



XXII Encontro Nacional de Pesquisa em Ciência da Informação – XXII ENANCIB

ISSN 2177-3688

GT-11- Informação & Saúde

MODELO DE MAPEAMENTO SEMÂNTICO ENTRE TERMINOLOGIAS DE SAÚDE

MAPPING MODEL BETWEEN HEALTH TERMINOLOGIES

Fabício Amadeu Gualdani. UNESP.

Leonardo Castro Botega. UNESP.

Nelson Júlio de Oliveira Miranda. USP.

Modalidade: Trabalho Completo

Resumo: A Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde e a Nomenclatura Sistematizada de Medicina são terminologias que visam a transparência dos dados. Terminologias possuem diferenças em suas composições, sendo necessário um mapeamento entre esses termos para que um sentido possa ser obtido, buscando aprimorar o cotidiano de profissionais da saúde com os seus pacientes por um modelo que estruture as informações de forma compreensiva de maneira sintática e semântica. O objetivo desta pesquisa é desenvolver um modelo para o mapeamento semântico entre estas terminologias de saúde. Trata-se de uma pesquisa exploratória, um estudo de caso realizado no Hospital das Clínicas da Faculdade de Medicina de Marília, que forneceu os códigos da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde registrados nos prontuários para a realização do mapeamento, visando migrar os dados armazenados que se encontravam em um banco de dados relacional para uma rede internacional de estrutura e compartilhamento de dados. Os resultados evidenciaram que há quatro tipos de situações durante a realização do mapeamento: exatidão semântica entre as terminologias, uso de expressões que tornam a condição de saúde genérica, termos que não são exatamente equivalentes, no entanto possuem aproximação semântica, assim como uma variedade de termos para representar uma única condição de saúde. Conclui-se que é possível desenvolver um modelo replicável que preserve a camada semântica dos termos entre a Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde e a Nomenclatura Sistematizada de Medicina.

Palavras-Chave: Prontuário eletrônico do paciente. Classificações em Saúde. Terminologias em saúde. Mapeamento semântico.

Abstract: The International Statistical Classification of Diseases and Health Problems and the Systematized Nomenclature of Medicine are terminologies that aim for data transparency. Terminologies have differences in their compositions, and a mapping between these terms is necessary in order to obtain a meaning, seeking to improve the daily life of health professionals with their patients through a model that structures the information in a comprehensive syntactic and semantic way. The goal of this research is to develop a model for semantic mapping between these health terminologies. This is an exploratory research, a case study carried out at the Hospital das Clínicas da Faculdade de Medicina de Marília, which provided the International Statistical Classification of Diseases and Health-Related Problems codes recorded in the medical records for the mapping, aiming to migrate the stored data that were in a relational database to an international structure and



data sharing network. The results showed that there are four types of situations during the mapping process: semantic accuracy between terminologies, use of expressions that make the health condition generic, terms that are not exactly equivalent but are semantically close, and a variety of terms to represent a single health condition. It is concluded that it is possible to develop a replicable model that preserves the semantic layer of terms between the International Statistical Classification of Diseases and Health Problems and the Systematized Nomenclature of Medicine.

Keywords: Electronic patient record. Health Classifications. Health Terminologies. Semantic mapping.

1 INTRODUÇÃO

O prontuário do paciente consiste na documentação da qual um médico ou um profissional de saúde preenche todas as informações relativas ao estado de saúde física, mental, assim como todas as condições sociais vivenciadas por um paciente. Esse registro pode tanto ser efetuado em seu formato tradicional em papel (analógico) como em suporte eletrônico graças ao avanço das tecnologias da informação que possibilitam a presença de sofisticados sistemas para a realização desses registros (SILVA, 2021). No entanto, ao se realizar o processo de anamnese, isso é, a entrevista, o atendimento prestado a um paciente por um profissional de saúde, uma série de dificuldades em se realizar o preenchimento adequado dos campos oferecidos pelo prontuário podem ocorrer, sendo os mais conhecidos: erros ortográficos, gramaticais e o uso excessivo de siglas e abreviações (CARVALHO, 2018). Como é possível observar, esses problemas dizem respeito aos termos inseridos no prontuário, portanto são obstáculos de natureza terminológica e justamente visando amenizar esses problemas, trazendo uma maior padronização para o preenchimento dessas informações, assim como um uso mais organizado dos termos referentes às diferentes condições do paciente, nascem as terminologias, que consistem em uma espécie de vocabulário controlado para a organização e a estruturação de termos de saúde, podendo conter termos que dizem respeito a: doenças e enfermidades, medicamentos, cirurgias, procedimentos, estruturas fisiológicas dentre muitos outros aspectos voltados ao campo da saúde (SHIVERS et al., 2021). Duas das principais terminologias que atualmente atendem as necessidades da área da saúde, consistem na Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10) desenvolvida pela Organização Mundial da Saúde e mantida pelo Ministério da Saúde brasileiro oferecendo termos especificamente como o seu próprio nome sugere, de doenças e enfermidades. Enquanto a *Systematized Nomenclature of Medicine* (SNOMED-CT) trata-se da terminologia clínica mais abrangente do



