

Brett Oppegaard, Talea Anderson, and Suzanne James-Bacon

From Sights to Sounds

A New Model for Integrating Audio Description into Library Digital Image Collections

Brett Oppegaard is a professor in the University of Hawai'i at Mānoa School of Communication and Information, email brett.oppegaard@hawaii.edu (<mailto:brett.oppegaard@hawaii.edu>), ORCID: <https://orcid.org/0000-0002-5778-1464> (<https://orcid.org/0000-0002-5778-1464>). Talea Anderson is the digital collections librarian at Washington State University Libraries, email: talea.anderson@wsu.edu (<mailto:talea.anderson@wsu.edu>), ORCID: <https://orcid.org/0000-0002-060-0426> (<https://orcid.org/0000-0002-0760-0426>). Suzanne James-Bacon is the electronic resources and MASC technical coordinator at the Washington State University Libraries, email: sbacon@wsu.edu (<mailto:sbacon@wsu.edu>), ORCID: <https://orcid.org/0009-0007-1329-5601> (<https://orcid.org/0009-0007-1329-5601>).

© 2025 Brett Oppegaard, Talea Anderson, and Suzanne James-Bacon.

Libraries and other cultural heritage institutions have grown their image collections through large-scale digitization projects, including, for example, the US Library of Congress, which just in the past five years has added 21 petabytes of data. These digital materials are a boon to historians, educators, and the general public, but they raise concerns about exacerbating an accessibility gap.

Not everyone, in other words, has been able to equivalently celebrate, access, and appreciate the increasing availability of these new resources. This is not a trivial concern for heritage institutions when considering that a sizable part of their audiences—about 50 million Americans—report difficulty using visual media.¹ Audio Description (AD) is the preferred remediation process for sight-related organizations throughout the country, such as the American Council of the Blind (ACB), because it offers more details and depth in its descriptions than typical alt text.² However, because millions of images circulate without proper AD, blind and low-vision people cannot access library collections in full.

Concerns about the accessibility of common library platforms, tools, and interfaces are not new,³ but a recent Department of Justice final ruling established that the Web Content Accessibility Guidelines (WCAG) Version 2.1, Level AA will be the technical standard for all state and local governments by 2026 or 2027, depending on the size of the jurisdiction. Combined with the well-established guidelines of Section 508 of the Rehabilitation Act of 1973, government agencies will be required to go beyond simple alt-text descriptions in the suddenly near future and to provide instead what the spirit of the law always has recommended, which was equivalent or comparable access to information.⁴

To proactively respond to this upcoming reckoning, we have documented one library's process for incorporating AD into a sample image collection. We share this process not only to encourage more institutions to add AD to their collections and to become legally compliant, but also to demonstrate how CONTENTdm and similar software could be improved with a few simple changes that create outsized benefits for community accessibility. In practice, AD does not typically get incorporated into library collections for a variety of reasons. Limited staffing, funding, and other resources, such as lack of AD training, present significant challenges for many institutions because the work of describing images requires expertise and attention, which can be labor intensive. Even a straightforward AD involves countless creative choices and interpretations by the describer. In describing people in a photo, for example, the describer has to choose how to represent social identity concepts such as age, gender, race, and ethnicity about each person, plus describe the setting, the actions, the visual context, and so on. Multiply all of those choices by each cut in a video, and the situation seems dire.

Artificial Intelligence (AI) has helped to address similar issues with captioning for people who are deaf or hard of hearing, but AI tools for automating AD haven't yet reached an adequate level of sophistication and might never be fully up to the task.⁵ Limitations in content management systems (CMSes) also can present additional challenges. In our case, CONTENTdm—a CMS that serves thousands of libraries and cultural heritage institutions—has no out-of-the-bo

options for incorporating AD. For our project, we wanted to do more than provide minimal alt-text for our images, but our CMS did little to support AD.



