

**Rochester Institute of Technology  
B. Thomas Golisano College  
of  
Computing and Information Sciences**

**Master of Science in Human-Computer Interaction**

**~ Project Proposal Approval Form ~**

**Student Name:** Rahul Jaiswal

**Project Title:** Evaluating Using LaTeX to help improve the accessibility of an open-access journal

Project Area(s):    HCI                      Accessibility                      HTML/PDF  
(Ö primary area)

~ MS Project Committee ~

Name

Signature

Date

Prof. John-Paul Takats  
Chair

Prof. Stephen Jacobs  
Committee Member

# **Publishing for all: Using LaTeX to help improve the accessibility of an open-access journal**

**By: [Rahul Jaiswal](#)**

Project submitted in partial fulfilment of the requirements for the degree of Master of Science in **Information Sciences and Technologies**

**Rochester Institute of Technology**

**B. Thomas Golisano College  
of  
Computing and Information Sciences**

**Department of Information Sciences and Technologies**

# Table of Contents

## Contents

Table of Contents.....	3
Abstract.....	4
Introduction.....	4
Problem Statement.....	5
Project Objectives.....	5
Literature Review.....	6
Accessible formats.....	7
PDF vs HTML.....	8
Approaches used for generating accessible articles.....	9
Indesign.....	13
Adobe Acrobat Reader.....	16
Online Conversion Tools.....	17
Latex to HTML.....	17
Usability Testing.....	18
Challenges.....	25
Limitations.....	26
Conclusions and Future Scope.....	26
References.....	32

# Publishing for all: Using LaTeX to help improve the accessibility of an open-access journal

## Abstract

A screen reader is a vital tool that helps individuals who are blind or have low vision read digital text. Unfortunately, not all file formats receive the same level of support from screen readers. For example, while PDF files have accessibility features that can be used, they are often not the preferred file format for screen reader users. Between line breaks, multiple columns, symbols, and images, screen readers often struggle with academic journal articles in certain file formats. Open@RIT collaborated with an open-access journal and their combined goal is to improve accessibility and readership for all. This proposal will showcase an approach to make these submissions more accessible using open-source tools and how adding more features to this solution could make it more accessible to the end-users when combined with the usability testing results for the software.

**Keywords:** Accessibility, Screen reader, Latex, PDF,

## 1. Introduction

According to a report by IAPB [1], 43 million people are blind, and 295 million people experience moderate to severe vision impairment. This is a big number, still, most of the research articles published online mostly target the audience with no visual impairments. Accessibility of the document is sometimes overlooked by the visual look and the format of the documents. The visually challenged audience mostly uses screen readers, which can convert text to reading the articles. As per the accessibility developer guide [2], “NVDA (72.4%) and JAWS (61.7%) are the most popular desktop/laptop screen readers, followed by Voiceover/macOS (47.1%).” The screen readers best work with the HTML web pages because of the tagging.

Although these screen readers do a good job in reading these articles, there are a few issues which screen readers cannot solve in most cases if not explicitly taken care of while editing the document in the required format. Some of the HTML files still have many issues that made them difficult for screen readers to read, such as pauses and interruptions. Also, in a few scenarios when the pdf is passed through conversion tools which can convert pdf into HTML format, some important factors like tagging, heading levels and reading order were not interpreted properly. This results when passed through the screen readers like NVDA and

JAWS, which might be really confusing for the visually impaired audience.

To solve this problem, we decided on using latex to convert non-accessible word/pdf into more accessible pdf. The next step was to deliver an HTML page for a screen reader. This HTML page is converted from the pdf file (exported from latex) using the scripts written in javascript. Some internal CSS is also applied to the script to ensure the consistency in the article's look and feel for the other users who are not visually impaired while keeping in mind that the accessibility of the document is not affected.

### **Problem Statement:**

**Project Objectives:** To convert inaccessible pdf/word into accessible pdf and HTML for the visually impaired audience with minimal changes in the existing design of the JESSED conference format.

1. Increase accessibility features in the existing solution.
2. Conduct usability testing on the proposed solution which converts Docx/pdf to an accessible HTML webpage.

## 2.Literature review

There have been a lot of efforts in the past for making the conference paper accessible to visually impaired users. Adobe Acrobat Portable document format (PDF) has introduced an accessibility tool within the application to make pdf files accessible, it also has an inbuilt accessibility checker which gives accessibility reports helpful for fixing accessibility bugs. In a 2011 article [3], the author showed in detail how you can Adobe PDF introduced accessibility in the PDF formats and how it can be used to make the pdf accessible.

In the 2015 article [4], the author tried to display the comparison of accessibility of papers published in conferences of CHI, ASSET, W4A(technical) and W4A(communication) in the year 2012, 2013, and 2014. The results were shocking, apart from a paper published in ASSET conferences, all the papers in the other conference were not even completely tagged. The percentage was below 20 per cent. It also showed even the paper published in conferences like ASSET is not completely accessible. This article explained that most of these accessibility issues are due to the need for the manual work the researcher must do to make an article accessible. It showed how using Latex templates and accessibility packages will help generate better accessible pdf. This paper also explained using HTML for screen readers will be better than using a pdf.

