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

Developing An Automated Mapping Algorithm for the ICNP in Japanese

Kimikazu Kashiwagi [a], Masayo Kashiwagi [b]

[1] [a] National College of Nursing, Japan
[b] Tsukuba University, Japan

Abstract

OBJECTIVES: The purpose of this study is to develop the mapping method from terms which are used in clinical settings to the Japanese Translation of the ICNP version 1. METHOD: We used n-grams for approximate matching and calculated the similarity scores between terms by the length of the n-grams and the frequency of n-grams in the ICNP. Using this method, we were able to create the lists sorted by the similarity scores. RESULTS: The proposed method seems to be useful in the mapping process.

Keywords:

ICNP, translating, terminology, International Council of Nurses

Introduction

The Japanese translation of the ICNP version 1 (International Classification for Nursing Practice) was published in July 2006. But one of the reason is that the concepts in ICNP is that some translation terms are rarely used in clinical settings.

In order to use the ICNP as a reference terminology, it is necessary for users to find the ICNP concepts from Japanese terms easily.

The simple word matching algorithm was not appropriate if the results suggest more than 20 terms.

Methods

At first, we extracted 20838 n-grams from the Japanese translation of the ICNP. To calculate the similarity scores between a query string and terms in the ICNP, we used following equation.

$$\sum_{n \in Q} \frac{Length(n)}{Freq(n)} \quad (1)$$

where

Q are the N-grams made by a query string

Length(n) is the length of the N-gram

Freq(n) is the number of documents containing the N-gram

To evaluate the method, the Japan Nursing Practice Standard Master (JNPSM) that includes the names of nursing activities and is built by the Medical Information Center Development Center (MEDIS-DC);

There was not full agreement in the mapping between the experts. We used 48 sets of the mapping as a gold standard.

Conclusion

The proposed method seems to be useful in the mapping process.

References

[1] Shannon CE, Weaver W. The Mathematical Theory of Communication. Urbana and Chicago: University of Illinois Press; 1949.
[2] Heja G, Surjan G. Using n-gram method in the decomposition of compound medical diagnoses. Int J Med Inform. 2003 Jul;70(2-3):229-36.