Technical Perspectives on Knowledge Management System in Bioinformatics Environment

Nurul Haslina Hairudin, Rusli Abdullah, Rodziah Atan, Hamidah Ibrahim, Suhaimi Napis and Siti Hernazura Hamidon @ Md Jamil

Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, 43400 UPM Serdang
Tel: 03-84966575 Fax:-
Email: nurulhaslina@yahoo.co.uk

ABSTRACT

The knowledge growth in the Bioinformatics environment needs a system that can organize and manage the biological knowledge. Knowledge Management System is appropriately for acquiring, storing, applying and disseminating that biological knowledge. In order to develop the system, the technical perspectives of Knowledge Management should take into account. In this paper, we discuss the technical perspectives of Knowledge Management System and the suitable method and technologies to be implemented in the system.

Keywords

Knowledge Management, Bioinformatics, Knowledge Technology

1.0 INTRODUCTION

Knowledge Management (KM) is a concept of acquiring, storing, applying and disseminating knowledge among Community of Practice (CoP). KM can be viewed from three perspectives which are Techno-centric, Organizational, and Ecological.

1.1 Techno-centric: A focus on technology, ideally those that enhance knowledge sharing/growth.

1.2 Organizational: How does the organization need to be designed to facilitate knowledge processes? Which organizations work best with what processes?

1.3 Ecological: Seeing the interaction of people, identity, knowledge and environmental factors as a complex adaptive system.

Bioinformatics is a field that deals with research, development or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data (Haas, 2001) The rapid growth of bioinformatics environment not only provides the vast amount of biological data with heterogeneity of data types but also presents the big challenge in relation to integration and management of knowledge in order to supply the meaningful knowledge to the CoP in bioinformatics environment.

Everyday, emerge new findings that contribute to bioinformatics domain. These findings have to be shared among the CoP so that the biological knowledge can fully utilized and not been wasted. By using Knowledge Management System (KMS) in Bioinformatics, CoP can work collaboratively and construct a new finding that can be shared and manipulated in producing a high quality product. For instance, the herb plantation in Malaysia provides new findings that can benefit the human in body health aspects, medicine, spices and cosmetics. People do not have the exposure that these herbs products give many advantages especially for the long term health body aspects. This becomes the reason to use herbs plantation as the main content of the KMS in Bioinformatics (Siti, 2007)

However, there are lacks of research on technical perspectives instead of organizational perspectives for KMS in Bioinformatics environment. Technical perspectives on KMS in Bioinformatics provide an in depth overview of

the methodologies, technologies, product features and product performances.

The success of KMS in Bioinformatics can be achieved by the assistance of Knowledge Technology. Knowledge Technology is a part of KM, refer to an unclear set of tools including languages and software enabling better representation, organization and exchange of information and knowledge. Among the existence Knowledge Technologies are knowledge mapping, collaborative technologies, semantic technologies and social computing tools.

This paper will discuss the technical perspectives on KMS in Bioinformatics that focus on technology, ideally those that enhance knowledge sharing and growth in bioinformatics domain especially for herbs plantation in Malaysia

2.0 LITERATURE REVIEW

Based on Marwick (2001), selected technologies that contribute to KM solutions is reviewed using Nonaka's Model of organizational knowledge creation as a framework. Nonaka Model (Nonaka and Takeuchi, 1995) focuses on tacit knowledge, which is featured in three of the four processes and thus on people and their use of technology.

Table 1: Examples of technologies that can support or enhance the transformation of knowledge (Marwick, 2001)

	TACIT	EXPLICIT
TACIT	➢ E-meetings ➢ Synchronous collaboration (Chat)	➢ Answering questions ➢ Annotation
EXPLICIT	➢ Visualization ➢ Browsable video/audio of presentations	➢ Text search ➢ Document categorization

Knowledge portals are single-point-access software systems intended to provide easy and timely access to information and to support communities of knowledge workers who share common goal. Mack (2001) described the role knowledge portals play in supporting knowledge work tasks and the component technologies embedded in portals, such as the gathering of distributed document information, indexing and text search, and categorization

Knowledge mapping is a consciously designed communication medium using graphical presentation of text, stories, models, numbers or abstract symbols between map makers and map users. Knowledge maps are excellent ways to capture and share explicit knowledge in organizational contexts. (Wexler, 2001)

BioJava Project (www.biojava.org) is an open-source that provide Java Tools for processing biological data, including objects for manipulating sequences, file parsers, dynamic programming, and simple statistical routines(BioJava.Org,2003)
Bio-Soft project at the University of Wisconsin-Parkside (UWP) is intended to develop robust Web-based software to facilitate bioinformatics research at UWP. This project develops Java software to run on web application servers such as the IBM Websphere. (Chalasani, 2004)

Adak (2002) presented a comprehensive bioinformatics KM framework called e2e that provide an early glimpse of the wide potential of an integrated KM solution for bioinformatics with a specialized text-mining system called MedMesH summarizer. E2e facilitate the data integration for the bioinformatics by exposed a common semantic view of inter-relationship among biological concepts in the form of an Extensible Markup Language (XML) representation called eXpressML, while internally, it can use any data integration solution to retrieve data and return results corresponding to the semantic view.