mundo, contendo não apenas termos relacionados a doenças, como também procedimentos, cirurgias, medicamentos, situações, lugares, alimentos, objetos, etc. (SHIVERS et al., 2021). Essas duas terminologias possuem diferenças intrínsecas em suas composições e nos objetivos que elas pretendem atingir, portanto não há uma equivalência sintática, isso é, no que diz respeito a composição, a forma de um termo, e nem uma equivalência semântica, que diz respeito ao significado, ao entendimento do conteúdo daquele termo. Por isso, se torna necessário a realização de um mapeamento entre esses termos seja por meio de algoritmos e outras técnicas computacionais ou até mesmo procedimentos manuais realizados por uma equipe multiprofissional contando com a presença de profissionais da área da Ciência da Informação, Ciência da Computação e Ciências da Saúde para tratar essas informações, buscando gerar uma melhor organização e estruturação das mesmas (de maneira sintática e semântica) para que assim seja possível em um momento futuro recuperar essas informações para as mais diversas utilizações, principalmente pensando em um melhor diagnóstico e tratamento direcionado ao paciente graças ao bom uso dessas informações clínicas pelo profissional de saúde durante o atendimento prestado. A problemática desta pesquisa se encontra na necessidade do Hospital das Clínicas da Faculdade de Medicina de Marília de realizar uma migração dos dados oriundos de seus prontuários eletrônicos armazenados em um banco de dados tradicional Oracle, da qual obedeciam a terminologia CID-10 em sua versão de 2019, assim como o seu arquétipo está estabelecido pelo Conjunto Mínimo de Dados de Atenção à saúde (CMD) para uma rede internacional de estrutura e compartilhamento de dados O Common Data Model (OMOP) pertencente a iniciativa Observational Health Data Science Informatics (OHDSI) que possui como seu vocabulário padrão a terminologia SNOMED-CT, o que induz a realização do mapeamento dos termos entre ambas as terminologias.

A realização desta pesquisa se justifica pelo fato de que se torna necessário explorar a camada semântica dos prontuários eletrônicos, visando gerar uma compreensão tanto por parte da máquina (o computador) como para o ser humano no aspecto de representar e recuperar essas informações clínicas, auxiliando tanto em um diagnóstico mais preciso do paciente como em uma melhor tomada de decisão no que diz respeito ao seu tratamento. Da mesma forma que promover a interoperabilidade no nível semântico de forma internacional



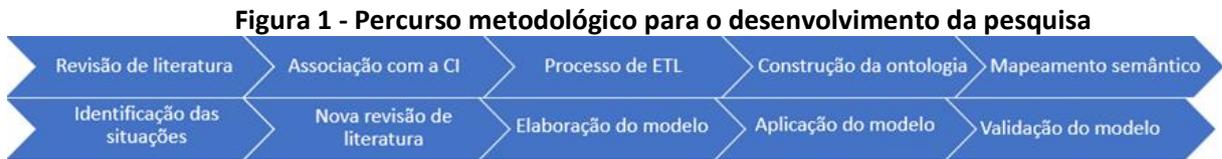
e padronizada, trata-se de uma iniciativa fundamental para uma melhor utilização dos dados gerados no campo da saúde.

O objetivo geral deste estudo consiste portanto na construção de um modelo que possa considerar a camada semântica de informações clínicas de prontuários disponibilizados em formato eletrônico pelo Hospital das Clínicas da Faculdade de Medicina de Marília, descrevendo de forma minuciosa o procedimento de mapeamento entre a classificação CID-10 em sua versão de 2019 e a terminologia clínica SNOMED-CT para a realização da migração destes dados para a rede internacional de estrutura e compartilhamento de dados (o OMOP) pertencente a iniciativa OHDSI .

