USE OF COMPUTATIONAL TOOLS AS SUPPORT TO THE CROSS-MAPPING METHOD BETWEEN CLINICAL TERMINOLOGIES

Denilsen Carvalho Gomes[1]
Lucas Emanuel Silva e Oliveira[1]
Marcia Regina Cubas[1]
Claudia Maria Cabral Moro Barra[1]

[1]Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Tecnologia em Saúde, Curitiba, PR, Brasil

ABSTRACT

Objective: to reflect on the use of computational tools in the cross-mapping method between clinical terminologies.

Method: reflection study.

Results: the cross-mapping method consists of obtaining a list of terms through extraction and normalization; the connection between the terms of the list and those of the reference base, by means of predefined rules; and grouping of the terms into categories: exact or partial combination or, in more detail, similar term, more comprehensive term, more restricted term and non-agreeing term. Performed manually in many studies, it can be automated with the use of the Unified Medical Language System (UMLS). Obtaining the terms list can occur automatically by natural language processing algorithms, being that the use of rules to identify information in texts allows the expert's knowledge to be coupled to the algorithm, and it can be performed by techniques based on Machine Learning. When it comes to mapping terms using the 7-Axis model of the International Classification for Nursing Practice (ICNP®), the process can also be automated through natural language processing algorithms such as POS-tagger and the syntactic parser.

Conclusion: the cross-mapping method can be intensified by the use of natural language processing algorithms. However, even in cases of automatic mapping, the validation of the results by specialists should not be discarded.

DESCRITORS: Terminology. Nursing. Informatics. Controlled vocabulary. Methods.

USO DE FERRAMENTAS COMPUTACIONAIS COMO AUXÍLIO AO MÉTODO DE MAPEAMENTO CRUZADO ENTRE TERMINOLOGIAS CLÍNICAS

RESUMO

Objetivo: refletir sobre o uso de ferramentas computacionais no método de mapeamento cruzado entre terminologias clínicas.

Método: estudo de reflexão.

Resultados: o método de mapeamento cruzado consiste na obtenção de listagem de termos, por meio de extração e normalização; ligação entre os termos da listagem e os da base de referência, mediante regras previamente definidas; e agrupamento dos termos em categorias: combinação exata ou parcial ou, de maneira mais detalhada, termo similar, termo mais abrangente, termo mais restrito e termo não concordante. Realizado manualmente em muitos estudos, pode ser automatizado com a utilização do *Unified Medical Language System* (UMLS). A obtenção da listagem de termos pode ocorrer de forma automática por algoritmos de processamento de linguagem natural, sendo que a utilização de regras para identificação de informação em textos permite que o conhecimento do especialista seja acoplado ao algoritmo, podendo ser realizada por técnicas baseadas em *Machine Learning*. Quando se trata de mapeamento de termos utilizando o modelo de sete Eixos da Classificação Internacional para a Prática de Enfermagem (CIPE®), o processo também pode ser automatizado por meio de algoritmos de processamento de linguagem natural, como o *POS-tagger* e o *parser sintático.*

Conclusão: o método de mapeamento cruzado pode ser intensificado pelo uso de algoritmos de processamento de linguagem natural. No entanto, mesmo em casos de mapeamento automático, a validação dos resultados por especialistas não deve ser descartada.

DESCRITORES: Terminologia. Enfermagem. Informática. Vocabulário controlado. Métodos.

USO DE HERRAMIENTAS COMPUTACIONALES COMO AYUDA AL MÉTODO DE MAPEO CRUZADO ENTRE TERMINOLOGÍAS CLÍNICAS

RESUMEN

Objetivo: reflexionar sobre el uso de herramientas computacionales en el método de mapeo cruzado entre terminologías clínicas.

Método: estudio de reflexión.

Resultados: el método de mapeo cruzado consiste en la obtención de listado de términos, por medio de extracción y normalización; conexión entre los términos del listado y los de la base de referencia, mediante reglas previamente definidas; y agrupación de los términos en categorías: combinación exacta o parcial o, de manera más detallada, término similar, término más amplio, término más restringido y término no concordante. Realizado manualmente en muchos estudios, puede ser automatizado con el uso del *Unified Medical Language System* (UMLS). La obtención del listado de términos puede ocurrir de forma automática por algoritmos de procesamiento de lenguaje natural, siendo que la utilización de reglas para identificación de información en textos permite que el conocimiento del especialista sea acoplado al algoritmo, pudiendo ser realizado por técnicas basadas en *Machine Learning*. Cuando se trata de mapeo de términos utilizando el modelo de siete Ejes de la Clasificación Internacional para la Práctica de Enfermería (CIPE®), el proceso también

puede ser automatizado a través de algoritmos de procesamiento de lenguaje natural, como el *POS-tagger* y el *parser sintático*.

