JOURNAL OF EDUCATIONAL AND INSTRUCTIONAL STUDIES
IN THE WORLD
May 2015, Volume: 5 Issue: 2 Article: 07 ISSN: 2146-7463

E-LEARNING MOBILE APPLICATION FOR ARABIC LEARNERS

Eng, Abdelghani Karkar
Qatar University
QATAR
a.karkar@qu.edu.qa

Prof. Dr. Jihad Mohamad Alja'am
Qatar University
QATAR
jaam@qu.edu.qa

Prof. Dr. Mohamad Eid
Nyu Abu Dhabi
ABU DHABI
mohamad.eid@nyu.edu

Prof. Dr. Andrei Sleptchenko
Qatar University
QATAR
andrei.sleptchenko@qu.edu.qa

Abstract

Digital devices may bring people modern aspect of learning environment. Due to the huge amount of obtainable applications, learners can use them on their mobile devices for entertainment, social network communication, exchanging multimedia, and learning. We present in this paper an educational mobile system that provides explanatory details about terms in Arabic stories characterized by semi-impulsive generated multimedia and generates semantic questions about them. In order to augment the text with these supplementary details, different processing phases have been accomplished which involve: the extraction from expanded educational ontology, word-to-word relationships, supplementary annotation, and querying search engines. The generation of semantic questions depends substantially on the results of semantic queries and reasoning techniques bounded by contrastive semantic rules. The fundamental target of our proposed work is to develop a mobile-based Arabic learning tool that adopt Bloom's Taxonomy for education, including comprehension, analysis, evaluation, knowledge, etc.

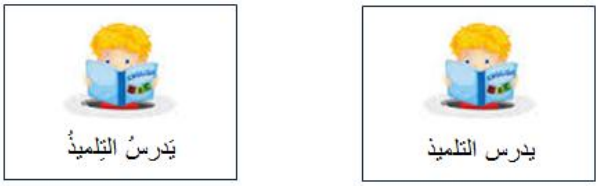
Key Words: Mobile Learning; Arabic Natural Language Processing; Ontology; Multimedia; Engineering Education.

INTRODUCTION

E-Learning is significant for learning languages. The Arabic language is different from Germanic languages in some characteristics such as language direction, looking characters, alteration of character form according to its place in the word, and mark of diacritics to recognize how to pronounce words. Any Arabic eLearning system has a considerable challenge in providing users the ability to use particular Arabic features especially if it is designed for children or non-native Arabic learning adults. Young children need assistance to improve their Arabic skills whereas non-Arabic grown-ups rely on the knowledge of their own language to learn another one.

An Arabic sentence is a chain of words coordinated from right-to-left. An Arabic word is a chain of letters that can have annotated diacritics. A letter with particular diacritic remains one letter but however its pronunciation differs, such as "تَـ" pronounced as "tà", while "تِـ" is pronounced as "tee". There are many symbols to show

particular combination of letters. In fact, there are more than 900 different illustrations used for Arabic scripting (Mosa et al. 2013). Figure 1-A shows the Arabic word "student" with its illustration. In some cases, diacritics can be ignored as shown in Figure 1-B. This happened when writing simple letters and reports.



A B

Fig. 1: The word "Student studying"

Phonetics, which is a branch of linguistics that teaches pronunciation of words and sentences, can be incorporated in e-Learning Arabic learning tools to provide the pronunciation of Arabic words and sentences. Dictionaries and textbooks sort the direction of phonetics from left-to-right (Habash et al., 2007). Few of them apply phonetics to present the Arabic word in their native language like "Al-Telmiz – التلميذ".

In effect, handheld digital devices like a smart mobile phone or a tablet can assist users to improve their language skills. They provide pervasive computing and display to make learning available anywhere and any time. Displaying illustration to describe Arabic words can help in improving learning capabilities for learners. We propose in this paper an e-Learning mobile-based application that displays illustrations for Arabic children stories and generate questions according to the developed Arabic educational ontology. It allows children and non-Arabic grown-ups to understand Arabic vocabulary that are semantically related, like "Student is studying – يدرس التلميذ". The educational ontology is developed with the appropriate editing tool and can be updated to include additional semantic information.