Os procedimentos metodológicos adotados nesta pesquisa consistem do ponto de vista de sua natureza em uma pesquisa aplicada devido ao fato de que possuem como objetivo gerar conhecimentos para a aplicação prática dirigidos à solução de problemas específicos, envolvendo verdades e interesses locais. Do ponto de vista de seus objetivos trata-se de uma pesquisa exploratória, pois ela buscou proporcionar mais informações sobre a temática estudada (no caso, a representação e a recuperação das informações provenientes de prontuários disponibilizados em formato eletrônico) possibilitando a sua definição e o seu delineamento. Quanto aos procedimentos técnicos empregados, foi realizada uma pesquisa bibliográfica, da qual foram consultados artigos científicos nacionais e internacionais de diferentes bases de dados, livros, teses de doutorado, artigos publicados em congressos e eventos científicos, assim como atas, guias, manuais e documentos oficiais publicados por instituições, iniciativas e equipes de saúde. Para a realização desta pesquisa bibliográfica foram consultadas bases de dados pertencentes a Ciência da Informação, Ciência da Computação e Ciências da Saúde tais como: BRAPCI; Scopus; Scielo; Springer Link; IEEE Xplore e PubMed no período de abril até outubro de 2020 com as seguintes palavras-chave: prontuário eletrônico do paciente* electronic health record* terminologias* terminologies* recuperação semântica da informação* semantic information retrieval* interoperabilidade* interoperability* . Outro procedimento técnico empregado de fundamental importância para a elaboração desta pesquisa consiste no estudo de caso, que envolve o estudo profundo e exaustivo de um ou poucos objetos de maneira que permita o seu amplo e detalhado conhecimento, descrevendo a situação do contexto em que está sendo feita determinada investigação (PRODANOV; FREITAS, 2013).



Na figura abaixo, é possível visualizar de forma geral cada procedimento a ser executado para a concretização deste estudo.



Fonte: Elaborado pelos autores (2022).

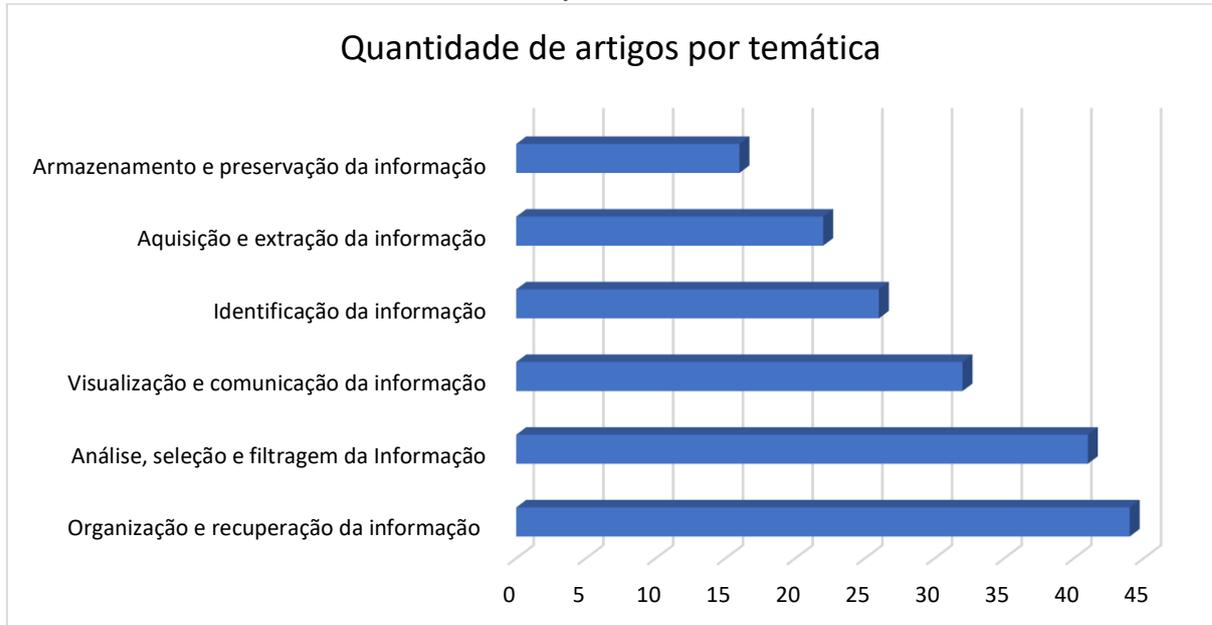
2 REVISÃO DE LITERATURA PARA A IDENTIFICAÇÃO DAS CONTRIBUIÇÕES DA CIÊNCIA DA INFORMAÇÃO PARA O PRONTUÁRIO DO PACIENTE

Para a realização desta pesquisa, foram considerados 56 estudos, sendo que 44 abordaram o tópico de organização e recuperação da informação, sendo esse o aspecto da qual apresentou uma maior quantidade de estudos considerados. Em seguida, temos como o segundo tópico mais abordado a análise; seleção e filtragem de informação, da qual se mostrou recorrente em 41 dos 56 estudos. O terceiro tópico da qual apresentou uma presença mais significativa, consistiu no tópico de visualização e comunicação da informação, da qual se fez presente em 32 dos 56 estudos. O quarto tópico da qual apresentou maior presença perante a revisão de literatura realizada, trata-se da identificação da informação, com 26 dos 56 estudos incluídos, no quinto tópico, aquisição e extração da informação, 22 de 56 estudos atenderam a esta temática. O sexto e último tópico é a criação, armazenamento e preservação da informação, se fez presente em 16 dos 56 estudos.

