Tendências da Pesquisa Brasileira em Ciência da Informação

UMA ANÁLISE DE DOMÍNIO SOBRE ONTOLOGIAS NO CONTEXTO DO PERIÓDICO *KNOWLEDGE ORGANIZATION*[1]

A DOMAIN ANALYSIS OF ONTOLOGIES IN THE CONTEXT OF THE JOURNAL KNOWLEDGE ORGANIZATION

Caliel Cardoso de Oliveira[2]
Maurício Coelho da Silva[3]
Thiago Henrique Bragato Barros[4]
Ana Maria Mielniczuk de Moura[5]

Resumo: Constrói-se uma Análise de domínio terminológica-conceitual sobre a temática de ontologias com o objetivo de compreender a evolução do tema nos estudos publicados no periódico *Knowledge Organization* nos últimos dez anos. Os dados foram coletados na base de dados *Web of Science* e sua análise e apresentação foram desenvolvidas utilizando o software de análise semântica *SketchEngine*. Dos principais resultados, destaca-se: o crescimento de trabalhos sobre ontologia; a predominância de estudos de caráter epistemológico e aplicado, principalmente nas temáticas de recuperação e organização do conhecimento; as ontologias de domínio como o tipo de artefato mais comumente estudado e uma maior aproximação do estudo das ontologias conforme pesquisado na área da Computação.

Palavras-Chave: Ontologia. Análise de Domínio. *Knowledge Organization.*

Abstract: *A terminological-conceptual domain analysis is built on the theme of ontologies in order to understand the evolution of the studies published in the journal Knowledge Organization in the last ten years. Data were collected from the Web of Science database and its analysis and presentation were developed using the SketchEngine semantic analysis software. Among the main findings, the following stand out: a growth of studies on ontology; the predominance of studies of an epistemological and applied nature,*

[1] Texto ampliado a partir do artigo submetido, avaliado, aprovado, apresentado e premiado no XXII ENANCIB.

[2] Mestrando em Ciência da Informação. Universidade Federal do Rio Grande do Sul. E-mail: caliel.co@gmail.com. Orcid: https://orcid.org/0000-0002-8974-0579.

[3] Mestrando em Ciência da Informação. Universidade Federal do Rio Grande do Sul. E-mail: mauriciocoelho.hlp@gmail.com. Orcid: https://orcid.org/0000-0002-7923-9457.

[4] Pós-Doutor em Ciência da Informação. Docente da Universidade Federal do Rio Grande do Sul. E-mail: bragato.barros@ufrgs.br. Orcid: https://orcid.org/0000-0001-7439-5779.

[5] Doutora em Comunicação e Informação. Universidade Federal do Rio Grande do Sul. E-mail: ana.moura@ufrgs.br. Orcid: https://orcid.org/0000-0002-7811-4660.

mainly in the themes of knowledge organization and retrieval; domain ontologies as the most commonly studied type of artifact, and a closer approach to the study of ontologies as researched in the area of Computing.

Keywords: *Ontology*. *Domain Analysis*. *Knowledge Organization*.

1 INTRODUÇÃO

A temática das ontologias na Ciência da Informação perpassa diferentes áreas do conhecimento: além da própria CI, entram em voga as áreas da Ciência da Computação, na qual a definição moderna do termo ontologia se originou, e a Filosofia, onde surge a definição clássica do termo Ontologia. Assim, a definição de Ontologia[6] difere de acordo com a área de estudos. Se partirmos da Filosofia, as ontologias tornam-se teorias com as quais se busca entender os entes do mundo ou de dado domínio; mais factivelmente, as chamadas ontologias de fundamentação buscam, baseadas em alguma teoria metafísica, fornecer as bases teóricas que permitam a conceptualização de dado universo de entidades ou domínio do conhecimento. Já na Computação, as Ontologias são entendidas como artefatos de representação do conhecimento, frutos da aplicação das teorias metafísicas ao desenvolvimento de recursos aplicações *Web*, e são utilizadas tanto para permitir a realização de operações de inferência quanto no desenvolvimento de sistemas de informação complexos, baseados em ontologias de fundamentação. Finalmente, na Ciência da Informação, entende-se que a aplicação da Ontologia especialmente no que se relaciona aos seus fundamentos terminológicos está relacionado às atividades da Organização do Conhecimento, desde os sistemas mais tradicionais, como taxonomias e tesauros, até modelos conceituais mais complexos, também chamados

[6]Cabe informar que, ao longo do texto, Ontologia, com a inicial maiúscula, refere-se à disciplina do conhecimento, enquanto ontologia, com a inicial minúscula, refere-se aos artefatos de representação do conhecimento elaborados a partir da aplicação da referida disciplina em áreas como a Ciência da Computação e a Ciência da Informação

de ontologias, que, ao serem traduzidos para uma linguagem formal, compreendida por máquinas, tornam-se aqueles artefatos desenvolvidos pela Computação. Essa relação complexa atesta a multidisciplinaridade e a transversalidade de saberes que incorrem ao se estudar a temática da Ontologia Aplicada, ou simplesmente das ontologias.

Para esse estudo, buscou-se realizar uma análise de domínio, tendo como objeto de estudo a comunidade de autores que publica no periódico *Knowledge Organization*, periódico de referência na área da Organização do Conhecimento, ligado diretamente à *International Society for Knowledge Organization* (ISKO)[7]. O KO, por atrair uma comunidade diversa de autores, veicula trabalhos de diferentes temáticas dentro do escopo da Organização do Conhecimento, sendo uma das temáticas mais predominantes as ontologias. Por ser um periódico especializado e que costuma publicar resultados de pesquisa sobre ontologias, mostra-se relevante compreender como a conceituação sobre ontologias têm sido compreendida no contexto da comunidade. Considerando o KO como um periódico engajado com diversas temáticas dentro da Organização do Conhecimento, desde a construção de sistemas de organização do conhecimento, a recuperação da informação ordenada e a aplicação da Ontologia para a classificação e ordenação dos saberes, uma análise de domínio como esta permite contextualizar e compreender a importância das ontologias para os estudos em Ciência da Informação e, especificamente, para os pesquisadores interessados em Organização do Conhecimento. Nesse contexto, o presente estudo parte do seguinte questionamento: Como tem evoluído a temática de ontologias a partir dos estudos publicados no periódico *Knowledge Organization*

[7]Mais informações sobre o periódico podem ser encontradas no *site* da ISKO: https://www.isko.org/ko.html.

(KO) nos últimos dez anos?

Estabeleceu-se ainda os seguintes objetivos que norteiam este estudo: Analisar de que forma a temática de ontologias tem sido estudada pela comunidade de autores que publicam na KO e discutir os tipos de compreensão terminológica-discursiva têm sido aplicadas nos estudos de ontologias dentro do domínio da KO. Quanto à metodologia, foram utilizados a abordagem epistemológica, terminológica e discursiva da análise de domínio, conforme apresentado por Hjorland (2017), que define um domínio como sendo um corpo de conhecimento definido por uma comunidade que compartilha de entendimentos sobre a natureza e a epistemologia de dado interesse comum a seus membros. A partir do uso do software *SketchEngine*, buscou-se alcançar os objetivos do trabalho a partir de procedimentos de análise semântico-discursiva. A seguir, são apresentadas algumas considerações teóricas acerca do o termo "ontologia" traz, conforme mencionado na introdução deste trabalho.

