

Composicionalidade e sobreposição em terminologias biomédicas: alternativas para interoperabilidade em saúde

Livia Marangon Duffles Teixeira

Doutora; Fundação Centro de Pesquisa e Desenvolvimento em Telecomunicações, Campinas, SP, Brasil;
liviamarangon@gmail.com

Mauricio Barcellos Almeida

Doutor; Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil;
priv.mba@gmail.com

Resumo: Uma importante fonte de informação para a prática clínica é o prontuário de paciente, um documento que já existe no formato eletrônico. A ambiguidade inerente ao discurso médico tem dificultado a integração automática desejável entre prontuários eletrônicos de pacientes uma vez que, no atual estágio de desenvolvimento tecnológico, os sistemas automáticos não estão aptos a lidar com variações sintáticas e semânticas da linguagem humana. Mesmo as terminologias padronizadas, criadas para lidar com tais variações, não resolvem o problema à medida da multiplicidade de opções. Essas, além dos objetivos e propósitos diversos, têm resultado em redundância, retrabalho e sobreposição entre termos. O presente artigo parte da premissa de que o problema da interoperabilidade é complexo e precisa ser atacado em diversas frentes. Uma dessas frentes, no âmbito dos prontuários eletrônicos de pacientes, envolve situações de caráter conceitual, a saber, a sobreposição epistemológica e a composicionalidade de termos. Vale-se aqui de um modelo baseado em ontologias criado em pesquisa anterior para projetar e testar um método que objetiva atenuar os efeitos dessas situações. O teste é realizado em fragmento da Classificação Internacional de Doenças e conclui-se que as situações de natureza conceitual mencionadas são aspectos relevantes a combater na busca por interoperabilidade.

Palavras-chave: Ontologia. Representação do conhecimento. Terminologias clínicas. Interoperabilidade.

1 Introdução

A adoção - nos últimos 10 ou 20 anos - de Prontuário Eletrônico do Paciente (PEP) por unidades de saúde brasileiras trouxe como consequência a

necessidade de conectar sistemas diversos em diferentes horizontes temporais e geográficos. Paradoxalmente, o sucesso dos sistemas de informação em saúde é comprometido não pela falta de padrões, mas pela multiplicidade deles. Além disso, é claro o desconhecimento de noções de básicas de classificação pelos criadores desses padrões. Observam-se facilmente, dentre a variedade de iniciativas terminológicas, problemas como sobreposição, redundância, erros de classificação etc., que ocasionam dificuldades de integração entre PEPs.

Essas dificuldades na integração de PEPs e sistemas médicos em geral, conhecidas popularmente por **problema da interoperabilidade**, afetam os cuidados à saúde em todo mundo. Uma busca simples do termo *healthcare interoperability*¹ no Google Scholar (GOOGLE, [2019]) retorna mais de cinco mil artigos, em que se pode encontrar uma variedade de **soluções**, como ontologias; padrões, como o *Health Level 7* (HL7) e *HL7 Fast Healthcare Interoperability Resources* (FHIR), que alegam ter a solução (SIWICKI, 2019); *big data*; abordagens do *blockchain*²; iniciativas de *data mining*; de internet das coisas e de *machine learning*, para citar algumas.

A grande maioria dessas soluções se baseia puramente em tecnologia, mas apenas usando ontologia já é possível lidar com a questão da interoperabilidade semântica, do ponto de vista conceitual e tecnológico (ALMEIDA; RIBEIRO; MACULAN, 2019). Isso significa que, no contexto da Ciência da Informação, as ontologias combinam tanto os princípios metafísicos da Filosofia quanto as técnicas computacionais da Ciência da Computação (ALMEIDA; RIBEIRO; BARCELOS, 2020). Nesse sentido, a disciplina denominada *Ontologia Aplicada* consiste no uso de princípios metafísicos para criar boas ontologias enquanto artefatos tecnológicos.

É sabido que a solução para os problemas de interoperabilidade não é trivial, se é que é possível. Defende-se aqui, portanto, a possibilidade de mitigá-los. Nesse contexto, o artigo aborda pelo menos dois aspectos conceituais aos quais se atribui relevância quando se consideram os fatores nocivos à interoperabilidade: (i) a questão da “sobreposição epistemológica”, um termo cunhado por Bodenreider, Smith e Burgun (2004); e (ii) a questão da

“composicionalidade de termos”, assunto da Filosofia da Linguagem, mas pesquisado desde a década de 1990 (ELKIN, 1998).

Bodenreider, Smith e Burgun (2004) explicam que as terminologias clínicas estão cheias de afirmações “disfarçadas” de classes. Essas afirmações são, de fato, informações complementares, denominadas **aspectos epistêmicos**. Para entender a questão, considere a declaração **severidade de uma dor**, em que o termo **severidade** é vago e não pode ser mantido com significado único na comunicação entre sistemas de informação. No contexto da padronização da linguagem médica, aspectos epistêmicos causam a **sobreposição epistemológica**, fenômeno em que informações indesejadas que não representam classes reais são adicionadas à terminologia. Muita informação desse tipo tem caráter subjetivo e não contribui para fins de interoperabilidade.

Nos estudos sobre composicionalidade³, Elkin *et al.* (1998) definem “termo atômico” como “[...]uma noção [...] que não pode ser dividida em partes sem a perda de significado” (ELKIN *et al.*, 1998, p. 660, tradução nossa). Nesse sentido, uma solução seria separar, de um lado, os termos atômicos, atribuindo a eles suas características em um modelo ontológico e, de outro lado, os termos compostos com qualificadores. Por exemplo, a declaração *Coronary Artery Disease Status Post CABG*⁴ é composta por vários termos atômicos com significados próprios.

Considerando o contexto apresentado, o presente artigo dá continuidade à pesquisa iniciada em Teixeira (2019), em que a questão subjacente é o uso de terminologias clínicas padronizadas para lidar com a falta de interoperabilidade semântica entre sistemas de informação, inclusive PEPs, em unidades médicas.

No presente artigo, pretende-se demonstrar que arranjos conceituais, objetos de pesquisa da Ciência da Informação (nem sempre soluções tecnológicas), são essenciais na busca por um nível de interoperabilidade em instituições médicas, a saber:

- a) a redução da sobreposição epistemológica em terminologias;
- b) o favorecimento do uso de termos atômicos em detrimento das composições sem critério;

- c) a construção de ontologias como artefato com classes representativas de entidades do mundo real.

Apresentam-se aqui testes empíricos, ilustrando possibilidades de mitigar os problemas na interoperabilidade por via conceitual. Espera-se evidenciar como a composicionalidade e a sobreposição em terminologias clínicas geram problemas aos sistemas.

O restante do presente artigo está organizado como segue: a Seção 2 traz um *background* dos grandes temas envolvidos, de forma a possibilitar o entendimento da pesquisa; a Seção 3 descreve a metodologia de pesquisa para realização de testes empíricos com a Classificação Internacional de Doenças (CID), bem como os resultados obtidos; a Seção 4 discute esses resultados à luz da necessidade de aplicações úteis para as instituições médicas, oferece considerações finais e sugere perspectivas de pesquisa futuras.

2 Visão geral: ontologias, terminologias clínicas, composicionalidade terminológica

Tem sido um real desafio a busca pela precisão na recuperação da informação em um contexto em que já se tornou lugar-comum falar em crescimento exponencial de dados, como na área de medicina, de biomedicina e de cuidados à saúde. Em função do grande avanço das tecnologias digitais, como a web semântica, novas necessidades se mostram prementes.

A necessidade de melhor padronizar o discurso em unidades médicas já é percebida há quase dois séculos com as primeiras terminologias especializadas (BOWKER, 1996; LAURENTI *et al.*, 2013; UJVARI; ADONI, 2014). Um cenário de diversidade de iniciativas gerou uma variedade de terminologias no contexto médico, nomeadas como ontologias, terminologias, vocabulários controlados, classificações, cabeçalhos de assunto, dentre outras. Sistemas de classificação e terminologias na área médica envolvem uma ampla gama de recursos desenvolvidos ao longo do tempo. Exemplos são a Classificação Internacional de Doenças; os *Medical Subject Headings* (MeSH); a *Gene Ontology*; a *Systematized Nomenclature of Medicine* (SNOMED CT) e o *Foundational Model of Anatomy* (FMA).

