A Concept Analysis Scheme of Simple Stories for Learning Resources through Extraction of Domain-based Multimedia Elements

M. U. Mwinyi
Qatar University, Doha, Qatar
m.udi@qu.edu.qa

J. M. Alja'am
Qatar University, Doha, Qatar
jaam@qu.edu.qa

S. A. ElSeoud
British University in Cairo, Cairo, Egypt
selseoud@yahoo.com

Abstract—Children with special needs face major difficulties to understand the script of stories. They cannot comprehend the words' meaning properly. Instructors use different techniques to explain them. They read the script several times, use glossy images, symbols, and short clips. The instructors spend a lot of time searching for these elements. The preparation of every tutorial requires several hours of hard work. The objective of this work is to build a system to assist the instructor finding multimedia elements in real time that can be used during the learning sessions. The proposed system is based on a domain-oriented architecture focusing on animals and food. We use different techniques for text processing including concept analysis, key words extraction and sentence ranking. We can then determine the most significant words and sentences and generate the corresponding multimedia elements. Instructors and children can use smart devices in learning inside and outside the classroom.

Keywords—Keywords Extraction, Multimedia, Ontology, Special Education.

1 Introduction

Children with special needs face a significant difficulty to understand the script of simple stories [13]. The instructors use different methodologies to explain the story' words to the children that include, reading the sentences several times; designing glossy images of the objects and concepts of the story; generating sounds of charac-

ters (i.e., birds, dogs, camels, etc.) and/or objects (e.g. cars, motorcycles) [1,4,5,9]. Finding or developing these items is typically complicated and requires significant efforts. The instructors spend hours to find these elements which also may need suit all the children in the same classroom where additional elements need always to be generated. Developing a system that can receive the story script as an input and automatically generate the corresponding multimedia elements will be of great importance for instructors, children and for parents who can use it at home. In this paper, we propose a new system based on a domain-oriented architecture to analyse the text and generate the multimedia. We create a large corpus that contains hundreds of sentences used in stories and categorizes them automatically based on their domains. We propose a new algorithm that determines the keywords of the script and selects the most significant sentences that represent the story. Instructors can then focus mainly on these sentences to explain the script. We use several natural language processing techniques [6,7,21] to extract the keywords and we link them with multimedia elements (e.g., images and short clips) fetched from Google search engine. The search is performed through Google's APIs features [8]. The proposed system is intended to improve the understanding and communication skills of children with SP and to help them understand stories in an enjoyable and entertaining manner. The system is also accessible through smart devices (i.e, iPad).

The remaining of the paper is organized as follows: section 2 discusses the system main components; section 3 presents the methodological approach; section 4 conclude the paper.

2   System Components

2.1   Building a Corpus of Stories

This corpus will group stories keywords with their associated multimedia elements and allows retrieval of the multimedia elements corresponding to a story words and sentences. The corpus is enriched automatically by the addition of new scripts. The goal of this corpus is to have a resource that groups all terms used in stories along with a weighting index indicating their importance to a domain. Whenever a new story script is added to the corpus, the terms' weights are calculated automatically and updated.

2.2   Information Selection and Presentation

Readers of stories are always looking for the information to flesh out the main themes and concepts of the story. Determining the importance of this information is an approach that readers use to distinguish between significant data and that which is not essential for an understanding of the story, and to decide which parts of the story need most attention and focus [3]. Instructors need to systematically educate children on how to extract the most significant information of stories they read. The approach

proposed in this work provide new means of identifying the main points during reading stories and automatically select the most important sentences of the story for presentation to the reader. We can therefore allocate more focus on these sentences to be disseminate to children.

