram que a “noção de contexto usada em disciplinas como lógica filosófica, psicologia cognitiva e linguística” (BOUQUET *et al.*, 2003, p. 484) são diferentes da concepção aplicada ao KRR, mas que este é ainda um tema que deve ser mais bem estudado. Segundo os autores, em Benerecetti, Bouquet e Ghidini (2000), nota-se uma primeira tentativa em evidenciar tais diferenças. Bouquet *et al.* (2003, p. 458) ainda declaram que pesquisadores interessados nessa temática devem procurar explorar os resultados dos trabalhos de “Akman e Surav (1996), para modelos formais de contexto, e Brézillon (1999), para aplicações de contexto”.

No campo dos estudos em BDI (Belief-Desire-Intention), que é um modelo baseado na teoria filosófica do raciocínio prático (*Theory of Human Practical Reasoning*), e que tenta retratar as atitudes humanas de crenças (contexto), desejos (conjunto de estados, selecionados do contexto, que se deseja atingir) e intenções (sequência de ações a executar para atingir os objetivos a atingir) para variadas finalidades, há o trabalho de Gelaim (2016). O autor apresenta um modelo BDI baseado em um Sistema Multi-Contexto, que integra um modelo computacional de emoções, no qual as

crenças são subdivididas em distintos graus de certeza, denominadas como especiais (julgamentos), dando ao agente a capacidade de intermediar a relação entre confiança e emoções. Gelaim (2016) trabalhou na definição dos seguintes contextos, cuja integração é realizada por um conjunto de regras de ponte (ligação): contexto de comunicação (interface entre contextos), contexto de emoções (manipulação e representação de emoções), contexto de crenças (conhecimento sobre o mundo, em gradações de certeza de sua credibilidade), contexto de confiança (grau de confiança em outras entidades, tais como, em relação aos contextos de crenças e de confiança), contexto de desejos (objetivos gerais, classificados como positivos ou negativos), contexto de intenções (sequências de ações geradas com base em suas crenças e desejos) e contexto de planejamento (sintetização dos planos que satisfazem aos desejos, com base em lógica de primeira ordem restrita a cláusulas de Horn). No trabalho de Gelaim *et al.* (2018), os autores ampliam o sistema original de Gelaim (2016), generalizando o sistema para ser capaz de criar novos contextos.

Brezolin (2017) apresenta um modelo conceitual para arquitetura de BDI, considerando aspectos do fenômeno cognitivo, entre os níveis simbólico e perceptual. Como pressuposto, o autor considera que o reconhecimento de objetos pode ser realizado por um processo de formação de conceitos, com a abordagem do paradigma dos espaços conceituais⁴, com base no trabalho de Gärdenfors (2000). Os autores desenvolveram uma proposta conceitual para agentes BDI para ser colocada em uma aplicação voltada ao auxílio de pessoas cegas, para as quais a aprendizagem baseada na observação é limitada. Para pessoas cegas, a construção de um mapa mental sobre determinado ambiente ocorre a partir, muitas vezes, da memória e de descrições feitas por outras pessoas sobre localização, distância e cor, por exemplo. A implementação do protótipo para a captação de imagens foi realizada em um dispositivo móvel (celular), em um sistema dividido em dois módulos: o módulo usuário (captura de imagens) e o módulo do sistema inteligente (classificação dos objetos do contexto). Como limitação, o autor verificou que o “sistema, por vezes, não acompanha a velocidade com que o usuário realiza seus movimentos, e ocorreram casos em que as mensagens que estavam sendo lidas já não refletiam mais o posicionamento do usuário” (BREZOLIN, 2017, p. 90), ainda que tenha sido avaliado positivamente pelos usuários. Para o autor, essa limitação demonstra a necessidade de continuidade dos estudos tendo em vista estabelecer recursos para facilitar a mobilidade de pessoas cegas.

⁴ Gärdenfors (2000) destaca que o espaço conceitual é um agrupamento de conceitos relacionados, usado para organizar experiências sensoriais, imaginativas ou de memória em diferentes contextos.

Também tratando de BDI, Arnhold (2018) desenvolveu um *framework*, estabelecendo procedimentos de operacionalização, que foi implementado com base em Java Virtual Machine. O *framework* foi “projetado para a atuação de um único agente”, sendo que os “sensores e atuadores estão ligados ao agente, não sendo definida uma interface para ambientes, o que facilita a integração do agente com outras entidades” (idem, p. 35-36). Essa proposta é semelhante à de Bordini e Hübner (2006). Segundo Arnhold (2018), foi criado um *Parser* para fazer a conversão dos arquivos fontes em uma estrutura de dados executável, utilizando uma gramática estabelecida por Gelaim et al. (2018). Essa gramática é composta de distintas classes para a representação e tratamento de texto, sendo o processo de tradução subdividido em análise léxica e análise sintática (ARNHOLD, 2018). O modelo do autor foi composto por contextos (sendo obrigatório o contexto de comunicação) dos tipos funcionais (*logicalContext*) ou de estado mental (nome e fórmulas), e regras de ponte. O autor apresentou muitas possibilidades para trabalhos futuros, indicando diversos caminhos para o aprimoramento do *framework* proposto.

O trabalho informacional de criação de representações que possam ser manipuladas em sistemas computacionais tem papel fundamental como elemento comunicativo, razão pela qual a próxima seção discute sobre esse tema.