The remaining parts of the paper are organized as follow: in section 2, we present some existing learning systems. Section 3 describes the proposed e-Learning system for Arabic learning. In section 4, we demonstrate the usability of the proposed mobile application. Finally, in section 5, we conclude the paper and provide perspectives for future work.

BACKGROUND

Learning Arabic vocabulary and grammar is a challenging task. Abdelnasser et al. (2014) proposed a question-answer system called "Al-Bayan" to answer particular question related to the holy Quran. It processes the Arabic text, understands semantics and answers users questions depending on reliable Islamic resources like tafseer book. Trigui et al. (2012) introduced the Arabic language at Conference and Labs of the Evaluation Forum (CLEF). They proposed a system that can answer questions with many answers from small Arabic paragraphs. The system uses the modern standard Arabic where its overall accuracy reached 0.19. Mahgoub et al. (2014) proposed an approach for query expansion based on ontology created using Wikipedia. They considered different features to make users utilize semantic search instead of keyword search, by handling: 1) generalization, morphological terms, concepts matching, and providing semantics in the proper sense. Abouenour et al. (2010) proposed a system based on Arabic WordNet in order to enhance question/answering for the Arabic language. Synonyms of terms in WordNet are semantically linked and provide the ability of extracting relevant terms according to the given questions.

Mosa and Kakehi (2014) introduced a multimedia Arabic e-Learning system called (EtaJWa) to help non-Arabic speakers in pronouncing Arabic letters and reading correctly. It uses Roman phonetics code to coincide the Arabic language direction. The proposed work presented significant statistical analysis. Wastam et al. (2010) have developed a system to help children enhance their knowledge in a specific topic. The system is composed of two components: the first one allows the instructor to select the topic of a story that will be illustrated by flashcards, while the second one asks the children to arrange the scenes according to the topic of the story.

Erradi et al. (2012) have proposed a multimedia-learning platform called ArabicTutor that provides Arabic spelling, gives meanings of words accompanied by simple multimedia. Wuang et al. (2011) have built a multimedia courseware system that is based on learning theories. All these systems are based on static contents with different learning objectives. They are generally addressed to adult users with a computer machine.

Educational ontologies are developed to improve the learning concepts in a particular area. They can be used in different domain to facilitate the access to information. Zemmouchi and Ghomari (2013) have developed a higher education reference ontology called 'HERO' which can help universities CEO in setting strategic development efficiently. John (2014) has developed an ontology to learn the Java programming language where concepts are linked in a hierarchical mode. Raju and Ahmed (2012) have proposed an enabling technology to create next-generation learning object repository. They have shown how semantic web and ontologies can be used to evolve and provide sharable learning objects. Sawsaa and Lu (2014) have proposed an ontology of information science for the process of software development life cycle. The formulation of concepts depends on recognizing the information science notions and coinciding them into a hierarchical structure view based on their classifications.

We propose an educational mobile system based on Arabic educational ontology and multimedia technology to teach children and non Arabic grown-ups in an attractive way. Our proposal generates the multimedia tutorials dynamically by using Arabic text processing, entities relationship extraction, educational ontology, and extraction of multimedia contents fetched from online search engines (i.e., Bing or Google).

PROPOSED SYSTEM

Our approach consists of generating the multimedia educational content dynamically. We use Arabic text processing tool for knowledge extraction (i.e., actors, action, event, location), educational ontology to retrieve details and associated multimedia elements, and web search engines to get additional contents. The instructor can send to the system the educational text and get the multimedia tutorials. The dynamic tutorials would become more time efficient due to the machine learning process of the automated system. The system will statistically learn the preferred customization per student and will generate automatically the customized tutorials.

A. Procedure

In order to develop our system, we have created first the general system artwork, set the end user graphical user interface, design the semantic model that will store all semantic information about terms, and collect educational stories and analyze them. We have gathered 30 educational stories, annotated terms, and associated some illustrations manually. Illustrations were gathered from the Internet and educational CDs. The system is developed while adopting the standard software development life cycle, as depicted in Figure 2 below.

