MANHATTAN COLLEGE

Computational Technique To Assure Quality of Medical Terminologies

by

Kalden Yugyel Dorji

Computer Science

A thesis submitted in partial fulfillment of the requirements for the

Science Honors Program, Manhattan College.

This manuscript by Kalden Yugyel Dorji has been read and accepted by Dr. Ankur Agrawal in satisfaction of the thesis requirements for the Science Honors Program at Manhattan College.

Faculty Thesis Advisor:

May 2, 2024 — Ankur

Date — Signature

Acknowledgements

I would like to take this opportunity to extend my appreciation to Dr. Ankur Agrawal for his guidance and mentorship. His expertise and encouragement were instrumental in shaping this work and navigating its challenges.

I am also profoundly grateful to my family for their unwavering support and encouragement throughout this challenging journey, amidst my numerous other commitments. Their constant cheerleading bolstered my spirits and kept me motivated every step of the way.

Abstract

This study introduces a computational method utilizing Word2Vec to analyze medical terms and identify potential information gaps within standardized terminologies, specifically SNOMED CT. Extracted from PUBMED articles, the primary objective is to establish firm word associations and calculate lexical similarity among medical concepts to discover semantic relationships. By identifying high-scoring pairs, the analysis ensures consistency in groupings and attribute relationships. Detected discrepancies in these relationships illuminate significant gaps within SNOMED CT, which is widely employed in Electronic Health Record (EHR) systems. The findings of this research substantially contribute to refining medical terminologies, thereby facilitating enhanced communication among healthcare professionals and elevating the quality of patient care.

Table of Contents

1. Introduction

The field of healthcare relies heavily on standardized medical terminologies, such as SNOMED CT (Systematized Nomenclature of Medicine–Clinical Terminology), to ensure accurate communication and information exchange among healthcare professionals [1]. However, the effectiveness of these terminologies depends on their comprehensiveness and semantic coherence. In this context, computational methods offer a novel approach to analyzing medical terms and uncovering potential gaps within existing terminologies. This study introduces a methodology leveraging Word2Vec [2] to delve into medical concepts present in SNOMED CT, through the help of extracted PUBMED articles, with a primary aim to discern latent semantic relationships and identify discrepancies. By closely examining word associations and computing lexical similarity, this research sheds light on the gaps in SNOMED CT, which are fundamental components within Electronic Health Record (EHR) systems. The implications of these findings are significant, as they hold the potential to refine medical terminologies, foster clearer communication among healthcare professionals, and ultimately enhance the quality of patient care.

a. Research Purpose

The primary aim of this research is to develop a computational method harnessing Natural Language Processing (NLP) techniques to construct an algorithm adept at comparing the similarities between two medical concepts. By developing and utilizing this algorithm, the research aims to detect inconsistencies within medical terminologies, particularly focusing on identifying missing information. Subsequently, the project seeks to enhance the quality of SNOMED CT (Version: 2023-07-31), by illuminating and addressing the identified gaps. Through this endeavor, the research aims to raise awareness regarding missing information within medical terminologies and propose potential solutions to rectify or augment them, ultimately contributing to the refinement and improvement of medical communication and healthcare outcomes.

b. Background Knowledge

EHR systems play a crucial role in modern healthcare by digitizing and consolidating patient records, facilitating enhanced communication among healthcare providers, empowering patients, and improving the quality and efficiency of healthcare delivery.

Standardized medical terminologies, such as SNOMED CT, are essential for organizing and categorizing medical concepts within EHR systems. These terminologies enable interoperability between healthcare facilities and computer systems, facilitating seamless data exchange and ensuring consistency in medical documentation.

SNOMED CT is a comprehensive repository of scientifically validated medical terms, offering a robust framework of codes, terms, synonyms, and relationships crucial for clinical documentation and reporting. It boasts an extensive vocabulary comprising over 360,000 concepts and approximately 1.37 million semantic relationships, serving as a cornerstone for encoding clinical information.

In addition to its vast repository of medically validated terms and extensive vocabulary, SNOMED CT includes a diverse range of relationships between concepts, useful for clinical documentation and reporting. Among these relationships are hierarchical relationships, which establish parent-child connections between concepts. This hierarchical structure allows for the organization of terms into broader categories and more specific subcategories, facilitating efficient data retrieval and analysis. Moreover, SNOMED CT incorporates semantic relationships, including groups and attributes, which provide further contextual information. Groups encapsulate single or multiple attribute relationships, while attribute relationships offer insights into specific characteristics or properties associated with a concept. For instance, attribute relationships can denote the location of a finding, such as the "heart" or "myocardium," enhancing the precision and granularity of clinical information encoded with SNOMED CT. These multifaceted relationships contribute to the comprehensive nature of SNOMED CT, empowering healthcare professionals with a nuanced framework for encoding and interpreting clinical data.

c. Problem Statement

Despite its comprehensiveness, SNOMED CT faces significant challenges, notably the presence of missing connections between concepts. These gaps in the terminology can have significant implications for the effectiveness of Electronic Health Record (EHR) systems and ultimately compromise the quality of patient care.

Missing connections between concepts in SNOMED CT result in incomplete patient records within EHR systems, hindering healthcare providers' access to crucial medical information and potentially leading to diagnostic errors or incomplete treatment plans. Moreover, interoperability issues arise when incomplete or inconsistent data exchange occurs between different healthcare systems, fragmenting patient care and impeding collaboration among providers.

Accurate diagnosis and treatment depend on comprehensive and well-organized medical information, but information gaps in SNOMED CT can hinder healthcare providers' ability to document and retrieve diagnostic information accurately. Additionally, clinical decision support tools integrated into EHR systems may be compromised by incomplete data, potentially leading to suboptimal treatment outcomes or patient safety concerns [3].

