

---

# An Uninteresting Tour Through Why Our Research Papers Aren't Accessible

**Jeffrey P. Bigham**

Carnegie Mellon University  
Pittsburgh, PA 15213  
jbigham@cs.cmu.edu

**Erin Brady**

IUPUI  
Indianapolis, IN 46202  
brady@iupui.edu

**Cole Gleason**

Carnegie Mellon University  
Pittsburgh, PA 15213  
cgleason@cs.cmu.edu

**David A. Shamma**

Yahoo Labs  
San Francisco, CA 94103  
aymans@acm.org

**Anhong Guo**

Carnegie Mellon University  
Pittsburgh, PA 15213  
anhongg@cs.cmu.edu

**Abstract**

Our research is delivered as Portable Document Format (PDF) documents, and very few include basic metadata to make them accessible to people with disabilities. As a result, many people are either unable to read them efficiently or at all. Over the past few years, we have tried everything from writing guidelines and giving accessibility feedback, to enforcing accessibility standards and volunteering to make PDFs accessible ourselves. The problem with making PDFs accessible is in part due to the lack of good tools, but the complexity of the PDF format makes improving tools difficult. Making accessible research papers is as much about our choices as a community: our choice of publication format, and our choice to make accessibility a voluntary task for authors. In this paper, we overview the context in which PDFs became our publication format, the difficulty in making PDF documents accessible given current tools, what we have tried to make our PDFs more accessible, and potential options for doing better in the future.

**Author Keywords**

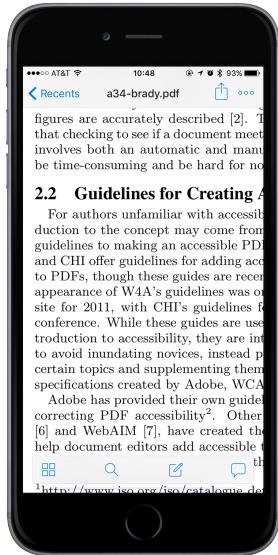
Research papers; accessibility; PDF.

**ACM Classification Keywords**

H.5.m [Information interfaces and presentation (e.g., HCI)]

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.  
*CHI'16 Extended Abstracts*, May 7–12, 2016, San Jose, CA, USA.  
Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-4082-3/16/05 ...\$15.00.  
<http://dx.doi.org/10.1145/2851581.2892588>



**Figure 1:** Reading a PDF on a mobile device is similar to a person with low vision trying to read the document with magnification tools.

## Introduction

Our research is delivered as Portable Document Format (PDF) documents. PDFs form the bulk of the ACM Digital Library. They are what we post on our web sites. They are what we publish at conferences and in journals. We hand out printed PDFs in research seminars. When a student asks for relevant reading on a topic of their research, we send them links to PDF documents. Almost each line in our curriculum vitae ultimately points to a PDF document somewhere. For the most part, this works.

And yet, our PDFs exclude a great many potential readers because they are inaccessible. Accessibility has a few different definitions. In this document, we use the term “accessibility” to generally mean that something is able to be used by people with disabilities. In the case of PDF documents, the most relevant disabilities are usually those affecting vision and reading. To make PDFs accessible, visual information should be represented in another form (often, plain text) so that it can be consumed in another way by blind or visually impaired readers. This requires the document’s structure to be separated from its visual appearance so it reads in a sensible order and can be efficiently navigated by people who cannot see the tacit visual layout of a template. The document should allow for easy magnification for people with low vision. While one may easily mistake this for a zoom feature in a PDF reader, which the current format does not due to (among other things) its two column format. To be made easier to read by people with reading disabilities, the document should be flexible, allowing for not only magnification but changing of fonts, colors, and layout. Creating documents accessible for every possible reader is more interesting and complicated than this handful of items, but if we were to make our PDFs accessible even in only these ways it would be a large step forward.

Technically, the PDF format can be made to do most of these things. In practice, we know that relatively few PDFs are produced in an accessible way in CHI and other conference proceedings [5]. In fact, we know that in CHI 2014, a year in which considerable effort was spent giving author’s feedback on the accessibility of their documents, only 26.8% included any document tags at all (up from 20.6% the year before). Even in 2014, only 10.9% specified a language (strongly advised so screen reading software will speak in the right language). In the years before and since, the accessibility of papers has been worse. Given such low presence of accessibility at all, we haven’t bothered to look at more depth at how *well* authors added accessibility to their papers.

