

FUNDAMENTOS DA ORGANIZAÇÃO DO CONHECIMENTO

Resumo: A evolução dos Sistemas de Organização do Conhecimento (SOC) tem sido balizada pelos problemas que a Ciência da Informação/Organização do Conhecimento (CI/OC) se propõe a resolver. Quando a CI surge nos anos 1950-60 o problema era a “explosão da informação”, a proliferação desordenada da literatura científica, endereçado pelos primeiros Sistemas de Recuperação de Informações (SRI) e tesouros voltados para controlar os vocabulários usados para consultar estes sistemas. O problema atual, mais abrangente afetando a sociedade atual como um todo, vem sendo chamado de Big Data. Questão: a CI/OC dispõe de teorias, fundamentos, metodologias e tecnologias para desenvolver SOC a altura dos desafios colocados pela sociedade com a emergência do Big Data e da Web Semântica? Que mudanças, impactos teóricos e práticos a emergência destas questões trará? A OC hoje vem sendo chamada a construir modelos de diferentes domínios, os vocabulários semânticos. Que coisas existem em um domínio, como representá-las? Estas são perguntas da Ontologia e da Semiótica. Objetivos: Este trabalho sugere que teorias abrangentes do que sejam Informação e Conhecimento, objetos das CI/OC, não podem ser obtidos pelas CI/OC com suas próprias bases teóricas desenvolvidas até agora. Metodologia: autores fundacionais de CI/OC são revistos em relação aos fundamentos, teorias e objetos das CI/OC. Suas colocações são cotejadas com teorias gerais de alto nível como a Semiótica e a Ontologia. Conclusões: é sugerido que a Semiótica de Peirce, com suas categorias de percepção e consequente Epistemologia, e a Metafísica, em especial a Ontologia enquanto teoria dos tipos de coisas que existem, incluindo aí a realidade social criada pela cultura humana, podem ser estas teorias.

Palavras-chave: Fundamentação; organização do conhecimento; ciência da informação; ontologia; semiótico.

FUNDATIONS OF THE KNOWLEDGE ORGANIZATION

Abstract: The evolution of Knowledge Organization Systems (KOS) has been guided by the problems that Information Science/Knowledge Organization (IS/KO) proposes to solve. When IC emerged in the 1950s and 1960s, the problem was the “information explosion”, the disorderly proliferation of scientific literature, addressed by the first Information Retrieval Systems (IRS) and thesaurus aimed at controlling the vocabularies used to consult these systems. The current, broader problem affecting today's society as a whole has been called Big Data. Question: Does IS/KO have theories, fundamentals, methodologies and technologies to develop KOS that are up to the challenges posed by society with the emergence of the Semantic Web and Big Data? What changes, theoretical and practical impacts will the emergence of these issues bring? The OC is currently being called upon to build models from different domains, the semantic vocabularies. What things exist in a domain, how to represent them? These are questions from Ontology and Semiotics. Objectives: This work suggests that comprehensive theories of Information and Knowledge, objects of IS/KO, cannot be obtained by IS/KO with their own theoretical bases developed so far. Methodology: foundational authors of IS/KO are reviewed in relation to the foundations, theories and objects of IS/KO. His statements are collated with high-level general theories such as Semiotics and Ontology. Conclusions: It is suggested that Peirce's Semiotics, with its categories of perception and consequent Epistemology, and Metaphysics, especially Ontology as a theory of the types of things that exist, including the social reality created by human culture, may be these theories.

Keywords: Foundations; knowledge organization; information science; ontology; semiotics.

Carlos H. Marcondes

Doutor em Ciência da Informação pela Universidade Federal do Rio de Janeiro (UFRJ). Professor Titular do Departamento de Ciência da Informação e do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense (UFF). Pesquisador 1D do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).
ch_marcondes@id.uff.br

1 INTRODUÇÃO

Se olharmos a trajetória dos Sistemas de Organização do Conhecimento (SOC) vemos que o grande salto de qualidade em seu desenvolvimento se deu para enfrentar a chamada “explosão da informação” (SARACEVIC, 1996), fenômeno fundador da CI. Assim é denominada a proliferação desordenada da literatura científica, em especial a partir dos anos 50-60 do século XX. Neste contexto os SOCs, antes sistemas de classificação como CDD e CDU, com a função também de ordenarem livros pelas estantes ou sistemas de cabeçalhos de assunto para permitir desdobrar as fichas catalográficas, se tornaram ferramentas fundamentais para tratar os registros de trabalhos científicos que começaram a ser armazenados em bases de dados computadorizadas. SOCs eram utilizados aí para descrever os registros e atribuir assuntos de forma padronizada na entrada dos SRI, e na sua saída, na recuperação de informações propriamente dita, para auxiliar usuários na escolha de termos padronizados que representassem suas necessidades de informação ao fazerem uma consulta.

Um momento semelhante é o que estamos vivendo hoje. Juntamente à “explosão da informação” científica que só vem se acelerando, temos uma nova “explosão” ou “dilúvio de dados” (HEY; TREFETHEN, 2003), o Big Data (HJØRLAND, 2018), a produção crescente de dados, resultado e subproduto da sociedade cada vez mais complexa em vivemos. Esta sociedade necessita cada vez mais de novos mecanismos para articular suas ações, muitas das quais são agora realizadas através da Internet – interações pessoais nas das redes sociais, transações econômicas e seus dados, dados coletados automaticamente através de dispositivos de sensoriamento, monitoramento e vigilância, dados coletados através de diferentes dispositivos que alimentam as atividades científicas (SONAWANE et al., 2019); estes, segundo alguns (HEY; TANSLEY; TOLLE, 2009), alterariam a própria metodologia científica como vem sendo praticada desde Bacon. O fenômeno atual, mais abrangente por incluir a sociedade atual como um todo e não só a Ciência, vem sendo chamado de Big Data, e na Ciência, eScience etc.¹

Qual a relação entre estes dois momentos para a CI e OC? O que teriam de semelhantes? Na primeira “explosão da informação” e no lastro da experiência de desenvolvimento de SOCs tradicionais para recuperação de informação em acervos de bases de dados e bibliotecas e das

¹ A quantidade de dados produzida em somente um minuto pelas redes sociais pode ser vista no diagrama publicado pela empresa Smartinsights. (CHAFFEY, 2021).

tradições teóricas metodológicas e técnicas da Documentação e da Biblioteconomia, os SOCs originários, para acervos em papel, tiveram que evoluir. Novos SOCs avançados foram concebidos e construídos para lidar com a recuperação de informações em sistemas computacionais, os tesouros voltados para esta finalidade: “the information retrieval thesaurus” (DEXTRE CLARKE, 2016, p. 138).

No momento atual a questão chave para endereçar a gestão e propiciar o reuso em larga escala dos Big Data gerados continuamente na Internet é atribuir semântica a estes dados, especialmente semântica computacional, permitindo que programas possam compreender seus conteúdos e serem mobilizados para organizá-los e recuperá-los. Esta é a proposta da Web Semântica (BERNERS-LEE, 2001), que vem se desenvolvendo desde o início dos anos 2000, e também a proposta da nova disciplina Data Science² ou Data Analytics,³ que tenta identificar padrões que possam ter um significado ao processar grandes conjuntos de dados em formato digital (HU, 2020, p. 14). A exemplo da “explosão da informação” das décadas de 50-60, quando os SOCs emergiram como instrumentos de padronização semântica para os emergentes SRIs, os SOCs agora podem ter um papel significativo, o de atribuir sentido, semântica, tanto para pessoas quanto para máquinas, para a profusão de dados gerados por nossa sociedade. Sem esta semântica legível por máquinas não seria possível a consulta da Web como se fosse uma base de conhecimento, o acesso, processamento em larga escala e reuso do conhecimento que pode emergir desses dados. (WHAT..., [2021]).

Representar documentos e seus assuntos sempre teve um papel fundamental nas práticas desenvolvidas pela CI/OC; em especial quando, ao contrário de hoje, não se tinha acesso a documentos completos em formato digital e a representação a descritiva e temática dos documentos era um mecanismo fundamental nos processos de intermediação e avaliação de relevância levados a cabo na recuperação de informações. A OC sempre representou domínios de conhecimento ao construir SOCs como vocabulários controlados/padronizados, cabeçalhos de assunto, tabelas de classificação. Os SOCs dos anos emergentes da CI/OC, como tesouros, tinham como objetivo viabilizar a recuperação por assuntos no contexto dos SRIs, porque seus registros eram representações de objetos que tinham assuntos como uma de suas propriedades, ou seja, documentos. Hoje trata-se não só de recuperar documentos (ou suas representações) mas representações digitais de qualquer coisa, a “Internet das coisas”, IoT – “Internet of

² Ver https://pt.wikipedia.org/wiki/Ci%C3%A2ncia_de_dados

³ Ver <https://en.wikipedia.org/wiki/Analytics>

Things.”⁴ Se a Documentação (OTLET, 2018) e depois a CI significaram a autonomização da “informação”, separando-a dos livros, a Web Semântica é também, num certo sentido, a autonomização do “conhecimento”⁵; este não mais somente inserido em textos para ser interpretado por humanos, mas registrado diretamente nas triplas RDF (W3C, 2014), formando descrições/representações de “coisas”, que podem ser processadas, computadores podem fazer inferências sobre estas representações. A Web assim torna-se uma grande base de conhecimento, que pode ser consultada a respeito das “coisas” assim representadas (SPARQL 1.1 QUERY LANGUAGE, 2013).

