AUTOMATIC AID SYSTEM TO ENHANCE THE READABILITY OF SCIENTIFIC PAPERS

João Ulisses, Rebeca P. Díaz Redondo, Ana Fernández Vilas

Information & Computing Lab, AtlantTIC Research Center, Universidade de Vigo (SPAIN)

Abstract

Since the amount of scientific papers is continuously growing, researchers and students are usually overwhelmed trying to be updated in their fields. In this paper it is presented the architecture of a new tool, coined Enhanced Paper Reader (EPR), that will support readers in this work by providing smart summaries of pdf technical documents. These smart summaries will combine text and images obtained from the documents in order to provide helpful information for them to decide if the document should be read in more detail or what sections are the most relevant. With the idea of adapting the EPR tool to the students, it would use personal information (learning background, other previous readings, etc.) together with learning preferences (text, image, voice, etc.) in order to display the smart summaries in the most efficient way.

Keywords: Scientific paper readability, enhancing readability, automatic paper analysis, text mining, text-to-speech, education, text analysis.

1 INTRODUCTION

Reading and comprehension is a common problem, especially in adolescence [1], which can diminish their ability to learn. At home, computers and .PDF documents are the most common way to read scientific papers. This can be considered boring, time consuming and, if a large volume of files is to be analysed, also overwhelming for students.

However, students also have more access to technologies such as mobile devices that often compliment their daily life. These devices are commonly used while traveling along with laptops. While using both devices for reading, these require vision only and do not make use the hearing sense, reading a large volume of words at uncomfortable distances can cause eye strain. E-book readers function like smartphones and can provide the most comfortable reading however they lack applications that could enhance the reading of scientific papers further. e-books often have the voice add-in where users can listen instead of reading and some entertainment narrative book services even have professional voice actors to enhance the reading experience.

Thus, with a proper applications, laptops, smartphones and compatible e-book readers (Android) could help keep the students updated in different ways by delivering the content in a more oriented or summarized way. Many techniques can be explored in order to enhance the user interest, readability and information retention. Most techniques are visual as it is easier to have more information at the same time. It can include images with context, more text focused with text summary, special techniques from psychology that give special emphasis to certain parts of the text. If the visual sense is overused, text-to-speech can be used as a non-visual sense, voice techniques can be used to increase attention, and in some cases effectively replace visual content. Moreover, new visual experiences can be created through the usage of image or new navigation types such as virtual reality.

Personal data such as research interests can provide important information on how to adapt certain techniques to the user for a better result. This data can also help tackle specific user issues in readability such as eye strain, blindness and loss of interest. While each specific problem is complex and may require a deeper solution, it may be possible that a more generic simple solution give the ability for the more problematic cases to be able to access the main information easier. This personal data includes different types of data acquisition, the user privacy is a question depending on the types of data needed, the user needs and the user agreement for this data to be gathered, stored and used.

This paper aims to explore the possibilities provided by different techniques and strategies analysing them through testing via use of technology. The objective is to maximize learning and information retention. In order to do so, the work will be focused on developing and evaluating readers and obtaining feedback from multiple reading scenarios.
We present Enhanced Paper Reader (EPR) and the methodology it uses to enhance paper reading in Section 3, followed by the architecture, modules in Section 4. Conclusions and future work are presented in Section 5.

2 RELATED WORK

At a very basic level, e-readers are admired for their convenience, affordability, and accessibility; however, they are also criticized for their complicated features, cause of eyestrain, and general uncertainty when relying on technology. Some researchers taking a more educational angle-fear that the increasing use of e-readers will cause users to merely "skim" rather than read and will cause reader distraction rather than knowledge absorption or contemplation [2].