Latex is another editor who is very popular among researchers nowadays. Latex can export the writings in pdf. In 2019 article [5], Accessible LaTeX Based Mathematical Document Authoring and Presentation (ALAP) tool and in 2020 an extension to ALAP [6], Automated Generation of Accessible PDF(AGAP) was introduced by a group of researchers which helped visually impaired audience to read and write mathematical document. These tools are one of the biggest steps taken toward making an editor which can help authors to write accessible documents. This tool also helped the user to inform about accessibility errors and warnings during the compile time. Although ALAP and AGAP are a great tool for accessibility, it has some drawbacks. AGAP uses the “accessibility” package for Latex which was deprecated in 2020. Also, this software’s supported only on windows and had to be installed on the system before use.

There have been many other promising ideas for making the research publication accessible to most of the audience which includes the visually impaired audience too. Our idea is to generate an accessible HTML page as an output because the screen reader read the HTML page better than the pdf because of the already present tagging in HTML files. This tagging is also easy to edit/modify using JavaScript and CSS.

## 3.Accessible formats

This study uses HTML and PDF as formats for testing the accessibility of the document. We have used these two formats because we are focusing on articles published in research

conferences. We surveyed to find out which format is mostly used by the researchers when they search for research articles online. Below is the result that got from the survey among 43 responses.

What is the format you see most often when you search for a research article?

43 responses

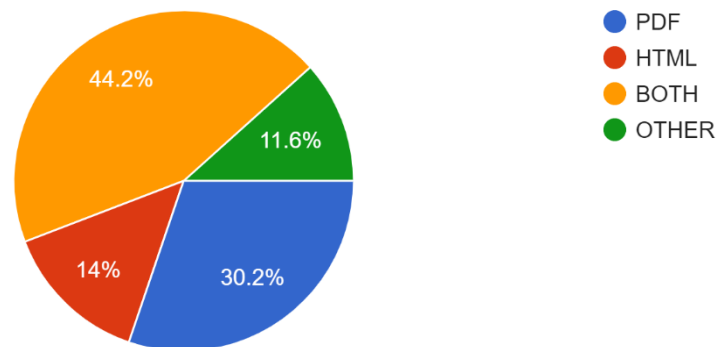


Figure 1: Survey result showing what format researchers mostly use to read their documents online.

We found out that 88.4 per cent of the people either read the research article in PDF or HTML format.

### 3.1 PDF vs HTML

PDF is the most common format among researchers. There are different merits and demerits of using PDF as your primary format.

PDF is not the best option friendly when it comes to mobile-friendliness since most of them are not responsive while Most of the HTML nowadays is responsive and the format can be adjusted to any screen.

Also, most of the PDFs available on the internet today are not tagged and tagging those documents requires knowledge of the Adobe acrobat tool. The abundance of tools to convert untagged PDFs into tagged PDFs makes the job much more difficult. While HTML documents are mostly tagged and easily read by browsers and screen reader tools like NVDA.

## 4. Approaches used for generating accessible articles for screenreader

So, our goal was clear that is to provide accessible file formats that are accessible for all our readers. We started exploring different approaches that can help us achieve that:

- Adobe InDesign.
- Adobe Acrobat Reader

- Latex and HTML

## 4.1 InDesign

We started with Adobe InDesign as a tool to make the articles accessible. (Either in PDF or HTML format).

Adobe InDesign is a tool which RIT publication was already using for most of the printing material. But the challenges that we faced using this InDesign was that the process involved different steps and if any one of the steps is missed, we might have to start doing things all over.

Pros:

- Efficient in generating accessible PDF

Cons:

- Not very efficient in generating accessible HTML.
- InDesign requires a long sequence of steps which is needed to be followed for making an accessible PDF.
- If any of the steps are missed or done wrongly, it might not generate an accessible document.
- It has a complex learning curve.
- Software is paid and expensive (US\$239.88/yr.)
- Have to be installed before use.

