

Jeanne Louize Emygdio

Universidade Federal de Minas Gerais
jeanne.emygdio@gmail.com

Maurício Barcellos Almeida

Universidade Federal de Minas Gerais
mba@eci.ufmg.br

REPRESENTAÇÕES FORMAIS DO CONHECIMENTO APLICADAS À INTEROPERABILIDADE SEMÂNTICA DE TERMINOLOGIAS CLÍNICAS

RESUMO

Apresenta-se uma pesquisa em andamento cujo propósito é o de demonstrar por meio de um experimento, que abordagens ontológicas podem ser melhores do que as de caráter puramente linguístico para fins de interoperabilidade semântica. A metodologia envolve uso de métodos léxicos e estruturais para combinar, mapear, alinhar e integrar terminologias clínicas, sob o enfoque linguístico e sob o enfoque ontológico. Prevê-se uma análise qualitativa dos dados, através da mensuração do grau de eficiência de cada abordagem para prover integração entre as terminologias clínicas. Espera-se obter contribuições teóricas que permitam: i) a descoberta e distinção dos aspectos ontológicos e epistêmicos inerentes às terminologias clínicas, ii) dos desafios epistêmicos à interoperabilidade semântica e iii) a relevância da abordagem ontológica para mitigar tais problemas. Em última instância, espera-se obter uma ontologia computacional de integração entre o SNOMED CT e a CID no âmbito obstétrico, a ser distribuída publicamente, em formato aberto e documentada.

Palavras-Chave: Interoperabilidade semântica, Terminologias clínicas, Representação do conhecimento, Ontologias formais.

ABSTRACT

We present here an ongoing research whose aim is to demonstrate, through an experiment, that ontological approaches may be better than purely linguistic approaches for semantic interoperability purposes. The methodology involves the use of lexical and structural methods to combine, map, align and integrate clinical terminology, under the linguistic approach and under the ontological approach. We perform a qualitative analysis of the results, by measuring the efficiency of each approach to provide integration between clinical terminologies. Theoretical contributions are expected to: i) to foster the discovery and distinction of the ontological and epistemic aspects inherent to clinical terminologies, ii) to check the epistemic challenges to semantic interoperability and, iii) to observe the relevance of the ontological approach to mitigate such problems. Eventually, we hope to develop a computational ontology of integration between SNOMED CT and ICD to be publicly distributed, in an open and documented format.

Keywords: Semantic interoperability, Clinical Terminologies, Knowledge Representation, Formal ontologies.

Universidade Federal de Minas Gerais

Correspondência/Contato
Av. Antônio Carlos, 6627
Pampulha: 31270-901
BELO HORIZONTE - MG

Escola de Ciência da Informação da UFMG

1. INTRODUÇÃO

A busca por interoperabilidade semântica em domínios complexos prescinde da capacidade de identificar e distinguir os elementos heterogêneos que compõem estes domínios, bem como as práticas de interoperabilidade utilizadas para garantir a recuperação da informação. O desenvolvimento de estratégias de articulação destes elementos visa o acesso consensual à dados igualmente heterogêneos, gerados durante os processos organizacionais, favorecendo assim o surgimento de inovações em benefício da sociedade.

Uma alternativa viável à organização do conhecimento nestes domínios remete aos estudos filosóficos relacionados às ontologias. Por meio destes estudos busca-se compreender a realidade dos seres, sua essência, princípios e causas (SEARLE, 1998), baseando-se em exaustivas classificações dos elementos do mundo real visando identificar seu tipo (categoria), estrutura, propriedades, eventos, processos e as inter-relações estabelecidas entre estes elementos (SMITH, 2003).

Modelos ontológicos baseados nestes estudos são concebidos com significativo grau de formalidade, em função do uso de axiomas¹ e lógica formal, sendo úteis para o refinamento de modelos de representação do conhecimento menos precisos. Sua conversão em artefatos computacionais, os caracterizam como complexos informacionais de alto nível, aptos portanto, a prover comunicação unívoca e autônoma entre humanos e máquinas, a despeito do domínio de conhecimento para o qual tenham sido projetados (GRUNINGER et al., 2008; SCHULZ e JANSEN, 2013; SMITH, 2003;2008; LEE e SIEGEL, 1996).

A concepção de estratégias baseadas no uso de ontologias para a articulação dos elementos heterogêneos presentes em ecossistemas de saúde poderá viabilizar o acesso consensual aos dados, mitigando os obstáculos à construção de Prontuários Eletrônicos de Pacientes (PEPs) capazes de melhor atender a sociedade. Tal é o objeto de reflexão desta pesquisa.