Text analysis offers crucial data for the tasks that increase readability. One of them is Text to speech (TTS) can enable the system read more like a human, or by simply being more linkable with more contexts and relatable, which may help the user retain information. TTS technology can improve reading performance reading rates and vocabulary as demonstrated in [3]. In [4] students with special needs in this study clearly benefited from the routine use of TTS software to accommodate access to reading passages. In [5] 73 students from across the state of Iowa in United States of America participated in a 23 week study of the impact of the use of a text reader software program on multiple measures of academic performance and concluded that enhanced performance was observed in the areas of reading fluency and comprehension. In [6] tests using TTS on students with learning difficulties were made and results concluded that students in high school benefitted the most from TTS. Additionally, if students with learning difficulties in reading can learn material more efficiently and thoroughly with TTS, perhaps more students will be able to handle higher level curriculum [6].

The structure of the paper gives important pointers on how to classify the importance of the content and can be a strong starting point for the text analysis. A substantial amount of evidence has accumulated indicating that headings in printed texts aid readers' representations of the text topics and their hierarchical organization. However, some might provide simple section labels such as "Introduction" or "Summary", attention to such headings might not produce benefits in note-taking or cued recall. It might also not be necessary to pay attention to the headings in certain types of text, such as narrative texts. The authors in [7] conducted experiments with voice change and how headings were rendered in TTS with the objective of obtaining the attention of the student. In the conclusion the authors suggest that a hybrid rendering that combines both discursive and prosodic rendering is a worthy candidate for future research. The participants in the experiment were college students who presumably were above average in their reading and studying abilities and more experienced with expository texts than other possible populations of participants. Participants who heard a version of the text in which the headings were rendered included more of the text's main ideas in their notes and portrayed better memory for main ideas than participants who heard the text version in which the headings were not rendered. The way in which headings were rendered did not reliably influence note-taking or recall but it did influence how frequently participants paused while listening to the text. Listeners paused less often in the Voice Change condition than in either of the other two conditions in which the headings were rendered, evidently as they could process headings more efficiently when an auditory contrast signaled a heading than when there was no auditory contrast. The ability to control the audio recording could lessen the importance of correctly rendering audio headings. There is a possibility that listeners could capture all headings effectively even without any special rendering strategy if they are allowed to rewind or repeat the audio recording.

LearnSec is framework for full text analysis, which incorporates domain specific knowledge and information in regard to the content of the document sections to improve the classification process with propositional learning and relational learning. To demonstrate the usefulness of the tool, the authors processed a scientific corpus based on OSHUMED, generating an attribute/value dataset for relational learning, and a First Order Logic dataset in Inductive Logic Programming format for propositional learning [8]. The authors used WordNet, Gene Ontology ISPELL to check technical terms and Porter Stemmer algorithm to remove inflectional affixes of words reducing the words to their root, normalizing several variants into the same form [8]. In later similar works, word dependency was analyzed, using tools like Genia Tagger, UMLS MetaMap and Spacy [9].

Images are also another useful source of information in scientific papers, [10] demonstrates that scientific paper’s illustration can help readers to focus their attention on explanatory information in text and to reorganize the information into useful mental models.
Tools like Carrot2 [11], a software of information recovery, clustering and visualization, can help the user understand better the context by allowing the user to amplify or modify the structure of search information. Carrot2 has a collection of algorithms of aggregation such as Lingo, K-means. These facilitate the exploration of thematic context of documents recovered by search engines such as Google, Yahoo, Wiki, PubMed and others.

In [12] has proposed five functions of text illustrations: (a) decoration—illustrations can help the reader enjoy the textbook by making it more attractive (but without being relevant to the text); (b) representation—illustrations can help the reader visualize a particular event, person, place, or thing (such as found commonly in narrative passages); (c) transformation—illustrations can help the reader remember key information in a text; (d) organization—illustrations can help the reader organize information into a coherent structure; and (e) interpretation—illustrations can help the reader understand the text.