XML is use to store and implement organizational data definitions, thus providing a synergetic framework for leveraging the potential of KM tools. (Otto, 2001) By the use of Data Type Definition (DTD), XML can allow exchange of knowledge across different database, organization and domain. These will facilitate the knowledge integrated environment. (Figure 1)

```
<?xml version="1.0"?>
<!DOCTYPE note [
  <!ELEMENT note    (to,from,heading,body)>
  <!ELEMENT to      (#PCDATA)>
  <!ELEMENT from    (#PCDATA)>
  <!ELEMENT heading (#PCDATA)>
  <!ELEMENT body    (#PCDATA)>
]>
<note>
<to>Biologists </to>
<from>Researchers</from>
<heading>Notification Alert</heading>
<body>Keep Update with the new findings!</body>
</note>
```

Figure 1: An Example of XML document with internal DTD

Davies (2003) applies ontologies to electronically available information to improve the quality of knowledge management in large and distributed organizations. Ontologies are formal theories supporting knowledge sharing and reuse. They can be used to explicitly represent semantics of semi-structured information. These enable sophisticated automatic support for acquiring, maintaining and accessing information.

3.0 METHODOLOGY

In order to build a KMS in Bioinformatics, several steps have been taken. The design, development, implement and assessment of KMS in Bioinformatics took place at selected bioinformatics organization in Malaysia.

First, the data collection has been done through questionnaires and interviews in order to obtain the experts' opinion in design and develop the KMS in Bioinformatics. Besides that, the data collection process also important to understand the issues pertaining to the use of the technologies in KM processes.

After analyze the data obtained, the model of KMS is constructed. Based on this model, we developed the system that can contribute to knowledge sharing and growth among CoP.

Then, after finishing the development phase, the system will be implemented in selected organization. The users will test the system on the functionality of the system. If there are some changes, the system will be refined. This process is iterating until the system is completely ready to be used by the CoP.

4.0 CASE STUDY OF KNOWLEDGE MANAGEMENT SYSTEM IN BIOINFORMATICS

A KMS in Bioinformatics has been developed to facilitate knowledge growth and sharing among CoP that involve in herb plantation R&D in Malaysia The system is still in the development phase and will be estimated to finish on June 2008. The importance of the research is to contribute several implications or benefits for the organization or institution that tend to be initiate and implement a KMS in Bioinformatics for herb plantation in Malaysia in achieving of mission statement

The system will support the collaboration among CoP to produce new findings in the herbs plantation R&D. With these new findings, people around the world can get benefit especially for their body health care.

5.0 RESULTS AND DISCUSSION

Based on data analysis, a model of KMS in Bioinformatics (Figure 2) is constructed. The model demonstrates the KM method involve and the appropriate technologies for the each of the method (Table 2). The involve methods in the KMS in Bioinformatics are Knowledge Acquisition, Knowledge Storage, Knowledge Dissemination and Knowledge Application.

Knowledge Acquisition is a process to determine the source and type of knowledge. Knowledge is categorized and mapped into specific requirement.

Knowledge Storage is a process where the knowledge is kept in the repositories. These can be in a form of categorized several types of documents.

Knowledge Dissemination is a process to distribute the knowledge to the user that request the knowledge.

In the Knowledge Application process, the user can utilized the knowledge in the system to solve problem, produce new findings and make decision.

Table 2: Technologies for Every KM Method for KMS in Bioinformatics.

KM Method	Technologies	Descriptions
Acquisition	Portal	The system will acquire the knowledge by using portal.
Storage	XML, MySQL	XML is used for integrated environment and MySQL is an object oriented database.
Dissemination	Java, Apache-Tomcat	Java is distributed and portable suitable for disseminating the knowledge and Apache-Tomcat is the suitable server that can work with Java technology.
Application	Java	The tool to apply the knowledge is developed by using Java.

This system also uses combination of Java and XML technology to make possible of integrated environment and accessible everywhere with the help of web technology. Java is a simple, object-oriented, distributed, interpreted, robust, secure, architecture neutral, portable, multithreaded, and dynamic.

The development of KMS in Bioinformatics portal (Figure3) demonstrates the role of technologies in order to acquire and disseminate knowledge among the CoP.

The use of MySQL database as the system repository, with the help of XML technology in the system is to store the knowledge in different types of documents.

XML is heavily used as a format for document storage and processing, both online and offline. The hierarchical structure is suitable for most types of documents. It is platform-independent, thus relatively immune to changes in technology. However, the distinction between content and attributes in XML lead to design of XML data structures harder. Besides that, XML is also inefficient to securely transfer large amount of data.