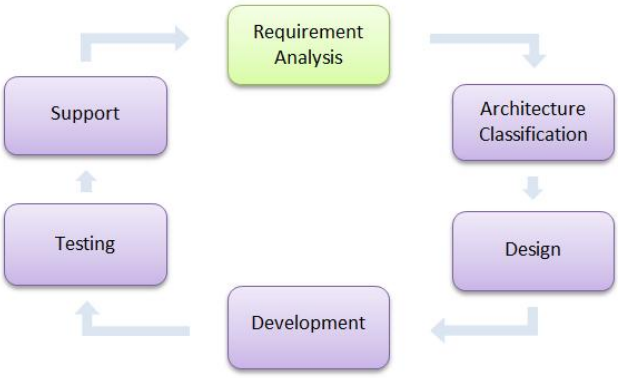


Fig. 2: Software development life cycle

B. Ontology Building

An ontology is a formal explicit description of concepts in a domain of discourse (classes, sometimes called concepts), properties of each concept describing various features and attributes of the concept (slots, sometimes called roles or properties), and restrictions on slots (facets, sometimes called role restrictions). Ontology together with a set of individual instances of classes constitutes a knowledge base (Gruber, 1993). Some reasons of using the ontology are: 1) To share common understanding of the structure of information among people or software agents; 2) To enable reuse of domain knowledge, to make domain assumptions explicit; 3) To separate domain knowledge from the operational knowledge; and 4) To analyze domain knowledge. Since ontologies provide a shared understanding of a domain of interest, they have become a key technology for semantics knowledge extraction and integration. The USDA National Animal Genome Research group has developed standardized trait ontology STO for farm animals. It contains information about cattle, pig, chicken and other species. Using the existing ontologies in the same domain was the most important of our considerations. Since the field of our study is education, we preferred the creation of a new ontology in a simplified manner to facilitate the delivery of information to learners. In order to build our ontology, we used the technique proposed in (Noy and McGuinness, 2001). Figure 3 shows the different phases while developing the ontology.

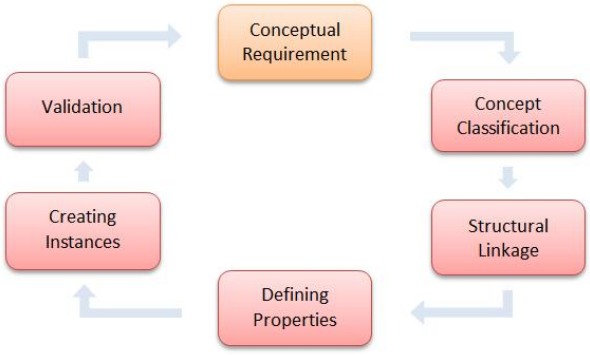


Fig. 3: The phases while creating the ontology

We defined all relevant terms that are expected to be in the ontology and we provide an index of specific children's tutorial (i.e., toys, animal, food, etc.). We created (either present or past tense should be used) different questions to extract the most important terms from educational stories, like “What is the benefit of a book?”, “What is composed from?”. We have also added detailed information about terms and categorized them.

We start with the definition of the most general concepts in the domain and subsequent specialization of the concepts. From the list created previously, we select a group of terms and ask what they have in common and what 'siblings' there might be. These concepts will be arranged in a hierarchical taxonomy into the ontology. Figure 4 shows a part of the class hierarchy of our *Noun* grammar ontology.

We defined also the facets of the slots, sometimes called role restrictions (Noy and McGuinness, 2001). In the Noun (الإسم) class, we used the different restrictions to prevent conflicts while using the inference technique. For instance, an elevated noun (مرفوع) cannot be at the same time genitive (مجرور). After defining the ontology, we have evaluated it through the system. We have also discussed its appropriateness with experts in the domain. In order to retrieve the corresponding multimedia instances from the ontology, we used semantic SPARQL[1] queries. Once a SPARQL query is executed a list of multimedia instances that satisfy the query is retrieved. They are then proposed to the instructor. Complimentary queries are sent to the search engines.