### Step 1: Defining typographic styles for text elements on the page.

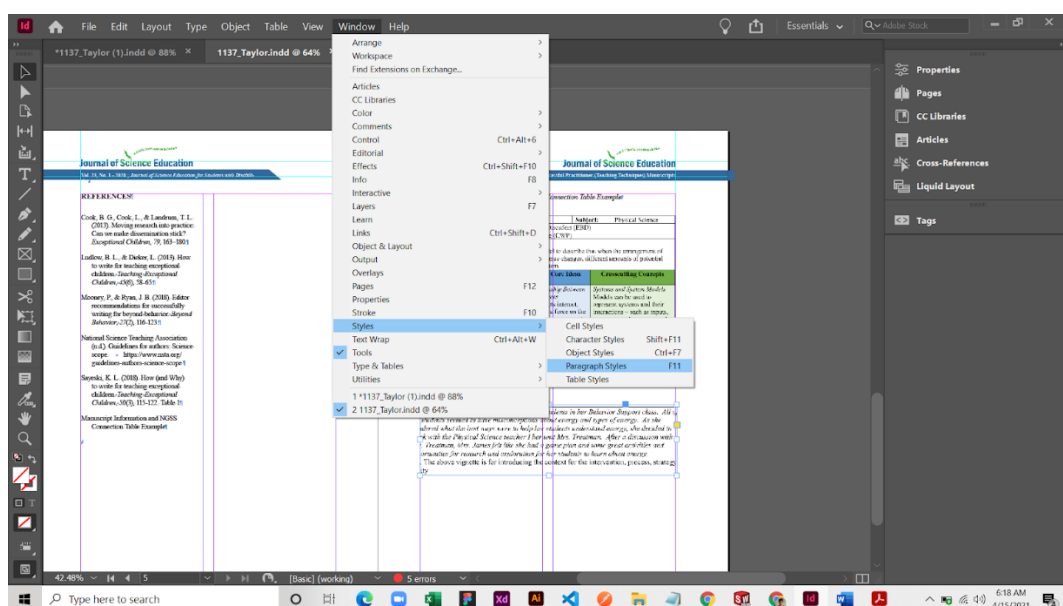




Figure 2: Defining paragraph styles

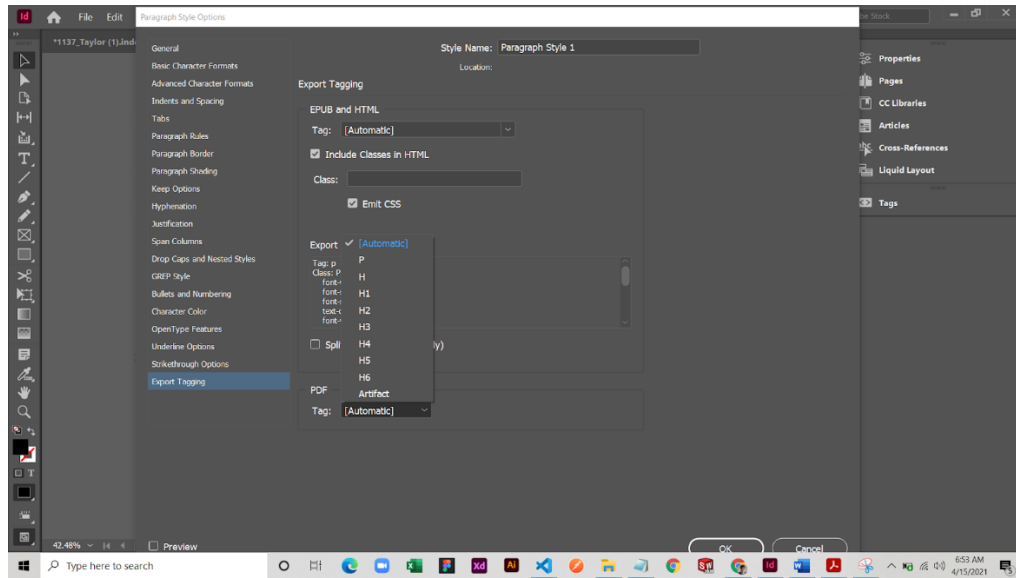


Figure 3: Mapping user-defined tags to the pre-defined HTML tags.

## Step 2: Creating tags

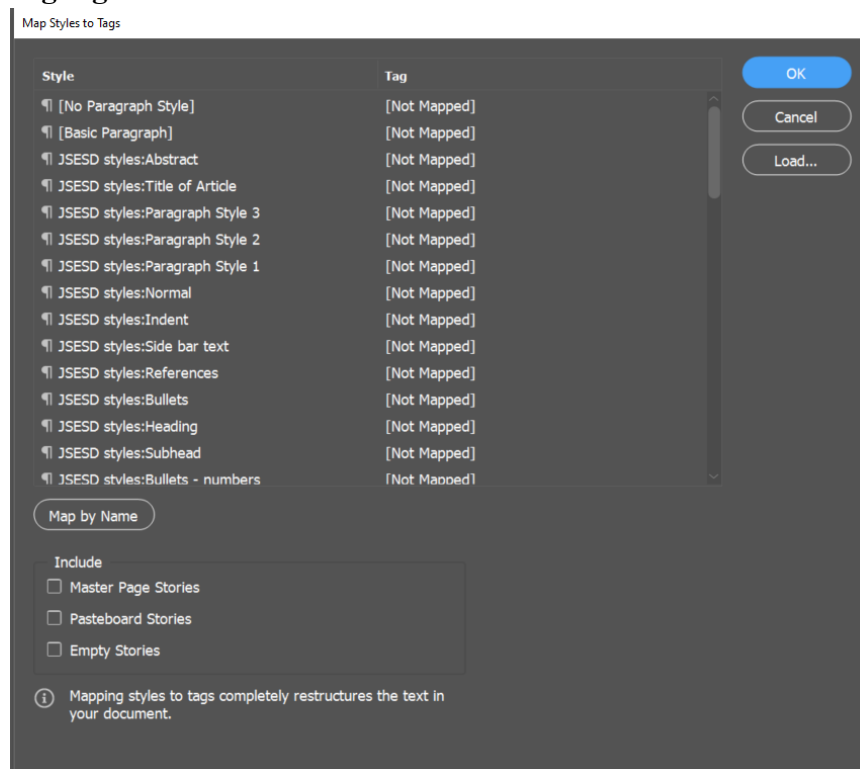


Figure 4: List of unmapped tags in Indesign.

