



HAL
open science

READ: A (Research) Platform for Evaluating Non-visual Access Methods to Digital Documents

Laurent Sorin, Julie Lemarié, Mustapha Mojahid

► **To cite this version:**

Laurent Sorin, Julie Lemarié, Mustapha Mojahid. READ: A (Research) Platform for Evaluating Non-visual Access Methods to Digital Documents. 17th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2015), Oct 2015, Lisbonne, Portugal. pp.431-432. hal-02494132

HAL Id: hal-02494132

<https://hal.science/hal-02494132>

Submitted on 28 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in:

<http://oatao.univ-toulouse.fr/24774>

Official URL

DOI : <https://doi.org/10.1145/2700648.2811338>

To cite this version: Sorin, Laurent and Lemarié, Julie and Mojahid, Mustapha *READ: A (Research) Platform for Evaluating Non-visual Access Methods to Digital Documents*. (2015) In: 17th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2015), 26 October 2015 - 28 October 2015 (Lisbonne, Portugal).

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

READ: A (research) Platform for Evaluating Non-Visual Access Methods to Digital Documents

Laurent Sorin
University of Toulouse
UPS & UT2J; IRIT
118 route de Narbonne
31062 Toulouse
(+33)5.61.55.77.02
sorin@irit.fr

Julie Lemarié
University of Toulouse
UT2J; CLLE
5 allées Antonio Machado
31100 Toulouse
(+33)5.61.50.35.41
lemarie@univ-tlse2.fr

Mustapha Mojahid
University of Toulouse;
UPS; IRIT
118 route de Narbonne
31062 Toulouse
(+33)5.61.55.63.18
mojahid@irit.fr

ABSTRACT

READ (which stands for **D**ocuments **A**rchitecture **R**estitution, in French), is a software allowing *augmented* access of tagged documents content. It was originally developed for a project aiming at evaluating informational and functional equivalents to text visual formatting [8] and was very recently tested in a user study (results unpublished yet). The design flexibility of this platform allows an easy implementation of input and output modalities, as well as navigation functions. We wish to publish this application for noncommercial reuse in the short run; the main goal is here to easily allow researchers to evaluate non-visual access methods to documents, since current assistive technologies are known to raise multiple issues [7]. A live demonstration of READ will enable attendees to experience the implemented output modalities and navigation functions, as well as the platform configuration and extension potential.

Keywords

Digital accessibility; non-visual interactions; documents accessibility; research platform

1. INTRODUCTION

READ was initially developed inside the frame of a project aiming at finding new access methods to restore information and functions provided by text formatting during visual reading [8], in order to improve digital documents accessibility. The terms restoration and restitution used here refer to the possibility for visually-impaired users to access the same information and as efficiently as non-disabled users. Regarding documents accessibility, much research work remains to be done.

Indeed, if we take the example of the screen readers as the main tool used by visually-impaired people to access digital content, even though they present overall an extensive list of functions, their use remains frustrating and poorly efficient [6], even when accessing content following accessibility guidelines [7].

Audio and tactile linearity inherent to those modalities are often mentioned as an important part of the overall problem. Several initiatives emerged to provide non-linear access to content, such as concurrent speech [5] or aural glancing of texts [4]. Multimodality is also pointed as a solution to improve content accessibility regarding the potential information overload raised by linear access to information [3]. This kind of approach is also applied for accessing complex content such as mathematics [1].

The evaluation of those paradigms implies developing dedicated tools or modifying existing open source screen-readers. For the latter option, since they need to handle interactions with the operating system, screen-reader have a complex architecture to build on; READ can provide a simpler alternative when the focus is on documents accessibility. What we propose with READ is a software to reuse as a base for those evaluations, requiring few to none developments depending on the project. It provides a controlled MS-Windows environment for documents non-visual access, with customizable inputs and outputs. READ will be published as soon as possible under the CC BY-NC-SA¹ license.

2. SYSTEM DESIGN

Our general aim was to develop a fully configurable system with a modular design, based on XML² configuration files and plugins.

-Technical Specifications: The system was developed in C#, as .NET provides a library to use installed synthetic voices following the SAPI³ norm. Any computer with a Windows OS with the framework .NET 3.5 (or more recent) can run the application. Application configuration is defined in XML files; XML was an obvious choice for flexibility and interoperability.

-Global architecture: The application was developed using a “blocks and events” paradigm; the document is turned into blocks of information, each block including one or several event, which can be played sequentially or in parallel depending on the configuration.

