

Construction of an ontology-based nursing knowledge system[†]

Discussion

Shi-Fan Han^{a,*}, Rui-Fang Zhu^a, Jia Xue^a, Qi Yu^b, Yan-Bing Su^b, Xiu-Juan Wang^c

^aFirst Hospital of Shanxi Medical University, Taiyuan, Shanxi 030001, China

^bSchool of Management, Shanxi Medical University, Taiyuan, Shanxi 030001, China

^cSchool of Nursing, Shanxi Medical University, Taiyuan, Shanxi 030001, China

Received: 18 May 2018; Accepted: 10 September 2018; Published: 20 December 2018

Abstract: This study proposes the establishment of a knowledge-system ontology in the nursing field. It uses advanced data mining techniques, digital publishing technologies, and new media concepts to comprehensively integrate and deepen nursing knowledge and to aggregate sources of knowledge in specialized technical fields. This study applies all forms of media and transmission channels, such as personal computers and mobile devices, to establish a knowledge-transmission system that provides knowledge services such as knowledge search, update retrieval, evaluation, questions and answers (Q&As), online viewing, information subscription, expert services, push notifications, review forums, and online learning. In doing so, this study creates an authoritative and foundational knowledge service engine for the nursing field, which provides convenient, flexible, and comprehensive knowledge services to members of the nursing industry in a digital format.

Keywords: nursing • knowledge system • ontology construction • big data • text mining • review

© Shanxi Medical Periodical Press.

1. Basic overview of ontology construction for the nursing field

In recent years, with the rapid development of computing, network technology, and information sciences, various industries have started implementing vigorous reforms aimed at high-level application of information technology; the nursing industry is no exception in this regard.¹ In the 2014 Nursing Knowledge: Big Data & Science for Transforming Health Care Conference, hosted by the International Council of Nurses, it stated that “the future of nursing work will focus on changing nursing

practices, research, and education in the context of big data, and will emphasize development and training related to capabilities of obtaining and integrating information.” On March 6, 2015, the Chinese State Council published the *Outline for the Planning of the National Medical and Health Service System (2015–2020)*, which clearly proposed the “active utilization of new technologies such as mobile Internet, Internet of Things, cloud computing, and wearable devices to promote health information services and smart medical treatment services which benefit all citizens, in addition to promoting big data applications and gradually transforming modes of service in order to improve service capabilities and management quality.” In March 2015, Joyce Sensmeier, the Vice President of Informatics at the Chicago-based Healthcare Information and Management Systems Society, published a specialized article in *Nursing Management*, expressing the challenges faced by nurses in

[†]This study was supported by National Natural Science Foundation of China (No. 71573162) and Shanxi Province Soft Science Research Program (No.2018041029-3).

* Corresponding author.

E-mail: 13903409675@139.com (S.-F. Han).

 Open Access. © 2018 Shi-Fan Han et al., published by Sciendo.  This work is licensed under the Creative Commons Attribution NonCommercial-NoDerivatives 4.0 License.

the big data era, in addition to their thoughts on the rational utilization of resources. In July 2015, the National Institute of Nursing Research (NINR), a member of the United States National Institute of Health, organized a conference on big data in the nursing field in an attempt to promote and encourage progress in this area.² On November 18, 2016, China's National Health and Family Planning Commission stated clearly in its *National Nursing Development Plan (2016–2020)*³ that the rapid development of information technology had created favorable conditions for advances in the field of nursing.

Big data will accelerate the penetration of new technologies beyond Internet applications and toward a broader range of fields and across all industries. Nursing is an independent medical discipline; in addition to nursing being an extremely high-skilled discipline, nursing work content extends throughout the process of health management. The main group responsible for nursing quality is nursing staff themselves; nurses depend on complete, accurate, and timely data and information in order to make the best clinical decisions, and these data must be presented according to standardized structures in order for the data to be shared and compared. However, inconsistent data standardization methods and different nursing terminologies are important barriers to the sharing of nursing information and data. When the sharing of information between systems is restricted, nursing big data become incompatible and incomparable. In addition, information technology education in the nursing field in China began relatively late, meaning that information technology in the nursing field in China is underdeveloped and that Chinese nursing workers lack sufficient knowledge of concepts and characteristics of big data and cloud computing. These factors have produced significant effects on the pace of development of big data in China's nursing field. In summary, with the arrival of the big data era, it is crucial that big data be integrated organically with nursing information, that nursing disciplines be instructed on how to more effectively retrieve and utilize big data, and that a nursing knowledge system be constructed rapidly and effectively.

