MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Mapping of Korean Healthcare Vocabulary to the UMLS

Yoo-Kyung Boo[a], Seung-Hee Kim[a], Mi-Sook Kwak[a], So-Yong Yoon[a], Yoon Kim[a]

[a] *R & D Center for Interoperable EHR, Korea*

Abstract

The UMLS is used for research and development across wide range of different applications. In this paper, we evaluated the coverage of the UMLS as mapped with healthcare vocabularies in Korea. Of the 216,697 terms, 42% were matched with the UMLS concepts exactly. We then analyzed unmatched terms. There were various reasons why 58% of our terms were unmatched with the UMLS concepts.

Keywords:

vocabulary, controlled; Korean language; UMLS

Introduction

There is an increasing need for successful terminology mapping to achieve compatibility among terminologies and to allow clinicians and researchers familiar with one terminology to search and use other terminology schemes [1]. The Unified Medical Language System (UMLS) has incorporated over 100 sources into one database [2].

In this study, we evaluated the coverage of the UMLS as mapped our terms to the UMLS Metathesaurus (MT) using string matching tool. Then, we analyzed mapping results.

Methods and results

Mapping to the UMLS MT

The National Health Information Standard Committee collected healthcare vocabularies which are used in twelve departments (or domains)—Clinical Findings and Problems, Diagnoses, Operations and Procedures, Nursing, Laboratory, Imaging, Medical Device, Dentistry, Oriental Medicine, Drug, Public Health, and Health Statistics—for standardization of healthcare vocabulary.

We mapped 216,697 healthcare vocabularies to the UMLS MT using string matching tool which was developed by the Vocabulary Server Development team in the Interoperable EHR Center. The mapping results are provided in Table 1.

Analysis mapping results

Of the total terms, 91,056 (42%) terms matched with the UMLS MT exactly. The Diagnoses had highest mapping percentage and the Medical Device had lowest mapping percentage. Un-mapping terms were the Korean alphabets (Hangul), composite concepts, misspellings, and domain specific terms. There were some of un-mapping terms which were included conceptually in the UMLS MT, because we did only string matching. From this result, we could know coverage of the UMLS MT. The UMLS MT has many terms which are related to diagnoses, laboratory, operations and procedures, and clinical findings and problems. On the other hand, the UMLS MT has a few terms which are related to medical device, imaging, drug, and oriental medicine.

Table 1 – Mapping Result to the UMLS MT

Domain	Matched terms (%)	Unmatched terms (%)	Total
Clinical Findings & Problems	39,927 (61.0)	25,523 (39.0)	65,450
Diagnoses	24,453 (84.8)	4,381 (15.0)	28,834
Operations & Procedures	5,661 (74.7)	1,915 (25.0)	7,576
Nursing	93 (12.8)	634 (87.0)	727
Laboratory	13,256 (79.5)	3,412 (20.0)	16,668
Imaging	5 (0.1)	3,888 (99.9)	3,893
Medical Device	0 (0.0)	130 (100.0)	130
Dentistry	5,867 (27.1)	15,763 (73.0)	21,630
Oriental Medicine	76 (2.1)	3,534 (98.0)	3,610
Drug	395 (0.6)	65,045 (99.0)	65,440
Public Health	1,313 (54.0)	1,117 (46.0)	2,430
Health Statistics	10 (3.2)	299 (97.0)	309
Total	**91,056 (42.0)**	**125,641 (58.0)**	**216,697**

Conclusions

The high percentage of medically relevant unmatched terms is attributed to a variety of factors. The UMLS MT has incomplete synonyms and absent concepts. Also, healthcare vocabularies have different term structures to the UMLS MT and misspellings.

Acknowledgments

This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (A050909).

References

[1] Vizine-Goetz D, Hickey C, Houghton A, and Thompson R. Vocabulary Mapping for Terminology Services. Journal of Digital Information. 2004 Mar [cited 2006 Mar 13]; 4(4): [about 23 p]. Available from http://jodi.tamu.edu/Articles/v04/i04/Vizine-Goetz/

[2] UMLS Knowledge Sources. Bethesda (MD): National Library of Medicine, 2006AA.

Address for correspondence

E-mail: shscool@snu.ac.kr