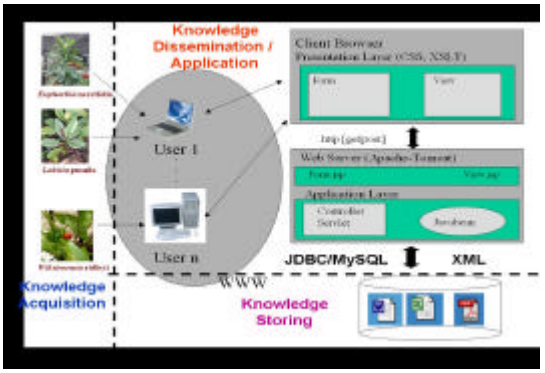

Figure 2: Model of Knowledge Management System in Bioinformatics With Collaborative Environment

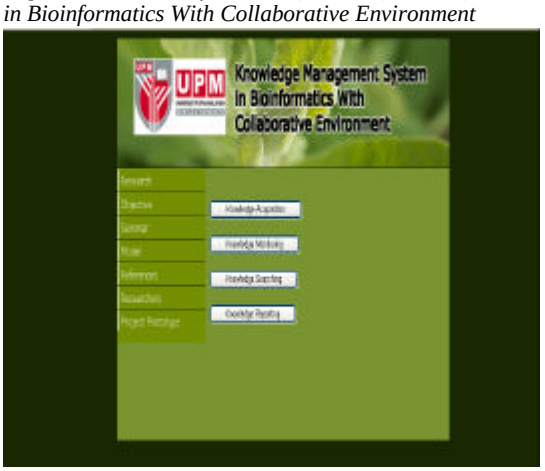

Figure 3: Knowledge Management System in Bioinformatics with Collaborative Environment Portal

Knowledge Mapping tool in the system will help the user of the system to map the document consists of certain knowledge to their category. For instances, the system user will key in the category of the knowledge that want to submit to the system. The category can be chose from existence categories or user construct a new category but with the system administrator permission. The permission in constructing a new category for the submitted knowledge is to avoid redundancies happened in the databases.

Besides that, Knowledge Ontology as part of semantic technology also one of technology to be implemented in the KMS in Bioinformatics with Collaborative Environment but currently, the Knowledge Mapping and Knowledge Ontology is still in development phase.

6.0 CONCLUSION

KM can be viewed in technical, organizational and ecological perspectives. People always concern about organizational perspectives instead of technical perspectives of KM. In this paper, addressed the knowledge methods and technologies consist in the KMS in Bioinformatics that involve in R&D processes especially for herb plantation in Malaysia. The contribution of those knowledge methods and technologies to bioinformatics domain also discussed. In order to develop a successful of KMS, the bioinformatics institution must ensure the proper requirement steps of implementation of KMS is ready and try to adapt as much as possible of the technologies that are suitable for.

REFERENCES

Adak, S., Batra, V., S., Bhardwaj, D., N., Kamesam, P., V., (2002) A System for Knowledge Management System in Bioinformatics. *CIKM'02*, Nov 4-9, 2002.

Chalasani, S., and Barber, R., (2004) Architectures for Java-based Bioinformatics Applications. *Industrial Management & Data System*, Volume 104, Number 7, 2004.

Davies, J., Fensel, D., and Harmelen F., V., (2003) *Towards the Semantic Web: Ontology-driven Knowledge Management* ISBN 0470848677

Haas, L., Schwarz, P., Kodali, P., Kotlar, E., Rice, J., and Swope, W., (2001). DiscoveryLink: A system for integrated access to life sciences data sources. *IBM Systems Journal*, Volume 40, Number 2, 2001.

Mack, R., Ravin, Y., and Byrd, R., J., (2001) Knowledge Portals and The Emerging Digital Knowledge Workplace. *IBM Systems Journal*, Volume 40, Number 4, 2001

Marwick, A., D., (2001) Knowledge Management Technology. *IBM Systems Journal*, Volume 40, Number 4, 2001.

Nonaka, I., and Takeuchi, H., The Knowledge Creating Company, *Oxford University Press*, Oxford, UK (1995).

Otto, J., R., Cook, J., H., and Chung, Q., B., (2001) Extensible Markup Language and Knowledge Management. *Journal of Knowledge Management*, Volume 5, Number 3, 2001.

Siti H., H., Rusli, A., Hamidah, I., Rodziah, A., Suhaimi, N., and Nurul H., H., (2007) BioKnowledgeWebs: A Knowledge Portal for Bioinformatics Environment. *3rd Malaysian Software Engineering Conference'07*, Dec 3-4, 2007.

Wexler, M., N.,(2001) The Who, What and Why of Knowledge Mapping. *Journal of Knowledge Management*, Volume 5, Number 3, 2001.