This has very real consequences for who can view our research and the impact that it can have. Imagine reading through a research paper, attempting to learn from it, and not having access to any of the figures, diagrams, or tables. Imagine the paper jumping from the left column to the right and back again as you read, apparently at random. This is what it’s like to read one of our PDFs with a screenreader. Or, just think of how difficult it was to attempt to read one of our PDF documents on your mobile device—once you zoomed in far enough to be able to read it, you have to constantly scroll from left to right, while trying not to scroll too far into the other column (Figure 1). This is what it’s like to read one of our PDFs if you have low vision.

The title of this paper is “An Uninteresting Tour” because sadly these problems exist with our papers despite the fact that we know how to solve all of them. Most could even be solved using the PDF format itself, although as will be discussed the tools for doing so with PDFs are complicated to use and expensive. Unlike much of the rest of the publishing process, we rely on authors to fight through this

complicated process to make their papers accessible. Instead of being a technical problem, the challenge comes down to choices that we have made as a community, the incentives we have given authors, and the poor usability of the tools we expect authors to use. As researchers (even human-centered ones), we are not always the best equipped to deal with problems like this. The problem isn't sexy. It's partially a matter of engineering, partially a matter of access to tools, and partially a matter of practice; and it's all too easy to ignore or dismiss.

This paper outlines the current state of making PDF documents accessible, describes the several different ways we have attempted to make PDF documents more accessible at various ACM conferences, and suggests several possible solutions.

### **PDFs and Accessibility**

According to an early whitepaper, the PDF's origins come from a need to replace an inaccessible object doing digital document transfers: the FAX machine.

The popularity of FAX machines has given us a way to send images around to produce remote paper... What industries badly need is a universal way to communicate documents across a wide variety of machine configurations... These documents should be viewable on any display and should be printable on any modern printers. [25]

By the PDF 1.0 1993 documentation, object annotations and hyperlinking were added to the specification [1], so the format had made a step beyond a simple FAX machine. However, by PDF 1.3, some 7 years later, the core utility and focus was still the printed page: "PDF is now the industry

standard... of printed material in electronic prepress systems for conventional printing applications." [2] By 2006, the PDF had become an open standard with accessibility support and listed as a feature in the manual's preface [3]. Today's PDF aspires to be a single container for all media, but, we assert, its core, and the surrounding supporting legacy information systems, remain a better FAX.

The PDF format has been extended by additional standards, largely driven by the needs of the archival and accessibility communities. PDF/A-1a was introduced mostly to facilitate visually reproducible documents for archival, but it also required the text of the document to be included in the metadata [11]. In 2012, accessibility features were formalized when PDF/Universal Accessibility (PDF/UA) was released [10]. PDF/UA provides a set of criteria for a PDF file to be considered accessible, including proper tags in a logical reading order and alternative texts for images. It also requires that the document supports reflowing the text into smaller screens or screen magnifiers. Both of these standards place the burden of making the document accessible on the author, as there is no tool to automatically verify documents conform to these standards.

### **How to make PDFs Accessible**

The biggest problem at this point is the difficulty of adding even basic accessibility metadata to the PDFs given the tools available. CHI authors most often use Word for Windows, Word for Mac, or  $\text{\LaTeX}$  to create their PDF documents. The CHI accessibility community has put together instructions<sup>1</sup> and one of the authors of this paper maintains the SIGCHI Github templates<sup>2</sup> to help make PDF document accessible. The tools required to follow these instructions are either Word for Windows (for Word