2.3 Visualization and Multimedia to Represent the Story

A data visualization approach can be used to attract the attention of learners and keep them engaged for a longer time. It can also influence their thinking and embed the ideas of the story in their mind. In fact, half of the human brain is dedicated to processing visual information, meaning that "whenever we present a reader with information injected with some images, we are reaching them through the mind's highest-bandwidth pathway" [10]. Absorption of important tales via data visualization technique is a very powerful manner to increase the understanding of stories [16,20]. The approach proposed in this work consists of linking the main concepts of the story with various multimedia elements. Instructors can use these elements to explain the story more effectively. In addition, they can obtain commentary data whenever needed from Google search engine using its APIs to strengthen the understanding of the concepts. In fact, additional resources are always helpful for understanding new ideas.

3 Related Work

In the section, we will discuss some works related to sentences ranking and section, image retrieval and filtering and conversion of text to multimedia.

3.1 Sentence Ranking and Selection

Most of the works on sentence selection revolves around the summarization of documents to assist the readers in understanding the main points present in the core text with a minimum effort. Automatic text summarization systems use the four methods to determine the sentence weights [17]: (1) term frequency weight; (2) cue method; (3) position method; and (4) title method. Some of the trainable document summarizers use the following features to calculate the weight of the sentence [18]: (1) sentence length; (2) cue method; (3) paragraph feature (i.e., position method); and (4) uppercase word feature (given that upper-case words are often thematic).

3.2 Image Retrieval and Filtering Algorithm

Most of the previous work on image retrieval from search engines depends mainly on humans' selection of the relevant images. While real-time Google and live image search, re-ranking rely on images' annotations or tags. These approaches lead to ambiguous and some noisy results [19]. In fact, images annotations and keywords used in queries are not properly matching. We can then get inappropriate images that should be eliminated.

3.3 Understanding of Stories Through Multimedia

Shize et al. [15] proposed an approach that generates image-text timelines for news events based on evolutionary image-text summarization. Alternatively, Dhiraj et al. [11] presented an approach that automates story picturing based on the mutual reinforcement principle. In their work, semantic keywords are extracted from the story text and an annotated image database is searched to form an initial picture pool. While most of the existing work relies on translating the whole text to relevant images, this work selects the most important sentences given the domain in discourse, identifies then the semantic relationships between entities, and translates them into images. We give then the most representative images to the story script. Our work consists of the following components:

1. *Automatic creation of domain-term database:* This component intends to create a large corpus that hosts children stories and categorizes them based on their domains. Our algorithms can analyze new stories, identify new terms and calculate their weights in relation to the domain of discourse. The main goal of this approach is to find a way of building the domain-oriented corpus automatically and without any human intervention.
2. *Sentence ranking and selection:* One of the most significant processes required to understand a story is the identification of the most important ideas in the story script based on the domain of discourse. To identify the most important concepts in the story, our algorithm computes the weights of all sentences in a story and rank them. The weight of a sentence depends totally on the weight of terms within the story and their relevance to the domain. The calculation of term weight depends on two major factors:
   (a) Historical information: the importance of the terms to the story domain; and
   (b) Current information: the importance of the terms within the story.
3. *Image retrieval and filtering algorithm*: The approach used in this work allows translation of semantic relationships to multimedia using a new image filtering algorithm. Since the proposed system will be used in education, it is highly important to ensure that, during the retrieval of images, certain filters must be applied:
   (a) Sensual and pornographic images must be discarded.
   (b) Images that are not related to the topic must be eliminated.
   (c) Images retrieved must appropriately represent the actions of the story.

4 Our Approach

This In this section, we will discuss in detail the proposed approach for understanding simple stories. This approach is divided into two parts:
   (a) Automatic creation of domain-term indexing: building a domain-oriented dataset along with methods of feeding it.
   (b) Story text processing and multimedia mapping: using sentence ranking and selection algorithm, pattern recognition, image retrieval and processing together in the form of a pipeline process to increase the quality of results.

