

Redução de ambiguidade em sistemas de informação clínicos: ontologias aplicadas à terminologias

Livia Marangon Duffles Teixeira; Mauricio Barcellos Almeida¹

¹ PPGGOC – ECI/UFMG – Belo Horizonte – MG – Brazil

liviamarangon@gmail.com; mba@eci.ufmg.br

***Abstract.** This paper presents an ongoing PhD project, which will be submitted to the examination board in February 2020, whose research question is the inherent ambiguity of medical terminologies. The methodology consists basically in anchoring to ontological resources, terms representing medical entities that exhibit epistemological features. There will be a double binding process with the aim of reducing terms ambiguity, which we believe essential to enable semantic interoperability between medical systems, a very important kind of issue in the ontology research.*

***Resumo.** Este artigo apresenta um projeto de pesquisa em andamento, com previsão de defesa para fevereiro de 2020, cujo problema de pesquisa é a ambiguidade de terminologias clínicas. A metodologia consiste basicamente em ancorar termos representativos de entidades médicas, que exibam características epistemológicas, a recursos ontológicos. O duplo processo de ancoragem proposto objetiva reduzir a ambiguidade dos termos e expressões, de forma a mitigar a falta de interoperabilidade semântica entre sistemas médicos, uma questão amplamente pesquisada na área de ontologias.*

1. Introdução

A proliferação de sistemas de informação na área de medicina e cuidados à saúde ocorre simultaneamente à evolução das diversas terminologias clínicas. Um dos aspectos responsáveis pela dificuldade na comunicação entre sistemas é a sobreposição de aspectos epistêmicos aos termos das terminologias que representam entidades do mundo real (Bodenreider, Smith, Burgun, 2004). Nesse contexto, a presente pesquisa consiste em elaborar e validar uma metodologia capaz de eliminar a sobreposição epistemológica presente em terminologias clínicas por meio da integração da SNOMED CT (enquanto terminologia desenvolvida para conectar sistemas de informação médica) ao OpenEHR (modelo de informação clínica para desenvolvimento de prontuários eletrônicos de paciente) e utilizando-se das teorias e práticas das ontologias. Entende-se que as ontologias, com suas capacidades de explicitar o conhecimento em linguagem de representação lógica, podem suprir o hiato na conexão de uma terminologia padrão a um modelo de informação clínico, provendo a interoperabilidade semântica entre sistemas de prontuários. A contribuição da pesquisa consiste no avanço das técnicas de modelagem baseadas em sistemas de classificação, amplamente pesquisadas na Ciência da Informação. Os resultados obtidos serão aplicados aos sistemas de informação clínicos na busca por mitigar problemas de interoperabilidade verificados em um hospital universitário brasileiro. Em última instância, buscam-se melhorias no atendimento e no

cuidado continuado ao cidadão, objetivo último da Ciência da Informação enquanto ciência social. Esse trabalho tem sido desenvolvido no contexto de pesquisa de doutorado, com previsão de defesa para fevereiro de 2020.

2. SNOMED CT

A SNOMED-CT é uma terminologia clínica global que inclui todos os domínios da saúde. Ela foi criada para representar, de forma consistente e processável, a informação encontrada em prontuários eletrônicos do paciente (Barra e Sasso, 2011). A terminologia pode ser implementada de diferentes formas, dentre elas, em aplicações de prontuários médicos. Outro fator de destaque é que a terminologia tem sido traduzida para OWL, o que possibilita a redução de ambiguidades. A SNOMED CT foi estabelecida como terminologia médica oficial brasileira (Brasil, 2011) e em 2018 o Brasil passou a integrar o grupo de países membros do SNOMED Internacional.

