Collection of Web Multimedia Teaching Resources Using Ontology

Mohamed Alhawiti and 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.

Index Terms—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.

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 II provides an overview of the used technologies like Ontology, Ontology tagging, and Conceptual mapping. In Section III, the design of the proposed framework is presented. In Section IV, the implementation aspects of the proposed system are specified. The results of the implemented prototype are evaluated and discussed in Section V, and finally, conclusions and ideas for future work are discussed in Section VI.

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 [8].

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 [9].

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 [10].

B. Applications of Ontology in the Educational Domain

Ontology is considered to be a highly suitable means of supporting educational-technology systems [11]-[13]. The increasing importance of the Semantic Web, which is based on ontology, is strengthening this argument [14]. 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 [15].

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 [16].

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 [17].

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. [18].

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 [19].

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 [20].

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 [21].

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 [22]

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 Fig. 1, the proposed system consists of four main components: Ontology extender, Topics tagging component, Web resources collector, and Subject Web-resources browser.
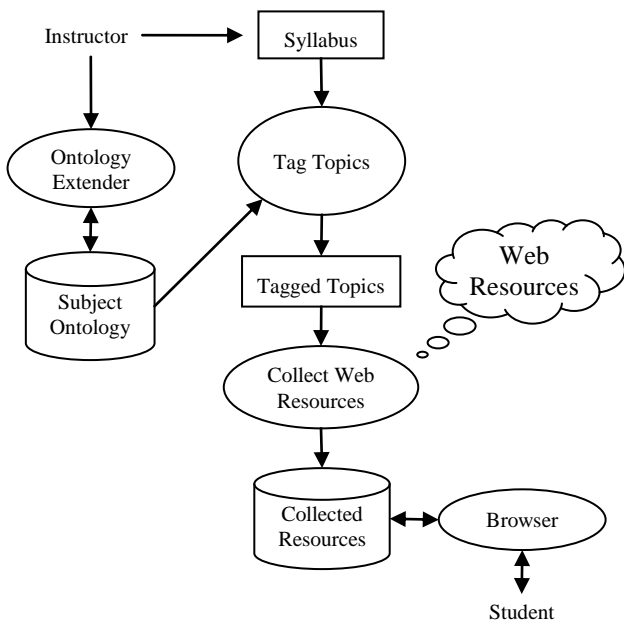

Fig. 1. Architecture of the proposed system.

A. Ontology Extender

Ontology is a major component of domain knowledge, and knowledge is dynamic by its nature as research and innovations expand domain knowledge and introduce new concepts and new relations to the domain ontology.

It is practically important to provide a facility for domain experts, and those who maintain ontology with a tool to add or modify the contents of the domain ontology.

The ontology extender provides this function, and helps the domain experts and instructors to extend the current concepts and relations with new instances.

B. Topics Tagging Component

A taxonomy is a hierarchical structure used for classifying or categorizing concepts in a specific domain. The most general category is at the root of the hierarchy. Each node, including the root node represents some object in the real world that is being modeled. Links represent relations between nodes as a "sub-class-of" relation or a "super-class-of" relation.

The function of the topics tagging component is to use the concept nodes of the taxonomy to assign tags that describe

the content of the topics. The generated tags are to be used later by the web resources collector as search keys for each topic.

The output of this component is an index of the topics that constitute the syllabus of the course, attached with the relevant generated tags.

C. Web Resources Collector

The web resources collector searches the web for the multimedia resources related to each topic using the attached tags. The used keywords are extracted from the index generated in the previous stage. The output of this component is a repository of links to web resources related to each topic.

D. Subject Web-Resources Browser

This component provides the facility of browsing and viewing the collected web resources related to each topic of the syllabus of a selected subject.

IV. IMPLEMENTATION OF THE PROPOSED SYSTEM

This section discusses the implementation aspects of the proposed system. The main modules to be discussed are the Topics Tagging Module, and the Web Resources Collector module.