2 ONTOLOGIAS: CONTEXTUALIZAÇÃO NA CIÊNCIA DA INFORMAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO

O termo "ontologia" é polissêmico; embora sejam correlatos, seus sentidos variam de acordo com a área do conhecimento na qual o assunto está sendo abordado. Na Filosofia, o termo denota um subcampo preocupado com as características e propriedades dos seres-em-si, ou seja, dos entes desprendidos de classificações sistemáticas nos quais pudessem estar inseridos, delimitados apenas pelos atributos que os compunham (Sowa, 2001). Até o século 15, quando o termo é cunhado a partir da junção das expressões gregas *ontos* ("ser") com *logos* (estudo, raciocínio lógico), a categorização dos entes a partir de suas características próprias

era objeto de estudo da Metafísica, área da Filosofia Clássica explorada notavelmente pelos filósofos gregos Aristóteles e Porfírio (Sowa, 2001; Schiessl; Bräscher, 2012). Diversas contribuições oferecidas por ambos os filósofos, como as noções de Categoria, Gênero dos entes e Herança de atributos, o processo de Diferenciação, e as operações lógicas de inferência de atributos herdados a partir de máximas lógicas conhecidas como Silogismos se tornaram bases para as atividades de classificação e organização do conhecimento, dando origem à Ontologia Aplicada (Sowa, 2001). Almeida (2020) ainda cita outras linhas teóricas importantes dentro do estudo da Ontologia, para além das proposições de Aristóteles, como a Ontologia de Kant e a de Husserl, da qual surgiram os conceitos que eventualmente se tornariam conhecidos como ontologias de alto nível e ontologias de domínio. A polissemia do termo "ontologia" se reflete ainda na expressão "Ontologia Formal", que pode possuir dois sentidos: um deles, mais ligado à Filosofia, une os métodos analítico-descritivos desta área com o raciocínio lógico da matemática a fim de observar os entes em diferentes níveis de granularidade e abstração, atendo-se às suas formas, propriedades e naturezas (Herre *et al.,* 2006).

Já outro significado para Ontologia Formal é aquele que busca aliar os preceitos lógico-filosóficos ao desenvolvimento de sistemas enriquecidos semanticamente através da representação de domínios em linguagens formais, como veremos a seguir ao tratar da Ontologia Aplicada à Computação (Guarino, 1995). Esclarecidos acerca dos significados que ontologia possui quando mais próximo da Filosofia, passemos à sua apropriação por outras áreas do conhecimento.

Na Computação, a Ontologia Aplicada resultou no desenvolvimento de artefatos de software específicos, voltados para a redução de incertezas e a representação fidedigna do conhecimento. Foi G. H. Mealy, teórico da Computação,

o primeiro a utilizar o termo “ontologia” fora de um contexto ligado exclusivamente às disciplinas da Filosofia, definindo uma ontologia como sendo uma estrutura de termos formalmente organizada (Almeida, 2020). A partir desse primeiro uso, “ontologia” adquire um sentido novo, para além de seu significado como campo da Filosofia: o sentido de um objeto, um artefato informacional fruto de um processo de organização de conceitos para fins de aplicação computacional. A partir dos anos 1980, torna-se amplamente aceito entre os pesquisadores da Inteligência Artificial (IA) que, para um desenvolvimento apropriado de sistemas computacionais baseados em conhecimento, seria necessária antes a “conceptualização do domínio” para o qual aquele sistema estaria sendo desenvolvido (Vickery, 1997). Essa conceptualização se faz necessária para que o conhecimento existente no mundo real possa ser apreendido de forma compreensível por máquinas, que não dispõem de capacidades intrínsecas para captação do saber real ou criação de conhecimento a partir da dedução e experimentação. Nesse contexto, "conhecimento" passa a significar tudo aquilo que pode ser representado em uma linguagem formal, legível por máquinas, podendo ser tomado como verdade pelo sistema (Santos; Costa, 2015). Para se representar este conhecimento, adotam-se princípios ontológicos para a modelagem conceitual, levando à construção de sistemas baseados na representação do domínio e na lógica; isso permite a operacionalização de inferências automáticas por parte dos sistemas, além de garantir a construção de sistemas de informação bem-estruturados, enriquecidos semanticamente através das conceptualizações desenvolvidas para representar o domínio (Almeida, 2020). Essa apropriação dos princípios ontológicos pela Computação ocorre já na década de 1980, em especial pelos estudiosos da IA, antes de se tornar amplamente aceita na década

de 1990, com teóricos como Guarino e Gruber[8], e em especial a partir de conferências realizadas a partir de 1996, que preconiza a importância das ontologias na solução de problemas que dificultariam a interoperabilidade entre diferentes sistemas e agentes e, por consequente, a manipulação automática de dados e do conhecimento formalizado (Almeida, 2020).

Já na CI, tornou-se comum associar o estudo das ontologias a outros sistemas de Organização do Conhecimento (SOCs), como taxonomias e tesauros. Segundo Santos (2014), o artefato ontologia pode ser entendido como um SOC pois "atua em um domínio específico, representando seus conceitos e a relação entre eles", e permite a visualização estruturada desse domínio. Na CI, portanto, as ontologias se caracterizam como artefatos de representação de conteúdo, que visam à diminuição da ambiguidade na linguagem natural e, através de semântica formal, a facilitar a recuperação da informação (Santos; Correa; Lapa, 2013). Para Almeida (2020), a importância das ontologias na CI caminha de mãos dadas com a transição das atividades de organização do conhecimento para os ambientes digitais, regidos pela lógica formal e pelas linguagens estruturadas, especialmente no contexto da *Web* Semântica[9] e do enriquecimento dos recursos *web* via atribuição de sentido.

3 O PERIÓDICO KNOWLEDGE ORGANIZATION

[8]Nicola Guarino e Thomas Gruber são dois eminentes pesquisadores das áreas de Representação do Conhecimento e Engenharia Ontológica, contribuindo com uma produção teórica e metodológica reconhecida tanto na Ciência da Informação quanto na Computação, desde a década de 1990 (Café; Santos; Barros, 2015)

[9]A Web Semântica é um modelo proposto pelo teórico e cientista da computação Tim Berners-Lee para descrever um paradigma futuro da Internet, onde o conteúdo presente nas páginas da Web seria legível não apenas para humanos, mas também para as máquinas, permitindo que estas extraíssem significado daquilo que está ali descrito e operassem de forma mais ou menos autônoma por meio de programas ou aplicativos automatizados (Berners-Lee; Hendler; Lassila, 2001)

O periódico científico *Knowledge Organization* (KO) trata-se de um dos mais prestigiados periódicos da área de Organização do Conhecimento. De âmbito internacional, a KO é vinculada a ISKO e foi fundada em 1973 pela pesquisadora e doutora Ingetraut Dahlberg, primeira presidente da ISKO, a partir de um conselho de editores com representantes de diferentes regiões do mundo, as diferentes subáreas da classificação e suas respectivas temáticas (ISKO, 2022). Trata-se de um fórum para todos os interessados na organização do conhecimento em escala universal ou a nível de um domínio específico, usando abordagens analítico-conceituais ou de síntese-conceitual, bem como metodologias quantitativas e qualitativas. A produção do periódico aborda também a compilação e utilização intelectual e automática de sistemas de classificação e tesauros em todas as áreas do conhecimento, com especial atenção para os problemas de terminologia.

