Information Research - Vol. 31 No. iConf (2026)

'Metadata is wonderful, but there's a coldness to it': exploring the visual accessibility of cultural heritage through participatory design

Victoria Van Hyning and Pramod Chundury

DOI: https://doi.org/10.47989/ir31iConf64240

Abstract

Introduction. Ensuring access to digital visual cultural heritage (VCH) for people who are blind or low vision (BLV) requires dedicated effort. Many digitised collections lack alt text, transcriptions, or descriptions that would enable independent access.

Method. We conducted participatory design sessions with BLV participants, cultural heritage experts, and stakeholders at the Library of Congress (LOC) to explore how access to VCH collections might be improved. Mixed groups of BLV and sighted participants completed three activities: 1) identifying categories of information valued in image descriptions, 2) evaluating descriptive sources through a Wizard-of-Oz method, and 3) brainstorming to envision future systems.

Analysis. We analysed category rankings from Activity 1, evaluations of descriptive sources from Activity 2, and qualitatively examined insights from one study session of Activity 3.

Results. Participants emphasised the importance of descriptions that extend beyond metadata, including vivid visual details, human presence, and contextual information. Human-provided descriptions were most valued, AI-generated descriptions showed potential but required oversight, and metadata-based descriptions were seen as accurate yet insufficient.

Conclusion. Our findings underscore the need for layered descriptive strategies combining factual precision with interpretive richness, ensuring equitable access for blind users.

Introduction

Since the 1990s cultural heritage objects such as manuscripts, archival materials, prints, photographs, film, and AV materials have been digitised and placed online. It is commonly said that these efforts make collections accessible or democratise access, yet digital access does not automatically mean that content is accessible to people with disabilities. For digitised cultural heritage materials to be accessible, they must, in the case of documents, be described and transcribed for people who cannot see them; and captioned, in the case of AV, for people who cannot hear them, to give just two examples. Moreover, the digital infrastructure hosting these materials must be accessible to assistive technologies through meaningful and consistently applied structure, accessible fonts, color schemes, and so on. Scholars and practitioners have identified significant accessibility barriers in digitised special collections and archives, including lack of alt text for images, lack of transcriptions for documents, and poorly structured pages that are unnavigable by assistive technologies such as screen readers and refreshable braille displays (Southwell and Slater, 2012; Van Hyning et al., 2024). Gracen Brilmeyer documents the harmful representation of disabled people in the archive or their complete erasure in some instances (Brilmyer, 2020). These issues are compounded by the inaccessibility of library discovery systems more widely (Xie et al., 2021; Yoon et al., 2016).

A new rule added to Americans with Disabilities Act (ADA) Title II in 2024, coming into effect in 2026/2027 (Department of Justice, 2024a, 2024b), will require public entities to ensure web and app-based content complies with WCAG 2.1 AA by 2027. While content management systems hosting digital archival and special collections must meet these standards, it remains unclear whether existing digitised content within these systems will be exempt. Regardless, BLV users both desire and deserve access to these materials for research and leisure, on par with sighted users.

This study, '*Enhancing Accessibility of Visual Cultural Heritage through Participatory Design*' was conducted in 2025 amid impending legal changes, but was driven by a longstanding desire to better meet the needs of BLV users of online visual cultural heritage (VCH) from the Library of Congress (LOC). The LOC, one of the world's largest cultural heritage organisations, is also home to the National Library Service (NLS) for the Blind and Print Disabled. The NLS provides access to published content for people with print disabilities. Our interdisciplinary team includes researchers from the LOC, NLS, the University of Maryland, and both blind and sighted contributors.

The LOC was an early adopter of digitisation, and like many cultural heritage organisations, hosts digital images that often lack alt-text or rely on metadata schemas unfamiliar to non-experts. To address this gap, the LOC has pursued crowdsourcing initiatives such as By the People (Van Hyning et al., 2022), and a Flickr project for generating image descriptions of photographs and prints (Library of Congress, 2024); with resulting content treated as a supplement to curatorial metadata (e.g., subject headings). While crowdsourcing has been central to improving engagement and accessibility of visual cultural heritage (VCH), the LOC is also beginning to explore AI-based approaches to enhance collection description. This study was designed to explore the following questions within this specific organisational context, but with the hope that the findings would apply more broadly to digital VCH:

- **RQ1:** What kinds of information do BLV people want to know about visual cultural heritage (VCH), and how much description is desired?
- **RQ2:** What sources of information do BLV people want to access for descriptions of VCH? How do BLV and sighted participants rank the relative merits of off-the-shelf AI-generated descriptions, versus human crowdsourced transcriptions, versus existing curatorial metadata, versus web-searches for additional context or reverse-image search?