O gráfico a seguir nos apresenta uma visualização da quantidade de estudos que representaram determinado tópico sobre as necessidades informacionais do prontuário em seu formato eletrônico, que podem ser atendidas pela Ciência da Informação. Portanto, o gráfico está estruturado da seguinte forma: no lado esquerdo, é possível verificar a contribuição, enquanto no lado direito, as barras correspondem à quantidade de estudos inseridos nesta pesquisa que correspondem ao tópico descrito.



Gráfico 1 - Contribuições da Ciência da Informação para o prontuário eletrônico do paciente



Fonte: Elaborado pelos autores (2022).

3 MODELO DE MAPEAMENTO SEMÂNTICO ENTRE AS TERMINOLOGIAS

O grupo de estudos Health Artificial Intelligence Study (HAIS) pertencente ao Hospital das Clínicas da Faculdade de Medicina de Marília buscou identificar qual era o modelo de composição, de estruturação das informações disponibilizadas no prontuário em formato eletrônico oferecido pelos sistemas do hospital, chegando à conclusão de que se tratava do Conjunto Mínimo de Dados de Atenção à Saúde, assim como os termos usados pertenciam a terminologia CID-10.

Essas descobertas foram feitas por meio de um processo de extração, transformação e carregamento dos dados (ETL- Extract, Transform, load) que estavam presentes em um banco de dados relacional Oracle. As ferramentas utilizadas em meio a esse procedimento consistiram no Oracle e o Pentaho Data Integrator. Uma vez feito o procedimento de ETL, buscando tratar a camada sintática dos dados, foi necessário a construção de uma ontologia no software Protégé que representasse, organizasse visualmente e estruturalmente os dados propostos pelo Modelo de Informação do CMD do SUS que são utilizados nos prontuários eletrônicos do Hospital das Clínicas da Faculdade de Marília (HCFAMEMA). Como é possível observar na figura abaixo como exemplo, as classes pertencentes a ontologia.



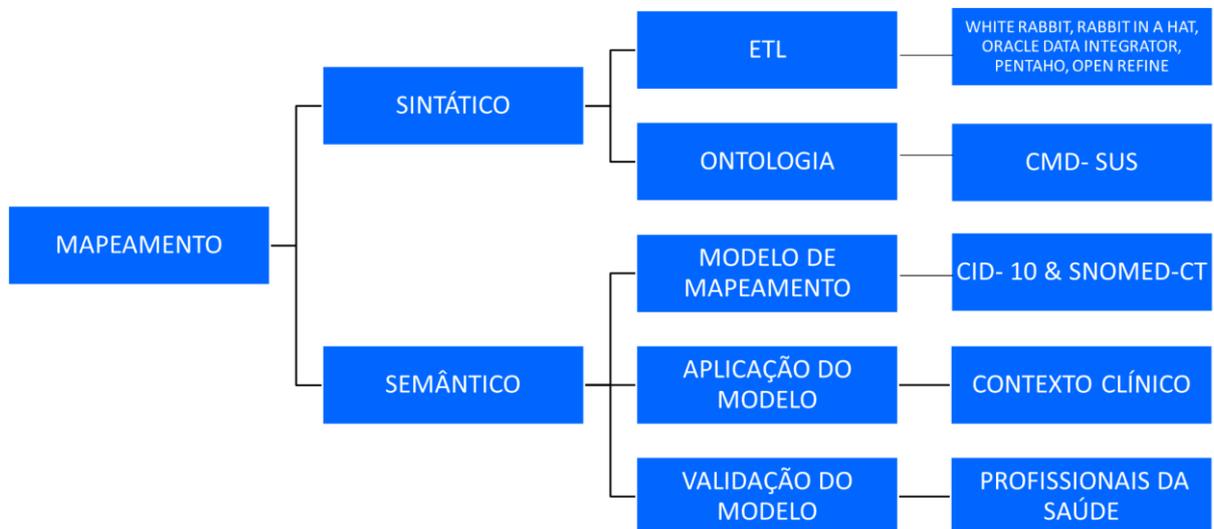
Figura 2 – Classes da ontologia do Conjunto Mínimo de Dados de Atenção à Saúde



Fonte: Elaborado pelos autores (2022).