Esses sistemas, em geral reunidos sob o rótulo de sistemas de organização do conhecimento, foram criados para acompanhar a evolução da medicina e a evolução tecnológica, mas cada um manteve sua especificidade de propósitos. Alguns deles objetivam lidar com o controle de endemias, outros objetivam a recuperação de textos médicos e outros ainda servem para modelos de aplicações computacionais. Vale aqui ressaltar que na literatura da informação médica e informática em saúde, esses sistemas já mencionados, como terminologia, tesauro, cabeçalho de assunto, ontologias, vocabulários, nomenclaturas e classificações, dentre outros, são, via de regra, confundidos e empregados de forma equivocada. Por exemplo, Rubin, Shah e Noy (2008) explicam que essas estruturas de representação do conhecimento são muitas vezes denominadas como vocabulários, tesouros, terminologias e ontologias, como se fossem termos sinônimos. Schulz e Martinez-Costa (2013) unificam todas essas estruturas sob o nome de “vocabulários”, sejam terminologias, tesouros, ontologias ou classificações. Wang (2018) corrobora os entendimentos anteriores ao afirmar que diferentes nomes são dados aos sistemas de termos ou conceitos padronizados. O mesmo autor ainda indica que esses sistemas podem ser denominados como terminologia, vocabulário ou léxico e qualquer um desses pode ser qualificado como “controlado”. Está além dos objetivos do artigo discutir individualmente cada denominação, sendo que os interessados podem consultar a literatura sugerida acima. Aqui se adota o termo **terminologia clínica**.

A proliferação de padrões, entretanto, não se tornou necessariamente profícua, pois resultou na redundância e na sobreposição de escopos, comprometendo as iniciativas de interoperabilidade e as possibilidades para sistemas de apoio à decisão (FREITAS; SCHULZ; MORAES, 2009). Corroborando essa constatação, o Quadro 1 traz uma amostra da grande diversidade de vocabulários médicos criados para propósitos dos mais variados, gerando redundância e sobreposição. Nesse mesmo sentido, o Quadro 2 mostra a heterogeneidade dos dados com relação à forma como são organizados em diferentes vocabulários médicos, o que torna a interoperabilidade e o uso no apoio à decisão uma tarefa árdua.

Baseando-se na premissa de que a interoperabilidade é um problema complexo e que deve ser atacado em diferentes frentes, descrevem-se no restante desta seção três temas para fundamentação da metodologia de pesquisa:

- as terminologias clínicas;
- ontologias e distinção entre ontologia e epistemologia;
- composicionalidade e a atomicidade em terminologias clínicas.

Quadro 1 - Diversidade de vocabulários médicos para diferentes propósitos

Vocabulário médico	Instituição responsável	Propósito / Uso
<i>International Classification of Diseases (ICD)</i>	<i>World Health Organization</i>	Classificar doenças para obtenção de dados estatísticos
<i>WHO Drug Dictionary</i>	<i>Uppsala Centre for International Drug Monitoring</i>	Classificar drogas
<i>Anatomical Therapeutic Chemical Classification System (ATC)</i>	<i>WHO Collaborating Centre for Drug Statistics Methodology</i>	Comparar estatísticas sobre o uso de drogas
<i>RxNorm</i>	<i>United States National Library of Medicine</i>	Normalizar nomes de drogas e fármacos
<i>Dictionary of Medicines and Devices</i>	<i>World Health Organization</i>	Prover códigos para identificação de remédios e dispositivos
<i>NCI Thesaurus</i>	<i>National Cancer Institute</i>	Prover terminologia de referência sobre câncer
<i>Logical Observation Identifiers Names and Codes (LOINC)</i>	<i>Regenstrief Institute</i>	Prover terminologia para comunicação entre laboratórios
<i>Medical Dictionary for Regulatory Activities (MedDRA)</i>	<i>Northrop Grumman</i>	Prover terminologia médica para atividades regulatórias
<i>Medical Imaging and Technology Alliance (DICOM)</i>	<i>Medical Imaging and Technology Alliance</i>	Prover descrições para imagens médicas
<i>Medical Subject Headings (MeSH)</i>	<i>United States National Library of Medicine</i>	Indexar literatura médica
<i>Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT)</i>	<i>International Health Terminology Standards Development Organisation</i>	Prover terminologia clínica para documentação médica

Fonte: Almeida (2013)⁵.

Quadro 2 - Diversidade de formas de organização de dados provenientes de diferentes vocabulários médicos

Termo 1	Relação	Termo 2
Aspirin	<i>prevents</i>	<i>Myocardial_infarction</i>
Blood_Plasma	<i>narrower_than</i>	<i>Blood</i>
Cancer	<i>causes</i>	<i>Weight_Loss</i>
Cell	<i>has_part</i>	<i>Cell_Membrane</i>
Diabete Mellitus	<i>has_prevalence</i>	<i>5,8%</i>
Diclofenac	<i>has_side_effect</i>	<i>Gastrointestinal_Bleeding</i>
ELM-2	<i>is_a</i>	<i>Protein</i>
Fever	<i>symptom_of</i>	<i>Malaria_Tropica</i>
Hepatitis	<i>has_location</i>	<i>Liver</i>
Hyperthermia	<i>has_synonym</i>	<i>Fever</i>
Smoking	<i>causes</i>	<i>Cancer</i>
WHO	<i>located_in</i>	<i>Geneva</i>
Aspirin	<i>is_a</i>	<i>Salicylate</i>
hypertension	<i>is_a</i>	<i>Cardiovascular_Risk_Factor</i>

Fonte: Almeida (2013).

2.1 Breve visão das terminologias clínicas

As causas de morte constituíam umas das necessidades informacionais do governo francês desde o século XIX. Dois médicos estatísticos, William Farr⁶ e Jacques Bertillon⁷, iniciaram as listas de classificação de doenças culminando na “Classificação das Causas de Morte de *Bertillon*” (LAURENTI e colaboradores, 2013). Já no século XX, houve grande progresso científico da medicina tendo, como consequência, maior produção científica e maior volume de informações relativas aos atendimentos médicos. Tal progresso impulsionou o desenvolvimento de diferentes sistemas classificatórios e terminologias, como as iniciativas já mencionadas CID, MeSH, SNOMED CT, dentre outras.

Neste artigo, o recorte para a CID e para SNOMED CT se justifica por: (i) a obrigatoriedade do uso da CID no Brasil; (ii) a ampla abrangência da SNOMED CT e sua adoção como terminologia oficial brasileira.

A CID é um padrão de codificação de diagnóstico para fins clínicos, estatísticos e de pesquisa. Seu escopo contempla mortalidade e morbidade. Sua estrutura está baseada em um extenso esquema classificatório de eixo variável tanto para fins práticos quanto epidemiológicos. Esse esquema contém códigos alfanuméricos organizados em capítulos, agrupamentos, categorias e subcategorias, e é nas subcategorias que estão os códigos mais específicos, que devem ser atribuídos em um prontuário ou declaração de óbito. Embora a CID se intitule como uma **classificação**, a organização de seu conteúdo não corresponde cientificamente a esse tipo de instrumento de representação.

A SNOMED CT é um padrão terminológico para suporte à decisão clínica. Seu escopo abrange estruturas corporais, achados clínicos, localizações geográficas, eventos, entidade observável, organismos, produtos farmacêuticos, força física, objeto físico, procedimentos, artefato de registro, situação com contexto explícito, contexto social, espécime, estágio e escala e substância. As informações estão organizadas hierarquicamente por meio de conceitos, descritores e relacionamentos, permitindo expressões. Embora a SNOMED CT se declare uma ontologia, enquanto instrumento de representação do conhecimento, ela é uma terminologia sem princípios ontológicos bem definidos.

Entretanto, nem toda representação presente e descrita em terminologias clínicas é adequada para a nova realidade digital. Na verdade, essas terminologias foram criadas em diferentes épocas, por diferentes equipes e com diferentes objetivos (JUPP *et al.*, 2007). A CID, por exemplo, não foi criada considerando aplicabilidade aos modernos sistemas de informação. A SNOMED CT, por sua vez, sofre de deficiências se avaliada a partir de pressupostos da ontologia como disciplina (SCHULZ; MARTÍNEZ-COSTA; MIÑARRO-GIMÉNEZ, 2017). O HL7 se autodeclara um padrão, mas se limita a mensagens e fluxos (SMITH, 2018).