- **RQ3:** What are some new ways of interacting with VCH that participants would like libraries to offer?
- **RQ4:** What are BLV participants' priorities and their contexts for using VCH?

Methods

We used a participatory design approach to engage stakeholders from the LOC, the public sector, and BLV communities in exploring more accessible futures for cultural heritage, building on prior work with people with disabilities (Guzman-Orth et al., 2023; Udoewa, 2022; Valencia et al., 2021). This study received ethics approval from the University of Maryland (IRB package 2241698-2). It was funded by a seed grant from the Maryland Initiative for Digital Accessibility in 2025. In 2025, we conducted three participatory design sessions: two online via Zoom, and one in person at the LOC—each lasting three hours and structured around three activities adapted to the session format. Each session included two BLV participants and two sighted cultural heritage experts from Title II organisations or NGOs. Between 3 and 5 research team members facilitated the sessions, serving as group leads, hosts, or '*Wizards*' for Activity 2.

Activity 1: 'building descriptions'

We split participants into small groups of one BLV person, one sighted cultural heritage expert, and one researcher. We asked BLV participants to imagine they were writing a blog post or presentation for which they wished to include visual content for a target audience of BLV and sighted people. They could ask their sighted cultural heritage expert collaborator for as much or as little information as they needed about pre-selected images we provided for the activity. BLV participants were asked to consider: '*what do I want to know about this image and how much information is useful to me?*'

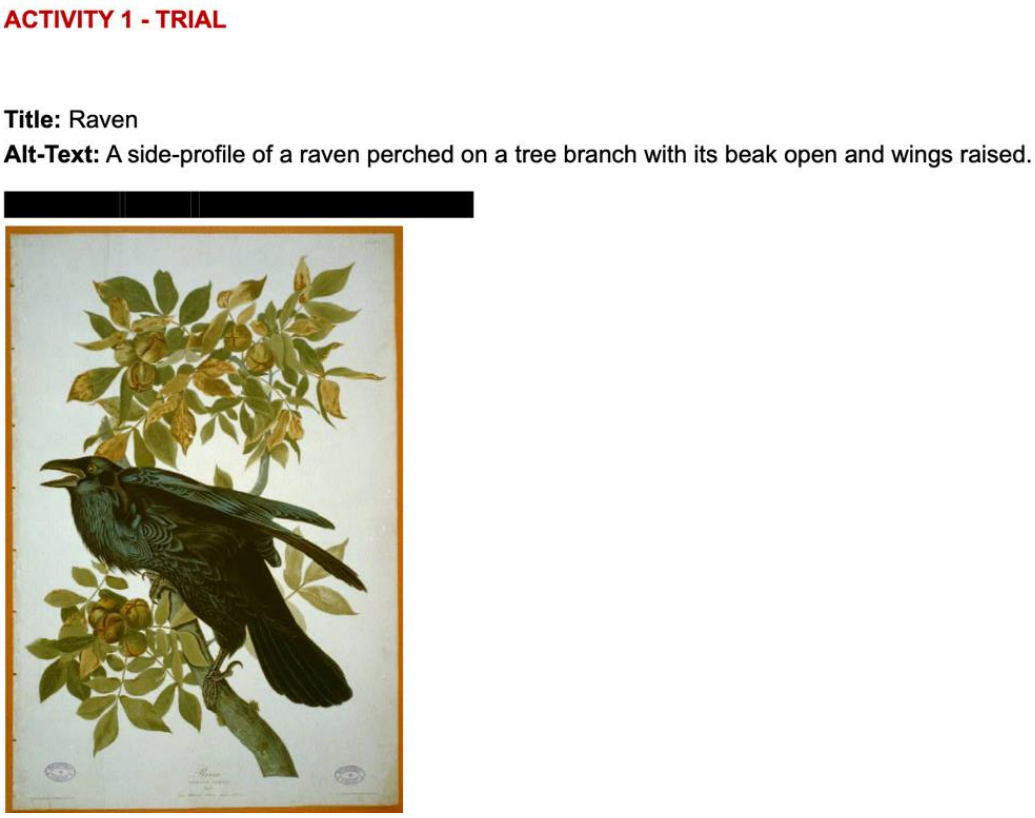

Figure 1. Activity 1 practice image for all sessions. Raven.