Segundo Buckland (2012, p. 8) vivemos hoje em uma “document-pervaded society”. González de Gómez (2001) afirma que a CI/OC ampliou seu campo de atuação a partir das décadas de 198/1990, passando a incluir a gestão da informação (de documentos) em contextos corporativos, com interfaces com as ciências da administração e da computação (GONZÁLEZ DE GÓMEZ, 2001), sem, no entanto, ampliar suas bases teóricas. A CI/OC extrapolou o ambiente de atuação e a problemática da informação em C&T e passou a atuar em outros ambientes, notadamente na gestão da “informação” em ambientes corporativos, sem, no entanto, ampliar suas bases teóricas. A CI/OC extrapolou o ambiente de atuação e a problemática da informação em C&T e passou a atuar em outros ambientes, notadamente na gestão da “informação” em ambientes corporativos (GONZÁLEZ DE GÓMEZ, 2001): Por outro lado, as questões informacionais são reformuladas com novas ênfases e em novos territórios sociais, aproximando-se de outras questões como a aprendizagem, de outros atores, como as organizações e de outros cenários, como o das empresas e os negócios. Segundo Kebede (2010) a CI poderia estar tendo um papel mais destacado, teórico, no desenvolvimento da Gestão do Conhecimento. Jashapara (2005, p. 136) afirma: “Information science has played a limited role in providing fresh insights into the emerging interdisciplinary discourse of knowledge management”. Áreas como Gestão da Informação/Conhecimento passaram a desenvolver teorias próprias, sem ligação/integração com as teorias da CI/OC desenvolvidas anteriormente, aumentando a confusão e dispersão conceitual. Ao que parece estas novas teorias têm como ponto de contato a noção de “informação como coisa” (BUCKLAND, 1991), de “coisas” ou objetos potencialmente informativos, os assim chamados ativos informacionais. O

⁴ Ver https://pt.wikipedia.org/wiki/Internet_das_coisas

⁵ Pelo menos de um tipo específico dele, o conjunto de proposições descrevendo uma “coisa”, representada como um “recurso informacional”.

conceito de informação tem esbarrado constantemente, ao longo da história da CI (e também da OC) na questão da falta de uma definição abrangente.

Vários autores perceberam e discutem esta questão: “Information is notoriously a polymorphic phenomenon and a polymorphic concept so” (FLORIDI 2019). Entre nós, Pinheiro (2004) afirma que informação é um objeto obscuro.

Hoje a CI/OC estão também diante da questão de ampliar seu escopo para abranger a Web Semântica e o Big Data. Estão a CI/OC preparadas em termos de fundamentos, metodologias e tecnologias para desenvolver SOC a altura dos desafios colocados pela sociedade atual, em especial, pela emergência da Web Semântica e do Big Data? Que mudanças e impactos teóricos e práticos a emergência da Web Semântica e do Big Data trarão para a OC?

Nos SRI/SOC computadorizados das décadas de 1960/70 a tradições teóricas e metodológicas da representação temática se somam à da representação descritiva, da catalogação e elaboração de fichas catalográficas, dos formatos bibliográficos para representar documentos. O uso mais ou menos informal e os padrões de representação (de documentos) desenvolvidos com esta finalidade que a CI/OC vêm fazendo poderá dar conta de representar outras “coisas” que não sejam documentos na “IoT”?

Em que consiste representar? Quais suas finalidades práticas e qual seu papel na cognição e na transferência do conhecimento? Quais os fundamentos da representação? Que “coisas” existiriam em um determinado domínio? Como representá-las formalmente em RDF ou OWL de modo que máquinas possam “compreender” e fazer inferências sobre estas representações, para viabilizar a IoT?

Existiriam teorias mais gerais, que fundamentariam e integrariam definições, fundamentos e áreas da CI/OC que se desenvolveram separadamente como a representação descritiva e a representação temática, os assim chamados paradigmas físico, cognitivo e sociocognitivo (CAPURRO 2003), as diferentes teorias de classificação, ou, a Teoria do Conceito (DAHLBERG, 1978), a Terminologia (CABRÉ, 2005), a comunicação científica e a recuperação da informação, ou a representação da profusão de objetos simbólicos que constituem a realidade da cultura humana, sua produção, transmissão, acumulação e recuperação e utilização?

Este trabalho sugere como hipótese que uma teoria abrangente do que sejam Informação e Conhecimento, pretensos objetos contidos nos próprios nomes das disciplinas CI/OC, não podem ser obtidos pelas CI/OC, “ontologias regionais” (BUNGE, 2015, p. XIII), com suas

próprias bases teóricas desenvolvidas até agora. É necessário ir além, buscar teorias de base para a CI e OC. Furner (2014, p. 146) coloca que

In the philosophical literature, authors typically make their ontological assumptions well known, especially if those assumptions form the foundations on which are built understandings of the concepts under analysis. In LIS, on the other hand, such views are not frequently made explicit, notwithstanding their equal importance for the development of cohesive and powerful conceptual frameworks.

Representar sempre jogou um papel fundamental nas atividades da CI/OC, embora com um enfoque prático, como um pressuposto que não se discutia muito. Mas o que é representar? Santaella e Vieira (2008) colocam em sua obra a Semiótica, a teoria de todas as formas e sistemas de representação, como uma metaciência, associando-a a Ontologia e a Teoria dos Sistemas. Coincidentemente o *Tratado sobre Filosofia Básica* de Mario Bunge (2015) é subdividido no que o autor considera as teorias mais fundamentais da Ciência; nele existem 2 volumes para a Semântica, 2 para Ontologia, 1 para Teoria dos Sistemas e 2 para Epistemologia, além de 1 para Ética: “1 SEMANTICS I Sense and Reference, 2 SEMANTICS II Interpretation and Truth, 3 ONTOLOGY I The Furniture of the World, 4 ONTOLOGY II A World of Systems, 5 EPISTEMOLOGY I The Strategy of Knowing, 6 EPISTEMOLOGY II Philosophy of Science, 7 ETHICS The Good and the Right”.

A OC hoje em sendo chamada a modelar diferentes domínios de conhecimento para construir os novos vocabulários semânticos. Que coisas existem em um domínio? Como representá-las? Estas são as perguntas da Ontologia e da Semiótica. Que teorias como a Semiótica de Peirce – a teoria de todos os tipos de representação –, com suas categorias de percepção e conseqüente Epistemologia, e a Metafísica, em especial a Ontologia enquanto teoria dos tipos de coisas que existem, podem ser estas teorias.

O trabalho está organizado como se segue. Após esta introdução a seção 2 revê a trajetória e os desafios dos SOC na atualidade. A seção 3 apresenta as opções metodológicas seguidas. A seção 4 discute os indícios e justificativas porque a Semiótica de Peirce, a Ontologia e a teoria da construção da realidade social seriam estas teorias. A seção 5 discute especificamente como estas teorias podem agregar esquemas explicativos sobre os objetos e teorias da CI/OC. A seção 6 apresenta as considerações finais e conclusões.

2 TRAJETÓRIA DA CI/OC

Nesta seção vamos apresentar a trajetória, contexto, tradições, desafios e, em especial os pressupostos da CI/OC que, acreditamos, têm um impacto nos seus fundamentos teóricos e metodológicos.

No contexto da “explosão informacional” no qual surgiram SOC como os tesouros, paralelamente surgem os SRIs, dos quais os SOC eram um de seus componentes. O desenvolvimento dos SRIs bebeu em fontes de outras tradições da Biblioteconomia e da Documentação, a catalogação. As fichas catalográficas e as entradas das bibliografias serviram de modelos para os formatos bibliográficos computacionais em projetos como o Machine Readable Cataloging (MARC)⁶ – baseado no padrão de catalogação AACR2 – e o *Manual de referência do UNISIST* (DIERICKX; HOPKINSON, 1981). Estes formatos serviram de modelo para as bases de dados dos catálogos de bibliotecas e para os assim chamados serviços de indexação e resumo.

No que diz respeito à representação temática, toda uma base teórica e metodológica, desde as teorias de classificação, a Teoria da Classificação Facetada (RANGANATHAN, 1967), passando pelo Classification Research Group,⁷ a Teoria do Conceito (DAHLBERG, 1978), a Terminologia (CABRÉ, 2005), foram desenvolvidas para darem conta do desenvolvimento de SOCs. Esta tradição teórica e metodológica da OC se encontra, com o surgimento da Web Semântica, com uma tradição da Ciência da Computação, em subáreas como modelagem de sistemas e inteligência artificial, de desenvolvimento de ontologias computacionais, entre outras; estas são compreendidas como um dos fundamentos da proposta da Web Semântica. Na atualidade muitos destes novos SOCs são desenvolvidos por profissionais de computação e por cientistas de diversas áreas ou especialistas – biomédicos, estatísticos, curadores de acervos digitais em Memória e Cultura etc. As palavras de Hjørland (2008, p. 86) destacam e alertam sobre esta aproximação com outras áreas: “(LIS) is the central discipline Of KO in this narrow sense (although seriously challenged by, among other fields, computer science).” Será que a OC se limitará a desenvolver SOCs tradicionais e deixará este espaço para os especialistas em Ciência da Computação?

⁶ Ver em <https://www.loc.gov/marc/>

⁷ Ver em https://en.wikipedia.org/wiki/Classification_Research_Group

As tradições técnicas e padrões desenvolvidos pela CI/OC para dar conta da “explosão da informação” acabaram por fixar pressupostos dos SRIs/SOCs que perduram até hoje e que, na maioria das discussões da área, estão tão implícitos que se torna difícil explicitá-los, considerá-los e analisar suas consequências. Todas as teorias e metodologias da CI/OC mencionadas trazem estes pressupostos: os SRIs representam em suas bases de dados computadorizadas documentos, o MARC e os formatos bibliográficos surgidos a partir do Manual de Referência do UNISIST (DIERICKX; HOPKINSON, 1981) especificam metadados que representam documentos, e os SOCs associados a eles são instrumentos de padronização terminológica para a propriedade assunto, o campo assunto, dos registros das bases de dados computadorizadas dos SRIs. Estes registros representam objetos que têm a propriedade assunto, objetos simbólicos, documentos. Estes pressupostos implícitos respondem pela divisão que ocorre no ensino e na prática da Biblioteconomia e OC entre representação descritiva e representação temática, ou de assuntos, de um documento. Até que ponto estes pressupostos se sustentam hoje, ou são suficientes para dar conta dos desafios da era da Web Semântica, do Big Data e da Internet das “Coisas”?

Uma das diferenças marcantes entre a Web tradicional e a Web Semântica é que, enquanto aquela é uma rede de documentos hipertextuais interligados por “links” que permitem a navegação entre documentos, a Web Semântica é uma rede de representações de “coisas” as mais diversas, denominadas “recursos”, desde recursos natos digitais como uma fotografia, um vídeo, até representações digitais de mim, ou de você, de objetos físicos os mais diversos, a chamada IoT. Outra diferença marcante é que na IoT essas representações vão além dos hipertextos, são legíveis não só por pessoas, mas também por máquinas.