Essas terminologias clínicas, esses sistemas de classificação, esses padrões e outros recursos dessa natureza tentam replicar a linguagem natural e o jargão médico. Por melhor que seja a intenção, computadores atuais ainda não lidam com as ambiguidades e idiosincrasias da linguagem humana. A seção seguinte (Seção 2.2) levanta questões em que a linguagem natural não se mostra razoável nem para processamento por máquinas, nem para fins de interoperabilidade.

2.2 A distinção ontologia-epistemologia

Ontologias em sistemas de informação são um assunto que é estudado desde a década de 1980 (WAND; WEBER, 1990) e que, pelo grande número de publicações recentes, dispensa introduções. Ontologias são uma proposta de sucesso para fins de representação e recuperação da informação (ANDRADE, 2013) amplamente utilizada na área médica. Cabe, de imediato, distinguir entre “ontologia como disciplina”, que compreende princípios metafísicos adotados na construção de modelos e “ontologia como artefato”, que diz respeito a um artefato de *software*, uma estrutura de representação do conhecimento usada na web semântica (ALMEIDA; FARINELLI, 2017).

A representação do conhecimento no âmbito da ontologia como disciplina se fundamenta em problemas milenares sobre a identidade e a existência, bem como questões filosóficas subjacentes (ALMEIDA; TEIXEIRA, 2019). A ontologia no âmbito da metafísica busca resolver tais questões por meio de entidades abstratas, invariantes na realidade, identificadas ainda por

Aristóteles como universais (FINE, 1995). Os universais remetem aos particulares, também um termo de origem aristotélica, usado para se referir a coisas que existem na realidade espaçotemporal (SMITH; MUNN, 2008).

Algumas características dos particulares são denominadas **essenciais**, por exemplo: uma característica comum, que poderia ser a fotossíntese, é que faz com que vegetais sejam do tipo **vegetal**. O termo **tipo** remete ao termo **categoria**, utilizado pela primeira vez, tecnicamente, por Aristóteles. Falando de outra forma, o vegetal é do tipo vegetal porque o **universal vegetal** inclui todas as características essenciais para algo ser daquele tipo, essencialmente a fotossíntese. A árvore em frente da minha janela é um **particular árvore** porque está associada ao **universal árvore** por meio da instânciação. As folhas da árvore têm a cor verde porque possuem o “universal verde” (ALMEIDA; TEIXEIRA, 2019). Toda essa teoria tem sido aplicada há pelo menos dois milênios, mas há cerca de 20 anos envolve também a ontologia como artefato tecnológico, no que resulta a *Ontologia Aplicada* (SMITH; MUNN, 2008).

A *Ontologia Aplicada* agrega a ontologia como disciplina e a ontologia como artefato de forma a produzir uma representação que possibilite comunicação entre máquinas sem intervenção humana. Esse nível de comunicação não envolve apenas questões de padronização técnica, mas diz respeito a uma questão de caráter conceitual denominada “interoperabilidade semântica” (UKOLN, 2005). Para tornar operacional esse nível de comunicação e de automatização, é preciso manter “significados” uniformes, restritos a interpretações especificadas via axiomas em uma ontologia como artefato, para uso por máquinas (ALMEIDA; SOUZA; FONSECA, 2011).

O estudo da ontologia, de forma simples, busca responder às questões sobre objetos, ou de forma mais genérica, entidades: (i) Qual o modo de existência das entidades? (ii) Quais relações existem entre essas entidades? (iii) Como analisar a estrutura da realidade?

Por outro lado, como é amplamente sabido, a epistemologia reflete a relação entre um **sujeito** e o **objeto** de sua apreensão. Dessa relação surgem respostas para perguntas um pouco diferentes, por exemplo: (i) Como sabemos o que outras pessoas querem dizer quando conversam conosco? (ii) Como

sabemos que as declarações que fazemos sobre a realidade são verdadeiras? (iii) Como verificamos tais declarações? (HESSEN, 2003).

De fato, ao lidar com entidades do mundo para fins de classificação, é necessário entender a natureza dessas entidades, o que inclui abordagens ontológicas e epistêmicas (GUBA, 1990). Isso envolve, num primeiro momento, identificar entidades subjetivas e objetivas ao considerar pelo menos quatro sentidos possíveis (SEARLE, 2000):

- a) o sentido **ontológico-objetivo**, por exemplo, montanhas, galáxias, pessoas etc.;
- b) o sentido **ontológico-subjetivo**, por exemplo, dores, cócegas, coceiras;
- c) o sentido **epistêmico-objetivo**, por exemplo, “Vicent Van Gogh morreu na França”;
- d) sentido **epistêmico-subjetivo**, por exemplo, “Van Gogh pinta melhor do que Monet”.

No âmbito da prática clínica, a distinção ontologia-epistemologia pode ser explicada, de um lado, pela identificação das coisas (como os pacientes) e, por outro lado, pelo que as pessoas conhecem sobre as coisas (como os sinais identificados pelo médico após o exame físico [OGMS, 2019a]). Por exemplo, existe uma doença chamada daltonismo⁸, em função da qual as cores vermelha e verde são trocadas. Ainda que o daltônico perceba o verde como vermelho em um sinal de trânsito, o verde no sinal de trânsito continua sendo verde.

Existe assim uma diferença entre as coisas em si e o que é fruto da percepção. No exemplo do daltonismo, deficiências no aparato visual da pessoa acarretam conhecimentos incompletos, ambíguos ou vagos, bem como interpretações subjetivas que nem sempre possibilitam representações fidedignas da realidade. Bodenreider, Smith e Burgun (2004) exibem casos que envolvem a distinção ontologia-epistemologia aplicada às terminologias clínicas. Por exemplo, a **severidade** de uma dor pode ter diferentes sentidos epistêmico-subjetivos e percepções distintas para as pessoas. Nesses casos é inútil buscar por uma representação consensual que restrinja o significado proporcionado pela ontologia como artefato no contexto das limitações das máquinas atuais. No

exemplo anterior das cores, a representação adequada seria considerar o melhor que a ciência nos diz a respeito — que verde é uma cor com esse comprimento de onda — e não diferentes percepções e interpretações sobre um fenômeno.

Essas questões podem causar controvérsias. Está além dos objetivos do presente artigo discutir implicações de filosofia da biologia, mas os interessados podem consultar Smith e Ceusters (2010) e Merrill (2010). Para fins práticos, relacionados aos objetivos do presente artigo, adotam-se aqui visões que explicam o fenômeno ontologia-epistemologia em vocabulários da medicina (RECTOR, 1999). Nesse contexto foi cunhado o termo “sobreposição epistemológica”⁹ (BODENREIDER; SMITH; BURGUN, 2004), o qual é adotado aqui para referir uma declaração que, além de receber *status* ontológico via ciência, vem acompanhada de termos adicionais, ou à qual vêm se **sobrepor** esses termos adicionais, que representam percepções, muitas vezes subjetivas, do que pessoas diferentes podem pensar ou julgar.

Embora tais termos possam ser importantes para o profissional de saúde no atendimento clínico e na tomada de decisão, a sobreposição epistemológica não é adequada nem para a construção de ontologias como artefatos, nem para a interoperabilidade que ontologias podem fornecer (TEIXEIRA, 2019). As sobreposições geram declarações que não correspondem a entidades da realidade, ou seja, nem aos tipos naturais nem a agrupamentos criados pela demarcação humana, como as classes. Mesmo as terminologias padronizadas, criadas para lidar com tais variações, não resolvem o problema, na medida em que a multiplicidade de opções, além dos objetivos e propósitos diversos, tem resultado em redundância, retrabalho e sobreposição entre termos.