Each sighted participant had a Qualtrics survey with images and metadata that they read aloud, starting with the title and alt-text description that we generated for each image based on the LOC web content rules. BLV participants could ask for additional information from the official metadata or for information based on their partner's powers of description. For the online sessions BLV participants used a Qualtrics survey to record their responses to the activity, while in-person participants used a tactile voting board made with braille Legos and small objects to represent different categories of description. We provided 15 necessarily broad categories, including '*Animals*' represented with curly wool, '*Clothes*' represented by cloth and a button, '*Format*' represented with a small oil painting etc. These categories are not standard LOC metadata: they come from a previous study conducted by the LOC team with BLV people, to identify common categories of visual information of interest.

Every time a BLV participant heard a piece of information they felt was important to their understanding of the image, they clicked a star or added a Lego brick (see Figure 2) next to the category that best captured it. Participants could mark the same category multiple times per image. It was not necessary to mark each piece of information shared, only that which was interesting or useful until BLV participants reached saturation for the description. This activity gives us a proxy for the desirable thickness or robustness of descriptions.

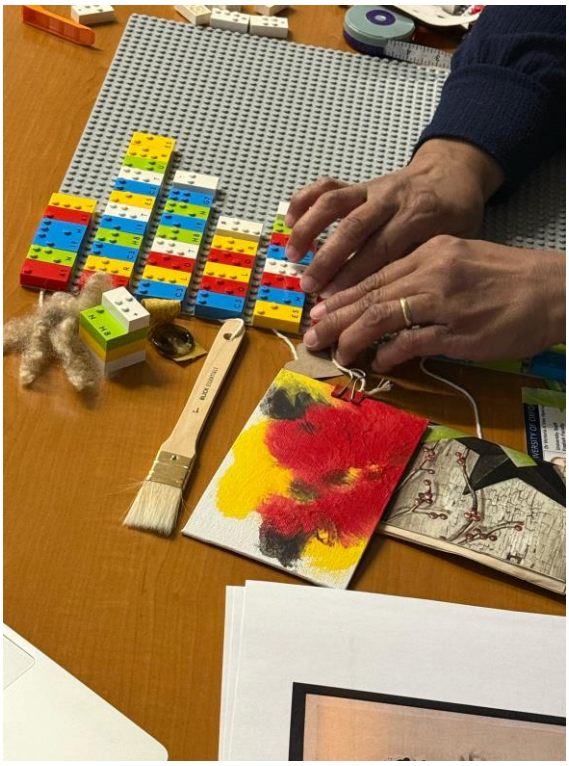

Figure 2. A blind participant runs her hands over the braille Lego categories on the tactile voting board.

Figures 3 and 4 below show the results of Activity 1.

Activity 2: the 'Wizards of Oz'

After Activity 1 and a short break, all participants reconvened in the main session (Zoom or around a table) for Activity 2. We adapted the '*Wizard of Oz*' (singular) method from HCI research (Nielson Norman Group, 2024), where designers act out system functionality, into a '*Wizards*' approach (plural) that explored how humans and technology could describe images. BLV participants took turns seeking information from several '*Wizards*', three of whom could provide '*open*' responses, meaning answers generated impromptu during the sessions, and one of whom (The Curator) could only provide '*closed*' or predetermined answers.

