International Journal of Advances in Computer Science & Its Applications– IJCSIA
Volume 5: Issue 1 [ISSN : 2250-3765]

Publication Date : 30 April, 2015

Collection of Web Multimedia Teaching Resources Using Ontology

[Mohamed Alhawiti, Yasser Abdelhamid]

***Abstract*— The World Wide Web has a huge amount of e-learning resources in the various branches of science, and it is available in the form of textbooks, presentations, video tutorials, pictures and audio lectures. No doubt that these resources would help students understand academic courses easily, especially those courses that need training and practical activities, like the computer science courses. It also helps the instructor to clarify his ideas in an interesting and innovative way.**

Searching for the available and suitable resources in the World Wide Web is perhaps one of the most difficult and time consuming tasks, as it requires the exact specification of keywords that characterize each topic in the course syllabus. Collecting such material manually from scratch for each course in a specific domain of knowledge is considered expensive and time consuming effort.

Ontology is the explicit formal specifications of the terms and relations among them in a specific domain. It defines a common vocabulary for researchers who need to share information in a specific domain. It is perhaps the key solution to the problems related to the knowledge sharing and reuse, as it includes machine-interpretable definitions of basic concepts in that domain and relations among them.

In this paper, a system is proposed to enable instructors to collect e-learning multimedia resources from the World Wide Web and link them automatically to the topics listed in the syllabus of the intended course using the Ontology of the domain of knowledge related to the that course.

***Keywords*— Automatic Retrieval of Educational Multimedia Resources, Information Retrieval, Knowledge Sharing and Reuse, Ontology-Based Annotation.**

I. Introduction

The process of creating a course, which includes identifying the conceptual map, creating appropriate text, all the required multimedia object materials, and setting up the tools for evaluating student performance, is a very tedious and time consuming process. Anyone who has to create learning materials from scratch knows just how labor intensive and time consuming the process can be [1]. Course creation is, hence, an expensive affair in terms of the amount of time and the expert knowledge required by the author/s.

Mohamed M. Alhawiti
Community College, University of Tabuk
Saudi Arabia

Yasser Abdelhamid
Community College, University of Tabuk
Saudi Arabia

Reusing suitable text and multimedia objects from the existing courses can considerably reduce the time and effort required for creating new courses. Additionally, if the author has the option of considering relevant material that exists across Internet documents and other sources to form course content it will extend the source material pool. The author can then choose the appropriate material from this relevant extended collection, thereby enhancing the quality of the course content.

By the introduction of Web 2.0 features like Blogs, Wikis, Social networking, and Web applications, every user of the web became an author, having all the tools and facilities to publish and share his creations on the web [2]. The result is a huge amount of resources and multimedia content available on the web.

To overcome the process of accessing all these proliferations of information on the web, and to automate the process of accessing these tons of resources, imposing semantic abstraction of information is needed. The main barrier to information automatic access is that all existing information is represented freely by different information providers and that concepts in the same domain are very often expressed using different methods. The consequence is that the semantics of information are not understandable for search engines and that knowledge cannot be shared between data sources [3].

Ontology is the explicit formal specifications of the terms in a specific domain and relations among them [4]. It defines a common vocabulary for researchers who need to share information in a specific domain and includes machine-interpretable definitions of basic concepts in that domain and relations among them. Ontology is becoming of increasing importance to a large number of application types like knowledge-based systems, information exchange, semantic web, and application integration. Many disciplines have developed their own standard ontologies that domain experts can use to share and annotate information in their fields. SNOMED [5] is an example of a large structured and standardized vocabulary in the domain of medicine. UMLS [6] is a semantic network of Unified Medical Language System in medicine. UNSPSC [7] is also an ontology which provides common terminology for products and services.

In this paper, a proposed system is presented for assisting course authors to automatically collect course multimedia learning materials out of the available related web resources using ontology.

The paper is organized as follows: Section 2 provides an overview of the used technologies like Ontology, Ontology tagging, and Conceptual mapping. In section 3, the design of the proposed framework is presented, and the implementation is specified. Section 4 demonstrates a sample run of the

implemented prototype and discusses the results, and finally, conclusions and ideas for future work are discussed.

II. Background

The World Wide Web has extremely affected the way of information exchange. Users of the web can electronically publish their ideas and innovations. This produced a massive amount of documents and resources, and this amount is rapidly growing. This exponential increase of information makes it hard to directly find useful information from Web sources [3].