---

<sup>1</sup><https://chi2016.acm.org/wp/accessibility/>

<sup>2</sup><https://github.com/sigchi/Document-Formats/wiki>

<p><b>[1]</b> books [5]. These guidelines are often complex and only detailed - the Adobe XI accessibility guidelines, when totaled, are 188 pages long, and McCall's digital book is over 800 pages. For a novice, this volume of information may be overwhelming.</p>	<p><b>[2]</b> to 2014, but even in 2014 only a quarter of the documents were tagged. ASSETS, which in 2011 and 2012 had extremely high rates of document tagging, has now slightly decreased from a high of 92% of documents being tagged in 2012 to only 71% in 2014. This may be the result of the growth of the ASSETS community, as new researchers join who are unfamiliar with how to make their documents accessible. W4A has recently improved, going from having no tagged documents in 2011 (the year the guidelines were introduced) to 100% tagging over both communications and technical papers last year.</p>
<p><b>[3]</b> Academic Research on PDF Accessibility</p> <p>With web-based guidelines, other research has looked at the accessibility of PDFs and the impact of inaccessible documents on screen reader users. Hewson and Tonkin used automated tools to evaluate accessibility of a repository of academic documents and found that 10% of the documents had used PDF tags to provide structure [3]. However, to our knowledge our paper is the first to discuss PDF accessibility specifically within the context of ACM publications.</p> <p>Other work has examined how the accessibility of PDFs impact screen reader users. A study with 100 blind screen reader users found that inaccessible PDFs were one of the major causes of frustration when browsing the web [4]. Lazar et al. point out in the paper that PDFs with PDFs is not a lack of solutions for accessibility problems in the format, but instead a lack of knowledge or prioritization of accessibility by content authors.</p>	<p><b>[4]</b> Manual Accessibility Check</p> <p>After running automated tests, we examined the accessibility of papers from the W4A, ASSETS, and CHI conferences and found that 73.1% of papers had no proper tab order specified. These high levels of non-compliance indicate that most authors understand the importance of this accessible information or navigation aids, and will take the time to add them in. However the use of structural tags was haphazard - for example, only 11.5% of papers had the title tagged with an H1. This shows that, while documents may appear accessible due to the presence of tags, they are not always correctly applied, and thus cannot be used as a good indicator of accessibility.</p>
<p><b>[5]</b> ACCESSIBILITY OF RECENT CONFERENCES</p> <p>To examine the accessibility of recent conference proceedings, we performed a two-fold analysis: (i) a large-scale automated check for accessibility of 1811 papers from the last four years of W4A, ASSETS, and CHI conferences, and (ii) a manual examination of accessibility for 26 papers from last year's W4A and ASSETS conferences. This dual analysis is suggested by the PDF/UA compliance checks (Section 2.1).</p> <p><b>3.1 Automated Accessibility Check</b></p> <p>For the automated tests, we selected all the papers from 2011 to 2014 of the conferences of CHI, ASSETS, and W4A (both the technical and non-technical tracks). We generated metadata from the conference proceedings PDFs using the Accessibility Checker<sup>3</sup>, a tool which allows one access to the metadata generated by the accessibility process. Excluding a negligible number of files which failed to work with the Accessibility Checker, we had a collection of 1811 PDFs.</p> <p>Using metadata from the PDFs, we were able to perform automated checks to see if papers were tagged at all, if any</p>	<p>commonly used features in Word for Windows can only be fixed using Adobe Acrobat. For example, Word has a very nice feature that allows rich charts and figures to be copy and pasted into Word documents from other Microsoft Office programs like Excel. Screen readers accessing the Word documents directly can access the underlying data in these rich objects, but there is no way to associate an alternative description (let alone data) in an exported PDF. Tables with multiple or nested headings can also only be structurally tagged by using Acrobat.</p>

**Figure 2:** PDF does not preserve the reading order of the text in a document, so screenreaders often jump from one column to another. This screenshot from Adobe Acrobat shows the order assigned incorrectly. Acrobat allows a user to fix this, but it must be done manually each time the paper is updated.

documents) or the Adobe Acrobat Application (a PDF maker and editor) which offers a 94 page manual just for accessibility [4].

**Word for Windows** Making PDFs accessible using Word for Windows requires adding alternative text to all figures, marking table headers, generating a tagged PDF, and using Adobe Acrobat to verify the default language as well as set the tab order of all pages. This offers an approach that works for generating accessible PDFs. A number of

commonly used features in Word for Windows can only be fixed using Adobe Acrobat. For example, Word has a very nice feature that allows rich charts and figures to be copy and pasted into Word documents from other Microsoft Office programs like Excel. Screen readers accessing the Word documents directly can access the underlying data in these rich objects, but there is no way to associate an alternative description (let alone data) in an exported PDF. Tables with multiple or nested headings can also only be structurally tagged by using Acrobat.