A base da IoT é a representação dessas coisas através de descrições composto de proposições sobre as mesmas feitas em Resource Description Framework (RDF). RDF é uma linguagem de marcação de conteúdos que tem como modelo básico descrever coisas através de proposições sobre as mesmas, formadas por triplas Sujeito, Predicado, Objeto. O RDF é a base de representações mais “ricas”, isto é, fiéis ou mais correspondentes à coisa representada, as ontologias computacionais. Ontologias computacionais são construídas/representadas em Ontology Web Language (OWL), uma extensão do RDF que incorpora construções para especificar relacionamentos classe-subclasse, domínio e escopo, cardinalidade de cada classe envolvida em um relacionamento, regras ou restrições para relacionamentos, entre outras. Desta maneira são especificadas com muito mais detalhes as coisas em um domínio, como elas

funcionam, quais suas relações, enfim, o que estas coisas (suas representações computacionais) são, como funcionam. Estas construções pré-definidas têm uma semântica a priori que permite assim que inferências automáticas sejam feitas; por exemplo, se uma classe e subclasse de outra, automaticamente infere-se que as propriedades da superclasse serão herdadas pela subclasse. No contexto da Web Semântica novos SOCs, muitos dos quais desenvolvidos em outras áreas que não a CI/OC, vêm sendo chamados Linked Open Vocabularies (LOV) (VANDENBUSSCHE et al., 2017). Seus conceitos são identificados/acessados através de links persistentes com abrangência global e servem de predicados e objetos nas triplas RDF (W3C, 2014).

Este esquema permite representar não mais os assuntos de um documento, para que este seja representado, recuperado e lido por um usuário em um SRI, para que este usuário o processe e extraia o conhecimento aí contido, mas sim este próprio conhecimento (pelo menos, em uma de suas formas), sob a forma descrições construídas a partir de um conjunto de proposições que registram as propriedades do objeto sendo representado, como triplas RDF, Sujeito, Predicado, Objeto.

Com relação ao fenômeno do Big Data, a newsletter da empresa Cognizant deixa claro a extensão do problema pelo seu título: *Making Sense of Big Data in the petabyte age*. Para viabilizar o reuso em larga escala é necessário que eles tenham sentido, semântica, não uma semântica informal, de compreensão local, mas sim padronizada, única, válida universalmente, formal, para que possam ser processados por programas (ALBALA, 2011). A CI/OC tem larga experiência em instrumentos de controle semântico, os vocabulários, tanto de dados como de metadados (ZENG, 2019).

3 OPÇÕES METODOLÓGICAS

The sciences are regional ontologies and ontology is a general science (BUNGE, 2015).

Este trabalho usa o método dedutivo, procurando identificar objetos e teorias da CI/OC aos conceitos de teorias de fundamentação como a Semiótica e a Ontologia. Utiliza como procedimento uma pesquisa bibliográfica e documental. Está orientado pelos desafios enfrentados pela CI/OC na atualidade, conforme delineados na seção 1 e aprofundados na seção 2, pelas questões e hipóteses que emergiram das discussões levantadas nestas seções,

formuladas ao final da seção 1. A hipótese é desenvolvida e endereçada na seção 4, a partir das questões formuladas e deixadas em aberto pelos autores da CI/OC e dos aportes das teorias na seção 5, Semiótica, Metafísica e Ontologia.

Este trabalho está orientado pelos desafios enfrentados pela CI/OC na atualidade, conforme delineados na seção 1 e aprofundados na seção 2, pelas questões e hipóteses que emergiram das discussões levantadas nestas seções, formuladas ao final da seção 1. O material principal são textos basilares de autores da CI/OC que discutem as mesmas questões aqui propostas. É considerado que a OC/CI é uma “ontologia regional” no sentido atribuído pela citação de Bunge. Os textos de autores da CI/OC são confrontados com textos de disciplinas muito gerais como Semiótica, Metafísica e Ontologia, e Fundamentos da Linguagem, de autores da Filosofia da Ciência como Peirce, o autor mais abrangente da Semiótica, Searle (1995), autor que desenvolve a teoria da realidade socialmente construída e instituída, que pode explicar o papel social dos documentos não só como objetos portadores de mensagens, mas como articuladores instituídos das ações sociais e Bunge que trabalha as questões da Ontologia e Semiótica na Ciência. Usa o método dedutivo, procurando identificar objetos e teorias da CI/OC aos conceitos de teorias de fundamentação como a Semiótica e a Ontologia. É considerado que a OC/CI é uma “ontologia regional” no sentido atribuído pela citação de Bunge; embora abranja um leque vasto de problemas, para os quais criou teorias específicas, mas não integradas. Como procedimento metodológico usa a pesquisa bibliográfica e documental.

4 FUNDAMENTOS DA CIÊNCIA DA INFORMAÇÃO E DA ORGANIZAÇÃO DO CONHECIMENTO

Em trabalhos anteriores já discutimos os pressupostos e fundamentos da CI/ OC (MARCONDES, 2001, 2010, 2015). Aqui estamos mais interessados em identificar que teorias mais gerais, muitas delas já citadas nestes artigos, poderiam fundamentar a CI/OC, como na citação de Bunge, e nos indícios de que a Semiótica de Peirce, a Ontologia e a teoria da construção social da realidade, de Searle (1995), são estas teorias. Nesta seção discutiremos estes indícios e justificativas a partir de duas fontes: as conceituações dos objetos da CI/OC - 4.1, e a literatura que discute estes indícios conforme sugerimos - 4.2.

4.1 OS OBJETOS DA CI E OC

A CI e a OC são disciplinas cujo surgimento e evolução estão fortemente imbricados. Hjørland (2008, p. 86) como vários outros autores, relaciona OC (KO) com “Library and Information Science”, afirmando ser esta “the central discipline of KO in this narrow sense”.

Como se verá adiante nas citações desta seção, informação e conhecimento são dois conceitos frequentemente usados para definir os objetos de uma ou outra das disciplinas, muitas vezes de forma circular (BUCKLAND, 1991; WERSIG; NEVELING, 1975), o que não contribui para seu avanço e consolidação teórica.

Outra questão é o fato de o termo informação ter sido apropriado também fora do domínio científico, por empresas e consultores, interessados mais em uma marca que em esclarecerem ou definirem o fenômeno informação, conforme Buckland (2012, p. 2).

The word ‘information’ has been used so much that it has come to dominate discourse (Day, 001). One information school website recently contained two striking statements: ‘161 exabytes of new information are created each year’ (They mean digital bits) and ‘Information: The power to transform the world’ (They don’t mean digital bits). Vagueness and inconsistency are advantageous for slogans and using ‘chameleon words’ that assume differing colors in different contexts allows flexibility for reader to perceive what they wish. However, when clarity is sought more careful definitions are needed.

As denominações Organização do Conhecimento e Ciência da Informação foram escolhas “políticas” intencionais, feitas em contextos histórico-sociais definidos (HJORLAND, 2008; SHERA, 1980), que privilegiavam “informação” ao invés de “documento”. No entanto, ambas as disciplinas não puderam ignorar e deixar de problematizar questões que estão em suas próprias denominações ao longo de suas trajetórias epistemológicas, ou seja, o que é conhecimento e o que é informação. “For a science like information science (IS) it is of course important how fundamental terms are defined; and in IS, as in other fields, the question of how to define information is often raised.” (CAPURRO; HJORLAND, 2003, p. 347).

Identificou-se aqui alguns textos marcantes da literatura de ambas as disciplinas em que a discussão dos objetos da CI/OC acontece.

4.1.1. HJORLAND

Segundo Hjørland (2008, p. 86) a OC poderia ser compreendida a partir de um enfoque, restrito, pragmático, e de um enfoque mais geral. Em seu enfoque restrito a OC teria como atividades: “... Knowledge Organization (KO) is about activities such as document description, indexing and classification performed in libraries, bibliographical databases, archives and other kinds of “memory institutions ... Library and Information Science (LIS) is the central discipline of KO in this narrow sense.”

Ainda comentando este enfoque restrito, Andersen, citado por Hjørland (2008, p. 96), descreve “o quê” é organizado na OC:

The description (indexing) and organization (classification) for retrieval of messages representing knowledge, texts by which knowledge is recorded and documents in which texts are embedded. Knowledge itself resides in minds and brains of living creatures.

Its organization for retrieval via short- and long-term memory is a principal topic of cognitive science. Library and information science deals with the description and organization of the artifacts (messages, texts, documents) by which knowledge (including feelings, emotions, desires) is represented and shared with others. These knowledge resources are often called information resources as well. Thus ‘knowledge organization’ in the context of library and information science is a short form of ‘knowledge resources organization’. This is often called ‘information organization’.

4.1.2 CAPURRO E HJORLAND

Os autores enxergam mensagens como os objetos da CI. Constatam que informação é um conceito amplo, interdisciplinar, interessando a diversas disciplinas científicas: “Almost every scientific discipline today uses the concept of information within its own context and with regard to specific phenomena.” (CAPURRO; HJORLAND, 2003, p. 357). No artigo analisam diferentes visões formuladas por diversas disciplinas, com uma preocupação epistemológica, segundo os autores, à luz da Filosofia da Ciência. Nestas diferentes visões Capurro e Hjørland (2003, p. 366) destacam um diferencial importante; embora critiquem o aspecto meramente sintático da visão da Teoria da Informação de Shannon, chamam a atenção para o seguinte aspecto:

Nevertheless, we note that Shannon retains a basic aspect of the modern concept of information in the sense of knowledge communication, namely selection. When dealing with the meaning of a message we discuss interpretation; that is, the selection

between semantic and pragmatic possibilities. To interpret a message means, in other words, to introduce the receiver's perspective — her or his beliefs and desire.

Destacando o aspecto central da interpretação como processo cognitivo que pressupõe um agente com intencionalidade, os autores concluem:

In our view, the most important distinction in the concepts of information is the distinction between information as an object or a thing (e.g. number of bits) and information as a subjective concept, information as a sign, i.e. as depending on the interpretation of a cognitive agent...This means the inclusion of interpretation processes as a *conditio sine qua non* of information processes. As we have shown several times in this article, this task of interpretation is essentially multidisciplinary and an interdisciplinary one. (CAPURRO; HJORLAND, 2003, p. 389).