Enquanto para que fosse possível trabalhar com os dados em um nível semântico, foi necessário a realização do mapeamento entre os dados que estavam descritos de acordo com a terminologia CID-10 para o formato SNOMED-CT que é o vocabulário padrão da base de dados OMOP-CDM. Atividade esta, que consistiu no foco desenvolvido para esta pesquisa, responsabilidade do GIHC. É possível visualizar o desenvolvimento da pesquisa como um todo na ilustração abaixo, sendo que a camada sintática da pesquisa se concentrou na responsabilidade do HAIS, enquanto a camada semântica foi de responsabilidade do GIHC.

Figura 3- Visão geral da pesquisa desenvolvida

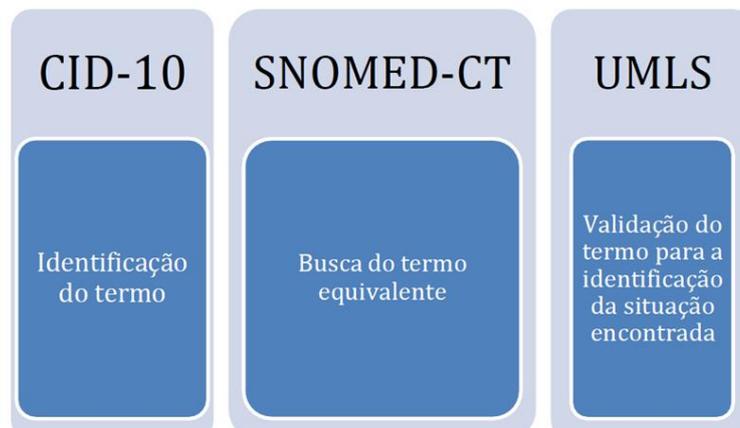


Fonte: Elaborado pelos autores (2022).



Como resultado do processo de mapeamento semântico entre os termos, da qual foi preciso realizar uma equivalência entre os termos da CID-10 e a SNOMED-CT, foi possível conceber um modelo que sintetiza o processo de identificação de cada termo entre as terminologias para logo em seguida, gerar uma validação na ferramenta Unified Medical Language System (UMLS) que consiste em um conjunto de arquivos e softwares que reúnem uma série de vocabulários e padrões de saúde para permitir interoperabilidade entre sistemas. A figura a seguir apresenta visualmente este modelo.

Figura 4- Modelo de mapeamento semântico entre as terminologias de Saúde CID-10 & SNOMED-CT



Fonte: Elaborado pelos autores (2022)

Dos 1608 termos disponibilizados, 212 até o momento foram mapeados. Os termos se referem a doenças e outros problemas de saúde, sendo que medicamentos, cirurgias e procedimentos não estão incluídos, pelo fato de se tratarem de termos codificados em CID-10. As colunas da tabela obedeceram a seguinte ordem: a primeira referente aos códigos disponibilizados em CID-10 em sua versão de 2019, a segunda com o seu código equivalente em SNOMED-CT, (SCTID) a terceira e a quarta, com o conteúdo, isso é, a enfermidade em si que determinado código da CID-10 e SNOMED-CT se refere, respectivamente, a quinta e a sexta coluna referem-se ao código CUI disponibilizado pela UMLS, isso é, o código interno da ferramenta para os termos sejam eles da CID-10 ou SNOMED-CT e a última coluna apresenta qual das quatro situações: exatidão semântica entre as terminologias, uso de expressões genéricas como “outros ou não especificado” aproximações semânticas entre as



terminologias e diversos termos para uma única condição de saúde. A figura a seguir apresenta um recorte desta tabela.

Tabela 1 – Mapeamento semântico entre as terminologias de saúde “Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde e a Nomenclatura Sistematizada de Medicina

| Códigos CID-10 Versão: 2019 | SCTID- SNOMED-CT SNOMED International 2021 v3.15.1 | Problema/ Enfermidade/ Condição de Saúde- CID-10 | Problema/ Enfermidade/ Condição de Saúde- SNOMED-CT | Código UMLS (CID-10) | Código UMLS (SNOMED -CT) | Situação |
|-----------------------------|--|--|---|----------------------|--------------------------|---|
| K42 | 396347007 | Umbilical hernia | Umbilical hernia (disorder) | C0019322 | C0019322 | (1) Exatidão semântica entre as terminologias |
| K46 | 52515009 | Unspecified abdominal hernia | Hernia of abdominal cavity (disorder) | C0178282 | C0178282 | (2) Uso de expressões genéricas como "outros" ou "não especificado" |

Fonte: Elaborado pelos autores (2022)

A primeira situação encontrada, referente a exatidão semântica entre as terminologias, como o próprio nome sugere, consiste em um "match", isto é, uma equivalência perfeita tanto sintática quanto semântica entre os termos de ambas as terminologias, não se fazendo necessários outros procedimentos. Dos 212 termos mapeados, 70 se encaixaram nessa categoria. Como é possível observar na figura abaixo, a equivalência exata do termo “Paranoid Schizophrenia” encontrado tanto na CID-10 como na SNOMED-CT.