Conclusión: el método de mapeo cruzado puede ser intensificado por el uso de algoritmos de procesamiento de lenguaje natural. Sin embargo, incluso en casos de asignación automática, la validación de los resultados por expertos no debe descartarse.

DESCRIPTORES: Terminología. Enfermería. Informática. Vocabulario controlado. Métodos.

INTRODUCTION

In the development of terminologies in health it is necessary to harmonize concepts to ensure the interoperability of the data and to inform the researchers of the area about possible updates to be performed.[1] The process of elaboration, development and harmonization of terminologies in health comprises a great effort on the part of its developers, who have limits to effect it in an individual way. The International Standards Organization (ISO) and the International Health Terminology Standards Organization (IHTSDO) are worldwide organizations that are dedicated to this process,[2–3] while the former invests efforts towards the standardization for the development of terminologies in the field of health,[2] the second is responsible for the development of the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT), coordinating projects for harmonizing this nomenclature with other terminologies, controlled vocabularies and classifications.[3–4]

The SNOMED CT comprises a global clinical terminology covering several specialties, disciplines and requirements. For this reason, it minimizes the use of different terminologies or clinical systems, which allows a greater sharing and reuse of structured clinical information.[3] In Brazil, the Ministry of Health, through the Ordinance nº 2073, of August 31, 2011, defined its use for the codification of clinical terms and mapping of national and international terminologies in use in the country, aiming to support the semantic interoperability between the systems.[5]

Specifically in nursing field, the use of standardized terminology provides a clear method for documenting its practices, it provides guidance and support for nurses in their clinical reasoning, and names the phenomena of interest in the profession, contributing for the construction of specific knowledge.[6] Thus, the implementation of terminology in care settings presupposes that a comparison between the records of the patient's medical recordand the standardized language should be made in advance, which can be made through the cross-mapping methodology.[7]

Considering the recommendation of the Ministry of Health regarding the employment of SNOMED CT, the cross-mapping between this terminology and those of nursing can broaden the representativeness of nursing phenomena in national databases, with the possibility of comparison with international bases. In addition, the use of this method contributes to the evolution and dissemination of terminologies by the different countries and specialties of nurses,[8] and their results collaborate so that professionals may reflect on the terms they use every day and are not registered in a uniform way.

Cross-mapping is a methodological procedure referenced by the nursing field since the 1990s, having as main objective the determination of similarities and differences between terms[9] and being one of the steps for the construction of subsets of nursing diagnosis, outcomes and interventions of the International Classification for Nursing Practice (ICNP®).[10]

Several national and international surveys that aim to contribute to the implementation of a standardized language through terminologies have been using this methodology.[7,11–15] A research that carried out the cross-mapping of the terms of ICNP® 1.0 and SNOMEDCT identified that 80% of the terms of that are present in this.[11] In this sense, the International Council of Nurses (ICN)

currently provides tables of equivalence between the statements of diagnosis, outcomes and nursing interventions of the ICNP® and of the SNOMED CT.[16–17]

In the studies mentioned, the cross-mapping was performed manually and, gradually, computational tools were incorporated to support its operation, in order to reduce the time and to reduce human inconsistencies. Another justification for the inclusion of computational tools to support the cross-mapping is the use of terminologies in multiple languages, which results in the demand for studies aiming at automatic or semiautomatic translation of a set of data.[18–19] On the other hand, computational tools are important resources for the elaboration and improvement of terminological subsets and creation of glossaries and complete terminological ontologies.[20–21]

Despite the incorporation of computational resources for the cross-mapping between terminologies,[22] between clinical texts and terminologies[23] and between elements present in archetypes for terminologies,[24] the potential of the tools is not yet fully used and there is no consensus on the automation of the method, nor on its effectiveness, which justifies the reflection proposed in this article, which aims to reflect on the use of computational tools in the cross-mapping method between clinical terminologies.