To solve the above problems, information should be communicated between applications and software agents at both the syntactic level and the semantic level. XML has proved its sufficiency for supporting information communication at the syntactic level as it allows application developers to describe different information elements by using domain specific terms [22].

Ontology is the candidate technology to describe the semantics of the underlying information. It defines the kinds of things that exist in an application domain. In the computing context, an ontology is a framework for representing concepts (things, or ideas about things) and the relationships that exist between those concepts [8].

A. Ontology

The most common definition of ontology is "a formal, explicit specification of a shared conceptualization" [4]. Where conceptualization, refers to building an abstract model of the concepts and relations that people use in a specific domain of knowledge like education, agriculture, medicine, etc.

An explicit specification means that the names of concepts and relationships in a specific domain are assigned explicit names in addition to a definition that describes the meaning of these concepts and relations.

Formal means that the concepts and relations are encoded in a formal language, that guarantees removing ambiguity, and assures understandability by both human being and machine, usually logic-based languages are used for this purpose, in addition to other representations drawn from the community of Artificial Intelligence.

Shared means that the ontology is to be used and reused by different applications and software agents.

Software applications integration is one of the common usages of ontology, where an enterprise authors a common ontology to be used for integrating its software applications.

Ontology also can be used as a common interchange format to translate from/to different software applications with different formats.

Another usage of ontology is "Ontology-Driven Software Engineering", where ontology of a given domain is created and used as a basis for specification and development of some software.

Finally, Ontology-based Search is used to facilitate search, where an ontology is used for indexing information repository [9].

B. Applications of Ontology in the Educational Domain

Ontology is considered to be a highly suitable means of supporting educational-technology systems [10]-[12]. The increasing importance of the Semantic Web, which is based on ontology, is strengthening this argument [13]. It has a promising role in the field of instructional design and the development of course content, as it can represent knowledge about content, supports course authors in creating content and provides easy accessibility of course content by students. Hence, it is strongly expected that Ontology might be useful in the domain of education [14].

The following is survey on selected applications of ontology in the domain of education.

1) Web-Based Educational Systems (WBES)

A new knowledge communication protocol which implements content awareness by ontology technology is proposed in order to make intelligent communication individuals in different area can understand the information between them, and by the new communication model built on the high layer, intensifies the understanding of communication content, optimizes the quality of network communication [15].

2) Ontologies for Education (O4E)

Consists of concepts like technological perspective, application perspective. Technological perspective defines the knowledge organization, knowledge inference, information, information visualization, information navigation, information querying, subject domain ontology, instructional knowledge. Application perspective defines sub concepts in knowledge construction, knowledge externalization, knowledge communication, architectural knowledge [16].

3) Ontologies for the Use of digital learning Resources and semantic Annotations on Line (OURAL)

This research project defines ontology in the e-learning domain which includes problem-situation, problem solving, critical analysis, case study, debate, cyber quest, project, exercise etc. [17].

4) Ontologies for E-Learning Systems in Higher Education

A Semantic Web that defines educational ontology consisting of user profile ontology, the person ontology, the contact ontology and the activities ontology. This was implemented using Protégé-OWL ontology editor [18].

5) Information and Communication Technologies (ICT) in Education Ontology

Provides a central repository of classified knowledge of ICT in the education domain. ICT ontology consists of concepts like Ontology of ICT curriculum, ontology of ICT Job, ontology of ICT Skill and ontology of ICT Research [19].

6) European Credit Vocational System

Uses ontologies in the construction of the educational resources library, to provide a common access to the information regarding the qualification systems of nine European countries [20].

C. Tagging of Resources

Tagging is the process of adding special annotations or marks that attach a piece of information to a resource or an object for future referral. There are many purposes for tagging objects, users usually tag objects for the purpose of future retrieval of the objects by themselves or by others. They also use tags for sharing their ideas about objects with other users, or to attract attention to specific resources, or to put self-referential tags to mark their contributions, or to put an evaluation for a specific object [21].

III. Architecture of the Proposed System

The main goal of the proposed system is to assist the instructor or the course designer to search for the related multimedia learning material on the web, and attach them automatically to each of the topics that comprise the syllabus of the intended course. This section provides an overview of the architecture of the proposed system, and describes the function of each of its components.

The input to the proposed system is the suggested syllabus provided by the instructor, which specifies the contents of the intended course and the topics included in that course. The output is an index of the syllabus topics with links to multimedia resources related to each topic in that syllabus with the facility to browse the course contents along with the attached multimedia resources.