It is indicated that 90% of all data and information generated over the past 12,000 years of human history were produced within the past 5 years. Furthermore, the rate of data growth doubles every 2 years.⁴ Large volumes of unstructured numbers do not appear to follow any rules, although they actually contain a wealth of information and huge potential value waiting to be extracted. Since 2009, when the United States Congress passed the HITECH Act,⁵ the attractiveness of big data in medical and health fields has become apparent. In July 2015, the United States NINR hosted a conference on big data in the field of nursing and attempted to promote the development of big data in nursing.²

The development of big data in China's nursing and health-care fields lags behind that of other countries, although some developments have taken place. For example, in September 2015, China's State Council published the *Action Outline for Promoting the Development of Big Data*, which provided policy guidance for the use of big data in medical and health services and in social security services, in addition to conceptualizing the establishment of a big data application system for medical and health management and services. The abovementioned outline also encourages relevant departments to develop practical and innovative research related to big data.⁶ This shows that as we approach the big data era, the development of research on big data has become an important trend in academic development. Large volumes of data in different forms and arriving through different channels provide nurses with new methods and ways of thinking. Big data have valuable applications in areas as diverse as nurse evaluation, improving the quality of nursing practices, disease monitoring, nursing research, and clinical decision-making support.

2. Overall objectives of constructing a nursing ontology

By constructing a nursing knowledge system, it is possible to better understand the context of nursing knowledge and reveal the composition, hierarchy, and sources of nursing knowledge. This system deepens nursing knowledge and facilitates the establishment of a method for navigating knowledge classification and information resource positioning in the nursing field, in addition to providing a foundation for a nursing knowledge service system, thereby supplementing shortcomings in the establishment of a nursing knowledge system and facilitating the promotion of information transmission and exchange in the nursing field. At the same time, this helps to improve the social and commercial influence of nursing information services and creates a positive information-transmission effect that directs public attention toward nursing information, finally achieving an organic integration of big data and nursing information. This optimizes the nursing information system and allows big data to be applied in clinical work via a nursing information platform, thereby improving the work efficiency of nurses; reducing hospital costs; improving nursing quality; and providing efficient, high-quality, and safe nursing services for patients, while also playing an active role in disease prevention and remote nursing and monitoring.

First, the construction of a nursing knowledge-system ontology can help to achieve a structured digitized production process for complete processes, including compilation, review, management, and publishing, and can

accelerate the digitization of nursing knowledge, thereby maintaining a leading position in technological and market competition without losing this status over time.

Next, promoting the construction of a universal media knowledge service platform on the foundation of a digital publishing platform and constructing a nursing knowledge service system on a digitized and fragmented foundation will allow publishers to shift from being providers of traditional resources to providers of knowledge services, thereby achieving added value with regard to resources and knowledge.

Finally, by establishing a universal media knowledge base for the nursing industry, it is possible to achieve the in-depth clinical application of knowledge services from academic journals to medical nursing and scientific learning processes, providing key knowledge and intellectual support for addressing problems and integrating research and clinical care, thereby promoting the process of transformation and upgrading.

3. Functional objectives of the nursing knowledge system construction

Based on all varieties of existing information resources, the construction of a nursing knowledge system, semantic annotation, and log analysis in order to construct a nursing knowledge system are studied in this article. By developing a system for constructing a nursing knowledge network, it is possible to manage the process of obtaining, processing, integrating, and structuring resources. According to structured data, this can provide functions such as queries, statistical analysis, visualization, report generation, and individualized services. The construction of a nursing ontology provides a knowledge system in the service of nursing scientists and provides topical and updated services for nursing researchers. In addition, it provides services for nursing disciplines, topical trends, disease forecasting, and nursing information assistance services for clinical nursing departments.

4. Core requirements for the construction of a nursing ontology

The construction of a big data nursing ontology is based primarily on providing nurses with multimedia content such as teaching materials, nursing literature, nursing books, nursing guides, dictionaries, images, audio and video content, library content, network content, and other multimedia resources. The construction of a nursing knowledge system and service system is supported by a digital resource-management system. It primarily includes the following: a universal media-management

system and ontology-construction system, a knowledge-based publishing platform, an encyclopedic service system, an application platform, a unified user-management and query-control system, an information-aggregation system, and a mobile-end publishing platform.