A. Topics Tagging Module

The input to the topics tagging module is the syllabus of the intended course. Usually the syllabus include information about the course including objectives, outcomes, time table, and the most important part which is the course plan that lists the topics to be taught throughout the course. Fig. 2 displays a sample database course syllabus.

```
<?xml version="1.0"?>
- <Syllabus>
    + <Chapter>
    - <Chapter>
        <Title>DATA MODEL</Title>
        <Topic>Business Rules</Topic>
        <Topic>Evolution of Data Model</Topic>
        <Topic>Hierarchy model</Topic>
        <Topic>Network model</Topic>
        <Topic>Relational model</Topic>
    </Chapter>
    - <Chapter>
        <Title>ENTITY RELATIONSHIP (ER) MODEL</Title>
        <Topic>Entity, Attributes, Relationship</Topic>
        <Topic>Types of attributes, null values and keys</Topic>
        <Topic>Connectivity, Cardinality</Topic>
        <Topic>Relationship Strength</Topic>
    </Chapter>
    + <Chapter>
</Syllabus>
```

Fig. 2. Sample database course syllabus.

Topics tagging module traverses the domain concepts taxonomy related to the subject of the course, and searches for each concept in the list of topics of the input course. Fig. 3, is an excerpt of the used ontology concepts hierarchy, where concepts are represented by ovals and relations are represented by arcs. Synonym relation indicates that two concepts have the same meaning and can be used interchangeably. For example, "ER Model" is a synonym concept to "Entity Relationship Model" concept.

Child relation indicates that a concept is more general than its child. For example, "Entity" concept is a more general concept of "Weak Entity" concept.

Fig. 4 is a pseudo code that describes the process of

tagging syllabus topics with domain concept tags. It starts with reading the syllabus and ontology xml files, then for each concept in the ontology concepts list, it searches for this concept in each topic in the syllabus, if a concept is found in a topic, a tag record is appended to the TopicTags XML file.
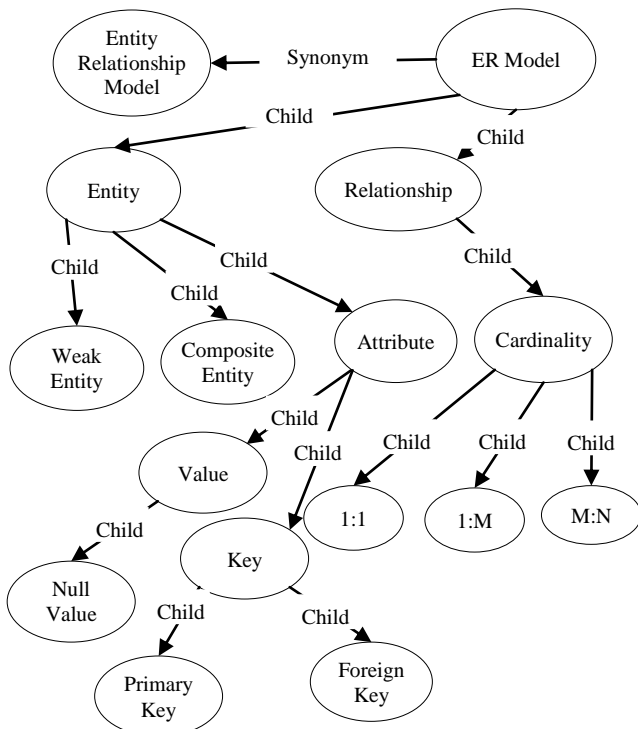

Fig. 3. Sample of the concepts hierarchy of the database domain.

```
Input:  S: Syllabus XML file
        O: Ontology XML file
Output: TopicTags.XML
Steps:
            Read S
            Read O
            Get LC //List of all concepts  in O
            Get LT // List of all topics in S
            for each concept C in LC {
                for each T in LT {
                    if ( C in T )
                        Append C to T TagList in TopicTags.XML
                }
            }
```
Fig. 4. Topic tagging pseudo code.

