Implementation of Automated Knowledge-based Classification of Nursing Care Categories

Shihong Huang, Subhomoy Dass, Sam Hsu, Abhijit Pandya
Department of Computer & Electrical Engineering and Computer Science
Florida Atlantic University
Boca Raton, FL U.S.A
(shihong, sdass, samh, pandya) @ fau.edu

Abstract

An important aspect of a health care delivery system is nursing. The use of technology is a vital aspect for delivering an optimum and complete nursing care to individuals; and also for improving the quality and delivery mechanism of nursing care. The model proposed for Nursing Knowledge Management System is a novel knowledge-based decision support system for nurses to capture and manage nursing practice, and further, to monitor nursing care quality, as well as to test aspects of an electronic health record for recording and reporting nursing practice. As part of ongoing collaborative research of the nursing school and the department of computer science, a prototype toolset was developed to capture and manage nursing practice in order to improve the quality of care. This paper focuses on implementing a web based SOA based solution for Automated Classification for Nursing Care Categories, based on the knowledge gained from the prototype for nursing care practice.

Keywords: Nursing care, LSI, Knowledge Management, Classification, Natural Language Processing, Web Services, and SOA.

1. Introduction

Florence Nightingale demonstrated the importance of proper and systematic nursing care as early as 1859. However, it was only in the late 1970's that the importance of nurses was finally recognized. International Council of Nurses defines nursing as development of other components of hospital information system.

Traditional role of nurses was to provide care and comfort to patients. However, the role of nurses has evolved and now includes health promotion and illness prevention [1]. America's Nurses Association has recently defined nursing information practice as facilitation of data, information and knowledge to support patients in their decision making [2]. Vast amount of data with useful as well as redundant information is generated in day-to-day nursing practices. Proper organization and application of this information is essential in the dynamically changing healthcare scenario. Many nursing records are not well written or maintained and are filled with redundant information [2]. Automated data gathering will prove very useful in systematic gathering, storage and utilization of data qualitatively and quantitatively, promoting the implementation of the reliable and efficient health care system.

This paper focuses on implementation, integration and analysis aspect of the above system with respect to automated nursing classification search. A clinical database and analysis tool enables nurses to document and improve characteristics of their practices. This information is vital if nurses and medical practitioners expect to achieve success in the changing U.S. health care system. The need to record and manage nursing data is not only due to federal requirements or out of the need to recall the treatment that had been rendered, but, is now an important tool in communication between healthcare professionals.

Nursing knowledge consists of the data collected during the interaction of the nurses and the patients; the data is captured in natural language and textual format. The data sets are then analyzed by the nurses and classified using their knowledge according to standardized nursing language. The nursing data is classified according to four broad categories i.e. Caring, Respect, Wholeness and Connection by nursing domain expert. The classification search index uses an LSI (Latent Semantic Indexing) [3] based indexing scheme using the classified data based on the above categories, for indexing. The search is performed based on the data provided by end user(s)/nurses and provides the classification to the end user. A plan of care is documented using the classification. The plan of care helps in providing a more holistic care to the patient, as it also focuses on non physical aspects or nursing care, which is also crucial for patient recovery.

IEEE IRI 2011, August 3-5, 2011, Las Vegas, Nevada, USA
978-1-4577-0966-1/11/$26.00 ©2011 IEEE