5. Core technologies for the establishment of a nursing ontology

5.1. Domain knowledge-extraction techniques

In order to transform domain knowledge into an ontology, structured vocabularies are input and a generated initial ontological model is output. Key processes involved include the transformation of concepts within structured vocabularies into ontological classes, transformation of the hierarchical structures of concepts into the relationships between classes in the ontology, and descriptions of concepts into attributes of ontological classes.

The first type of structured vocabulary includes the basic concepts of a domain, in addition to various attributes of these concepts. In order to make full use of the domain knowledge contained in vocabularies, it is necessary to include each attribute of vocabulary concepts within the ontology. Because organized vocabularies do not have the same number of names of attributes and attributes themselves, and because concepts are not located in the same place, in order to provide a relatively universal management process, the system provides a code-based custom format vocabulary-input function. Users need to only enter the vocabulary format in order for the vocabulary to be automatically entered into the ontology.

After the conversion of basic concepts in the vocabulary is completed, it is necessary to automatically convert the hierarchical structures of concepts into relationships between ontological classes. Because vocabularies using this type of format also provide codes, each class code includes both a code for itself and a superclass code. As such, codes can be interpreted to determine the hierarchical relationship between classes.

The second type of vocabulary structure only includes the concepts themselves and the hierarchical structure between concepts. During processing, concept codes are generated according to the hierarchical structure of the concept in order to facilitate the conversion of hierarchical structures into ontological class relationships.

The third type of vocabulary structure-processing method is basically the same as the second type of vocabulary structure. First, concepts are converted into the second vocabulary format and then encoded; the encoding process is the same as mentioned earlier.

After encoding, ontology input is executed, and finally the hierarchical structure of concepts is converted into the relationships between ontological classes.

5.2. Knowledge relationship-extraction techniques

The inputs for knowledge relationship-extraction functions include specialized teaching materials and literature resources, whereas the output is ontological class attribute values. Specialized teaching material and literature inputs must undergo a simple formatting process in order to increase the efficiency and accuracy of knowledge extraction.

With regard to attributes and attribute values, in order to extract knowledge relationships, one must first import texts relevant to specialized teaching materials and literature to serve as the value of attribute 1 of the class for input into the ontology. During input, data blocks serve as the unit for data input, and one data block is processed at a time. After relevant text from specialized teaching materials and literature is inputted into the ontology, the value of a concept's second attribute is extracted from attribute 1, and knowledge relationships are thereby extracted.

5.3. Attribute feature-extraction techniques

5.3.1. Feature extraction based on structured and semi-structured text fragments

Considering that initial ontology data are sourced from structured text materials, such as manually edited authoritative teaching materials, the quality of these teaching materials has great significance for the construction of this initial ontology. The extraction of attributes from structured and semi-structured documents is an important element for the large-scale automated construction of initial ontologies.

5.3.2. Unguided feature extraction from multiple documents

With regard to knowledge-intensive text fragments, knowledge descriptions in individual documents are often limited. Although a certain hierarchy exists in the structure of content expression and concepts have certain scopes of inclusion, limited knowledge descriptions for individual texts are often unavoidable, as individual texts cannot reflect the most recent descriptions of knowledge across texts of the entire domain. Often, knowledge classification, editing, and sorting are dependent on large amounts of manual work, as there is no single effective knowledge description

and feature-extraction mechanism for knowledge-intensive text fragments. Updating knowledge also typically requires manual work, as it is not possible to satisfy the requirements for updating representations of knowledge in large volumes of knowledge-intensive text fragments. As such, the system uses an unguided multi-document feature-extraction method to accurately extract vocabulary from texts on a large scale, thereby reducing the cost of manually organizing knowledge.

5.4. Semantic annotation and semantic indexing techniques

The basic thinking underlying semantic annotation is that one should not follow the content of the entire document when indexing resources. One reason for this is that descriptions of a concept in an entire document are too broad and may touch on many different elements of a concept. Another reason for this is that the semantic content of the entire document may be related to multiple concepts, making it difficult to summarize the basic semantic content. This study follows the two dimensions of paragraphs and sentences to conduct semantic annotation. When annotating a paragraph, the semantic indexing of the entire section is emphasized, whereas there is a greater focus on the extraction of relevant concepts when annotating sentences.