In [13] experiments student performance was evaluated on appropriate tests, such as explanatory recall and problem solving, and on inappropriate tests, such as non-explanatory recall and verbatim retention. Illustrations that concretized the changes in status of parts within the system (explanative illustrations) were compared to illustrations that did not (non-explanative illustrations). Results provide evidence concerning the characteristics of the models portrayed in effective explanatory illustrations. The ineffective illustrations failed to visually portray either system topology i.e. steps illustrations or component behavior i.e. parts illustrations whereas the effective illustrations i.e. parts-and-steps illustrations portrayed both. In particular, the parts-and-steps illustrations used a series of two or more frames to show the state of the components within the system at various points in the operation of the system. The explanatory illustrations may also be effective in improving problem solving for low prior-knowledge students.

In [14] the authors present 10 practical suggestions for educators considering text-accompanying illustrations:

1. Select pictures that overlap with text content. Learning benefits occur when pictures and text provide congruent, or supporting, information. Decorational illustrations may help to make the text more attractive or more marketable, but they are unlikely to enhance desired outcomes related to understanding, remembering, or applying the text content.
2. Easy-to-follow texts that are highly concrete and engaging (e.g., interesting narrative passages) readily elicit visual imagery in students and therefore are unlikely to yield additional cognitive benefits from the inclusion of pictures.
3. Prerequisite basic reading skills are required on the part of the student for positive effects of pictures to emerge. At the same time, young children or other students lacking such skills can improve their listening comprehension and recall with well-selected pictorial accompaniments.
4. Choose pictures with an eye toward the desired functions they are to play, representational (to make the text more concrete), organizational (to make the text more coherent), interpretational (to make the text more comprehensible), or transformational (to make the text more codable— and more memorable), in light of the desired learning outcomes.
5. In general, the more complex the text, the more likely that pictures are helpful, in particular explanatory pictures function as useful mental models if (a) the text describes a cause-and-effect system or complex process and (b) the learners are relatively inexperienced in the content domain.
6. To yield the maximum benefits from pictures as text adjuncts, direct students to do something with the picture that yields a controllable product, such as labeling the features of the illustration or structuring the process thereby the students are certain to be constructing veridical pictorial representation of the passage.
7. Computer software that uses integrated or pop-up displays may be more effective than those using split displays in which the picture and text appear in separated locations on the screen.
8. You may also want/need to consider students’ individual learning styles.
9. Realize that even professionally designed pictures and illustrations in textbooks are not necessarily perfect, nor easy for students to comprehend or remember. Thus, even though a particular textbook illustration may be designed to be cognitively useful, it may turn out to be functionally useless unless the learner perceives the illustrated content or process in the intended manner.
Consider the use of transformational (mnemonic) pictures as pictorial adjunct aids to text. An important research priority is to identify the situational characteristics that distinguish between the successes and failures so that more specifically “useful” mnemonic illustration guidelines can be developed.

The technology advances allow us to find much more information as well new forms to deliver this information to students. We then introduce Enhanced Paper Reader (EPR) which will start from the culmination of state-of-the-art methods presented and seeks to test them in an application. Through a usage and feedback system it seeks to improve itself, the methodology is explained in the next Section.

3 METHODOLOGY

In this chapter the methodologies of EPR are presented and how they can improve the reader experience.

3.1 Techniques to decide if a paper is worth reading

In grey literature, PhD program introductions and PhD professors, it is generalized that you should not read every paper that you find, but rather do a first quick analysis to see if the paper is worth being read, before committing to deep reading. This is due the amount of time fully reading and analysing a paper takes. This first analysis may use different techniques, which may also have slightly different objectives, however the main objective of this first analysis is always the same, to see if a paper is worth being fully read. Normally these tasks are generalized to be in the first phase only the title and keywords, abstract, conclusion matter. If a paper is uninteresting after this phase it is discarded. The second phase is the images where the user tries to see if the content is understandable and usable in a meaningful way. Most scientists read the abstract first. Experts in the field prefer to skip right from the title to the visuals as the visuals, what kinds of experiments were done and what results were obtained [15]. The next phases to read a paper are meant when you accept the paper as important for the research and as such the methodology here varies depending on the area, experiments and objectives.