3 A complexidade das representações

Conforme descrito em dicionário, etimologicamente o termo “complexo” tem origem no latim *complexus* (cercado, abarcado, compreendido, abrangido). Por sua vez, o termo “complexidade” tem origem no latim *complexus* + *plecto, plexi, complector, plexus* (trançado, tecido; enlaçado, entrelaçado) e deriva do termo latino *plicare* (dobrar, desdobrar) (FERREIRA, 2009).

A noção emergente do fenômeno da complexidade veio contrariar a concepção de representar a realidade de maneira simplificada (racionalidade mecanicista), uma vez que

[...] temos para nós um dos princípios mais significativos da complexidade, que é o que afirma não existir apenas um único plano de realidade. O mundo, e também nós no mundo, faz parte de múltiplos planos simultâneos de realidade integrada nos novos limites da ciência diante do infinitamente grande e do infinitamente pequeno. Esta é a grande questão e o mais importante princípio estruturante da complexidade (LIMA, 2006, p. 141).

Para o autor, desde o conhecimento sobre a mecânica quântica que a concepção cartesiana de um único nível de realidade foi colocada em questão. Segundo ele, no campo informacional, ao longo dos últimos

cinquenta anos, tivemos aceleração dos meios computacionais com a microinformação digital e a microinformação genética. Estes elementos “operam apenas na escala do micrão e, mesmo assim, enfrentam atualmente novos dilemas complexos que ainda nos deixam meio atordoados, diante das suas implicações no mundo organizacional em nossas sociedades” (LIMA, 2006, p. 142), tal como os vivenciados pelo projeto Genoma, e, mais recentemente, pelos impactos da nanoinformação. O autor acrescenta que convivemos também com microparadigmas (especializações) que

[...] acontece, por exemplo, quando nos referimos apenas a subparadigmas econômicos na sociedade, tipo sociedades agrárias (materialidade da terra), sociedades industriais (materialidade das mercadorias e produtos industrializados) e sociedades de informação (imaterialidade das ideias, dos símbolos, dos ícones, da informação-imagem, da estética e do conhecimento). Vivemos no mundo da ciência e do conhecimento, imersos numa transição macroparadigmática, migramos a passos largos do paradigma cartesiano-newtoniano para o paradigma da complexidade (LIMA, 2006, p. 144).

Construir representações da realidade, de abstrações do mundo real, em ambiente de complexidade é, assim, um desafio a ser enfrentado; sobretudo para a sua manipulação por sistemas computacionais.

No âmbito da Ciência da Informação, temos diferentes tipos de instrumentos de representação de domínios de conhecimento, que podem ser delimitados como: científicos, profissionais e corporativos, de lazer, setores econômicos, áreas temáticas, processuais, de atividades. Esses instrumentos são agrupados sob a denominação de sistemas de organização do conhecimento (SOC), ou, em inglês, *Knowledge Organization Systems* (KOS). Os SOC são desenvolvidos com diferenciados níveis de estruturação e formalidade, a partir de distintas abordagens, atendendo a uma multiplicidade de aspectos (institucionais, público-alvo, recursos humanos e financeiros), sejam eles ideológicos, teóricos ou metodológicos. Têm como principal objetivo apoiar processos para tornar gerenciável, acessível e recuperável pelo usuário o recurso informacional que é disponibilizado em um Sistema de Recuperação da Informação (SRI). São ferramentas compostas por conceitos (terminologia), que representam entidades e/ou objetos, com características e metodologias próprias de construção, incluindo certas semelhanças que geralmente se aplicam a contextos e públicos específicos.

Cumprindo seus propósitos particulares, na construção de SOC buscamos resolver problemas de ambiguidade (polissemia, sinonímia),

próprias da linguagem natural, fazendo o controle da terminologia utilizada. Com isso, garante-se a consistência e minimizam-se dificuldades na comunicação dos significados dos termos que representam os conceitos (CINTRA *et al.*, 2002).

Dahlberg (2006), discorrendo sobre a definição de conhecimento, afirma que o que uma pessoa sabe sobre um fato ou um assunto não é transferível, visto que é um conhecimento tácito dela. Contudo, a autora esclarece que as pessoas são capazes de expressar as coisas do mundo empregando suas habilidades linguísticas. Sendo assim, é possível transferir qualquer conhecimento pessoal em uma linguagem de representação “não apenas para compreensão, mas, também, para comparação da comunicação entre pessoas com representações diferentes” (DAHLBERG, 2006, p. 12). Para a autora, o conhecimento pode ser representado em quatro graus diferenciados de complexidade, que têm relação com seus referentes no mundo real ou abstrato: (a) elementos constitutivos, que estabelecem as características de conceitos que predicam as propriedades (genéricas e específicas, por exemplo) sobre referentes; (b) unidades de conhecimento simples – conceitos – como uma síntese das declarações das características necessárias e suficientes sobre os referentes, que são representados por um signo (palavra, nome, termo, código); (c) unidades de conhecimento compostos e complexos, que são combinações de conceitos (em declarações, definições ou textos); e (d) sistemas de conhecimento, que são conceitos organizados em estruturas coesas e planejadas segundo um propósito.

Dahlberg (2006) acrescenta que a representação do conhecimento envolve expressá-lo em uma linguagem verbal, assim como ações de agrupamento, tendo em vista evidenciar as relações existentes entre as distintas representações, em duas básicas aplicações de organização: (a) construção de sistemas conceituais; e (b) mapeamento de unidades do conhecimento desse sistema conceitual com objetos do mundo real.

