STS on Accessible and Inclusive Digital Publishing

Introduction to the Special Thematic Session

Reinhard Ruemer and Valentin Salinas López

1 Verein BookAccess, 4040 Linz, Austria
reinhard.ruemer@bookaccess.at
2 Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria
valentin.salinas_lopez@jku.at

Abstract. The special thematic session on Accessible and inclusive Digital Publishing consists of a wide range of publications in this area. It will show efforts undertaken to understand the role of accessibility within a company, practical methods for accessing and structuring digital content and ways of improving the evaluation of digitally created content.

Keywords: Digital publishing · Accessibility · World Wide Web · EPUB · PDF

1 Introduction

In a perfect world, we would see the considerations regarding the accessibility of the published content considered in every step of the publishing process. From the authoring and design phase to the technical implementation and all stages of production and marketing.

In the real world, we are far away from such a situation. Digital publishing still has a way to go to become an inclusive process. Although, there are several well-documented resources available regarding the accessibility e.g. the WCAG-Guidelines [1] and there is also commitment from governments regarding the accessibility of digital published content [2, 3] in the day to day production process accessibility is still not as much implemented as it could be. There are many reasons why this is happening. At the top is the lack of knowledge on the practical side of accessible publishing by the responsible actors. As a result, a large share of the digitally available documents is not - or only on a very shallow level - accessible. Just to illustrate the magnitude of the problem, in 2017 the European Blind Union estimates that the share of accessible books for visually impaired persons (VIP) ranged between 7% and 20% in the EU. In recent times, a number of initiatives with different approaches have emerged to address this issue. One of them is the co-funded by the Erasmus + Programme of the EU: Supporting Inclusive Digital Publishing through Training (SIDPT). The aim of this project is supporting the publishing industry in the creation and delivery of accessible digital publications, in order to help it to comply with new legislative context of accessibility in digital publications. To this end, it focuses on the provision of practical
accessibility training material for professionals working in the publishing sector. It covers most of the publication process from content creation and design to production and marketing of digital works. Due to the complexity of the publishing process and the wide variety of actors involved, the SIDPT’s partners conducted a survey through publishers in Austria, France and the Netherlands to get a better understanding of the current situation and actual needs. Below, there are some interesting findings regarding to the actual digital publishing, accessibility awareness in the sector and topics in which they are interested the most.

Digital publishing is widespread through the publishing industry due to more than three-quarters of respondents said the organizations they work for already publish digital works. Another noticeable fact is the popularity of digital formats, which raise some accessibility concerns, like PDF or EPUB2 (Fig. 1).

![Fig. 1. Popularity of digital publishing format among respondents.](image)

Over 50% of participants said their organization had not implemented yet digital accessibility in their workflow, but they were working on it. The most hopeful point is that many of them already implemented actions that affect positively on the accessibility of their outcomes (Fig. 2).

![Fig. 2. Tasks that respondents already complete and affect positively to accessibility.](image)
Regarding to accessibility awareness, over 50% consider accessibility to be a social and moral responsibility. However, 30% answered that they did not start to integrate accessibility in the publishing workflow yet (Fig. 3).

Only a few publishers have received some form of accessibility training, and these are mostly respondents intervening towards the end of the book production phase. This fact is depicted in Fig. 4, where is possible to notice the few people who answered they got some sort of training and most of them work in production (Fig. 5).

**Fig. 3.** Status of the production of accessible publication through publishers.

**Fig. 4.** Accessibility trainings taken by respondents and professional profiles.
Finally, the SIDPT project decided to produce training material tailored for the different professional profiles and present it ordered and labeled in order to provide quick and easy access to lessons and resources. The distinctive feature of the project is that the production of training material is based on the findings of the described survey. It will target the different needs of the different user groups identified. In the coming months, content will be produced with the core goal of being a simple and comprehensive resource for learning to publish accessible digital works.

In this STS we will hear other different approaches to tackle the issue of accessible digital publishing and the access to the published material. The following section will

---

Fig. 5. Most interesting topics though participants.
give a short overview on how the authors tackle different issues in the process of making content available to people with special needs.

2 Content

2.1 Practitioner’s Perception of Web Accessibility: Roles, Process and Scope

Providing usable web information and services to as many people as possible confronts web practitioners with a demanding task. The present study provides insights on how Web Accessibility is perceived in practice. A total of 163 web practitioners in various professional roles reported their perspective on roles primarily responsible for Web Accessibility, on key phases in the development process and on types of disabilities mainly considered. Results show that non-technical roles are perceived to be less involved, that Web Accessibility is mainly considered in the conceptual design and development phase only, and that efforts focus on people with visual impairments. Based on these findings, recommendations for research and practice are discussed.

A Multi-site collaborative sampling for Web Accessibility Evaluation

Many sampling methods have been used for web accessibility evaluation. However, due to the difficulty of web page feature extraction and the lack of unsupervised clustering algorithm, the result is not very well. How to optimize the manual workload of different websites under the premise of ensuring that the overall manual workload remains the same during multi-site collaborative sampling is an important issue at present. To solve the above problems, we propose a Multi-site collaborative sampling method to obtain the final sampling result of each website. The effectiveness of the two sampling methods proposed in this paper is proved by experiments on real website datasets [4].