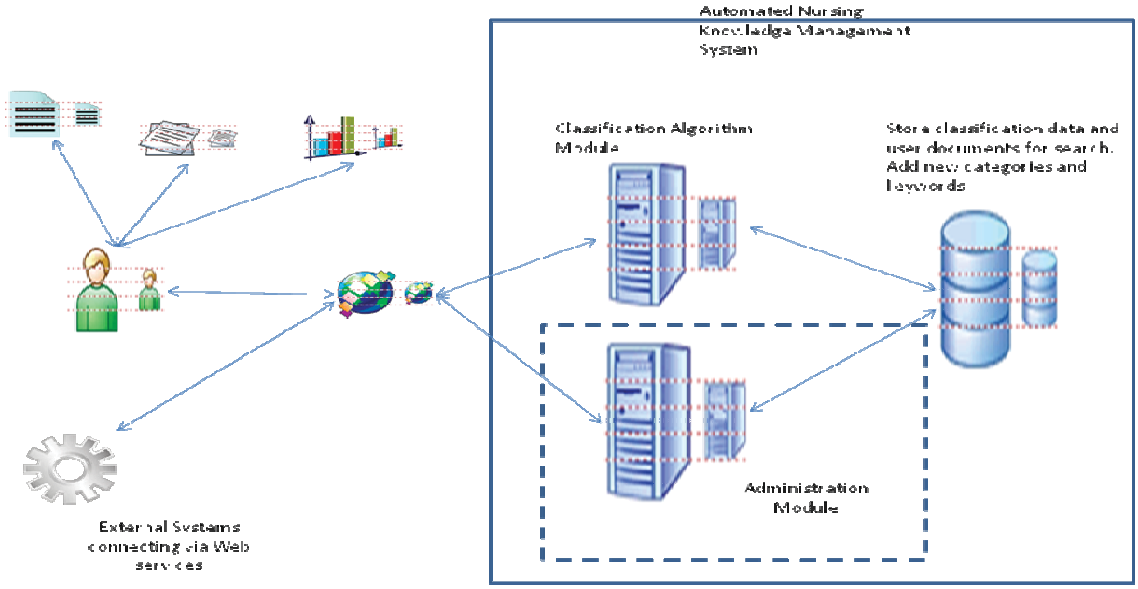

Figure 1. High level system architecture

2. Background/Related Work

The Nursing care practice as defined above consists of two aspects, Qualitative Analysis and visualization [4] and automated classification of nursing data based on categories defined above [5]. There are some general approaches to nursing knowledge management; some representative methodologies for natural language classification are described in Section 3.1. For example use of printed forms for collecting and analyzing nursing knowledge is laborious and time consuming and has logistical and storage issues. Also, the mapping to various nursing categories is done manually by the nurses or the nursing domain experts. There are some CASE tools available which can be also used for classification; this method allows for classification by nurses or nursing domain experts in a semi automated way. In this method, data is generally stored in an electronic format. The category assignment is manual in this case.

Some of the technologies generally used for qualitative data analysis are Semantic Web, Naïve Bayes Classification and Latent Semantic Indexing (LSI). The above methods provide for a way to classify data according to the nursing classification and to retrieve search. This paper focuses on applying LSI indexing algorithm for automated nursing classification using J2EE technologies. The implementation is web based and the framework uses service oriented architecture (SOA) [6], which enhances the usability, portability and interoperability of the solution, thus providing seamless way of integrating the system with other products and solutions as another component. LSI uses Singular Value Decomposition matrix which is a subset of the original document term matrix, on which searching can be done. LSI is able to correlate semantically related terms in a data or document set. A proof of concept for automated classification using the LSI approach was developed using Microsoft technologies. The proof of concept was a client server based application; and could not be enhanced or extended. This paper is a continuous work from [5], and focuses on implementation of automated nursing classification algorithm and advances the previous work by implementing a scalable, multi user, J2EE based, SOA oriented system using open source components.

3. High Level System Architecture

This paper is structured based on Tilley's [7] description of the three categories of canonical activities that are characteristic of reverse engineering for program understanding, i.e. data gathering, knowledge management and information exploration. The high level system architecture for the proposed system is shown in Figure 1. The data from various sources will be converted by available converters into text format. The converted text can be then automatically classified by running the data against the LSI [3] [5] based nursing classification indexing algorithm. The search results can then be used for further processing.

3.1 Data Gathering

The data is gathered from various input sources like field input [8], focus group [9], doctor's input, patient's

encounter forms [10] or voice to text devices [11]. Voice input may include nurse's assessment of a patient as well as discussions between doctor and nurse or the nurse and the patient and the patient's family members. Other forms of data gathering techniques have been extensively described in [10]. The input can thus be in any form which can be converted into .doc, .pdf or a text format. The inputs from the different sources will be stored in the database. Data from external sources can also be directly pulled in to the system via customized data adapters. The data can then be classified by the domain experts.