Most journals now require a structured abstract with separate subheadings, allowing the reader quickly to identify the important parts of the study. Most structured abstracts contain the following sections: background or hypothesis, methods, results, and conclusions. While some will read the abstract from the beginning, an efficient reader will begin with the abstract’s conclusions first [16]. In [17] the suggested order to read is Abstract, Discussion, Introduction, Results, (Methods).

With these techniques in mind, EPR can use its modules to simulate or permit the user choose these techniques in order to aid the user to see if the paper is worth reading or not. There are two different ways to do this. If the user wants to perform this task faster in which some EPR capabilities may help or if the user wants to do this in more detail fully using the capabilities of Enhanced Paper Reader. Allowing the user to customize the way he wants to do this may help the user be more productive, for this further customization should be added as EPR evolves and more feedback is gathered.

3.2 Automatic information retrieval

From the title, abstract, keywords, visuals (especially figure and table titles) and first sentence or the last one or two sentences of the introduction it is possible to extract the paper main points [15]. Algorithms such as in [18] can help retrieve the topics from the abstract. However, papers often contain keywords section which provide that data more accurately. In the methodology section of papers, it can be useful to extract software and algorithm names through entity name recognition methods to tune the previous search or for the user summary.

3.3 Summaries

As mentioned in Section 3.1 the user may want to quickly understand what the paper concerns. While abstract and images may provide answers, the user may want to filter content. For this a quick list generated by keywords and important names such as algorithms or software used may be generated. This way not only the user understands what the paper concerns but can quickly discard if the user is not interested in a particular method.
3.4 Finding new research

A tool that recommends new related research inside EPR could be helpful for the user. However, there are some limitations such as some papers are not publicly available. The best way to guarantee a paper is available is to search for it on search engines and dedicated engines such as Google Scholar, which searches most scientific paper databases. The search could start by trivially looking at cited papers and see which are available through google scholar and then do an analysis, at first just looking into keywords, to see if they are appropriate. This method could also be used to search and add information from cited papers into the paper that is being read. As mentioned in Section 3.3, the user could filter the content in order to discard content that he does not want to see, tuning the search to his needs.

3.5 Information Adder/Completer

As mentioned in Section 3.3 and Section 3.4 it is possible to add more information from other similar papers, specially the cited papers. The information gathered would all be tuned to fit the user input information and research topics as mentioned before. This way the reader would rarely lose available nearby information that could be in another document. Other sources of information could be Wikipedia, Balbelnet and Carrot2 which can enrich the content. This is especially useful in certain areas for example to allow to see how an algorithm works, fully understand math formulas or concepts and then diving back into the paper. This can currently be done by using search engines manually, but it can be disruptive and distracting. The mentioned tools provide ways to link content into more interesting relevant content that can be helpful. As for being disruptive a more immersive way to present this type of content is needed therefor the reader does not have to leave the paper to find the proper explanations he needs.

3.6 Slow Strong information retention through Virtual Reality

In the last two years Virtual Reality (VR) received a boon in terms of research interest and usage, more devices and technology became available to the public with reduced cost and more applications were available. For EPR this enables an opportunity to deliver content in a new powerful meaningful way, to explore visual memory to enhance how paper content such as images or math, chemistry and other formulas can be explored. By escaping the monotony of the 2d papers the readers may feel more immersed and create more memorable experiences which benefits their memory and information retention. It may also enable to have discussions with other VR simulation users in the same simulation, discussing about the paper, classroom in regards the paper, pointing out parts of the paper to others.

VR interfaces and navigation are often tied to game design. For EPR the interface can be similar to most PDF/Book readers, where you click to go to next page, zoom, search and other utilities. However, there is the drawback of taking out the novelty feel that comes with VR and for some others it may feel less productive since it is the same interface. The navigation can also be done in a 3D generated space, like a classroom, which is also better for multi users this is a more natural way to navigate but may be less comfortable as the user is not able to lay down. Another possible navigation is to limit what the user virtual location to just the paper and its additional information that can be extracted from related papers or searches as mentioned in Section 3.5. This way to navigate can be more comfortable and allowing the user to lay down and look to the roof while in VR he is seeing all the content. Other possibility that can be mixed is to navigate in a similar way to Pretzi where the user jumps into other page or content.