Fig. 5. Displays a sample of the output TopicTags.XML file.

Fig. 5. Sample output of the topics tagging module.

B. Web Resources Collector module

The web resources collector module searches for tags in TopicTags.XML file and generates a search key for each tag, the search key is the full path of concept tag as indicated in the domain concepts hierarchy. For example, if the tag is "Attribute" then the search key will be "ER Model Entity Attribute". The output of the web resources collector module is the list of multimedia resources attached to each topic after removing redundant occurrences of the search results.

Fig. 6 is a pseudo code of the web resources collector module.

```
Input:  TT: TopicTags.XML
        O: Ontology XML file
Output: TopicLinks.HTML
Steps:
            Read TT
            Read O
            for each T in LTopics {
                Get LTG //List of tags of T
                for each TG in LTG{
                    Get P = Path (TG) // Full path of the tag concept
                    Get SK = SearchKey(P)
                    Get SR = WebSearch(SK) // Search Result
                    Append SR to TopicLinks.HTML
                }
            }
```
Fig. 6. Web resources collector pseudo code.

V. RESULTS AND DISCUSSIONS

This section discusses the feasibility of using ontology for searching for web multimedia resources related to the topics of the selected syllabus.

We have two points to test: the first is the accuracy of syllabus topics tagging process, the second is the effect of using ontology in the search process.

A. Tagging Test

The approach used for testing the process of assigning tags to the topics of the syllabus is to analyze type I and type II errors. Type I errors or unmatched true results happen when a concept tag is not included in the tag list of a topic although it is in the description text of the topic. Type II errors or matched false results happen when a concept tag is included in the tag list of a topic although it is actually not matching the topic.

To narrow the scope of the analysis, the test is done on a subset of the whole syllabus that is the "Entity Relationship Model" which is a common part in all the database course syllabi found on the web. Three syllabus documents were tested and the results were as shown in Fig. 7.

Syllabus	Correct Matching	Type I errors	Type II errors
Syllabus 1	12	5	2
Syllabus 2	3	4	0
Syllabus 3	4	2	0

Fig. 7. Type 1 and type 2 errors test.

By analyzing type 1 errors, it was found that there are two reasons for these errors: the first is the incompleteness of the used ontology, the second is the need for some linguistic processing for the terms to match for example singular and plural terms.

Type 2 errors occurred because of the existence of the

same term in two different places in the ontology hierarchy with different meanings, and this should be managed while building the ontology.

B. Feasibility Test

Text-based search means using the topic description text directly as a search key to look for the matching multimedia resources, which is the currently used method for searching the web, while ontology-based search means using the domain ontology in formulating the search key.

The approach used for testing the feasibility of using ontology to formulate search keys instead of direct text-based search is to compare the number of correct/incorrect results for both types of search for the first 20 search results as shown in Fig. 8.

Topic	Text-based Search		Ontology-based Search	
	Correct	Incorrect	Correct	Incorrect
Entity, Attributes, Relationship	20	0	20	0
Types of attributes, null values and keys	16	4	20	0
Connectivity, Cardinality	1	19	20	0
Relationship Strength	0	20	20	0
Weak Entities and Composite Entity	18	2	20	0

Fig. 8. Feasibility test.

The results reveal that text-based search often get incorrect results, as it depends only on the terms found in the topic text, without having background knowledge about the context of the topic. For example, ss it appears in Fig. 8, the search for the topic "Relationship Strength" resulted in 0% correct results, as all the results were related to human relationships, instead of entity relationship model.

On the other hand, ontology-based search results were consistently correct, because it augments the context of the topic in the search key using the hierarchy of relevant domain concepts.

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.

Future research in this topic will resume finding solutions for filtering and grouping results, search for multiple types of resources, and enhancing the interface of the system.

REFERENCES

[1] Casey and M. McAlpine, "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," in *Proc. the 16th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2012)*, San Sebastián Spain, IOS Press, vol. 243, pp. 1899-1908, Frontiers in Artificial Intelligence and Applications, September 2012.