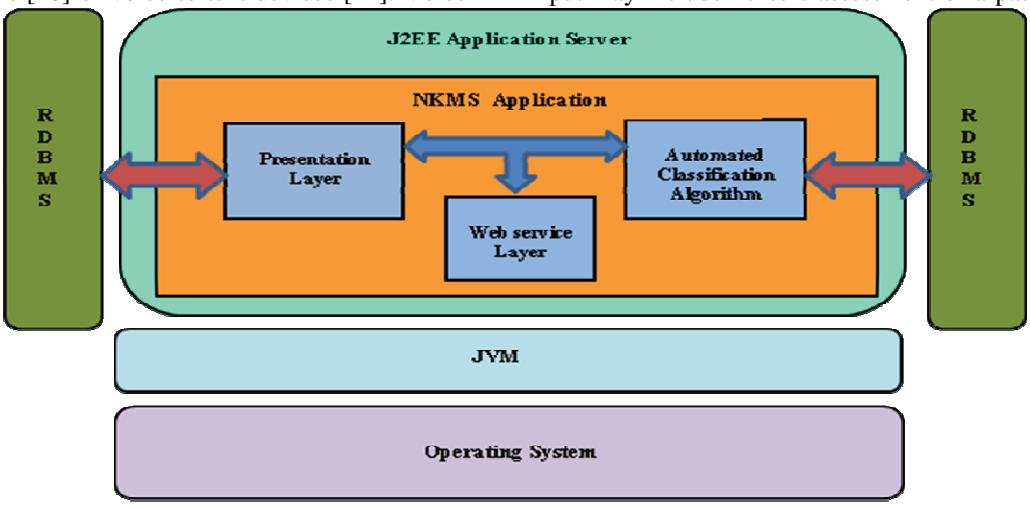

Figure 2. Automated nursing knowledge management system components

3.2 Knowledge Management

Knowledge management encompasses the data and the algorithms like LSI and natural language process; which provide the end user a means of analyzing the results of the transformations and make decisions or process the results further. The results can be accessed by other applications via web services, which make the system an easily pluggable solution and can be leveraged by other systems with minimum changes.

The framework of proposed Automated Knowledge Management system is a web application based on J2EE technology [12] following Service Oriented Architecture (SOA) and consists of components as shown in Figure 2. The user can access the system via web services or using browser. The web application can be concurrently accessed by multiple users via the web. Since the application follows SOA, the application is interoperable and can be accessed by different applications running on diverse platforms or using different technologies. The rationale behind using J2EE instead of Microsoft technologies is to enable the solution to be portable and deployable on Windows, Mac and UNIX platforms without any modifications to code.

3.2.1. System Components

The proposed system consists of the Presentation Layer, Web Service Layer and the Classification Algorithm; also the interactions between the components are as shown in Figure 2. The system components are based on JAVA language [13] and the proposed deployment is based on Tomcat 6 [14]. However the application can be deployed on any J2EE application server deployed on Java 6 or above. Since the system is Java based, a wide variety of operating systems can be chosen based on need. The system components are developed using Apache Struts which is used for developing presentation layer [15], Apache Axis 2 [16] which enables the classification algorithm to be accessed as a web service, Apache Lucene [17] and JAMA (Java Matrix Package) [18] are the basis of implementing the LSI based search algorithm. The database of choice used for storing nursing classification information is MySQL [19] community edition.

3.2.2. Knowledge Classification and Indexing

In the proposed model, nursing information is classified into four main categories i.e. Caring, Respect, Wholeness and Connection. The information for each category is collected from different data sources as outlined in data gathering. The information is then categorized into categories mentioned above by nursing domain experts. The categorized word documents are then uploaded to the system using the GUI and stored against the corresponding categories. The information stored for each of the categories consists of complete definition

along with description and key terms and expressions which define the category.

The category data is indexed using Apache Lucene after removing irrelevant terms like English stop/noise words in additional to set of terms, and applying stemming algorithm(s). An M × N term document matrix is created. This matrix consists of M terms found in the document set and N documents entries for each term the frequency of term within the document divided by the number of times the term occurs across all documents are stored. Based on the term document matrix a Singular Value Decomposition term matrix is created, and terms are ranked. The search expression(s) are then matched against the term matrix and the matches are presented to the user(s). The user(s) will be presented the view of the search expression(s) matched against the categories.