4.1.3 CAPURRO

No seu texto Capurro (2003) aprofunda o que seriam, os três paradigmas epistemológicos da CI, ou as diferentes maneiras de enxergar informação, seu objeto. Visão semelhante é encontrada também em Capurro e Hjørland (2003, p. 392, 394-395). Capurro critica a visão de paradigma de Kuhn como esquemática, afirmando que a multiplicidade e coexistência de paradigmas são características da ciência normal. Toma então o conceito de paradigma para analisar a evolução epistemológica da CI. Segundo o autor estes paradigmas seriam:

Em primeiro lugar expõe-se o paradigma físico. Tomando como ponto de partida a teoria de Shannon e Weaver, mencionam-se os experimentos de Cranfield e a teoria ‘informação-como-coisa’ de Michael Buckland. Em segundo lugar analisa-se o paradigma cognitivo, representado dentre outros por B.C. Brookes, Nicholas Belkin, Pertti Vakkari e Peter Ingwersen. Por fim, expõe-se o paradigma social que tem suas origens na obra de Jesse Shera, atualmente representado pelas teorias de Bernd Frohmann, Birger Hjørland, Rafael Capurro e Søren Brier.

A referência a paradigma por Capurro, um conceito formulado por Kuhn ao analisar o que este autor chama de as revoluções científicas, remete a sucessivas conceituações do seu objeto ao longo da trajetória epistemológica da CI. Kuhn chama de revolução científica a superação de um paradigma – uma teoria explicativa de um fenômeno ou conjunto de fenômenos – por outro, em direção a uma visão mais acurada, mais precisa, deste fenômeno. Assim Kuhn descreve a superação do paradigma geocêntrico de Ptolomeu pelo paradigma heliocêntrico, de Copérnico.

Capurro reconhece a importância dos três paradigmas para a compreensão dos processos de recuperação da informação e mesmo para uma sociedade informatizada. Levanta outros aportes epistemológicos em disciplinas afins à CI que poderiam lançar nova luz sobre a RI: “A corrente epistemológica relacionada com a filosofia externalista da linguagem wittgensteiniana permite conceber os sistemas de recuperação da informação, não sob a divisa física do best matching, mas como um tipo de conversação...”. Menciona os “speech acts de John Austin (1962)” para introduzir a questão da hermenêutica, a teoria da interpretações, como aporte teórico para a CI. Conclui que: “A diferença entre mensagem, ou oferta de sentido, e informação, ou seleção de sentido, é, ao meu ver, a diferença crucial de nossa disciplina entendida assim como teoria das mensagens e não só como teoria da informação.” (CAPURRO, 2003).

Na verdade, a proposta de Capurro não é uma superação ou integração dos paradigmas tradicionais da CI, mas sim a apresentação da hermenêutica como mais um aporte teórico que poderia contribuir para a CI.

4.1.4 WERSIG E NEVELING

Segundo estes autores, analisando a literatura de comunicação técnica e documentação constata-se como o termo “informação” é vago e pode ter várias interpretações. Os autores classificam estas interpretações: “The most extreme case of polysemy in the technical communication of information and documentation is the term, 'information. A semantic analysis made by one of the authors showed that there are at least six different approaches to the use and understanding of the term in the whole field of disciplines.” (WERSIG; NEVELING, 1975, p. 22).

Estas seis interpretações ou enfoques seriam: – O enfoque estrutural, orientado à matéria (a) Structure approach (matter-oriented); neste enfoque a informação seria uma questão física, material, independente de observadores humanos. – O enfoque de informação enquanto conhecimento (b) Knowledge approach); neste enfoque informação seria igualada a conhecimento gerado a partir da percepção da realidade. – O enfoque de informação enquanto mensagem (c) Message approach), mensagens enquanto fenômenos comunicacionais, calcado na Teoria Matemática da Comunicação de Shannon e Weaver. – O enfoque do significado de uma mensagem (d) Meaning approach (characteristic of message-oriented). E o enfoque do

efeito da informação nos usuários (e) Effect approach (recipient-oriented), a redução da incerteza por um usuário ao agir informado.

4.1.5. BUCKLAND

Outro autor que reconhece diferentes aspectos do fenômeno informação é Buckland (1991 p. 351). Em seu texto o autor considera que informação poderia ser: “1. Information-as-process”, a alteração do estado de conhecimento de um agente como efeito da informação, “2. Information-as-knowledge” e “3. Information-as-thing”. Buckland também faz um esforço para integrar os três aspectos:

A key characteristic of ‘information-as-knowledge’ is that it is intangible: one cannot touch it or measure it in any direct way. Knowledge, belief, and opinion are personal, subjective, and conceptual. Therefore, to communicate them, they have to be expressed, described, or represented in some physical way, as a signal, text, or communication. Any such expression, description, or representation would be ‘information-as-thing’.

4.1.6. BROOKES (1980)

Brookes (1980, p. 131) propõe definir de Informação através da sua equação fundamental da Informação.

[...] I expressed this relationship by what I called the ‘fundamental equation’: $K[S] + K = K[S + S]$, Which states in its very general way that the knowledge structure $K[S]$ is changed to the new modified structure $K[S + S]$ by the information I , the S indicating the effect of the modification.

Thellefson, Sorensen e Thellefson (2014) comentam que a definição fornecida pela equação fundamental da Informação envolve um pressuposto metafísico, o de que é possível formalizar uma definição de informação através de uma equação matemática. Este é o sonho de toda a ciência, em especial no momento que Brookes escreveu seu artigo, em que a CI buscava seu status de ciência. Na prática a equação fundamental da Informação não vai além de propor uma mera relação entre Informação e Conhecimento, muito longe de um formalismo como o alcançado pela Física. Muitos autores, a exemplo de Buckland (1991) e outros, sugerem que Informação e Conhecimento são dois aspectos do mesmo processo cognitivo humano, o

que sugere o quão complexo é o fenômeno Informação. Uma definição abrangente de Informação precisa ir além destes aspectos.

4.1.7 BELKIN E ROBERTSON

“A Text (in information science) is a collection of signs purposely structured by a sender with the intention of changing the image-structure of recipient. Information (in information science) is the structure of any text which is capable of changing the imagestructure of a recipient” (BELKIN; ROBERTSON, 1976, p. 201). A CI é fortemente marcada pelo contexto histórico de seu surgimento e pelos problemas que endereça. Segundo um dos seus mais destacados pesquisadores, Saracevic, a CI surge para endereçar o problema do excesso de informação característico das sociedades modernas. Ela só poderia ser compreendida a partir do problema que endereça; este seria seu objeto: “In this sense information science, as any other field, is defined by the problems it has addressed and by the methods it has chosen to solve them over time” (SARACEVIC, 1995). Para as finalidades deste trabalho é interessante notar que o autor complementa sua afirmação: “Like any other field information science cannot be understood by lexical definitions or ontology alone”.

Outro autor que define a CI pelos problemas que endereça é Borko (1968, p. 3), um dos seus pioneiros; ele afirma: “Information Science is that discipline that investigates the properties and behavior, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability.”

Embora a CI reconheça o escopo amplo dos fenômenos que são do seu interesse: ser informado, conhecer, aprender -, vários autores consideram que seu escopo aplicado diz respeito praticamente a mensagens, em especial mensagens mediadas/transferidas e registradas, à economia destas mensagens. Para Capurro (2003) a CI é “a message theory”. Comentando Machlup, Capurro e Hjørland (2003) afirmam que a CI diz respeito à troca de mensagens entre seres humanos: “In short, for Machlup, information is a human phenomenon. It involves individuals transmitting and receiving messages in the context of their possible actions.” Belkin e Robertson (1976, p. 198) coincidem com a visão de Machlup ao afirmarem que “information science is specifically concerned with information in the context of human communication”.

Floridi propõe a Lógica como base para explicar o fenômeno informação. Ao discutir a Definição Geral de Informação (GDI) (FLORIDI, 2019) vincula dados, que seriam os elementos básicos da percepção, com informação: informação seria igual a dado + semântica.

A definição, aparentemente simples, para as questões levantadas neste trabalho, abre outras perspectivas. São bastante frequentes as menções na literatura sobre dados, suas relações com informação e conhecimento (BUCKLAND, 1991), frequentemente chamada de hierarquia Dados, Informação, Conhecimento, Sabedoria, em inglês DIKW – Data, Information, Knowledge, Winsdon – (ROWLEY, 2007). Embora Floridi não seja o único autor a fazê-lo, sua definição conecta dados, o elemento primário da percepção, com informação. A noção de semântica remete a compreensão, sentido, em especial, no contexto da comunicação humana. Sem falar que a semântica, enquanto resultado da ação de signos, é estudada profundamente pela Semiótica.

Percebendo talvez uma contradição entre este seu escopo aplicado e um propalado objeto tão amplo, vários autores se esforçaram por definir informação (MARCONDES, 2015). Buckland (1991, p. 358) afirma: “If anything is, or might be, informative, then everything is, or might well be, information. In which case calling something ‘information’ does little or nothing to define it. If everything is information, then being information is nothing special”. Esta visão é corroborada por Brookes (1980, p. 132) ao afirmar: “I use the perceptron only to emphasise that potential information is everywhere”. Além do que dizem os autores da CI, segundo outros autores, fora deste escopo, informação é um conceito muito genérico, de escopo muito amplo, que interesse a várias disciplinas científicas além da CI, (DAVIES; GREGERSEN, 2010; MAARTENS, 2007). Naturalmente que, dado este escopo tão amplo, é natural que a CI, com seu foco voltado à informação enquanto documento, registros etc., não tenha conseguido formular um conceito amplo e unificado de informação.

As origens e o contexto de surgimento da CI na década de 1950 também explicam a denominação de uma “ciência da informação”. Este período histórico é marcado pela percepção social do fenômeno da “explosão da informação” científica, pela disseminação da Teoria Matemática da Informação de Shannon e Weaver, pelas primeiras experiências de uso de computadores para processar textos (LUHN, 1958, 1966). Jesse Shera (1980), já citado, relata como, em uma conferência da Associação Americana de Documentalistas, foi decidida a troca de nome da entidade para American Society of Information Science, uma decisão claramente política (e não epistemológica) no sentido de valorizar a profissão de seus membros.

Retomando Buckland, se qualquer coisa pode ser informativa para os seres humanos, então talvez fosse útil deslocar o foco do que possa ser informação e passar a considerar a capacidade dos seres humanos, individual e coletivamente, de se informarem. As ciências

cognitivas e a biologia atuais, inclusive com experimentos práticos, corroboram este enfoque, oferecendo explicações detalhadas, que vão confirmando o que afirmavam algumas correntes epistemológicas, sobre a aquisição do conhecimento:

Desde um ponto de vista prático, a memória dos homens e dos animais é o armazenamento e evocação de informação adquirida através de experiências; a aquisição de memórias denomina-se aprendizado [...] O aprendizado e a memória são propriedades básicas do sistema nervoso; não existe atividade nervosa que não inclua ou não seja afetada de alguma forma pelo aprendizado e pela memória (IZQUIERDO, 1989).