When annotating digital resources, one must first consider the source of the digital material. Generally, domain documents are sourced from domain texts that have been manually proofread. These texts have structured and semi-structured features, and annotated information is relatively accurate. Another portion of digital resources comes from network domain texts and has unclear structural features, and the annotated corpus is generally used for reference. Next, digital resources undergo corpus processing, forming a corpus used for annotation. During annotation, a text vector space model is used as a fundamental model for domain corpus analysis, wherein relevant concepts and concept attributes in document fragments are given higher annotative weights. After annotation, the large-scale domain corpus forms the structure for semantic indexing; the keys to indexing are domain concepts and concept attributes, whereas the value of citation is the location of the document, document content, and the semantic information of the document. Because the annotation of domain texts consumes a relatively large amount of computing resources, and because semantic annotation is generally an offline computing process, after a text is annotated, the citation structure of a domain text is saved in a certain form to facilitate knowledge searches.

5.5. Mutual conversion techniques for ontology publishing and ontology construction

In order to give full play to the knowledge-management functions of the knowledge-management systems, the system uses ontology data to adapt to the database structure of ontology content publishing and proofreading systems and to interact with them.

After the initial ontology is constructed, it is necessary to publish the ontology publishing and proofreading systems to facilitate maintenance and management by domain experts and for the reference of ontology users. When interacting with the ontology content publishing and proofreading system, ontology classes serve as the basic unit of interaction, and each class is a topic within the knowledge-management system. This topic includes all the attributes of the class and the superclass and subclass of the class. After the ontology is published to the ontology content publishing and proofreading system, the ontology synchronizes data with the ontology content publishing and proofreading system database with each operation to ensure that

the ontology database and ontology content are published and that the data in the proofreading database are synchronized.

5.6. Resource integration, precision analysis, and extraction techniques

The analysis of various resources extracted locally and across networks provides tools for distinguishing between words, sentences, and paragraphs and allows for granular analysis and structuring of resources according to sections, chapters, paragraphs, or sentences. With regard to types of literature, the system supports the standardized analysis of books, periodicals, theses, and patents; with regard to content, the system supports the analysis of texts, images, and tables.

6. Developing the nursing ontology-construction system

Subsystems that still need to be developed include the following (Figure 1).

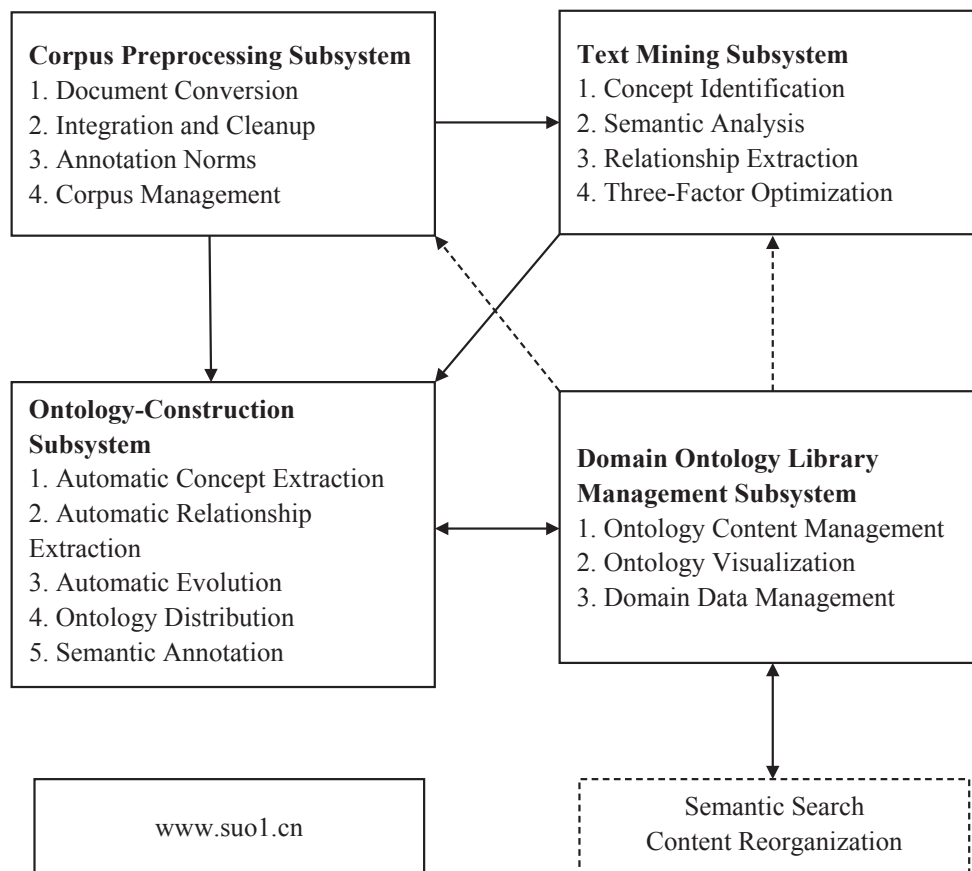


Figure 1. Nursing knowledge-construction system in nursing field.