- Curator—played by a sighted researcher or participant, restricted to reading official metadata created by collections specialists.
- Crowdsourcerer—a sighted cultural heritage expert who described images based on their sight and their contextual or historical knowledge
- AI Wizard—a researcher who submitted participants' questions or images to ChatGPT to generate automated descriptions
- Webmaster—a cultural heritage expert whose strength was web search strategies

The Curator always began by reading the official curatorial image title, without any additional description or alt text. BLV participants then led questioning, choosing which Wizard to consult, in any order and as often as desired. Once a participant felt their questions were answered, the other could continue. When both were satisfied, the group paused for 3–5 minutes to complete a Qualtrics survey rating the usefulness of each Wizard's responses:

> *The **Curator Wizard** provided information about existing LOC metadata created by people who work at the LOC. Which of the following statements best reflects your experience?* **[usefulness score—ordinal value]**
>
> 1. *I didn't ask this Wizard any questions.* **[0]**
> 2. *I didn't understand the information this Wizard gave me.* **[1]**
> 3. *I understood the information this Wizard gave me but it did not help me understand the visual characteristics of the object.* **[2]**
> 4. *The information this Wizard gave me helped me understand the visual characteristics of the object, but it was too little information on its own to get a full picture.* **[3]**
> 5. *The information this Wizard gave me helped me understand the visual characteristics of the object, but there was too much information.* **[4]**
> 6. *The information this Wizard gave me helped me understand the visual characteristics of the object, and it was a good amount of information.* **[5]**

We created two sets of 10 images for activities 1 and 2. Additionally, we had TRIAL images prepared which gave participants a chance to understand both activities. Images from each of the sets were randomly assigned to the participants during the activities.

Activity 3: brainstorming and debrief

Activity 3 consisted of brainstorming and discussion in the same small groups as Activity 1, followed by a brief full-group discussion. We used Padlet boards for the online sessions, and post-its and sharpies for the in-person session. Sighted team members served as scribes during the in-person session. Brainstorming prompts encouraged participants to imagine future ways of accessing information about visual cultural heritage using any of the sources of information explored in the previous activities, and indeed others not yet discussed. We hoped that this activity would help all participants synthesise and process thoughts that may have arisen during earlier activities, as well as prompt reflection on their experiences of accessing VCH more broadly.

Analysis

For Activity 1 (Building Descriptions), BLV participants selected from 15 descriptive categories while collaborating with a sighted partner via Qualtrics or a tactile voting board. To analyse the voting data, we transformed a wide-format table into a long-form representation of categories

and their counts across participants and images. This allowed us to compare overall category frequencies, identify differences across images (IMG-ID), and explore variation among participants (P-ID). Visualisations such as bar charts, heatmaps, and grouped bars were used to surface patterns at multiple levels: across all responses, within specific images, and between participants.

For Activity 2 (The '*Wizards of Oz*'), we analysed the modified SUS questionnaire responses filled by blind participants after interacting with a wizard. We mapped each Wizard's rating to an ordinal usefulness scale (1 to 5) to allow descriptive analysis as follows: 1—'*didn't understand the information*', 2—'*did not help me understand the visual characteristics*', 3—'*helpful but too little information*', 4—'*too much information*', 5—'*good amount of information*'. The response of '*I didn't ask*' was excluded from our analysis.

For Activity 3, we qualitatively summarise takeaways from one study session. A detailed qualitative analysis of all sessions is planned as future work.

Results

In this section, we report the results from all three activities after 3 study sessions—with 2 blind participants per session.

For Activity 1, after excluding TRIAL images, the total votes by category were: Color = 27, People = 25, Identity = 19, Time = 19, Emotion = 17, Man-made objects = 17, Landscape = 16, Format = 12, Other = 11, Clothes = 9, Genre = 9, Actions = 8, Texture = 8, Animals = 6, Size = 4, Architecture = 2. Color and People remain most salient with Identity and Time close behind; architecture and Size were least selected. To some extent the number of votes per category reflects the dataset of images we chose, and we should be careful about drawing wider conclusions about value of any category over others, however, the amount of information about the visual qualities of each image sought by all BLV participants goes significantly beyond alt-text and much standard metadata. Our BLV participants asked for enough information to furnish a '*long description*', typically more than 125 characters—see guidance from A11y Canada (Government of Canada, 2025) conveying details such as facial expressions, activities, colors, image style and composition, etc.