REFLECTION

In the context of standardized languages, the cross-mapping consists of a method that allows the comparison of a standardized language with the language used in daily health services or between different classification systems.[25] The method consists of obtaining the list of terms, through the extraction and normalization; connection between the terms of the list and those of the reference base (structured terminologies), by means of previously defined rules; and grouping the terms into categories. The terms extracted should represent the breadth of the nursing practices in a given care space; therefore, the researches in this domain use different databases and temporalities. Being it a human process, it is prone to failures due to the amount of data being processed.

As an example of the different bases and temporalities, for the mapping between the ICNP® 1.0, the nursing diagnosis contained in children's records and the nomenclature of nursing diagnosis and interventions in the city of Curitiba, PR, it was necessary to manually retrieve 20% of the medical records of the patients treated in six months - it was considered a consultation to each medical record selected, a total of 80.[7]

The complete transcription of the information contained in nursing records, through deep and exhaustive reading, was reported in a study that mapped nursing diagnosis of patients of an Intensive Care Unit (ICU) with the North American Nursing Diagnosis Association International (NANDAI). The database consisted of 256 records of patients who were hospitalized in the ICU in a period of six months.[12] In turn, similar studies used a computational tool, called Poronto[21] for the extraction of terms of nursing evolutions contained in an electronic health record of a university hospital[26] and for the identification of terms in scientific articles related to the practice of nursing aimed at children and adolescents *in situ*ation of domestic violence.[27] In the first study, a database of 115,760 patient evolutions and a temporality of two years were used, and 257,893 terms were extracted from the records.[26] In the second, the database was composed of 40 articles in total, of which 17,365 terms were extracted.[27]

The automatic extraction of terms from the texts is a task of the Natural Language Processing (NLP) algorithms, which involves solving simple and compound terms and that can be based on statistics, linguistics and/or knowledge.[28] Among the tools* available for use, it is possible to mention:

---

* CoGrOO NLP tools - http://cogroo.sourceforge.net/download/current.html; Natural Language Toolkit (NLTK) - http://nltk.org; OpenNLP - https://opennlp.apache.org/; Stanford Core NLP - https://stanfordnlp.github.io/CoreNLP/; GATE - https://gate.ac.uk/

CoGrOO,Natural Language Toolkit (NLTK), OpenNLP, Stanford Core NLP and GATE. Therefore, the inclusion of a computational tool allows the processing of a large number of texts in search of information, in a shorter time when compared to the manual activity. In the automatic process, using the patient records as an empirical basis, it is possible to perceive the use of databases and greater temporalities, when compared to the extraction of terms by the manual process. In addition, the quantification of terms, that is, the frequency with which a term appears in the analysis corpus, is performed automatically by the tool, demonstrating the relevance of a term or concept of nursing, in a given care space.

The quantification of the terms of large databases is a complex activity if performed manually, due to the number of occurrence of the terms; for example, nursing evolutions of a university hospital presented more than 50,000 occurrences of the terms "time" and "abdomen", in a total of 115,760 records.[26] This, in part, explains the use of smaller databases by studies that extract terms through the manual process. However, it should be emphasized that the larger the database and the temporality, the greater the possibility of representing the phenomena of the nursing practice.

The process of content normalization consists of the withdrawal of duplicate terms and their appropriateness regarding the spelling, gender, number and verbal time.[12,26–27] In the automatic form, the methods of normalization and adequacy of texts originate from the stages of pre-processing used in NLP algorithms. The preprocessing and tokenization, in turn, are related to the transformation of the input text into something that the computational algorithm can understand and manipulate, which may include the removal of stopwords (words without relevant function in a given context, not being necessary for the processing of texts) and the capitalization of the text (uppercase to lowercase), the normalization of words by means of linguistic reducers or other forms of standardization and, finally, the separation of sentences and words from the text into individual units, called tokens.[29]

The NLP algorithms can use morphological, syntactic, semantic and pragmatic analysis. Regarding the morphological analysis of the texts, it is possible to mention the POS-Tagger,[29] which defines the morphology of words and their grammatical classes; for example, in the phrase "patient reports reduced pain", the algorithm would define "patient" as a singular masculine noun, "report" is a verb in the present indicative in the third person singular, and so on. Also, common cases of morphological normalization include changes in the grammatical class of words (noun, adjective, and verb), changes in their variations (verbal tense, gender, number and grade) or even the lexical reduction of a set of words that represent a similar meaning in a single term, using stemming or lemmatization (example: the words "organize", "organized" and "organizing" are transformed into "organ").[29–34]

When performed automatically, the morphological normalization can reduce the time spent by the researcher during this step. However, the specialist's knowledge is extremely important, as is the normalization of the terms "right" and "patient rights" - the first refers to one location and the second to a nursing attention focus -; in case of automatic standardization, the two terms would be normalized to "right".[26] This refers to the reflection that the normalization process needs to be performed in a semiautomatic way, that is, the knowledge of the expert is necessary for the semantics of the terms to be preserved.