Ou então:

[...] não existe experiência pura no sentido do empirismo e os fatos só são acessíveis quando assimilados pelo sujeito, o que pressupõe a intervenção de instrumentos lógico-matemáticos de assimilação construtora das relações que enquadram ou estruturam esses fatos e do mesmo modo os enriquecem (PIAGET, 1983, p. 29).

O foco no indivíduo, no agente, dentro do contexto da comunicação humana, em que este traz suas bagagens cultural e social adquiridas ao longo da sua história de vida, tem como elementos chave os processos de interpretar e informar-se. Estes processos são contemplados no escopo da Semiótica, pelo que Santaella (2008b) chama de Epistemologia Semiótica.

As relações entre CI e OC são largamente discutidas na literatura, reconhecidas desde há muito tempo e relativamente consensuais. Para Saracevic (1996, p. 48) a CI tem relações com a Biblioteconomia. Em outro artigo Saracevic (1995, p. 3) afirma que a CI se ocupa em responder a “1. How to organize information intellectually?”.

Segundo Hjørland (2008) a OC poderia ser compreendida a partir de 1) um enfoque, restrito, pragmático, e 2) um enfoque mais geral. O mesmo, de alguma maneira, acontece à CI. Ao mesmo tempo suas denominações sugerem um escopo e objeto amplos a ambas – informação e conhecimento. Ambas demonstraram realizações práticas significativas no escopo pragmático, restrito, ao passo que não alcançaram aportes significativos quando se trata dos escopos mais amplos, envolvendo informação e conhecimento. Estas tentativas teóricas mais amplas, cujo caso exemplar são os três “paradigmas” da noção de informação – paradigma físico, cognitivo e sociocognitivo (CAPURRO, 2003) – demonstram uma visão metafórica da noção de paradigma tomada de Kuhn, não têm ainda uma explicação de como se articulam nem como, eventualmente, se integram. Esta questão é notada por Raber e Budd (2003, p. 517) quando discutem os “paradigmas” da CI, afirmarem que: “Each paradigm, in its own way, mistakes the part for the whole”.

Em relação a aparente incompatibilidade dos “paradigmas” da CI, Gnoli (2018, p. 1226), se apoiando na Teoria dos níveis de existência, afirma que a OC lida com “mentefatos”, objetos que perpassam diferentes níveis de existência: “Both the cognitive approach and the sociological approach, if taken in isolation, prove to be cases of philosophical monism as they emphasize a single level over the others.”.

Raber e Budd (2003, p. 517) concluem; “there is reason to conclude that information science might be regarded as a form of semiotics in that it must embrace the relations of signification and value that produce meaning in order to realize its project”.

Embora sempre repetida a afirmação de Saracevic (1995) que considera que a CI possui relações interdisciplinares com outras disciplinas, necessárias segundo o autor para enfrentar o problema da “explosão da informação”, é necessário buscar relações com disciplinas mais gerais, independentes de domínios específicos como são a Ontologia e a Semiótica, para definir objetos tão gerais e abstratos como informação e conhecimento, relacionados à cognição humana.

A resposta a estas e a outras questões com as quais vêm se defrontando a CI e OC ao longo do seu desenvolvimento possivelmente não serão encontradas no interior destas próprias disciplinas, mas sim, elevando o olhar, como se um drone permitisse ver seu desenvolvimento de uma perspectiva ampla. Ao mesmo tempo, no desenvolvimento de outras ciências, em especial, as mais semelhantes como Comunicação Social, Economia e Psicologia.

4.2 INDÍCIOS NA LITERATURA

A questão da falta de definição de um conceito claro para informação, que tão frequentemente assombra as discussões epistemológicas da área, foi endereçada por uma proposta importante no texto de Capurro, Fleissner e Hofkirchner (1999), cujo título é “Is a Unified Theory of Information Feasible? A Dialogue.”

Para organizar artefatos como mensagens, textos, documentos dos acervos das bibliotecas, se recorria originalmente aos modelos de classificação de disciplinas científicas, propostas desde a antiguidade por diversos filósofos. A classificação dos saberes (POMBO, 1998) evolui desde a proposta de Aristóteles, passando pela classificação das disciplinas das universidades medievais – o Trivium e o Quadrivium – até a classificação da estrutura departamental das universidades européias e norte-americanas do século XIX, que serviu de

base às classificações bibliográficas como a CDD. A OC herda e dá continuidade a esta tradição, de onde emerge no século XX a noção de disciplina científica ou domínio científico. Esta noção teve que ser revista, já em fins do século XX, como surgimento dos domínios interdisciplinares (LÓPEZ-HUERTAS, 2015), que não se encaixavam nas rígidas classificações biblioteconômicas enumerativas.

Também sob o impacto recente da Ciência da Computação que modelava e representava não somente domínios disciplinares, a noção teve que ser ampliada para diferentes domínios de ação humana.

A noção de domínio, em especial a partir de Hjørland e Albretchesen (1995) vem sendo cada vez mais incorporada à OC tanto em aspectos teóricos quanto nas metodologias de representar/modelar domínios de conhecimento (CAMPOS, 2004).

Neste aspecto a OC se ocupa em identificar que termos são representativos para denominar as entidades em determinado domínio, organizá-los sistematicamente como taxonomias de conceitos mais genéricos, mais específicos, além de outros tipos de relacionamentos, com vistas a padronizar um vocabulário através do qual documentos (suas representações) neste domínio pudessem ser recuperados.

Segundo Hjørland e Hartel (2003, p. 239) um domínio (de ação humana) deve considerar três dimensões:

- (1) Ontological theories and concepts about the objects of human activity;
- (2) Epistemological theories and concepts about knowledge and the ways to acquire knowledge, implying methodological principles about the ways objects are investigated; and
- (3) Sociological concepts about the groups of people concerned with the objects.

Estas três dimensões circunscreveriam as entidades em determinado domínio. Identificar as “coisas”, entidades, em um domínio é uma questão ontológica. Entre os diferentes “approaches” para a OC em sentido mais amplo, Hjørland (2008, p. 88) relaciona:

1. The traditional approach to KO expressed by classification systems used in libraries and databases, including DDC, LCC and UDC (going back to about 1876).
2. The facet-analytical approach founded by Ranganathan about 1933 and further developed by the British Classification Research Group.
3. The information retrieval tradition (IR) founded in the 1950s.
4. User oriented and cognitive views gaining influence from the 1970s.

5. Bibliometric approaches following Garfield's construction of the Science Citation Index in 1963.
6. The domain analytic approach (first formulated about 1994).
7. Other approaches (among recent suggestions are semiotic approaches, "critical-hermeneutical" approaches, discourse-analytic approaches and genre based approaches.

Embora mencionando em 7. a Semiótica, juntamente com outras abordagens, Hjørland não aprofunda esta relação e a coloca no mesmo nível de outras, e mesmo das seis anteriores. Ora, a Semiótica tem um escopo muitíssimo mais amplo e geral que os outros mencionados por Hjørland, que são teorias no escopo e contemplando objetos da CI/OC. Em continuação serão analisados autores da CI/OC que mencionam a Semiótica.

Hjørland (2008, p. 96) relaciona a Teoria do Conceito de Dahlberg com a Semiótica: "Documents are more related to the concept and theory of semiotics (the field about signs), which may turn out to be a more fruitful theoretical frame for KO". De fato, a relação do Triângulo do Conceito de Dahlberg com o Triângulo Semiótico é imediata.

Embora não mencione Peirce como uma de suas bases teóricas, Dahlberg (1978) também usou um triângulo para sistematizar os componentes do Conceito, a unidade básica de conhecimento que compõe os SOC, como é percebido por diversos autores (FRIEDMAN; THELLEFSEN, 2011; VELTMAN, 2004). A autora propõe uma teoria objetiva de conceito adequada aos propósitos da OC de acesso, troca e reutilização de conhecimento, excluindo os aspectos epistemológicos da Semiótica Peirceana ao explicar a percepção (SANTAELLA 2008a).

O Triângulo do Conceito explica as relações entre três elementos do conceito. É, de forma semelhante a Peirce (também), uma relação triádica entre um objeto – o referente –, um conjunto de características observáveis que descrevem o objeto e um termo linguístico acordado que resume suas características. De acordo com Peirce, o Triângulo Semiótico expressa a semiose, uma relação triádica necessária entre signo, objeto e interpretante, ao contrário de uma relação diádica como em Saussure (COELHO NETTO, 1980, p. 21) somente entre signo e objeto. Para Peirce (1935, CP 1.339): epistemológicos da Semiótica Peirceana ao explicar a percepção (SANTAELLA 2008a).

O Triângulo do Conceito explica as relações entre três elementos do conceito. É, de forma semelhante a Peirce (também), uma relação triádica entre um objeto - o referente -, um conjunto de características observáveis que descrevem o objeto e um termo linguístico acordado que resume suas características.

De acordo com Peirce, o Triângulo Semiótico expressa a semiose, uma relação triádica necessária entre signo, objeto e interpretante, ao contrário de uma relação diádica como em Saussure (COELHO NETTO, 1980, p. 21) somente entre signo e objeto. Para Peirce (1935, CP 1.339):

The easiest of those which are of philosophical interest is the idea of a sign, or representation.†2 A sign stands for something to the idea which it produces, or modifies. Or, it is a vehicle conveying into the mind something from without. That for which it stands is called its object; that which it conveys, its meaning; and the idea to which it gives rise, its interpretant.

Dahlberg omitiu o vértice interpretante do Triângulo Semiótico de Peirce, ou seja, os aspectos cognitivos e subjetivos que são de interesse para Peirce, para quem todo pensamento é um signo (PEIRCE, 1931 CP 1.538). A Teoria Analítica do Conceito é centrada no referente – o objeto de Peirce. Baseia sua objetividade na coleta de um conjunto de características observáveis do referente, ou seja, o objeto de Peirce. Além de ser uma teoria, é indissociável de uma metodologia de desenvolvimento de SOCs em que os conceitos são produzidos a partir da coleta de suas características acordadas; estas garantem a validade do conceito e representam um acordo/consenso entre a comunidade de agentes em um domínio. Estas características são as bases para organizar conceitos, relacionando-os uns com os outros dentro de sistemas de conceitos, ou seja, SOCs, no que Dahlberg (1981, p. 16) chama de “campo conceitual” de um conceito.

