

# JLSC

ISSN 2162-3309 | JLSC is published by the Iowa State University Digital Press | <http://jisc-pub.org>

**Volume 11, 1 (2023)**

## **Actually Accessible Data: An Update and a Call to Action**

Randy Colón, Abigail Goben & Sebastian Karcher

Colón, R., Goben, A. & Karcher, S. (2023). Actually Accessible Data: An Update and a Call to Action. *Journal of Librarianship and Scholarly Communication*, 11(1), eP15449. <https://doi.org/10.31274/jlsc.15449>



© 2023 The Author(s). This is an open access article distributed under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

## THEORY ARTICLE

# Actually Accessible Data: An Update and a Call to Action

Randy Colón

*University of Illinois Chicago*

Abigail Goben

*University of Illinois Chicago*

Sebastian Karcher

*Qualitative Data Repository, Syracuse University*

## ABSTRACT

As funder, journal, and disciplinary norms and mandates have foregrounded obligations of data sharing and opportunities for data reuse, the need to plan for and curate data sets that can reach researchers and end-users with disabilities has become even more urgent. We begin by exploring the disability studies literature, describing the need for advocacy and representation of disabled scholars as data creators, subjects, and users. We then survey the landscape of data repositories, curation guidelines, and research-data-related standards, finding little consideration of accessibility for people with disabilities. We suggest three sets of minimal good practices for moving toward truly accessible research data: 1) ensuring Web accessibility for data repositories; 2) ensuring accessibility of common text formats, including those used in documentation; and 3) enhancement of visual and audiovisual materials. We point to some signs of progress in regard to truly accessible data by highlighting exemplary practices by repositories, standards, and data professionals. Accessibility needs to become a main-stream component of curation practice included in every training, manual, and primer.

**Keywords:** digital accessibility, FAIR data, research data

Received: 06/23/2022 Accepted: 05/25/2023



© 2023 The Author(s). This is an open access article distributed under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

## IMPLICATIONS FOR PRACTICE

1. To ensure that research data is accessible, accessibility must be built into curation and deposit guidelines for data repositories as well as applicable primers and standards.
2. Accessibility requires awareness and training; educating researchers and curators is an important part of making data accessible.
3. Ensuring actually accessible research data is an ongoing process of evaluating and improving tools, resources, and workflows. That process must involve disabled people as central actors at all steps: “Nothing about us without us” ([Charlton, 1998](#)).

## INTRODUCTION

In a 2015 article, Walker and Keenan highlighted the importance of providing “truly accessible research data,” i.e., research data that are not merely available but accessible to all users, including users with disabilities ([Walker & Keenan, 2015b](#)). In the 8 years since, the conversation about accessible research data that Walker and Keenan hoped to start has, as we will show in this article, mostly not occurred. We hope to reenergize and refocus this conversation.

We begin our discussion by defining accessibility. In contrast to the widely used understanding of “accessible” for research data in the FAIR (i.e., findable, accessible, interoperable, reusable) Principles, we understand “actually accessible” data to be data that is easy to locate, obtain, interpret, use, share, and analyze for everybody, including disabled people. We then discuss the importance of such accessibility: many users and re-users of data are disabled and rely on digital accessibility of data. Moreover, without explicit attention to their accessibility, digital resources, including research data, default to being inaccessible (see, e.g., [Azadbakht et al., 2021](#), on the parallel case of open education resources). We highlight this concern by considering three different areas related to research data and accessibility: 1) the dearth of consideration for accessibility in the existing literature of data curation; 2) attention to accessibility in research-data-related standards and guidance; and 3) the digital accessibility of data repositories themselves.

Having outlined both the importance of the topic and the overall poor state of accessibility for research data, we turn toward an agenda for action. We suggest three high-impact, reasonable steps to drastically improve accessibility of research data in the short-term: 1) ensuring that data repositories meet Web accessibility standards; 2) implementing curation or oversight so that commonly used data formats are accessible; and 3) adding supplementary information

to aid the accessibility of audiovisual data. We end by reiterating a call to action for key stakeholders and pointing to future areas of research.

## Defining accessibility

In the context of research data, “accessible” is frequently defined in the context of the FAIR Principles (Boeckhout et al., 2018). For the FAIR Principles, “accessible” means that either the data can be accessed using a standard automated protocol (such as hypertext transfer protocol [HTTP] or secure file transfer protocol [SFTP]) or that instructions for access, e.g., for restricted data, are clearly specified in machine-readable metadata. However, FAIR’s notion of accessibility is incomplete. Although it purports to describe data accessible for humans and machines, it fails to account for the diversity of human users and makes no mention of deliberately including disabled people. In this paper, we understand “actually accessible data” to be data that are easy to locate, obtain, interpret, use, share, and analyze for everybody, including disabled people. In the United States (US), this notion of accessibility could be umbrellaed under the idea of compliance with the Americans with Disabilities Act (ADA) and the Rehabilitation Act. Internationally, many countries, including all European Union-member states (although not the US) are signatories to the United Nations’ Convention on the Rights of Persons with Disabilities (CRPD). More narrowly, digital accessibility is codified in the US in Section 508 of the Rehabilitation Act and in the European Union (EU) with the Web Accessibility Directive (EU 2016/2102) for the public sector and the more recent European Accessibility Act (EU 2019/882) for the private sector. These international legal frameworks for digital accessibility differ in important ways, but they reference a single technical framework, the Web Content Accessibility Guidelines (WCAG).<sup>1</sup> As a consequence, although this article most closely investigates the US case, most of its lessons apply globally. Within libraries, as identified by Kumbier and Starkey, the American Library Association’s Core Values of Librarianship may be expanded to the integration of disability justice with an obligation to provide both equitable and equal access to the various forms of information made available in our collections (Kumbier & Starkey, 2016).