No caso da construção de sistemas conceituais, Dahlberg (2006) relata que os conceitos (como unidades de conhecimento) podem ser organizados em uma dada ordem sistemática, a partir de seu próprio conteúdo (elementos conceituais; características essenciais, acidentais e individualizantes). Para a autora, usando uma abordagem teórico-conceitual na construção de sistemas de conceitos, é possível construir sistemas relacionados a quaisquer tipos de referentes. Ela acrescenta que a maior parte dos sistemas conceituais existentes são também sistemas de classificação, uma vez que representam classes de objetos e de conceitos com suas respectivas posições na estrutura do sistema (notações; códigos), estabelecendo, assim, as relações conceituais. Sobre a construção de classificações, Hayakawa e Hayakawa (1990) afirmam que

O que chamamos de coisas e onde traçamos a linha entre uma classe de coisas e outra dependem dos interesses que temos e dos propósitos da classificação. Por exemplo, os animais são classificados de uma maneira pela indústria da carne, de uma maneira diferente pela indústria do couro, de uma maneira diferente pela indústria de peles, e ainda de maneira diferente por um biólogo. Nenhum deles é mais certo que o outro; cada um é útil para o seu propósito (HAYAKAWA; HAYAKAWA, 1990, p. 122).

Sendo assim, as classificações estarão sempre alinhadas a algum objetivo e abordagem epistemológica, sendo necessário que sejam revisados sempre que a realidade que representam sofrer transformações.

Dahlberg (2006) menciona que usando uma abordagem matemático-conceitual é possível visualizar as relações em rede, a partir da teoria dos grafos e da teoria de redes. Essa abordagem inclui uma definição matemática (formal) das entidades conceituais a ser representadas em um fluxo de pontos (nós ou vértices), direta ou indiretamente conectados por meio de linhas (arestas).

Segundo a sua natureza, os conceitos podem se referir a três categorias básicas de coisas: (a) objetos inanimados, (b) seres vivos, e (c) eventos (interações entre seres e objetos) (SZOSTAK, 2004). Souza (2002) destaca que há grande variação no nível de abstração dos conceitos, pois o nosso conhecimento sobre as coisas (dados) pode ser construído por observações e percepções, de modo direto ou indireto, ou até mesmo ser inferido a partir de múltiplas observações diretas e indiretas.

Discorrendo sobre dados observáveis, Szostak (2004) afirma que eles podem ser classificados tendo por base os cinco sentidos do ser humano: visão, audição, paladar, tato e olfato. O autor afirma que a ciência confia bastante nos dois primeiros. Segundo o autor, qualquer um desses sentidos não é infalível, e talvez seja preferível classificar dados observáveis a partir da identificação de sua natureza (objetos, seres, eventos). Szostak (2004) salienta que a capacidade humana de se comunicar é grande, ainda que se tenha que levar em conta os erros na transmissão das mensagens, sobretudo quando as pessoas envolvidas têm diferentes origens de formação e contextos culturais diversos. Para o autor, a maior parte dos eventos não se refere a fenômenos, sendo que muitos deles podem ser evidenciados por ligações causais: alguns são contáveis (*discrete*) e outros não contáveis (*continuous*).

Dada a multiplicidade de tipos de dados existentes e finalidades às quais eles serão empregados, representar esses dados em classificações, agrupando-os segundo critérios consistentes, não é uma tarefa trivial. Sem serem processados, atividade que pode ser realizada a partir de diferentes técnicas, os dados são inúteis. Muitas vezes, cada parte dos dados deverá ser classificada de maneira específica, para serem

armazenados em distintos tipos de bases de armazenamento, para que possam estar acessíveis e recuperáveis nos sistemas. Nesse sentido, os dados têm que ser considerados como conceitos comunicáveis, que têm significados para um domínio em especial.

Hjorland (2009) discute os diversos paradigmas que coexistem e disputam espaço, afirmando que a determinação do conceito de *conceito* é que evidenciará a epistemologia/metodologia adotada. No seu estudo, o autor especifica quatro, afirmando que a formação de um conceito e as relações estabelecidas em um sistema de conceitos podem ser fundamentadas em: (1) empirismo: pesquisa empírica, a partir de observações de conceitos concretos; (2) racionalismo: deduções lógicas, com a construção de conceitos, propriedades e relações *a priori*, em geral com métodos lógicos, analíticos e formais; (3) hermenêutica ou historicismo: caráter evolutivo e histórico, com o conceito construído ao longo do tempo, dinamicamente, em escala mundial; e (4) pragmática: alinhado a propósitos, objetivos e meios de consegui-lo, com conceitos formados, observando onde, como e por que será empregado.

O autor declara que a "literatura sobre conceitos é enorme e confusa, estando espalhada ao longo de muitas disciplinas e conectada a profundos problemas teóricos e filosóficos" (HJORLAND, 2008a, *on-line*). Partindo dessa premissa, em minha tese de doutorado, elaborei um quadro-síntese sobre o conceito de *conceito*, tendo em vista evidenciar algumas acepções epistemológicas recuperadas na literatura, conforme Quadro 1.