Regarding the syntactic analysis of the texts, it is possible to mention the syntactic parser,[29] which is divided into two categories: the constituency parser (Figure 1), which demarcates the structure of the sentences of a text, and the dependency parser (Figure 2), which establishes the relations of dependence between the words of a text.

Another aspect of normalization of extracted content that can be performed automatically, in addition to the removal of stop words and capitalization, is the expansion of the abbreviations used by nursing professionals when recording their activities.[29,31]
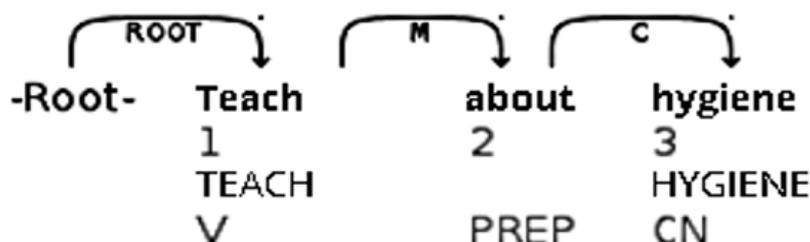
Figure 1 - Demarcation of the structure of a sentence by the constituency parser
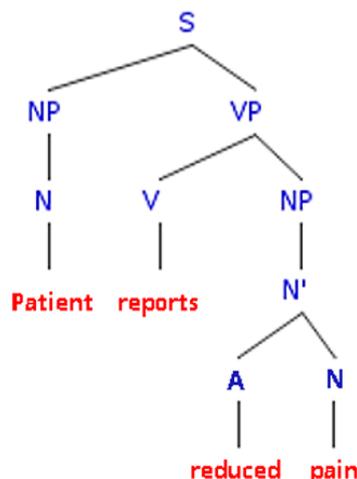
S: sentence; VP: verbal phrases; NP: nominal phrases; N: names - nouns; V: verb; A: adjective; S: sentence; N': nominal constituent sub phrase.

Source: Adapted from constituency parser, 2017.

Figure 2 - Relation of dependence between words by the dependency parser

M: modifier; C: complement; V: verb; PREP: preposition; CN: common name.

Source: Adapted from dependency parser, 2017.

The semantic analysis consists of discovering the meaning of words or concepts in the middle of the text. Among the problems that the semantic analysis aims to solve, it should be highlighted the resolution of ambiguity[35] and the recognition of nominated entities,[36] which includes the identification and classification of entities, such as names of people, organizations and locations, in a text.

The rules for the cross-mapping can be determined according to the study design, based on the characteristics of the data structure of the information system and the terminology to be used.[37] The number of rules, if on the one hand guarantees the accuracy of the mapping, on the other, demands from the researcher an effort and knowledge that goes beyond their specialty: it needs a theoretical and practical basis of the classificatory system and semantic and transcultural equivalences.

Studies that performed the cross-mapping manually between terms and nursing diagnosis contained in patient records and NANDA I [12,38] included rules such as: guarantee the meaning of the terms, check the context and the meaning and not only the words; compare the terms to the statements of diagnosis and the focus of attention; compare terms to defining characteristics and related risk factors; identify and describe the possible concepts nursing diagnosis; and to map the nursing diagnosis in NANDA I domains and classes.

Similarly, studies that have performed the manual cross-mapping between nursing interventions and the Nursing Interventions Classification (NIC)[13] included rules such as: use the verbs of the interventions to perform the mapping for the NIC; map the intervention from the NIC intervention title to the activity; maintain the consistency between the mapped intervention and the definition of intervention in the classification; use the title of the more specific NIC intervention; and map interventions that had two or more verbs to two or more NIC interventions corresponding to them.

The use of rules for the automatic identification of information in the text (rule-based information extraction) is a methodology widely used in computational tools, since it allows the expert's knowledge to be incorporated into the algorithm. This approach has some known limitations[30] and can be enhanced if used in conjunction with statistical-based techniques such as the Machine Learning (ML).