[1] http://www.w3.org/TR/rdf-sparql-query/

JOURNAL OF EDUCATIONAL AND INSTRUCTIONAL STUDIES
IN THE WORLD
May 2015, Volume: 5 Issue: 2 Article: 07 ISSN: 2146-7463

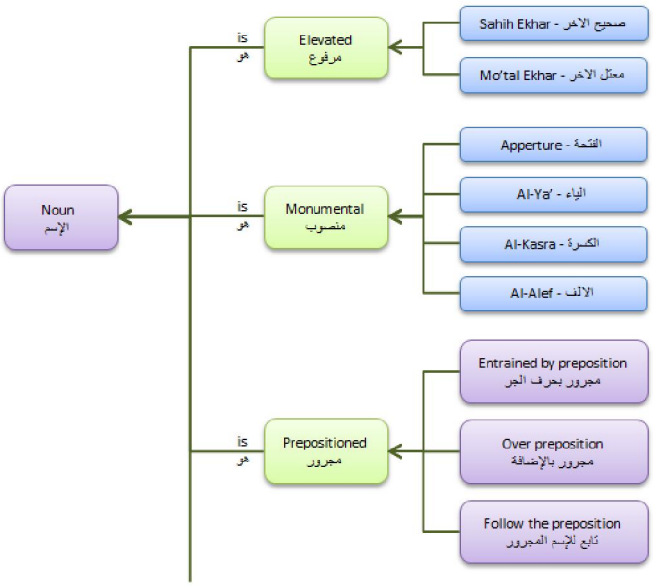


Fig. 4: A part of the Noun grammar ontology

C. Inference

We have used the inference technique in order to improve the quality of semantic knowledge through discovering new relationships between instances, analyzing the content of information, and identifying inconsistencies in data. For instance, the fork is an instance from Kitchen_Tool, fork is used with plate, subsequently if particular nutriment uses fork then the plate instance is required too.

D. Arabic Text Processing

We have used the recent version from "Stanford Parser[2]" to process the natural text. Stanford parser is a natural language processing tool that marks up the grammatical structure of the sentence. It annotates words according to their part-of-speech (POS) and generates their structure tree. It uses different Treebank to adapt languages (e.g., Penn Arabic Treebank, Chinese Treebank, etc.). The result of the parser is analyzed to store most frequent terms, and used in order to generate multimedia elements. Figure 5 shows the result after processing the natural text.

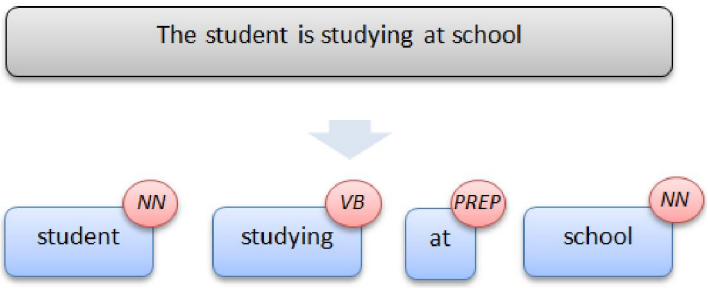


Fig. 5: Text Processing Phases

[2] http://nlp.stanford.edu/software/lex-parser.shtml

JOURNAL OF EDUCATIONAL AND INSTRUCTIONAL STUDIES
IN THE WORLD
May 2015, Volume: 5 Issue: 2 Article: 07 ISSN: 2146-7463

E. Semantic Model Architecture

The semantic model is developed using "Protégé editor[3]", a free open source ontology editor developed by Stanford. The semantic model is composed from many classes that are referred to as concepts. Each concept can have different properties that describe the characteristics of the concept itself (e.g., tool_color, tool_size, etc.). It is possible to create instances from each class where each instance will have the structure of its model concept. Figure 6 shows essential concepts in the educational ontology.

Fig. 6: The grammar educational ontology hierarchy

F. Generation of Question

Generation of questions uses the defined semantic model in order to generate random answers. Answers are generated randomly by accessing the same properties of other instances. After each attempt, the mobile application performs statistical analysis to report the questions that are answered incorrectly. According to the statistical analysis, wrong answered questions will be asked repeatedly to make sure that the user is able to answer them properly.