In this paper, we define “accessibility” as the ease with which people, including those with disabilities, can locate, obtain, interpret, use, share, and analyze data or information. This inclusive understanding of accessibility aligns with various international legal frameworks and guidelines such as the ADA, Rehabilitation Act, CRPD, and WCAG.

---

<sup>1</sup> Specifically, Section 508 of the Rehabilitation Act in the United States references WCAG 2.0 and the Harmonized European Standard for digital accessibility WCAG 2.1.

## The importance of accessibility

This paper offers advice related to accessibility in the context of research data and digital resources. We focus primarily on blind and low-vision users, who often encounter significant obstacles to access digital resources on the Web and frequently rely on both screen readers and keyboard navigation. However, by emphasizing standard compliance, many of our recommendations apply more broadly. E.g., screen readers are also used by some researchers with dyslexia and other cognitive disabilities (WebAIM, 2021), and keyboard navigation is also essential for users with a range of mobility issues. Considering that there are an estimated 1 billion people with disabilities worldwide, addressing accessibility is crucial to ensure equitable access to information for all. Despite existing standards and guidelines, a significant portion of online resources and databases are not fully accessible, which raises concerns for academic libraries and highlights the need for continued efforts to improve accessibility in digital resources.

Although some progress has been made, the access needs of disabled users have not yet been fully met (Goggin, 2021; Kazuye Kimura, 2018). For example, although there have been accessibility standards in place for decades, the vast majority of online resources and databases are not compliant: the annual WebAIM survey of the most popular million websites finds that less than 5% of sites fully comply with Web accessibility standards (WebAIM, 2023). This raises concerns for academic libraries, which have a variety of roles related to content access and specific legal and ethical obligations to facilitate access to all users (Wentz et al., 2023). Small-scale studies in which testers with disabilities attempt to use digital library content still find accessibility issues in library services (Beyene, 2018; Mulliken, 2019). This is particularly concerning with regard to library databases, which are relied upon by students and researchers with disabilities for access to scholarly materials for coursework and scholarship. Additionally, some vendors have been slow to respond to accessibility feedback, even when it is publicly available. For example, the Library Accessibility Alliance contracted out accessibility checks of 99 e-resources and made the results public (Library Accessibility Alliance, 2023b). Although vendors were given the opportunity to respond to the feedback and make improvements, many have not done so. This is illustrated by the case of JSTOR, which was found to have screen reader issues but did not respond to the report (Library Accessibility Alliance, 2023a). The fact that there is a need for such accessibility checks in the first place is a clear indication of the ongoing problem of inaccessibility in digital resources.

There is nothing inherently inaccessible about digital data itself; in fact, many people with disabilities are proficient with the technologies needed to access data (Beyene, 2018). Instead, decisions regarding the organization, formation, and portrayal of data can determine whether the content will be accessible (Kazuye Kimura, 2018). For while accessibility has become

a growing concern in recent years, it is important to recognize that the Web in general was not designed with accessibility in mind (Kazuze Kimura, 2018). As a result, we are often left with starting points that need to be retrofitted to be made accessible later rather than starting out accessible from the outset. One of the reasons posited for the lack of accessibility in digital settings is the insufficient inclusion of people with disabilities in digital social settings, digital public spheres, and digital cultures (Goggin, 2021). Going forward, it is critical that we work to make things accessible from their creation and include disabled people from the beginning rather than having to retrofit or retroactively consult them. Additionally, new digital technologies are often lauded for their innovative features, but, frequently, accessibility has not been considered. This can make past gains in accessibility irrelevant if they cannot be transferred to the new systems. By recognizing these issues and taking proactive steps to address them, we can work to make digital data and data repositories that are accessible to all regardless of disability status or impairment.

Moreover, disability advocates are calling for large institutions such as universities and libraries to only work with vendors who adhere to accessibility guidelines (Kazuze Kimura, 2018). This, coupled with changing accessibility policies worldwide, and the potential to face litigation when accessibility guidelines are not being met, make it beneficial pragmatically to proactively work toward greater accessibility (Wentz et al., 2023). Although accessibility is generally thought of as benefitting the disability community, resources created with disability needs in mind can also benefit the public at large (Goggin, 2021; Vollenwyder et al., 2019). For instance, making digital repositories easier to access visually can increase the usability of the platform overall, not just for those with visual disabilities. Additionally, many people using screen readers prefer content to be as streamlined as possible, making it easier for them to access material using keyboard (rather than mouse) navigation. At the same time, many users, especially users with learning disabilities, express that online resources are overly complex and difficult to navigate (Beyene, 2018). Thus, thinking through how to streamline content for a screen reader user may not only benefit the user themselves but also improve the overall functionality of the content.