In the case of ML algorithms with supervised learning, the knowledge of the expert is passed to the algorithm by means of annotation of data in the texts, being morphological, syntactic or semantic. This process is very time-consuming and costly, and it requires specialists engaged in task execution, well-defined annotation guidelines, and computational tools that accelerate and support the process.[39]

When it comes to unsupervised learning algorithms, the expert's knowledge is not necessary because the algorithm itself can group data by similarity and extract the necessary information. In addition, the use of statistical methods can extend the scope of the algorithm, not limited to the knowledge of the expert and the generation of rules.

The establishment of categories for the arrangement of terms, the last phase of the mapping, should follow criteria capable of making possible subsequent comparisons or the re-use of results. In general, when the term found corresponds exactly to the term of the classification system, it is categorized as an exact combination and, when it presents similar, synonymous or related concepts, as a partial combination.[13]

In the nursing domain, it is common to use the criteria established by Leal,[40] which indicate more detailed categories for the mapping, among them: similar term, when there is no agreement of the spelling of the term, but the meaning is identical; broader term, when the term identified has a greater meaning than that of its terminology; more restricted term, when the term identified has a more limited meaning than that of its terminology; and non-agreeing term, when there is no agreement between the identified term and that of its terminology.

Regarding the cross-mapping between nursing terminologies, the use of the Unified Medical Language System (UMLS) can anchor the performance of an automatic mapping, since it comprises a knowledge source that integrates hundreds of terminologies or health-related classifications using a unified platform.[1] In addition, there is already an initiative for the translation of UMLS into Brazilian Portuguese.[41]

The UMLS uses several processes to integrate terminologies, such as the use of lexical tools, for the normalization of concepts and preservation of meanings and relations in the source vocabularies.[42] However, the automated process may have limitations. The automatic mapping by the UMLS between the Logical Observation Identifiers Names and Codes (LOINC), terminology for laboratory tests and clinical observations, and SNOMED CT has proved to be unsatisfactory, although the two terminologies cover both the domain of laboratory procedures and use similar knowledge representation formalisms. The study found that to improve the performance of the automatic mapping process, additional techniques are required.[43]

Inaccurate correspondences were also observed in the mapping between nursing terminologies, indicating a series of complexities to be addressed in the UMLS, requiring collaboration between specialists to solve problems in semantic mappings.[1] This fact was approached in a cross-mapping between ICNP® and the Classification of Clinical Care (CCC), and between ICNP® and NANDA I, in which there were 97% exact matches when the mapping was performed by specialists; when

processed by the UMLS, the comparison analysis presented an overall precision of 33.6% in the semantic mapping.[1]

On the other hand, when it comes to mapping terms using the ICNP®7-Axis model, the process can be automated through NLP algorithms such as the POS-tagger and the *syntactic parser*.[29] This is justified by the fact that, in the 7-Axis model, the terms belonging to the focus axis consist, for the most part, of nouns; the terms of the judgment axis correspond to adjectives; and the terms of the action axis refer to verbs in the infinitive, which enables a discipline for the semantics of the terms.

Considering that the analysis of the context of the terms extracted from empirical bases is of extreme importance in terminological work,[44] being often necessary to consider excerpts from nursing records to identify the context of nursing terms[38] the *dependency parser* is a tool that can support the cross-mapping methodology, since the dependencies between the words help to understand the context in which the terms are inserted.

Although the POS-tagger, the *syntactic parser*[29] and other NLP techniques are identified as facilitators for the cross-mapping method, with tools widely available for use, studies that have used them are still not identified, which is a limit for the reflection proposed in this article.

CONCLUSION

The operation of the cross-mapping methodology can be hampered by the amount of data from the empirical bases and by the human limitation in the comparison process. In this sense, computational tools are resources to maximize time and minimize manual inspection errors, but are supportive of the expert.

It is necessary that the researchers who are dedicated to the development of terminologies of the nursing field know computational tools able to support the process of cross-mapping, so that they can evaluate them and use them potentially.

In addition, the steps of obtaining and normalizing terms are the ones that most exploit the potential of computational resources, and the cross-mapping method can be intensified by the use of NLP algorithms. However, even in cases of automatic mapping, the validation of the results by specialists should not be disregarded, especially regarding the cross-cultural equivalence.