SYSTEM IMPLEMENTATION

A. Proposed Modules

The proposed system is composed of different modules as shown in Figure 7. The *Natural Language Parser* module is responsible for providing all text processing features, which include paragraph segmentation, words tagging, diacritics annotations, and so. The *Corpus* module is responsible for reading and storing new processed Arabic stories. In addition, it stores statistics about the frequently used words and their respective downloaded illustrations. The *Semantic models* component provides the functionality to access required ontology to acquire semantic knowledge attached to particular terms, inference techniques, and reasoning. The *multimedia loader* provides required services to translate words into illustrations. It uses the Search Engine module to get the required illustrations if it is not found in the local media database. The Search Engine accesses online search engines API to provide different search features, which include search for images, text, translation, and so.

[3] http://protege.stanford.edu

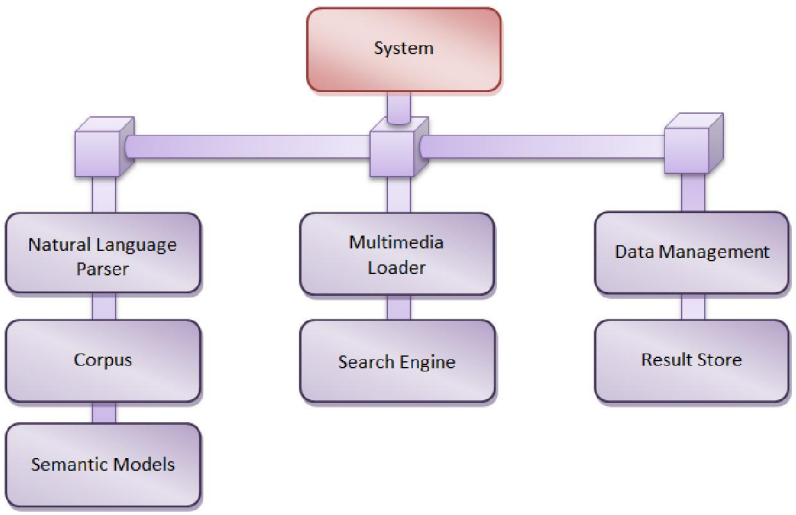


Fig. 7: System implementation modules

B. System Architecture

The architecture of the implemented system is based on distributed computing structure model as shown in Figure 8. The advantage from this architecture is to separate server logic from the client one. The server is responsible for all processing phases, like processing clients' requests, performing semantic analysis, maintaining usage log, and so on. The server is developed with Java language. It uses different components (e.g., search engine library, semantic library, etc.) to fulfill the requirements of processing the text. On the other hand, the mobile application is developed with HTML5 markup language. It uses different JavaScript libraries to display information in an enjoyable style (e.g., Jquery Mobile, JSTree, etc.). The tremendous benefit from this technology is to support the diverse type of mobile-based systems.

Fig. 8: System architecture

C. Generation of Multimedia

The obtained result from the Text parser module will be used to classify terms and place them in a hierarchical structure. A search query will be formed using the template of the created structure (e.g., ADJ-NN: Blue Cup, NN-VB: Water is boiling, etc.). The query will be used in the search engine, like Google, using its API to download related images that represent the sentence. The administrator can store download images in the corpus for future usage.

JOURNAL OF EDUCATIONAL AND INSTRUCTIONAL STUDIES
IN THE WORLD
May 2015, Volume: 5 Issue: 2 Article: 07 ISSN: 2146-7463