A própria SNOMED CT, desenvolvida para aplicação em sistemas de PEP e autodeclarada **ontologia**, não está livre de aspectos epistêmicos apresentando redundâncias e sobreposições. Um exemplo é a história de uma reação adversa à penicilina. Aspectos como a gravidade (“severo”) e a certeza sobre a causa da reação adversa atribuída à penicilina (“Provavelmente presente”) são aspectos epistêmicos e não consistem em verdades invariáveis (MARCO-RUIZ *et al.*, 2017, p. 37):

161632004 |História do fator de risco|:
 { 246090004 |Descoberta associada| =
 (281647001 |Reação adversa|:
 {246112005 |Gravidade|= 24484000 |Severo|,
 42752001|Devido a|=6369005|Penicilina|}),
 408729009|Finding context|=410592001|Provavelmente presente|}

Como exemplos complementares de sobreposições epistemológicas em diferentes terminologias clínicas, apresentam-se ainda: (i) na CID 10: insuficiência renal aguda; insuficiência renal crônica; endometriose do peritônio pélvico; pólipos do corpo do útero; pólipos do colo do útero; dermatite alérgica de contato devida a metais; dermatite alérgica de contato devida a adesivos e dermatite alérgica de contato devida a alimentos em contato com a pele; (ii) na SNOMED CT: otorragia da orelha esquerda; otorragia da orelha direita; febre hemorrágica da dengue, grau I; febre hemorrágica da dengue, grau II; alta umidade e baixa umidade.

Outro exemplo também comprova e pode ajudar no entendimento das sobreposições epistemológicas: certas terminologias incluem **pressão arterial em pé** e **pressão arterial sentado(a)**. Enquanto, de fato, a posição da pessoa afeta na medição de sua pressão, a posição não cria uma nova entidade, ou seja, a pressão arterial consiste em uma única entidade e as duas “pressões” medidas dizem respeito a duas instâncias da mesma entidade¹⁰. A duplicidade entre **posição em pé** ou **posição sentado(a)** expressa um tipo de sobreposição, como se definiu aqui. Exemplos desse tipo são comuns em terminologias clínicas (Quadro 3), e exemplos reais foram obtidos no contexto brasileiro (Quadro 4).

Quadro 3 - Fragmento de análise de declarações em terminologias

Terminologia	Declaração original	Comentários sobre sobreposição
SNOMED CT	“aumento da susceptibilidade a infecções”	não se trata de uma classe ou de um universal; na verdade, se refere a um achado clínico
CID 10	“tuberculose do olho”	a tuberculose é uma doença e a menção do local olho não a modifica como entidade da realidade
CIPE ¹¹	“magreza excessiva”	o termo “excessiva” denota uma característica subjetiva, sujeita a diferentes interpretações

ICF ¹²	“funções genitais e reprodutivas, outras especificadas e não especificadas”	apesar de esse tipo de construção objetivar exaustividade, não é adequado para fins de interoperabilidade por sua inerente vagueza
ICHI ¹³	“Biópsia de osso de outro local e não especificado”	indica o local do procedimento, não uma entidade inequívoca da realidade

Fonte: Elaborado pelos autores.

Quadro 4 - Fragmento de tipos de sobreposições

Caso 3	<i>termos que incluem localização ou forma de realizar diagnóstico</i> anotar como: <u>localização da doença</u> termo identifica local	exemplo: “ <i>tuberculose de glândulas adrenais</i> ” identifica local, mas a bactéria da tuberculose pode também afetar o pulmão e outros órgãos
Caso 5	<i>Termos que incluem a forma de descoberta da doença e seu estágio</i> anotar como: <u>forma de descoberta fortuita de doença</u>	exemplo: “ <i>colelitíase assintomática</i> ” nomeia a condição na qual cálculos biliares são descobertos por ultrassonografia, tomografia ou raio X, sendo que os exames foram prescritos por outra observação, por exemplo, sangue na urina
Caso 6	<i>termos que incluem modalidade ou possibilidade</i> anotar como: <u>modalidade da doença</u> presença de termos que se referem a modalidades das doenças anotar como: <u>possibilidade de doença</u> presença de termo que indica uma suposição do médico	exemplo: termos “definitivo” e “provável” em “ <i>abscesso definitivo tubo-ovariano</i> ” e “ <i>provável abscesso tubo-ovariano</i> ” exemplo: “ <i>suspeita de tuberculose</i> ”

Fonte: Teixeira (2019, p. 120).

Outros casos de sobreposição epistemológica são amplamente exibidos na literatura (RECTOR, 1999; BODENREIDER; SMITH; BURGUN, 2004; STENZHORN *et al.*, 2008; FREITAS; SCHULZ; MORAES, 2009; MARCO-RUIZ *et al.*, 2017). Um desafio nesse contexto consiste em identificar e tratar a presença de sobreposições epistemológicas nas terminologias para depois propor meios de integração entre elas.

2.3 Composicionalidade e atômidade em terminologias clínicas

A ideia de dividir termos compostos em termos atômicos para entender seus significados não é nova. Ela tem origem na Filosofia da Linguagem, remete a Frege¹⁴ no século XIX e é conhecida como **composicionalidade**. Um sistema é composto se o significado de cada expressão complexa naquele sistema depende

apenas da estrutura sintática da própria expressão e do significado de cada parte atômica da respectiva expressão (MURPHY, 2002).

No contexto das terminologias clínicas, a tentativa de tornar completo o conjunto de termos especializados resulta no uso de termos, atômicos ou compostos, para abranger as situações que o profissional enfrenta na prática. O efeito dessa tentativa de exaustividade, entretanto, é um considerável nível de sobreposição entre termos e conceitos. Um termo atômico consiste em “[...] uma representação de um conceito que não é composto de outras representações conceituais [...] em uma terminologia específica” (ELKIN, 2012, p. 57).

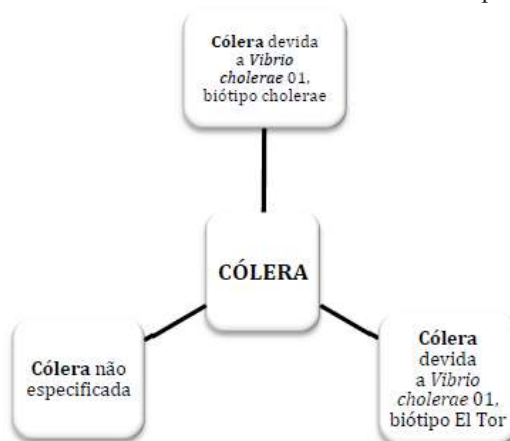
Os termos compostos, em geral, carregam sobreposições epistemológicas, gerando extensas terminologias cheias de redundância. Tais terminologias são impossíveis de gerenciar, mesmo com um uso intensivo de computadores. Estudos ainda do início dos anos 2000 corroboram a possibilidade dessa situação explosiva ao demonstrar a amplitude de termos existentes: testes de interoperabilidade foram derivados de 200 sistemas terminológicos diferentes (CAMPBELL; OLIVER; SHORTLIFE, 1998). Em 2019, sistemas pós-coordenados como a SNOMED CT já ultrapassam os 300.000 termos.

Ontologias são uma alternativa para lidar com esse tipo de situação, uma vez que permitem realizar um inventário das entidades. A ontologia como artefato consiste em uma estrutura formal, especificada em lógica descritiva¹⁵, livre de aspectos epistêmicos como modificadores e qualificadores (EMYGDIO; FELIPE; ALMEIDA, 2019). Uma definição bem conhecida de ontologias sugere: “Uma ontologia é uma especificação conceitual que descreve o conhecimento sobre um domínio que é independente de estados epistêmicos e situações [...]” (GUIZZARDI, 2005, p. 83).

Se os estados epistêmicos não contam para a ontologia como artefato, considera-se uma oportunidade criar e aplicar um método baseado na atomicidade das declarações em terminologias clínicas para fins de redução de erros, de redundância e de complexidade inerente e, em última instância, para fins de busca por interoperabilidade (TEIXEIRA, 2019).

Dessa forma, utiliza-se aqui a expressão “termo atômico” para designar o termo que não pode mais ser decomposto (ZINGANO, 2013). A busca pela atomicidade dos termos resulta em termos livres de aspectos epistêmicos. A Figura 1 exemplifica declarações do agrupamento **A00 Cólera** da CID, as quais derivam do mesmo termo atômico **cólera**. Tais declarações contêm informações complementares, além de conter os códigos específicos que serão atribuídos aos pacientes nos seus prontuários. A busca por termos atômicos está relacionada à questão, já mencionada, da sobreposição epistemológica, visto que essa última ocorre em termos compostos.