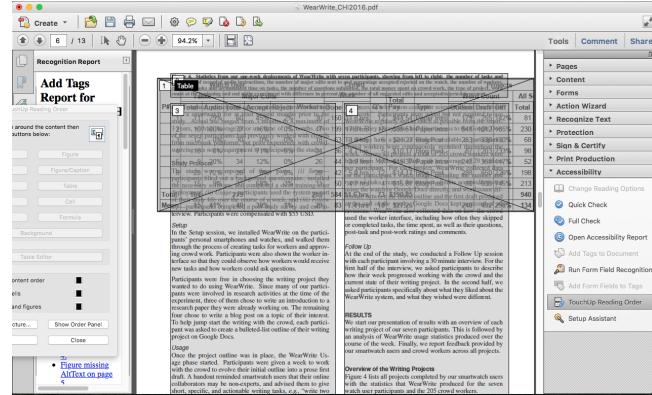
**Word for Mac** Accessibility support for Word on the Mac when exporting to PDFs is non-existent; leaving the SIGCHI website to suggest “The best choice is for authors to open their completed Word file on a Windows machine.” While Word for Windows has gradually supported greater accessibility for Word documents exported to PDF over the past few years, Word for Mac exports none of the accessibility metadata to which it has access. Authors can insert alternative text for images and add structure, but none of that will make it into PDFs created with Word for Mac. While there are clearly a number of different incentives at play for Microsoft in choosing what features make it into their Word product for Mac, it is unfortunate that accessibility of PDFs has been excluded.

**L<sup>A</sup>T<sub>E</sub>X** If you use L<sup>A</sup>T<sub>E</sub>X to typeset your document, the ACM Guidelines advise you to use Adobe Acrobat to make the resulting PDF accessible. This is somewhat disappointing because L<sup>A</sup>T<sub>E</sub>X source itself is the document's marked up structure, including much of the metadata that could be used to make the documents accessible (e.g, headings, floats, etc.). However, the two attempts we found to make accessible PDFs are incomplete implementations at best.

**Adobe Acrobat** Acrobat<sup>3</sup> is a product from Adobe that allows PDF documents to be read and edited in various ways. While some other tools exist to help make PDF documents accessible, none are as developed as Adobe Acrobat. To make documents accessible using Adobe Acrobat requires first “Auto-Tag” the document. The original PDF format, having been originally developed to represent documents to be printed, had no way to represent structure. “Tagging” a document introduces a separate tag tree (kind of like a Document Object Model) that can include structure. The auto-tagging procedure associates text that should be considered to be a cohesive block, identifies headings, figures, and reading order. The CHI two-column format often causes problems for this, as text and figures are often grouped incorrectly, with text from figures appearing in the middle of a paragraph (according to the automatically inferred structure). Additionally, the reading order of the text (as seen by screenreaders) often jumps between the columns (Figure 2). Acrobat includes features to fix these elements (TouchUp Text, Reading Order), but these features are difficult to find and use.

Adding alternative text to images often exposes weird bugs in Acrobat. For instance, it is quite common for an image to change size (e.g., explode) in the PDF upon adding alternative text or for tables to lose their position when tagged (Figure 3). Acrobat does not include an undo function for any of this functionality, and so if users make a mistake, they must start over. Inserting alternative text for images also often introduces a new font that is not embedded in the document. All text needing an associated font, even text that doesn’t appear anywhere visually, is probably a vestige of PDF’s past life as a printing format. The easiest way to embed fonts in Acrobat requires being able to visually select them (an inaccessible feature of the

<sup>3</sup><http://acrobat.adobe.com> (as of 1/2016)



**Figure 3:** Making a table accessible in Adobe Acrobat often messes up its position in the paper. In this case, the table has been placed above text on the page. We are not sure why this happens, but with no ‘undo’ feature in Acrobat, we imagine many authors have simply given up in frustration after it happened. We now make a habit of saving our documents after every change so we can close and reopen the document upon hitting a bug.

tool), and so this cannot be done for this font. Lately, CHI has given Sheridan permission to allow this font not to be embedded.