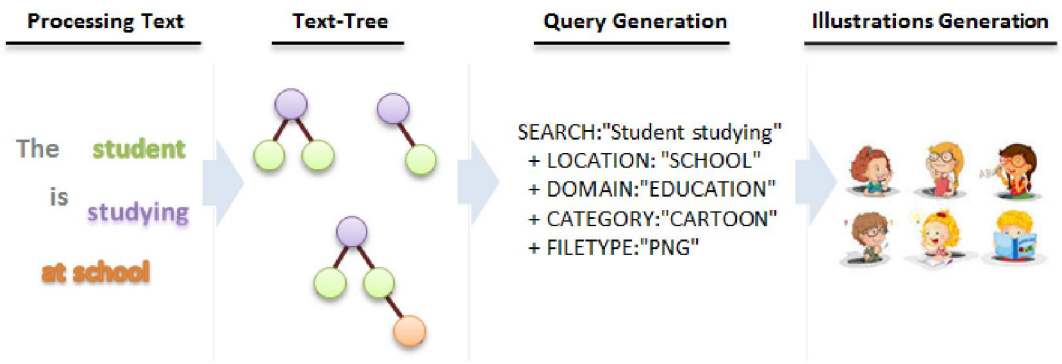


Fig. 9: Illustration generation process

D. System Operation

The mobile application enables the user to access different learning sections (e.g., Grammar, Reading, Assessment, etc.). It uses both Arabic and English languages. Three snapshots of the developed application are shown in Figure 10 (labeled as screens A, B, and C). In screen A, the English version of the home page is displayed which allows the user to navigate in the mobile application. In screen B, the user can navigate through the Noun grammar tree and select the node to get more information about it. Finally, in screen C, the user can check the result of the assessment he/she makes.

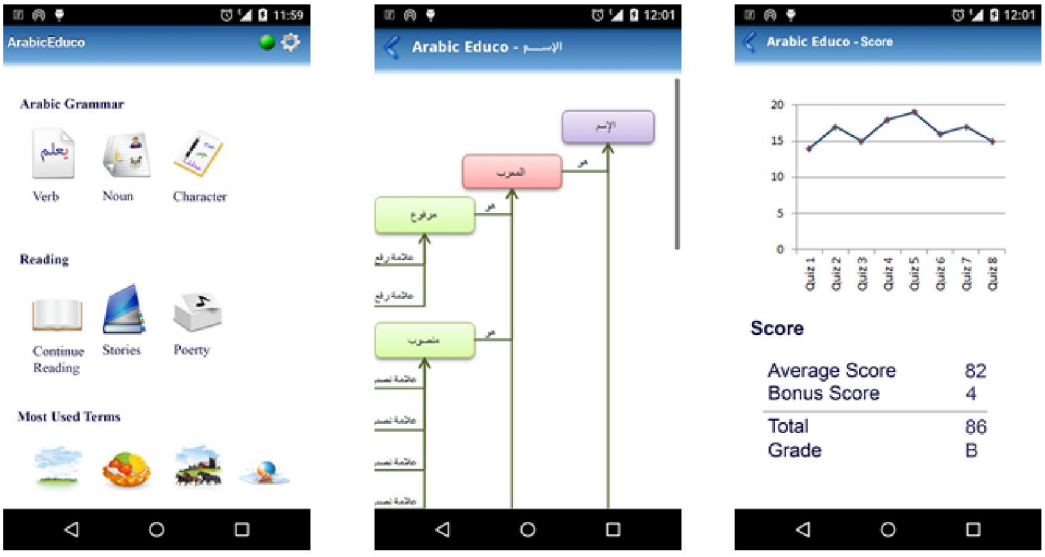


Fig. 10: Application screenshots: Home page, Noun Tree, Assessment results (from left-to-right)

CONCLUSION AND FUTURE WORK

In this paper, we present an e-Learning mobile-based system that can generate illustrations for Arabic stories through knowledge categorization and extraction using natural language processing and semantic models. The e-Learning system can provide explanatory details about Arabic vocabulary and can display the corresponding images. It generates questions according to the developed educational ontology. The system can also be used

by children and non-Arabic grown-ups to learn Arabic in an attractive style. The system can develop several learners' skills such as recognition, understanding, and memorization of Arabic words and phrases.

In the future, we plan to improve the generation accuracy of images. In order to improve the accuracy of generating images, methods to filter appropriate images from non-useful ones will be considered. We plan also to develop a new assistive writing device to teach the learners how to write Arabic words properly.

Acknowledgement: This publication was made possible by a grant from Qatar University under its graduate teaching assistantship award. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of Qatar University.

WJEIS's Note: This article was presented at 6th International Conference on New Trends in Education and Their Implications - ICONTE, 24-26 April, 2015, Antalya-Turkey and was selected for publication for Volume 5 Number 2 of WJEIS 2015 by WJEIS Scientific Committee.