REFERENCES

1. Kim TY, Coenen A, Hardiker N, Bartz CC. Representation of nursing terminologies in UMLS. Amia Annu Symp Proc [Internet]. 2011 [cited 2016 Dec 20];2011:709-14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3243214/

2. International Standards Organization (ISO). ISO/TS 17117 de 2002: controlled health terminology: structure and high-level indicators. 2002 [cited 2016 Dec 06]. Available from: https://www.iso.org/obp/ui/#iso:std:iso:ts:17117:ed-1:v1:en

3. International Health Terminology Standards Development Organisation (IHTSDO). Systematized Nomenclature of Medicine Clinical Terms - SNOMED CT [Internet]. Starter Guide. 2014 [cited 2017 Jan 02]. Available from: http://www.ihtsdo.org/snomed-ct/what-is-snomed-ct

4. Campbell JR, Brear H, Scichilone R, White S, Giannangelo K, Carlsen B, *et al.* Semantic interoperation and electronic health records: context sensitive mapping from SNOMED CT to ICD-10. Stud Health Technol Inform [Internet]. 2013 [cited 2016 Dec 21];192:603-7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23920627

5. Ministério da Saúde (BR). Portaria 2073 de 31 de agosto de 2011: regulamenta o uso de padrões de interoperabilidade e informação em saúde para sistemas de informação em saúde no âmbito do Sistema Único de Saúde, nos níveis Municipal, Distrital, Estadual e Federal, e para os sistemas

privados e do setor de saúde suplementar. Brasília (DF): MS; 2011 [cited 2016 Dec 21]. Available from: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt2073_31_08_2011.html

6. Carvalho EC, Cruz DALM, Herdman TH. Contribuição das linguagens padronizadas para a produção do conhecimento, raciocínio clínico e prática clínica da enfermagem. Rev Bras Enferm [Internet]. 2013 [cited 2016 Dec 22];66(esp):134-41. Available from: https://doi.org/10.1590/S0034-71672013000700017

7. Luciano TS, Nóbrega MML, Saparolli ECL, Barros ALBL. Cross mapping of nursing diagnoses in infant health using the International Classification of Nursing Practice. Rev Esc Enferm USP [Internet]. 2014 [cited 2016 Dec 23];48(2):250-6. Available from: https://doi.org/10.1590/S0080-623420140000200008

8. Barra DCC, Dal Sasso GTM. The nursing process according to the international classification for nursing practice: an integrative review. Texto Contexto Enferm [Internet]. 2012 [cited 2017 Mar 13];21(2):440-7. Available from: http://dx.doi.org/10.1590/S0104-07072012000200024

9. Delaney C, Moorhead S. Synthesis of methods, rules and issues of standardizing nursing intervention language mapping. Nurs Diagn. 1997;8(4):152-156. https://doi.org/10.1111/j.1744-618X.1997.tb00471.x

10. Carvalho CMG, Cubas MR, Nóbrega MML. Brazilian method for the development terminological subsets of ICNP®: limits and potentialities. Rev Bras Enferm [Internet]. 2017 [cited 2017 Mar 17]; 70(2):430-35. Available from: http://dx.doi.org/10.1590/0034-7167-2016-0308

11. Park HA, Lundberg C, Coenen A, Konicek D. Evaluation of the Content Coverage of SNOMED CT Representing ICNP Seven-axis Version 1 Concepts. Methods Inf Med [Internet]. 2011 [cited 2016 Dec 20];50(5):472-8. Available from: http://dx.doi.org/10.3414/ME11-01-0004

12. Ferreira AM, Rocha EN, Lopes CT, Bachion MM, Lopes JL, Barros ALBL. Nursing diagnoses in intensive care: cross-mapping and NANDA-I taxonomy. Rev Bras Enferm. [Internet]. 2016 [cited 2016 Dec 18]; 69(2):307-15. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-71672016000200307&lng=en&nrm=iso&tlng=en

13. Silva TG, Santana RF, Souza PA. Nursing interventions for elderly who aged in psychiatric institutions: crossed mapping. Rev Eletr Enf [Internet]. 2016 [cited 2016 Dec 19];18:e1185. Available from: https://www.revistas.ufg.br/fen/article/view/39049/21988

14. Kim TY, Hardiker N, Coenen A. Inter-terminology mapping of nursing problems. J Biomed Inform [Internet]. 2014 [cited 2016 Dec 19];49:213-20. Available from: http://dx.doi.org/10.1016/j.jbi.2014.03.001