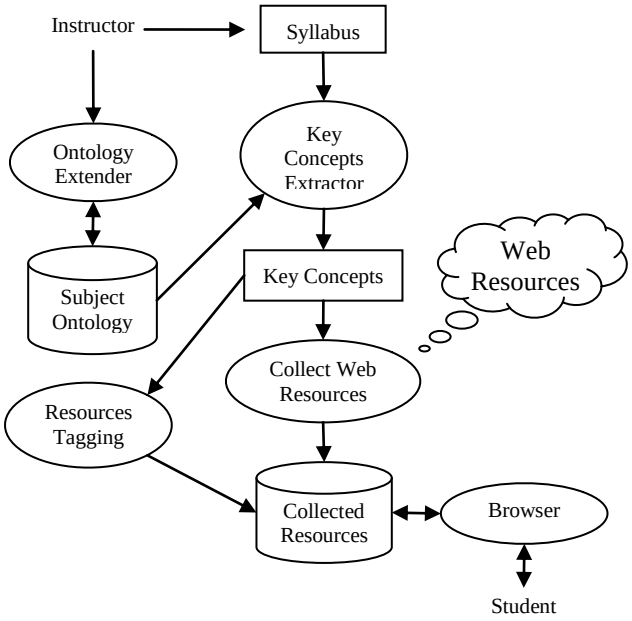
As it appears in Figure 1, the proposed system consists of five main components: Key concepts extractor, Web resources collector, Resource tagging component, Ontology extender, and Subject Web-resources browser.

D. Key Concepts Extractor

The main function of this component is to analyze the topics of the input syllabus to determine the key concepts that characterize each topic. The concepts are drawn from the domain ontology provided by the instructor and related to the domain of knowledge of the intended course. The output of this component is an index of the topics that constitute the syllabus of the course, with the relevant domain concepts related to each topic. The known techniques of generalization and specialization are used to explore the taxonomy of the domain ontology concepts to specify the related domain concepts to each topic.

E. Web Resources Collector

The web resources collector searches the web for the multimedia resources related to each topic using the key concepts that characterize these topics. The used keywords are extracted from the topics domain concepts index generated in the previous stage. The output of this component is a repository of links to web resources related to each topic.



1. Architecture of the proposed system

F. Resource Tagging Component

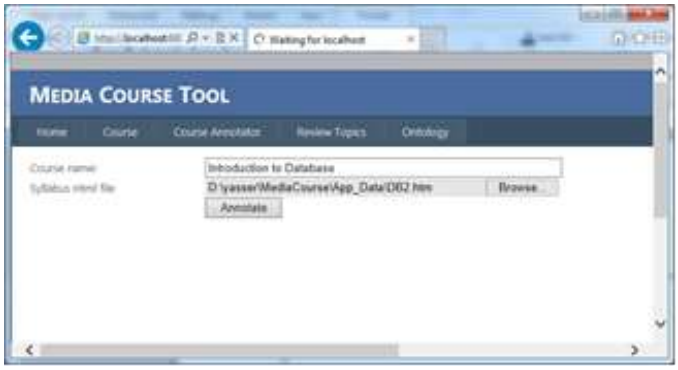
This component analyzes the collected resources and assigns to each of them tags from the key domain concepts that have been prepared before.

G. Ontology Extender

The instructor uses this component to build the hierarchy of the subject ontology. This tool allows the instructor to put the main concepts related to the subject, and to put relations between them, so that these concepts are arranged in a hierarchy according to each type of relation.

H. Subject Web-resources browser

The target of the proposed system is to provide the facility of browsing web resources related to each topic of the syllabus of a selected subject. This component provides this facility. Through this component the user/student will be able to go through the collected web resources of a selected topic.