### **One step toward accessibility**

One paper cannot address every aspect of accessibility needs for researchers who are seeking to work with innumerable formats of data. Among the most obvious issues that we do not address is that of physical access, in which barriers in the built environment prevent people with disabilities from being able to travel to archives, special collections, museum collections, or physical samples. However, working beyond that, researchers, data librarians, and data curators are faced with extensive varieties of digital data in which a lack of accessibility creates impenetrable barriers. Some familiar examples are spreadsheets without intelligible metadata,

digital scans of texts without an optical character recognition (OCR) or text layer, images without descriptive metadata, or digital humanities exhibits that fail to implement universal or inclusive design.

As a beginning place for those engaged in digital data curation, we propose two groups to consider: people with visual disabilities who often use screen readers or enhancers; and users who predominantly use a keyboard to navigate through websites, spreadsheets, and other data formats.

## **THE STATUS QUO: LIMITED PROGRESS TOWARD ACCESSIBILITY**

Since Walker and Keenan's 2015 call to action, we have seen no substantial discussion of research data accessibility in any mainstream publication. In this section, we hope to reopen the discussion on data accessibility by describing progress along three dimensions: 1) the literature on research data and research data curation; 2) major standards for data repositories; and 3) the accessibility of data repositories and repository software.

### **Accessibility in the data curation literature**

As of this writing, Google Scholar finds a total of five citations to the article by Walker and Keenan. A range of searches in Library, Information Science, and Technology Abstracts (LISTA); Google Scholar; and Scopus did not yield any publication directly addressing accessibility of research data for people with disabilities or the role of data curation for such accessibility (example searches are provided in the supplementary materials), which aligns with the broader challenge of identifying research generally on disability studies (Brunskill, 2021). In academic library literature, the focus of papers related to accessibility either centered the idea of availability or were focused on library website accessibility or service provision (Brunskill et al., 2021; Liaw et al., 2021; Máchová et al., 2018; Pulverer, 2015; Volentine et al., 2015, Yang et al., 2020).

When addressing accessibility, little attention has been paid to research data specifically. Nevertheless, work on the accessibility of library resources and institutional repositories may hold relevant lessons. Prior to their call for accessible data, Walker and Keenan had explored the accessibility of two commercial products that serve as institutional repositories or hold digital collections (Walker & Keenan, 2015a). McLaughlin and Hoops (2021) focus on accessibility for institutional repositories, although they do not address data but instead focused on manuscript submissions, primarily in Microsoft Word or portable document format (PDF) format. Similarly, Rodriguez (2021) describes the process of identifying and improving the accessibility of audiovisual materials in an institutional repository.

Data librarians and data curators can draw on and expand the more general efforts to make the digital content available through libraries more accessible.

However, when writing about research data directly, digital accessibility is rarely considered. Several major academic works on data curation make no reference to data curation for accessibility (Hudson-Vitale et al., 2017; Johnston, 2016), nor do any of the documents on data preparation and/or curation published by social science data repositories (including that co-authored by one of the co-authors of this paper) make any mention or provide any guidance on making research data truly accessible (Demgenski et al., 2021; Van den Eynden et al., 2011; Inter-university Consortium for Political and Social Research [ICPSR], 2021).

There are some notable exceptions to this bleak picture. The Data Curation Network has recently published a dedicated primer for accessibility (Oxford & Woodbrook, 2023), and the curation primer for R (Kellam et al., 2019) contains specific sections on data accessibility, although none of the other curation primers contain such a section as of this writing.

An important area that has seen significant advances over the last few years is the accessibility of data visualizations; however, this is arguably the least accessible display of data for people who have visual impairments. Even basic steps such as providing alternative text (alt text) for data visualization remains rare, as documented, e.g., by Canelón and Hare focusing on “tidy Tuesday” visualizations (Canelón & Hare, 2021), and may not be regularly included in training materials such as the checklists provided by Stephanie Evergreen (Evergreen et al., 2022). The lack of usable alt text for data visualizations extends to high-profile publications. For example, the journal *Cell* includes only “Figure thumbnail fx1” as the alt text for figures, including those without any meaningful caption.

Another area of innovation is tactile or sound representations of data visualizations. Since the publication of the landmark *Guidelines and Standards for Tactile Graphics* (BANA & CBA, 2011), a growing number of such resources exist, although, measured as a share of the total amount of data, they remain a rarity. Providing tactile options is of particular importance where alternative modes of displays of data (such as tables for qualitative data) are not readily available. Additionally, when tables are available, they should be formatted to be readable by a user’s screen reader (e.g., with headers in the top row and left column).

### **Accessibility in standards for data repository and research data**

Despite requirements for data sharing, disciplinary focus on reproducibility, and concern about the reliability of scholarly data and the findings arising from them, little attention

has been paid, either in federal policy in the United States or in widely used standards, on the need to consider researchers with disabilities.

This lack of explicit consideration is also apparent in the data curation and data repository standards that are becoming common shorthand for researchers. The FAIR Principles (Wilkinson et al., 2016) were developed as a set of minimal practices toward data sharing and reuse and have been quickly codified into data policies both in the United States and in the European Union (GO FAIR, 2022). These guidelines are primarily from the perspective of facilitating big data and improving automated access to data sets, with limited consideration for the need to protect and moderate data access or to avoid data exploitation. “Accessible,” in FAIR, is therefore defined primarily in terms of whether a human or a computer can access a resource through standardized protocols. Neither the FAIR Principles nor any of their operationalizations that we are aware of make any mention of accessibility of research data to disabled users.