O presente artigo encontra-se organizado da seguinte forma: a seção 2 contextualiza a pesquisa no âmbito da literatura, a seção 3 caracteriza a pesquisa em termos metodológicos, e a seção 4 oferece algumas considerações possíveis no estágio atual da pesquisa.

¹ Proposições a serem definidas em uma base de conhecimentos para representar os fatos de um domínio (RUSSELL e NORVIG, 1995).

2. CONTEXTUALIZAÇÃO

Os ecossistemas de saúde constituem-se em um exemplo de domínios complexos e, assim como qualquer instituição, possuem duas dimensões básicas: a dimensão interna (microambiente) e a dimensão externa (macroambiente).

No microambiente as comunidades de profissionais da saúde e profissionais técnicos realizam as práticas estabelecidas a partir de processos de negócios sustentados por arquiteturas tecnológicas, visando uma convergência para atendimento aos propósitos das instituições de saúde. Três pilares sustentam esta dimensão: pessoas, processos e tecnologias.

No macroambiente encontram-se uma série de aspectos dignos de nota: i) uma comunidade composta por consumidores dos serviços de saúde (pacientes), os fornecedores de produtos e serviços de apoio às práticas internas da instituição, os parceiros, os concorrentes e os órgãos reguladores; e ii) os atores políticos, econômicos, científicos e a comunidade internacional cujas atividades e descobertas provocam impactos nas práticas do microambiente. Tais impactos ocorrem por exemplo, através das modificações legais e tributárias, da ampliação ou redução de investimentos, nos aspectos de adoção de normas internacionais de regulamentação ou na descoberta de inovações (ESLAMI ANDARGOLI et al., 2017; GUTIERREZ; MORENO; REBELO, 2017; MOBIN et al., 2018; SALUSE et al., 2010; SCHULZ; STEGWEE; CHRONAKI, 2019).

A comunicação humana dentro do ecossistema é norteada tanto pela linguagem, legislação e contratos, quanto pelo jargão clínico e padrões terminológicos adotados. A comunicação técnica, por sua vez é norteada por linguagens as vezes artificiais e especializadas, além de padrões técnicos, modelos de representação, acordos organizacionais e acordos sobre conteúdos (FREITAS; SCHULZ; MORAES, 2009; SCHULZ et al., 2017).

Como resultado da interação entre os atores durante as práticas de saúde, um grande volume de dados heterogêneos é gerado, distinguindo-se quanto à organização em: estruturados e não estruturados; quanto à natureza em: imagens, sons, vídeos e mensagens e quanto ao estado em: incompletos, incorretos, ambíguos, vagos, em silos, etc. (ARP; SMITH; SPEAR, 2015; SMITH, 2008).

Várias alternativas de interoperabilidade concorrem dentro do ecossistema para que os diversos recursos de informação possam cooperar de forma eficiente e produtiva, tais como: i) a interoperabilidade técnica, ii) semântica, iii) organizacional,

política e humana, iv) inter-comunitária e v) internacional (UKOLN, 2015). No entanto, tais práticas são muitas vezes pouco conhecidas e até mesmo pouco compreendidas (ABNT, 2016; SCHULZ *et al*, 2017). A visão de um ecossistema de saúde é apresentada na Figura 1.

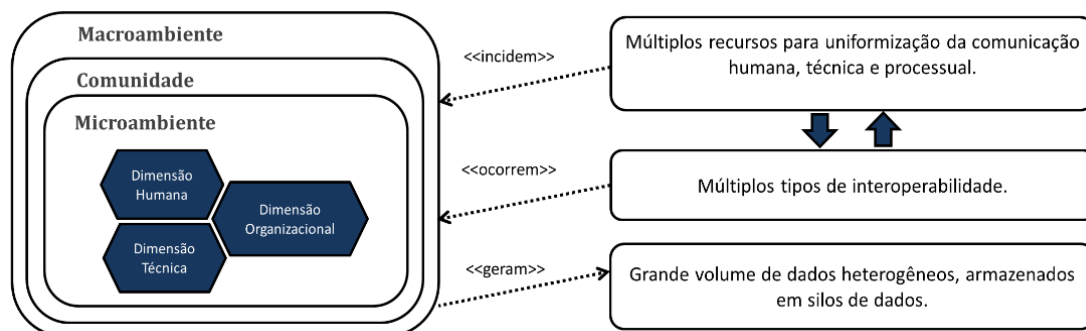


Figura 1. Ecossistemas de saúde

Com o propósito de uniformizar a comunicação no contexto médico, diversas organizações internacionais provêm padrões, destacando-se dentre elas, a *International Organization for Standardization (ISO)*², a *Health Level Seven International (HL7)*³, o *European Committee for Standardization (CEN)*⁴, a *International Health Terminology Standards Development Organisation (IHSTDO)*⁵, a *openEHR Foundation*⁶, e o *IHE International Committee (IHE)*⁷. Por esta razão, diversos padrões concorrem para alcance dos mesmos fins, como por exemplo: *Systematized Nomenclature of Medicine–Clinical Terms (SNOMED CT)*⁸, *International Classification of Diseases (ICD)*⁹, HL7, openEHR¹⁰, *Digital Imaging and communications in medicine (DICOM)*¹¹, *Logical Observation Identifiers Names and Codes (LOINC)*¹², dentre outros.