The final problem with using Acrobat to make papers accessible follows from it being a step after writing the paper. Authors write their paper in Word or L<sup>A</sup>T<sub>E</sub>X, produce a PDF, and then make it accessible in Acrobat. There is no way to save the work you’ve done to make your PDF accessible. This means that when you notice a typo, or Sheridan comes back with a requested change, you have to do all your work again. Those of us who have been through this before have learned to at least keep a copy of the alternative text descriptions that we have written, but the rest of the process must be manually redone. That means

doing the same work once for submission (while we submit multiple versions of a submission just in case, only the last one is generally accessible), then (usually) multiple times for camera-ready.

### Will People Make Their Papers Accessible?

Despite the technical hassles listed above, it is possible (though time-consuming) to make PDFs accessible. Yet, it is generally found that PDFs are a common challenge in accessibility across online domains [15], and particularly in academic contexts [22, 24]. Why is it that authors are still not making their documents accessible? Our failures in making PDFs accessible are representative of our general failures in accessibility in a range of domains, and exploring those failures may provide some intuition of how to solve this problem. We know how to make websites accessible, or at least we have very clear standards for web accessibility, that we could use as a guide. Standards like the Web Content Accessibility Guidelines (WCAG) [7] have been in place for over fifteen years, providing high-level focus areas and actionable 'success criteria' which can be used to judge whether a website is accessible before it is released. Despite these guidelines, many websites remain inaccessible [12], and an analysis of web accessibility over time showed that a random sample of websites only grew more inaccessible as time went on and the owners added increasing complexity to their pages [9].

It is rare that these inaccessible websites are due to malice, or explicit discrimination against people with disabilities. Instead, they are the inevitable consequence of the development process. Web developers are under numerous competing constraints to finish a project, with time constraints being one of the biggest issues cited that was more pressing than accessibility issues [16]. They also may lack the knowledge or formal training to make websites

accessible [16], and the guidelines are not always specifically directed at developers (a strategy which could greatly improve their usefulness [14]).

This leads to a dilemma: how can we retroactively make this content accessible? The volume and speed at which new content on the web is created is overwhelming. Because of the rate of new materials, and the inherent inequality of publishing materials for people with disabilities after the general population has access to them, scholars have encouraged the creation of *born-accessible* content to deal with this deluge of information [26]. But content authors who lack the know-how or resources to make their content accessible will likely not feel compelled to prioritize accessibility over other constraints without pressure.

Legal requirements or the fear of lawsuits may be one motivator for generating accessible documents, and law has been proposed as a framework for improving the accessibility of CHI conferences [13]. After Target Corporation settled with the National Federation for the Blind to avoid legal repercussions for having an inaccessible digital storefront, other large corporations started to pay attention to accessibility and improve their sites [8]. Government information technologies are required to be accessible under Section 508 of the Rehabilitation Act Amendments of 1998, and as a result [9] found that government websites, unlike the random sample, do not degrade in accessibility over time. While a fear of legal action may work for large corporations with resources to enact change, the expense and time involved with enforcing accessibility through the law means that individual authors are unlikely to give much regard to the potential of legal action.

There is also a fundamental gap in knowledge between people with specific disabilities and people without them

### Excerpt from UIST email

(We) need your help to generate descriptions for your images and figures, which they will insert into your PDF files to make them accessible to screen reader users. This should only take on few minutes of your time. You will provide a description for each image in plain text via email. The goal of the description is to convey the information that you intend a sighted person to get from the image. A generic description, e.g., “A picture of our prototype” is not very helpful. Much better is a description with relevant details, e.g., “A picture of our smartwatch prototype demonstrating how it projects additional information on the user’s arm to extend the viewable area.”...

which may contribute to bad accessibility. Beyond not knowing how to implement standards correctly, people without a certain disability often misjudge or do not recognize crucial issues which may be problematic for people with disabilities. In a study of mixed-ability workplaces, Branham and Kane found that blind employee’s co-workers were not able to accurately identify the majority of the problems that they encountered in the workplace [6]. This misunderstanding of priorities can also lead to accessibility issues when developers or content authors don’t realize the significance of certain standards or features.

Content creators are, in general, so bad at making their materials accessible that the bulk of academic research has focused on ways to work around their errors. Screenreaders are very forgiving to structural issues of websites, and numerous tools have been created to improve or supplement the original site [23]. Despite these advances, most guidelines suggest manual checks or require some level of human-generated content like alternative text.