QUADRO 1 – Síntese de acepções epistemológicas sobre o conceito de conceito

ESTUDOS / EPISTEMOLOGIAS	TRATAMENTO DO CONCEITO	DESCRIÇÃO DO CONCEITO
ESTUDOS DA FILOSOFIA CLÁSSICA		
Sócrates	Essência	Formação de universais reais, a partir de ideias abstratas e concretas, para obter o caráter genérico comum à multiplicidade de exemplares.
Platão	Essência	Formação de universais (formas) reais, que estão fora da mente e fora dos objetos, obtidos pela razão.
Aristóteles	Substância	Formação de universais que não existem por si, mas encontram-se nas coisas e são percebidas pelo intelecto (forma e matéria), compostos por conteúdos imutáveis.

ESTUDOS / EPISTEMOLOGIAS	TRATAMENTO DO CONCEITO	DESCRIÇÃO DO CONCEITO
Porfírio	Elemento individual	Formação do objeto conhecido feita pela mente e individualmente; mas não há definição a respeito de um objeto ser criado na mente ou ser real.
ESTUDOS DA IDADE MÉDIA		
Abelardo	Função	Formação do objeto a partir de sua denominação, com duas funções: uma significativa (mental) e outra referencial.
ESTUDOS DA IDADE MODERNA		
Descartes	Essência (<i>cogito</i>)	Formação a partir de uma interpretação epistemológica.
Kant	Função	Formação pela ordenação de ideias, obtendo uma representação <i>universal</i> ou <i>refletida</i> , a partir de uma construção lógica, sendo abstrato em relação ao seu uso, e cresce na escala de concretude de acordo com a quantidade de características enunciadas a seu respeito.
Hegel	Processo	Formação iniciada na mente, passando pelo objetivo, com abstrações realizadas pelo sujeito, atingindo uma ideia manifesta em um domínio específico.
ESTUDOS DO SÉCULO XIX		
Frege	Função	Formação de conteúdo mental, de categorias de entidades mutuamente excludentes e exaustivas, com caráter operacional, a partir de abstrações intelectuais.
ESTUDOS DO SÉCULO XX		
Bachelard	Processo de construção	Formação pela razão, construída no processo de conhecer.
Wittgenstein	Função/Regra	Formação com base nos jogos de linguagem, em que o sentido só tem sentido se incorporado pelo sujeito em suas interações com o contexto de uso, a partir de regras que norteiam a composição de proposições sobre a sua função ou papel na atividade, nesse dado contexto.

ESTUDOS / EPISTEMOLOGIAS	TRATAMENTO DO CONCEITO	DESCRIÇÃO DO CONCEITO
Bakhtin	Função	Formação do enunciado a partir de um processo interacional social, que sofre influência do contexto, da ideologia dominante e da luta de classes.
Vygotsky	Função	Formação a partir da socialização dos indivíduos, com função comunicativa mediadora.
Kuhn	Regra	Formação por meio de um processo abduativo, com base no que é conhecido sobre a realidade, chegando-se a uma verdade provável, como um construto social que gera um comportamento conceitual.
Granger	Função científica	Formação iniciada na sintaxe e complementada pela semântica, combinando teoria e prática.
Rosch	Processo analítico	Formação por meio de um conjunto de características e atributos que serão os traços nucleares de um exemplar escolhido como representativo de uma classe, a partir de abstrações, independentemente dos contextos de uso.
Medin	Elemento constituinte	Formulação a partir de um conjunto de propriedades (traços ou atributos) que mantém relação com outros conceitos em determinado contexto (denominado por ele de teorias), sendo o exemplar parte constituinte desse contexto.
LINGUÍSTICA E TERMINOLOGIA		
Saussure	Função	Formação do significado (plano da forma), por meio de uma representação mental ou ideia, em um processo que une o significado ao significante, de maneira interdependente, compondo o signo de forma arbitrária, como resultado de uma convenção, em razão da função que desempenha.
Wüster	Unidade do pensamento	Formação a partir de um conjunto de atributos do objeto, que precede a expressão do signo linguístico; algo universal e imutável.
Cabré	Unidade terminológica	Formação a partir da materialização de um conjunto de características, com origem em uma situação comunicativa intensional, em dado contexto e tempo.

BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO

Dahlberg	Unidade de conhecimento	Formação unívoca de um conjunto de enunciados (características ou atributos) verdadeiros, a partir de representações individuais, sem considerar influências sociais nesse processo.
Ranganathan	Unidade de pensamento	Formação individual, a partir de associações de percepções na mente, usando raciocínio lógico.

ACEPÇÕES EPISTEMOLÓGICAS

Conceitualismo	Representação mental	Formação de universais na mente, sendo derivadas e assemelhadas às coisas reais.
Empirismo	Elemento da experiência	Formação a partir do agrupamento de objetos semelhantes, com a observação objetiva da realidade e evitando definir classes por propriedades teóricas.
Historicismo	Ente genético	Formação a partir de elementos genealógicos e explicitando as suas relações com teorias e discursos.
Nominalismo	Etiquetas ou rótulos	Formação de universais a partir da designação de etiquetas ou rótulos para representar objetos ou eventos particulares, nominando-os arbitrariamente.
Pragmatismo	Função	Formação a partir de elementos e de classes considerados mais adequados a um determinado propósito, aplicação e contexto de uso.
Racionalismo	Essência/ Substância	Formação de universais e primitivos, a partir dos traços ou dados essenciais do objeto, a partir de conhecimento anterior, estabelecendo relações de semelhança com outros objetos.
Realismo	Ente concreto	Formação de universais estáveis, usando a razão para chegar à verdade real, que existe, <i>a priori</i> , fora da mente e do objeto.