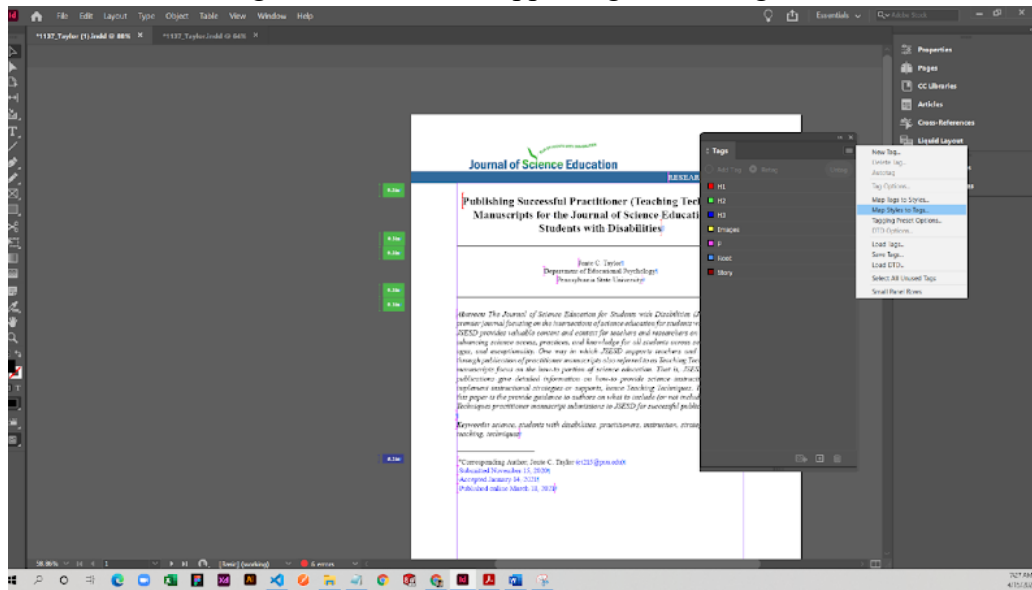


Figure 5: Mapping tags to the document elements.

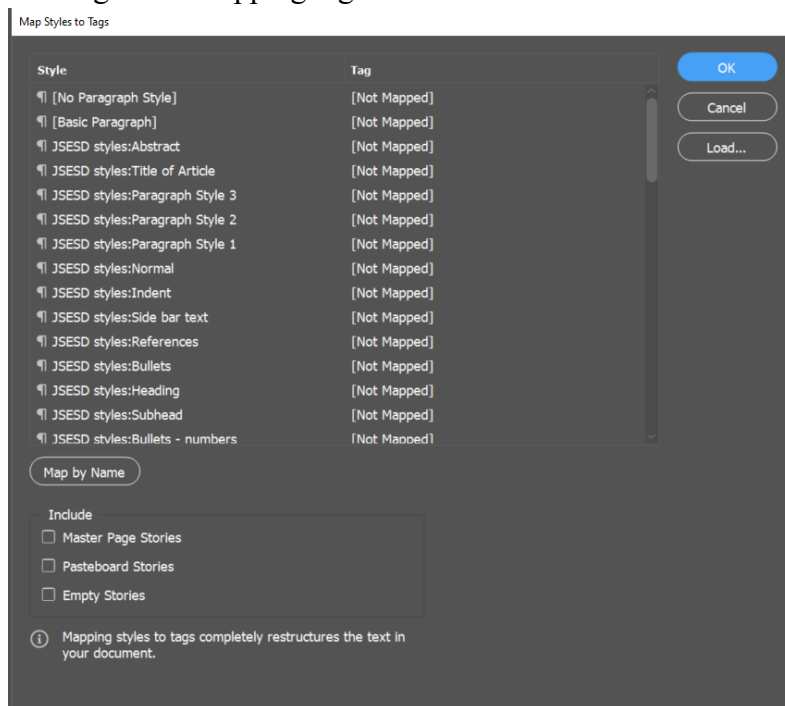


Figure 6: List of tagged document elements.

### Step 3: Ordering the content

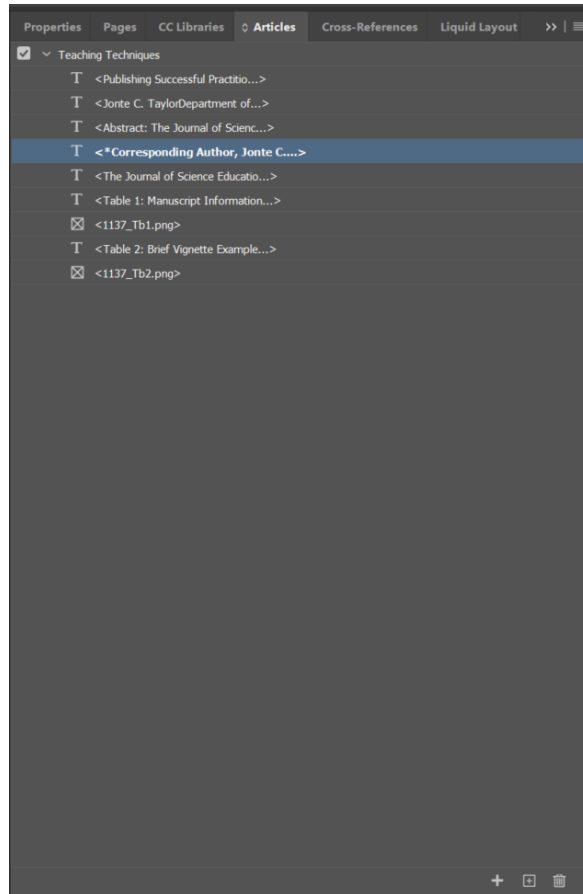


Figure 7: Document order tab.