Figure 1, shows the four components of our approach highlighted in blue color which are: *(1)* building a domain-oriented; *(2)* sentence weighting and selection; *(3)* pattern recognition and role identification; and *(4)* Google image retrieval and selection.
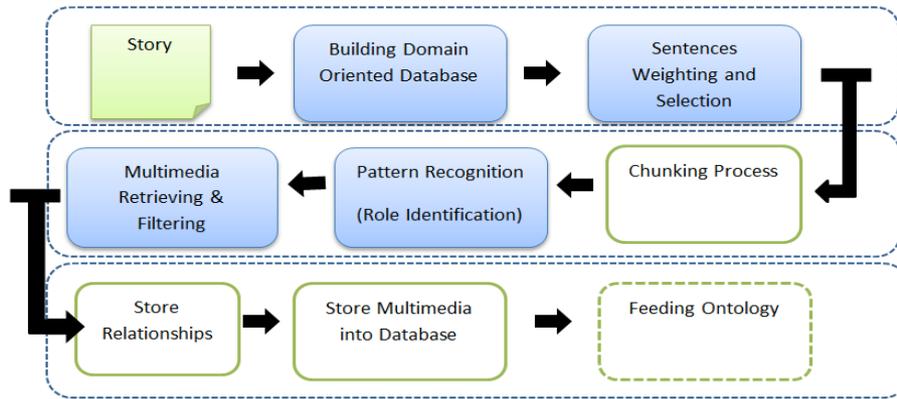

Fig. 1. The story script processing phases.

4.1 Automatic domain-term creation

This component creates a corpus that hosts a large set of categorized stories based on their domains. This component can analyze a new story automatically, extract its terms, calculate their weights, and update the corpus as depicted in Figure 2. A major challenge when building a corpus is the determination of domain of the selected terms to be inserted in the corpus. We build first a simple corpus where the terms' weights are calculated manual or in a semi-automatic manner. Several questions remained regarding the determination of the ranking of terms, including: what are the terms that represent the story script? How can we assign weights marking the importance of these terms to the domain? Based on our proposal, these questions are addressed using the document frequency term weighting with slight modification of the formula.
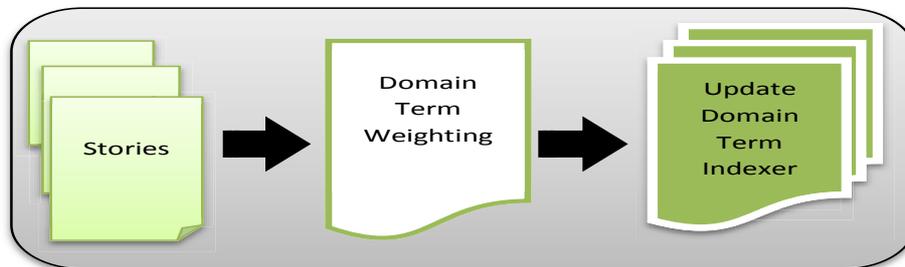

Fig. 2. Terms weighting and indexing.

Existing terms in documents need to be stored in the INDEXER (i.e. in the corpus table) and assigned weights that express their importance within a domain as show in Table 1.

Table 1. Example of a corpus of keywords with weightings based on domain annotation

Term	Domain	Weight
Eat	Animal	1.5
Eat	Food	2
Meat	Food	2
Zoo	Animal	2
Kid	Food	1

A commonly used term weighting method is the normalized document frequency, which consists of assigning a high weight to a term if it occurs frequently throughout the document. Additionally, a none stop word term that occurs in nearly all documents within the specific domain has more power and will be given a greater weight. Before calculating the normalized document frequency weight of a term in a domain, it necessary to know: *(1)* how many documents are available in the domain (N = number of documents); and *(2)* how many documents of the collection the term occurs in (the document frequency = DF). The following formula is then applied:

Domain Weight (DW) = 1 + DF / N

4.2 Sentence ranking algorithm

One of the most important processes in our approach is the determination of the most important sentences of a story given the domain of discourse [31]. The proposed ranking algorithm is based on a sentence weighting and ordering technique. We assumes that a given document d is represented by a set of terms $\{t_, t_2 ..., t_n\}$ in which every t_i is a word that occurs in the document d_i, and n represents the total number of stemmed words in d_i. The algorithm proceeds as follows:
(a) Given a document d and a domain D.
(b) Documents = bags of [unordered] words; a document is composed of terms $t_,$ $t_2, ..., t_n$
(c) The calculated weight of each term t_i in the document $\approx$ *weight (t,d,D)*.
(d) The function therefore requires three inputs to calculate the weight which are: (1) *Term(t);* (2) *Document(d)*; and (3) *Domain(D)*.