Em decorrência deste fato, são comuns sobreposições semânticas e sintáticas recorrentes, às quais, associadas à natureza distinta de cada padrão, comprometem a interoperabilidade no contexto. A título de ilustração, apresenta-se na Tabela 1, uma visão geral do número de padrões disponíveis no repositório de padrões biomédicos Bioportal, mantido pelo *National Center for Biomedical Ontology (NCBO)*¹³.

² Maiores informações em: <https://www.iso.org/home.html>. Acesso em: 16/09/2019.

³ Maiores informações em: <https://www.hl7.org/>. Acesso em: 16/09/2019.

⁴ Maiores informações em: <https://www.cen.eu/Pages/default.aspx>. Acesso em: 16/09/2019.

⁵ Maiores informações em: <https://www.ihtsdo.org/>. Acesso em: 16/09/2019.

⁶ Maiores informações em: <https://www.openehr.org/>. Acesso em: 16/09/2019.

⁷ Maiores informações em: <https://www.iheusa.org/ihe-international>. Acesso em: 16/09/2019.

⁸ Maiores informações em: <http://www.snomed.org/>. Acesso em: 16/09/2019.

⁹ Maiores informações em: <https://www.who.int/classifications/icd/en/>. Acesso em 16/09/2019.

¹⁰ Maiores informações em: https://www.openehr.org/what_is_openehr. Acesso em: 06/07/2019.

¹¹ Maiores informações em: <https://www.dicomstandard.org/>. Acesso em: 08/02/2019.

¹² Maiores informações em: <https://loinc.org/>. Acesso em: 08/02/2019.

¹³ Leland Stanford Junior University, California, USA.

Tabela 1. Visão geral do número de padrões biomédicos disponíveis.

Padrões biomédicos	Nº de Padrões hospedados	Tipos de estruturas
Repositório BioPortal	992	Taxonomias, Tesouros, Meta-tesouros, Frameworks, Padrões Técnicos e Ontologias.

Devido à impossibilidade de um único padrão atender a todas as especialidades de uma instituição de saúde, abre-se uma demanda natural de integração dos diversos padrões existentes, a despeito da falta de entendimento comum sobre este fato (ACCESS CT, 2016; ABNT, 2016).

As ontologias oferecem alternativas viáveis para soluções de interoperabilidade semântica em contextos diversos. Por incluírem relações hierárquicas, ou seja, taxonomias, classificações e vocabulários, sustentam numerosas aplicações como a busca semântica, integração de dados, mineração de textos, inferências e anotação semântica de recursos, elementos imprescindíveis às práticas de interoperabilidade esperadas. Tal fato justifica a delimitação da pesquisa aqui apresentada.

2.1 Delimitação da pesquisa

A presente pesquisa investiga questões capazes de gerar impactos negativos sobre a busca por interoperabilidade semântica no contexto médico, sendo portanto estabelecidas três premissas básicas:

Premissa 1: padrões terminológicos caracterizam-se como instrumentos para representação e organização do conhecimento, recebendo no contexto médico, denominações diversas, como: cabeçalho de assunto, classificação, nomenclatura, vocabulário, ontologias, terminologias biomédicas, entre outros (TEIXEIRA e ALMEIDA, 2019). Está além dos propósitos desta pesquisa discutir tais denominações, adotando-se o termo genérico "terminologias clínicas" (SCHULZ et al, 2017), e fazendo-se distinção por variantes quando necessário.

Premissa 2: a Ciência da Computação desenvolve volume expressivo de pesquisas sob este enfoque (OCHIENG e KYANDA, 2018; OTERO-CERDEIRA, RODRÍGUEZ-MARTÍNEZ e GÓMEZ-RODRÍGUEZ, 2015) buscando prover técnicas e ferramentas de alinhamento semântico de terminologias médicas (EUZENAT e SHVAIKO, 2007; NOY, 2009); a Ciência da Informação dá ênfase a descoberta e a distinção dos aspectos ontológicos e epistêmicos que permeiam as classificações, aos desafios que a epistemologia impõe ao alcance de interoperabilidade semântica (RECTOR, 1999; BODENREIDER, SMITH e BURGUN, 2004) e a forma como os fundamentos on-

tológicos podem ser empregados para mitigar tais problemas (JANSEN, 2008; SMITH, 2003).