4 ARCHITECTURE MODULES

The architecture is defined as Figure 1 displays, to cover all the aspects we would like to tackle with the aim of helping students and researchers to process scientific documents.
EPR modules are made to deal with the amount of data scientific papers have in different phases. These modules can be improved as new techniques are found, each independently or removed depending the objectives of the analysis or output. Not all modules are used everytime. Now it is presented a brief description of the most important modules.

**Paper PDF** – The .PDF file is to be analyzed by PDF reader. It always contains Authors, keywords, abstract and references. The body can be divided into sections, these have multiple formats but can be detected by PDF reader, their titles and divisions can help detect the subject and when that subject starts and ends. Normally these sections include introduction, the section right after abstract and conclusion or discussion, often located right before the references. These sections are of the utmost importance to characterize the paper.

**PDF Reader** – PDF code to read the raw PDF and detect portions of text, such as sections. PDF Reader will create divisions to better help the automatic analysis of the paper.

**Folder of Papers** – Multiple Papers in PDF format, each folder normally has one topic depending on the user organization. A special analysis on these papers can be made if they are analyzed in bulk rather than one by one.

**Personal Data** – Each user has multiple folders of papers depending on their research topics. Each user can also have their keywords of interest which work like hashtags and can be more useful than the research topics for analysis purposes. This data can change the way the papers are analyzed and give more tailored data to the user.

**Automatic Paper Review** – The papers can be analyzed to give a paper review. For this, two parts are necessary, Text Analysis and Semantic Analysis. The text analysis includes grammatical errors with the usage of a grammar library, however another way can be done without one, which is to catch the most common mistakes. Early students often commit the same mistakes such as poor formal language, usage of the first person and mistakes in the reference formats. These errors can be reported, in which they can serve as part of paper review. The second part of semantic must be done by human, however summarization or division done in PDF reader can help the user review faster the paper, there are many techniques for this such as reading the introduction, then conclusion then confirming the methodology. Language processing libraries can also be of use, and the semantic analysis can also use a lot of context analysis and feeds from image/video for a faster analysis by the user. These are important as nowadays students and reviewers are required to read an increasing number of papers, which can cause eyestrain and loss of information.

**Generate Text Summary** – Using the paper divisions given by PDF reader a summary can be more easily created, matching with the interests of the user, or alternatively what is most common topic or
keyword in the paper. Contexts can be built from the analysis of the text to help generate a text summary.

**Reading Experience** – The information that the Text Summary outputs should be read by the user. This text should be using at least one reading technique that benefits the user. Psychology claims there are many techniques to be used, which depends if its visual text or voice, such as reading only paragraphs with keywords, reading only sections with keywords, reading only sentences with keywords, reading only certain sections and other. There are also small programs using Data Mining techniques from the text directly that ignoring stop words, such as K-NN can help understand what a certain section concerns without knowing the title. Machine learning techniques also exist; however, these require either a lot of human data to start with, and then create variations, which then again must be tested by human.

**Context Tagger** – Creates tags of context for the text in Text summary, which can help find similarities in context. These can come from Wiki, Babelnet, Wordnets, Carrot2. Google scholar can also be of help depending on the type of context and similar papers. These context tags can also generate short descriptions which can also be of use for the Text Summary, or also for the voice Text-to-Speech.

**Text-to-Speech** – The non-visual way to send data to the user. This method is slow but can have bigger impact when the visual sense is saturated, or the user eyes are tired. Must not be overused or the voice will read a large text slowly making the user lose focus or annoyed. Webspeech API or Android have good text-to-speech libraries that can be used, some even provide multiple voice actors, which can be useful to regain attention of the user.