2.1. Sobreposição de Aspectos Epistêmicos

Uma característica de diversas terminologias clínicas, inclusive da SNOMED CT, é sobreposição de aspectos epistemológicos a termos usados para representar o mundo real. A sobreposição epistêmica pode ser definida como a presença de dados adicionais aos termos da terminologia que, embora sejam cruciais para atendimento clínico, não constituem uma representação fidedigna da realidade no domínio médico. Isso quer dizer que é possível observar dados complementares na descrição das classes, que nada tem a ver com a definição ou categorização daquela classe, como por exemplo (Bodenreider, Smith, Burgun, 2004): i) “Convulsão afebril” - não captura a essência ou origem da ocorrência da convulsão e o termo “afebril” não é adequado para representação; ii) *"Tuberculose das glândulas adrenais, bacilos tuberculosos não encontrados (no escarro) por microscopia, mas encontrados por cultura bacteriana"* - fornece dados complementares sobre como o médico obtém conhecimento sobre a doença; iii) *"Fratura de crânio sem lesão intracraniana"* - comunica que a possibilidade de lesão intracraniana foi verificada e descartada, o que não é necessário para a definição; iv) *"Abscesso tubo-ovariano possível"* - reflete uma convicção momentânea do médico, dentre outros. De fato, “é importante por uma série de razões que as classes denotadas por termos clínicos representem o mais próximo possível das classes naturais que existem na realidade” (Bodenreider, Smith, Burgun, 2004)¹

3. OpenEHR

O OpenEHR disponibiliza padrões para o desenvolvimento de Prontuários Eletrônicos de Pacientes (PEPs), buscando a comunicação e o compartilhamento sem perda de informações e significados. Os arquétipos são o principal diferencial entre o OpenEHR e os demais modelos de informação (Beale e Heard, 2007). Arquétipos representam o conhecimento clínico por meio de classes de informações que configuram a entrada de dados em um sistema de informação. Para operacionalizar a interoperabilidade, é essencial que outros sistemas baseados em OpenEHR sejam construídos a partir dos mesmos arquétipos. Além disso, é fundamental a utilização de uma terminologia clínica

¹ It is important for a number of reasons that classes denoted by biomedical terms represent as closely as possible the genuine classes which exist in reality.

para a representação das informações dos pacientes. A partir desses fatos, alguém poderia dizer que o OpenEHR lembre mais a proposta de um sistema federado, onde há acordos anteriores entre as partes, e apenas através desse tipo de acordo é obtida alguma integração. Nesse contexto, é importante entender a conexão entre ontologias e terminologias (seção 3.1) e do que trata a interoperabilidade semântica (seção 3.2).

3.1. Papel das Terminologias e Ontologias

Bacelar e Correia (2015) explicam que os reais benefícios dos PEPs dependem da estruturação e codificação dos dados usualmente descritos em texto livre. Assim, cada termo do arquétipo pode ser vinculado a uma terminologia de forma a melhorar a sua compreensão, reduzir a ambiguidade e incrementar o uso de dados no processo decisório. Para o compartilhamento de dados dos prontuários é essencial explicitar o significado dos termos. A maioria dos sistemas terminológicos e classificatórios não foi criada para o propósito de automação e sim, para uso manual, de forma que têm sido colocados à prova nesse sentido (Andrade, 2013). A semântica bem definida de um termo é caracterizada pelo seu significado construído por regras simples e por meio dos seus componentes e conectivos (Touretzky, 1986). As ontologias são capazes de tornar explícitos os compromissos ontológicos, determinando significado restrito para o termo específico (Gómez-Pérez et al., 2004). Vincular ontologias aos modelos de informação clínica não é uma tarefa trivial (Rector et al., 2009), uma vez que estes são desenvolvidos e apoiados na prática clínica baseada em linguagem natural e as ontologias em lógica descritiva.

3.2. Interoperabilidade Semântica

A interoperabilidade semântica é a capacidade dos sistemas se comunicarem, integrando desde o registro clínico até normas e diretrizes, sem que haja perda de informações e significado ou interferência humana. Ao assegurar sua viabilidade, garante-se que as mesmas conclusões sejam obtidas quando diferentes partes interessadas interpretarem um conjunto de dados (Marco-Ruiz et al., 2017). As terminologias nasceram para propósitos de padronização, mas ainda assim carregam ambiguidades. Apenas ontologias, sem interferência de aspectos epistemológicos, pode reduzir tais ambiguidades.