The FAIR approach to data has been highly influential. For example, the most recent US federal agency to release their guidance and requirements for data management and sharing is the National Institutes of Health (NIH). The NIH’s guidelines specifically refer to the need for data to be FAIR but do not require researchers to consider how they will enhance access across abilities (NIH, 2020a, 2020b). We find a similar lack of reference to accessibility for all users in the desirable characteristics for data repositories put forth by the US White House Office of Science and Technology Policy (Nelson, 2022), as well as by the requirements for the Core Trust Seal, the most widely used certification for research data repositories (CoreTrustSeal Standards and Certification Board, 2019).

### **Accessibility of research data repositories**

A step toward accessible data is accessible repositories. In spite of existing standards for Web accessibility and the legal requirement to implement these in many countries (e.g., under Section 508 of the Rehabilitation Act in the United States or the European Union’s Web Accessibility Directive [2016/2102]), the websites of many academic institutions (Acosta-Vargas et al., 2017), libraries (Spina, 2019), and databases provided through academic libraries (Falloon & O’Reilly, 2020; Rysavy & Michalak, 2020) remain inaccessible.

Research repositories present a mixed picture in terms of accessibility. For a snapshot of the state of accessibility in data repositories, we looked at seven leading data repositories: four generalist repositories; and three leading social science data archives. We ran automated accessibility checks using both the WAVE and AXE tools for a data page and a search page for each repository. Such automated accessibility evaluation tools have significant shortcomings.

Most notably, their coverage of WCAG 2.0 success criteria is poor; independent benchmarking studies find that any individual tool only covers between 12% to 35% of WCAG 2.0 success criteria (Abduganiev, 2017; Vigo et al., 2013). That said, most automated tools have relatively low levels of false positives and perform comparatively well on the most critical (WCAG level A) criteria (Vigo et al., 2013). Thus, automated tests do provide a reasonable preliminary benchmark for assessing the relative accessibility of websites, even though that must always be complemented by manual testing to ensure the accessibility of any individual website.

No repository passed all automated tests (Table 1). Five repositories (figshare, Dryad, Harvard Dataverse, ICPSR, UK Data Archive) have less than 20 combined errors in WAVE, as well as less than 10 critical errors in AXE,<sup>2</sup> suggesting some attention to accessibility in the development process. Two repositories (Zenodo, GESIS) have a large number of errors, including critical errors, on each page type.

We also reviewed development guidance for six widely used open-source data repository systems. Four projects had dedicated documentation related to accessibility. Fedora and DSpace provide detailed descriptions of WCAG goals and related coding approaches. Dataverse had a more general description of goals and automated approaches for testing accessibility. Samvera provides little explicit guidance for accessible development but is the only

Repository	Errors in WAVE	Critical Issues AXE	Serious Issues AXE	Most Serious Error
Figshare	1	5	0	ARIA roles contained in parents
Dryad	1	0	6	Empty heading
ICPSR	5	22	6	Alt text/form element
Zenodo	297	67	168	Alt text/form element labels
Harvard Dataverse	17	2	24	Empty links
UKData	17	4	4	Form elements labels
GESIS	839	9	93	Select element must have accessible name

**Table 1.** Total errors found by AXE and WAVE accessibility checks for seven popular data repositories. Errors combined for search results and data landing page. Errors include low contrast errors. See supplementary materials for the full table including tested pages.

<sup>2</sup> Manual inspection of the AXE results for ICPSR suggests a number of false-positive errors, likely caused by misinterpreting html elements generated using react.js.

repository describing dedicated accessibility testing (including expert testing using screen readers). CKAN and Eprints make only cursory references to accessibility in their developer documentation (see supplementary materials).

## **TOWARD MORE ACCESSIBLE DATA: OPPORTUNITIES FOR DATA REPOSITORIES**

The amount of research data deposited in repositories is growing rapidly, and making all of these data accessible at the highest standards of accessibility is a daunting task: so daunting that we worry it may deter repositories from even taking meaningful first steps. We suggest three pragmatic steps to get started and to make the majority of our resources more accessible. To be clear, partial accessibility is not sufficient, but given the task at hand, we believe identifying and prioritizing high-impact efforts is nonetheless useful.

### **Making data repositories accessible**

Making the Web platforms through which we provide access to data fully accessible (defined as level AA of WCAG 2.1) should be a minimal, and short-term achievable, standard. Resources for internal accessibility audits are readily available, e.g., from the A11y group or the WAVE Web Accessibility Evaluation Tool ([The A11Y Project, n.d.](#)). As part of this, repositories should have a dedicated accessibility policy clearly linked from the homepage that will ideally include an assessment of Web accessibility, such as a completed Voluntary Product Assessment Template (VPAT) particularly for general and domain repositories serving multiple institutions ([O'Reilly, 2020](#)). The accessibility policy and VPAT provided by ICPSR are exemplary in that regard ([ICPSR, 2020](#)).

Additionally, for people with visual disabilities, providing options on the website, such as the ability to change font or color scheme, can increase accessibility and enhance the user experience. This may also help people with learning disabilities and other print disabilities better navigate the material ([Beyene, 2018](#)).