15. Park HA, Lundberg C, Coenen A, Konicek D. Mapping ICNP version 1 concepts to SNOMED CT. Stud Health Technol Inform [Internet]. 2010 [cited 2017 Feb 19];160(Pt 2):1109-13. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20841856

16. International Council of Nurses (ICN). ICNP para SNOMED CT: equivalency table for diagnosis & outcome statements. [Internet] Geneva: ICN; 2018. [cited 2019 Jan 18]. Available from: https://www.icn.ch/sites/default/files/inline-files/ICNP-SCT%20Cross-Mapping%202018-DC.pdf

17. International Council of Nurses (ICN). ICNP para SNOMED CT: equivalency table for intervention statements [Internet]. Geneva: ICN; 2018 [cited 2019 Jan 18]. Available from: https://www.icn.ch/sites/default/files/inline-files/ICNP-SCT%20Cross-Mapping%202018-IC.pdf

18. Cimino JJ, Barnett GO. Automated Translation between Medical Terminologies using Semantic Definitions. MD Comput [Internet]. 1990 [cited 2017 Mar 10];7(2):104-9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/2186251

19. Silva J, Chaves T. An ontology-based approach for SNOMED CT translation [Internet]. 2015 [cited 2017 Mar 05];1-5. Available from: http://ceur-ws.org/Vol-1515/regular4.pdf

20. Matney SA, Warren JJ, Evans JL, Kim TY, Coenen A, Auld VA. Development of the nursing problem list subset of SNOMED CT. J Biomed Inform [Internet]. 2012 [cited 2017 Mar 05];45(4):683-8. Available from: http://dx.doi.org/10.1016/j.jbi.2011.12.003

21. Zahra FM, Carvalho DR, Malucelli A. Poronto: ferramenta para construção semiautomática de ontologias em português. J Heal Informatics [Internet]. 2013 [cited 2017 Mar 05];5(2):52-9. Available from: http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/232/167

22. Lamy JB, Tsopra R, Venot A, Duclos C. A semi-automatic semantic method for mapping SNOMED CT concepts to VCM icons. Stud Health Technol Inform [Internet]. 2013 [cited 2017 Mar 12]; 92:42-6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3861883/

23. Stenzhorn H, Pacheco EJ, Nohama P, Schulz S. Automatic mapping of clinical documentation to SNOMED CT. Stud Health Technol Inform [Internet]. 2009 [cited 2017 Mar 12];150:228-32. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19745302

24. Yu S, Berry D, Bisbal J. Clinical coverage of an archetype repository over SNOMED-CT. J Biomed Inf [Internet]. 2012 [cited 2017 Mar 24];45(3):408-18. Available from: http://www.sciencedirect.com/science/article/pii/S1532046411002152

25. Nonino FOL, Napoleão AA, Carvalho EC, Petrilli Filho JF. A utilização do mapeamento cruzado na pesquisa de enfermagem: uma revisão da literatura. Rev Bras Enferm [Internet]. 2008 [cited 2018 Jan 25];61(6):872-7. Available from: https://doi.org/10.1590/S0034-71672008000600013

26. Gomes DC, Cubas MR, Pleis LE, Shmeil MAH, Peluci APVD. Terms used by nurses in the documentation of patient progress. Rev Gaúcha Enferm [Internet]. 2016 [cited 2017 Jan 08]; 37(1):e53927. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1983-14472016000100412&lng=en&nrm=iso&tlng=en

27. Albuquerque LM, Carvalho CMG, Apostólico MR, Sakata KN, Cubas MR, Egry EY. Nursing Terminology defines domestic violence against children and adolescents. Rev Bras Enferm [Internet]. 2015 [cited2017 Jan 08]; 68(3):393-400. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-71672015000300452&lng=en&nrm=iso&tlng=en

28. Conrado MS, Felippo A, Pardo TAS, Rezende SO. A survey of automatic term extraction for Brazilian Portuguese. J Brazilian Comput Soc [Internet]. 2014 [cited 2017 Jan 08]; 20(12):1-28. Available from: http://conteudo.icmc.usp.br/pessoas/taspardo/JBCS2014-ConradoEtAl.pdf

29. Jurafsky D, Martin JH. Speech and language processing. 2nd ed. New Jersey, US: Pearson Prentice Hall; 2009.

30. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc [Internet]. 2011 [cited 2017 Feb 05];18(5):544-51. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3168328&tool=pmcentrez&rendertype=abstract