Figura 5 – Exatidão semântica entre terminologias

F20.0 Paranoid schizophrenia

Paranoid schizophrenia is dominated by relatively stable, often paranoid delusions, usually accompanied by hallucinations, particularly of the auditory variety, and perceptual disturbances. Disturbances of affect, volition and speech, and catatonic symptoms, are either absent or relatively inconspicuous.

Paraphrenic schizophrenia

Excl.: involuntal paranoid state (E22.8)
paranoia (E22.0)

Paranoid schizophrenia (disorder)
SCTID: 64905009
64905009 | Paranoid schizophrenia (disorder) |
en Paranoid schizophrenia (disorder)
en Paranoid schizophrenia
en Paraphrenic schizophrenia

Fonte: Elaborado pelos autores (2022)

A segunda situação consiste no uso de expressões que tornam a doença ou a condição de saúde muito mais genérica, como o uso da expressão “outros” ou não “especificado.” Essa



é uma possibilidade oferecida pela terminologia CID-10 da qual não se faz presente de forma correspondente pela SNOMED-CT, a CID permite que um médico ou outro profissional de saúde ao não conseguir identificar adequadamente em uma anamnese qual é determinado problema que um paciente está passando, inserir a opção “outros.” Supondo-se que um paciente reclame de dor na região do abdômen, no entanto não fica exatamente claro para o médico ou outro profissional de saúde a partir do seu conhecimento prévio a causa ou o tipo de dor que aquela condição se refere, logo, a terminologia abre uma lacuna para que se possa preencher no prontuário, “Outros tipos de dores abdominais” enquanto a SNOMED-CT por mais que seja uma terminologia extremamente complexa, robusta e completa não oferece essa opção mais aberta, gerando um conflito, já que uma série de códigos oferecidos pela equipe do HAIS apresentaram essa possibilidade, sendo que não há uma equivalência para esse tipo de situação na SNOMED-CT.

A solução encontrada foi mapear todos os códigos que apresentaram termos como “outros” ou “não especificado” para a forma mais comum de se referir a determinado problema de saúde, por exemplo, caso o código apresente como no exemplo citado “Dor abdominal não especificada” o mapeamento equivalente na SNOMED-CT seria “Dor abdominal” ou seja, a forma genérica e comum possível do termo, como apresentado na figura abaixo.

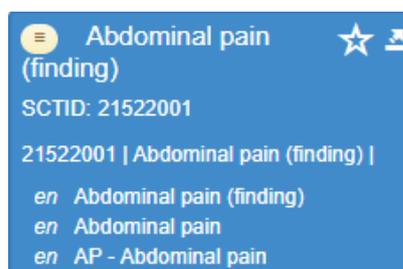
Figura 6- Uso de expressões genéricas como “outros” ou “não especificado”

R10.4 Other and unspecified abdominal pain

Abdominal tenderness NOS

Colic:

- NOS
- infantile



Fonte: Elaborado pelos autores (2022).

A terceira situação encontrada consiste nas aproximações, ou seja, termos que não exatamente equivalentes, no entanto possuem uma aproximação semântica. Sendo assim, o critério utilizado para vincular os termos entre as terminologias CID-10 e SNOMED-CT foi



observar o quanto o título do termo e a sua descrição em si eram equivalentes, aproveitando ao máximo a similaridade entre ambos. É possível visualizar na figura abaixo, em que o primeiro termo “*Malignant neoplasm of base of tongue*” presente na CID-10, encontra o seu correspondente na SNOMED-CT “*Primary malignant neoplasm of base of tongue*” na qual possui um maior grau de aproximação não somente pelo título das condições de saúde, mas pela descrição e composição semântica informacional presente na terminologia SNOMED-CT.

Figura 7- Aproximações semânticas entre as terminologias

| | |
|---------------|---|
| C01 | Malignant neoplasm of base of tongue |
| Incl.: | Dorsal surface of base of tongue Fixed part of tongue NOS Posterior third of tongue |
| |  |

Fonte: Elaborado pelos autores (2022).

Dos 212 termos mapeados, 88 se encaixam nesta categoria, sendo este o tópico mais representativo deste estudo.