Fonte: Maculan (2015, p. 101-104).

Analisando o quadro, podemos dizer que, da perspectiva da Ciência da Informação, e na prática com o insumo informação, a construção de conceitos e de sistemas de conceitos em SOC pode adotar quaisquer dos tipos de tratamento do conceito expostos, desde que de maneira coerente e consistente. É possível observar que após a metade do século XX as teorias iniciam um movimento de caráter social, incluindo o sujeito no

contexto discursivo para a formação dos conceitos, evidenciando o papel ativo desse sujeito em função dos contextos de uso (situacionais). Quando inserimos um conceito em uma dada hierarquia classificatória, estamos determinando nossa visão de mundo e nossas atitudes em relação àquele conceito, que também é alinhado ao conhecimento de uma dada época e lugar (contexto). Souza (2002) destaca que quando as desordens mentais foram classificadas como possessão diabólica, por exemplo, não havia a indicação de tratamentos específicos, que hoje existem, uma vez que agora são consideradas patologias. Para o autor, “dependendo da classificação dos fenômenos, as condutas podem ser irracionais, preconceituosas, supersticiosas, ou mais gerais e mais úteis como as que a ciência busca” (SOUZA, 2002, p. 693).

Outro ponto de que não devemos nos esquecer se refere, por exemplo, às representações de termos equivalentes em diferentes idiomas (ou traduções). Para conseguir uma relação consistente e coerente de equivalência, será preciso levar em conta que os termos têm estreita ligação com o conceito por trás das palavras que o expressam. No italiano, por exemplo, temos a palavra *culaccino*, que não tem uma tradução literal em português brasileiro, cujo conceito somente pode ser explicado a partir da paráfrase “marca de líquido do fundo de um copo na mesa”. Isso vai ocorrer em muitas situações, pois cada povo e cultura (contexto cultural) se expressa de uma maneira particular.

Essa circunstância pode ser exemplificada no filme denominado *A Chegada* (*Arrival*, em inglês), do diretor franco-canadense Denis Villeneuve. O filme apresenta uma questão que buscou a ajuda de uma especialista em linguística comparada para interpretar a linguagem utilizada pelos alienígenas, tendo em vista descobrir se vinham à Terra em paz ou não. Focando nos aspectos linguísticos e comunicacionais abordados, esse tema discutido no filme sugere que o nosso modo de “compreender a realidade depende diretamente da gramática, da sintaxe e da semântica de nossa linguagem”, nos levando a “refletir sobre a importância da comunicação” (COSTA, 2017, p. 35) e dos significados expressos pelos termos que representam os conceitos. Isso porque a estrutura da língua de toda cultura é ao mesmo tempo complexa e específica, e sofre transformações ao longo do tempo.

Também temos que considerar os fenômenos de homonímia na representação de conceitos, quando temos níveis diferenciados de semelhança (fonologia; signo) entre termos que têm significados distintos. Um exemplo para um homônimo é *banco*, palavra que tem diversos significados: *banco* (tipo de cadeira; mobiliário), *banco* (instituição financeira) e *banco* (porção de mar de menor profundidade; baixio). Em sistemas de recuperação de informação orientados por palavras para o processamento de textos escritos, esses casos de polissemia precisam ser

desambiguados, tendo em vista que podem ocorrer problemas de relevância.

Os estudos filosóficos primários da lógica que trataram sobre o significado das palavras (que expressam conceitos) chamavam a atenção para duas classes básicas: termos categoremáticos e sincategoremáticos (KANT, 2014). A primeira classe inclui os termos que indicam um tal conteúdo semântico que pode ter múltiplas aplicações, como, por exemplo, a palavra "cadeira": "Uma mulher está na cadeira de vime na varanda da casa", "A cadeira tem encosto de madeira", "A estilista trocou todas as cadeiras de metal da sala". Em todas as frases, "cadeira" indica uma categoria lógica, possuindo o mesmo significado de representação, ainda que cada uma delas tenha propriedades semânticas diferenciadas. Temos ainda termos categoremáticos definíveis e os indefiníveis; por exemplo, em "o cão é preto", "cão" é um termo categoremático definível (animal, mamífero) e "preto" é um termo categoremático indefinível (o significado da cor somente é apreendido pelos sentidos, pela referência que temos dela no mundo real). Já a segunda classe se refere a termos que vêm junto aos termos categoremáticos e determinam a sua especificidade, restringindo ou ampliando o seu sentido: "a cadeira de couro marrom"; "cadeiras de vime".

Contudo, é o contexto que dará o significado categorial ao termo, pois, em geral, funcionam apenas dentro de um dado discurso de especialidade. Na construção de hierarquias em SOC, poderíamos pensar que o significado do termo sincategoremático "de vime" é expresso pela junção de um termo categoremático de sentido mais amplo "cadeiras", que poderia ser trocado pelos termos categoremáticos "cesta" ou "painel".

Pelo exposto até este ponto, considero que é possível compreender a complexidade envolvida na construção de representações acerca do conhecimento em diferentes domínios, tendo em vista criar um ambiente propício para a comunicação. Uma linguagem comum e consensual entre os membros de um dado domínio é pré-requisito essencial para uma comunicação eficaz.