### Step 3: Exporting to PDF/HTML

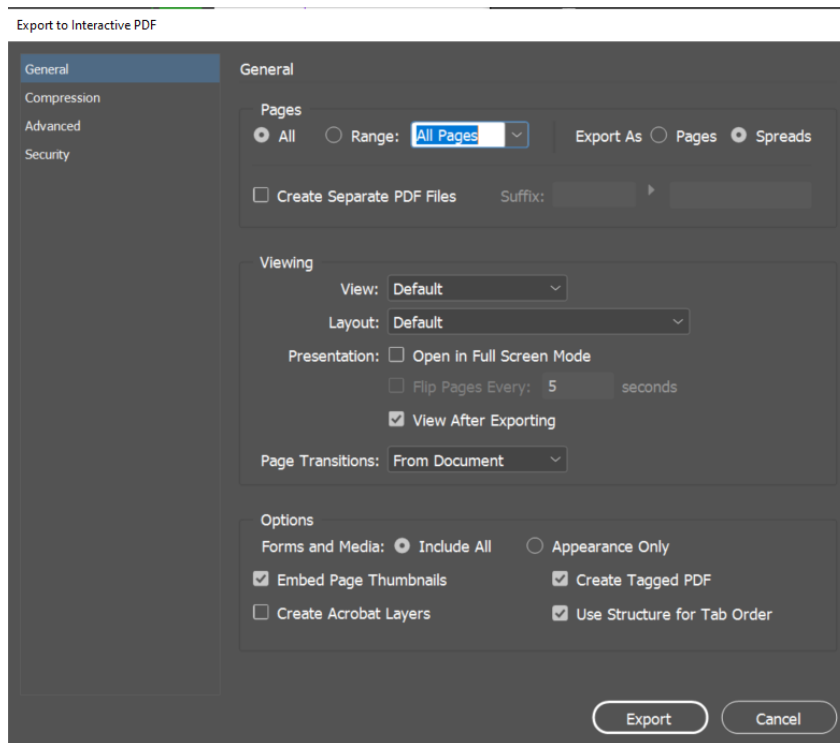


Figure 8: Export settings to maintain the document order.

## 4.2 Adobe acrobat reader

The next tool that we used was the Adobe Acrobat Reader. This was again a great tool but the issue with this approach is that it requires an additional edit after the writing. The editor has to go through all over the document and has to make sure about reading order and tagging. Some additional problems that we noticed are :

Pros

- Efficient in generating accessible PDFs.
- Easy to use.

Cons

- Not very efficient in generating accessible HTML.
- Software is paid and expensive (US\$179.88/yr.)
- Have to be installed before using the accessibility feature.

## 4.3 Online Conversion Tools

There are different online conversion tools which can convert a word or a text document into PDF or a PDF document to HTML. Although we were getting the desired converted formats like PDF and HTML, the accessibility was still a huge problem associated with the approach. Some problems that we noticed are :

Pros

- Free to use.

Cons :

- The document was not ordered properly for the screen reader.
- Tables and images do not have proper alt text.
- A complex tag hierarchy makes it harder for the screen reader to compile the document.

#### 4.4 Word to PDF Using Overleaf/LATEX

We then used the “Latex” to make our articles accessible. Since Latex is also used by a lot of researchers worldwide and this was free to use. So, we tried to make a Latex general template which writers can use to edit their documents on Latex. The pdf generated from this template was most accessible. There are still some challenges that we are working on to make this pdf more accessible. But at the same time to give our reader an additional option we tried generating HTML using the tex and the pdf file from the latex.

The workflow with LaTeX is simple:

- A submitted paper—in the form of a PDF—is pasted into a .tex template and turned into a .tex file.
- This .tex template is an edited version of the Association for Computing Machinery (ACM) .tex template.

Then tex2html—the conversion tool built by Open@RIT—is applied to the .tex file at uses an open-source LaTeX converter called LaTeXXML to convert it to HTML finally.

- The resulting HTML file shows significant improvement with screen readers.



Figure 15: Proposed solution to convert word documents to accessible pdf and HTML format.

The Latex template is available at the following link publicly [Overleaf Link](#).

The above workflow is almost completed during phase 1 of the development of the “Latex” template and “LateXML” tool. [LaTeXML](#) is an open-source Latex to XML/HTML/MathML Converter. We fixed some minor errors and enhanced the accessibility of the HTML file with client-side JavaScript.

```
let footnotes = document.getElementsByClassName("ltx_note ltx_role_footnote");
let footnote;
for(let i in footnotes){
  footnote = footnotes[i];
  footnote.innerHTML =
    `
```

Figure 9: Snapshot of the code showing how we are extracting the latex content using the JavaScript library.