Segundo Alves, Dalessandro e Santos (2019) o escopo do periódico abrange diferentes aspectos dentro da OC e recebe contribuições de uma comunidade de pesquisadores diversa que inclui autores de diferentes domínios, das mais variadas áreas do conhecimento e que acabam convergindo para estudos da OC. Entre os aspectos da OC abordados pelo escopo do periódico estão os estudos da própria história da OC, estudos de fundamentação teórica envolvendo aspectos filosóficos do conhecimento e seus artefatos e das bases teóricas de classificação, as operações práticas relacionadas à indexação e classificação, problemas de terminologia gerais e específicos e aplicações de sistemas de classificação e tesauros, indexação manual e mecânica para a organização e recuperação do conhecimento, incluindo as ontologias (ISKO, 2022). A KO, assim como a ISKO num todo, serve como espaço em que a Ciência e a Filosofia convergem ao redor das temáticas atreladas à classificação e recuperação da informação e do conhecimento.

Conforme a proposta do estudo, as ontologias têm sido uma temática recorrente entre as pesquisas identificadas na KO. A proeminência das ontologias como temática e o interesse de pesquisadores da OC pelos seus estudos tiveram seu avanço ao longo do século XXI, contrastando com a percepção dos pesquisadores na década de 1990 quando os estudos em CI começaram a incluir a temática de ontologias. Esse contraste, segundo Vickery (1997) ocorreu porque havia um desencontro entre os teóricos da Informação e os da Computação: os pesquisadores da OC recebiam com questionamentos o ineditismo que se apresentava com a temática das ontologias em razão dos seus princípios, como a estruturação precisa do domínio a ser representado, já serem frequentemente aplicados na construção de SOCs, enquanto os pesquisadores da Ciência da Computação tradicionalmente não possuíam preocupações em relação a questões de controle terminológico e classificação, intrínsecas às atividades de organização do conhecimento especializado.

Ao longo do tempo as ontologias encontraram um campo fértil na OC e no periódico da ISKO dedicado à temática, sendo exploradas pelos autores que nele publicam. Para Alves, Dalessandro e Santos (2019) os pesquisadores que publicam na KO configuram comunidades discursivas que se articulam entre si e compartilham do objetivo de compreender a construção e o impacto do conhecimento científico a partir de diferentes domínios e abordagens científicas, incluindo a análise de domínio sobre os aspectos implícitos de determinadas comunidades científicas. O estudo de Alves, Dalessandro e Santos (2019, p. 142) também demonstrou a interlocução e a intensidade da internacionalização entre os pesquisadores que publicaram em coautoria na KO, mostrando que a comunidade acadêmica do periódico estabelece relações com países como “Dinamarca, Estados Unidos, Canadá, Kuwait, Singapura,

Brasil, Índia, Colômbia, Grécia, Espanha, Itália, Catar, Taiwan, Israel, Reino Unido e China".

Segundo Silva, Evangelista e Guimarães (2019) o número de publicações de pesquisadores brasileiros aumentou exponencialmente no periódico KO ao longo dos últimos anos, redimensionando o alcance da produção brasileira acerca de temáticas da OC em nível internacional e criando grupos temáticos que propiciam a interlocução entre a produção brasileira de ciência e pesquisas desenvolvidas em outros países. Nesse panorama o periódico KO configura um espaço significativo para análise de comunidades discursivas e suas respectivas produções científicas sobre determinadas temáticas da OC, contexto no qual esse estudo privilegiou a temática de ontologias e cujo os resultados serão apresentados e discutidos nas seções a seguir.

4 PROCEDIMENTOS METODOLÓGICOS E RESULTADOS

Trata-se de um estudo de análise de domínio, de caráter descritivo-exploratório, que analisou tematicamente e semanticamente os artigos publicados no periódico KO nos últimos dez anos (2010-2020). Dessa forma, buscou-se discutir como tem ocorrido a evolução dos estudos sobre ontologias nos trabalhos publicados no periódico a partir da análise dos temas e das abordagens que os autores têm dado aos estudos das ontologias. Barros e Laipelt (2021, p. 444) contextualizam a análise de domínio como um paradigma teórico e metodológico na organização do conhecimento e dão ao domínio a definição de especialidade, configurando-a como "um conjunto de pessoas que trabalham em prol de um objetivo específico" ou "um agrupamento que possui substância e é definido em uma conjuntura institucional".

Para Smiraglia (2012), um domínio pode ser composto por um grupo de pesquisadores que compartilha da mesma epistemologia, viabilizando a identificação

de correntes teóricas de determinadas áreas e/ou subáreas científicas e as possíveis relações dos autores que compõem o grupo. O domínio deste estudo é composto pelos autores que publicaram no periódico KO no período de 2010 a 2020, contexto no qual privilegia-se os estudos sobre ontologias para compor nossa discussão. Acentua-se ainda o trabalho de Barros e Laipelt (2021) como marco teórico-metodológico para a realização deste trabalho, sendo que algumas de suas etapas metodológicas foram reproduzidas tendo em mente a especificidade da temática de ontologias.

De maneira a compreender melhor os autores que compõem o domínio desse estudo, foi feito ainda um levantamento dos perfis acadêmicos e institucionais dos autores mais produtivos a partir de plataformas como o Currículo Lattes (no caso de pesquisadores brasileiros), o *ResearchGate* e também o site das instituições às quais os autores são filiados. Essa busca permitiu identificar dados como o nível de titulação desses pesquisadores, seu vínculo institucional atual/mais recente e o grau de especialização que possuem em relação à temática de Ontologias. Constatou-se que quatro pesquisadores possuem três ou mais publicações no periódico KO nos últimos 10 anos, enquanto os demais pesquisadores que compõem o domínio estudado possuem 2 ou 1 publicação. Nesse sentido, foi analisado o perfil dos autores com mais de três trabalhos em Ontologias nos últimos 10 anos a partir das plataformas mencionadas e de maneira a compreender sua relação com as temáticas de ontologias.

Em relação a análise de domínio propriamente, foi realizado em um primeiro momento um levantamento dos trabalhos publicados na KO na temática de ontologias. Para isso, utilizou-se a base de dados *Web of Science* onde foram feitas buscas utilizando as expressões "*Ontologies*" e "*Ontology*", uma vez que os trabalhos

publicados no periódico são em língua inglesa. A base de dados permite filtrar pela recuperação de trabalhos de um ou mais periódicos específicos, de modo que a pesquisa foi refinada de maneira a recuperar somente trabalhos publicados na KO. Foi ainda utilizado o filtro de tempo para trabalhos publicados de 2010 a 2020, de maneira a recuperar somente trabalhos dos últimos dez anos. Nesse primeiro momento foram analisados os títulos e resumos dos trabalhos de maneira a confirmar de que se tratam de trabalhos sobre ontologias.