Premissa 3: duas abordagens são relevantes para os propósitos de integração de terminologias clínicas: i) a de mapeamento retrospectivo, utilizada pelo Sistema de Linguagem Médica Unificado (UMLS), baseada na identificação de relações de sinonímia entre termos constituintes de seu acervo, gerando benefícios para a indexação e recuperação de documentos. Em contrapartida, existe uma demanda contínua de manutenção e evolução dos mapeamentos em função da ausência de uma arquitetura comum entre as fontes de vocabulários e; ii) a de padronização prospectiva, desenvolvida pela OBO *Foundry* (SMITH et al., 2007) através do BioPortal aderente a princípios compartilhados para construção, distribuição e evolução de ontologias.

Estudos relacionados à adoção dos princípios ontológicos em terminologias clínicas são encontrados na literatura científica, tendo por foco a melhoria na representação do conhecimento a partir da:

- i. correção de anomalias estruturais, definição de conceitos, relações e meta-classes (RECTOR, BRANDT e SCHNEIDER, 2011; SCHULZ; SUNTISRIVARAPORN e BAADER, 2007; BODENREIDER; SMITH e BURGUN, 2004);
- ii. melhoria na eficiência para descoberta de combinações terminológicas e economia de recursos de processamento em função do alinhamento a partir de ontologias de alto nível (ZHANG e BODENREIDER, 2005);
- iii. melhoria na identificação de fronteiras de conhecimento compartilhado entre terminologias clínicas visando sua harmonização para fins de mapeamento e evolução, tendo por base a definição de axiomas em linguagem OWL (RODRIGUES et al, 2015);
- iv. aspectos relacionados à avaliação de qualidade e interoperabilidade de terminologias clínicas indicam que o uso de técnicas híbridas, ou seja, as que usam métodos léxicos aliados a métodos estruturais, são efetivos para a detecção de erros na SNOMED CT, além de sugestivos para correção. Além disso, os aspectos de governança são relevantes para a definição e uso concorrente de terminologias clínicas de um ponto de vista de gestão (BODENREIDER, 2018; SCHULZ et al 2019);

- v. definição de métodos para a reorganização da classificação internacional de doenças (TEIXEIRA e ALMEIDA, 2019).

Apresenta-se então a pesquisa em andamento cujo objetivo é **demonstrar** por meio de uma **prova de conceito** que abordagens ontológicas podem ser mais efetivas para fins de **interoperabilidade semântica** se comparadas com as que adotam práticas puramente linguísticas. Atualmente, a pesquisa ainda se encontra em uma fase qualitativa, motivo pelo qual se definiu exatamente em que aspectos tal “efetividade” vai se manifestar.

A metodologia envolve uso de métodos léxicos e estruturais para combinar, mapear, alinhar e integrar terminologias clínicas, precedidos ou não do escrutínio ontológico. Prevê-se uma **análise qualitativa dos dados**, através da observação do grau de eficiência de cada abordagem para prover integração entre as terminologias clínicas.

Espera-se obter **contribuições teóricas** e práticas que permitam a descoberta e distinção dos aspectos ontológicos e epistêmicos inerentes às terminologias clínicas, destaquem os desafios epistêmicos à interoperabilidade semântica e demonstrem a relevância da abordagem ontológica para mitigar problemas de interoperabilidade. Espera-se obter uma ontologia computacional de integração entre o SNOMED CT e a CID no âmbito da obstetrícia, a ser documentada e distribuída publicamente em formato aberto.

3. PESQUISA EM ANDAMENTO

A presente pesquisa busca respostas para a seguinte questão: **de que forma ontologias podem contribuir para resolver situações que impedem a interoperabilidade semântica?** O objetivo geral é testar a integração de terminologias clínicas a partir de uma abordagem ontológica e a partir de uma abordagem linguística, verificando qual é a mais efetiva para fins de interoperabilidade semântica.

Como objetivos específicos busca-se:

- i. **Identificar e distinguir os aspectos ontológicos e epistêmicos** presentes nas terminologias clínicas;
- ii. **identificar âncoras léxicas** que estabeleçam fronteiras compartilhadas de conhecimento entre as terminologias clínicas;

- iii. **identificar relações semânticas** que possibilitem uma representação explícita do conhecimento que permeia as conexões entre as âncoras léxicas através da extração de relações hierárquicas;
- iv. **identificar âncoras estruturais** através da comparação das relações semânticas previamente identificadas;
- v. **demonstrar o grau de efetividade** alcançado por cada tipo de abordagem, tendo como base o número de combinações terminológicas encontradas aliado ao percentual de evidências positivas para interoperabilidade que estes resultados representam.