### Our Experience

We have attempted to make PDF documents accessible in several ways that we have categorized into *asking*, *requiring*, *volunteering*, and *being paid*.

1. *Asking*: As a community, we have helped to produce guides for authors to follow to make their PDF documents accessible, and we have asked authors to follow them [5]. While we believe we could be even better about asking authors to follow the guidelines, compliance was remarkably low even at CHI 2014 in which each author received multiple reports regarding the accessibility of their papers. We suspect, like the web developers mentioned above, authors are

constrained by a lack of time and knowledge of how to make their papers accessible.

2. *Volunteering*: For CHI 2015, we volunteered to make the camera-ready versions of papers accessible for any authors who emailed them to us. We advertised this service on social media. We made 25 of 486 PDFs accessible this way. For this service, we required authors to do nothing but send us their PDF document. We returned an accessible PDF back to them within a few hours (usually, much faster than that). The turn-around time was important because authors had to meet the real camera-ready deadline. This approach worked for the PDFs that we were sent. Of course, some authors chose not to send us their PDFs, but given that we did it for free we could probably encourage that in the future. The biggest challenge is the cost in time. Each PDF required between 20 and 30 minutes to make accessible, with the most time spent writing alternative text, embedding fonts, and dealing with bugs in Acrobat, e.g., images changing size after adding alternative text to them.
3. *Requiring*: The Program Chair of ACM ASSETS<sup>4</sup>, Jeffrey Bigham, was able to require and enforce all authors to produce accessible PDFs. Each of the 30 accepted submissions was manually checked for accessibility, and if it did not meet the requirements was sent back to the authors. Because it was ASSETS, it mostly worked.
4. *Being Paid*: UIST 2015 paid us to make all of their PDFs accessible, although we ended up sponsoring the conference for the same amount. We gave them a very discounted rate, which we couldn’t extend to

<sup>4</sup><http://www.sigaccess.org/assets/>

more conferences; it was still fairly expensive (\$1600, ~\$23 per article). In an attempt to minimize our labor, the Program Chairs of UIST 2015 (Tovi Grossman and Bjoern Hartmann) emailed the UIST authors on our behalf to request alternative text for images with the following email. See the sidebar to the left for an excerpt.

Authors of 53 of the 70 papers sent us a total of 563 alternative text descriptions. Unfortunately, the majority (~60%) of alternate text was the same as the image caption text. This can sometimes be appropriate, but only rarely when the image really is simply a supplement for the caption text. The goal of alternative text is to convey whatever information is conveyed visually, which captions rarely do because if you are sighted having such detailed captions might be seen as redundant. Less than 25% of the alternate text given was descriptive enough to understand the point the author was attempting to make with the image.

Authors may need additional guidance on writing alternate text that is descriptive and useful to someone who can't see the figures. A paper recently published in TACCESS explored the use of templates to help crowd workers better create alternative text [18], but may also be useful for paper authors as it is clearly not straightforward. The authors of that paper have encapsulated their work into a web-based wizard<sup>5</sup>.

## Options Moving Forward

The most frustrating part about the lack of accessibility of research papers is that all we would need to do to make

them accessible is to decide to do so. Here are some options for how we could do it:

1. Pay someone to make PDFs accessible as part of the publishing strategy. This is likely the only near-term solution for making papers accessible, but it requires money. The amount required is likely to be sizeable, but not out of line with other important expenses that CHI and other conferences incur. Hiring a group to make the content accessible should also result in alternate text that is consistently high quality.

To see if this would be cost-effective, we contacted a few PDF accessibility companies to see what they would charge to make typical CHI papers completely accessible. The responses ranged from \$85 to \$175 per document, although this did not take into account any bulk discounts. To find a cheaper option, SIGACCESS spoke with Sheridan about making the documents accessible on a very basic level (*i.e.*, no alternative text, just simple tags). Their estimate was about \$3 per page (*i.e.*, \$30 for a 10-page CHI paper), which is much cheaper than a fully-accessible document. Having at least this level of accessibility guaranteed in all documents would be a step in the right direction.

2. Wait for and/or actively encourage better tools for authors—our caution here is that we've been waiting for this for nearly a decade, and even Microsoft Word for Mac doesn't produce accessible PDFs. It seems like an easy problem, but those who have investigated it find it's a messy problem that requires a lot of engineering. Even with better tools, it is unclear authors will voluntarily make their papers accessible, or that they will do it well.