Como exemplificação, a próxima seção apresenta os aspectos que devem ser considerados na criação de SOC para o domínio jurídico.

3.1 Representações no campo de informações jurídicas

A informação jurídica é constituída por diferentes documentos normativos (proposições, emendas, pareceres, relatórios) e legislativos, e aquelas de natureza pública vêm sendo disponibilizadas para consulta por acesso aberto. Para que isso seja feito de modo eficaz, é preciso que essa informação esteja em um formato estruturado e, de preferência, padronizado, o que exige a existência de um vocabulário de referência (SOC).

Guimarães (1993) alega a existência de três fontes jurídicas principais, que engloba a doutrina, a legislação e a jurisprudência. Segundo Miranda e Santos (2018), o

Direito possui terminologia própria que desenvolve conceitos e termos com definições muito particulares. Além disso, o discurso jurídico, muitas vezes, traz implícitas, informações que são propositadamente apresentadas de maneira indireta e obscura, tornando a leitura técnica e o trabalho de representação da informação ainda mais desafiador (MIRANDA; SANTOS, 2018, p. 301).

Para as autoras, o Direito possui uma linguagem de especialidade, com definições e conceitos empregados em conformidade com cada uma de suas ramificações. Elas ainda acrescentam que a linguagem do Direito é “considerada por muitos codificada, rebuscada, prolixa e formal a tal ponto que é entendida por alguns como elitizada e até causa de segregação do conhecimento jurídico e do acesso à justiça” (MIRANDA; SANTOS, 2018, p. 306).

As representações no campo jurídico devem atentar para o fato de que a linguagem jurídica adota muitas palavras do cotidiano popular com sentido técnico muito específico (REALE, 2001). Por exemplo, o termo “competência” (cujo significado na linguagem natural é “capacidade” ou “habilidade” ou “experiente”, entre outros) na linguagem jurídica, significa, conforme o *Novo Código de Processo Civil (Novo CPC, Lei nº 13.105/15)*, os limites estabelecidos em lei sobre o poder de julgar – jurisdição ou autoridade – de cada órgão ou grupo de órgãos jurisdicional. A competência pode ser atribuída em razão de quatro elementos: (a) do valor (rito ordinário, rito sumário, rito sumaríssimo); (b) da matéria (devido à recorrência de casos semelhantes – varas de direito bancário, civil, de família, entre outras); (c) da função (graus de judicatura – Juízes, Desembargadores e Ministros); (d) do lugar/territorial (domicílio do Réu; vários juízos competentes – Prevenção). Serão considerados nulos todos os atos julgados por jurisdição ou autoridade sem competência para tais questões julgadas.

Sow (2013) alerta para outro caso, de que a informação jurídica é relacional (dependente de outras fontes) e sua representação necessita levar em consideração o ordenamento jurídico brasileiro, a técnica legislativa empregada para elaborar (redação) as leis, assim como a atualização dos textos normativos.

Também é problemático para a representação da informação jurídica o fato de que o contexto da linguagem jurídica é carregado de “ideologia, amalgamada de significados e de sentidos próprios, que são muitas vezes apresentados de forma aparentemente neutra” (MIRANDA; SANTOS, 2018, p. 309). Nesse universo de incertezas, conhecer essas

características possui papel importante na garantia da segurança jurídica das informações representadas, armazenadas e disponibilizadas.

Miranda e Santos (2018) alertam ainda para a importância do reconhecimento da fonte produtora da informação jurídica, para determinar as competências jurídicas que devem estar alinhadas com a representação. Outro aspecto importante a ser considerado é que os documentos jurídicos possuem “modelos padronizados em que a estrutura textual do documento segue padrão e ritos procedimentais comuns” (p. 304), o que exige saber onde estão as informações relevantes.

Martinho (2006, p. 91) sintetiza diversos elementos que caracterizam a informação jurídica:

- a) grandes volumes de informação e rapidez da sua desatualização, face a um constante crescimento e criação de novas fontes (novas leis, novas áreas do Direito, constante produção de doutrinas);
- b) grande diversidade de fontes de informação;
- c) proliferação de suportes de informação (bases de dados, CD-ROMs, WWW); Interação com outras áreas do conhecimento (Filosofia, Sociologia, Economia, História, Informática, Política, etc.);
- d) público-alvo exigente e diversificado (pluralidade de finalidades da informação pretendida);
- e) necessidade de grande rigor e precisão da sua conservação no documento jurídico, qualquer que seja o seu suporte;
- f) necessidade de grande rapidez e rigor na sua transmissão, de modo a assegurar a sua correta utilização e aplicação, de acordo com as últimas alterações ocorridas e a garantir a fiabilidade da mensagem transmitida.

Todas essas características da informação jurídica tornam envolta em complexidade a atividade de representação para a construção de SOC, ou seja, de um sistema de conceitos jurídicos.

Pelo breve conteúdo exposto, é possível perceber que os conceitos jurídicos têm um ciclo de vida específico (longevidade) e categorias de dados específicas (conceitos criados *a priori*, para objetos concretos ou não). Tipos especiais de relações conceituais são necessários para descrevê-los, quando devem ser aplicadas tanto a abordagem sincrônica quanto a diacrônica na definição dos conceitos jurídicos e na construção do sistema de conceitos do instrumento.