The next phase includes the following tasks:

### **Creating better accessible output formats**

This part of the next phase focus on enhancing the tool for generating better accessible HTML and pdf output file. This can be achieved by adding more accessibility features to the existing solution. The current tool does not support any accessibility support that deals with images, figures, tables or mathematical notations and formulas. So, working on these areas will help create better accessible output formats.

### **Increasing ease of use.**

This first phase of this project was mostly focused on the functional requirements. Usability and user centre approaches were not considered while developing the tool. In the coming phase, we will conduct usability testing at the end of the project. Based on the feedback gathered during the testing phase, our goal for this part is to go through a couple of iterations and work on some of the areas which can help users to use this tool easily and efficiently.

### Using the SE Instructional Model in an Online Environment with Pre-service Special Education Teachers

Delinda van Garderen\*  
University of Missouri  
Mary Decker  
University of Missouri  
Rachel Juergensen  
University of Missouri  
Haba Abdelaaby  
University of Missouri

**Abstract:** In this practitioner article, we describe the innovative way the SE Instructional Model was used in an online, hybrid special education undergraduate course to prepare pre-service teachers to teach academic content to their students with disabilities. We provide a rationale for the use of the model in the course, describe how we implemented the model in the course, pre-service teachers' perceptions about the model as a way to facilitate and model the process of learning for themselves and students, and discuss implications for practice.

**Keywords:** SE Instructional Model, Online learning, Inquiry, Teacher preparation, College teaching

\*Corresponding Author, Delinda van Garderen ([vangardend@missouri.edu](mailto:vangardend@missouri.edu))  
Submitted October 18, 2019  
Accepted February 6, 2020  
Published online March 3, 2020

#### INTRODUCTION

The National Council for Accreditation of Teacher Education (2010) state that faculty and instructors in preservice teacher education programs should model instructional practices to enhance learning and best prepare preservice teachers for their future classrooms. Explicit modeling with reflection and connection to theory is a way for teacher educators to intentionally structure their instruction so that preservice teachers (1) attend to the model used, (2) model the practice appropriately, (3) explicitly connect the model to theory, and (4) allow for reflection as to how the model may affect them and the application to their future classrooms (Moore & Bell, 2019). The use of explicit modeling in connection to theory and reflection can encourage student growth in practice while leveraging the affordances of already known best practices (Lunenburg, Korthagen, & Swennen, 2007). Given this recommendation and the challenge we were recently faced with of creating a hybrid course focused on teaching methods in science and social studies for pre-service special education teachers at a large research university, we decided to use the SE Instructional Model as our form of explicit modeling.

In this practitioner article, we will (a) explain why we used the SE Instructional Model and its benefits for students with disabilities, (b) describe the way we implemented the SE Instructional Model in an online format as a part of a hybrid course, (c) share the pre-service teachers' perceptions about the use of the SE Instructional Model as a way to facilitate and model the process of learning for themselves and students with disabilities, and (d) wrap up the article with final thoughts and implications for practice.

#### SE Instructional Model for Teaching and Learning for ALL Learners

Within science education a well-researched and widely cited instructional model is the SE Instructional Model (Bybee, 2015). (Table 1 provides an overview of the SE Instructional Model) The SE Instructional Model has been demonstrated to be grounded in sound educational theory about learning (Bransford, Brown & Cocking, 1999; Bybee, 2015). As a result, a central argument, among a few (see Abell & Volkman, 2015), for the use of the SE Instructional Model is that the structure facilitates learning in a meaningful and powerful way (Abell & Volkman, 2006; Bybee, 2015). This type of "learning" is one that is focused on developing understanding as opposed to just learning facts, where facts are connected and organized around important concepts that can support transfer of ideas rather than only recall (Bransford, Brown, & Cocking, 2000).

An implication to learning with understanding is the recognition that this type of learning is constructed from experiences and that students should be actively involved in that process (Bybee, 2015). This does not mean, however, that there is no teacher involvement or guidance in that process as has been suggested by some (e.g., Rizzo & Taylor, 2016). Rather, the teacher plays an integral and critical role in ensuring that systematic and carefully designed learning experiences are provided. The strength of the SE Instructional Model is that it provides a structure and function (for each component of the instructional model) for teaching to generate learning experiences to enhance student inquiry (Bybee, 2015).

Findings from research supports the effectiveness of an instructional model such as the SE specifically for improved student (at any

Figure 10: Proposed solution to convert word documents to accessible pdf and HTML format.

## 4.5 LATEX to HTML JavaScript

Using LaTeX, our script can then convert the PDF file into an HTML web page which can be used by screen readers. The script is ongoing development as new articles incorporate new elements (images, tables, equations, etc.).

We used the LaTeXML library for converting a document from latex to HTML, LaTeXML as the name suggests first converts latex to an intermediate XML before converting to any other document formats.

LateXML renders a highly accessible HTML and overcomes some of the issues related to figures and equations that we previously had.