3.1 Caracterização da pesquisa

A pesquisa é de **natureza** aplicada visando gerar conhecimentos para realização prática de interoperabilidade de dados que demandem a integração das terminologias clínicas SNOMED CT e CID na área obstétrica.

A **forma de abordagem** ao problema é qualitativa tendo em vista que a relação entre o mundo (práticas clínicas) e o sujeito (terminologias clínicas) não pode ser expressa em números.

Em relação aos **objetivos** a pesquisa é **explicativa** e **experimental** uma vez que objetiva identificar a relação terminológica entre o SNOMED CT e a CID para atendimento à demanda de cruzamento de dados, determinando por fim as representações terminológicas formais capazes de refiná-las. A caracterização desta pesquisa baseia-se em Gil (1989) e Marconi e Lakatos (2003).

3.2 Coleta de dados

O universo da pesquisa compreende duas terminologias clínicas já mencionadas, e a unidade de análise compreende os conjuntos de termos relacionados à área obstétrica e neonatal existentes nas terminologias clínicas SNOMED CT, CID e as demais vinculadas à iniciativa *OBO Foundry*. Tal forma de coleta de dados, caracteriza-se como a documental (MARKONI e LAKATOS, 2003).

O domínio obstétrico e neonatal foi selecionado por conveniência, seja por conta de pesquisas anteriores, seja pelo acesso a um Depto. de Obstetrícia.

Foram identificados 32 terminologias clínicas e três repositórios de ontologias de interesse para a pesquisa, conforme apresentado na Tabela 2.

Tabela 2. Terminologias clínicas e repositórios pesquisados.

Qtd.	Tipo	Quais	Semântica
5	Taxonomias	ATC/DDD, CPT-4, DSM, HCPCS, ICD/CID, NANDA, ISO9999	Linguagem natural
6	Tesauros	ICECI, ICF, ICHI, MedDRA, MeSH, NCIt	Linguagem natural
1	Meta-tesauro	SNOMED CT	Mista (linguagem natural e formal)
3	Frameworks	openEHR, HL7, IHE-PIX	Formal
4	Padrões técnicos e terminológicos	TISS, DICOM, LOINC, ISBT 128	Mista (linguagem natural e formal)
13	Ontologias	BWW, DOLCE, UFO, GFO, SOWA, BFO, OntoNeo, GO, CL, FMA, IAO, OBI, OGMS.	Formal
3	Repositórios de ontologias	OBO Foundry, UMLS, Biportal	Formal

Selecionou-se a SNOMED CT, a CID e a OntoNEO (ALMEIDA e FARINELLI, 2017) em função da perspectiva de contribuição social real ao país a partir da: i) possibilidade de replicação do experimento base (ZHANG e BODENREIDER, 2005); ii) da maturidade do repositório *OBO Foundry* em construção de ontologias biomédicas; iii) da ortogonalidade da ontologia OntoNEO ao repositório mencionado e de sua tradução para o português, e, iv) da presença de especialistas para validação dos resultados.

Foram identificadas até o momento conjuntos de dados estratificados por classes e propriedades, compreendendo: i) 3 capítulos (XV, XVI, XVII); 222 categorias (O00-O99; P00-P96; Q00-Q99); 1.381 sub-categorias da CID versão 10; ii) 1.767 classes e 471 propriedades da OntoNEO e iii) estudos em andamento sobre a SNOMED CT.

O processo de execução, em aderência aos objetivos específicos, encontra-se representado na Figura 2.

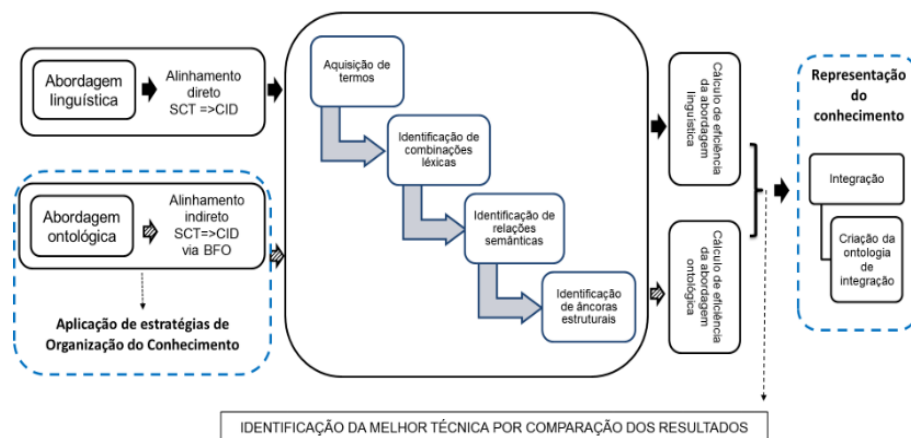


Figura 2. Processo de execução da pesquisa

As duas abordagens terão sua efetividade calculada como uma medida da proporção de evidências positivas para alinhamentos sobre o número total de combinações encontradas, conforme a equação:

$$\text{"Eficiência} = (\text{número de evidências positivas para alinhamentos} \div \text{total de combinações encontradas}) \times 100\text{"}$$

Apresenta-se a seguir, um exemplo da identificação e distinção de elementos ontológicos e epistêmicos encontrados no estudo das terminologias e os mapeamentos léxicos identificados:

Na SNOMED CT, encontra-se a seguinte doença:

"Malformation of central nervous system of fetus".