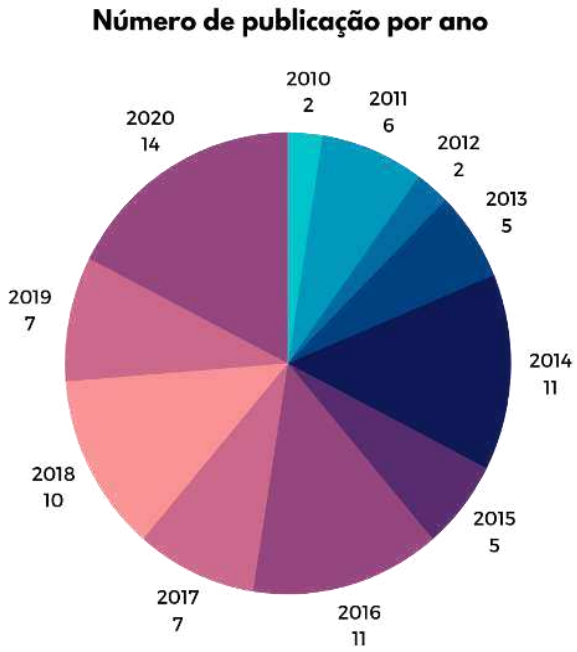
Ambas as buscas pelos termos "*Ontology*" e "*Ontologies*" na *Web of Science* recuperaram 80 trabalhos publicados no periódico KO do ano 2010 até o ano de 2020. Os resultados de ambas as buscas foram exportados da própria base de dados em uma planilha Excel, onde posteriormente foram comparados para excluir resultados duplicados. Identificou-se que ambos os termos recuperaram os mesmos 80 resultados, uma vez que foram todos duplicados, fazendo com o que o número total de trabalhos que compõem essa pesquisa seja 80, sendo este o nosso *corpus* de análise discutido na seção seguinte.

5 ANÁLISE DOS DADOS E DISCUSSÃO

Primeiramente, analisou-se os dados de publicação por ano para compreender a evolução da temática dentro do periódico nos últimos dez anos. A partir dos dados organizados na planilha, identificou-se que nos últimos 5 anos (2016-2020) aumentou o número de publicações com a temática de ontologias no periódico, com um total de 49 trabalhos publicados nesse período de tempo, enquanto no período anterior (2010-2015) foram publicados 31 trabalhos, sendo o ano de 2020 o com maior número de publicações (14 trabalhos). Isso demonstra que a temática vem ganhando

relevância dentro do periódico e ganhando maior atenção dos autores que publicam na KO.

Figura 1: Número de trabalhos publicados por ano na KO que abordam ontologias



Fonte: elaborada pelos autores.

Também foram analisados os autores mais produtivos em ontologias na KO nos últimos 10 anos, de maneira a compreender quem são os autores que compõem o domínio que esse estudo se propõe a analisar. Conforme destacam Barros e Laipelt (2021), o domínio se dá a partir de um conjunto de pessoas que trabalha por um objetivo em comum e/ou que possui uma substância em comum, que nesse caso é o empenho no desenvolvimento do estudo de ontologias no escopo da Organização do Conhecimento. Nesse sentido, foram identificados os autores que publicaram três ou mais trabalhos sobre ontologias na KO nos últimos dez anos e os resultados encontram-se no Quadro 1.

Quadro 1: Autores com mais de três produções em Ontologias na *Knowledge Organization* no período de 2010 a 2020.

Autor	Nº de publicações	Instituição de afiliação	Titulação máxima
Smiraglia, Richard P.	7	Universidade de Wisconsin - Milwaukee e Institute for Knowledge Organization and Structure	Doutor
Campos, Maria Luiza de Almeida	3	Universidade Federal Fluminense	Doutora
Das, Subhashis	3	University of Arkansas	Doutor
Wei, Wei	3	Big Data Institute, Zhongnan University of Economics and Law, China	Doutora

Fonte: elaborado pelos autores.

No Quadro 1 foram inseridos os autores com mais de três publicações em Ontologias entre os trabalhos publicados na KO no período 2010-2020 e que compõem o *corpus* deste estudo. Nota-se que somente quatro autores possuem mais de três publicações, sendo que os demais publicaram duas ou uma vez no periódico nos últimos dez anos. Observa-se que o pesquisador Richard P. Johnson se destacou com sete trabalhos, seguido por outros três autores com três trabalhos cada. Destaca-se o nível de especialização e dedicação a temática dos autores mais produtivos em relação às ontologias, sendo que o autor Richard P. Johnson atuou como *editor-in-chief* da KO e também é conhecido como um pesquisador proeminente no campo da Organização do Conhecimento, sendo também um membro do *Institute for Knowledge Organization and Structure* e membro do grupo de pesquisa de Organização da Informação da Universidade de Wisconsin.

A pesquisadora Maria Luiza de Almeida Campos é líder do grupo de pesquisa EOOCI: Estudos Ônticos e Ontológicos em Contextos Informacionais: representação, recuperação e métricas e também leciona na área de Organização e Recuperação da Informação. Isso, juntamente da titulação máxima de todos os pesquisadores destacados (doutor), demonstra o alto nível de especialização da temática de Ontologias no Escopo da KO. Além disso, também foi possível identificar que pelo menos 14 autores dos que compõem o domínio deste estudo são brasileiros, demonstrando contribuições de pesquisadores em âmbito nacional para a KO e também que a temática de ontologias tem sido interesse de estudo de pesquisadores brasileiros nos últimos dez anos.

Após a coleta dos trabalhos, passamos para a etapa da análise de domínio propriamente dita. O domínio, neste contexto, é composto pela comunidade de autores que publicam acerca de ontologias na *Knowledge Organization,* representada pelos textos que recuperamos em nossa busca, especialmente quando consideramos o recorte temporal aplicado naquela etapa. Nesse sentido, o esforço de análise foi empreendido a partir da análise sintático-semântica do *corpus* textual selecionado, processado através do uso do software *SketchEngine*. O *SketchEngine* é uma ferramenta que analisa aspectos linguísticos de um *corpus* textual usando algoritmos capazes de analisar textos a fim de identificar na linguagem utilizada no texto ocorrências raras, emergentes ou incomuns, bem como ocorrências frequentes. A partir da análise do *corpus* textual, o programa permite ao usuário gerar gráficos que ilustram essas ocorrências no texto (Barros; Laipelt, 2021). Para esse trabalho, inserimos no *SketchEngine* os arquivos originais dos 80 trabalhos recuperados, gerando o *corpus* textual a ser analisado.