2.2 A Multi-site Collaborative Sampling for Web Accessibility Evaluation

Many sampling methods have been used for web accessibility evaluation. However, due to the difficulty of web page feature extraction and the lack of unsupervised clustering algorithm, the result is not very well. How to optimize the manual workload of different websites under the premise of ensuring that the overall manual workload remains the same during multi-site collaborative sampling is an important issue at present. To solve the above problems, we propose a Multi-site collaborative sampling method to obtain the final sampling result of each website. The effectiveness of the two sampling methods proposed in this paper is proved by experiments on real website datasets [5].
2.3 Towards More Efficient Screen Reader Web Access with Automatic Summary Generation and Text Tagging

Readers with 20/20 vision can easily read text and quickly perceive information to get an overview of the information within a text. This is more challenging for readers who rely on screen readers. This study investigated factors affecting successful screen reading in order to shed light on what contributes towards the improvement of screen reading access. Text extraction, summarization, and representation techniques were explored. The goal of this work leads to the development of a new summarization technique, referred to as On Demand Summary Generation and Text Tagging (ODSG&TT). This technique makes use of a summarization algorithm and a text tagging algorithm developed by Algorithmia, which enables on the fly and on-demand summarization of text and keyword generation. The focus of the screen reader is transferred to the keywords using a button control. The intention is to provide summaries with minimum user navigation effort to simplify the screen reading process [6].

2.4 A Series of Simple Processing Tools for PDF Files for People with Print Disabilities

This paper presents a simple tool to process PDF files for people with print disabilities. They consist of the following three titles: “PDFcontentEraser”, “PDFfontChanger” and “PDFcontentExtracter.” PDFcontentEraser is a tool to remove a certain type of elements in a PDF file. PDFfontChanger is a tool to replace the selection of fonts in a document. By default, all fonts are replaced by “Universal Design Font for Digital Textbooks” (UD font) developed by Morisawa Inc. PDFcontentExtracter is a tool to retrieve the components of a PDF file. It can output in the following three forms: (1) character information, image location and path location in XML, (2) image files, and (3) the shapes of paths, the shapes of characters and the locations of images in SVG [7].

2.5 Layout Analysis of PDF Documents by Two-Dimensional Grammars for the Production of Accessible

This paper proposes the use of two-dimensional context-free grammars (2DCFGs) for layout analysis of PDF documents. In Japan, textbooks are considered to be the primary teaching instruments that are necessary to guarantee the quality of education, and audio textbooks have been available for students with print disabilities in compulsory education. In order to create any type of audio textbooks, it is necessary to obtain the information of structure and the reading order of documents of regular textbooks in PDF. It is not simple task because most PDF files only have the information how to print them out, and page-layouts of most textbooks are complex [8].
3 Conclusion

As we could see, there are many different approaches to make published content accessible. Still many of the presentations and the underlying work done is connected to the adaptation and processing of already published material. We would like to see more research and efforts put into the actual information creation process. The potential benefit of the principles of multi-channel publishing put into practice is much greater and can save a lot of work and time spent for making content afterwards accessible.

References

1. Web Content Accessibility Guidelines (WCAG) 2.0 Homepage. https://www.w3.org/WAI/GL/WCAG20/. Accessed 18 June 2020
2. Section 508 Homepage. https://www.section508.gov. Accessed 18 June 2020
3. Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled. Homepage https://wipolex.wipo.int/en/text/301019. Accessed 18 June 2020
4. Vollenwyder, B., Opwis, K., Brühlmann, F.: Practitioner’s perception of web accessibility: roles, process and scope. In: Miesenberger, K., Covarrubias Rodriguez, M., Penaz, P. (eds.) ICCHP 2020, LNCS, vol. 10896. Springer, Heidelberg (2020)
5. Yu, Z., Bu, J., Shen, C., Wang, W., et al.: A multi-site collaborative sampling for web accessibility evaluation. In: Miesenberger, K., Covarrubias Rodriguez, M., Penaz, P. (eds.) ICCHP 2020, LNCS, vol. 10896. Springer, Heidelberg (2020)
6. Sarwar, U., Eika, E.: Towards more efficient screen reader web access with automatic summary generation and text tagging. In: Miesenberger, K., Covarrubias Rodriguez, M., Penaz, P. (eds.) ICCHP 2020, LNCS, vol. 10896. Springer, Heidelberg (2020)
7. Nakamura, S., Kohase, K., Fujiyoshi, A.: A series of simple processing tools for PDF files for people with print disabilities. In: Miesenberger, K., Covarrubias Rodriguez, M., Penaz, P. (eds.) ICCHP 2020, LNCS, vol. 10896. Springer, Heidelberg (2020)
8. Kohase, K., Nakamura, S., Fujiyoshi, A.: Layout analysis of PDF documents by two-dimensional grammars for the production of accessible textbooks. In: Miesenberger, K., Covarrubias Rodriguez, M., Penaz, P. (eds.) ICCHP 2020, LNCS, vol. 10896. Springer, Heidelberg (2020)