**Visual Content** – Images and Videos can provide a more pleasing way to read a paper, specially to younger students. Text causes eye strain and can scare some younger students. Images and Videos can replace some text making the information less bulky and easier to attain, faster and can invite the user to read more. The right image or video can come from context and text summary and it can provide a powerful way to drive how the text is displayed to the user. Other idea is to display the visual content similar to Pretzi and use a virtual reality ambient type of navigation.

**Finding New Research** – New research can be found with citations, keywords, algorithm names and method names. These can be searched on google scholar and other scientific paper repositories to find publicly available. PDFs.

**Finding New Content** – New content can be found through finding new research module results, or simply through keywords and context tags. These can be used on Wikipedia, Babelnet and Carrot2 to demonstrate interesting new information to add.

**Flow of Content** – The way the information should be displayed to the user in order to keep the user interested in reading, but also fast enough for the user to not be bored. Images can provide some variety for the eyes, while voice text-to-speech can be complimentary optional to help capture the most important content. Each segment should help guide to the next one, image or text and each should be small enough, however opportunity for the user to explore a certain topic that he chooses is of utmost importance and the whole experience needs to be reshaped, as that new information can be crucial to understand what the user wants to know more about. The navigation of content depends of the content type. Virtual reality can offer a unique way to set the flow of content with less bounds.

### 5 CONCLUSIONS AND FUTURE WORK

Our proposal is enriching the framework of smartphones, laptops or even e-book readers devices to help students to keep updated in different ways to deliver scientific content in a more personal and comfortable form. Many techniques can be used in order to enhance the user interest, readability and information retention and the way and form the information should be displayed is important. Technologies such as text-to-speech may offer complimentary help or replace some parts. Videos may also help deliver the information fast, therefor the user is more focused in a short period of time. How to create these from the information available in scientific papers is precisely explored in this paper. What is more, personal data can be collected, such as research interests or previous readings, which can provide important information on how to adapt certain techniques to the user for a better result. Readers are unable to read a vast majority of papers depending on their languages. Some of those papers may have crucial information for their research. Removing the language barrier by translating the paper helps solve this issue. For this automatic translator such as google translator
may be the fastest and additionally cheapest solution, however the accuracy is not perfect and
important parts may be lost in the translation. To increase this accuracy automatically a more in-depth
study of the paper context must be made in order to help the translator make the right decisions when
translating. Image processing may allow to receive more information from the images in a paper and
link it with other images to give the user more data. However, an expert may not benefit as much, and
this is more tailored to enrich the unexperienced user exploration. Paper reviewers may also benefit of
EPR and most of its methodologies and techniques described in section 4, which in turn may provide
quick feedback on how to improve the tool or techniques used. Multiple scenarios for students,
readers and reviewers were detected, in which some may prefer comfort, others deep analysis and
others just a quick information check, however each requires a different set of techniques and
methods which may prove difficult to solve all at once. In an era of multitasking and travelling, hearing
ability can be exploited as a means to gather new information on the go or while performing other
tasks without any impairment. This is especially common in researchers that must read a considerable
number of scientific papers for their research or conference paper reviews and are often travelling to
conferences or multitasking. With the proposed architecture, EPR seeks to facilitate paper reading and
enhance reading experience with new methods while also mixing with well-known well established
already existent methods. Through experience and feedback from users EPR techniques may be
improved. The current state of virtual reality allows users to experience virtual worlds on mobile
devices and this can be used to enrich knowledge gathering experiences for students and
researchers.

REFERENCES

fluency and vocabulary difficulties among adolescents struggling with reading comprehension,” J.

Michigan, vol. 27, no. 1, p. 9, 2011.


study,” in Handbook of Research on Human Cognition and Assistive Technology: Design,


[6] M. C. Young, “The Effects of Text-to-Speech on Reading Comprehension of Students with

2018.

“LearnSec: A Framework for Full Text Analysis,” in International Conference on Hybrid Artificial


1981.

Psychol., vol. 82, no. 4, p. 715, 1990.