Inicialmente, fizemos uma busca no *SketchEngine* pelos substantivos mais comuns no *corpus* analisado, identificando os 50 termos mais frequentes no *corpus*. Devido aos limites deste formato de publicação, apresentamos apenas os 10 primeiros resultados:

Figura 2: 10 substantivos mais frequentes no *corpus* de análise

	Lemma	Frequency
1	ontology	5,327
2	knowledge	4,562
3	[number]	4,114
4	information	3,477
5	organization	2,632
6	concept	2,135
7	system	2,067
8	domain	2,042
9	term	1,901
10	classification	1,510

Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

É possível observar, além da já esperada predominância do termo *ontology*, outros termos que confirmam que a visão predominante acerca de ontologias dentro da KO ainda as trata como SOCs. Termos como *concept*, *organization* e *classification* demonstram preocupações de caráter epistemológico acerca das ontologias, enquanto a presença de *term* e *system* indica estudos de caráter aplicado. A presença de *domain* pode indicar que as ontologias de domínio são a tipologia mais comum de artefato estudado pela comunidade, como sistemas para a organização de um domínio específico do conhecimento ou uma tarefa específica, conforme a definição de ontologia de domínio da Enciclopédia da ISKO (Biagetti, 2021); cabe apontar, no entanto, que "*domain*" também pode surgir independentemente de ser qualificador de "*ontology*". É interessante notar a ausência de *representation*, termo que, unido a

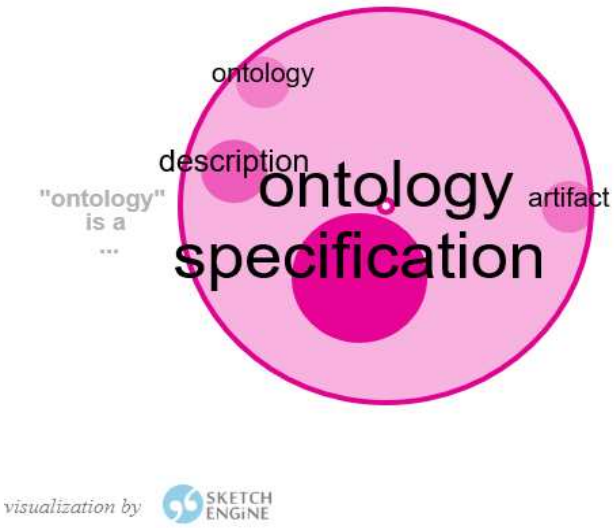
knowledge, indicaria estudos preocupados com questões advindas da Representação do Conhecimento e o estudo de IAs, originadas na Ciência da Computação (Almeida, 2020). Desta forma, podemos observar que a comunidade que compõe o domínio analisado não se afasta muito da visão própria da CI acerca das ontologias, trabalhando dentro do nicho da Informação e não da Computação. Na figura 1, o termo "[*number*]" trata-se de uma poluição dos dados gerados pelo programa, uma vez que ele faz a leitura de todo texto dos trabalhos, incluindo elementos técnicos do periódico e folha de rosto.

Outro uso do *software* foi durante as buscas por sintagmas e expressões que indicassem a relevância de diferentes tipos de ontologias dentro do domínio analisado. A busca por expressões como *domain ontolog**[10] (265 ocorrências), *top-level ontolog.** (54 ocorrências) e *foundational ontolog** (19 ocorrências) reforça que os estudos acerca do tema ainda exploram aspectos aplicados a domínios específicos, no caso do primeiro termo, e epistemológicos, no caso dos dois últimos. Cabe apontar também a ocorrência de formal *ontolog** (196 ocorrências), cuja alta frequência pode indicar que há mais considerações sendo feitas acerca de aspectos computacionais das ontologias do que inicialmente observado, embora também possa se referir à definição de formal em seu outro sentido, como postula Guarino (1995), referindo-se à forma dos entes, suas propriedades e relações, para além de sua simples existência dentro do domínio sendo representado.

Também foi utilizada a função *word sketch*, que permite gerar gráficos a partir de sentenças simples ou de modificadores que alteram um substantivo. Nesse sentido, o gráfico abaixo foi produzido.

[10]O conjunto de caracteres .* tem a função de deixar o restante do termo em aberto no *SketchEngine*, recuperando assim variações do termo desejado

Figura 3: Definições mais comuns para ontologia a partir da sentença *ontology-is-a*

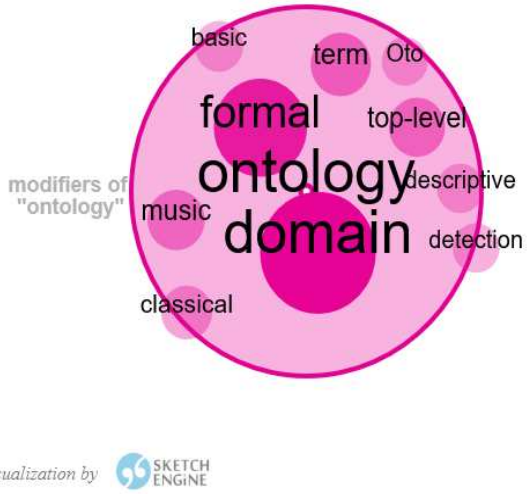


Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

O gráfico expõe quais substantivos aparecem como objeto da sentença *ontology* is a ("ontologia é uma…"); ou seja, expõe definições comuns para *ontology* no corpus estudado. Nesse sentido, pode-se observar que *specification* e *description* aparecem com mais frequência que outros substantivos; isso demonstra como, mesmo na CI, o cerne do entendimento acerca de ontologias se mantém próximo à Computação, onde elas são vistas como artefatos (outro termo que consta no gráfico) de delimitação de um contexto e redução de incertezas (Sowa, 2001).

Situação similar ocorre ao se buscar pelos adjetivos que atuam como modificadores de *ontology*, conforme gráfico a seguir.

Figura 4: Adjetivos mais comuns que modificam o *lemma ontology*



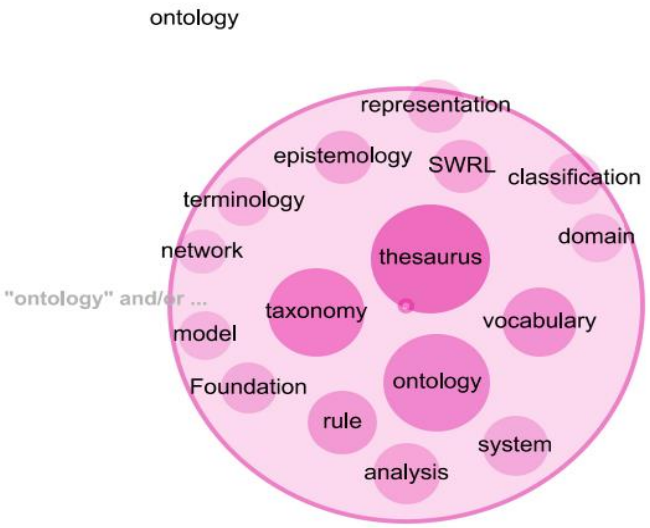
Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Aqui, podemos ver que *formal* é quase tão frequente quanto *domain*, o que demonstra que, embora prevaleçam discussões sobre ontologias aplicadas a domínios específicos, também figuram discussões acerca da natureza das ontologias em seu aspecto formal. Aqui, é possível que *formal* esteja se referindo tanto ao sentido de as ontologias servirem como modelos rigorosos para a estruturação de domínios do conhecimento dentro de contextos legíveis por máquinas quanto ao sentido de se referirem à forma dos entes dentro dos contextos que busca representar, conectando a presença deste termo com a de *domain*. O termo *descriptive* também pode servir como indício de que é este segundo sentido que figura com mais frequência no corpus estudado. Já a presença de *music* indica a aplicação das ontologias de domínio, tendo em vista a prevalência de modelos como a *Music Ontology*[11]. Finalmente, o restante dos termos (como *top-level*, *classical* e *basic*) aponta para indagações de caráter mais epistemológico.