---

<sup>5</sup><http://valeriemorash.com/imageDescription.html>

3. Force authors to make their submissions accessible by enforcing PDF/UA conformity. Currently, when you submit your final document into the Precision Conference System (PCS) you will get an error if your fonts are not properly embedded in the document. PCS could throw a similar error if a document does not contain alternative text for an image or the text is not available to a screenreader. Sheridan and SIGCHI currently provide instructions to ensure font embeds, but, similar to the accessibility instructions, it requires commercial software. Also, while not impossible, this might pose a barrier to many authors as the tools are not turn key and friendly (as we have discussed).
4. Get rid of PDF as the publication format. This is likely the only long-term solution. PDF is an objectively terrible format [19, 20], designed for printing and not for reading, and most of the problems we have with making papers accessible are either due directly to it being a terrible format or are made worse because of it. It treats the document markup and text as second-class citizens compared to visual layout, which implies that it will never be a great format for people with visual impairments or anyone who needs custom styling. Switching to another document format that preserves text order and allows the user to easily change font size and type would make this problem simple. For historic reasons this is difficult, as our community is very familiar with distributing PDFs. Researchers have already spent decades making other document publishing formats accessible by default, most notably websites. Using HTML and CSS to format academic publications is a great way to solve the problem. PubCSS[21] could be used to render the document in a way similar to the current

proceedings template<sup>6</sup>, if we want articles to still look like the printed page. Another approach would be to use a document converter like Pandoc[17] to produce a traditional PDF from a better markup language, e.g., using HTML and CSS or plain text. Using a document format that screenreaders can access implies that it is easily machine readable, which also benefits archival and analysis of our field. This is something the community seems to be interested in, as CHI recently changed the reference format partially to make it easier for machines to recognize papers and authors<sup>7</sup>.

Moving forward, we can also add animations, videos, and demonstrations in the paper itself, since it's all supported by HTML5. The reader could more easily explore referenced datasets inline with the text. When printing, a static version would be used. Creating web templates of academic publications can also facilitate collaborative writing of research papers, if the template could be embedded into popular online tools like Google Docs. It is sometimes suggested that an accessible version of the paper be supplied along with the PDF. In practice, this does not work because it is likely that the PDF will be shared, but not the supplementary information.

## Conclusion

All of our research is produced as PDFs and very few of these are made in an accessible way, which means some people cannot access this research or have difficulty doing so. Creating accessible documents is a solved problem

---

<sup>6</sup><http://thomaspark.co/project/pubcss/demo/acm-sig-sample-web.html>

<sup>7</sup><http://sigchi.tumblr.com/post/127563985260/a-rose-by-any-other-name>

technically, but is difficult to achieve in practice due to our choice to publish documents as PDFs (and the resulting limited selection of high-quality tools), our choice to have authors voluntarily make their PDFs accessible, and the lack of author incentives to do so (especially given the difficulty of the process). Accessible documents also provide a broader benefit beyond the visually impaired as they provide better contexts to read and view documents across a plurality of devices. We hope authors will take steps to make their own papers accessible, encourage better tools for producing documents, and advocate for accessible publication formats at ACM conferences.

### Acknowledgements

We would like to thank our reviewers for their helpful comments, and the many people who have contributed to helping us make PDFs more accessible over the past few years. This work was made possible with funding from the National Science Foundation and the National Institute on Disability Independent Living and Rehabilitation Research.