A quarta e última situação consiste na presença de uma condição de saúde com diversos termos para representá-la, como por exemplo, na terminologia CID-10 temos a condição “*Sinus, fistula and Cyst Of Branchial cleft*”, na qual é possível perceber que se trata de um quadro clínico muito específico, sendo que esta situação pode ser dividida em múltiplos termos que se encontram de formas isoladas na terminologia SNOMED-CT. A figura abaixo ilustra essa situação.



Figura 8- Diversos termos para uma única condição de saúde

Q18.0 Sinus, fistula and cyst of branchial cleft

Branchial vestige

| | |
|---|---|
|  Sinus (morphologic abnormality) ☆ |  Branchial cleft cyst (disorder) ☆ |
| SCTID: 419351001 | SCTID: 59857007 |
| 419351001 Sinus (morphologic abnormality) | 59857007 Branchial cleft cyst (disorder) |
| <i>en</i> Sinus (morphologic abnormality) | <i>en</i> Branchial cleft cyst (disorder) |
| <i>en</i> Sinus | <i>en</i> Branchial cleft cyst |
| | <i>en</i> Branchial cyst |
| | <i>en</i> Cyst of branchial cleft |

Fonte: Elaborado pelos autores (2022)

Dos 212 termos mapeados, esta categoria se demonstrou como a menos representativa com apenas 10 termos identificados.

3 CONSIDERAÇÕES FINAIS

A Ciência da Informação graças a sua natureza interdisciplinar e colaborativa possibilita que os seus principais métodos e processos possam atender as demandas da área da saúde, mais especificamente do prontuário, que consiste em um objeto informacional para o auxílio de um médico ou uma equipe multiprofissional realizar o registro das informações referentes a um paciente a respeito do seu estado de saúde física, mental e emocional. Desde o seu formato em papel (analógico) o prontuário possui inúmeras dificuldades no que diz respeito aos processos de se representar adequadamente as suas informações de forma lógica e estruturada, para que logo em seguida, fosse possível realizar de forma adequada a sua devida análise, sintetização e extração dessas informações.

Com a disponibilização de seu novo formato (o eletrônico) os problemas ainda permanecem, sendo necessária a elaboração de ferramentas tecnológicas para lidar com essas situações. As terminologias surgem nesse cenário como uma alternativa para disponibilizar termos para um preenchimento mais padronizado do prontuário, sendo as duas mais utilizadas na área médica a CID-10 e a SNOMED-CT. Visando realizar um mapeamento que explorasse a camada semântica da informação entre as terminologias CID-10 e SNOMED-CT, o Grupo de Interação Humano- Computador pertencente a Unesp Campus de Marília e *Health Artificial Intelligence Study* grupo de estudos estabelecido no Hospital das Clínicas da



Faculdade de Medicina de Marília buscaram migrar os dados da base de dados atual do hospital (banco de dados relacional Oracle) para uma base de dados que possui uma integração e interoperabilidade global de seus dados, o OMOP-CDM oferecido pela iniciativa OHDSI assim como uma representação de suas informações, utilizando como vocabulário padrão a terminologia SNOMED-CT.

Com a realização da primeira etapa de todo o procedimento metodológico envolvido, a revisão de literatura, foi possível encontrar quais são as principais contribuições necessárias da área da Ciência da Informação ao alinhar os seus métodos com as demandas do prontuário eletrônico do paciente, sendo possível identificar os processos de: identificação; criação, armazenamento; preservação; análise; seleção; filtragem; organização; categorização; aquisição; extração; visualização e comunicação da informação. Sendo que a primeira conclusão a ser obtida por meio do desenvolvimento deste trabalho é que o tópico mais recorrente apresentado na literatura, foi o de organização da informação, assim como outros processos muito próximos como a seleção, categorização, filtragem e análise são de igual importância para um bom processo de recuperação da informação, no entanto, o tópico menos considerado em torno da literatura está ligado ao ato de se preservar a informação, característica essa de grande valor, justamente devido ao fato de que se em momentos futuros for necessário recuperar essa informação, ela deve estar devidamente armazenada. Logo em seguida, com a construção do mapeamento entre as terminologias, foi possível observar quatro tipos de situações que ocorreram com frequência: exatidão semântica entre as terminologias, quando dois termos eram exatamente iguais tanto em suas camadas sintáticas como semânticas em ambas as terminologias (CID-10 e SNOMED-CT). O uso de expressões genéricas como “outros” ou “não especificado” que constantemente se faz presente na CID-10, no entanto não se trata de uma opção disponibilizada pela SNOMED-CT, aproximações, termos que não exatamente equivalentes, no entanto possuem uma aproximação semântica. E a presença de diversos termos para uma única condição de saúde.