31. Indurkhya N, Damerau FJ. Handbook of Natural Language Processing. [Internet] USA: CRC Press Book; 2010, Available from https://karczmarczuk.users.greyc.fr/TEACH/TAL/Doc/Handbook%20Of%20Natural%20Language%20Processing,%20Second%20Edition%20Chapman%20&%20Hall%20Crc%20Machine%20Learning%20&%20Pattern%20Recognition%202010.pdf

32. Singh J, Gupta V. A systematic review of text stemming techniques. Artificial Intelligence Review. Springer Netherlands [Internet]. 2017 [cited 2018 Jan 25];48(2):157-217. Available from: https://doi.org/10.1007/s10462-016-9498-2

33. Liu H, Christiansen T, Baumgartner WA, Verspoor K. BioLemmatizer: A lemmatization tool for morphological processing of biomedical text. J Biomed Semantics [Internet]. 2012 [cited 2018 Jan 25];3(1):1-29. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3359276/

34. Soares MVB, Prati RC, Monard MC. Improvements on the Porter's Stemming Algorithm for Portuguese. IEEE Lat Am Trans [Internet]. 2009 [cited 2018 Jan 25];7(4):472-7. Available from: https://doi.org/10.1109/TLA.2009.5349047

35. Ranjan Pal A, Saha D. Word Sense Disambiguation: a Survey. Int J Control Theory Comput Model [Internet]. 2015 [cited 2018 Jan 25];5(3):1-16. Available from: https://doi.org/10.5121/ijctcm.2015.5301

36. Goulart RRV, Strube de Lima VL, Xavier CC. A systematic review of named entity recognition in biomedical texts. J Brazilian Comput Soc [Internet]. 2011 [cited 2018 Jan 25];17(2):103-16. Available from: https://link.springer.com/article/10.1007/s13173-011-0031-9

37. Moorhead S, Delaney C. Mapping nursing intervention data into the nursing interventions classification (NIC): process and rules. Nurs Diagn [Internet]. 1997 [cited 2016 Dec 19];8(4):137-44. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9624991

38. Tosin MHS, Campos DM, Blanco L, Santana RF, Oliveira BGRB. Mapping Nursing language terms of Parkinson's disease.Rev Esc Enferm USP [Internet]. 2015 [cited 2017 Mar 03];49(3):409-16. Available from: https://doi.org/10.1590/S0080-623420150000300008

39. Oliveira LES, Gebeluca CP, Silva AMP, Moro CMC, Hasan SA, Farri O. A statistics and UMLS-based tool for assisted semantic annotation of Brazilian clinical documents. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017 [cited 2018 Jan 25]; Kansas City, USA, p. 1072-8. Available from: https://doi.org/10.1109/BIBM.2017.8217805

40. Leal MT. A CIPE® e a visibilidade da enfermagem: mitos e realidade. Lisboa (PT): Lusociência; 2006.

41. Oliveira LES, Hasan SA, Farri O, Barra CMCM. Translation of UMLS ontologies from European Portuguese to Brazilian Portuguese. In: Anais do XV Congresso Brasileiro de Informática em Saúde, 2016 Nov 27-30, Goiânia, Brazil. p. 373-9.

42. National Library of Medicine (NLM). UMLS® Reference Manual. Bethesda, MD: NLM; 2009.

43. Bodenreider O. Issues in mapping LOINC laboratory tests to SNOMED CT. AMIA Annu Symp Proc. [Internet] 2008 [cited 2016 Dec 19];2008:51-5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2655945/

44. Pavel S, Nolet D. Manual de terminologia [Internet] 2002 [cited 2017 Mar 02]. Available from: http://publications.gc.ca/collections/collection_2011/tpsgc-pwgsc/S53-28-2001-spa.pdf

NOTES

CONTRIBUTION OF AUTHORITY
Study design: Gomes DC, Oliveira LES, Cubas MR.
Writing and / or critical review of content: Gomes DC, Oliveira LES, Cubas MR.
Review and final approval of the final version: Cubas MR, Barra CMCM.

ETHICS COMMITTEE IN RESEARCH
Not applicable.

CONFLICT OF INTEREST
No any conflict of interest.

CONSENT OF USE OF IMAGE
Not applicable.

CORRESPONDENCE AUTHOR
Denilsen Carvalho Gomes
deni.gomesc@gmail.com