Although WCAG specifications provide a useful and important benchmark for reaching accessibility, when testing websites and content for usability, it is highly recommended to include people with disabilities in the testing and design process ([Beyene, 2018](#); [Goggin, 2021](#); [Grenon et al., 2021](#); [Kazuye Kimura, 2018](#); [Vollenwyder et al., 2019](#)). In part, this is because websites can be technically accessible but still contain accessibility issues, and, conversely, some technical issues may not significantly impact the effectual accessibility of the content ([Kazuye Kimura, 2018](#)). Therefore, it is recommended to involve people with disabilities in the design process as much and as early as possible. This can aid in

troubleshooting concerns that are not captured by technical standards while also helping website developers prioritize potential concerns.

### **Ensuring accessibility of common data formats**

Traditional quantitative data often meet minimal accessibility standards by virtue of being highly structured, plain-text formats (e.g., comma-separated values [csv]). Beyond that, and thankfully for digital curators, the accessibility of data formats correlates strongly with its suitability for digital preservation (we thank Kate Flynn, personal communication, for highlighting this): the less widely used, less standardized, and more proprietary a format, the less accessible it tends to be. However, even within common data formats, there is significant variation. Videos, PDFs, Microsoft Word documents, and presentations all come in more or less accessible formats. Even for quantitative data sets, codebooks and documentation are often formatted in Microsoft Word, Microsoft Excel, or PDF and may therefore be more or less accessible. Therefore, we recommend curating repositories to use best practices (and existing tools) when archiving common file formats.

- Microsoft Office tools include built-in accessibility checks for Microsoft Word (.docx), Excel (.xlsx), and PowerPoint (.pptx) files (Microsoft, 2021). For most Office files found on data repositories, such as documentation or codebooks, accessibility is relatively easy to ensure. Accessibility becomes more challenging as files become more complex, especially when they include complex formatting, images, and other graphical elements.
- PDFs are designed to ensure visual consistency across devices. This principally graphical approach to displaying information can make accessibility difficult. However, since 2012 a dedicated International Organization for Standardization (ISO) standard (updated 2014, confirmed 2020) specifies “universally accessible” PDFs (PDF/UA; ISO, 2014), which, since 2015, is also the accepted standard for Section 508 in the United States. Adobe Acrobat includes an accessibility check tool that can check for and assist with PDF/UA compliance (Adobe, 2019).

Automated checks cannot fully replace human/expert checks on the accessibility of documents and websites, but they can serve as a baseline for accessibility and prevent a large number of common errors, often with relatively limited effort.

### **Providing supplementary information for multimedia files**

Multimedia files are increasingly common in data repositories, both as qualitative data and as documentation. They pose particular challenges for users who are blind or have low vision

as well as those who are deaf or hard of hearing. Wherever possible, repositories should include auxiliary information that makes these files more accessible. For audio files containing spoken words, transcripts are strongly recommended. For video files, closed captions should be included in the file itself. Automated transcription has made significant progress over the past years, but even the best automated transcripts still have significant shortcomings and should, whenever possible, be reviewed by humans or replaced by human transcripts. Depending on the nature of the material, providing separate transcript files in addition to captions on the videos themselves can be advisable and can, as a secondary benefit, also increase findability.

Image files do not have built-in accessibility features such as videos. Creating separate image description files for every image included in a data set can become unwieldy. Curators can consider including image descriptions in the file metadata, although they may conflict with other purposes of descriptive metadata. An intriguing suggestion by Nell Chitty is to include “alt-text” in the image file property’s description (Chitty, 2016). This strategy can make the alt text harder to discover, but it keeps it connected to the file itself, facilitating automated workflows and consistency.

## **SIGNS OF PROGRESS**

We started this article noting the lack of recent and significant action toward improving accessibility. And, indeed, progress toward universally accessible research data over the last 8 years has been unsatisfactory. Nevertheless, there are several examples of repositories and other organizations that have taken exemplary steps worth highlighting and emulating.

### **ICPSR: Website accessibility**

The footer of ICPSR’s website includes a prominent link to the repository’s website accessibility statement. The statement commits to Section 508 compliance, explains how that is achieved technically, and provides contact information for issues with accessibility. ICPSR goes further than most sites by publicly providing a copy of their completed VPAT, which provides detailed information about compliance with WCAG standards (ICPSR, 2020).

### **Cornell University repository: Requiring accessible data**

Cornell University’s eCommons repository requires all new data deposits to conform to current accessibility guidelines. The repository’s guidance specifies necessary steps for common file formats such as Microsoft Word, PDF, and EPUB (Cornell eCommons, 2022b). In addition, eCommons offers a dedicated controlled vocabulary that signals the accessibility

of individual data projects. The metadata describe existing accessibility features (such as “largePrint” or “captions”), as well as accessibility hazards such as “flashing” (Cornell eCommons, 2022a).

### **Data curation primers: Mainstreaming accessibility in curation**

The Data Curation Network’s (DCN) data curation primers are a unique, open, community-generated resource that seeks to establish a corpus of evolving standards to help curators understand and better curate increasingly numerous and complex data formats. As noted earlier, few primers contain specific advice on accessibility. In one example, Emily Oxford added an explicit section on “Accessibility Considerations” to the R-data curation primer (Kellam et al., 2019). The Wordpress.com primer includes some accessibility considerations as part of its general discussion of “FAIRness.” However, Oxford and Woodbrook (2023) recently released an Accessibility Data Primer, which suggests some initial steps for foregrounding accessibility during the curation and preservation process across any data curation process. Although accessibility is presently only covered in a small number of the DCN primers, we believe that this is an important step; the mainstreaming of accessibility in curation will only happen if accessibility becomes a standardized step in curation workflows.