As características dos conceitos de Dahlberg também são chamadas de propriedades na literatura de Ontologia. Propriedades (também chamadas de “atributos”, “qualidades”, “características”, “características”, “tipos”) são aquelas entidades que podem ser predicadas de coisas ou, em outras palavras, atribuídas a elas. Por exemplo, se dissermos que aquela coisa ali é uma maçã e é vermelha, provavelmente estamos atribuindo as propriedades “vermelha” e “maçã” a ela.

Relações, por exemplo, amorosas e entre marido e mulher, também podem ser vistas como propriedades e, mais geralmente, podem ser tratadas em muitos aspectos em pé de igualdade com propriedades. Na verdade, eles podem até ser vistos como tipos de propriedades. (SWOYER; ORILIA, 2011).

Uma forma especial de obter características úteis para fins de SOCs é observar objetos, analisar suas características e sintetizá-las em definições.

Definições uma classe especial de descrições textuais com uma estrutura específica. As definições especificam classes de entidades declarando suas características. De acordo com Dahlberg (1983), existem três tipos de definição. As Definições Genéricas especificam a classe genérica a que pertence uma entidade, acrescida de uma diferença específica – uma característica – que caracteriza uma subclasse dessa classe; as Definições Partitivas especificam

as partes de uma entidade; e as Definições Funcionais especificam a função de uma classe de entidades e todas as outras entidades e seus papéis envolvidos em tal função.

Organizar os objetos de um domínio pela hierarquia de classes-subclasses obtidas pela sistematização de suas definições produz a estrutura taxonômica do domínio, “taxonomic structure” (GUARINO, 1998). Outras relações usadas em SOCs, como os diferentes tipos de relações associativas, podem ser obtidas dos outros tipos de definições propostos por Dahlberg, constituindo o “campo conceitual” (DAHLBERG, 1981) de um conceito.

Brier (2006) é um autor que há muito, através de vários trabalhos, ressalta as limitações teóricas da Biblioteconomia e Ciência da Informação. Segundo o autor: “the lack of a full theoretical scientific foundation for the practice of librarianship in the age of the computer”. (BRIER, 2006). Para superar esta questão o autor propõe a Semiótica Peirceana como esta base mais consistente: “The theoretical foundation of LIS in the IR-area must be replaced by a broader foundation that incorporates the semantic production of meaning”.

O autor formula uma compreensão ampla de informação que pressupõe um processo de interpretação a partir de dados da sensibilidade por um agente, com uma individualidade, com uma história de experiências e conhecimentos prévios e inserido em relações socioculturais que lhe permitem interpretar dados, conceituá-los e compartilhar seu significado. No contexto da Recuperação da Informação, que é onde o autor aplica sua teoria, documentos seriam dados passíveis de serem interpretados por usuários e cumprirem sua função de comunicarem mensagens. O autor ressalta: “This is where semiotics both as a general and as a social science of meaning generation and interpretation can contribute to the informational view of LIS”.

Douglas Raber e John M. Budd (2003, p. 507) afirmam que a Semiótica pode integrar diferentes visões aparentemente conflitantes usadas em nossa área: “Of particular interest is the way semiotics suggests ways to bridge the theoretical gap between information as thing and information as cognitive phenomenon by positing information as a cultural phenomenon”.

Os autores desenvolvem sua análise a partir da da linguagem baseada em Saussure, na qual a sua dimensão sintagmática, ou seja, sequencial na enunciação de uma declaração se associam outras dimensões paradigmáticas relacionando os diferentes signos da assertiva a outras dimensões, como..., fornecendo o contexto capaz de gerar significados. Apontam “the loss of contextual meaning in automated information retrieval processes.” (RABER; BUDD, 2003, p. 519) como a grande deficiência dos SRI atuais e que a Semiótica poderia ajudar a superar.

Os autores Alon Friedman e Martin Thellefsen (2011) discutem a Teoria do Conceito, de grande interesse para a OC, relacionando-a com a Semiótica. Ao comparar as duas teorias os autores afirmam que a Semiótica é uma teoria ampla, geral e não específica sobre signos, ao passo que a Teoria do Conceito de Dahlberg é aplicada à análise e representação de conceitos em SOC.

Embora a questão da representação do conhecimento seja amplamente discutida na área, Friedman e Thellefsen afirmam que nenhuma destas discussões apresenta seus fundamentos teóricos e que diferentes autores discutem esta questão a partir de tradicionais escolas de pensamento. Um enfoque alternativo seria examinar a questão a partir da filosofia da linguagem. Aí entra a Semiótica para os autores. Após uma extensa exposição da Semiótica Peirceana, os autores passam a examinar a Teoria do Conceito de Dahlberg. Esta é também baseada na linguagem, uma vez que as unidades presentes nos SOC – os conceitos – são expressos através de termos linguísticos.

Os autores percebem várias similaridades entre a Semiótica Peirceana e a Teoria do Conceito de Dahlberg, como estarem inseridas nas teorias da linguagem, usarem o triângulo para relacionar seus elementos e suas interações, incorporarem tipologias e subclassificações destes elementos e endereçarem a questão da representação do conhecimento. Enquanto em Peirce a percepção/cognição é um aspecto importante da sua teoria e um fundamento para se pensar na representação do conhecimento, este aspecto não é endereçado na Teoria do Conceito; nesta os aspectos comunicacionais e consensuais são destacados. Na teoria de Peirce o significado é o resultado final do processo de semiose, o aspecto consensual de um conceito, que Peirce chama de interpretante final. Na teoria da Dahlberg este aspecto consensual, o significado do conceito, é dado pelos diferentes tipos de relações entre conceitos, obtidas das proposições verdadeiras feitas sobre o objeto e que sugerem estas relações entre conceitos. Segundo os autores, estas relações permitem, na Teoria do Conceito, classificar e relacionar conceitos.

5 A SEMIÓTICA DE PEIRCE E A ONTOLOGIA COMO TEORIAS DE BASE PARA A CI/OC

Bunge (2005, p. 24) descreve como as leis de Galileu referentes à queda dos corpos podem ser derivadas da mecânica Newtoniana e como esta, por sua vez, pode ser derivada da

Teoria da Relatividade de Einstein. Os exemplos de Bunge são esclarecedores de como teorias científicas de amplo escopo englobam teorias de escopo mais restrito.

A citação do projeto Understanding Science: how Science works, da Universidade da Califórnia, Berkeley, é bastante ilustrativa dessa questão.

Some theories, which we'll call over-arching theories, are particularly important and reflect broad understandings of a particular part of the natural world. Evolutionary theory, atomic theory, gravity, quantum theory, and plate tectonics are examples of this sort of over-arching theory. These theories have been broadly supported by multiple lines of evidence and help frame our understanding of the world around us. Such over-arching theories encompass many subordinate theories and hypotheses... (UNIVERSITY OF CALIFORNIA, 2013).

Muitos autores da área de CI/OC vêm a Semiótica apenas como mais um (entre vários) fundamentos teóricos para a área, não atentando para a abrangência e escopo gerais da Semiótica de Peirce. Furner (2014, p. 147) propõe um esquema para classificar as definições de informação, de acordo com sua filiação epistemológica; este esquema consiste em:

(i) a semiotic family in which distinctions are drawn among real-world situations, mental representations of those situations, and linguistic expressions of those representations; (ii) a socio-cognitive family in which the emphasis is on action and process, and especially processes by which people become informed or inform others; and (iii) an epistemic family in which conceptions are developed with the aim of providing an account of the properties that an information resource must have if the beliefs that are generated upon interpreting the content of that resource are seen to be justified.

A Semiótica Peirceana abrange um escopo de fenômenos que vão desde, como na classificação de Furner (i) processos epistêmicos de cognição e emergência individual de signos mentais, e destes aos linguísticos, as percepções conceitualizadas de objetos do mundo real; (ii) processos de comunicação mediados por signos compartilhados socialmente; e (iii) processos de interpretação visando a ação no mundo, a partir destes signos compartilhados.

Segundo Santaella (1983, p. 13), “Semiótica é a ciência que tem por objeto de investigação todas as linguagens possíveis, ou seja, que tem por objetivo o exame dos modos de constituição de todo e qualquer fenômeno como fenômeno de produção de significado e sentido”. Ela explica e integra fenômenos que vão desde as percepções individuais do mundo real que nos afetam, as “categorias mais universais de todas as experiências possíveis” (SANTAELLA 2008a, p. 7), (PEIRCE 1868), passando pela sua elaboração mental até se tornarem compreensíveis/inteligíveis e reconhecíveis com o aparato cultural o qual cada um de

nós é dotado ao longo de sua vida cultural, até sua comunicação enquanto conhecimento a outros seres humanos: “the necessary conditions of the transmissions of meaning by signs from mind to mind” (PEIRCE 1931, CP 1.444).

É ainda Santaella (2008a, p. 9) que acrescenta: “O que disso pode ser provisoriamente concluído é que a semiótica peirceana é uma teoria lógica e social do signo. A objetividade do interpretante é, por natureza, coletiva, não se restringindo aos humores e fantasias pessoais de um intérprete particular”. E ainda: “Para complementar, a teoria dos signos é, por fim, uma teoria sígnica do conhecimento. Todo pensamento se processa por meio de signos, qualquer pensamento é a continuação de um outro, para continuar em um outro. Pensamento é diálogo, inteligência, mente, crescimento, aprendizagem e vida”.

Representação é um conceito chave na Semiótica Peirceana. Entre outras definições⁸ representar é “To stand for, that is, to be in such a relation to another that for certain purposes it is treated by some mind as if it were that other.” (PEIRCE 1931, CP 2.273). Entre os SOC que emergiram a partir da proposta da Web Semântica estão as ontologias computacionais. Barry Smith, conhecido teórico das ontologias computacionais e coordenador do consórcio Open Biomedical Ontology (OBO), o esforço mais abrangente de padronizar e coordenar o desenvolvimento de ontologias para apoiar a pesquisa biomédica, ao discutir as ontologias computacionais, destaca tanto seu aspecto de representação de um domínio da realidade, quanto o aspecto de que este artefato de representação deve estar calcado em bases ontológicas sólidas:

An ONTOLOGY is a representational artifact, comprising a taxonomy as a proper part, whose representational units are intended to designate some combination of universals, defined classes, and certain relations between them. (SMITH et al., 2006, p. 61).