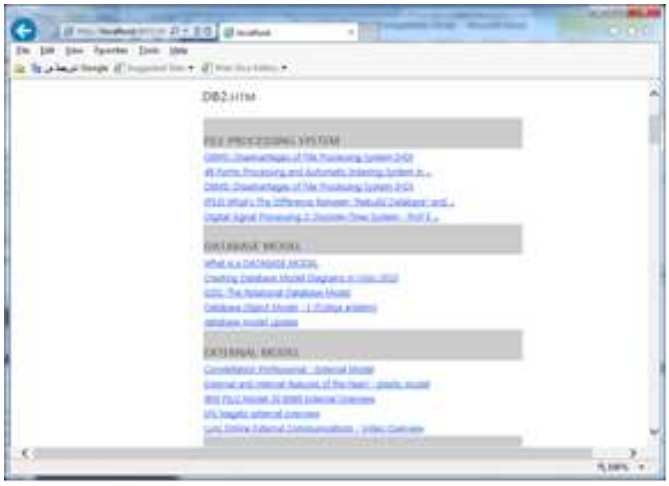


2. Course Syllabus selection form.

IV. Sample Run

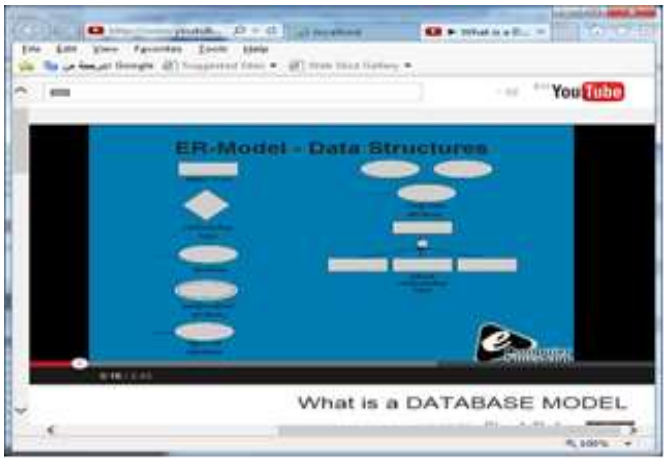
This section demonstrates a sample run of the implemented prototype of the proposed system. The input to the course annotator module is the syllabus of the intended course as an HTML file as instructors usually publish the syllabus of their courses on their web pages in HTML format. Figure 2, shows the form that the instructor uses for selecting the syllabus file.

The output of the proposed tool is a list of the topics found in the input syllabus file associated with links to the multimedia resources found on the web and related to each topic as shown in Figure 3.



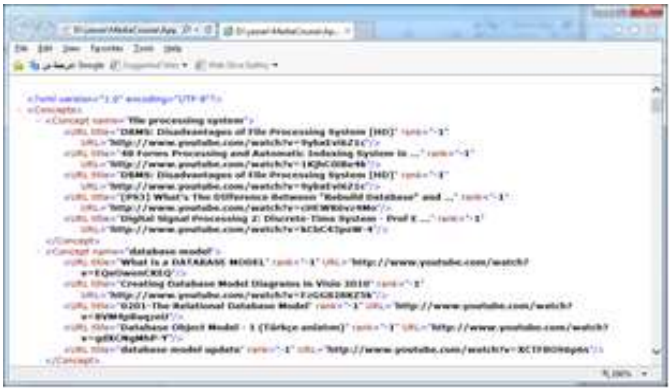
3. Syllabus topics with associated web multimedia resources.

Upon selecting one of the displayed links, the multimedia content of the selected link is displayed as shown in Figure 4.



4. Displaying selected multimedia content.

The internal representation of the output XML file that contains the syllabus and its topics along with the associated web multimedia resources is shown in Figure 5.



5. The generated XML output.

Figure 6, is a pseudo code that describes the process of reading the syllabus in an HTML format, reading the ontology XML file, and locating domain concepts in the syllabus, then extending them using the relations in the domain ontology, and finally calling the "Search-Annotate" module that both searches the web for multimedia resources using the set of generated keywords, and links the topics found in the syllabus file with links to the relevant web multimedia resources. Finally the annotated syllabus file is added to the database of annotated syllabus to be available later to the students to browse them, or the instructor to manipulate them using the course browsing utility.

```
Input: SHF: Syllabus HTML file
       OXF: Ontology XML file
       T: number of top URL
Output: Annotated syllabus
Steps:
        Get SHF file name
        Upload SHF
        Get LR  // List of all relations which exist in OXF Ontology
        Get LC //List of all concepts  in OXF ontology
        for each concept C exist in LC
                Append to LC all concepts related to C
        MC = Get Keywords MC; // set of OXF concepts that exist in
        SHF
        SearchAnnotate MC with top T of multimedia URLs related to
        C
        Add course name in courses.xml
```

6. Course annotation algorithm

```
Input: SHF: Syllabus HTML file
       KeywordL: keywords List
       T: number of top URL
Output: Annotated syllabus
Steps:
        For each Con in KeywordL
            Get M_URL media file for Con
            ScoreM = Score(M_URL.title,C.Line)
            If not exist M_URL
                    MediaList Add (M_URL,ScoreM)
            Get MediaList
            Annotate Con with top T in MediaList
        Save syllabus xml
```

7. Search-Annotate algorithm

The "Search-Annotate" algorithm described in Figure 7, uses the generated list of keywords that characterize each topic of the input syllabus to search for multimedia resources on the web, statistical approach is adopted to select the most matching resources, with the keywords of a topic. Only the top ranked resources are selected.