Figura 1 - Termo atômico ‘cólera’ na CID e composições



Fonte: Teixeira (2019, p. 127).

Trabalhos de caráter prático e de abrangência internacional, visando em última instância a melhor atendimento à população através de alinhamento de sistemas de informação médicos e PEP, exemplificam situações em que a sobreposição e a composicionalidade prejudicam a possibilidade de integração (MARCO-RUIZ *et al.*, 2017). Na seção seguinte (seção 3), apresenta-se um método para lidar com as duas questões.

3 Metodologia e resultados

Esta seção descreve a metodologia de pesquisa e os resultados do estudo. Apresenta-se o contexto de pesquisa para em seguida explicar-se o método por

meio de passos bem definidos, incluindo o seu teste em uma terminologia médica.

Em pesquisa anterior (TEIXEIRA, 2019), da qual a presente pesquisa é uma continuação, desenvolveu-se uma ontologia da CID. Essa ontologia, denominada Artefato Ontológico CID, é aplicada ao experimento do presente artigo. A CID foi analisada e se constatou a forte presença de sobreposição epistemológica. Foram selecionadas duas amostras por conveniência, partes do capítulo I da CID. A método aplicado será apresentado junto aos resultados (seção 3.1), de forma a subsidiar discussões posteriores (seção 3.2).

3.1 Procedimentos metodológicos, método e resultados obtidos

No restante desta seção, apresentam-se os passos estabelecidos, assim como seus respectivos resultados.

3.1.1 Passo 1: selecionar e identificar escopo

Analisar o escopo delimitado a partir de aspectos centrais do Artefato Ontológico CID, a saber: doença; diagnóstico; etiologia (agente e causas); local (anatomia humana), dentre outros (Quadro 5).

Quadro 5 - Fragmento de análise preliminar das declarações CID

Código	Título	Análise preliminar
A00-A09	Doenças infecciosas intestinais	"infecciosas" diz respeito ao tipo de desordem; "intestinais" diz respeito à anatomia (local)
A00	Cólera	
A00.0	-Cólera devida a <i>Vibrio cholerae</i> 01, biótipo <i>cholerae</i>	inclui um diagnóstico; um agente etiológico com duas variações; um local
A00.1	-Cólera devida a <i>Vibrio cholerae</i> 01, biótipo <i>El Tor</i>	
A00.9	-Cólera não especificada	inclui um diagnóstico; um local

Fonte: Teixeira (2019, p. 111).

3.1.2 Passo 2: identificar sobreposição de aspectos epistemológicos

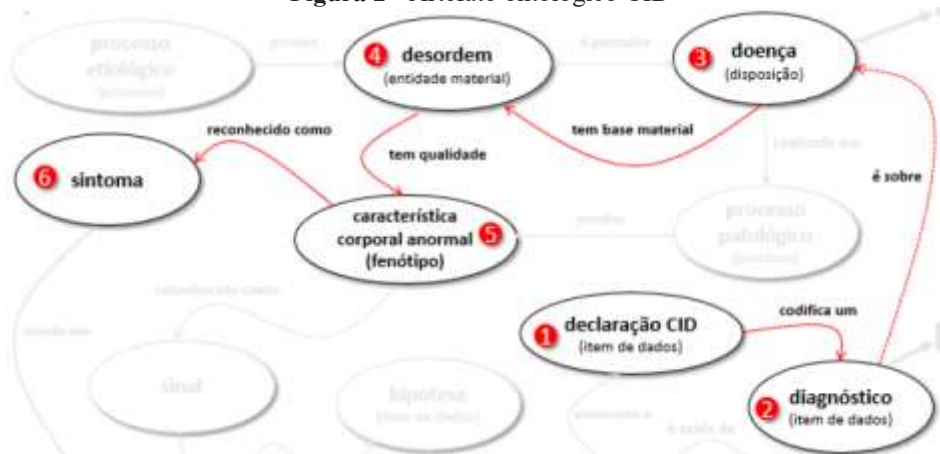
Analisar a declaração CID, de acordo com os casos de sobreposição identificados, registrados no Quadro 4 (seção 2.2).

3.1.3 Passo 3: decompor a declaração de acordo com o Artefato Ontológico CID.

Dividir a declaração de acordo com o modelo ontológico para identificar classes do mundo real (Figura 2) considerando as seguintes classes: declaração CID;

diagnóstico; doença; desordem; característica corporal anormal; sintoma; processo etiológico.

Figura 2 - Artefato ontológico CID



Fonte: Teixeira (2019, p. 117).

3.1.4 Passo 4: verificar e implementar termos atômicos

Neste passo, foram realizadas as divisões dos termos com a identificação das sobreposições epistemológicas, de forma a detectar os termos invariantes em uma declaração, que são os candidatos a termos atômicos. Os qualificadores e modificadores identificados são complementos ao redor de um termo atômico.

Este passo corresponde ainda à inserção das classes obtidas com o Artefato Ontológico CID no editor de ontologias *Protégé* (Figura 3). Para fins de otimização de tempo, de esforços demandados e visando a consistência e a qualidade da representação, a reutilização de ontologias é importante. Nesse sentido, foram importadas as seguintes ontologias:

- Ontology for General Medical Science* (OGMS)¹⁶;
- Human Disease Ontology* (DO)¹⁷;
- Information Artifact Ontology* (IAO)¹⁸;
- Basic Formal Ontology* (BFO)¹⁹.

Na Figura 3 é apresentado o processo de decomposição e ancoragem da declaração **A00.0 Cólera devida a *Vibrio cholerae* 01, biótipo *cholerae*** no *Protégé*. É possível observar a sequência que se inicia no destaque em azul. A **declaração CID** codifica o **diagnóstico de cólera**, que é sobre uma doença

cólera, que tem como base material a **célula infectada pela bactéria *Vibrio cholerae* 01**, que se apresenta como uma característica corporal anormal **anormalidade do trato gastrointestinal**, que é reconhecida por seu sintoma **cólica abdominal**, dentre outros. Assim, à medida que a declaração composta vai sendo decomposta em termos atômicos (ou quase atômicos), a interoperabilidade semântica é favorecida.

Figura 3 - Fragmento da representação esquemática da decomposição

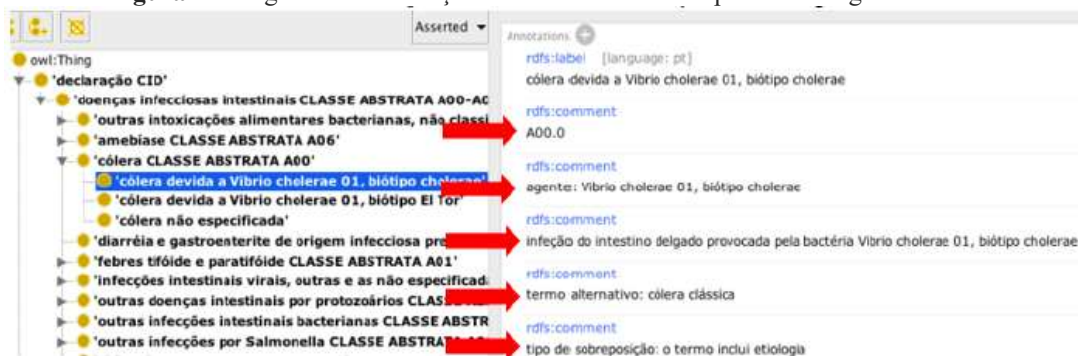


Fonte: Teixeira (2019, p. 118).

3.1.5 Passo 5: refinar o Artefato Ontológico CID

Buscou-se conteúdo na internet e em dicionários especializados para preencher as subclasses do artefato, para as quais a CID nem sempre fornece conteúdo, por exemplo: características corporais anormais, processo etiológico, desordem, dentre outras informações (Figura 4).

Figura 4 - Fragmento de anotações e conhecimento complementar registrado



Fonte: Teixeira (2019, p. 110).

3.1.6 Passo 6: validar os aspectos médicos da decomposição

A validação da clareza dos procedimentos e da viabilidade do método foram obtidas por meio de outro pesquisador de ontologias que, seguindo o mesmo processo, alcançou os mesmos resultados com um segundo agrupamento CID distinto. Já a validação dos aspectos médicos das decomposições foi realizada informalmente, com um profissional de saúde especializado nos assuntos e pertencente ao corpo docente de um hospital universitário.