### **CONCLUSION: A CALL TO ACTION**

Making data truly accessible requires making accessibility a standard component of data-related workflows. It requires action by and participation of data producers, data repositories, data curators, data librarians, information professionals shaping standards for data and repositories, and research policy makers deciding on standard implementation. As we face more data preservation and sharing requirements from funders, publishers, institutions, and our communities, the consideration and active planning for accessibility measures will only increase. “Accessibility” is used in a variety of ways to describe a variety of data-related activities. It is critical that we examine not only the definitions of accessibility that address discoverability and whether files can be downloaded but that we center the ability of all researchers to find, analyze, and use data. We must revisit and expand data standards such as FAIR to prevent actions or decisions that exclude disabled users. Compliance requirements such as WCAG, the US ADA, or the EU’s European Accessibility Act serve only as a threshold floor and should be considered as our starting place for exploring how to expand and ensure equitable access.

In addition, there are significant educational needs to improve accessibility for data sharing and curation. Among the audiences for training are data producers of all types, researchers and businesses who are engaging with the data, and curators. Screen reader and captioning accessibility, as well as other checkpoints, should be included in curation handbooks and

guidance. It is critical to foreground a definition of accessibility that promotes access to data by diverse users such as those who use screen readers, tabbed navigation, and transcript files in order to capture the greatest advances of knowledge from data. Exemplar data repositories have begun to provide accessibility information, but deliberate attention and training is needed to ensure that, in another 8 years, we are not still lamenting the need for foundational improvement.

As we begin to address these challenges, and when we can influence choices over which vendors we work with, it will be important to work only with vendors who provide accessible content. By doing so, we can not only protect our institutions from falling out of compliance with the ADA and other policies and standards, but also, by tying accessibility to purchasing decisions, we can use our buying power to advocate for more inclusive practices (Beyene, 2018).

Such changes in practices, procedures, and policies benefit from and require institutional commitment at all levels. They also require the active inclusion of disabled people, as researchers, developers, librarians, throughout. This is, in part, a pragmatic concern: where accessibility policies and technologies are developed without disabled people, they often result in “disability dongles,” Liz Jackson’s description of “[a] well intended elegant, yet useless solution to a problem we never knew we had” (Jackson et al., 2022). However, more fundamentally, “Nothing about us without us (...) forces political-economic and cultural systems to incorporate people with disabilities into the decision-making process and to recognize that the experiential knowledge of these people is pivotal in making decisions that affect their lives” (Charlton, 1998, p. 17). It is essential in recognizing disability rights as human rights.

## ACKNOWLEDGMENTS

The authors thank Emily Oxford and Rachel Woodbrook for early conversations on this topic; participants at the Research Data Access and Preservation (RDAP), International Digital Curation (IDCC), and STEM Librarians Collaborative for comments on a related presentation; and two anonymous reviewers for helpful comments on an earlier version of this paper. Karcher’s contribution is based upon work supported by the US National Science Foundation, grants 1823950 and 2116935.

## SUPPLEMENTARY MATERIALS

Supplementary materials are available from Zenodo as Colón, Randy, Goben, Abigail, & Karcher, Sebastian. (2023). Supplementary materials for: Actually accessible data: An update and a call to action [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7968784>

## REFERENCES

- Abduganiev, S. G. (2017). Towards automated web accessibility evaluation: A comparative study. *International Journal of Information Technology and Computer Science*, 9(9), 18–44. <https://doi.org/10.5815/ijitcs.2017.09.03>
- Acosta-Vargas, P., Luján-Mora, S., & Salvador-Ullauri, L. (2017). Web accessibility policies of higher education institutions. *2017 16<sup>th</sup> International Conference on Information Technology Based Higher Education and Training (ITHET)*, 1–7. <https://doi.org/10.1109/ITHET.2017.8067808>
- Adobe. (2019 November 24). *Using the Acrobat Pro DC Accessibility Checker*. <https://www.adobe.com/accessibility/products/acrobat/using-acrobat-pro-accessibility-checker.html>
- Azadbakht, E., Schultz, T., & Arellano, J. (2021). Not open for all: Accessibility of open textbooks. *Insights*, 34(1), 24. <https://doi.org/10.1629/uksg.557>
- BANA & CBA. (2011). *Guidelines and standards for tactile graphics*. Braille Authority of North America; Canadian Braille Authority. <http://www.brailleauthority.org/tg/web-manual/index.html>
- Bejene, W. M. (2018). Digital inclusion in library context: a perspective from users with print disability. *Journal of Web Librarianship*, 12(2), 121–140. <https://doi.org/10.1080/19322909.2018.1427657>
- Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: Fair enough? *European Journal of Human Genetics: EJHG*, 26(7), 931–936. <https://doi.org/10.1038/s41431-018-0160-0>
- Brunskill, A. (2021). Disability studies research literature: it's (mostly) not where we think. *Portal Libraries and the Academy*, 21(1), 81–97. <https://doi.org/10.1353/pla.2021.0006>
- Brunskill, A., Lantz, C., & Mundle, K. (2021). What information are we providing to users with disabilities? An analysis of ARL libraries' accessibility webpages. *College and Research Libraries*, 82(7). <https://doi.org/10.5860/crl.82.7.935>
- Canelón, S., & Hare, L. (2021 May 4). *Revealing room for improvement in accessibility within a social media data visualization learning community*. csv,conf.6. <https://github.com/spcanelon/csvConf2021>
- Charlton, J. I. (1998). *Nothing about us without us: Disability oppression and empowerment*. University of California Press. <http://www.jstor.org/stable/10.1525/j.ctt1pnqn9>
- Chitty, N. (2016 December 13). How to add ALT text to image metadata. *Through My Eyes*. <https://throughmyeyesinclusivedesign.wordpress.com/2016/12/13/adding-alt-text-to-image-metadata/>
- CoreTrustSeal Standards and Certification Board. (2019). *CoreTrustSeal trustworthy data repositories requirements 2020–2022*. <https://doi.org/10.5281/zenodo.3638211>
- Cornell eCommons. (2022a). *Creating metadata for batch upload: Accessibility metadata*. <https://guides.library.cornell.edu/ecommons/batchmetadata>