Like the scientist, the metaphysician looked and looks for unity in diversity, for pattern in disorder, for structure in the amorphous heap of phenomena- and in some cases even for some sense, direction or finality in the reality as a whole. Metaphysics and science have the same origin. However, they can be distinguished up to a certain point, namely by the scope of their problems. Whereas the scientific specialist deals with rather specific questions of fact, the ontologist is concerned with all the factual domain; he is a generalist not a fragmentarian.

John Batemann (2019, p. 127) apresenta uma análise da Linguagem baseada na ontologia de fundamentação DOLCE (BORGO; MASOLO, 2010). Instâncias da Linguagem, específicas verbalizações dos falantes, que correspondem ao conceito de parole de Saussure,

⁸ Peirce (2011).

são dependentes existencialmente de seus formuladores humanos; é o que é chamado de dimensão sintagmática da Linguagem. A Linguagem como um todo, é mais que uma específica verbalização, é mais que o conjunto das diferentes verbalizações; é, adicionalmente, o conjunto do vocabulário e das regras gramaticais socialmente convencionados, a chamada dimensão paradigmática, que permite que a Linguagem funcione como um dispositivo de comunicação.

6 CONSIDERAÇÕES FINAIS

Este trabalho busca bases teóricas que possam fundamentar a CI e a OC, que contribuam para minimizar a fragmentação teórica e conceitual da área, que tornem informação conceitualmente compatível com outras áreas. Procurou mostrar que uma teoria abrangente do que sejam Informação e Conhecimento, objetos das CI/OC, deve se fundamentar em teorias como a Semiótica de Peirce e a Ontologia.

Desde a publicação pela International Federation of Library Associations (IFLA) do modelo conceitual Functional Requirements for Bibliographic Records (FRBR), que modelos conceituais como representações de alto nível de abstração do domínio da documentação vem sendo usados. O Conceptual Reference Model (CRM) é mantido pelo International Council of Museums (ICOM), em 2006 se transforma na norma ISO 21127 e vem ganhando implementações em linguagens computacionais como OWL. Em 2017 a IFLA publica o Library Reference Model (LRM), a unificação em um único modelo conceitual do FRBR com o FRASAD e o FRSAD. Em 2019 o International Council of Archives (ICA), seguindo seus congêneres das áreas de bibliotecas e museus, publica o Records in Context Conceptual Model (RiC-CM), hoje em sua versão 2 e também sob forma de uma ontologia computacional em OWL (2007). A Library of Congress publica o modelo conceitual BIBFRAME (LIBRARY OF CONGRESS, 2012).

A criação de modelos conceituais de diferentes domínios específicos, mas envolvendo conceitos como documentos, “informação”, conhecimento, como são a interação de usuários com catálogos (FRBR/LRM), ou objetos de memória cultura (CIDOC CRM) ou registros/documentos (RiC-CM) envolve identificar as “coisas” ou entidades existentes em um domínio com categorias, ou tipos de “coisas”; esta é a questão da Ontologia enquanto área de investigação filosófica e científica. Atualmente, no contexto da Web Semântica, modelos conceituais e ontologias computacionais são profundamente relacionados, podendo ser

conversíveis entre si. A criação de um modelo conceitual ou ontologia envolve a representação de um domínio através de categorias como entidades, atributos e relacionamento, por exemplo. É através da necessidade de desenvolver modelos conceituais como os citados, que representam seus objetos, que a CI/OC pode se aproximar da Semiótica e da Ontologia.

Alguns dos objetos do que poderia ser uma ontologia regional da CI/OC seriam documento, usuário, necessidade de informação, relevância, procedência, recuperação da informação, valor de prova, texto. Uma futura pesquisa seria coletar definições destes conceitos e transcrevê-las com base nas categorias da Ontologia e da Semiótica.

REFERÊNCIAS

- ALBALA, M. Making Sense of Big Data in the Pentabyte Age. **Cognizant 20-20 insights**, june 2011. Disponível em: <https://www.cognizant.com/InsightsWhitepapers/Making-Sense-of-Big-Data-in-the-Petabyte-Age.pdf>
- BATEMAN, J. A. Ontology, Language, Meaning: Semiotic Steps Beyond the Information Artifact. In: VIEU, L.; BORGIO, S.; FERRARIO, R.; MASOLO, C. (ed.). **Ontology Makes Sense**. Amsterdam: IOS Press, 2019. p. 119-135.
- BORGIO, S.; MASOLO, C. Ontological Foundations of DOLCE. In: POLI, R.; HEALEY, M.; A. KAMEAS, A. (ed.). **Theory and Applications of Ontology: Computer Applications**, New York: Springer, 2010. p. 279-296.
- BORKO, H. Information science: what is it? **American Documentation**, New York, v. 19, n. 1, p. 3-5, 1968.
- BRIER, S. The foundation of LIS in information science and semiotics. **LIBREAS - Library Ideas**, v. 4, 2006. Disponível em <https://libreas.eu/ausgabe4/001bri.htm>. Acesso em: 11 abr. 2018.
- BUCKLAND, M. Information as thing. **Journal of the American Society of Information Science**, New Jersey, v. 42, n. 5, p.351-360, 1991. Disponível em:<http://www.sims.berkeley.edu/~buckland/thing.html>. Acesso em: abr. 13, 2009.
- BUCKLAND, M. What Kind of Science can Information Science Be? **Journal of Information Science and Technology**, v. 63, n. 1, p. 1-7, 2012. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.4060&rep=rep1&type=pdf>. Acesso em: 11 maio 2018.
- BUNGE, M. **Treatise on Basic Philosophy: Volume 3 Ontology I: The furniture of the World**. Dordrecht: D Reidel Publishing, 2015.

CABRÉ, M. T. A Terminologia, uma disciplina em evolução: passado, presente e alguns elementos de futuro. **Debate Terminológico**, n. 01, 2005. Disponível em: <https://www.seer.ufrgs.br/riterm/article/download/21286/15349>. Acesso em: 21 set. 2020.

CAMPOS, M. L. de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília. v. 33, n. 1, p. 22-32, abr. 2004. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652004000100003&lng=pt&nrm=iso. Acesso em: 23 jul. 2006.

CAPURRO, R. Angeletics - A Message Theory. *In*: DIEBNER, H. H.; RAMSAY, L. (ed.): **Hierarchies of Communication**. Karlsruhe: ZKM - Center for Art and Media, 2003. p. 58-71.

CAPURRO, R. Epistemologia e Ciência da Informação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. **Anais [...]**. Belo Horizonte: ANCIB/UFGM, 2003. Disponível em: http://www.capurro.de/enancib_p.htm. Acesso em: 8 mar. 2010.

CAPURRO, R.; FLEISSNER, P.; HOFKIRCHNER, W. Is a Unified Theory of Information Feasible? A Trialogue. *In*: HOFKIRCHNER, W. (ed.). **The quest for a unified theory of information**: proceedings of the Second International Conference on the Foundations of Information Science. Amsterdam: Gordon and Breach, 1999. p. 9-30. Disponível em: <http://www.capurro.de/trialog.htm>. Acesso em: 26 marc. 2004.

CAPURRO, R.; HJORLAND, B. The concept of Information. *Annual Review of Information Science and Technology*, v. 37, p. 343-411, 2003.

CHAFFEY, D. What Happens Online in 60 Seconds in 2021? **Smart Insights**, Leeds, UK, 2021. Disponível em: <https://www.smartinsights.com/internet-marketing-statistics/happens-online-60-seconds/>

COELHO NETTO, J. T. **Semiótica, Informação e Comunicação**. São Paulo: Ed. Perspectiva, 1980.

DAHLBERG, I. A Referent-Oriented, Analytical Concept Theory for Inter-Concept. **Knowledge Organization**, [S. l.], n. 5, p. 142-151, 1978.

DAHLBERG, I. Conceptual Definitions for Interconcept. **International Classification**, v. 8, n. 1, p. 16-22, 1981.

DAVIES, P.; GREGERSEN, N. (ed.). **Information and the Nature of Reality**: From Physics to Metaphysics. Cambridge: Cambridge University Press, 2010.

DIERICKX, H.; HOPKINSON, A. **Reference Manual for Machine-Readable Bibliographic Descriptions**. 2. ed. Paris: Unesco, 1981. Disponível em: http://biblio.cerist.dz/hrbdonf5214/ouvrages/00000000000000594806000000_2.pdf. Acesso em: 21 abr. 2012.

FLORIDI, L. Semantic Conceptions of Information. *In*: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Palo Alto: CLSI Stanford University, 2019. Disponível em: <https://plato.stanford.edu/archives/win2019/entries/information-semantic/>. Acesso em: 21 dez. 2019.

FRIEDMAN, A.; THELLEFSEN, M. Concept theory and semiotics in knowledge organization. **Journal of Documentation**, London, v. 67, n. 4, p. 644-674, 2011. Disponível em: https://www.academia.edu/download/58107871/concept_theory_and_semiotics.pdf. Acesso em: 25 nov. 2017.

GNOLI, C. Mentefacts as a missing level in theory of information Science. **Journal of Documentation**, London, v. 74, n. 6, p. 1226-1242, 2018. Disponível em: <https://doi.org/10.1108/JD-04-2018-0054>. Acesso em: 20 jun. 2020.

GONZÁLEZ DE GÓMEZ, M. N. Para uma reflexão epistemológica acerca da Ciência da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 6, n. 1, p. 5-18, 2001. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/download/433/243>. Acesso em: 15 jun. 2010.

GUARINO, N. Some ontological principles for designing upper level lexical resources. *In*: International Conference on Language Resources and Evaluation, 1., 1998, Granada. **Proceedings** [...]. Granada, Spain, 1998. Disponível em: <https://arxiv.org/pdf/cmp-lg/9809002>. Acesso em: 22 maio 2005.

HEY, T.; TANSLEY, S.; TOLLE, K. (ed.). **The Fourth Paradigm**: data intensive scientific research. Redmond, Washington: MicroSoft Research, 2009.