No campo jurídico, a maior parte de seus conceitos são formados pela construção de novas ideias advindas do conhecimento acumulado sobre a área, e não pela abstração de objetos empiricamente observados. A linguagem jurídica, na maior parte das vezes, é a representação de fenômenos legais. Quando uma nova lei muda a descrição de um fenômeno legal de uma lei anterior, a realidade legal jurídica muda.

Assim, na definição dos conceitos jurídicos, será preciso acrescentar informações específicas, que alertem para os fatos e terminologias que ficaram obsoletos, os que são atuais e os que entrarão em vigor no futuro, por exemplo.

4 Considerações Finais

Neste artigo apresentei uma visão geral sobre as questões que permeiam as pesquisas que desenvolvo e oriento na linha de pesquisa Arquitetura & Organização do Conhecimento do PPG-GOC. As temáticas envolvem as teorias do Conceito, da Classificação e da Terminologia, incluindo os princípios da Linguística e da Lógica, perpassando as soluções computacionais que podem dar suporte à automatização ou semiautomatização dos processos de tratamento da informação para apoiar a recuperação do recurso informacional disponibilizado. Como exemplo da complexidade na atividade de representação da informação para a construção de SOC, apresentei o domínio de representação da informação jurídica, temática que vem sendo recorrente nas pesquisas orientadas por mim, pela importância que é torná-la acessível pelos cidadãos, ainda que não seja apenas neste contexto de uso que se encontram os meus interesses de pesquisa.

Na elaboração de SOC de natureza terminológica, a sua abrangência conceitual será definida pelo tipo do serviço, pelo seu objetivo e escopo, levando em consideração o domínio (garantia literária), as especificidades das necessidades informacionais do usuário (garantia do usuário), assim como do seu comportamento de busca e o ambiente (garantia da instituição) no qual o instrumento será adotado. Todas as garantias refletem os diferentes contextos que devem ser considerados na modelagem conceitual de um SOC, minimizando ou eliminando a interferência dos fenômenos de ambiguidade da linguagem natural que dificultam o processo comunicativo da mensagem veiculada.

Referências

AKMAN, V.; SURAY, M. Steps towards formalizing context. *AI Magazine*, v. 17, p. 55-72, 1996.

ARNHOLD, V. L. H. *Implementação de um framework para o desenvolvimento de agentes como sistema multi-contexto*. 2018. 108 f. Trabalho de Conclusão de Curso (Graduação em Sistema de Informação) – Universidade Federal de Santa Catarina, Centro Tecnológico, Sistema de Informação, Florianópolis, 2018.

BENERECETTI, M.; BOUQUET, P.; GHIDINI, C. Contextual Reasoning Distilled. *Journal of Theoretical and Experimental Artificial Intelligence*, v. 12, n. 3, p. 279-305, 2000.

BORDINI, R. H.; HÜBNER, J. F. BDI agent programming in agent speak using Jason. In: TONI, F.; TORRONI, P. (Ed.). *Computational logic in multi-agent systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 143-164.

BOUQUET, P. et al. Theories and uses of context in knowledge representation and reasoning. *Journal of Pragmatics*, v. 35, n. 3, p. 455-484, March 2003.

BREZOLIN, J. M. L. *Espaços conceituais: uma proposta de uso de representações conceituais aplicada a sistemas multiagentes*. 2017. 103 f. Tese (Doutorado em Ciência da Computação) – Programa de Pós-Graduação em Ciência da Computação, Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, 2017.

BRÉZILLON, P. Context in problem solving: a survey. *The Knowledge Engineering Review*, v. 14, n. 1, p. 1-34, 1999.

CHAER, I.; ROCHA, R. L. A. Estabelecimento do contexto semântico em textos em linguagem natural. In: I CONGRESSO DE COMPUTAÇÃO DA GRANDE DOURADOS, 1., 3 a 11 de novembro de 2008 na cidade de Dourados/MS. *Anais do CCGD'2008...* Dourados/MS: UFGD, 2008, v. 1.

CINTRA, A. M. M. et al. *Para entender as linguagens documentárias*. 2. ed. rev. ampl. São Paulo: Polis, 2002.

COAN, M. A importância do contexto na atribuição de referência temporal: incursões no terreno do pretérito perfeito simples. *Working Papers em Linguística*, Florianópolis, n. 4, 2000.

COSTA, A. 'A chegada': decifrando o inteiramente desconhecido. *Em Cartaz - Crítica*, (Online), v. 9, n. 3, p. 34-35, 2017.

DAHLBERG, I. Teoria do conceito. Tradução Astério Tavares Campos. *Ci. Inf.*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.

DAHLBERG, I. Knowledge organization: a new science? *Knowledge Organization*, Frankfurt, v. 33, n. 1, p. 11-19, 2006.

DEY, A. K.; ABOWD, G. D. Towards a better understanding of context and context-awareness. In: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, XI Workshop on The What, Who, Where, When, and How of Context-Awareness, April 1-6, The Hague, Netherlands. *Proceedings of CHI 2000...* Netherlands, 2000.

DUCROT, O. *Dizer e não dizer: princípios de semântica linguística*. São Paulo: Cultrix, 1977.

DURANT, W. *História da civilização*. São Paulo: Companhia Editora Nacional, 1954.