REFERENCES

Abdelnasser, H., Mohamed, R., Ragab, M., Mohamed, A., Farouk, B., El-Makky, N., & Torki, M., "Al-Bayan: An Arabic Question Answering System for the Holy Quran", ANLP, 2014.

Abouenour, L., Bouzouba, K., & Rosso, P., "An evaluated semantic query expansion and structure-based approach for enhancing Arabic question/answering", International Journal on Information and Communication Technologies, vol. 3, no. 3, pp. 37-51, 2010.

Erradi, A., Nahia, S., Almerekhi, H., & Al-kailani, L., "ArabicTutor: A multimedia m-Learning platform for learning Arabic spelling and vocabulary,", In International Conference on Multimedia Computing and Systems (ICMCS), pp. 833-838, 2012.

Gruber, T. R., "Towards Principles for the Design of Ontologies used for Knowledge Sharing," In Journal of Human-Computer Studies. Florida Institute for Human and Machine, vol. 43, pp. 917-928, 1993.

Habash, N., Soudi, A., & Buckwalter, T., On Arabic Transliteration. In Arabic computational morphology (pp. 15-22). Springer Netherlands, 2007.

John, S., "Development of an Educational Ontology for Java Programming (JLEO) with a Hybrid Methodology Derived from Conventional Software Engineering Process Models," In International Journal of Information & Education Technology, vol. 4, no. 4, 2014.

Mahgoub, A. Y., Rashwan, M. A., Raafat, H., Zahran, M. A., & Fayek, M. B., "Semantic Query Expansion for Arabic Information Retrieval", ANLP, 2014.

Matsuda, K., & Lea, R., WebGL programming guide: interactive 3D graphics programming with WebGL. Addison-Wesley, 2013.

Mosa, A. and Kakehi, K., "THROUGH THE LOOKING-GLASS: Roman letters in phonics for Arabic as a part of multimedia support". 9th WSEAS International Conference on educational technologies (EDUTE '13). Kuala Lumpur, Malaysia, April 2 to 4. Proceedings Page Number 120-125, 2013.

Mosa, A., & Kakehi, K., "Letting Non-Arabic speakers read and pronounce Arabic sounds using roman phonetic codes in the mirrored form-a feature of ETaJWa, an Arabic e-learning system". Int. J. Inf. Technol. Comput. Sci.(IJITCS), pp. 2091-1610, 2014.

Noy, N., and McGuinness, D., "Ontology Development 101: A Guide to Creating Your First Ontology", Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, 2001.

Raju, P., & Ahmed, V., "Enabling technologies for developing next-generation learning object repository for construction", Automation in Construction, vol. 22, pp. 247-257, 2012.

Sawsaa, A., & Lu, J., "Building an Advance Domain Ontology Model of Information Science (OIS)", International Journal of Digital Information and Wireless Communications (IJDIWC), vol. 4, no. 2, pp. 258-266, 2014.

Trigui, O., Belguith, H. L., & Rosso, P. "DefArabicQA: Arabic definition question answering system", In Workshop on Language Resources and Human Language Technologies for Semitic Languages, 7th LREC, Valletta, Malta, pp. 40-45, 2010.

Trigui, O., Belguith, L. H., Rosso, P., Amor, H. B., & Gafsaoui, B. "Arabic QA4MRE at CLEF 2012: Arabic Question Answering for Machine Reading Evaluation", In CLEF Online Working Notes/Labs/Workshop, 2012.

Wastam, J., Awang Rambli, D. R., & Sulaiman, S., "A guided digital storytelling prototype system using illustrated flashcards," pp 1-6, 2010.

Wuang, Y. P., Chiang, C. S., Su, C. Y., & Wang, C. C. (2011). Effectiveness of virtual reality using Wii gaming technology in children with Down syndrome.Research in developmental disabilities, vol. 32, no. 1, pp. 312-321, 2011.

Zemmouchi-Ghomari, L., & Ghomari, A. R., "Translating Natural Language Competency Questions into SPARQLQueries: A Case Study", In the first International Conference on Building and Exploring Web Based Environments, pp. 81-86, 2013