3.3 Information Exploration

Information exploration provides for information to be used by third party applications or internal modules to present the data to the user for further analysis. The document focuses on Automated Nursing Management System which is one of the modules of the system. The original proof of concept was developed using .NET framework and related Microsoft technologies. The system was not portable, and was client server based application. Hence the proposed alternate system was developed which supports multiple users, is scalable and easy to integrate with existing systems or products. The system also uses a database for storing the classification data, which can be easily updated by the Admin module.

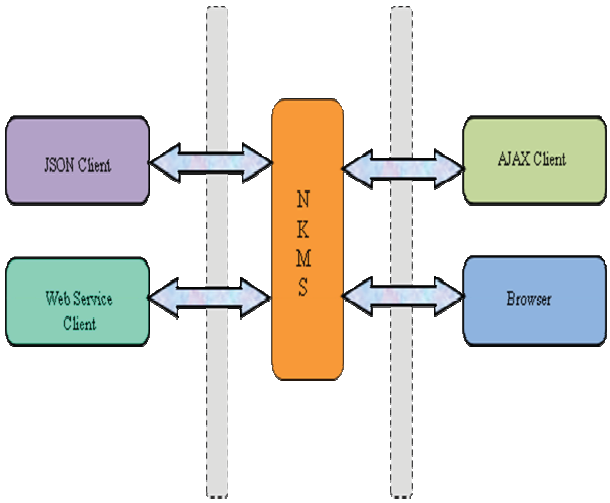

Figure 3. Representation of various clients accessing the system

The system is built using robust Relational Database and using J2EE technology and supports web services. The domain expert can classify the data by accessing the system via a browser from any machine. Users using any non java based web service client can access the web service by accessing the WSDL [20] provided by the system exposed as web services. As shown in Figure 3. the different types of clients can access the system without having any complex or expensive client and any time consuming client setup or installation involved.

The system has been designed keeping in mind inter operability, as this system will generally be used as a part of another encompassing system or it can also be used as a stand-alone system. There are no overhead costs involved in accessing the application, as developing a web service client or accessing the service via JSON [21] is simple and inexpensive. Since WSDL is auto generated, changes made are immediately available to consumers

The Relational Database used allows for any third party software or clients to access the data for data-mining purposes or for analyzing by querying the data. The option of exposing modules as a service allows for the application to be easily integrated with other systems. Also the use of Relational Database for storing information allows for other clients to be built based on end user on end user or system requirements. This ensures that the system is highly usable and accessible for end users or systems for exchanging information.

4. Results

The system tests focuses primarily on testing the SOA features of the system and validate the ease of integrating the system with any legacy or third party applications. Figure 4. shows the search results for searching the nursing classification index and viewing the results via browser. The system provides for a method of accessing the search functionality via web services using SOAP. In Figure 5. the SOAP request enclosing the search text is shown for searching the automated classification indexing algorithm. The search text contains text which needs to be classified by the automated classification indexing algorithm. Figure 6. shows the XML soap response for the search request from a web service client. The soap response results are categorized based on the categories, in which the search text is classified by the classification algorithm. The response then can be processed further by end user application or it could be input for another system.

The client can be implemented in any language supporting web services. The system has been tested with Java based XML and JSON web service client. The use of standards based JAX-WS [22] enabled web services in developing the system enables the system, to be accessed by other systems or end user applications developed in any language or tools which support web services.