The term weight is calculated through three steps using the following:

1. First, the *term frequency tf* (t_i, d_i) evaluates all *nouns* and *verbs* in the document except the stop words**.** This value represents the number of occurrences of the term in the document.

2. Second, the domain weight (*DW*) of the terms (nouns or verbs) is calculated. The domain weight shows the importance of the term within the domain as discussed previously.
3. Finally, w_i – the combination of these two entities – is calculated. Therefore, d_i can be represented as a specific *n*-dimensional vector d_i as follows: $d_i = \{w_1, w_2,..., w_n\}$.

$$W_i = tf(t_i, d_i) * DW(t_i, D)$$

The general framework of sentences ranking and extracting the those that represent the main ideas of the story is shown in Figure 3.
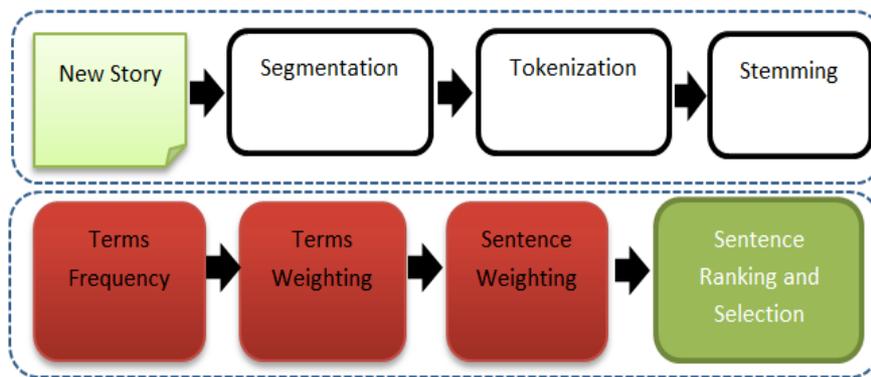

Fig. 3. Sentences ranking and extraction.

4.3 Pattern Recogantion

The main benefit of using pattern recognition is the ability to automatically identify hidden objects, types or relationships that exist in raw data; for instance, the recognition of location information in text.

Pattern recognition is the most important stage of the proposed system as it allows extraction of the following information that is needed during the multimedia translation process:

1. Actions.
2. Actors (i.e., subject and object).
3. Location.
4. Time.

Most previous systems relied on hand-fashioned patterns, while others learnt automatically or semi-automatically. In this work, we combine the two approaches by using the output that has been generated from the chunking process - that is, the chunking sequence. 'Chunking' is a process of splitting text into syntactically related words - for example, "a big tree" is considered as a noun phrase chunk (NP) [14].

The proposed system can identify the following pieces of information: objects, places, time and actions that describe the object or places. The goal is to create a classifier that does not depend on manual annotation. For instance, the following sentence:

The ugly duckling spent the winter in a marshy pond

becomes: NP-VP-NP-IN-NP after applying the chunking algorithm. The system assigns the role of each chunk as shown in table 2:

Table 2. Chunk abbreviation and associated role

Chunk	Role
NP	Object
VP	Action
IN	Empty
NP	Place

Sentences in any language fall into specific number of patterns that determine the function or role of nouns, verbs and adjectives. In the English language, there are approximately ten patterns that define a sentence structure [54]. The patterns are classified based on the type of verb used: *(a)* verbs of being such as 'is'; *(b)* linking verbs such as 'seem'; and *(c)* action verbs such as 'write'. Sentence patterns can be represented as a decision tree; Figure 4 shows the classification in a decision tree proceeding from top to bottom. The question asked at each node concerns the property of the pattern, and downward links correspond to possible values. Successive nodes are visited until a terminal or leaf node is reached, following which the pattern label is read. The decision tree can be expanded to include an indefinite number of levels. Each node in the decision tree contains three values: ChunkType; ChildID; and ParentID. For example, in VP (2,1), VP is a Verb Phrase chunk, 2 is a chunk ID (ChildID) within the chunk tree, and 1 is the ParentID of this chunk NP(1,0).
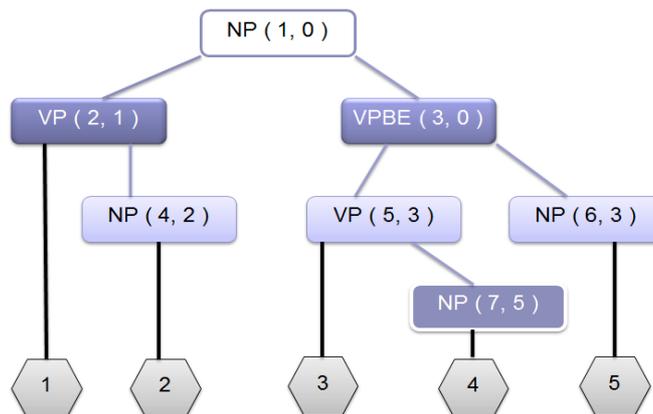
Fig. 4. Pattern recognition decision tree.

After the identification of the pattern ID, the system can assign the role of each chunk. Table 3 shows the data model architecture for the pattern recognition and identification process.

Table 3. Data Model Architecture for pattern recognition process for the role identification

Pattern ID	NP¹	VP	NP²
1	Subject	Action	N/A
2	Subject	Action	Object
3	Subject	Action	N/A
4	Subject	Action	Object
5	Subject	N/A	Adjective

4.4 Translating relationships to multimedia elements

The following filtering algorithm eliminated non-relevant and offensive images from the set of retrieved images through Google search engine.

Algorithm 1: Image Filtering Algorithm	
Inputs:	
	(Relationship: **subject**, **action**, **object**)
Outputs:	
	CANDIDATES
	// List of top five images
1:	**BEGIN**
2:	List ⬚ 0
3:	While list <= 5 do
4:	RETRIEVE image from Google;
5:	GET Image Annotation Value (Content);
6:	GET Image Annotation Value (Title);
7:	**IF (IMAGEMATCH(Content,** subject, action, object) **and IMAGEMATCH(Title,** subject, action, object))
8:	**{**
9:	CANDIDATES ⬚ IMAGE
10:	LIST ⬚ LIST + 1
11:	**}**
12:	**ELSE**
13:	**Next**
14:	**End while**
15:	**RETURN** CANDIDATES
16:	**END**

Algorithm 2: Matching between subject, action, object and Text	
Inputs:	
(Text, subject, action, object)	
Function IMAGEMATCH()	
0:	Matched ▯ **FALSE**
1:	TOKENSBUFFER ▯ **Tokenize** (Text);
2:	ImgcontentStemming ▯ **Stemming** (TOKENSBUFFER);
3:	NextSearch ▯ "subject"// *checking the value in sequence;*
4:	**FOR** Each Token in TOKENSBUFFER **DO**
5:	**IF** (NextSearch.equal("subject"))
6:	{
7:	**IF** (Stemming[Token] contains subject)
8:	nextsearch ▯ "action"
9:	}
10:	**ELSE**
11:	**IF** (nextsearch.equal("action"))
12:	{
13:	**IF** (Stemming[Token] contains action)
14:	nextsearch ▯ "object"
15:	}
16:	**ELSE**
17:	**IF** (nextsearch.equal("object")
18:	{
19:	**IF** (Stemming[Token] contains object)
20:	Matched ▯ **TRUE**
21:	**EXIT LOOP**
22:	}
23:	**END LOOP**
24:	**RETURN** Matched