HEY, T.; TREFETHEN, A. **The data deluge**: An e-science perspective. [*S. l.*: s. n., 2003]. To be published in “Grid Computing – Making the Global Infrastructure a Reality”, Wiley, January 2003. Disponível em: https://eprints.soton.ac.uk/257648/1/The_Data_Deluge.pdf. Acesso em: 10 out. 2020.

HJØRLAND, B. Data (with big data and database semantics). **Knowledge Organization** [*S. l.*], v. 45, n. 8, p. 685-708, 2018. Disponível em: <http://www.isko.org/cyclo/data>. Acesso em: 3 mar. 2019.

HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, [*S. l.*], v. 35, n. 2/3, 2008. DOI: 10.5771/0943-7444-2008-2-3-86.

HJØRLAND, B.; ALBRECHTSEN, H. **Toward a new horizon in information science: Domain-analysis**. *Journal of the American Society for Information Science*, New Jersey, v. 46, n. 6, p. 400-425, 1995.

HJØRLAND, B.; HARTEL, J. Ontological, epistemological and sociological dimensions of domains. **Knowledge Organization**, [*S. l.*], v. 30, n. 3-4, p. 239-245, 2003.

HU, Y-S. The impact of increasing returns on knowledge and big data: from Adam Smith and Allyn Young to the age of machine learning and digital platforms. **Prometheus**, [S. l.], v. 36, n. 1, p. 10-29, 2020. Disponível em: <https://www.jstor.org/stable/pdf/10.13169/prometheus.36.1.0010.pdf>. Acesso em: 14 maio 2021.

JASHAPARA, A. The emerging discourse of knowledge management: a new dawn for information science research? **Journal of Information Science**, [S. l.], v. 31, n. 2, p. 136-148, 2005.

KEBEDE, G. Knowledge management: An information science perspective. **International Journal of Information Management**, [S. l.], v. 30, n. 5, p. 416-424, 2010. Disponível em: http://www.academia.edu/download/49972348/Knowledge_management_An_information_scie20161029-12442-bwepnj.pdf. Acesso em: 17 maio 2019.

LIBRARY OF CONGRESS. **Bibliographic Framework as a Web of Data**: Linked Data Model and Supporting Services. Washington DC: Library Of Congress, 2012. Disponível em: <https://www.loc.gov/bibframe/pdf/marcl-d-report-11-21-2012.pdf>. Acesso em: 26 jun. 2016.

LÓPEZ-HUERTAS, M. J. Domain analysis for interdisciplinary knowledge domains. **Knowledge Organization**, [S. l.], v. 42, n. 8, p. 570-580, 2015.

LUHN, H. P. The automatic creation of literature abstracts. **IBM Journal of research and development**, v. 2, n. 2, p. 159-165, 1958.

LUHN, H. P. Keyword-in-Context Index for Technical Literature (KWIC Index). **Readings in automatic language processing**, [S. l.], v. 1, p. 159, 1966.

MAARTENS, W. **The basic elements of nature**: matter, energy, and information. [S. l.]: Authorsden.com, 2007. Disponível em: <http://www.authorsden.com/visit/viewarticle.asp?id=30331>. Acesso em: 15 jan. 2010.

MARCONDES, C. H. Análise ontológica de definições de informação: em busca da sua essência. **Transinformação**, Campinas, v. 27, n. 2, p. 105-122, 2015. Disponível em: <https://www.scielo.br/pdf/tinf/v27n2/0103-3786-tinf-27-02-00105.pdf>. Acesso em: 27 abr. 2016.

MARCONDES, C. H. Linguagem e documento: fundamentos evolutivos e culturais da ciência da informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 2, 2010. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1019>. Acesso em: 23 jan. 2011.

MARCONDES, C. H. Representação e economia da informação. **Ciência da informação**, Brasília, v. 30, n. 1, p. 61-70, 2001. Disponível em: https://www.scielo.br/scielo.php?pid=S0100-19652001000100008&script=sci_arttext&tlng=pt. Acesso em: 1 abr. 2003.

OTLET, P. **Tratado de Documentação**: o livro sobre o livro, teoria e prática. Brasília: Briquet de Lemos Livros, 2018. Disponível em: https://repositorio.unb.br/bitstream/10482/32627/1/LIVRO_TratadoDeDocumenta%C3%A7%C3%A3o.pdf. Acesso em: 3 fev. 2021.

OWL Ontology Web Language Overview, 2004. Disponível em: <http://www.w3.org/TR/owl-features/>. Acesso em: 15 maio 2007.

PEIRCE, C. S. Collected Papers of Charles Sanders Peirce. v 1. HARTSHORNE, C.; WEISS, P. (ed.). **Principles of Philosophy**. Cambridge: Harvard University Press, 1931. Disponível em: <https://colorysemiotica.files.wordpress.com/2014/08/peirce-collectedpapers.pdf>. Acesso em: 25 abr. 2020.

PEIRCE, C. S. **76 Definitions of The Sign by C. S. Peirce**. Collected and analyzed by Robert Marty. 2011. Disponível em: <https://arisbe.sitehost.iu.edu/rsources/76DEFS/76defs.HTM>

PEIRCE, C. S. On a new list of categories. **Proceedings of the American Academy of Arts and Sciences**, v. 7, p. 287-298, 1868. Disponível em: <http://www.peirce.org/writings/p32.html>. Acesso em: 28 jul. 2018.

POMBO, O. Da classificação dos seres à classificação dos saberes. **Revista da Biblioteca Nacional de Lisboa**, n. 2, p. 19, 1998. Disponível em: <http://www.educ.fc.ul.pt/hyper/resources/opombo-classificacao.pdf>. Acesso em: 14 jun. 2013.

RABER, D.; BUDD, J. M. Information as sign: semiotics and information science. **Journal of Documentation**, London, v. 59, n. 5, P. 507-522, 2003. Disponível em: https://www.researchgate.net/profile/John_Budd/publication/242338875_Information_as_sign_Semiotics_and_information_science/links/5d6d07cf92851c853885-as-sign-Semiotics-and-information-science.pdf. Acesso em: 11 abr. 2018.

W3C. **RDF 1.1 Primer**. [S. l.: s. n.], 2014. Disponível em: <https://www.w3.org/TR/rdf11-primer/>. Acesso em: 26 maio 2016.

ROWLEY, J. The wisdom hierarchy: representations of the DIKW hierarchy. **Journal of Information Science**, [S. l.], v. 33, n. 2, p. 163-180, 2007. Disponível em: <http://web.dfc.unibo.it/buzzetti/IUcorso2007-08/mdidattici/rowleydikw.pdf>. Acesso em: 14 jun. 2013.

SANTAELLA, L. **O que é Semiótica?** Rio de Janeiro: Brasiliense, 1983.

SANTAELLA, L. **A teoria geral dos signos**. São Paulo: Cengage Learning, 2008a.

SANTAELLA, L. Epistemologia semiótica. **Cognitio: Revista de Filosofia**, v. 9, n. 1, p. 93-110, 2008b. Disponível em: <https://revistas.pucsp.br/cognitiofilosofia/article/viewFile/13531/10042>. Acesso em: 12 nov. 2020.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>. Acesso em: 19 mar. 2012.

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da informação**, Rio de Janeiro, v. 24, n. 1, p. 36-41, 1995. Disponível em: <https://pdfs.semanticscholar.org/34a6/f583e2e23c8fddda7b9ba9bf6d7b01c09697.pdf>. Acesso em: 13 maio 2019.

UNIVERSITY OF CALIFORNIA. Museum of Paleontology. Science at multiple levels. *In*: UNIVERSITY OF CALIFORNIA. Museum of Paleontology. **Understanding Science: how Science Works**. Berkeley: University of California, 2013. Disponível em: https://undsci.berkeley.edu/article/0_0_0/howscienceworks_19. Acesso em: 7 jun. 2016.

SEARLE, J. **The construction of social reality (1995)**. New York: The Free Press, 1995.

SHERA, J. Sobre Biblioteconomia, Documentação e Ciência da Informação. *In*: GOMES, H. E. (org.). **Ciência da Informação ou Informática?** Rio de Janeiro: Editora Calunga, 1980. p. 91-105.

SMITH, B.; KUSNIERCZYK, W.; SCHOBER, D.; CEUSTERS, W. Towards a Reference Terminology for Ontology Research and Development in the Biomedical Domain. *In*: CEUR, 2006, Baltimore, MD. **Proceedings** [...]. Baltimore, MD, 2006. v. 222, p. 57-65.

SONAWANE, A. R. *et al.* Network medicine in the age of biomedical big data. **Frontiers in Genetics**, v. 10, p. 294, 2019. Disponível em: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00294/full>. Acesso em: 7 fev. 2021.

SWOYER, C.; ORILIA, F. Properties. *In*: ZALTA, E. N. (ed.). **The Stanford Encyclopedia of Philosophy**. Palo Alto: CLSI Stanford University, 2011. Disponível em: <http://plato.stanford.edu/archives/win2011/entries/properties/>. Acesso em: 14 set. 2013.

THELLEFSEN, T.; SØRENSEN, B.; THELLEFSEN, M. The information concept of Nicholas Belkin revisited—some semeiotic comments. **Journal of Documentation**, London, v. 70, n. 1, p. 74-92, 2014.

VANDENBUSSCHE, P-Y. *et al.* Linked Open Vocabularies (LOV): a gateway to reusable semantic vocabularies on the Web. **Semantic Web Journal**, v. 8, n. 3, p. 437-452, 2017. Disponível em: <http://www.semantic-web-journal.net/system/files/swj1127.pdf>. Acesso em: 15 fev. 2021.

VELTMAN, K. H. Towards a Semantic Web for Culture. **Journal of Digital Information**, v. 4, n. 4, p. 3-15, 2004. Disponível em: <http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Veltman/>. Acesso em: 24 nov. 2004.

WERSIG, G.; NEVELING, U. The phenomena of interest to information science. **The Information Scientist**, v. 9, n. 4, p. 127-140, 1975. Disponível em: <https://pdfs.semanticscholar.org/bc52/766117dcae45979915be0ed63542a11e08c8.pdf>. Acesso em: 26 dez. 2020.

WHAT is a knowledge graph? Sofia, Bulgaria: Ontotext, 2021. Disponível em: <https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph>

ZENG, M. L. Interoperability. *In: In: HJØRLAND, B. (ed.). ISKO Encyclopedia of knowledge organization.* [S. l: s. n.], 2018. Disponível em: <https://www.isko.org/cyclo/>. Acesso em: 18 set. 2019.