V. Discussion of Confronted Problems

During the development of the prototype of the proposed system a number of problems were revealed. This section discusses these problems along with the suggested solutions.

I. Misleading description of multimedia resources

YouTube is selected as the source of video feeds using Google API to narrow the scope of the multimedia resources. The title, and description of the feeds are selected for matching course topics ontology terms. But the initial results revealed that some of the video feeds have a description that does not match the content of the feed.

The proposed solution of this problem is to use matching at different levels of the Ontology hierarchy, and apply matching at more specific terms down the hierarchy of Ontology terms. The second solution is to restrict search for feeds in only authentic sites in YouTube, those related to the domain of the course subject feeds.

B. Replication of the same content with different descriptions

Feeds are sometimes repeated by different sites with different descriptions. This problem causes lengthy results of matching resources, and can be filtered using other parameters of the feed like type, length, etc.

VI. Conclusion

Ontology is becoming of a great importance in the field of information technology, it is the key for integrating applications, information sharing and information association. This paper introduced an approach for associating academic courses syllabus with the available multimedia resources related to its topics on the web. The proposed system has been designed and implemented to read the syllabus provided by the instructor and marking its topics with the domain concepts drawn from the domain ontology, then using these concepts for searching for multimedia resources on the web, and attaching them with each topic of the syllabus.

The most important result is the possibility of sharing and reusing domain ontology, as it can be used for annotating many courses in the same domain of knowledge, in addition to expediting the process of building the course content, as the proposed system collects the multimedia content of the intended course from the web automatically, reducing the burden of searching and collecting resources individually for each topic.

For the students, the proposed system allows them to browse the multimedia resources of the course indexed with the topics easily. The proposed tool also provided some amendment services, so that the instructor can edit the ontology used for a specific domain, and can also edit the resources allocated to each topic of the syllabus.

A prototype of the proposed Multimedia-Enabled Syllabus Browser has been implemented, and a sample run has been introduced that clarifies the idea of the proposed system. Future research will continue in providing alternative solutions to the confronted problems.

Acknowledgment

The contributors of this research work would like to express their gratitude to the deanship of scientific research in the University of Tabuk for their help and kind support.

References

1. Casey, and McAlpine M. "Writing and using reusable ducational materials: a beginner's guide." CETIS Educational Content Special Interest Group Publication, 2002.
2. Carlos Bobed and Guillermo Esteban and Eduardo Mena, "Ontology-driven Keyword-based Search on Linked Data", Proc. of the 16th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2012), San Sebastián (Spain), IOS Press, ISBN 978-1-61499-104-5, volume 243, pp. 1899-1908, Frontiers in Artificial Intelligence and Applications, September 2012.
3. Gardarin, G.; Kou, H.; Zeitouni, K.; Meng, X. & Wang, H., SEWISE: An Ontology-based Web Information Search Engine., in Antje Düsterhöft & Bernhard Thalheim, ed., 'NLDB' , GI, , pp. 106-119, 2003.