[11]Ontologia disponível em: http://musicontology.com/. Acesso em: 8 ago. 2022.

O caráter interdisciplinar que as ontologias possuem podem ser observados no gráfico a seguir, que expõe termos que aparecem próximos a *ontology* em nosso corpus via interações e/ou.

Figura 5: Termos que acompanham o *lemma ontology* em relações *and/or*



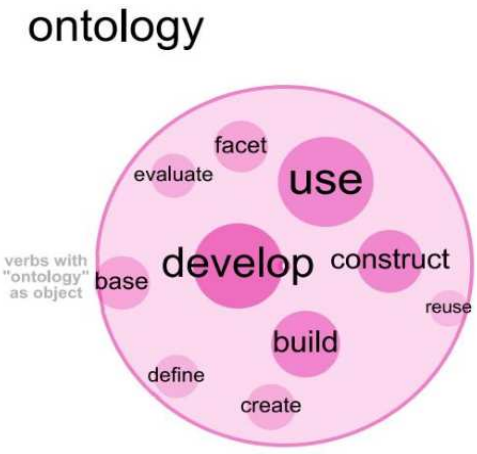
Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Inicialmente, podemos observar como as ontologias são percebidas como artefatos próximos com outros SOCs, como os tesauros e as taxonomias, representados por suas versões em língua inglesa; além disso, somam-se outros termos que remetem às atividades de Organização do Conhecimento, como *classification*, *epistemology* e *terminology*. A Epistemologia é elemento importante dentro da Organização do Conhecimento contemporânea, com suas preocupações acerca da natureza do conhecimento dependendo de como este conhecimento é abordado pelas áreas a qual é objeto de estudo; Hjørland (2008) aponta particularmente para a Análise de Domínio como sendo, dentro da OC, a abordagem que mais lida com problemas de caráter epistemológico em suas atividades de

Organização. Nesse sentido, a preocupação epistemológica manifesta-se nas atividades de construção de ontologias ao se considerar como diferentes especialistas veem dado domínio que se busque representar em um artefato ontológico. Já a relação das ontologias com o campo da Terminologia é ainda reforçado pela presença de *vocabulary*, e confirma-se na literatura através de contribuições como as de Santos e Costa (2015), que lançam mão de uma perspectiva terminológica que busque centralizar o sentido dos conceitos do saber especializado em termos discretos, que permitam a representação de um domínio do saber de determinado campo especializado. Finalmente, o restante dos termos indica aproximações de *ontology* com a Ciência da Computação, via *network*, *system*, *analysis* e *representation*, embora não destoem muito dos demais termos que mantêm *ontology* próximo à CI.

A percepção das ontologias como artefatos ou produtos de um esforço teórico-metodológico se reflete na figura 6, a seguir:

Figura 6: Verbos que têm *ontology* como objeto em suas sentenças

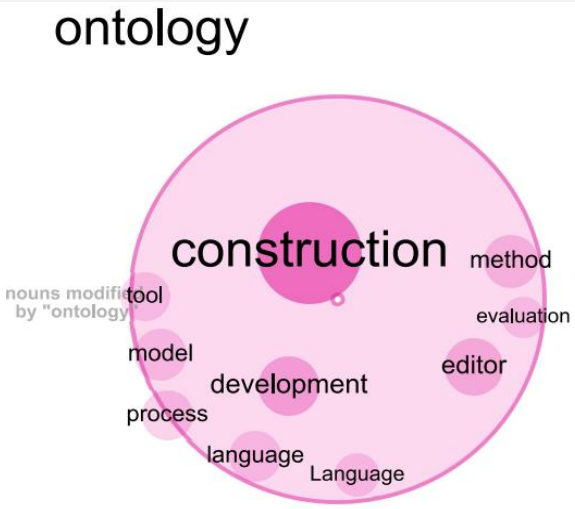


Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Aqui, excluindo *facet* e possivelmente *base*, que podem na verdade ser adjetivos entendidos como verbos pelo *SketchEngine*, o restante dos termos aponta para preocupações de eficiência na construção de ontologias, perspectiva importante tanto na CI quanto na Computação. *Develop*, *build*, *construct* e possivelmente *define* se referem todos à criação do artefato ontológico, em suas diferentes etapas. *Evaluate* e *reuse*, por sua vez, remetem tanto ao meio do ciclo de vida de um artefato ontológico quanto ao preparo que antecede a construção de uma nova iteração. Ontologias são, em maior ou menor grau, artefatos inerentemente iterativos, considerando que representam dado domínio em dado momento no tempo (Gonçalves, 2020), e a avaliação de sua eficiência na modelagem deve ser uma etapa de extrema importância em qualquer esforço de representação. Reuse, além de se referir à integração de ontologias já prontas na construção de novos artefatos, pode ainda se referir ao papel que as ontologias possuem no modelo da *Web* Semântica, facilitando a interoperabilidade entre diferentes aplicações.

Essa visão das ontologias enquanto produto é também consolidada pelos substantivos que são modificados por *ontology*.

Figura 7: Substantivos modificados por *Ontology*

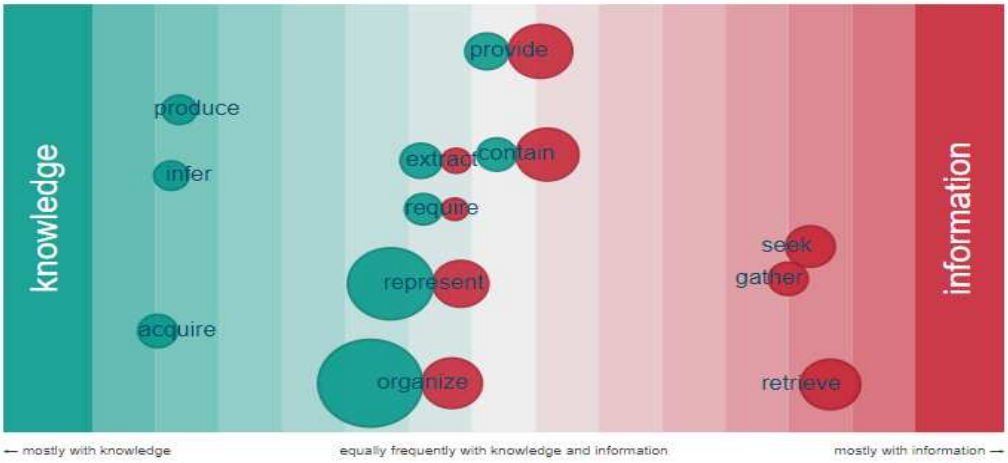


Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Aqui, tal como ocorreu nos resultados apresentados na Figura 6, estreitam-se as relações dos termos com o papel das ontologias na Computação. *Construction*, *development*, *processo*, *method*, *tool* e *editor* referem-se ao desenvolvimento dos artefatos ontológicos num todo, com editor podendo remeter tanto a *softwares* de construção de ontologias quanto ao profissional responsável por esta construção. Aqui, chama a atenção a presença de *language*, que provavelmente refere-se às linguagens de construção de ontologias, como o formato OWL (*Web Ontology Language*), embora também possa sugerir a compreensão das ontologias como próximas às linguagens documentárias comuns nas atividades de classificação e organização na CI.