### References

- [1] Adobe Systems Incorporated 1993. *Portable Document Format Reference Manual* (1 ed.). Adobe Systems Incorporated, Reading, Massachusetts.
- [2] Adobe Systems Incorporated 2000. *Portable Document Format Reference Manual* (2 ed.). Adobe Systems Incorporated, Reading, Massachusetts.
- [3] Adobe Systems Incorporated 2006. *Portable Document Format Reference Manual* (6 ed.). Adobe Systems Incorporated, Reading, Massachusetts.
- [4] Adobe Systems Incorporated 2012. *Adobe Acrobat XI Pro Accessibility Guide: Best Practices for PDF Accessibility*. Adobe Systems Incorporated, 345 Park Avenue, San Jose, CA 95110-2704 USA.  
<http://www.adobe.com/accessibility/products/acrobat.html>.
- [5] Erin Brady, Yu Zhong, and Jeffrey P. Bigham. 2015. Creating Accessible PDFs for Conference Proceedings. In *Proceedings of the 12th Web for All Conference (W4A '15)*. ACM, New York, NY, USA, Article 34, 4 pages. DOI: <http://dx.doi.org/10.1145/274555.2746665>
- [6] Stacy M Branham and Shaun K Kane. 2015. The Invisible Work of Accessibility: How Blind Employees Manage Accessibility in Mixed-Ability Workplaces. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 163–171.
- [7] Ben Caldwell, Michael Cooper, L Guarino Reid, and Gregg Vanderheiden. 2008. *Web content accessibility guidelines (WCAG) 2.0*. Vol. 11. W3C.
- [8] Jonathan Frank. 2008. Web accessibility for the blind: Corporate social responsibility or litigation avoidance?. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*. IEEE, 284–284.
- [9] Stephanie Hackett, Bambang Parmanto, and Xiaoming Zeng. 2004. Accessibility of Internet websites through time. In *ACM SIGACCESS Accessibility and Computing*. ACM, 32–39.
- [10] ISO 14289-1:2014 2014. *Document management applications – Electronic document file format enhancement for accessibility*. Standard. International Organization for Standardization.
- [11] ISO 19005-2:2011 2011. *Document management – Electronic document file format for long-term preservation*. Standard. International Organization for Standardization.
- [12] Shaun K Kane, Jessie A Shulman, Timothy J Shockley, and Richard E Ladner. 2007. A web accessibility report card for top international university web sites. In *Proceedings of the 2007 international*

*cross-disciplinary conference on Web accessibility (W4A).* ACM, 148–156.

- [13] Reuben Kirkham, John Vines, and Patrick Olivier. 2015. Being Reasonable: A Manifesto for Improving the Inclusion of Disabled People in SIGCHI Conferences. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 601–612.
- [14] Chris Law, Julie Jacko, and Paula Edwards. 2005. Programmer-focused website accessibility evaluations. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 20–27.
- [15] Jonathan Lazar, Aaron Allen, Jason Kleinman, and Chris Malarkey. 2007. 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.
- [16] Jonathan Lazar, Alfreda Dudley-Sponaugle, and Kisha-Dawn Greenidge. 2004. Improving web accessibility: a study of webmaster perceptions. *Computers in Human Behavior* 20, 2 (2004), 269–288.
- [17] John MacFarlane. 2013. Pandoc: a universal document converter. (2013).
- [18] Valerie S Morash, Yue-Ting Siu, Joshua A Miele, Lucia Hasty, and Steven Landau. 2015. Guiding Novice Web Workers in Making Image Descriptions Using Templates. (2015).
- [19] Jakob Nielson. 2001. Avoid PDF for On-Screen Reading. (2001). <https://www.nngroup.com/articles/avoid-pdf-for-on-screen-reading/>.
- [20] Jakob Nielson. 2003. PDF: Unfit for Human Consumption. (2003). <https://www.nngroup.com/articles/pdf-unfit-for-human-consumption/>.
- [21] Thomas Park. 2015. PubCSS: Formatting Academic Publications in HTML & CSS. (2015).
- [22] Jane Seale. 2013. *E-learning and disability in higher education: accessibility research and practice*. Routledge.
- [23] Hironobu Takagi, Shinya Kawanaka, Masatomo Kobayashi, Takanori Itoh, and Chieko Asakawa. 2008. Social accessibility: achieving accessibility through collaborative metadata authoring. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 193–200.
- [24] Cindy Tandy and Mike Meacham. 2009. Removing the Barriers for Students with Disabilities: Accessible Online and Web-Enhanced Courses. *Journal of Teaching in Social Work* 29, 3 (2009), 313–328.
- [25] John Warnock. 1991. The Camelot Project. (1991). <http://bit.ly/pdfwhitepaper>
- [26] Brian Wentz, Paul T Jaeger, and Jonathan Lazar. 2011. Retrofitting accessibility: The legal inequality of after-the-fact online access for persons with disabilities in the United States. *First Monday* 16, 11 (2011).