Na próxima seção, discutem-se os resultados obtidos a partir de considerações levantadas na revisão de literatura.

3.2 Discussão

A proposta metodológica do presente artigo se baseou no argumento de que a reconstrução ontológica de uma terminologia clínica poderia ser usada na melhoria da qualidade da informação médica em um PEP, para fomentar a interoperabilidade. Cabe destacar que métodos de alguma forma similares aos apresentados aqui, de refinamento a partir de princípios ontológicos, são encontrados na literatura brasileira de Ciência da Informação, como em Maculan (2014).

Em uma primeira perspectiva, foi possível observar a extensão da sobreposição epistemológica. A CID 10 possui cerca de 14.000 declarações com sobreposições. De fato, em uma única declaração CID — a qual deveria ser uma classe representativa de uma entidade no mundo real —, encontram-se mescladas informações sobre etiologia, local da doença ou procedimento, o tipo de manifestação, dentre outras. A pesquisa listou doze tipos de sobreposições epistemológicas, como forma de contribuição:

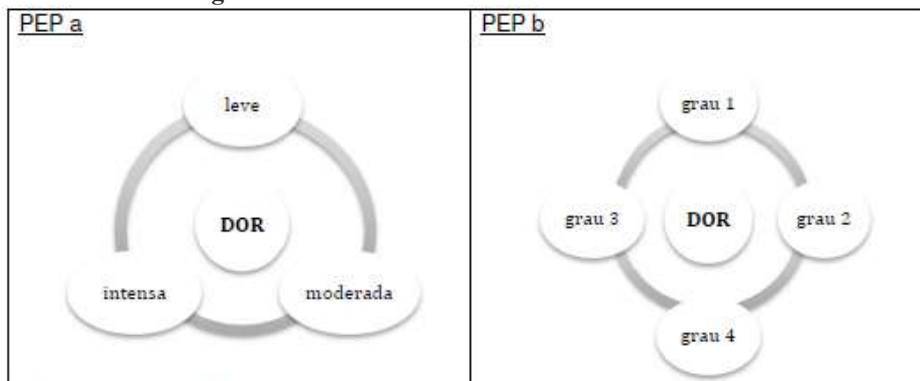
- a) termos que não se referem exatamente ao diagnóstico de uma doença;
- b) termos que incluem a técnica de identificação;
- c) termos que incluem localização ou forma de realizar diagnóstico;
- d) termos que incluem lesões, manifestações e ferimentos associados;
- e) termos que incluem a forma de descoberta da doença e seu estágio;
- f) termos que incluem modalidade ou possibilidade;

- g) termos que exibem vagueza;
- h) termos usados para obter uma partição completa do domínio;
- i) termos pós-coordenados;
- j) termos que combinam códigos;
- k) termos que representam diagnósticos e doenças sem correspondência;
- l) termos que incluem informações sobre as causas.

É provável que existam ainda outros tipos de sobreposição a identificar na CID. A etapa de identificação do tipo de sobreposição epistemológica presente na declaração, mais sua implementação em um modelo ontológico, conduziu à identificação de termos atômicos. Ao analisar as declarações CID após as respectivas decomposições em termos atômicos, observou-se que os agrupamentos consistem em variações de características da mesma doença.

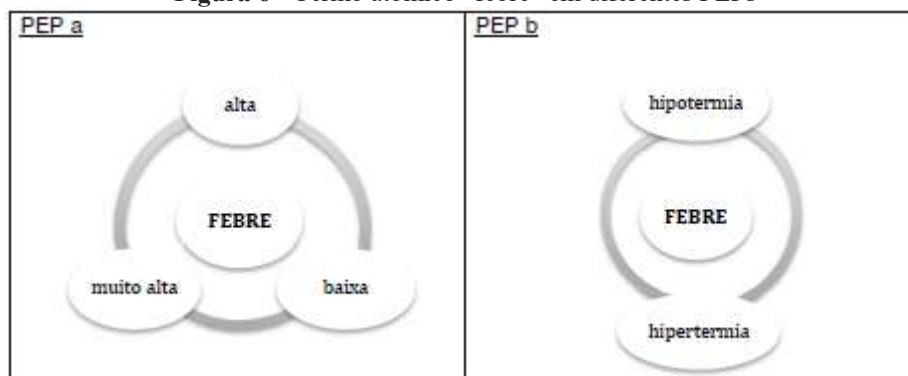
Pôde-se ainda verificar, como seria de esperar, que os qualificadores ou modificadores em um diagnóstico se concentram ao redor de um termo atômico. Um exame clínico, por exemplo, pode resultar em **dor severa e febre**. Em PEPs, esses qualificadores podem variar entre sinônimos e outros referenciais (Figuras 5 e 6). É possível pensar em interoperabilidade com relação aos termos atômicos, mas, como se pode perceber, não há possibilidade de integração direta automática no nível de composicionalidade não controlado. Na ontologia como artefato, esse tipo de informação adicional composta não é descartada, mas registrada no que se costuma denominar **anotação**.

Figura 5 - Termo atômico “dor” em diferentes PEPs



Fonte: Teixeira (2019, p. 129).

Figura 6 - Termo atômico “febre” em diferentes PEPs



Fonte: Teixeira (2019, p. 130).

Ao observarmos a Figura 5, dois questionamentos ontológico-objetivos são evidentes: (i) em que ponto a dor deixa de ser considerada **leve** para ser considerada **moderada**? (ii) como o **grau 3** da dor pode ser **moderado** ou **intenso**? Os mesmos questionamentos se estendem à Figura 6: (i) em que ponto a febre deixa de ser **alta** e é caracterizada como **muito alta**? (ii) quão **alta** ou **muito alta** uma febre precisa ser para ser identificada como **hipertermia**?

Pode-se afirmar que, na tentativa de integrar esses dois supostos PEPs com informações sobre os termos atômicos **dor** e **febre**, a interoperabilidade estaria comprometida, exceto se houvesse uma ontologia definindo formalmente os termos atômicos sem ambiguidades e provendo as relações semânticas.

4 Considerações finais

A presente pesquisa evidenciou a contribuição conceitual da Ciência da Informação para problemas de interoperabilidade semântica entre PEPs, destacando a necessidade de evitar abordagens puramente tecnológicas. Como contribuição científica, apresentou uma metodologia para reduzir a sobreposição em terminologias clínicas pela identificação de termos atômicos, favorecendo iniciativas de integração. A pesquisa ainda se desenvolve sob a premissa de que os diferentes tipos de PEPs podem interoperar.

Dentre os diferentes níveis de interoperabilidade possíveis (UKOLN, 2005), enfatizou-se aqui a interoperabilidade semântica. Cabe lembrar que a interoperabilidade semântica a que se faz referência ocorre em um contexto

automatizado, sem intervenção humana, propiciando troca de dados transparente entre diferentes sistemas médicos (ALMEIDA; SOUZA; FONSECA, 2011).

A sobreposição epistemológica cria barreiras à interoperabilidade semântica, uma vez que estabelece alinhamentos inadequados e redundâncias entre as classes de diferentes terminologias médicas. Nesse sentido, estabelecer um método de decomposição das declarações de terminologias clínicas — em partes que são ontológicas e partes que são epistemológicas — foi o primeiro desafio e gerou uma contribuição inédita para a Ciência da Informação. Em seguida, foram identificados diversos tipos de sobreposição, informações complementares que podem ser anotações e outras variáveis, o que levou à ênfase em termos atômicos, livres de aspectos epistêmicos, como forma de favorecer a interoperabilidade.

Como limitação da pesquisa, cita-se a falta de validação dos aspectos médicos das decomposições. Ainda assim, a ancoragem do termo atômico na classe **doença** a uma ontologia bem estabelecida com a *Human Disease Ontology* proporciona melhorias na representação do conhecimento médico por eliminar dados redundantes e variantes na declaração CID original.