Considerando sua importância para a temática de ontologia, buscamos realizar procedimentos junto ao *SketchEngine* envolvendo os termos “informação” e “conhecimento”. Dessa forma, a Figura 8, a seguir, mostra uma relação entre os verbos que possuem *knowledge* e os que possuem *information* como objetos em suas sentenças.

Figura 8: Verbos que possuem *knowledge* ou *information* como objeto em suas sentenças



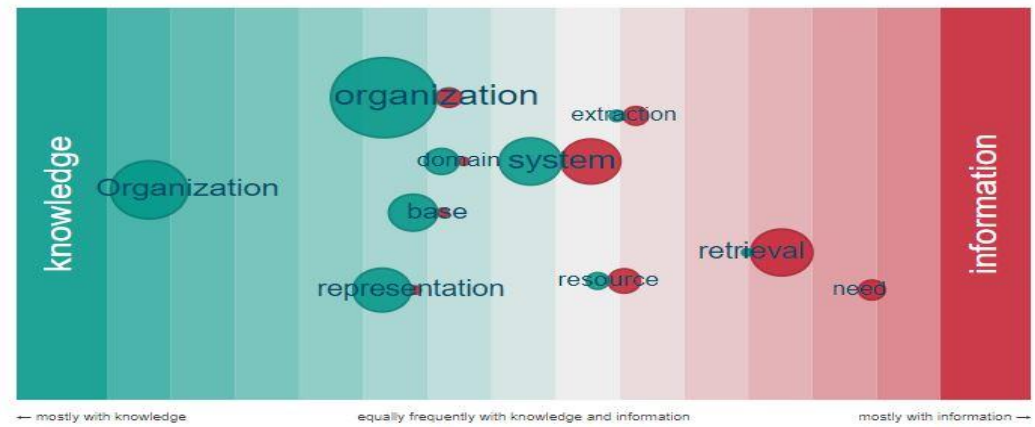
Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Aqui, podemos observar que os verbos mais frequentes junto a *information* (como *seek*, *gather*, *retrieve* e *provide*) possuem caráter mais prático, remetendo à função das ontologias como dispositivos de software que facilitam a busca e recuperação da Informação em contextos *web*. Verbos como *contain* e *provide*, que possuem mais ocorrências junto a *information* que a *knowledge*, repetem este padrão de preocupações práticas, possivelmente indicando que *information* surge no corpus aproximando as ontologias de seus usos na Computação. Os verbos que ocorrem apenas com *knowledge* indicam diferentes perspectivas: *infer* também remete ao uso das ontologias na Computação, considerando o papel destes artefatos no desenvolvimento de sistemas inteligentes capazes de operações de raciocínio via inferência. *Produce* e *acquire* podem indicar tanto inclinações à OC quanto à Computação, embora *acquire*, em particular, remete à aquisição de conhecimento que é etapa básica para a construção de ontologias e de SOCs num geral. Finalmente, *represent* e organize remetem tanto à Computação, quanto à OC, especialmente no

caso de organize, que também pode indicar discussões de caráter epistemológico, posto que o viés metafísico pelo qual se busca caracterizar e diferenciar os entes está no cerne da Ontologia clássica. Finalmente, *represent* pode vincular-se especificamente à construção de IAs na Computação, por exemplo.

A Figura 9 também apresenta comparações entre *knowledge* e *information*. Neste caso, expõe-se os substantivos que são alterados por um ou por outro termo.

Figura 9: Substantivos modificados por *knowledge* ou *information*



Fonte: elaborada pelos autores, dados extraídos com o *SketchEngine*.

Novamente, destaca-se o papel das ontologias como agentes de eficiência na recuperação da informação, conforme observável em (*information*) *retrieval*; nesse sentido, também destaca-se (*information*) *need* como termo que remete às necessidades informacionais. O termo *system*, no centro do gráfico, possui ocorrências quase equivalentes junto a *knowledge* e *information*. Nesse sentido, cabe especial destaque a (*knowledge*) *system*; considerando a função dos SOCs no ordenamento do saber especializado, em especial da sistematização das diferentes áreas do conhecimento humano em seções distintas (Currás, 2010), podemos inferir um destaque para o papel das ontologias nas atividades que compõem a Organização

do Conhecimento. Essa conclusão é reforçada pelo grande volume de ocorrências que *organization* possui junto a *knowledge* em comparação com *information.*

6 CONSIDERAÇÕES FINAIS

Considerando os objetivos deste trabalho de analisar como a temática de ontologias vem sendo trabalhada dentro do domínio de autores que publicam junto à *Knowledge Organization,* considera-se que os objetivos foram alcançados. Tomando por base o trabalho de Barros e Laipelt (2021), foi possível traçar estratégias de busca e visualização dos dados coletados. Dessa forma, constatou-se que a temática de ontologias de domínio ainda prevalece dentro da KO, estando a comunidade pautada pelo desenvolvimento das ontologias enquanto objetos-fim e mantendo preocupações fundamentais acerca da natureza do controle terminológico e dos processos de classificação. A presença de termos que remetem a questões de caráter epistemológico e teórico reforçam as aproximações da Ontologia Aplicada junto à Organização do Conhecimento, perspectiva que é reforçada ao considerar a presença de termos que remetem à Terminologia e a marcos teóricos da OC, como a aplicação da teoria da Classificação Facetada. Também o léxico utilizado em relação ao termo "*knowledge*" aponta para deliberações acerca de sua organização e representação, reforçando a noção apresentada de que, dentro da OC, as ontologias são compreendidas como SOCs. Somado à frequência com que surgem termos demonstrando ligações entre os artefatos ontológicos e outros SOCs, isso fornece pistas acerca de como as ontologias vêm sendo compreendidas junto à comunidade autora da KO. Ainda assim, as intersecções entre a OC e a Computação também puderam ser observadas, tanto a partir da prevalência de elementos lexicais de

caráter prático acerca do termo "*information*" quanto à proeminência do termo "*specification*" como definição para "*ontology*".

Desta forma, considera-se que o trabalho alcançou seus objetivos, expondo as tendências e perspectivas adotadas pela comunidade autora da KO em relação à temática das ontologias; em particular como as ontologias são compreendidas pela comunidade-domínio analisada e como as noções de conhecimento e informações são tiradas quando juntos da temática maior da Ontologia Aplicada à CI. Embora não tenha exaurido a temática, tampouco abordado o domínio estudado de todos os aspectos possíveis, este trabalho ainda se faz uma importante adição nas discussões da área, trazendo conclusões importantes acerca do que vem sido publicado na KO nos últimos dez anos, dentro da seara da Ontologia Aplicada e noções correlatas.