A ferramenta UMLS possibilitou um suporte, uma verificação de mais códigos para cada termo para alimentar a base de dados OMOP-CDM. Uma vez identificadas estas situações, é preciso definir critérios para lidar com cada tipo de problema de saúde encontrado, pois cada código possui as suas especificidades, sendo que uma vez que o modelo estiver finalizado e implementado no contexto clínico de um médico ou algum outro



profissional da área da saúde, ele precisará validar o modelo assim como realizar sugestões de aprimoramentos, processo esse chamado de DSL (*Domain Specific Language*) da qual é realizado uma “consulta” a um profissional específico de determinada área que valide a ferramenta, instrumento ou modelo proposto. Assim como é possível a elaboração de um algoritmo para auxiliar no trabalho dos profissionais da informação e computação quanto a definição desses critérios.

Os próximos passos a serem desenvolvidos para esta pesquisa, consistem na progressão do mapeamento, buscando a identificação de novas situações, a elaboração dos critérios em si sobre como proceder perante cada termo mapeado de saúde, na análise de mapeamentos já existentes, assim como uma introdução a décima primeira versão da CID que desta vez possui uma configuração pós-coordenada alterando a natureza do funcionamento de seus termos. Finalmente, é válido ressaltar estas considerações: a primeira é que o prontuário disponibilizado em seu formato eletrônico consiste em um objeto informacional riquíssimo para a exploração de todo o seu conteúdo informacional, pelas mais diversas áreas, e justamente a Ciência da Informação por possuir uma natureza colaborativa e interdisciplinar pode e deve atuar em parcerias com uma equipe multiprofissional de saúde assim como profissionais da área da computação buscando implementar mapeamentos, algoritmos e ontologias para uma melhor representação da informação disponibilizada, como é o caso específico desta pesquisa, da qual uma ontologia está em desenvolvimento para melhor representar o Conjunto Mínimo de Dados da Atenção à Saúde, que é o modelo de informação estrutural dos prontuários em formato eletrônico disponibilizados pelo Hospital das Clínicas da Faculdade de Medicina de Marília buscando alimentar a base de dados OMOP-CDM e pensando posteriormente na possibilidade de sua disponibilização online, contribuindo assim com outras pesquisas científicas assim como na formação de uma Web Semântica e em uma cultura de Medicina Baseada em Evidências. Conclui-se que é possível a realização de um mapeamento concentrado na camada semântica dos termos provenientes da terminologia CID-10 e SNOMED-CT.

REFERÊNCIAS

CARVALHO, R. C. Aplicação de mineração de dados em informações oriundas de prontuários de paciente. **INFORMAÇÃO EM PAUTA**, Fortaleza, v. 3, p. 161-181, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/106569>. Acesso em: 16 mai. 2022.



DA SILVA, C. R.; História do Prontuário Médico: Evolução do Prontuário Médico Tradicional ao Prontuário Eletrônico do Paciente–PEP. **PESQUISA, SOCIEDADE E DESENVOLVIMENTO**. v. 10, n. 9, pág. e28510918031-e28510918031, 2021.

GIHC. Grupo de Interação-Humano- Computador. Disponível em: <https://www.schoolandcollegelisting.com/BR/Makati/310439162788899/Grupo-de-Intera%C3%A7%C3%A3o-Humano-Computador---GIHC-Unesp>. Acesso em: 18 ago.2022

HAI. Health Artificial Intelligence Study. Disponível em: <https://hc.famema.br/hcfamema-implanta-o-hais/>. Acesso em: 18 ago.2022.

ICD-2019. International Statistical Classification of Diseases and Related Health Problems 10th Revision. Disponível em: <https://icd.who.int/browse10/2019/en#/>. Acesso em: 18 de ago.2022

OHDSI. Observational Health Data Sciences and Informatics. <https://www.ohdsi.org/data-standardization/the-common-data-model/>.

PRODANOV, C.C.; FREITAS, E.C. Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico. 2.ed. Novo Hamburgo: **FEEVALE**, 2013.

SHIVERS, J.; AMLUNG, J.; RATANAPRAYU, N.; RHODES, Bryn.; BIONDICH, P. Enhancing narrative clinical guidance with computer-readable artifacts: Authoring FHIR implementation guides based on WHO recommendations. **JOURNAL OF BIOMEDICAL INFORMATICS**, v. 122, p. 103891, 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34450285/>. Acesso em: 13 mai. 2022.

SNOMED-CT. Systematized Nomenclature of Medicine. Disponível em: <https://browser.ihtsdotools.org/?>. Acesso em: 18 ago.2022.