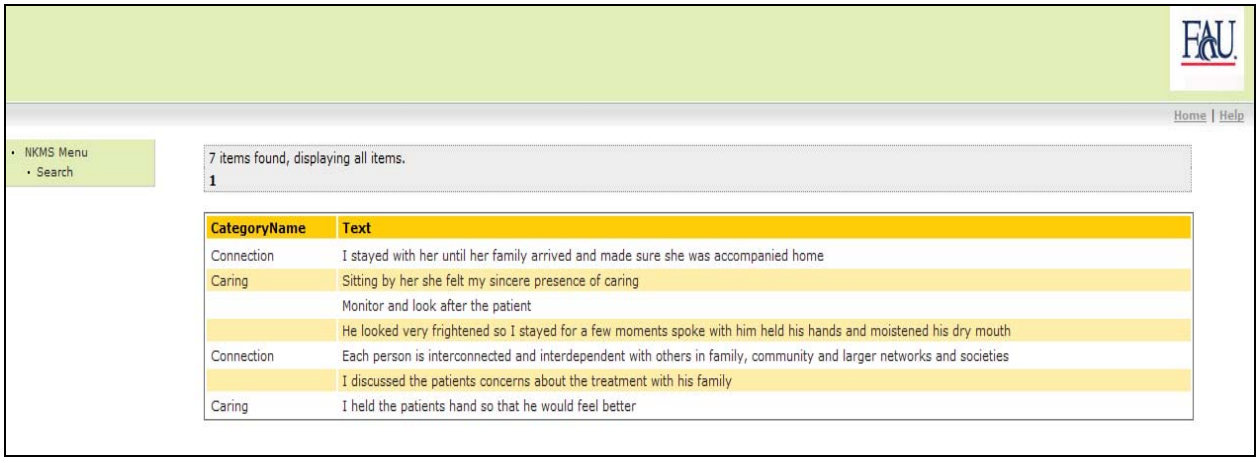

Figure 4. Search result classified based on user input.


```
<?xml version='1.0' encoding='UTF-8'?>
  <soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
    <soapenv:Body>
      <NKMSSearch>
       <searchText>Caring in nursing is nurturing well being of persons, families and communities based
      on knowing and understanding their values, interests, preferences and priorities. described caring as
      helping the other grow, and ingredients of caring as knowing, alternating rhythms, patience,
      honesty, trust, humility, hope and courage</searchText>
      </NKMSSearch>
    </soapenv:Body>
  </soapenv:Envelope>
```


Figure 5. XML SOAP search Request


```
<?xml version="1.0" encoding="UTF-8"?>
    <soapenv:Envelope xmlns:soapenv= "http://schemas.xmlsoap.org/soap/envelope/">
      <soapenv:Body>
        <dlwmin:NKMSSearchResponse xmlns:dlwmin="http://services.nkms.fau.edu">
          <dlwmin:Message>
            <statusCode>00</statusCode>
            <processedTime>2011-03-29 20:34:21 EDT</processedTime>
            <status>processed</status>
            <dlwmin:MessageData xmlns:xsi= "http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="dlwmin:nkmsContentVO">
              <categoryName>Caring</categoryName>
              <text>Caring in nursing is nurturing well being of persons, families and communities based on
      knowing and understanding their values, interests, preferences and priorities</text>
            </dlwmin:MessageData>
            <dlwmin:MessageData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="dlwmin:nkmsContentVO">
              <categoryName>Caring</categoryName>
              <text>described caring as helping the other grow, and ingredients of caring as knowing, alternating
      rhythms, patience, honesty, trust, humility, hope and courage</text>
            </dlwmin:MessageData>
          </dlwmin:Message>
        </dlwmin:NKMSSearchResponse>
      </soapenv:Body>
  </soapenv:Envelope>
```


Figure 6. XML SOAP search Response

5. Conclusion

The automated classification module implements LSI algorithm for searching the nursing classification data. The user can search the nursing classification index and the search data will be categorized based on four nursing categories i.e. Caring, Respect, Wholeness and Connection. The nursing domain expert can add more categories and add classification data which will enable the user to search on additional categories. The addition of data can be done by extracting data from various input sources, normalizing them and then storing them against

their respective category. The implementation is done using open source technologies and supports SOA methodology which makes it loosely coupled; hence any existing package can be customized according to application requirement. The application can be deployed on any platform supporting Java and can use any J2EE application server. The application is operating system independent and can be deployed on different platforms where Java is installed. The application and its modules can be easily integrated with any application, which need to use or integrate with the application leveraging interoperability feature provided by JAX-WS.

6. Future Work

The effectiveness of the automated nursing classification indexing algorithm depends on the quality and quantity of the data in the classification index. The nursing domain experts play a critical part in ensuring that the classification is done correctly and there is sufficient data for each of the categories. There also needs to be provision for extracting data from different systems which can provide nursing care related information.

The extracted data needs to be automatically categorized and then stored in the nursing classification database. The data which cannot be automatically classified need to be categorized by the nursing domain expert, so that they can be categorized or new categories could be created for storing the classification information. The application can also serve as an effective teaching tool in a scenario where extended nursing care is needed.