We tried to further improve the accessibility and the appearance of the rendered HTML document with the help of client-side JS.

Here is the code snippet for making the footnotes accessible, which were previously visible or accessible only on the hover state.

## 5. Usability Testing

We conducted usability testing for our proposed solution which converts word to accessible PDF and accessible HTML. Our primary users were in the editing department and who will work on these papers to convert them into PDF and HTML. Secondary personas will be the audiences who are in the research field and are involved in writing technical papers. Since it was hard to approach people in the editing field, we could only recruit 1 participant from the editing department. We tried doing the rest of the usability testing with the researcher who has some experience writing technical papers. We recruited three researchers. Below are some demographic details of the participants.

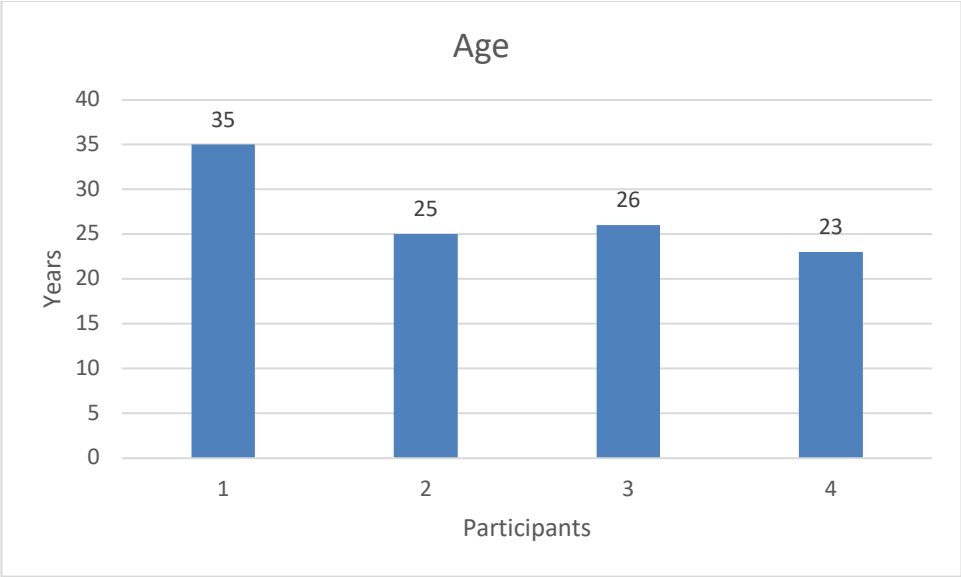


Figure 11: Chart showing age group of different participants

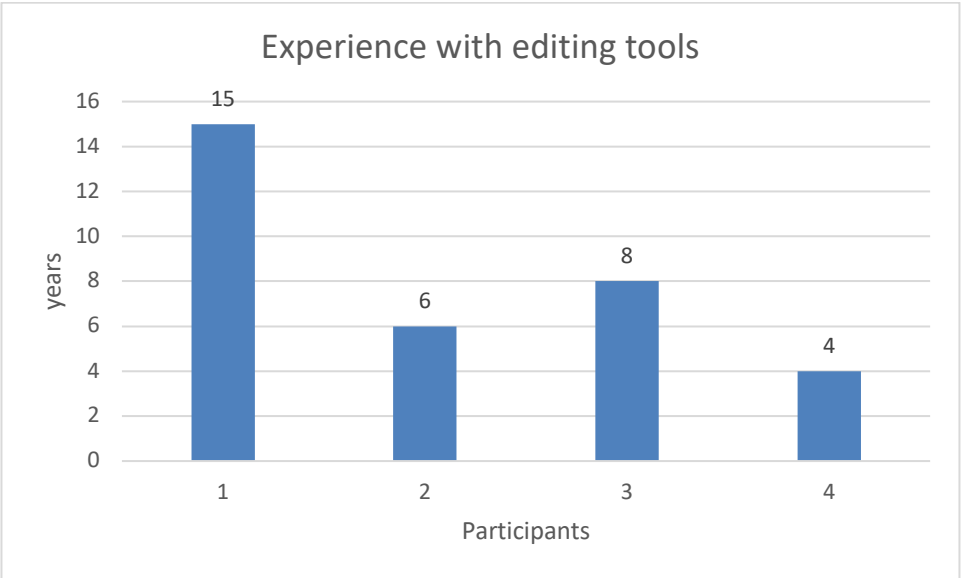


Figure 12: Chart showing participants' experience with the editing tools



The usability testing was divided into 3 sections. The first section included a pre-test survey questionnaire which allowed us to know more about the participants. A pre-test survey questionnaire was followed by the actual test and at the end, they were asked some post-test questions. There was a brief information session where the user was introduced to the system where they must paste the text. Also, they will be provided with the information that was required to use the software.

The usability test included two test scenarios:

**Scenario 1:** You are a researcher who has completed writing your research paper in word. Now you want to paste the same information into Overleaf. Your paper has the following information Title of the paper, author details, abstract and introduction. You will be provided with the word document which has the above-mentioned details.