¹ Creative Commons Attribution (by) NonCommercial ShareAlike

² Extensible Markup Language

³ Speech Application Programming Interface

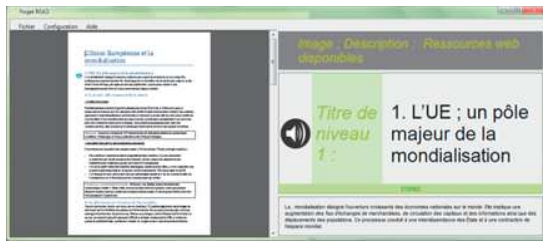


Figure 1. Screenshot of the READ system

The system also uses “relationships” between blocks, specified in a separate XML file; relationships can be visual relationships or discourse relationships, such as one block being an illustration of another. Those relationships can be used to implement navigation functions, for instance.

Finally several “modes” of access can be defined, by writing different configurations files. This allows switching between modes during navigation.

-*Configuration overview*: the configuration files (one by defined mode) basically specify how to render each tag of the original XML document (e.g. when encountering a `<p>` in the document, play a sound); note any vocabulary can be used to tag the original document (e.g. HTML or custom tags) as long as this vocabulary matches the one in the configuration files.

Configuration files also allows defining several types of events (reading text with Text-To-Speech, playing sounds, etc.) using implemented output capabilities with specific configurations (e.g. reading text with a specific voice, at a specific rate or pitch, etc.).

-*Implemented outputs modalities and functions*: Currently, READ can use any SAPI voice to read text, which can contain SSML⁴ tags to control its parameters. Sounds store in WAV files (e.g. auditory icons) can be played, and any output sound or voice can be spatialized, that is to say the sound is played as if it came from a certain location in space (see [2] for an example of spatialized sound use for accessibility). READ also provide a basic visualization of the events being played (*on Figure 1. the original document is on the left and current events on the right*).

Navigation functions to move between blocks, jump to a given block from its type or relationships, are also implemented. For data analysis purposes, every user and system action is logged into a text file.

-*Implementing new extensions*: If other input or output devices need to be implemented, the application provides a plugin mechanism to do so. One can then add C# plugins to define new types of events, which can be afterwards used in the configuration files. Plugins defining new navigation functions can also be added to the software. Writing a new plugin consists in writing a code file implementing the defined programming interfaces, and placing it in the right folder.

3. CONCLUSION

This platform could be used to implement and combine several input or output devices (e.g. haptic devices) for multi-modality or non-linear rendering of documents. The applications range from the evaluation of navigation performances to content comprehension and learning. We believe READ will be useful to

address this type of research questions in the frame of documents non-visual access.

Regarding the limitations of this software, the XML configurations files, the relationships files and the documents have to be manually edited. This can be a rather time-consuming process depending on the document size, yet it seems acceptable for a research usage.

We recently used READ to evaluate different rendering and navigation methods for textual documents in an experiment with 60 blindfolded and 12 blind participants; their main task was to understand a 1500 words structured text⁵. Even though the scope of the experiment was different from the software evaluation in itself (rendering methods were evaluated), it is important to note that READ was stable and bug-free for every participant. Besides, the usability scores obtained with a SUS⁶ questionnaire were on average around 90 (over 100) for the best methods, showing that READ can be used to implement usable interactions.

4. ACKNOWLEDGEMENTS

This research was funded by the PRES-Toulouse and the Région Midi-Pyrénées.

5. REFERENCES

- [1] Bates, E. and Fitzpatrick, D. Spoken mathematics using prosody, earcons and spearcons. *ICCHP 2010*, (2010), 407–414.
- [2] Cofino, J., Barreto, A., Abyarjoo, F., and Ortega, F.R. Sonically-Enhanced Tabular Screen-Reading. *International Technology and Persons with Disabilities Conference*, (2014).
- [3] El-glaly, Y.N. and Quek, F. Digital Reading Support for The Blind by Multimodal Interaction. *ICMI'14*, (2014), 439–446.
- [4] Gadde, P. and Bolchini, D. From Screen Reading to Aural Glancing : Towards Instant Access to Key Page Sections. *ASSETS'14*, (2014), 67–74.
- [5] Guerreiro, J. and Gonçalves, D. Text-to-Speeches : Evaluating the Perception of Concurrent Speech by Blind People. *ASSETS '14*, (2014), 169–176.
- [6] Lazar, J., Allen, A., Kleinman, J., and Malarkey, C. What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users. *International Journal of Human-Computer Interaction* 22, 3 (2007), 247–269.
- [7] Power, C., Freire, A., Petrie, H., and Swallow, D. Guidelines are only half of the story: accessibility problems encountered by blind users on the web. *SIGCHI Conference on Human Factors in Computing Systems*, ACM (2012), 433–442.
- [8] Sorin, L., Mojahid, M., Aussenac-gilles, N., and Lemarié, J. Improving the accessibility of digital documents for blind users : contributions of the Textual Architecture Model. *HCI International 2013*, Springer-Verlag Berlin Heidelberg (2013), 423–431.

⁵ The results will be published soon; the aim was to evaluate different text formatting restitution methods

⁶ Standard Usability Scale

⁴ Speech Synthesis Markup Language