4. Metodologia

A busca pela interoperabilidade semântica está originalmente associada à aplicação das terminologias clínicas aos PEPs. Porém, as terminologias sofrem de sobreposição epistemológica, mesmo naquelas que têm sido traduzidas para OWL. Nesse caso, a descrição de instâncias presentes nos PEPs não está relacionada às entidades (universais) do domínio. Torna-se então necessário o uso de ontologias, com sua capacidade de descrever partes da realidade com redução de ambiguidade, o que as torna mais estáveis para representar o vínculo *universal > instâncias particulares* no registro de saúde. A seguir, são apresentados os passos metodológicos para a condução empírica da pesquisa, que em seu decorrer ou ao final, serão passíveis de adaptação e generalização.

4.1. Passo 1: Selecionar termos organizados de acordo com o OpenEHR

Nesse primeiro passo, por meio de um recorte de universo a ser realizado, serão selecionados em conjunto com o especialista da área do conhecimento, campos reais dos

arquétipos os quais se apresentam com alguma ambiguidade, ou a não existência de semântica bem definida. Um exemplo são as diversas declarações de um modelo que encontram diversas possibilidades de mapeamento na terminologia.

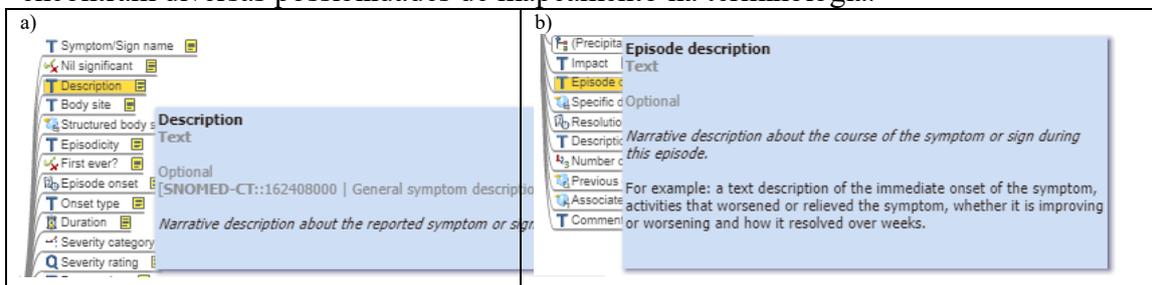


Figure 1 – (a) (b) Exemplos de descrição de entidades do Arquétipo Symptom

4.2. Passo 2: Ancorar entidades do OpenEHR à SNOMED CT

Ao realizar essa ancoragem (assinalada em vermelho na figura 2), será possível a duas unidades de saúde intercambiar dados entre sistemas com a garantia de que, pelo menos, os campos dos prontuários de pacientes têm o mesmo significado. Entretanto, o conteúdo do campo continua conectado ao preenchimento em texto livre pelos médicos.