- Cornell eCommons. (2022b). *Web accessibility*. <https://guides.library.cornell.edu/ecommons/accessibilitypolicy>
- Demgenski, R., Karcher, S., Kirilova, D., & Weber, N. (2021). *QDR curation handbook. Qualitative Data Repository*. <https://doi.org/10.5281/zenodo.4672678>
- Evergreen, S., Sanjines, S., Emery, A. K., & Lyons, J. R. (2022 February 8). Data visualization checklist. *Evergreen Data*. <https://stephanieevergreen.com/data-visualization-checklist/>
- Falloon, K. A., & O'Reilly, F. M. (2020). Prioritizing accessibility in the e-resources procurement lifecycle: VPATs as a practical tool for e-resource acquisitions and remediation workflows at academic libraries. *The Serials Librarian*, 78(1–4), 130–140. <https://doi.org/10.1080/0361526X.2020.1722020>
- GO FAIR. (2022). *FAIR Principles*. <https://www.go-fair.org/fair-principles/>
- Goggin, G. (2021). Disability, Internet, and digital inequality: The research agenda. In Hargittai, E. (Ed.), *Handbook of digital inequality*, pp. 255–273. Edward Elgar Publishing. <https://doi.org/10.4337/9781788116572.00025>
- Grenon, M. M., Ruel, J., Fougereyrollas, P., Normand, C. L., Moreau, A. C., Romero-Torres, A., & Gravel, S. (2021). Conceptualizing access to and understanding of information. *Universal Access in the Information Society*, 1–12. <https://doi.org/10.1007/s10209-021-00836-w>
- Hudson-Vitale, C., Imker, H., Johnston, L. R., Carlson, J., Kozlowski, W., Olendorf, R., & Stewart, C. (2017). *SPEC Kit 354: data curation*. Association of Research Libraries. <http://publications.arl.org/Data-Curation-SPEC-Kit-354/>
- ICPSR. (2020). *Accessibility*. <https://www.icpsr.umich.edu/web/pages/datamanagement/policies/access.html>
- ICPSR. (2021). *Guide to social science data preparation and archiving: Best practice throughout the data life cycle* (6<sup>th</sup> ed.). ICPSR. <https://www.icpsr.umich.edu/files/deposit/dataprep.pdf>
- ISO. (2014). *Document management applications—Electronic document file format enhancement for accessibility* (ISO 14289-1:2014). International Organization for Standardization. <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/45/64599.html>
- Jackson, L., Haagaard, A., & Williams, R. (2022 April 19). Disability dongle. *Platypus: The CASTAC Blog*. <https://blog.castac.org/2022/04/disability-dongle/>
- Johnston, L. (Ed.). (2016). *Curating Research Data*. Association of College and Research Libraries.
- Kazuye Kimura, A. (2018). Defining, evaluating, and achieving accessible library resources: A review of theories and methods. *Reference Services Review*, 46(3), 425–438. <https://doi.org/10.1108/RSR-03-2018-0040>
- Kellam, L., Koziar, K., & Pejša, S. (2019). R data curation primer. *Data Curation Network*. <https://github.com/DataCurationNetwork/data-primers>