6.1. The corpus pre-processing subsystem

This subsystem provides a corpus input-management function and allows for operations such as additions, deletions, revisions, and searches on all types of corpus resources. This subsystem also provides an information-extraction function. For situations in which no specialized vocabulary exists, the nursing dictionary corpus is used as a foundation for applying the automatic construction of vocabularies using semantic dictionaries. This vocabulary is then inputted into the ontology to generate the concept tree structure and conduct information extraction.

6.2. The text-mining subsystem

This subsystem provides concept-identification functions and uses the identification and determination of concepts related to specific domains to form conceptual groups within domains. The text-mining subsystem also provides semantic analysis and relationship-extraction functions and uses identification to determine key text within a corpus and to extract the inheritance relationships between concepts, in addition to the triple relationships such as synonymous relationships, attribute relationships, and example relationships. This subsystem also provides triple-component optimization functions by classifying and organizing inheritance relationships, synonymous relationships, attribute relationships, and example relationships, and identifying and eliminating meaningless triplet information.

6.3. The ontology construction subsystem

This subsystem provides functions for converting domain vocabulary into domain ontology. It also provides ontology-generation functions and converts triplet relationships into ontology. The subsystem makes deductions according to the existing knowledge and rules and extracts new information. Ontology learning is a process of semantic analysis and relationship extraction, and using automatic relationship-extraction techniques, it is possible to extract the relationships

between concepts in teaching materials and establish triplet relationships, thereby forming conceptual nodes connected by semantic relations in text resources.

6.4. Domain ontology library management subsystem

This subsystem manages domain data and ontology content and visually displays the ontology.

7. Conclusions

In this study, the use of natural language-processing techniques and machine-learning methods to construct a nursing knowledge-system ontology represents a novel data-processing technique. For example, this study reorganizes and utilizes professional vocabularies, professional thesauruses, specialized teaching materials, and authoritative works. It also applies semantic annotation techniques to semantically annotate literature resources and constructs a standardized, expandable, and sharable nursing big data system founded on the basis of a domain ontology. It achieves the integration of nursing domain knowledge resources and provides knowledge services as a foundation for innovative and transformative applications in the nursing field in the era of big data. As such, the construction of a nursing knowledge-system ontology not only benefits the standardization of nursing information services and the maximization of the value of nursing resources but also promotes the standardized and universal development of nursing information services and encourages advances in integrated publishing. Furthermore, it unifies social and economic benefits and creates richer nursing knowledge and interactive experiences for the public. Subsequently, after the system's construction is perfected, more functions will be designed on the platform according to the needs of different groups.

Conflicts of interest

All contributing authors declare no conflicts of interest.

References

1. Li BB, Wu XY. Thoughts on nursing information technology in the big data era. *J Nurs Sci*. 2016;31:91-92 (in Chinese).
2. NIH. NINR big data in symptoms research methodologies boot camp. <https://www.ninr.nih.gov/training/trainingopportunitiesintramural/bootcamp>. Accessed February 18, 2017.
3. Zhu RF, Yue XD, Han SF. Research progress on big data in nursing. *Chin Nurs Res*. 2017;31:1553-1557 (in Chinese).

4. EMC. Data growth, business opportunities, and the IT imperative. Accessed February 18, 2017. <https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>.
5. Strobel CD. American recovery and reinvestment act of 2009. *J Corporate Accounting Finance*. 2009;20:83-85.
6. Central Government of the People's Republic of China. Notice of the state council on issuing the action outline for promoting the development of big data (2015). Accessed February 18, 2017. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm (in Chinese).

How to cite this article: Han S-F, Zhu R-F, Xue J, Yu Q, Su Y-B, Wang X-J. Construction of an ontology-based nursing knowledge system. *Front Nurs*. 2018; 4: 241-248. <https://doi.org/10.1515/fon-2018-0035>.