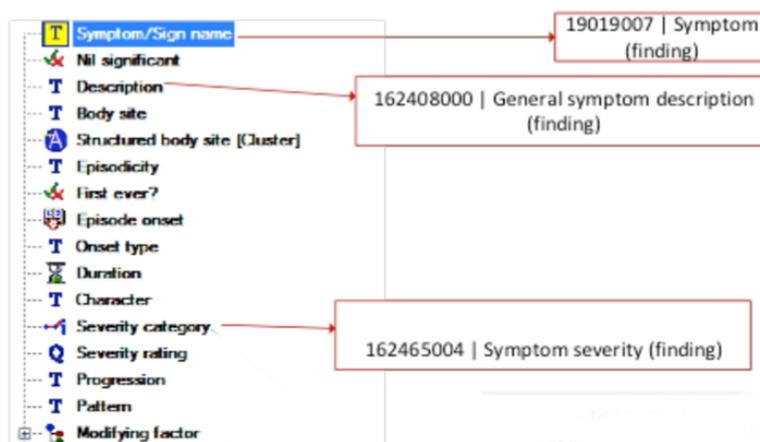


Figure 2 – Fragmento OpenEHR ancorado em termos do SNOMED CT

4.3. Passo 3: Identificar a sobreposição de aspectos epistêmicos na SNOMED CT

Identificar a sobreposição dentre os termos utilizados na ancoragem do passo 2. Esse passo é fundamental na metodologia e discute-se a possibilidade de automação, usando recursos de processamento de linguagem natural para marcar partes do discurso no texto das terminologias.

4.4. Passo 4: Ancorar termos da SNOMED CT à ontologias reais

Nesse passo, será feita uma nova ancoragem, dessa vez, da SNOMED CT a uma ontologia de alto nível livre de aspectos epistêmicos, a saber, a BFO (Smith et al., 2007). A ontologia de nível médio chamada de Ontology for General Medical Science (OGMS) (Scheuermann, Ceusters e Smith, 2009), desenvolvida abaixo da BFO, pode ser incluída para melhor clarificar o significado de termos. Ao realizar essa segunda ancoragem, será

possível avaliar as possibilidades de integração de sistemas, uma vez que serão obtidas conexões com redução de ambiguidades.

"Uma mulher de 88 anos apresentou-se na emergência em estado de confusão. Um ano antes, ela começou a ter episódios transitórios de confusão, tontura, tremor e ansiedade. Esses episódios eram imprevisíveis, tendo a duração de alguns minutos, e depois diminuindo espontaneamente. Entretanto, estavam se tornando mais frequentes. O paciente se sentia bem no intervalo entre os episódios e não relatava qualquer sensação anormal, alteração de peso, ou sintomas relacionados a refeições, jejum ou atividades físicas"

(1) Dados representando aspectos da realidade	Entidade da ontologia realista	(2) Dados que representam constructos úteis para a prática médica
Profissional de saúde	continuante BFO	... episódios transitórios de confusão, tontura, tremores e ansiedade ocorridos um ano antes Sem sensação anormal ... episódios são imprevisíveis Confusão Geral: boa aparência Peito: claro para auscultação Abdômen: macio e indolor
Mulher	continuante BFO	
88 anos de idade	qualidade BFO	
Relato do paciente	entidade informacional IAO	
Confusão, tontura, tremor	sintoma OGMS	
Duração dos episódios	região temporal BFO	
Tempo entre episódios	região temporal BFO	
Relato de alteração de peso	sintoma OGMS	
Aspirina	continuante BFO	
Aspirina tomada diariamente	regras AOI	
Resultado do exame físico	achado de exame físico OGMS	
Glicose	continuante BFO	
Diagnóstico de hipoglicemia	diagnóstico OGMS	
Insulinoma	continuante BFO	
(3) Dados que representam observações sobre a realidade		(4) Dados que representam observações sobre o entendimento dos profissionais de saúde
Frequência dos episódios Aumento na frequência dos episódios 36° de temperatura 76 batimentos por minuto 114/60 mmHg Concentração de glicose Dose de aspirina		Insulinoma causando hipoglicemia Relação entre sintomas e refeições

Figure 3 – Exemplo de análise ontológica com a BFO e OGMS de dados de prontuário

4.5. Passo 5: Analisar resultados da ancoragem

Os resultados serão analisados quanto à precisão e redução da ambiguidade de termos, bem como quanto a possíveis impactos para os sistemas de informação médica.