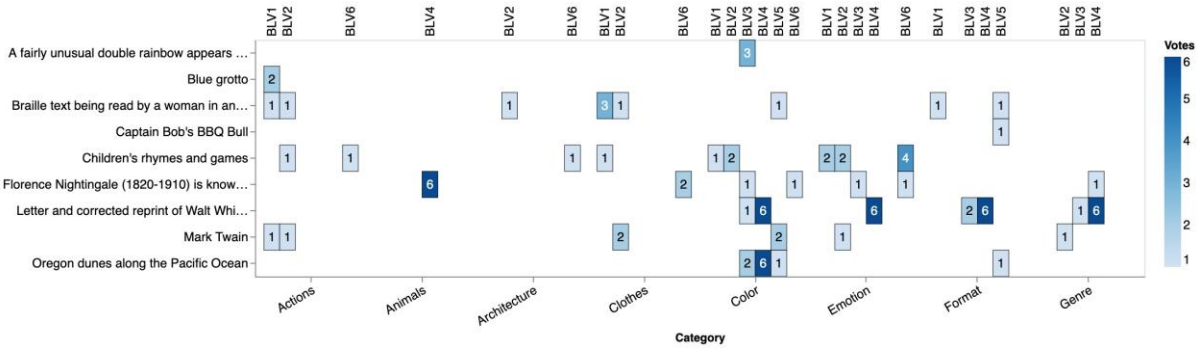

Figure 3. Total number of votes for first 8 categories across all images and all participants.

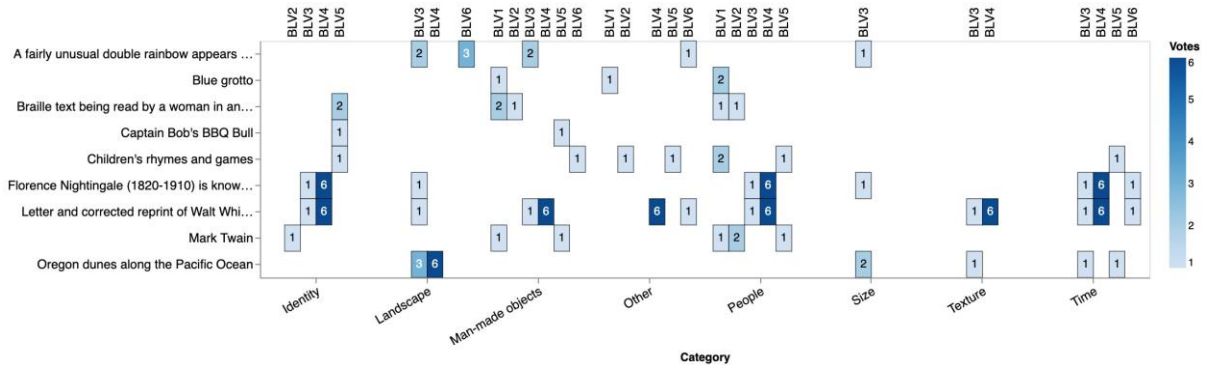

Figure 4. Total number of votes for next 8 categories (including the '*other*' category) across all images and all participants.

For Activity 2, after excluding TRIAL images and unused Wizard ratings, the Crowdsourcerer Wizard received the highest evaluations, with a mean of 4.8 (median 5.0, n = 11 uses). Figure 5 shows the ratings for the four wizards when used by a participant. Participants may have found this Wizard's flexible and human (played by a cultural heritage expert), context-rich descriptions to be '*a good amount of information*' in most sessions. In our in-person session the Crowdsourcerer was played by an LOC cultural heritage expert with over 20 years of experience creating metadata for LOC and other collections. This was a departure from the online sessions where this role was assumed by participants outside of LOC with little or no existing knowledge of the LOC collections. The LOC Crowdsourcerer shared a valuable insight at the end of the Session 3:

> Metadata is wonderful, but there's a coldness to it. I appreciate the opportunity of being the Crowdsourcer. That added warmth to the experience for me. That was a fun thing to be able to do, and it made me realise how data, how information, is communicated.

This reflection indicated that there is value to role playing and participatory design with curators and users or would-be users of their collections.