Furthermore, incomplete or inconsistent data resulting from missing connections in SNOMED CT can affect the quality of healthcare reporting and research, hindering efforts to improve healthcare quality and advance medical knowledge. Addressing these information gaps in SNOMED CT is essential to ensure the effective utilization of EHR systems and enhance the overall quality of patient care.

d. Significance of the Study

This research holds significance in the healthcare sector as it is an endeavor to develop a computational approach utilizing Natural Language Processing (NLP) techniques to examine, identify, and bridge gaps within SNOMED CT. By developing an algorithm adept at comparing the similarities between medical concepts and detecting inconsistencies, this study aims to illuminate missing information and establish vital connections between concepts within SNOMED CT. Through this innovative methodology, the research not only enhances the quality and completeness of SNOMED CT but also fosters its global adoption in electronic health record (EHR) systems. Ultimately, by addressing these terminological gaps and proposing solutions to rectify them, this research strives to advance medical communication standards and push towards the adoption of SNOMED CT as terminology in more countries, thus contributing significantly to improved healthcare outcomes and patient care on a global scale.

e. Research Objective and Hypothesis

The proposed research hypothesis is that concepts sharing lexical similarity will demonstrate parallel or closely similar modeling structures. To investigate this hypothesis, the study outlines several objectives. First and foremost, it aims to develop a robust computational technique leveraging Natural Language Processing (NLP) tools. This technique will utilize Word2Vec to establish semantic associations between medical concepts based on their contexts. The subsequent step involves training the model using a corpus of published PUBMED articles, ensuring a wide variety of contexts is utilized to build a versatile set of word associations. Once trained, the model will compare randomized concepts possessing identical tags, evaluating their similarity using a normalized score ranging from 0 to 1. Concept pairs exhibiting high similarity scores, surpassing the threshold of 90%, will be subjected to further analysis. Through these objectives, the research endeavors to not only validate the hypothesis but also to develop a method for detecting semantic similarities between medical concepts, thereby contributing to the advancement of computational techniques in medical informatics.

2. Literature Review

In the rapidly evolving landscape of healthcare informatics, the utilization of standardized terminologies plays a pivotal role in facilitating efficient communication, enhancing data interoperability, and ultimately improving patient care. Central to this endeavor is the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), a comprehensive terminology that encompasses a vast array of medically validated concepts, relationships, and hierarchies. As healthcare systems increasingly transition towards digitalization, the importance of SNOMED CT in electronic health record (EHR) systems and clinical decision support tools cannot be overstated.

The purpose of this literature review is twofold: firstly, to explore previous studies that have identified issues and challenges within SNOMED CT hierarchies, shedding light on areas requiring improvement; and secondly, to examine the manifold ways in which SNOMED CT can be leveraged to advance medical practice and healthcare outcomes. By delving into both the shortcomings and strengths of SNOMED CT, this review aims to provide a comprehensive understanding of its role in contemporary healthcare informatics and inform the direction of future research in the field.

a. Overview of Literature

SNOMED CT offers a robust framework for standardized clinical terminology. As healthcare systems continue to embrace digitalization, the role of SNOMED CT in facilitating efficient communication and enhancing data interoperability becomes increasingly paramount. However, the practical application of SNOMED CT is not without its challenges.

Previous studies have delved into the intricacies of SNOMED CT hierarchies, uncovering various issues that impact its usability in clinical settings. For instance, research by Rector, Brandt, and Schneider (2011) identified significant challenges related to incomplete modeling, inconsistencies, and overgeneralization within SNOMED CT hierarchies. These findings underscore the importance of understanding the nuances of SNOMED CT's structure and addressing key issues to optimize its utility in healthcare informatics.

Despite these challenges, SNOMED CT offers immense potential to revolutionize medical practice and healthcare outcomes. Its adoption in electronic health record (EHR) systems and clinical decision support tools has paved the way for standardized clinical documentation and improved patient care. Moreover, SNOMED CT serves as a vital tool for facilitating data exchange and interoperability across diverse healthcare settings, contributing to enhanced healthcare quality and patient safety.

Moving forward, it is essential to bridge the gap between the theoretical framework of SNOMED CT and its practical implementation in clinical practice. Future research endeavors should focus on exploring innovative solutions to address the identified challenges and maximize the benefits of SNOMED CT in healthcare delivery. By delving into both the strengths and shortcomings of SNOMED CT, this literature review aims to provide valuable insights into its role in contemporary healthcare informatics and guide the direction of future research in the field.

b. Key Theories and Concepts

In healthcare informatics, standardized terminologies like SNOMED CT serve as foundational pillars supporting the efficacy and functionality of electronic health record (EHR) systems. At the heart of this significance lies the principle of data consistency, where the utilization of standardized terminology ensures uniformity and coherence in the representation of clinical information across various healthcare settings, systems, and applications. This consistency not only enhances the comprehensibility and reliability of clinical data but also promotes interoperability, enabling the seamless exchange and sharing of information between different EHR systems and healthcare providers. Moreover, standardized terminology facilitates semantic interoperability by imbuing clinical data with precise and unambiguous meaning, thereby facilitating accurate interpretation and utilization by healthcare professionals.

Central to the adoption of standardized terminology is the realization of various benefits that come from its implementation. One such benefit is the rise in quality improvement initiatives within healthcare systems. By providing a common language for encoding clinical information, SNOMED CT enables healthcare organizations to conduct comprehensive data analyses, identify areas for improvement, and implement evidence-based interventions to enhance patient care and outcomes. Additionally, the utilization of standardized terminology lays the groundwork for advancements in clinical decision support systems, enabling the development of intelligent tools that leverage encoded clinical data to provide real-time guidance and recommendations to healthcare professionals at the point of care.

c. Discussion of Previous Studies

The study by Rector, Brandt, and Schneider (2011) [4] categorized seven major types of issues within SNOMED CT hierarchies, from simple errors to inconsistent modeling of complications. Analyzing the SNOMED Core Problem List Subset and the 31 January 2010 IHTSDO distribution, they highlighted complexities in using SNOMED CT for clinical documentation. Their findings stress the need to address these issues, as inconsistencies can have wide-ranging consequences in healthcare informatics. By employing specific SNOMED CT versions and ensuring consistency in classification, the authors offer insights into challenges faced by healthcare professionals. This study enhances understanding of SNOMED CT complexities and emphasizes ongoing efforts to improve its usability and accuracy. Moreover, the adoption and implementation of SNOMED CT in healthcare settings face several challenges, as identified through international surveys and feedback from SNOMED CT education and training series attendees in Korea.

These barriers include low awareness of SNOMED CT's benefits among users and healthcare institutions, absence of governance and infrastructure for SNOMED CT management, insufficient training programs to address user expertise, lack of support services for SNOMED CT use, and vendor capability issues in implementing SNOMED CT in EHR systems. To overcome these barriers, health terminology experts suggest an incremental step-wise plan, including clear government direction on health terminology use, increased awareness through use cases, implementation of SNOMED CT governance, development of supporting tools and services, and professional training and education programs. These facilitators aim to improve SNOMED CT usability and adoption in clinical practice and research [5].

3. Methodology

This study employs a quantitative approach to investigate the completeness of information within the SNOMED CT terminology database. By utilizing computational methods and manual auditing techniques, the research aims to identify and address potential gaps in the representation of medical concepts [6].

a. Research Design

A correlational design is adopted to explore the relationships between pairs of medical concepts within SNOMED CT. This design allows for the examination of similarities and differences in the relationship structures of concept pairs with high similarity scores, as determined by the computational algorithm developed for this study. The presence or absence of corresponding relationships between concept pairs after manual auditing serves as an indicator of potential missing information within the database.

b. Data Collection Methods

Data collection for this study involves multiple phases to ensure comprehensive coverage and accuracy of information. Initially, relevant medical concepts and their associated information are extracted from PubMed articles using Python scripts. These articles are key components necessary for the creation of the model for word associations. As such, they serve as the foundation for the subsequent analysis.

Next, the SNOMED CT terminology database is utilized by obtaining relevant concept, relationship, and description files, which are essential components of the terminology. Python scripts are employed to convert these files into SQLite3 databases, facilitating efficient referencing and manipulation of data during the analysis phase. This step streamlines the process of accessing and querying SNOMED CT data, enabling seamless integration with the computational algorithm developed for this study.

Once the data collection process is complete, the study proceeds to algorithmic computation and manual auditing phases, where concept pairs are used to generate similarity scores and are then subjected to examination of relationship structures within SNOMED CT. This dual approach, combining computational efficiency with manual precision, ensures a thorough assessment of information completeness and accuracy within the SNOMED CT terminology database.

c. Data Analysis Techniques

In the initial phase of the analysis, a function is executed to retrieve a specified number of concepts based on their tag or category within SNOMED CT. Subsequently, to minimize bias, another function is employed to randomize the retrieved concepts. Following this preparatory step, the concept_similarity function is utilized to assess the similarity between pairs of concepts. This function employs a Word2Vec-based computational method, wherein each word in the PubMed articles is assigned corresponding vectors for comparison. When comparing two phrases, individual similarity scores are computed for each word pair, and these scores are then averaged to derive an overall similarity score for the concept pair. High similarity scores indicate a strong association between concepts, prompting a detailed investigation into their relationship structures within SNOMED CT. Any discrepancies or inconsistencies in these relationships are analyzed to identify missing information or links, thereby offering valuable insights into the completeness and accuracy of SNOMED CT's information.

Furthermore, to facilitate manual auditing and further analysis, the concept similarity scores are saved into a CSV file along with the concept information. This file provides a comprehensive overview of the assessed similarities between concept pairs, enabling thorough examination and verification of the computational results.

4. Results

In total, 45 concept pairs were analyzed to assess the completeness and accuracy of information within SNOMED CT. These pairs were drawn from three distinct categories, totaling 15 pairs per category. Among the concept pairs examined, approximately 11 pairs exhibited similarity scores of 90% or greater, indicating a close relationship between the concepts. However, upon closer examination of the relationship structures within SNOMED CT, it was found that three pairs displayed notable discrepancies in terms of attribute relationships.

These findings are consistent with previous studies that have identified missing information gaps and incomplete connections within the SNOMED CT terminology. By highlighting these discrepancies, our study contributes to the growing body of literature on the usability and reliability of SNOMED CT in healthcare informatics.

In table 1, the similarity scores of a selection of randomized concept pairs are depicted, along with the results obtained after manual auditing. Notably, among the pairs chosen, approximately 24% of them displayed similarity scores of 90% or higher, indicating a high degree of semantic similarity between the concepts. Despite the high similarity scores, upon manual auditing, it was discovered that only 3 concepts within the pairs were missing essential information or attribute relationships within SNOMED CT. This discrepancy underscores the importance of manual validation in complementing computational analyses and highlights potential areas for improvement in the completeness and accuracy of SNOMED CT's information.

a. Presentation of Auditing Results

Pair	Similarity Score	Missing Information
C1: Multiple Joint Synovial Fluid C2: Shoulder Joint Synovial Fluid Tag: Specimen	94.8%	C1: Specimen source Topography Attribute C2: C1 children relationship
C1: Gastric Aspirate Specimen C2: Gastric Lavage Aspirate Specimen Tag: Specimen	93.2%	None
C1: Chlamydophila Psittaci Immunoglobulin G Level C2: Chlamydophila Psittaci Immunoglobulin M Level Tag: Procedure	91.0%	None
C1: Propafenone Adverse Reaction C2: Pentazocine Adverse Reaction Tag: Disorder	91.1%	None

Table 1: Similarity scores of a few randomized pairs and results after manual auditing

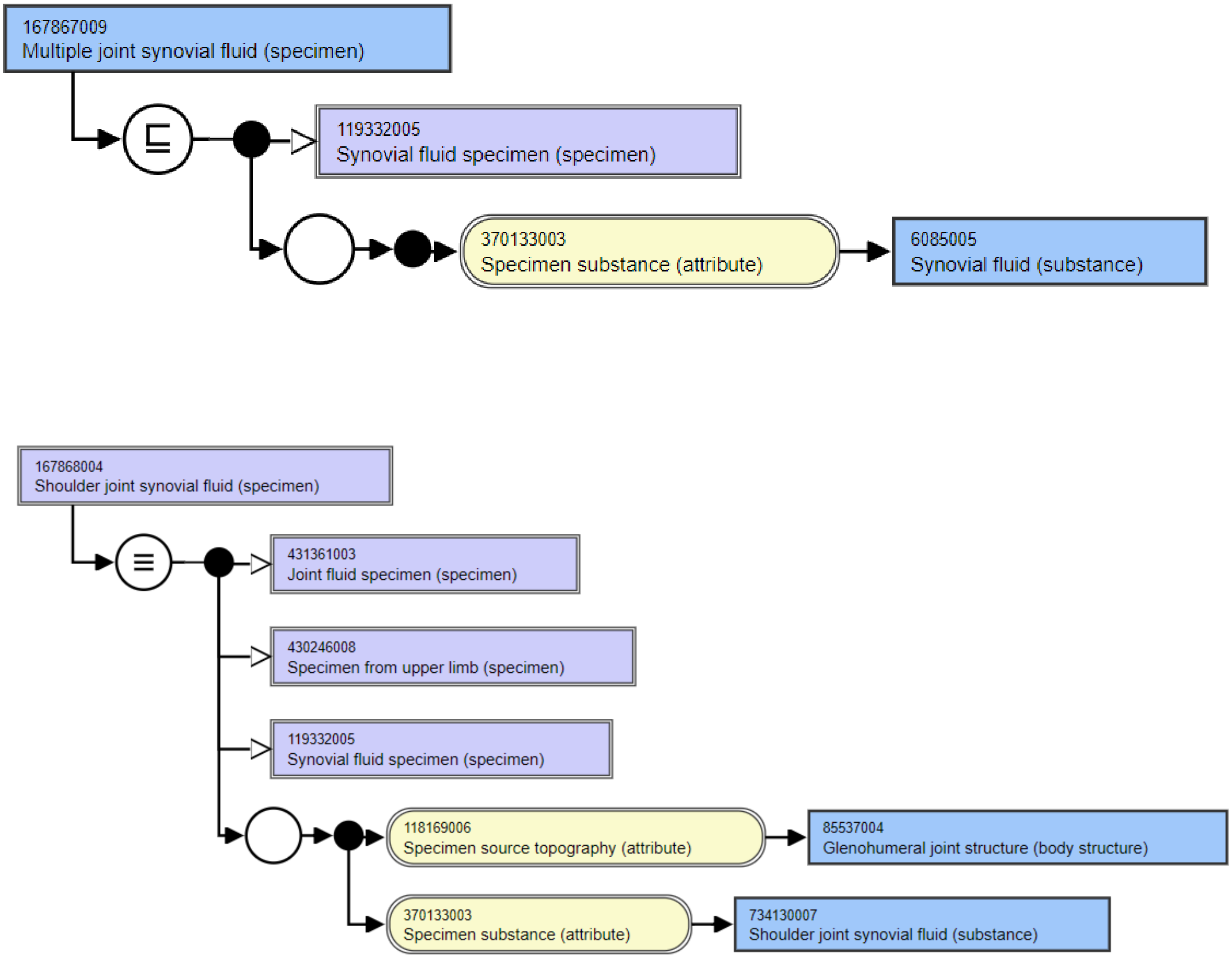
167867009
Multiple joint synovial fluid (specimen)
⊑
119332005
Synovial fluid specimen (specimen)
370133003
Specimen substance (attribute)
6085005
Synovial fluid (substance)
167868004
Shoulder joint synovial fluid (specimen)
≡
431361003
Joint fluid specimen (specimen)
430246008
Specimen from upper limb (specimen)
119332005
Synovial fluid specimen (specimen)
118169006
Specimen source topography (attribute)
85537004
Glenohumeral joint structure (body structure)
370133003
Specimen substance (attribute)
734130007
Shoulder joint synovial fluid (substance)

Figure 1: Structure of two concepts with high similarity scores with information gaps

The striking similarity in structural composition between concept pairs with high similarity scores is notable. Despite this, Figures 1, 2, and 3 reveal a discrepancy in which essential attribute relations are absent. This discrepancy underscores the importance of meticulous examination beyond similarity scores to ensure the integrity and completeness of information within SNOMED CT. Conversely, Figure 4 depicts concept pairs with equally high similarity scores and parallel structuring, showcasing the variability in completeness and accuracy within SNOMED CT.

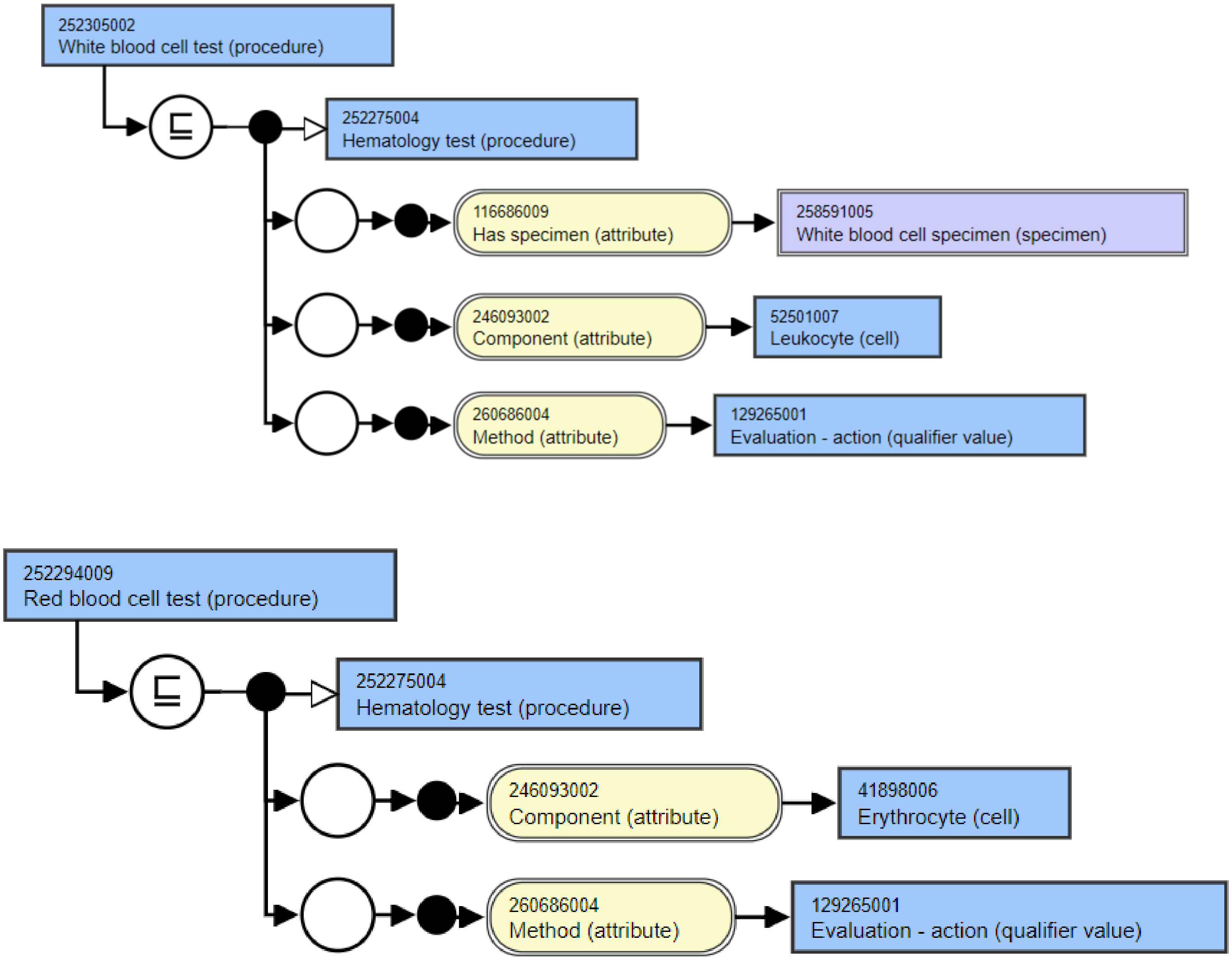

Figure 2: Structure of two concepts with high similarity scores with information gaps

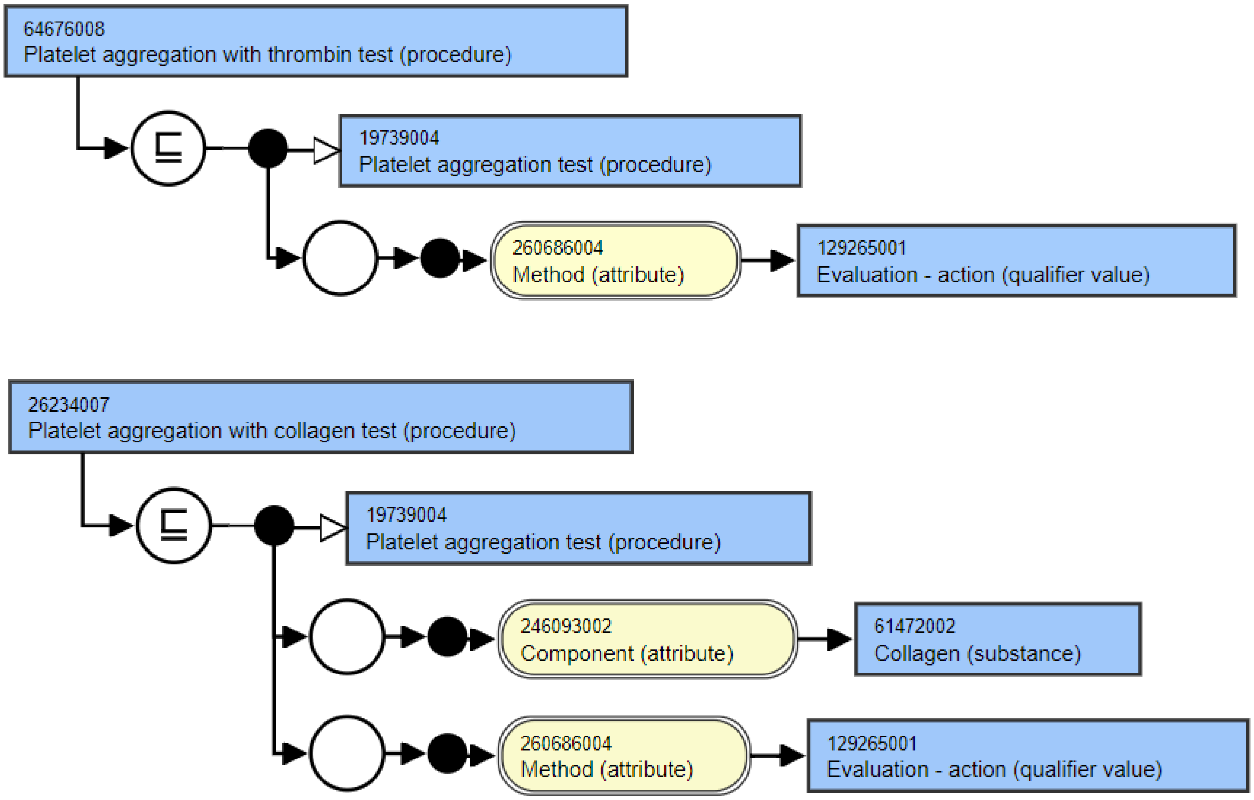
64676008
Platelet aggregation with thrombin test (procedure)
⊑
19739004
Platelet aggregation test (procedure)
260686004
Method (attribute)
129265001
Evaluation - action (qualifier value)
26234007
Platelet aggregation with collagen test (procedure)
⊑
19739004
Platelet aggregation test (procedure)
246093002
Component (attribute)
61472002
Collagen (substance)
260686004
Method (attribute)
129265001
Evaluation - action (qualifier value)

Figure 3: Structure of two concepts with high similarity scores with information gaps

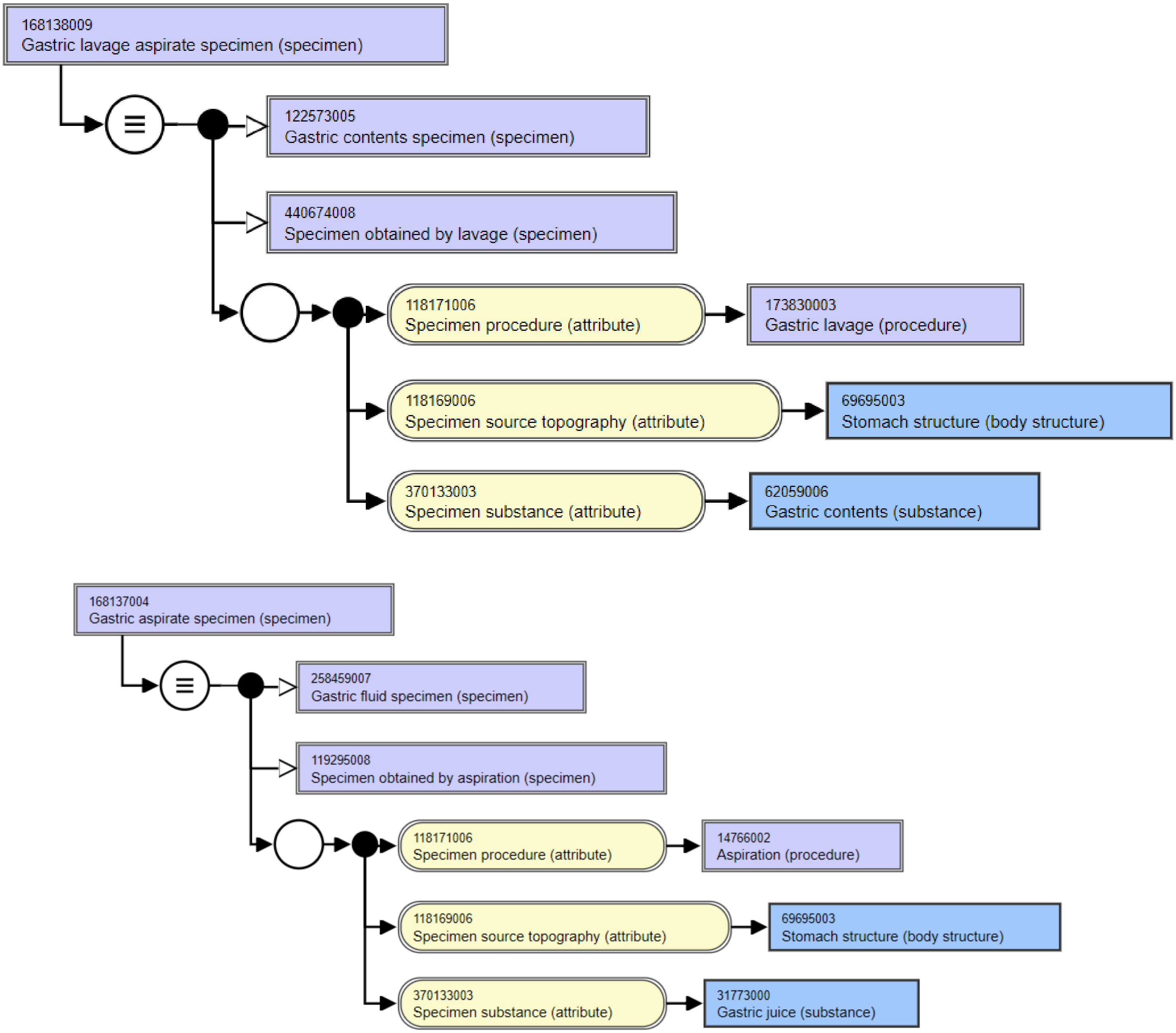
168138009
Gastric lavage aspirate specimen (specimen)
≡
122573005
Gastric contents specimen (specimen)
440674008
Specimen obtained by lavage (specimen)
118171006
Specimen procedure (attribute)
173830003
Gastric lavage (procedure)
118169006
Specimen source topography (attribute)
69695003
Stomach structure (body structure)
370133003
Specimen substance (attribute)
62059006
Gastric contents (substance)
168137004
Gastric aspirate specimen (specimen)
≡
258459007
Gastric fluid specimen (specimen)
119295008
Specimen obtained by aspiration (specimen)
118171006
Specimen procedure (attribute)
14766002
Aspiration (procedure)
118169006
Specimen source topography (attribute)
69695003
Stomach structure (body structure)
370133003
Specimen substance (attribute)
31773000
Gastric juice (substance)

Figure 4: Structure of two concepts with high similarity scores with parallel structure

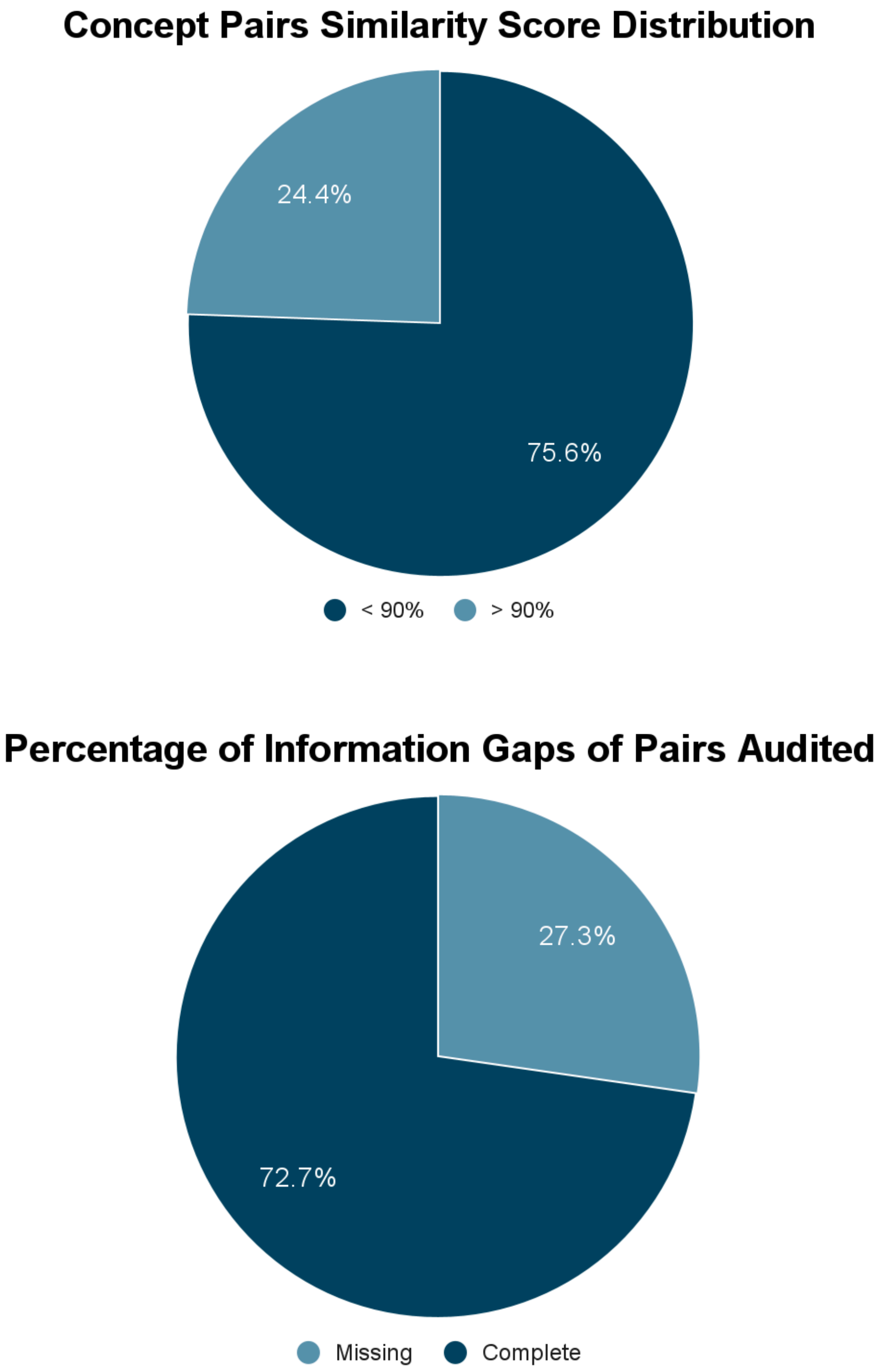

Figure 5: Distribution of Concept Pairs: Similarity Scores and Missing Information

5. Discussion

The analysis of similarity scores among 45 concept pairs revealed insights into the structure and completeness of information within SNOMED CT. Approximately 24% of pairs exhibited high similarity scores of 90% or greater, indicating strong semantic relationships. However, only 27% of these pairs were found to have missing attribute relationships upon manual auditing. This suggests that while most concept pairs demonstrate a high degree of structural similarity within SNOMED CT, issues related to missing information are relatively infrequent. However, it is essential to acknowledge that the findings of this study are based on a small subset of concept pairs within the SNOMED CT database.

a. Implications of Results

The relatively low incidence of missing attribute relationships within concept pairs exhibiting high similarity scores suggests that SNOMED CT's hierarchical structure effectively captures semantic relationships. This implies that SNOMED CT serves as a robust resource for accurately representing complex healthcare terminology, facilitating precise information retrieval and interpretation. However, given that these observations are based on a limited subset of only 45 pairs, caution must be exercised in generalizing these findings. Nevertheless, the identification of discrepancies in a subset of concept pairs underscores the ongoing need for vigilance in ensuring the integrity and comprehensiveness of SNOMED CT [7]. Addressing these discrepancies remains crucial to further enhancing SNOMED CT's usability and reliability in both clinical practice and research contexts. Overall, these implications emphasize the importance of continued efforts to maintain and refine SNOMED CT's content to support accurate and comprehensive data management and analysis in healthcare informatics.

b. Limitations

To optimize the accuracy of similarity scores, several enhancements can be implemented in future work. One approach involves expanding the methodology to include synonyms alongside main fully specified names during the calculation process. This would ensure a more comprehensive assessment of concept similarity, reducing the likelihood of overlooking relevant relationships within SNOMED CT. Additionally, diversifying the usage of models such as Doc2Vec and FastText offers a nuanced approach to semantic representation, providing complementary perspectives on concept relationships and similarity scores. By integrating multiple models, a more robust understanding of similarity scores can be obtained. Furthermore, extending the duration of the model training on the PUBMED corpus with a more powerful computer can mitigate memory limitations and improve overall performance and accuracy. This prolonged training period allows word associations to capture subtle semantic relationships, enhancing the overall efficacy of the model in producing reliable similarity scores.

However, it's important to note that in the current methodology, a threshold of 90% similarity was chosen, meaning that only concept pairs with a similarity score of 90% and greater were considered. While this threshold ensures a certain level of confidence in the identified relationships, it also makes the process somewhat selective, as all other concepts with similarity scores below 90% are ignored. In future research, exploring different threshold values and their impact on the analysis could provide valuable insights into the completeness and accuracy of the results.

c. Future Research

For future research, the focus will likely shift towards utilizing the latest version of the international SNOMED CT to track quality issues. Expanding the study scope to encompass a larger subset of concept pairs will provide a stronger understanding of SNOMED CT's structure and completeness. Additionally, future work may entail developing a more comprehensive algorithm capable of automatically detecting discrepancies based on the number of semantic relationships present, and incorporating the solutions to the limitations mentioned earlier, thereby reducing reliance on manual auditing and improving overall error detection efficiency and ability. Collaboration with healthcare professionals will be crucial to validate the relationships identified as missing by the algorithm, ensuring their relevance and accuracy in clinical practice. These endeavors aim to advance the understanding of SNOMED CT's integrity and usability, ultimately enhancing its effectiveness in healthcare informatics.

6. Conclusion

This study leverages computational methods to analyze the completeness and accuracy of information within SNOMED CT, a critical resource in healthcare informatics, mainly EHR systems. By examining similarity scores among concept pairs and conducting manual auditing, the research identifies discrepancies in attribute relationships, highlighting potential information gaps within SNOMED CT. While most concept pairs exhibit high similarity scores, indicating a strong semantic relationship, issues related to missing information are relatively infrequent. However, caution must be exercised in generalizing these findings, as they are based on a small subset of concept pairs.

The implications of these results underscore the importance of ongoing efforts to maintain and refine SNOMED CT's content. Addressing discrepancies in the terminology is crucial to enhancing its usability and reliability in clinical practice and research. Future research endeavors may involve utilizing the latest version of SNOMED CT, expanding the study scope, and developing comprehensive algorithms to automatically detect discrepancies. Collaboration with healthcare professionals will be essential to validate identified missing relationships, ensuring their relevance and accuracy.

Overall, this study contributes to the ongoing refinement of SNOMED CT and highlights its significance in supporting accurate and comprehensive data management and analysis in healthcare informatics. By addressing information gaps and proposing solutions to rectify them, this research strives to improve medical communication standards and ultimately enhance patient care outcomes on a global scale.

7. References

[1] SNOMED International. Retrieved from https://www.snomed.org/

[2] Rehurek, R. Word2Vec - gensim. Retrieved from https://radimrehurek.com/gensim/models/word2vec.html

[3] Agrawal, A. Evaluating lexical similarity and modeling discrepancies in the procedure hierarchy of SNOMED CT. BMC Med Inform Decis Mak 18 (Suppl 4), 88 (2018).

[4] Rector, A. L., Brandt, S., & Schneider, T. (2011). Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. Journal of the American Medical Informatics Association, 18(4), 432–440.

[5] Park, H. A., Yu, S. J., & Jung, H. (2021). Strategies for adopting and implementing SNOMED CT in Korea. Healthcare Informatics Research, 27(1), 3.

[6] Agrawal, A., Cui, L. Quality assurance and enrichment of biological and biomedical ontologies and terminologies. BMC Med Inform Decis Mak 20 (Suppl 10), 301 (2020).

[7] A. Agrawal and P. Revelo, "Analysis of the consistency in the structural modeling of SNOMED CT and CORE problem list concepts," 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 2017, pp. 292-296