5 The System

The proposed system proved its effectiveness to assist the instructors during the learning sessions. The instructors need only to input the text to the system and see the sentences with their multimedia elements. Every student is provided with an iPad where words and their corresponding multimedia can be seen. The instructor should validate the retrieved images before sending them to the children. They can get complimentary images by a simple clique on the words or sentences. The system allows personalized learning. In fact, every child will have a dedicated data sets where words and images that she/he learnt are stored together. Therefore, whenever a sentence occurs in the story text, the system will search first its occurrence in the children' dataset before fetching the images from Google search engine. This approach will help the children understand better the new story as the images are presented from their prior knowledge. Every time a child learn a new word, it will be stored in the data set with its pool of representative images.

6   Conculsion

In this paper, we have proposed a new computerized system that allows special education instructors to convert story scripts to multimedia. The algorithm for keyword extraction from story texts can select the most significant words and sentences based on the story domain. Our approach is markedly different from the formal concept analysis that was used for term extraction; instead, our system allows the user to create a new corpus containing story terms annotated with their weights based on their domains. This corpus can be enriched automatically by the addition of new stories which are analyzed and categorized automatically. Their new terms are inserted into the corpus along with their weights, while the weights of existing terms are recalculated upon each new addition. The text of each story is pre-processed and restructured based on the weights of its sentences, allowing the instructor to select the sentences that cover the main ideas of the story. These sentences are ranked and presented to the instructor to facilitate the learning process. We have also proposed an algorithm that can identify the actors and actions in story texts. This algorithm can identify, for instance, that a banana has been eaten by a monkey, or that a lion has attacked a zebra. This methodology thus allows queries to be sent to the Google search engine in order to retrieve the corresponding multimedia elements – i.e. those that show a real monkey eating a banana or a lion attacking a zebra. We have also proposed an algorithm to filter the retrieved images from Google in order to exclude offensive or irrelevant images. A graphical user interface has been proposed to facilitate the use of the system. The instructor can thus input any story text and get as output of corresponding multimedia elements, allowing them to simultaneously show the sentences and the multimedia elements to their pupil(s). This approach may facilitate the teaching methodology and strengthen the learning skills of the children. Initial feedback from instructors and readers has been promising. The system allows instructors to improve their teaching methodology and reduce the repetition of the lessons and concepts.

7   References

[1] J. Chelin, L. Kosseim, and T. Radhakrishnan, "Using Natural Language Processing to Assist the Visually Handicapped in Writing Compositions," in *LNAI 4013,* L. Lamontagne and M. Marchand, eds., pp. 300–311, 2006.

[2] D. G. Evans, L. Bowick, M. Johnson, P. Blenkhorn, "Using Iconicity to Evaluate Symbol Use," in *Proceedings of the 10th International Conference Computers Helping People with Special Needs*, Austria, LNCS 4061, Springer, pp. 874–881, 2006. https://doi.org/10.1007/11788713_127

[3] D. Debbie, "Comprehension Strategies: Summarizing, Synthesizing and Determining Importance," http://www.decd.sa.gov.au/northernadelaide/files/links/Summarising_Synthesing__DI.pdf, pp. 1–4, 2011.

[4] A. Elsaddik, "Interactive Multimedia Learning," ISBN: 3-540-41930-6, Springer, 2001. https://doi.org/10.1007/978-3-642-56790-2

[5] D. G. Evans, L. Bowick, M. Johnson, P. Blenkhorn, "Using Iconicity to Evaluate Symbol Use," in *Proceedings of the 10th International Conference Computers Helping People with Special Needs*, Austria, LNCS 4061, Springer, pp. 874–881, 2006. https://doi.org/10.1007/11788713_127