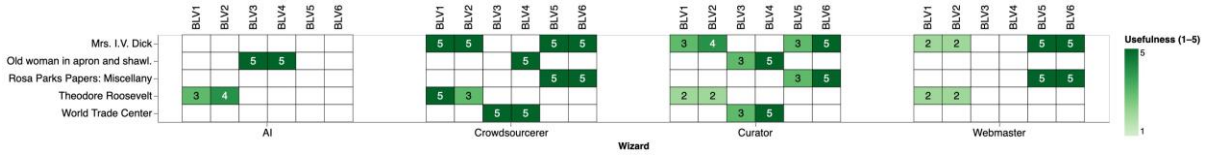

Figure 5. Usefulness ratings for the wizards across all images and participants.

The AI Wizard also scored positively when used, averaging 4.25 (median 4.5, n = 4 uses), though Activity 3 discussion reveals further nuance in sighted and BLV participants' assessments of this AI. Curator Wizard averaged 3.58 (median 3.0, n = 12 uses). This indicates participants generally found the metadata accurate but incomplete—'*helpful but too little*' information to convey the visual characteristics of an image. The Webmaster Wizard averaged 3.5 (median 3.5, n = 8). It provided a '*good amount of information*' in 4 uses; but in the other uses it was not helpful in understanding 'the visual characteristics of the object'.

Activity 3 revealed nuanced comparisons across information sources. In one session—BLV1 and BLV2 desired a system that presented multiple descriptive inputs for comparison. They noted that the AI description of the TRIAL image featuring dried plants, while detailed, '*describe[d] an already inaccessible visual construct, which is the shapes of leaves, using the shape of leaves of a different plant. If you don't know what one leaf looks like, a comparison with another won't help.*' In contrast, a crowdsourced description used basic geometric shapes and relative sizes, which participants found more effective—though cultural differences affect what counts as '*familiar*.'

Participants noted the overly '*flattering*' tone of many AI descriptions and the difficulty of forming an accurate understanding of images using AI alone. Where AI may be verbose, '*humans wear out faster than machines do*' as BLV6 put it, drawing on experiences of relying on humans for help with visual tasks, like reading a menu in a restaurant, or selecting groceries via online shopping: '*no human is going to want to sit there and read every cheese to you on the cheese shelf*'. In terms of RQ3, several participants preferred building on existing crowd-assisted services like BeMyEyes or BeSpecular (BeMyEyes, n.d.; BeSpecular, n.d.), rather than AI-only tools such as BeMyAI (BeMyAI, n.d.). One proposed solution is a dedicated librarian helpline; extending AskALibrarian services, where rotating information specialists could draw on metadata, web resources, and vetted AI descriptions.

Conclusion

When asked to imagine using visual cultural heritage (VCH) for specific purposes such as blog posts, BLV participants sought information often missing from the LOC metadata and more verbose than typical alt-text (RQ1, RQ4). Participants valued both human and machine-generated descriptions and preferred access to multiple information streams to support understanding (RQ2). Human contributors were seen as reliable but potentially unsustainable at scale, while automated tools, though efficient, pose greater risks for BLV users who cannot visually verify content. We argue that accessibility can be improved through expanded crowdsourcing, increased funding support for cultural heritage workers, and closer collaboration with BLV users. Collection-specific AI systems with human oversight, along with techniques such as handwritten text recognition (Nockels et al., 2024), may further enhance access. A limitation of this study was the small number of images assessed per group and limited time provided for reflection (especially in Activity 3). Future work should include slower-paced sessions with paper prototypes and extended Wizard-of-Oz testing involving both BLV and sighted users.

Acknowledgements

This research was supported by a seed grant from the Maryland Initiative for Digital Accessibility (MIDA) at the University of Maryland, College Park, USA (hereafter UMD). Our collaborators on this research included Dr Stephanie Valencia in the College of Information at UMD, and Rachael Bradley-Montgomery and Elizabeth Bottner from the Library of Congress. We are grateful for HR, ethics review, and other practical support from colleagues at both UMD and the Library of Congress. We are also grateful to all participants in this study.