- Kumbier, A., & Starkey, J. (2016). Access is not problem solving: Disability justice and libraries. *Library Trends*, 64(3), 468–491. <https://doi.org/10.1353/lib.2016.0004>
- Liaw, S. T., Guo, J. G. N., Ansari, S., Jonnagaddala, J., Godinho, M. A., Borelli, A. J., de Lusignan, S., Capurro, D., Liyanage, H., Bhattal, N., Bennett, V., Chan, J., Kahn, M. G., & de Lusignan, S. (2021). Quality assessment of real-world data repositories across the data life cycle: a literature review. *Journal of the American Medical Informatics Association*, 28(7), 1591–1599. <https://doi.org/10.1093/jamia/ocaa340>
- Library Accessibility Alliance. (2023a February 2). JSTOR. Library Accessibility Alliance. <https://libraryaccessibility.org/evaluation/jstor>
- Library Accessibility Alliance. (2023b March 13). Library Accessibility Alliance. <https://libraryaccessibility.org/node/1>
- Máchová, R., Hub, M., & Lnenicka, M. (2018). Usability evaluation of open data portals. *Aslib Journal of Information Management*, 70(3), 252–268. <https://doi.org/10.1108/AJIM-02-2018-0026>
- McLaughlin, M., & Hoops, J. (2021). Web accessibility in the institutional repository: Crafting user-centered submission policies. *The Serials Librarian*, 80(1–4), 40–45. <https://doi.org/10.1080/0361526X.2021.1868217>
- Microsoft. (2021 September 13). *Improve accessibility with the Accessibility Checker*. Microsoft Support. <https://support.microsoft.com/en-us/office/improve-accessibility-with-the-accessibility-checker-a16f6de0-2f39-4a2b-8bd8-5ad801426c7f>
- Mulliken, A. (2019). Eighteen blind library users' experiences with library websites and search tools in U.S. academic libraries: A qualitative study. *College and Research Libraries*, 80(2), 152–168. <https://doi.org/10.5860/crl.80.2.152>
- Nelson, A. (2022). *Ensuring free, immediate, and equitable access to federally funded research*. Office of Science and Technology Policy. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>
- NIH. (2020a). *Final NIH policy for data management and sharing (NOT-OD-21-013)*. National Institutes of Health. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>
- NIH. (2020b). *Supplemental information to the NIH policy for data management and sharing: Selecting a repository for data resulting from NIH-supported research (NOT-OD-21-016)*. National Institutes of Health. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html>
- O'Reilly, F. (2020). The VPAT as an e-resources assessment tool: Putting accessibility to the test. *Journal of Electronic Resources Librarianship*, 32(1), 52–63. <https://doi.org/10.1080/1941126X.2019.1709749>
- Oxford, E., & Woodbrook, R. (2023). *Accessibility data curation primer* (Data Curation Network Primers). Data Curation Network. <https://github.com/DataCurationNetwork/data-primers/blob/cea103babe4b95952967b52d743b46cac0ac36aa/Accessibility%20Data%20Curation%20Primer/accessibility-data-curation-primer.md>

Pulverer, B. (2015). Data accessibility and reproducibility: Moving to transparent publishing in the biosciences. *Information Services & Use*, 35(3), 185–188. <https://doi.org/10.3233/ISU-150780>

Rodriguez, D. (2021). Increasing accessibility of audiovisual materials in the institutional repository at Florida State University. *The Journal of Academic Librarianship*, 47(1), 102291. <https://doi.org/10.1016/j.acalib.2020.102291>

Rysavy, M. D. T., & Michalak, R. (2020). Assessing the accessibility of library tools & services when you aren't an accessibility expert: Part 1. *Journal of Library Administration*, 60(1), 71–79. <https://doi.org/10.1080/01930826.2019.1685273>

Spina, C. (2019). WCAG 2.1 and the current state of web accessibility in libraries. *Weave: Journal of Library User Experience*, 2(2). <https://doi.org/10.3998/weave.12535642.0002.202>

The A11Y Project. (n.d.). *Checklist: Check your WCAG compliance*. <https://a11yproject.com/checklist/>

Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). *Managing and sharing data: best practice for researchers*. UK Data Archive.

Vigo, M., Brown, J., & Conway, V. (2013). Benchmarking web accessibility evaluation tools: Measuring the harm of sole reliance on automated tests. *Proceedings of the 10<sup>th</sup> International Cross-Disciplinary Conference on Web Accessibility*, 1–10. <https://doi.org/10.1145/2461121.2461124>

Volentine, R., Owens, A., Tenopir, C., & Frame, M. (2015). Usability testing to improve research data services. *Qualitative & Quantitative Methods in Libraries*, 59–68.

Vollenwyder, B., Iten, G. H., Brühlmann, F., Opwis, K., & Mekler, E. D. (2019). Salient beliefs influencing the intention to consider web accessibility. *Computers in Human Behavior*, 92, 352–360. <https://doi.org/10.1016/j.chb.2018.11.016>

Walker, W., & Keenan, T. (2015a). Do you hear what I see? Assessing accessibility of digital commons and CONTENTdm. *Journal of Electronic Resources Librarianship*, 27(2), 69–87. <https://doi.org/10.1080/1941126X.2015.1029395>

Walker, W., & Keenan, T. (2015b). Going beyond availability: Truly accessible research data. *Journal of Librarianship and Scholarly Communication*, 3(2), Article 2. <https://doi.org/10.7710/2162-3309.1223>

WebAIM. (2023 March 29). *The WebAIM million*. <https://webaim.org/projects/million/>

WebAIM. (2021 June 30). *Screen Reader User Survey #9 Results*. <https://webaim.org/projects/screenreadersurvey9/>

Wentz, B., Gorham, U., & Jaeger, P. T. (2023). Academic libraries and their legal obligation for content accessibility. *First Monday*, 28. <https://doi.org/10.5210/fm.v28i1.12892>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M.,

Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ...Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>

Yang, W., Zhao B., Liu, Y. Q., & Bielefield, A. (2020). Are Ivy League library website homepages accessible? *Information Technology & Libraries*, 39(2), 1–18. <https://doi.org/10.6017/ital.v39i2.11577>