**Scenario 2:** After completing task 1 you realized that you must include 1 footnote in your introduction section. You will be provided with a picture which tells you what your text will look like once you include the footnote and the footnote text that you must include in the paper.

**Success criteria:** The pdf produced by the user was visually compared with the snapshot as mentioned in the picture below. If it matches with the snapshot, it is a success, if it doesn't then it is a failure.

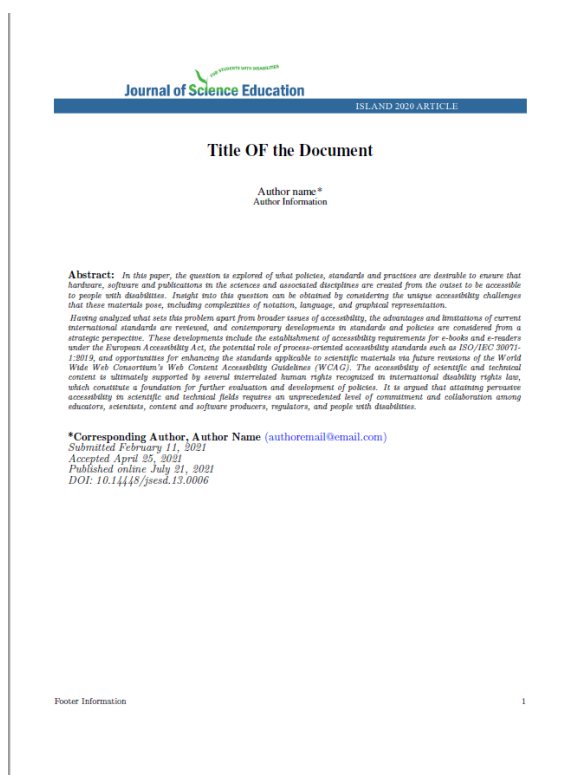


Figure 13: Snapshot of the desired result

## 6. Results

Major findings from the usability testing were:

- Users with some experience with Latex tools find this process easy.
- Users with no experience with Latex find the process a little difficult while performing Task 1 but they find finishing Task 2 easily.
- This solution still requires a little learning curve for the users who are not familiar with Latex.

## 7. Challenges

Some challenges that we faced during the process were:

- Transferring Content from “docx” to “LaTeX”.
- Styling Content.
- Formatting Layout.
- Two different end-product formats.
- Reading Order in LaTeX.
- Add accessible images and tables.

## 8. Limitations

Even though the approach used in this study can improve accessibility for the JESED articles, more work can be done further to make this solution generic, which is to make this tool helpful not just for JESED articles but for all other major conferences as well. Also, the proposed solution still requires some manual work where the user must copy-paste the submitted article into the Latex template (if the submitted format is not the latex files). This manual work can be automated in future work to make it more efficient for the users trying to make non-accessible documents into accessible documents.

## 9. Conclusions and Future Scope

The proposal aims to find a solution that can make research articles for the JESED conference more accessible to the users who are visually impaired and help them efficiently go through the information available in the article.

Future works include more refinement on the solution so that it requires fewer manual tasks for the users while converting it into an accessible format. Also, more work can be done in this tool to make a generic template for any Journal and not specific to JESED articles.

## References:

- [1] “Magnitude and Projections.” The International Agency for the Prevention of Blindness, <https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/>. Accessed 20 Aug. 2021.
- [2] “Relevant Combinations of Screen Readers and Browsers.” Accessibility Developer Guide, <https://www.accessibility-developer-guide.com/knowledge/screen-readers/relevant-combinations>. Accessed 20 Aug. 2021.
- [3] Devine, Heather, et al. “Making Accessible PDF Documents.” Proceedings of the 11th ACM Symposium on Document Engineering, Association for Computing Machinery, 2011, pp. 275–76. ACM Digital Library, <https://doi.org/10.1145/2034691.2034748>.
- [4] Brady, Erin, et al. “Creating Accessible PDFs for Conference Proceedings.” Proceedings of the 12th International Web for All Conference, Association for Computing Machinery, 2015, pp. 1–4. ACM Digital Library, <https://doi.org/10.1145/2745555.2746665>.
- [5] Manzoor, Ahtsham, et al. “ALAP: Accessible LaTeX Based Mathematical Document Authoring and Presentation.” Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, 2019, pp. 1–12. ACM Digital Library, <https://doi.org/10.1145/3290605.3300734>.
- [6] Zulfiqar, Shaban, et al. “Automated Generation of Accessible PDF.” The 22nd International ACM SIGACCESS Conference on Computers and Accessibility, Association for Computing Machinery, 2020, pp. 1–3. ACM Digital Library, <https://doi.org/10.1145/3373625.3418045>.
- [7] Experience, World Leaders in Research-Based User. “Why You Only Need to Test with 5 Users.” Nielsen Norman Group, <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>. Accessed 22 Aug. 2021.
- [8] C. Rana, “Building a Book Recommender system using time-based content filtering,” Wseas.org. [Online]. Available: <http://www.wseas.org/multimedia/journals/computers/2012/54-571.pdf>. [Accessed:

06-Nov-2020].

[9] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egypt. Inform. J.*, vol. 16, no. 3, pp. 261–273, 2015.