About the Authors

Victoria Van Hyning is an Assistant Professor of Library Innovation in the College of Information at UMD. She received her PhD in English Literature from the University of Sheffield in the UK. She conducts research and contributes to policy in the US, UK, and EU about accessibility of cultural heritage for people with disabilities, with a focus on users of carceral libraries, digital archives and special collections. She can be contacted at vvh@umd.edu

Pramod Chundury is a researcher in the College of Information at UMD. He received his Ph.D from UMD in 2025, and his research interests include multimodal accessibility for data-driven systems and data visualization. He can be contacted at cvspramod@gmail.com

References

BeMyAI. (n.d.). BeMyAI [Accessed 8 January 2026]. https://www.bemyeyes.com/ bme-ai/

BeMyEyes. (n.d.). BeMyEyes [Accessed 8 January, 2026]. https://www.bemyeyes. com/

BeSpecular. (n.d.). Bespecular - help the blind. https://www.iaccessibility.com/ apps/low-vision/index.cgi/product?ID=337

Brilmyer, G. M. (2020). 'it could have been us in a different moment. it still is us in many ways': Community identification and the violence of archival representation of disability. In A. Sundqvist, G. Berget, J. Nolin, & K. I. Skjerdingstad (Eds.), Sustainable digital communities (pp. 480–486). Springer International Publishing.

Department of Justice, U. S. (2024a). American with disabilities act, title ii, requirements for web and mobile accessibility., 28 cfr 35.200 [Accessed September 15, 2025]. https://www.ecfr.gov/current/title-28/chapterI/part-35/subpart-H/section-35.200

Department of Justice, U. S. (2024b). Fact sheet: New rule on the accessibility of web content and mobile apps provided by state and local governments [Accessed September 15, 2025]. https://www.ada.gov/resources/2024-03-08-webrule/

Government of Canada, D. A. T. (2025). Alternative text and long description – best practices [Accessed 8 January, 2026]. https://a11y.canada.ca/en/alternative-text-and-long-description-bestpractices/index.html

Guzman-Orth, D., Steinberg, J., & Albee, T. (2023). English learners who are blind or visually impaired: A participatory design approach to enhancing fairness and validity for language testing accommodations. Language Testing, 40(4), 933–959. https://doi.org/10.1177/02655322231159143

Library of Congress, U. S. (2024). Photographs on flickr from the library of congress collections. Library of Congress. https://guides.loc.gov/flickr

Nielson Norman Group. (2024). The wizard of oz method in ux [Accessed September 15, 2025]. https://www.nngroup.com/articles/wizard-of-oz/

Nockels, J., Gooding, P., & Terras, M. (2024). The implications of handwritten text recognition for accessing the past at scale. Journal of Documentation, 80(7),

148–167. https://doi.org/10.1108/JD-09-2023-0183

Southwell, K. L., & Slater, J. (2012). Accessibility of digital special collections using screen readers. Library Hi Tech, 30(3), 457–471. https://doi.org/10.1108/ 07378831211266609

Udoewa, V. (2022). Radical participatory design: Awareness of participation. Journal of Awareness-Based Systems Change, 2(2), 59–84. https://doi.org/10.47061/ jasc.v2i2.3816

Valencia, S., Luria, M., Pavel, A., Bigham, J. P., & Admoni, H. (2021). Co-designing socially assistive sidekicks for motion-based aac. Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, 24–33.

https://doi.org/10.1145/3434073.3444646

Van Hyning, V., Algee, L., Jones, M., Osborn, C., Owens, T., Seroka, L., & Shelton, A. (2022). By the people crowdsourcing datasets from the library of congress. Journal of Open Humanities Data, 8, 5. https://doi.org/10.5334/johd.67

Van Hyning, V., Jones, M., & Jordan, J. B. (2024). Does crowdsourced transcription data increase accessibility for blind and low vision users? ARPHA Proceedings, 6, 201–206. https://doi.org/10.3897/ap.e127213

Xie, I., Babu, R., Lee, H. S., Wang, S., & Lee, T. H. (2021). Orientation tactics and associated factors in the digital library environment: Comparison between blind and sighted users. Journal of the Association for Information Science and Technology, 72(8), 995–1010. https://doi.org/10.1002/asi.24469

Yoon, K., Hulscher, L., & Dols, R. (2016). Accessibility and diversity in library and information science: Inclusive information architecture for library websites. The Library Quarterly, 86(2), 213–229. https://doi.org/10.1086/685399