- FERREIRA, A. B. H. *Novo dicionário Aurélio da língua portuguesa*. 4. ed. Curitiba: Positivo, 2009.
- GÄRDENFORS, P. *Conceptual spaces: the geometry of thought*. Cambridge, Massachusetts: MIT Press, 2000.
- GELAIM, T. A. *Modelo de agentes e-BDI integrando confiança baseado em sistemas multi-contexto*. 2016. Dissertação (Mestrado em Ciências da Computação) – Universidade Federal de Santa Catarina. Florianópolis, 2016.
- GELAIM, T. A. et al. Sigon: a multicontext system framework for intelligent agents. *Expert Systems with Applications*, v. 119, p. 51-60, September 2018.
- GIUNCHIGLIA, F. Contextual reasoning. *Epistemologia*, v. 16, p. 345-364, 1993.
- GUIMARÃES, J. A. C. Formas da informação jurídica: uma contribuição para sua abordagem temática. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 26, n. 1/2, p. 41-54, jan./jun.1993.
- HAYAKAWA, S. I.; HAYAKAWA, A. R. *Language in thought and action*. 5. ed. Orlando: Harcourt Brace CO., 1990. Disponível em: https://archive.org/stream/in.ernet.dli.2015.30957/2015.30957.Language-In-Thought-And-Action_djvu.txt . Acesso em: 12 jun. 2019.
- HJORLAND, B. Concept in knowledge organization (KO). *The epistemological lifeboat*, 2008.
- HJORLAND, B. Concept theory. *Journal of the American Society for Information Science and Technology*, v. 60, n. 8, p. 1519-1536, 2009.
- KANT, E. *Crítica da razão pura*. Tradução de J. Rodrigues de Merege. Ciudad de México: Editora Acrópolis, 2014. (E-Book, Domínio Público).
- LEANDRO FERREIRA, M. C. *Da ambiguidade ao equívoco: da resistência da língua nos limites da sintaxe e do discurso*. Porto Alegre: Editora da Universidade, 2000.
- LIMA, G. Sociologia na complexidade. *Sociologias*, Porto Alegre, v. 8, n. 15, p. 136-148, 2006.
- MACULAN, B. C. M. S. *Estudo e aplicação de metodologia para reengenharia de tesouro: remodelagem do THESAGRO*. 2015. 345 f. Tese (Doutorado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação. Belo Horizonte, 2015.
- MARTINHO, A. M. O bibliotecário jurídico: identidade e competências profissionais. In: I ENCONTRO NACIONAL DE BIBLIOTECAS JURÍDICAS, 1., Faculdade de Direito da Universidade de Lisboa & Associação

Portuguesa de Bibliotecários, Arquivistas e Documentalistas, 2006.
Anais... Lisboa: Lisboa Faculdade de Direito, 2006.

MCCARTHY, J. Generality in Artificial Intelligence. *Communications of ACM*, v. 30, n. 12, p. 1030-1035, 1987.

MCCARTHY, J. Notes on Formalizing Context. *In: XIII INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 13.*, August 28 - September 3, Chambéry, France, 1993. *Proceedings...* Chambéry, France, 1993. p. 555-560.

MIRANDA, R.; SANTOS, C. A. C. M. Documentação jurídica: interfaces da leitura documentária, linguagem e análise de discurso no tratamento da informação. *RDBCI: Revista Digital Biblioteconomia e Ciência da Informação*, Campinas-SP, v. 16, n. 3, p. 299-316, set./dez. 2018.

ORTEGA, C. D. O conceito de documento em abordagem bibliográfica segundo as disciplinas constituintes do campo. *InCID: Revista de Ciência da Informação e Documentação*, Ribeirão Preto, v. 7, n. esp., p. 41-64, ago. 2016.

REALE, M. *Lições preliminares do Direito*. 25. ed. São Paulo: Saraiva, 2001.

SANTOS GARCÍA, D. *Fundamentos de la comunicación*. México: Aliat, 2012. Disponível em: <http://www.aliat.org.mx>. Acesso em: 12 jun. 2019.

SOUZA, M. F. As classificações e a construção do conhecimento na enfermagem. *Revista Brasileira de Enfermagem*, Brasília, v. 55, n. 6, p. 691-696, nov./dez. 2002.

SOW, M. M. *Informação jurídica: contribuições da arquitetura e organização da informação no âmbito da Câmara dos Deputados*. 2013. 47 f. Monografia (Especialização em Arquitetura e Organização da Informação) – Universidade Federal de Minas Gerais, Curso de Especialização em Arquitetura e Organização da Informação, Belo Horizonte, 2013.

SQUIRE, L. R.; KANDEL, E. R. *Memória: da mente às moléculas*. Tradução de C. Dalmaz e J. A. Quillfeldt. Porto Alegre: Artmed, 2003.

STICHT, T. G. Applications of the audread model to reading evaluation and instruction. *In: RESNICK, L.; WEAVER, P. (Ed.). Theory and practice of early reading*. Hillsdale, NJ: Erlbaum, 1975. v. 1.

STICHT, T.G. Comprehending reading at work. *In: JUST, M.; CARPENTER, P. (Ed.). Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum, 1976.

SZOSTAK, R. *Classifying science: phenomena, data, theory, method, practice*. Norwell, USA: Springer, 2004. (E-book).

TENNIS, J. T. Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero? *Brazilian Journal of Information Science*, Marília-SP, v. 6, n. 1, p. 3-15, jan./jun. 2012.

VAN DIJK, T. A. Contexto e Cognição. In: VAN DIJK, T.A. *Cognição, Discurso e Interação*. São Paulo: Contexto, 1999. (Kindle Edition).