[6] J. M. Alja'am, A. Jaoua, A. Hasnah, F. Hassan, M. Hassan, T. Mosaid, H. Saleh, F. Abdullah, and H. Cherif, "Text Summarization based on Conceptual Data Classification," *Chapter XI in the Book "Agent Technologies and Web Engineering: Applications and Systems"* G. Alkhatib and D. Rine, eds., IGI Global, USA, pp. 195–209, 2008. https://doi.org/10.4018/978-1-60566-618-1.ch011

[7] J. M. Alja'am, A. Jaoua, A. Hasnah, F. Hassan, M. Hassan, T. Mosaid, H. Saleh, F. Abdullah, and H. Cherif, "Text Summarization Based on Conceptual Data Classification," in *The International Journal of Information in and Web Engineering*, (44), pp. 22–36, 2006.

[8] Google Ajax Language API, http://code.google.com/apis/ajaxlanguage/, Last Access 20-Jul-2016.

[9] M. A. Garcia-Ruiz, S. A. El-Seoud, Arthur Edwards, J. M. Alja'am, R. Aquino-Santos, "Integrating the Sense of Smell in an Educational Human-Computer Interface," In *Proceedings of the ICL'8,* Villach, Austria, pp. 1–7, 2008.

[10] Using Visualizations to Tell Stories. http://datajournalismhandbook.org/1.0/en/delivering_data_4.html, Last Access 20-Jul-2016.

[11] J. Dhiraj, Z. James, and L. Jia, "The Story Picturing Engine: Finding Elite Images to Illustrate a Story Using Mutual Reinforcement," in *MIR'04*, New York, pp. 119–120, 2004.

[12] C. Jingyu, W. Fang, T. Xiaoou, "Real Time Google and Live Image Search Re-ranking," in *Proceeding of the 16th ACM international conference on Multimedia*, pp. 729–730, 2008.

[13] R. G. Richards, "Helping Children with Learning Disabilities Understand What They Read", pp. 1–3, 2008.

[14] Z. Tong, D. Fred, J. David, "Text Chunking based on Generalization of Winnow," in *Journal of Machine Learning Research* vol. 2, pp. 616–617, 2002.

[15] X. Shize, K. Liang and Z. Yan, "A Picture Paints a Thousand Words: A Method of Generating Image-text Timeline," in *CIKM'12*, USA, pp. 2511–2512, 2012.

[16] Data Visualization for Human Perception, http://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html, Last Access 20-Jul-2016.

[17] H. Edmundson, "New methods in automatic extracting,", in *Journal of the ACM*, 16(2), pp. 264–285, 1969. https://doi.org/10.1145/321510.321519

[18] J. Kupiec, J. Pedersen, and F. Chen, "A Trainable Document Summarizer," In *Proceedings of the 18th ACMSIGIR Conference,* pp. 68–73, 1995. https://doi.org/10.1145/215206.215333

[19] C. Jingyu, W. Fang, T. Xiaoou, "Real Time Google and Live Image Search Re-ranking," in *Proceeding of the 16th ACM international conference on Multimedia*, pp. 729–730, 2008.

[20] A. Schüler, J. Arndt, and K. Scheiter, "Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations?," *in Learning and Instruction*, vol. 35, pp. 62–72, February 2015. https://doi.org/10.1016/j.learninstruc.2014.09.005

[21] S. Aloraini, "The impact of using multimedia on students' academic achievement in the College of Education at King Saud University," in *Journal of King Saud University - Languages and Translation*, vol. 24, Issue 2, July 2012, pp. 75–82, 2012.

8 Authors

M. U. Mwinyi is with Qatar University, ITS Services Center P.O. Box 2713, Qatar (e-mail: author@ boulder.nist.gov).

J. M. Alja'am, is with the Department of Computer Science and Engineering, Qatar University, Doha, P.O. Box 2713 Qatar (e-mail: jaam@qu.edu.qa).

S. A. ElSeoud is with the Computer Science Department, British University in Egypt (e-mail: Samir.Elseoud@bue.edu.eg).