Figure 1: Sacking clean hops at Yakima Golding farms photograph

While libraries didn't design or create access issues online, they are now dealing with them regularly without much staffing, scholarly or technical support, or funds to address these issues. By default, libraries have become responsible for deciding who can or cannot directly use their materials based on the media's form, the patron's sensory abilities, or institutional remediation priorities. In two years, will we be talking about an enormous accessibility evolution in public resources or a shrinking landscape, where accessibility for all means access for the few or none at all, despite the legal ramifications? Cue the lawyers, who probably will need to be called in as a way to sort everything out. Read on if you would like to get ahead of such a legal morass.

A Sample Project: The Haas Collection

Like many other academic institutions, the Washington State University (WSU) Libraries provides direct or guided access to tens of thousands of photographs, maps, charts, illustrations, and other types of static visual media. As the focus for this project, we selected a set of 41 photographs showing the harvesting of hops—a plant used to make beer—during the 1940s. Pictured in the [Haas, Inc., collection \(https://cdm16866.contentdm.oclc.org/digital/collection/hops/search/searchterm/hops!image/field/all!type/mode/all!all/conn/and!and\)](https://cdm16866.contentdm.oclc.org/digital/collection/hops/search/searchterm/hops!image/field/all!type/mode/all!all/conn/and!and) were laborers who had most likely immigrated to Washington state as part of the braceros program, a federal initiative that brought Mexican farm workers to the United States to support farms during World War II. Images came from a company scrapbook that had minimal accompanying descriptions.

Object Description

Title	Laborers cutting hop vines at Yakima Golding farms
Date	1946
Creator	John L. Haas, Inc.
Synopsis	This black-and-white image was taken in central Washington state sometime between 1940 and 1945. It shows laborers cutting hop vines in a field and transferring them to a flatbed truck.
Description	A row of hop vines sweeps across the length of this image. The vines look like a leafy curtain, and they extend as high as 18 or 20 feet above the ground. A trellis has been set up alongside the hops consisting of wooden poles and wire lines. The plants derive support from the lines, which cross at around six-foot intervals. In this image, six laborers are working to cut the vines and transfer them into a truck with a wooden flatbed. At least two of the workers have long poles that they're using to reach the tops of the vines. Two are standing in the flatbed, assisting the process. All six figures either have their backs turned to the camera or they are obscured by leaves. Their features are washed out in the image, so one can only surmise that they are dressed in long-sleeved shirts, pants, and hats or head coverings.
Caption	Additional view -- cutting hop vines in the field.
About This Collection	To learn more about hops harvesting at Yakima Golding Farm, see the description of the Haas photograph collection: https://content.libraries.wsu.edu/digital/collection/hops/id/71/rec/1 .

Figure 2: Screenshot of the metadata for the sack cleaning hops image, including link to the audio description. View the image and metadata at <https://content.libraries.wsu.edu/digital/collection/hops/id/71/rec/1>.