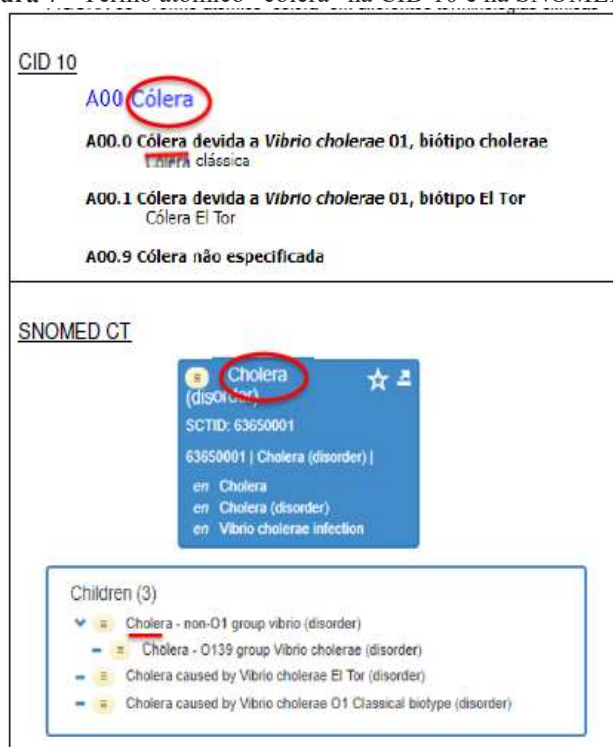
A pesquisa contemplou um amplo arcabouço científico, valendo-se de teorias e práticas da Ciência da Informação, da Informática em Saúde e da Medicina. Embora artigos seminais internacionais indiquem e abordem o problema, as possibilidades de soluções oriundas da Ciência da Informação eram, até então, inexistentes. Diante dos resultados obtidos, tanto a pesquisa quanto o artigo advogam que não há espaço para timidez da Ciência da Informação no âmbito da Informática em Saúde. Há muito a contribuir e os ganhos são para a sociedade como um todo, que poderá usufruir de melhor atendimento.

Como pesquisa futura, estão em andamento análises em que o método desenvolvido aqui é aplicado a outras terminologias, como a SNOMED CT. Desdobramentos da pesquisa buscam conciliar a descrição do paciente no SNOMED CT ao código de diagnóstico da CID, ensejando a sustentação das estruturas conceituais vias termos atômicos. Para ilustrar essa afirmação, a Figura 7 identifica o termo atômico **cólera** como aparece nas duas terminologias

clínicas citadas. A pesquisa que opera o vínculo entre ambas terminologias já estará em teste após a publicação do presente artigo.

Cabe destacar novamente que a identificação e a decomposição da sobreposição epistemológica da declaração CID, além da identificação do termo atômico nela contido, são passos para atacar frentes em que a falta de interoperabilidade precisa ser combatida, além de fomentar a busca por melhorias em sistemas que fazem uso de PEP.

Figura 7 - Termo atômico “cólera” na CID 10 e na SNOMED CT



Fonte: Teixeira (2019, p. 131).

Referências

ALMEIDA, M. B.; FARINELLI, F. Ontologies for representation of Electronic Medical Records: the obstetric and neonatal ontology. **Journal of the Association for Information Science and Technology**, [s. l.], v. 68, n. 11, p. 2529-2542, 2017.

ALMEIDA, M. B.; RIBEIRO, E. F.; BARCELOS, R. Toward a Document-Centered Ontological Theory for Information Architecture in corporations. **Journal of the Association for Information Science and Technology**, [s. l.], p. 1-19, 2020.

ALMEIDA, M. B.; RIBEIRO, E. F.; MACULAN, B. C. Ontologies and Classification: the unavoidable interplay between human reasoning and machine reasoning. *In: QUALITATIVE AND QUANTITATIVE METHODS IN LIBRARIES INTERNATIONAL CONFERENCE*, 11., 2019, Florence.

Proceedings [...]. Florence: QQML, 2019. p. 1-13.

ALMEIDA, M. B.; SOUZA, R. R.; FONSECA, F. Semantics in the Semantic Web: a critical evaluation. **Knowledge Organization Journal**, Edmonton, v. 38, n. 3, p. 187-203, 2011.

ALMEIDA, M. B.; TEIXEIRA, L. M. D. Revisitando os fundamentos da classificação: uma análise crítica sobre teorias do passado e do presente.

Perspectivas em Ciência da Informação, Belo Horizonte, p. 1-19, 2019. No prelo.

ANDRADE, A. Q. de. **A linguagem médica utilizada em prontuários e suas representações em Sistemas de Informação**: as ontologias e os modelos de informação. 2013. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

BFO. **Basic Formal Ontology**: home. [S. l.], [2019]. Acesso em: 7 nov. 2019.

BODENREIDER, O.; SMITH, B.; BURGUN, A. The Ontology-Epistemology divide: a case study in medical terminology. *In: INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY AND INFORMATION SYSTEMS*, 2004, Torino. **Proceedings** [...]. Torino: FOIS, 2004. p. 185-195.

BOWKER, G. C. The history of information infrastructures: the case of the international classification of diseases. **Information Processing & Management**, [s. l.], v. 32, n. 1, p. 49-61, 1996.

CAMPBELL, K. E.; OLIVER, D. E.; SHORTLIFE, E. H. UMLS: a collaborative approach to terminologic problems. **Journal of the American Medical Informatics Association**, [s. l.], v. 5, n. 1, p. 12-16, 1998.

DO. **Disease Ontology**: home. [S. l.]: Institute for Genome Sciences, [2019]. Acesso em: 7 nov. 2019.

ELKIN, P. L. *et al.* The role of compositionality in standartized problem generation. *In: WORLD CONGRESS ON MEDICAL INFORMATICS*, 9., 1998, Seoul. **Proceedings** [...]. Seoul: MEDINFO, 1998. p. 660-664.

ELKIN, P. L. Theoretical foundations of terminology. *In: ELKIN, P. L. (ed.). Terminology and terminological systems*. London: Springer-Verlag, 2012. cap. 4, p. 51-70.

EMYGDIO, J. L.; FELIPE, E. R.; ALMEIDA, M. B. Aspectos do ciclo de vida de dados em processos de construção de ontologias biomédicas. (2019). In: SANTOS, C. C. (org.). **Estudos interdisciplinares nas Ciências Exatas e da Terra e Engenharias**. Ponta Grossa: Athena, 2019. cap. 11, p. 113-125.

FINE, G. **On Ideas**: Aristotle's Criticism of Plato's Theory of Forms. Oxford: Oxford Scholarship, 1995.

FREITAS, F.; SCHULZ, S.; MORAES, E. Pesquisa de terminologias e ontologias atuais em biologia e medicina. **Revista Eletrônica de Comunicação Informação e Inovação em Saúde**, Rio de Janeiro, v. 3, n. 1, p. 8-20, 2009.

GOOGLE. **Google Scholar**. [S. l.], [2019]. Acesso em: 30 jul. 2019.

GUBA, E. **The paradigm dialog**. Newbury Park: Sage Publications, 1990.

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. 2005. Thesis (Doctorate in Computer Science) - Centre for Telematics and Information Technology, University of Twente, Enschede, 2005.

HESSEN, J. **Teoria do Conhecimento**. São Paulo: Martins Fontes, 2003.

IAO. **Ontobee**: Information Artifact Ontology. [S. l.]: University of Michigan Medical School, [2019]. Acesso em: 7 nov. 2019.

JUPP, S. *et al.* Document Navigation: ontologies or knowledge organisation systems. In: INTERNATIONAL WORKSHOP ON NETWORK TOOLS AND APPLICATIONS IN BIOLOGY, 7, 2007, Palermo. **Proceedings** [...]. Palermo: University of Manchester, 2007. p. 1-7.

LAURENTI, R. *et al.* A Classificação Internacional de Doenças, a Família de Classificações Internacionais, a CID-11 e a Síndrome Pós-Poliomielite. **Arquivos de Neuro-Psiquiatria**, São Paulo, v. 71, n. 9A, p. 3-10, 2013.

MACULAN, B. C. M. dos S. **Estudo e aplicação de metodologia de reengenharia de tesouro**: reformulação do THESAGRO. 2014. Tese (Doutorado em Ciências da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

MARCO-RUIZ, L. *et al.* Alignment of information models and domain ontologies. In: MARCO-RUIZ, L. *et al.* **Ontology-based terminologies for healthcare**: impact assessment and transitional consequences for implementation. Oslo: Norwegian Centre for E-health Research, 2017. cap. 8.