REFERÊNCIAS

ALMEIDA, Maurício Barcellos. **Ontologia em Ciência da Informação**: teoria e método. Curitiba: Editora CRV, 2020.

ALVES, Bruno Henrique; DALESSANDRO, Rafael Cacciolari; SANTOS, Fernanda Bochi Colaboração científica no periódico Knowledge Organization: elementos para caracterização de um domínio. *In*: CONGRESSO BRASILEIRO EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTo, 5., 2019, Belém. **Anais** [...]. Belém: ISKO Brasil, 2019. p. 137-144. Disponível em: https://brapci.inf.br/index.php/res/v/123278. Acesso em: 28 maio 2022.

BARROS, Thiago Henrique Bragato; LAIPELT, Rita do Carmo Ferreira. Uma análise de domínio da área de Organização e Representação do Conhecimento no contexto do periódico Em Questão. **Em Questão**, Porto Alegre, v. 27, n. 4, p. 438-468, 2021. Disponível em: https://seer.ufrgs.br/index.php/EmQuestao/article/view/116729. Acesso em: 26 maio 2022.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, *[s. l.]*, May 2001. Disponível em: http://www2.ic.uff.br/~bazilio/cursos/sistweb/material/Barners-Lee-Scientific-American-May-2001.pdf. Acesso em: 9 maio 2022

BIAGETTI, Maria Teresa. Ontologies as knowledge organization systems. *In*: HJØRLAND, Birger; GNOLI, Claudio (ed.). **ISKO Encyclopedia of Knowledge Organization**. *[S. l.]*: ISKO, 2021. Disponível em: https://www.isko.org/cyclo/ontologies. Acesso em: 28 maio 2022.

CAFÉ, Lígia; SANTOS, Raimundo Nonato Macedo dos; BARROS, Camila Monteiro. Os estudos de Gruber e Guarino sobre ontologias na ciência da informação e nas ciências da computação. **DataGramaZero**, *[s. l.]*, v. 16, n. 3, 2015. Disponível em: http://hdl.handle.net/20.500.11959/brapci/8169. Acesso em: 8 fev. 2023.

CURRÁS, Emilia. **Ontologies, taxonomies and thesauri in systems science and systematics**. Oxford: Chandos Publishing, 2010.

GONÇALVES, José Eugênio de Assis. **Método Ágil de integração semântica de dados científicos baseado em ontologias.** Tese (doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2020. Disponível em: http://hdl.handle.net/1843/34013. Acesso em: 26 abr. 2023.

GUARINO, Nicola. Formal ontology, conceptual analysis and knowledge representation. **International Journal of Human-Computer Studies**, *[s. l.]*, v. 43, n. 5-6, p. 625-640, 1995. Disponível em: https://www.sciencedirect.com/science/article/pii/S107158198571066X. Acesso em: 26 maio 2022.

HERRE, Heinrich; HELLER, Barbara; BUREK, Patryk; HOEHNDORF, Robert; LOEBE, Frank; MICHALEK, Hannes. **General Formal Ontology (GFO)**: Part I: Basic Principles. Leipzig, Alemanha: Institute of Medical Informatics: Statistics and Epidemiology: University of Leipzig, 2006. Disponível em: https://www.onto-med.de/sites/www.onto-med.de/files/files/uploads/Publications/2006/om-report-no8.pdf. Acesso em: 16 ago. 2022.

HJØRLAND, Birger. Domain Analysis. *In*: HJØRLAND, Birger; GNOLI, Claudio (ed.). **ISKO Encyclopedia of Knowledge Organization**. *[S. l.]*: ISKO, 2017. Disponível em: Disponível em: https://www.isko.org/cyclo/domain_analysis. Acesso em: 28 abr. 2023.

HJØRLAND, Birger. What is Knowledge Organization (KO)? **Knowledge Organization**, *[s. l.]*, v. 35, n. 2-3, p. 86-101, 2008. Disponível em: https://is.muni.cz/el/1421/jaro2016/VIKBA06/um/56249939/HJORLAND__Birger._What_is_knowledge_organization__KO_.pdf. Acesso em: 18 abr. 2023.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO). Knowledge Organization (KO). **ISKO**, *[s. l.]*, c2022. Disponível em: https://www.isko.org/ko.html. Acesso em: 1 maio 2022.

SANTOS, Monick Trajano dos. **Estudo do processo de apropriação da ontologia pela Ciência da Informação no Brasil.** 2014. 270 f. Dissertação (Mestrado em Ciência da Informação) – Centro de Artes e Comunicação, Universidade Federal de Pernambuco, Recife, 2014. Disponível em: https://repositorio.ufpe.br/handle/123456789/12945. Acesso em: 26 maio 2022.

SANTOS, Monick Trajano dos; CORREA, Renato Fernandes; LAPA, Remi Correia. Estudos sobre a apropriação da Ontologia pela Ciência da Informação. *In:* ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais** [...]. Florianópolis: UFSC, 2013. p. 1-7. Disponível em: https://brapci.inf.br/index.php/res/download/184410. Acesso em: 9 maio 2022.

SANTOS, Cláudia; COSTA, Rute. Domain specificity: semasiological and onomasiological knowledge representation. *In*: KOCKAERT, Hendrick; STEURS, Frieda (ed.). **Handbook of Terminology**. Amsterdam; Philadelphia: John Benjamins Publishing Company, 2015. p. 153-179.

SCHIESSL, Marcelo; BRÄSCHER, Marisa. Ontologia: ambiguidade e precisão. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 17, n. esp. 1, p. 125-141, 2012. DOI: https://doi.org/10.5007/1518-2924.2012v17nesp1p125. Disponível em:

https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17nesp1p125. Acesso em: 26 maio 2022.

SILVA, Manoela Ferreira da; EVANGELISTA, Isadora Victorino; GUIMARÃES, José Augusto Chaves. A presença da produção científica brasileira na revista Knowledge Organization no século XXI. **Informação & Informação**, Londrina, v. 24, n. 3, p. 28-51, 2019. DOI: https://doi.org/10.5433/1981-8920.2019v24n3p28. Disponível em: https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/38346. Acesso em: 1 maio 2022.

SMIRAGLIA, Richard P. Epistemology of Domain Analysis. *In*: SMIRAGLIA, Richard P.; LEE, Hur-Li (ed.). **Cultural frames of knowledge**. Wurzburg: Ergon, 2012. p. 111-124.

SOWA, John Floria. **Building, sharing, and merging ontologies**. *[S. l.: s. n.]*, 2001. Disponível em: http://www.jfsowa.com/ontology/ontoshar.htm. Acesso em: 26 maio 2022.

VICKERY, Brian Campbell. Ontologies. **Journal of Information Science**, v. 23, n. 4, 1997. DOI: https://doi.org/10.1177/016555159702300402. Disponível em: https://journals.sagepub.com/doi/10.1177/016555159702300402. Acesso em: 26 maio 2022.

AGRADECIMENTOS

Agradecemos à CAPES e à Universidade Federal do Rio Grande do Sul pelo financiamento e apoio a esta pesquisa.

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional.