



Harmonizing EHR Databases with SNOMED CT

*Presenter: Nor Azlinayati Abdul Manaf, Khalil Ben Mohamed (PhD), Dickson Lukose (PhD)
Artificial Intelligence Centre, MIMOS Berhad, Malaysia*

Audience

Staffs from National Release Centres or Healthcare Institutions who are involved in SNOMED CT adoption; vendors providing systems that allow encoding data with SNOMED CT concepts.

Objectives

To present and demonstrate Mi-Harmony tool that encodes structured and unstructured data from electronic health record databases with SNOMED CT codes. To present the use-case with EHR database from Hospital Serdang, Malaysia.

Abstract

Most of the existing Electronic Health Record (EHR) systems that are available on the market today are rather 'hospital-centric' than 'patient-centric'. Different vendors supplying the EHR systems use their own standards and vocabularies in storing the patient information in their databases resulting in difficulty to exchange healthcare information from different hospitals that uses different EHR systems. Furthermore, majority of important information is recorded and stored in the form of free text such as medical history, patient diagnosis, etc. This information is extremely difficult to manage on a large scale especially if to be used with a clinical decision support system (CDDS).

SNOMED CT is a systematized nomenclature of medical terminologies and concepts. Using SNOMED CT will enable EHR systems to communicate in a common language. This will result in increasing quality of patient care across many different provider specialties. However, given the complex structure of the nomenclature, incorporating SNOMED CT on the data entry level of EHR system could be a major challenge to the EHR vendors.

Mi-Harmony is a tool that encodes data stored in an EHR database with SNOMED CT codes. It encodes both structured data (e.g., column's name) and unstructured data (e.g., medical note in the form of free text), creating a standardized format for further exchange and processing by machines. This tool is designed to perform the encoding process at the database level, thus reducing the interruption on the data entry flow of existing EHR systems. It provides a user interface for semi-automated mapping of both column names and reference tables content. However, encoding of unstructured information within the database tables is done automatically using String manipulation and Natural Language Processing techniques. The encoded medical information is stored in a different database for further use such as analysis and reporting purposes. Mi-Harmony tool has first been tested on the EHR database from Hospital Serdang, Malaysia. The database consists of 2219 tables, including both system tables and patient related tables. Among the challenges we encountered were: no schema provided for the database; varieties of abbreviations and acronyms; and mixes of English and Malay terms in the clinical notes. Preliminary evaluation on the completeness and accuracy of the resulting codes shows 90% precision and 70% recall. It shows that most of the time the tool correctly encodes terms with SNOMED CT codes, but there are still 30% of the terms that are not recognized. Eventually, by encoding the EHR database using Mi-Harmony tool, we were able to overcome the aforementioned challenges and produced a coded EHR database standardized with SNOMED CT codes, thus enabling future information exchange and interoperability between different hospitals using different EHR systems.