MERRILL, G. H. Ontological realism: methodology or misdirection? **Applied Ontology**, Amsterdam, v. 5, p. 79-108, 2010.

MURPHY, G. **Conceptual Combination**: the big book of concepts. Cambridge: The MIT Press, 2002.

OGMS. **Summary**. [S. l.]: BioPortal, 4 May 2019a. Acesso em: 2 nov. 2019

OGMS. **Welcome to BioPortal [...]**. [S. l.]: BioPortal, 2019b. Acesso em: 7 nov. 2019.

RECTOR, A. L. Clinical terminology: why is it so hard? **Methods of Information in Medicine**, [S. l.], v. 38, n. 4, p. 239-252, 1999.

RUBIN, D. L.; SHAH, N. H.; NOY, N. Biomedical Ontologies: a functional perspective. **Briefing in Bioinformatics**, [S. l.], v. 9, n. 1, p. 75-90, 2008.

SCHULZ, S.; MARTINEZ-COSTA, C. How ontologies can improve semantic interoperability in health care. In: RIAÑO, D.; LENZ, R.; MIKSCH, S.; PELEG, M.; REICHERT, M.; TEN TEIJE, A. (ed.). **Process Support and Knowledge Representation in Health Care: AIME 2013 Joint Workshop, KR4HC 2013/ProHealth 2013**. Murcia: Springer International Publishing, 2013. p. 1-10. (Lecture Notes in Computer Science, v. 8268).

SCHULZ, S.; MARTÍNEZ-COSTA, C.; MIÑARRO-GIMÉNEZ, C. Lexical ambiguity in SNOMED CT. In: **ONTOLOGIES & DATA IN LIFE SCIENCES**, 2017, Bolzano. **Proceedings [...]**. Bolzano: [s. n.], 2017. p. 1-5.

SEARLE, J. R. **Mente, linguagem e sociedade**: filosofia no mundo real. Rio de Janeiro: Rocco, 2000.

SIWICKI, B. **How FHIR 4 will drive interoperability progress in healthcare**. [S. l.]: Healthcare IT News, 25 Apr. 2019. Acesso em: 22 abr. 2020.

SMITH, B. Administrative gender code system: the never-ending story. In: SMITH, B. **Blog HL7 WATCH**. [S. l.], 30 Jan. 2018. Acesso em: 2 nov. 2019

SMITH, B.; CEUSTERS, W. Ontological realism as a methodology for coordinated evolution of scientific ontologies. **Applied Ontology**, Amsterdam, v. 5, n. 3-4, p. 139-188, 2010.

SMITH, B.; MUNN, K.. **Applied Ontology**: an introduction. Frankfurt: Ontos-Verlag, 2008.

STENZHORN, H. *et al.* Adapting clinical ontologies in realworld environments. **Journal of Universal Computer Science**, [s. l.], v. 14, n. 22, p. 3767-3780, 2008.

TEIXEIRA, L. M. D. **Princípios ontológicos aplicados à classificação Internacional de doenças**: alternativas para a busca por interoperabilidade semântica entre sistemas de prontuários eletrônicos de pacientes. 2019. Tese

(Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

UJVARI, S. C.; ADONI, T. **A história do século XX pelas descobertas da medicina**. São Paulo: Contexto, 2014.

UKOLN. **Interoperability focus**: looking at interoperability. Bath, 2005. Acesso em: 4 jun. 2013.

WAND, Y.; WEBER, R. Mario Bunge's ontology as a formal foundation for information systems concepts. *In*: WEINGARTNER, P.; DORN, J. W. G. (ed.). **Studies on Mario Bunge's treatise**. Amsterdam: Rodopi, 1990. p. 123-153.

WANG, K. Standard lexicons, coding systems and ontologies for interoperability and semantic computation in imaging. **Journal of Digital Imaging**, [s. l.], n. 31, p. 353-360, 2018.

WIKIPÉDIA. **Daltonismo**. [S. l.], abr. 2017. Acesso em: 22 abr. 2020.

ZINGANO, M. Unidade do gênero e outras unidades em Aristóteles: significação focal, relação de consecução, semelhança e analogia. **Analytica**, Rio de Janeiro, v. 17, n. 2, p. 395-432, 2013.

Compositionality and overlapping in biomedical terminology: alternatives for health interoperability

Abstract: A paramount information source within clinical practice is the patient record, an information-filled document that already exists in electronic format. The inherent ambiguity of the medical discourse has hampered the much sought-after automatic integration between electronic health records since, at this stage, automated systems are unable to address the syntactic and semantic complexity of the human language. Even standard terminologies designed to address such complexity are unable to solve the issues involved because of the multiplicity of alternatives, in addition to the fact that they are created for different purposes and goals. All this has resulted in redundancy, rework, and overlapping between terms. This paper assumes that the issue of interoperability is complex and needs to be tackled on several fronts. One of these fronts, in the context of electronic healthcare records, involves two conceptual conditions: epistemological overlapping and term compositionality. After explaining those conditions, we make use of an ontology-based model created on previous research to design and to test a method that aims to compensate for the shortcomings of such conditions. The test was performed on a fragment of the International Classification of Diseases. We found that the aforementioned conceptual conditions are relevant aspects to be faced in seeking interoperability.

Keywords: Ontology. Knowledge representation. Medical terminologies. Interoperability.

Recebido: 11/11/2019

Aceito: 26/12/2019

Como citar:

TEIXEIRA, Livia Marangon Duffles; ALMEIDA, Mauricio Barcellos. Composicionalidade e sobreposição em terminologias biomédicas: alternativas para interoperabilidade em saúde. **Em Questão**, Porto Alegre, v. 26, n. 3, p. 196-223, set/dez. 2020. DOI: <http://doi.org/10.19132/1808-5245263>.



¹ Traduzido para o português: Interoperabilidade em saúde.

² O *blockchain* é um recurso digital incorruptível que tem sido utilizado principalmente em transações econômicas e financeiras.

³ Princípio segundo o qual o significado de uma expressão complexa é função do significado das suas partes e da forma como estão associadas.

⁴ Status da doença arterial coronariana após cirurgia de revascularização miocárdica (CAGB = *Coronary Artery Bypass Graft Surgery*).

⁵ Aula do professor Mauricio Almeida na University of Madeira, 2013.

⁶ William Farr, 1807-1883, médico epidemiologista britânico, considerado um dos fundadores da estatística médica.

⁷ Jacques Bertillon, 1851-1922, médico estatístico francês.

⁸ Daltonismo, discromatopsia ou discromopsia é uma perturbação da percepção visual caracterizada pela incapacidade de diferenciar todas ou algumas cores (WIKIPÉDIA, 2017).

⁹ Tradução nossa para a expressão do inglês “*epistemological overlapping*”.

¹⁰ Fala do professor Mathias Brochhausen em aula na University of Arkansas For Medical Science, 2018.

¹¹ Classificação Internacional para a Prática de Enfermagem.

¹² *International Classification of Functioning, Disability and Health*.

¹³ *International Classification of Health Interventions*.

¹⁴ Friedrich Ludwig Gottlob Frege, 1848-1925, filósofo alemão, autor do *Princípio Extensional da Composicionalidade*.

¹⁵ Um fragmento da lógica de primeira ordem amplamente utilizado na web semântica, criado por filósofos alemães nos primeiros anos de 1990.

¹⁶ A OGMS é uma ontologia de domínio que representa as etapas de uma doença, desde o processo etiológico até o tratamento (OGMS, 2019b).

¹⁷ A DO ([2019]) é uma ontologia de domínio criada para organizar ontologicamente doenças, diagnósticos, fenótipos e desordens, dentre outras entidades afins.

¹⁸ A IAO ([2019]) é uma ontologia de nível médio, de domínio neutro, para a representação de tipos de entidades de conteúdo de informação, como documentos e imagens.

-
- ¹⁹ A BFO é uma ontologia de alto nível, de domínio neutro, desenvolvida com o propósito de prover os níveis mais genéricos (ou fundamentais) para classificação de entidades do mundo, especialmente na área biomédica. Disponível em: <https://basic-formal-ontology.org/>. Acesso em: 7 nov. 2019.