Tradução: Má-formação do sistema nervoso central de fetos.

Na CID-10 Versão: 2016, encontra-se:

"Maternal care for (suspected) central nervous system malformation in fetus".

Tradução: Cuidados maternos para (suspeita) de má-formação do sistema nervoso central de fetos.

O termo ("*suspected*"), presente na definição da CID, cumpre naquela terminologia um papel específico dentro da gramática composicional, visando representação do conhecimento a partir da identificação de conceitos elementares, análogos aos universais ontológicos, baseados na linguística e na teoria da semiótica. Representa uma percepção do profissional, proveniente de seu conhecimento e experiência sobre um possível diagnóstico.

Analisando o mesmo termo sob a ótica ontológica, visando a construção de representações formais para integração das terminologias, percebe-se a inexistência de evidências factuais que possibilitem a construção de axiomas capazes de validar a definição original proposta na CID. Desta forma, em se tratando de uma informação relevante para a prática médica, sua classificação deverá ser modificada para outro nível de representação, como o de anotações relacionadas às ontologias, por exemplo.

Os demais termos podem ser ontologicamente classificados a partir da BFO da seguinte forma:

- i. Malformação ("*Malformation*") = instância de uma disposição ("*instance of disposition*"). Representa a concretização da disposição do organismo humano de desenvolver uma doença;

- ii. sistema nervoso central ("*central nervous system*") = instância de um local ("*instance of site*"). Representa o local de desenvolvimento da doença;
- iii. feto ("*fetus*") = instância de uma entidade material ("*object*"). Representa um objeto do mundo real que materializa uma entidade material.

Assim, basicamente três âncoras léxicas foram identificadas entre as terminologias em estudo: *malformation*, *central nervous system* e *fetus*. Os mapeamentos resultantes poderiam ser assim representados:

$$A_1T_1, A_2T_2, r$$

$$\langle \text{"Malformation"}, \text{"malformation"}, \equiv \rangle$$

$$\langle \text{"central nervous system"}, \text{"central nervous system"}, \equiv \rangle$$

Onde A_1T_1 é a âncora léxica 1, identificada na terminologia 1; A_2T_2 é a âncora léxica 2, identificada na terminologia 2; e r é a relação entre elas, que pode ser, conforme abaixo (simbologia da teoria dos conjuntos):

equivalência (\equiv), subsunção (\sqsubseteq) ou disjunção ($A \sqcap B \sqsubseteq \perp$)

Os próximos passos da pesquisa, em replicação ao experimento original, compreendem a identificação das relações semânticas entre os termos e sua representação formal; a identificação de âncoras estruturais, ou fronteiras de conhecimento entre as terminologias, tendo por base a identificação de âncoras vizinhas, que compartilhem relações de mesmo tipo. Ao término das análises, três tipos de mapeamentos serão revelados: i) os que não apresentem evidências para alinhamento das terminologias; ii) os que evidenciem positivamente aspectos de alinhamento e iii) os que evidenciem aspectos negativos de alinhamento, o que poderá representar anomalias de representação do conhecimento nas terminologias em estudo.

Como exemplo, supondo que os resultados finais alcançados fossem os demonstrados na Tabela 3 a seguir, seria possível comprovar que os alinhamentos indiretos de BFO-CID e BFO-SCT foram os mais eficientes em função do percentual de evidências positivas encontradas, 94,8% e 90,1% respectivamente. A Tabela 3 apresenta uma simulação dos cálculos de eficiência esperados:

	SCT-CID 715 combinações	BFO-CID 1.353 combinações	BFO-SCT 2.173 combinações
Sem evidência	62 (8,7%)	66 (4,9%)	205 (9,4%)
Evidência positiva	653 (91,3%)	1.283 (94,8%)	1.958 (90,1%)
Evidência negativa	0	4 (0,3%)	10 (0,5%)

4. CONSIDERAÇÕES FINAIS

O diferencial da pesquisa reside na oportunidade de explicar como elementos epistêmicos presentes nas terminologias clínicas conduzem a diversos obstáculos às práticas de alinhamento, mapeamento, integração e evolução.