4. Gruber, T.R. A Translation Approach to Portable Ontology Specification.Knowledge Acquisition 5: 199-220, 1993.
5. Price, C. and Spackman, K. SNOMED clinical terms. BJHC&IM-British Journal of Healthcare Computing & Information Management 17(3): 27-31, 2000.
6. Humphreys, B.L. and Lindberg, D.A.B. The UMLS project: making the conceptual connection between users and the information they need. Bulletin of the Medical Library Association 81(2): 170, 1993.
7. http://www.unspsc.org, web, Oct., 2013.
8. Uschold, M., & Gruninger, M. , Ontologies: principles, methods, and applications. Knowledge Engineering Review, 11(2), 1-63, 1996.
9. M. Uschold and M. Gru¨ninger.Ontologies and semantics for seamless connectivity.SIGMOD Record, 33(3), 2004.
10. Mizoguchi, R. & Bourdeau, J , Using ontological engineering to overcome common AI-ED problems. International Journal of Artificial Intelligence in Education, 11(2), 107-121, 2000.
11. Sampson, D. G., Lytras, M. D., Wagner, G., & Diaz, P. , Ontologies and the Semantic Web for E-learning. Journal of Educational Technology and Society, 7(4), 26-142, 2004.
12. Aroyo, L., Dicheva, D., & Cristea, A. , Ontological Support for Web Courseware Authoring. In Cerri, S.A., Gouarderes, G. & Paraguacu, F. (Eds.), Proceedings of Intelligent Tutoring Systems, Berlin: Springer, 270-280, 2002.
13. Berners-Lee, T., Hendler, J., & Lassila, O. , The Semantic Web. Scientific American, 284 (5), 34-43, 2001.
14. Boyce, S., & Pahl, C. , Developing Domain Ontologies for Course Content. Educational Technology & Society, 10 (3), 275-288, 2007.
15. Darina Dicheva. and Lora Aroyo, A Minimalist Approach to Support Ontology-driven Educational Information Systems Communication. In Proceedings of International Workshop on Applications of Semantic, 2004.
16. Lora Aroyo. and Darina Dicheva.. "The New Challenges for E-learning: The Educational Semantic Web". Journal of Educational Technology and Society, 59-69, 2004.
17. Monique Grandbastien., Faiçal Azouaou., Cyrille Desmoulins., Richard Faerber., Dominique Leclet., and Céline Quénu-Joiron, Sharing an ontology in Education: Lessons learnt from the OURAL project. In Proceedings of 7th IEEE International Conference on Advanced Learning Technologies, 2007.
18. Marian Bucos., Bogdan Dragulescu., and Marius Veltan, Designing a Semantic Web Ontology for Elearning in Higher Education. In Proceedings of IEEE 9th International Symposium on Electronics and Telecommunications, 415 – 418, 2010.
19. Kum Leng Chin. and Elizabeth Chang. A Sustainable ICT Education Ontology. In Proceedings of 5th IEEE International Conference on Digital Ecosystems and Technologies, Korea, 350-354, 2011.
20. Feng Yang. and Youquan Chen. Ontology based application framework for Network Education Resources Library. In Proceedings of 2nd International Workshop on Education Technology and Computer Science,423 – 426, 2010.
21. Gupta, M., Li, R., Yin, Z., and Han, J. Survey on social tagging techniques. SIGKDD Explor. Newsl. 12, 58–72, 2010.
22. XML.org web site. http://www.xml.org, 2003.

Mohammed M. Alhawiti was born in Tabuk, Saudi Arabia in 1983. Alhawiti earned his Bachelor degree in mathematics education in 2004 from Tabuk Teachers' College (currently Education & Arts College), Tabuk, Saudi Arabia. He earned his Master of Science degree and Ph.D. in educational technology from Indiana State University, Terre Haute, Indiana, USA in 2009 and 2011 respectively.
He worked as a math teacher between 2004 and 2005. Between 2005 and 2006, he worked as an instructor at the Educational Technology Department in Tabuk Teachers' College. Currently, he works as an assistant professor at the Education & Arts College and the Dean of Tabuk Community College, University of Tabuk, Tabuk, Saudi Arabia.
Dr. Alhawiti is a member of several honor societies including Phi Lambda Theta (International Honor Society and Professional Association in Education), the Golden Key International Honor Society, and the Honor Society of Phi Kappa Phi. He is also a member of the Association for Advancement of Computing in Education (AACE). Dr. Alhawiti is the chairman of the Permanent Committee of Distance Education in the University of Tabuk. He is also a member of several committees at the university level including Tabuk University Advisory Committee, the Permanent Committee for Developing University of Tabuk's Portal, and Tabuk University Council.

Yasser Abdelhamid was born in Giza, Egypt in 1962. Abdelhamid earned his Higher Diploma degree in Computer Science in 1986 from Institute of Statistical Studies and Research (ISSR), Cairo University, Cairo, Egypt. He earned his Master of Science degree and Ph.D. in computer science from ISSR, Cairo University, Cairo, Egypt in 1992, 1998 respectively.
He worked as a computer programmer, system analyst, and IT manager during the period from 1986 to 1994. He worked as a research assistant at the Central Lab for Agricultural Expert Systems (CLAES) during the period from 1994 to 1998, and as a researcher at (CLAES) during the period from 1998 to 1999. During the period from 1999 to 2002 he worked as an assistant professor at (ISSR), Cairo University,Egypt, then as an assistant professor at the Community College, Computer Science department King Abdualaziz University, Tabuk, Saudi Arabia during the period from 2002 to 2006, currently he is working as an associate professor at the Community College, Department of computer science, University of Tabuk, Tabuk, Saudi Arabia.
Dr. Abdelhamid has many publications in the domain of artificial intelligence and its applications in the domain of agriculture and education. And he is the chairman of computer science department in Community College, University of Tabuk.