4.6. Passo 6: Validação da metodologia de integração

Após análise da ancoragem, será estabelecida a validação da integração da SNOMED ao OpenEHR por meio de consulta a profissionais da área de saúde, de forma a verificar se as etapas realizadas para mostrar a possibilidade de integração, encontram respaldo e concordância dentre os especialistas. Ainda não está definida a amostra de termos, nem os profissionais que participarão dessa tarefa. Isso é o que se espera de resultado empírico para criar a metodologia, uma vez que integração real de sistemas médicos específicos não seria passível de teste no ambiente do hospital universitário em referência.

5. Considerações Finais

Os objetivos da ontologia ficam claros em uma explicação de Brachman, citado em Rector et al. (2009), onde as ontologias são tipos de “cabides conceituais” para outros registros, às vezes de natureza não ontológica. Guizzardi (2007) corrobora com essa ideia e exemplifica o aporte teórico da metodologia apontando a divisão entre ontologia e epistemologia: [...] uma ontologia é uma especificação conceitual que descreve o conhecimento sobre um domínio de forma **independente dos estados epistêmicos e estado de coisas**.²(Guizzardi, 2007, p.8 - grifo nosso). Essa afirmações de eminentes pesquisadores da área respaldam a busca por uma metodologia, teste e validação da

² An ontology is a conceptual specification that describes knowledge about a domain in a manner that is independent of epistemic states and state of affairs.

sobreposição epistemológica. A metodologia pretende estabelecer os passos para a revisão e melhorias no vínculo entre a SNOMED CT (que não é isenta de estados epistêmicos) e o OpenEHR, a partir de recursos baseados em ontologias. Diante do exposto, espera-se com a identificação da sobreposição de aspectos epistemológicos, ancorar os termos passíveis de ambiguidade em ontologias (ou recursos ontológicos), de forma a mantê-los claros.

6. Referências

- Andrade, A. Q. de. (2013). “A linguagem médica utilizada em prontuários e suas representações em Sistemas de Informação: as ontologias e os modelos de informação”. UFMG, Escola de Ciência da Informação.
- Bacelar, G.; Correia, R. (2015). “As bases do openEHR: v.1.0”. Porto: Virtual Care.
- Barra, D. C. C.; Sasso, G. T. M. D. (2011). “Padrões de dados, terminologias e sistemas de classificação para o cuidado em saúde e enfermagem”. *Rev. bras. enferm.*, v. 64, n. 6.
- Beale, T.; Heard, S. (2007). “An ontology-based model of clinical information”. *Stud Health Technol Inform*, v. 129.
- Bodenreider, O.; Smith, B.; Burgun, A. (2004). “The Ontology-Epistemology divide: a case study in medical terminology”. In: *Proceedings... FOIS*.
- Brasil. (2011). “Ministério da Saúde. Portaria n. 2073, de 31 de agosto de 2011”.
- Gómez-Pérez, A.; Fernández-López, M.; Corcho, O.(2004). “Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic web”. London: Springer-Verlag.
- Guizzardi, G. (2007). “On Ontology, ontologies, conceptualizations, modeling languages, and (Meta) models”. IOS Press, Amsterdam.
- Marco-Ruiz, L et. al (2017). “Alignment of information models and domain ontologies. In: *Ontology-based terminologies for healthcare: Impact assessment and transitional consequences for implementation - project report*”. cap.8. Norwegian Centre for E-health Research.
- Rector, A. L.; Qamar, R.; Marley, T. (2009). “Binding ontologies and coding systems to electronic health records and messages”. *Applied Ontology*, v. 4, n. 1.
- Scheuermann, R. H.; Ceusters, W.; Smith, B. (2009). “Toward an Ontological Treatment of Disease and Diagnosis”. *Proceedings... AMIA Summit on Translational Bioinformatics*.
- Smith, B. et al. (2007). “The OBO foundry: coordinated evolution of ontologies to support biomedical dataintegration”. *Nature Biotechnology*, v. 25, n. 11.
- Touretzky, D.S. (1986). “The mathematics of inheritance systems”. Morgan Kaufmman.