Além disso, espera-se demonstrar que o uso de princípios ontológicos, uma vez que são usados para orientar a construção de taxonomias livres de erros, possibilita um aumento nas chances de descobertas precisas de âncoras léxicas, de relações semânticas e âncoras estruturais que caracterizem fronteiras compartilhadas de conhecimento entre as terminologias clínicas.

Assim, torna-se possível uma economia de tempo e de recursos computacionais de processamento para a definição dos alinhamentos, tendo em vista que as comparações serão realizadas apenas entre termos ontologicamente identificados.

O artefato ontológico poderá ser utilizado de forma embarcada em sistemas de informação, visando prover interoperabilidade semântica sobre os dados dos pacientes, adquiridos durante os encontros médicos das mais diversas especialidades.

AGRADECIMENTOS

A pesquisa é subsidiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERÊNCIAS

ACCESS CT. **WP1 D1.4 Current and Future Use of SNOMED CT: Assessing SNOMED CT for Large Scale eHealth Deployments in the EU.** [s.l.] ASSESS CT, 2016. Disponível em:

<http://access-ct.eu/fileadmin/assess_ct/deliverables/final_submissions/assess_ct_ga_643818_d1.4.pdf>.

Acesso em: 13 mar. 2019.

ALMEIDA, M. B.; FARINELLI, F. (2017) Ontologies for the representation of electronic medical records: The obstetric and neonatal ontology. **Journal of the Association for Information Science and Technology.**

ARP, R.; SMITH, B.; SPEAR, A. D. **Building Ontologies with Basic Formal Ontology.** Cambridge, Massachusetts: The MIT Press, 2015.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT ISO/TR 12300** Informática em saúde: princípios para mapeamento entre sistemas terminológicos. 1. ed. Rio de Janeiro: ABNT, 2016.

BODENREIDER, O. **Evaluating the Quality and Interoperability of Biomedical Terminologies.** [s.l.] LHNCBC Board of Scientific Counselors, 2018.

- BODENREIDER, O., SMITH, B., BURGUN, A.. The Ontology-Epistemology Divide: A Case Study in Medical Terminology. **Formal ontology in information systems: Proceedings of the International Conference. FOIS (Conference)**, v.2004, p.185.
- ESLAMI ANDARGOLI, A. et al. Health information systems evaluation frameworks: A systematic review. **International Journal of Medical Informatics**, v. 97, p. 195–209, jan. 2017
- EUZENAT, J.; SHVAIKO, P.. **Ontology Matching**. Berlin, Heidelberg: Springer-Verlag, 2007.
- FREITAS, F.; SCHULZ, S.; MORAES, E. Survey of current terminologies and ontologies in biology and medicine. **RECIIS**, v. 3, n. 1, 11 mar. 2009.
- GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 2nd. ed. São Paulo: Atlas, 1989.
- GRUNINGER, M. et al. Ontology summit 2007 - ontology, taxonomy, folksonomy: Understanding the distinctions. **Applied Ontology**, 2008.
- GUTIERREZ, M. A.; MORENO, R. A.; REBELO, M. S. Information and Communication Technologies and Global Health Challenges. In: **Global Health Informatics**. [s.l.] Elsevier, 2017. p. 50–93.
- JANSEN, L.. Chapter 7: Classifications. In: **Applied Ontology: An Introduction [Hardcover]**. 8. ed. German: Katherine Munn e Barry Smith, 2008. v. 9p. 342, 2008.
- LEE, J. L.; SIEGEL, M. D. An ontological and semantical approach to source-receiver interoperability. **Decision Support Systems**, v. 18, n. 2, p. 145–158, 1 out. 1996.
- MARCONI, M. DE A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5a. ed. São Paulo: Atlas, 2003.
- MOBIN, N. S. M. et al. Factors Associated in the Implementation and Adoption of Electronic Health Records (EHR) in Private Healthcare. **International Journal of Engineering & Technology**, v. 7, n. 3.7, p. 257, 4 jul. 2018.
- NOY, N. F. Ontology Mapping. In: **Handbook on Ontologies**. 2nd ed. Springer-Verlag, 2009, p. 573–590.
- OCHIENG, P.; KYANDA, S.. Large-Scale Ontology Matching. **ACM Computing Surveys**, v. 51, n. 4, p. 1–35, 2018.
- OTERO-CERDEIRA, L.; RODRÍGUEZ-MARTÍNEZ, F. J.; GÓMEZ-RODRÍGUEZ, A. Ontology matching: A literature review. **Expert Systems with Applications**, 2015.
- RECTOR, A. L. Clinical terminology: why is it so hard? **Methods of Information in Medicine**, v. 38, n. 4–5, p. 239–252, dez. 1999.
- RECTOR, A. L.; BRANDT, S.; SCHNEIDER, T. Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. **Journal of the American Medical Informatics Association : JAMIA**, v. 18, n. 4, p. 432–40, 2011
- RODRIGUES, J-M., ROBINSON, D., MEA, V. D., CAMPBELL, J., RECTOR, A., SCHULZ, S., BREAR, H., ÜSTÜN, B., SPACKMAN, K., CHUTE, C. G., MILLAR, J., SOLBRIG, H., PERSSON, K. B. Semantic Alignment between ICD-11 and SNOMED CT. **Studies in Health Technology and Informatics**, v. 216, p. 790–794, 2015.
- SALUSE, J. et al. **Assessing the Economic Impact/Net Benefits of the Estonian Electronic Health Record System**. Tallinn: European Union European Social Fund, 2010.
- SEARLE, J.. **Mind, Language and Society: Philosophy in the real world**. New York: Basic Books, 1998.
- SCHULZ, S., SUNTISRIVARAPORN, B., BAADER, F.. SNOMED CT's problem list: Ontologists' and logicians' therapy suggestions. **Studies in health technology and informatics**, v. 129, p. 802–806, 2007.
- SCHULZ, S.; JANSEN, L. Formal Ontologies in Biomedical Knowledge Representation. **Yearbook of Medical Informatics**, v. 22, n. 01, p. 132–146, ago. 2013.

