

Michael Chee and Kari D. Weaver

# The great PDF debate

## Accessible or impossible?

One of the most basic products of the academic library, the electronic text document, has been the object of recent discussion among accessibility champions—specifically, the ideal accessible format in which text documents should be published online. Such text documents may serve a wide range of purposes, including instructional (e.g., worksheets), informational (e.g., staff directory), or promotional (e.g., annual report). Text documents may be published in a range of formats, but in the academic library context are usually published and made available as Microsoft Word documents, PDF documents, or HTML webpages. While the PDF has long been the preferred publication format, recent publications have questioned this, suggesting that PDF may be less accessible than alternatives like the Word document.<sup>1</sup> We must consider when and where PDFs make sense as a useful format that can be made accessible, and where we can engage in an effort to set standards for accessibility compliance.

### Document creation in the academic library

It is helpful to reflect on the most common ways text documents are produced within the academic library context. In our estimation, they are the following:

- Library workers open their institutional copy of Word to create a document in .docx format.
- An existing print resource is scanned, creating an electronic file; this file is typically exported as a PDF that is meant to be read.
- Content meant to be published straight to a website in HTML is, depending on local organization, written directly into a web editor, or more likely, drafted first in text editing software before being copied into the webpage either by the author themselves or a designated IT department liaison.

What this reflection illustrates is that text documents have a range of provenance, but that the main bias librarians share when considering document creation is that they likely start within their institutionally provided document text editing program (e.g., Word).

It is important to note that Word documents and HTML outputs are not born accessible.<sup>2</sup> The accessible format conversation often sets up a false dichotomy between PDF and other

---

Michael Chee is open education librarian, email: mchee@uwaterloo.ca, and Kari D. Weaver is learning, teaching, and instructional design librarian, email: kdweaver@uwaterloo.ca, at the University of Waterloo.

formats, ignoring that all formats require knowledge, time, and remediation to be made accessible. Using PDFs is a good accessibility strategy because all documents commonly created in academic libraries can be easily output in a standard format and remediated according to consistent practices.

## **Advantages of PDF**

The PDF was designed to be a universal and accessible format type. The foremost advantage of PDFs is that they may be read by anyone with a freely downloadable PDF reader software (or free browser plug-in). The PDF format is also flexible in offering a clear structure (like HTML) that dictates correct reading order and allows for robust alt-text programming.

Beyond these considerations, PDF offers a stability and permanence that HTML does not. The HTML of webpages can be downloaded and kept as a local copy, but the downloadability of PDFs is far more robust and recognized. Studies on scholar workflows indicate that scholarly materials must be readily available in stable, unalterable, downloadable formats.<sup>3</sup> This need is partially fulfilled by Word documents, but these are inherently presented within an editing software with a design that is not meant to dissuade downloaders from making their own edits.

## **Accessibility requires resources**

A key realization is that unremediated PDFs will always fail accessibility requirements: no work has been done on them yet! This realization reiterates a central aspect of accessibility work that cash-strapped “do more with less” academic libraries continue to grapple with: there is no magic solution; accessibility always requires resources. When properly remediated, PDFs may provide a highly accessible experience. Remediation may require academic libraries to make strategic investments in acquiring Adobe Acrobat Pro and in training employees, but this is the persistent requirement of accessibility compliance—not something that is resolved by jettisoning the PDF. It may be tempting to emphasize Word documents as the text document format of accessible choice because it aligns so well with existing expertise and workflows; however, it would be a disservice to the vision of greater accessibility to ignore the untapped potential of PDF documents.

## **PDF: Meeting users where they are**

A key benefit of PDF materials is their ease of use and ubiquity within the digital information environment. PDF is so valuable to users because of its flexibility and ease of use across any type of digital device (e.g., phones, tablets, and e-reading devices). We also know many of our vendor-supplied materials are provided in PDF form. By leaning into the possibilities of PDF, we have stronger positioning to make needed or requested accommodations for our learners and to advocate for better accessibility functions from publishers.

## **A new way forward**

The biggest complaint about PDFs is that they are not inherently accessible—not born that way. What is ignored in this conversation is that *no* format is born accessible. As a profession, we need to consider that PDFs are a useful format widely adopted across our sectors. Then we must consider how we can set standards within our organizations and wider li-

brary community for accessibility compliance. It is time for us to let go of our collective insistence on a singular, easily understood and implemented solution to accessibility and work toward a future of shared knowledge, tagging conventions, and accessible PDFs.<sup>4</sup> ❧

## Notes

1. Aneta Kwak and Jeffrey Newman, “An Accessibility-First Approach to Online Course Readers,” *Reference Services Review* 46, no. 3 (2018): 340–49, <https://doi.org/10.1108/RSR-04-2018-0046>; Julia Caffrey-Hill, Julia, Nathan Clark, Brent Davis, and William Helman, “PDF: The ‘P’ Stands for Problematic,” *Weave: Journal of Library User Experience* 4, no. 1 (2021), <https://doi.org/10.3998/weaveux.279>; Nosheen Fayyaz, Shah Khusro, and Shakir Ullah, “Accessibility of Tables in PDF Documents,” *Information Technology and Libraries* 40, no. 3 (2021): 1–20, <https://doi.org/10.6017/ital.v40i3.12325>.

2. M. Chee, Z. Davidian, and K. D. Weaver, “More to Do than Can Ever Be Done: Reconciling Library Online Learning Objects with WCAG 2.1 Standards for Accessibility,” *Journal of Web Librarianship* 16, no. 2 (2022), <https://doi.org/10.1080/19322909.2022.2062521>.

3. Ellen Collins, *Information Practices in the Physical Sciences: A Quantitative Study to Evaluate the Impact of Digital Technologies on Research in the Physical Sciences* (Bristol, UK; London: Research Information Network; IOP Publishing, 2015), <http://iopublishing.org/img/news/RIN-info-practices-report.pdf>; Monica Bulger, Eric T. Meyer, Grace de la Flor, Melissa Terras, Sally Wyatt, Marina Jirotko, Katherine Eccles, and Christine Madsen, “Reinventing Research? Information Practices in the Humanities,” Research Information Network Report, April 2011, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1859267](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1859267).

4. S. Voichita, J. Kholodova, M. Chee, and K. D. Weaver, “PDF Accessibility: Considerations and Best Practices for Learning Objects and Other Library Materials,” 2022 ACRL Distance and Online Learning Virtual Poster Presentation, April 25-29, <https://uwspace.uwaterloo.ca/handle/10012/18126>.