[3] G. Gardarin *et al*., "SEWISE: An ontology-based web information search engine," in A. Düsterhöft and B. Thalheim, ed., *NLDB, GI*, pp. 106-119, 2003.

[4] T. R. Gruber, "A translation approach to portable ontology specification," *Knowledge Acquisition*, vol. 5, pp. 199-220, 1993.

[5] C. Price and K. Spackman, "SNOMED clinical terms," *BJHC&IM-British Journal of Healthcare Computing & Information Management*, vol. 17, no. 3, pp. 27-31, 2000.

[6] B. L. Humphreys and D. A. B. Lindberg, "The UMLS project: Making the conceptual connection between users and the information they need," *Bulletin of the Medical Library Association*, vol. 81, no. 2, p. 170, 1993.

[7] Unspsc. (Oct. 2013). [Online]. Available: http://www.unspsc.org, web

[8] XML.org web site. (2003). [Online]. Available: http://www.xml.org

[9] M. Uschold and M. Gruninger, "Ontologies: Principles, methods, and applications," *Knowledge Engineering Review*, vol. 11, no. 2, pp. 1-63, 1996.

[10] M. Uschold and M. Grüninger, "Ontologies and semantics for seamless connectivity," *SIGMOD Record*, vol. 33, no. 3, 2004.

[11] R. Mizoguchi and J. Bourdeau, "Using ontological engineering to overcome common AI-ED problems," *International Journal of Artificial Intelligence in Education,* vol. 11, no. 2, pp. 107-121, 2000.

[12] D. G. Sampson, M. D. Lytras, G. Wagner, and P. Diaz, "Ontologies and the semantic web for e-learning," *Journal of Educational Technology and Society*, vol. 7, no. 4, pp. 26-142, 2004.

[13] L. Aroyo, D. Dicheva, and A. Cristea, "Ontological support for web courseware authoring," in S. A. Cerri, G. Gouarderes, and F. Paraguacu, Eds., *Proc. Intelligent Tutoring Systems*, Berlin: Springer, pp. 270-280, 2002.

[14] T. Berners-Lee, J. Hendler, and O. Lassila, *The Semantic Web. Scientific American*, vol. 284, no. 5, pp. 34-43, 2001.

[15] S. Boyce and C. Pahl, "Developing domain ontologies for course content," *Educational Technology & Society*, vol. 10, no. 3, pp. 275-288, 2007.

[16] D. Dicheva and L. Aroyo, "A minimalist approach to support ontology-driven educational information systems communication," presented at International Workshop on Applications of Semantic, 2004.

[17] L. Aroyo and D. Dicheva, "The new challenges for e-learning: The educational semantic web," *Journal of Educational Technology and Society*, pp. 59-69, 2004.

[18] M. Grandbastien, F. Azouaou, C. Desmoulins, R. Faerber, D. Leclet, and Q.-J. Céline, "Sharing an ontology in education: Lessons learnt from the OURAL project," presented at 7th IEEE International Conference on Advanced Learning Technologies, 2007.

[19] M. Bucos, B. Dragulescu, and M. Veltan, "Designing a semantic web ontology for elearning in higher education," in *Proc. IEEE 9th*

International Symposium on Electronics and Telecommunications, pp. 415–418, 2010.

[20] K. L. Chin and E. Chang, "A sustainable ICT education ontology," in *Proc. 5th IEEE International Conference on Digital Ecosystems and Technologies*, Korea, pp. 350-354, 2011.

[21] F. Yang and Y. Q. Chen, "Ontology based application framework for network education resources library," in *Proc. 2nd International Workshop on Education Technology and Computer Science*, pp. 423–426, 2010.

[22] M. Gupta, R. Li, Z. Yin, and J. Han, "Survey on social tagging techniques," *SIGKDD Explor. Newsl.*, vol. 12, pp. 58–72, 2010.

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. degree 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. degree 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.