SCHULZ, S., RODRIGUES, J.-M., RECTOR, A., CHUTE, C. G. Interface Terminologies, Reference Terminologies and Aggregation Terminologies: A Strategy for Better Integration. **Studies in health technology and informatics**, v. 245, p. 940–944, 2017.

SCHULZ, S., STEGWEE, R., CHRONAKI, C. Standards in Healthcare Data. In: Kubben, P.; Dumontier, M.; Dekker, A. (Eds.). **Fundamentals of Clinical Data Science**. Cham: Springer International Publishing, 2019. p. 19–36, 2019.

SMITH, B.. Ontology. In: **The Blackwell Guide to the Philosophy of Computing and Information**. Oxford: Luciano Floridi, 2003.

SMITH, B.. Ontology (Science). Formal Ontology in Information Systems: **Proceedings of the Fifth International Conference (FOIS 2008)**. Anais: 170. In: Formal Ontology In Information Systems Frontiers In Artificial Intelligence And Applications. IOS Press, 2008.

SMITH, B., ASHBURNER, M., ROSSE, C., BARD, J., BUG, W., CEUSTERS, W., GOLDBERG, L. J., EIL-BECK, K., IRELAND, A., MUNGAL, C. J., LEONTIS, N., ROCCA-SERRA, P., RUTTENBERG, A., SANSONE, S., SCHEUERMANN, R.H., SHAH, N., WHETZEL, P. L., LEWIS, S. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. **Nature biotechnology**, v. 25, n. 11, p. 1251, 2007.

TEIXEIRA, L. M. D.; ALMEIDA, M. B. Princípios ontológicos no suporte a terminologias clínicas: método e ontologia para reorganização da Classificação Internacional de Doenças. **Ciência da Informação**, v. 48, n. 1, 3 maio 2019.

UKOLN. (2015) **UKOLN - Interoperability Focus - About**. UKWA -UK Web Archive.

Disponível em:

<<https://www.webarchive.org.uk/wayback/en/archive/20150806095947/http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>>. Acesso em: 25 fev. 2019.

ZHANG, S.; BODENREIDER, O.. Alignment of multiple ontologies of anatomy: deriving indirect mappings from direct mappings to a reference. **AMIA ... Annual Symposium proceedings. AMIA Symposium**, p. 864–868, 2005.

Jeanne Louize Emygdio

Doutoranda em Gestão e Organização do Conhecimento (UFMG). Mestre em Desenvolvimento, Tecnologias e Sociedade (UNIFEI). Especialista em Produção de Software Livre (UFLA), em Análise, Projeto e Gerenciamento de Sistemas (PUC-Minas) e em Design Instrucional para EaD (UNIFEI). Tecnóloga em Processamento de Dados (FUMEC). Áreas de interesse em pesquisa: Representação do Conhecimento em ontologias com foco em interoperabilidade semântica.

Maurício Barcellos Almeida

Pesquisador da Informação em Saúde e Professor Associado da UFMG; Doutor em Ciência da Informação; Bolsista de Produtividade CNPQ; pós-doutorado na *State University of New York*; pós-doutorado na Faculdade de Medicina, *Senior Visiting Fellow* no *College of Medicine* da *University of Arkansas for Medical Sciences*; membro

do Ontology Research Group do Center of Excellence in Bioinformatics and Life Sciences at Buffalo.