The care related information can be searched against the nursing classification index and based on matches or mismatches; the data could be categorized and added to the classification index. The system will need to be updated with classification data on a continual basis for the system to perform optimally.

There also needs to be further studies to explore possibility of automating the classification. This would reduce burden on nursing domain experts and make the system more effective and user friendly.

References

[1] Cho, I. and Park, H.: 'Development and evaluation of a terminology-based electronic nursing record system', Journal of Biomedical Informatics, Vol. 36, pp. 304-312, 2003

[2] American Nurses Association, 2001. The scope of practice of nursing informatics and the standards of practice and professional performance for the informatics nurse specialist. Washington DC, 2001.

[3] Latent Semantic Indexing, http://www.knowledgesearch.org/lsi/lsa_definition.htm Last accessed on February 08, 2011.

[4] Huang, S., Parker, M., Pandya, A., Hsu, S., and Tripathi, S.: 'Knowledge-based evaluation of nursing care practice'. Proc. 10th IEEE international conference on Information Reuse & Integration, (IRI'2009), pp. 171-174, 2009.

[5] Huang, S., Chinchanikar, S., Pandya, A., Hsu, S., and Parker, M.: 'Automated Nursing Knowledge Management Using Indexing'. Proc. International Conference on Software Engineering & Knowledge Engineering, (SEKE'2009), pp. 93-96, 2009.

[6] Channabasavaiah, K., Tuggle, E., Jr., and Holley, K., Migrating to a Service-Oriented Architecture, http://www.ibm.com/developerworks/library/ws-migratesoa/ Last accessed on February 08, 2011

[7] Tilley, S.: 'The Canonical Activities of Reverse Engineering'. Annals of Software Engineering, Vol. 9, pp. 249–271, 2000.

[8] Nursing Assessment in Family Nursing Practice, http://www.scribd.com/doc/4430506/ Family-Nursing-Care-Plan Last accessed on February 28, 2011.

[9] Chamane, N. and Kortenbout, W.: 'Focus group interview as a data gathering tool: its application to nurses' understanding of HIV infection and AIDS'. Curationis, Vol. 19, No. 4, pp. 23-25, Dec 1996.

[10] Courtney, R. and Rice, C.: 'Using an encounter form to develop a clinical database for documenting nurse practitioner primary care', Journal of the American Academy of Nurse Practitioners, Vol. 7, No. 11, pp. 537-544, Nov 1995.

[11] Hamlet, R., A Voice Recognition Solution for Nursing Documentation, http://www.scribd.com/doc/6333073/ A-Voice-Recognition-Solution-for-Nursing-documentation-Ryland-Hamlet Last accessed on February 28, 2011.

[12] The Java EE 5 Tutorial, Chapter 1, http://download.oracle.com/javaee/5/tutorial/doc/bnaaw.html Last accessed on February 08, 2011.

[13] Java SE 6 Key Features, http://www.oracle.com/technetwork/java/javase/overview/features-jsp-140508.html Last accessed on February 08, 2011.

[14] Apache Tomcat 6.0, http://tomcat.apache.org/tomcat-6.0-oc/index.html Last accessed on February 08, 2011.

[15] Welcome to Struts 1, http://struts.apache.org/1.x/ Last accessed on February 08, 2011.

[16] Welcome to Apache Axis2/Java, http://axis.apache.org/axis2/java/core/ Last accessed on February 08, 2011.

[17] Apache Lucene – Overview, http://lucene.apache.org/java/docs/index.html Last accessed on February 08, 2011.

[18] JAMA: A Java Matrix Package, http://math.nist.gov/javanumerics/jama/ Last accessed on February 08, 2011.

[19] http://dev.mysql.com/doc/refman/5.1/en/introduction.html Last accessed on February 08, 2011.

[20] Christensen, E., Curbera, F., Meredith, G. and Weerawarana, S., Web Services Description Language (WSDL) 1.1, http://www.w3.org/TR/wsdl Last accessed on February 08, 2011.

[21] JSON in JavaScript, http://www.json.org/js.html Last accessed on February 08, 2011.

[22] JAX-WS Guide, http://axis.apache.org/axis2/java/core/docs/jaxwsguide.html Last accessed on February 08, 2011.