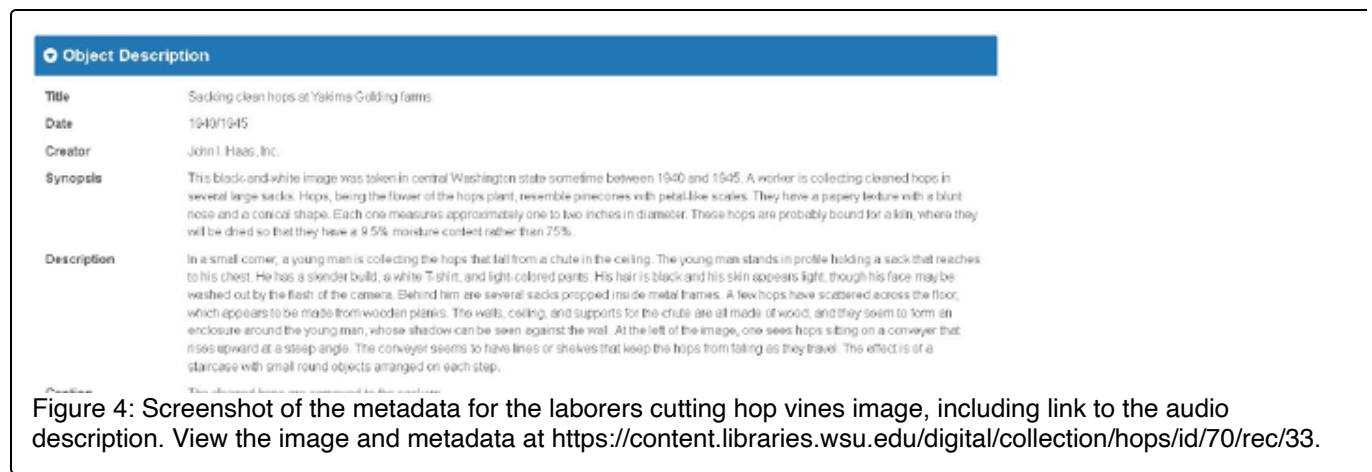
Over the course of about a month, a librarian at WSU Libraries wrote AD for these Haas images using open access software provided by [The UniDescription Project \(http://www.unidescription.org\)](http://www.unidescription.org) (UniD), based at the University of Hawai'i at Mānoa. During the past decade, the UniD research team has developed software and various other online support systems, such as best practice guidelines, that have been used by more than 200 cultural heritage organizations globally to produce AD of visitor guides and maps, mostly in service of US National Park Service sites.



Figure 3: Photograph of laborers cutting hop vines at Yakima Golding farms in 1946.

After the descriptions were written, WSU Libraries used the UniD project's built-in system to allow members of ACB to review each of the descriptions and to provide direct feedback based on guided questions. After all the reviews were gathered, representatives of UniD and WSU Libraries hosted a focus group for ACB reviewers, which revealed a showstopping problem in the delivery and contextualization of the descriptions in the CONTENTdm system. In other words, there was not a straightforward technical approach to this task, and without the involvement of representatives of the target audience, a major issue in the approach would have gone undetected. Instead, we were able to address this issue directly with our next steps.

At the beginning of this project, we had planned to directly port our newly created descriptions from UniD to the CONTENTdm CMS via a custom API. However, we soon learned that CONTENTdm does not provide both read and write API access to collections, so we shifted tactics and created a custom export option that delivered AD metadata in a CSV format. In the absence of an API, our plan was to update the Haas collection using a bulk revision process.



We encountered a second obstacle when we considered the setup of the Haas collection in CONTENTdm. Images had been imported as standalone files, but we wanted to provide both the text of each AD and an accompanying recording in which each AD had been synthetically voiced by the UniD software. As a result, we ultimately had to delete the existing collection and re-upload it with images and audio packaged together as “compound objects.” In the future, we will set up all new image collections and accompanying ADs using the compound-object formulation.

Our next consideration was the placement of the AD in the existing metadata framework. We decided to map the parts of the AD to multiple fields in CONTENTdm, including a short “Synopsis” description, a more in-depth “Description,” and a link to the MP3 file in a field called “Audio Description.” This approach was aligned in spirit with UniD’s research findings about the recommended layout of a description, beginning with an overview and followed by a richer description.

Once the descriptions were posted to CONTENTdm, we hired four reviewers who were blind to listen to five of the first 20 descriptions and answer nine questions about each. Questions included “What was the most interesting part of the description?,” “What is missing from this description?,” and “In what ways could this description be clearer?” The tenth question of our IRB-approved set was used as the opening prompt of the focus group that followed.

Although our sample size was too small to use for broad generalization, the feedback did provide a snapshot of responses by representative audience members, which we used for minor edits to the descriptions. As a usability test, though, the feedback alerted us to a key showstopping issue with the descriptions as presented in CONTENTdm, which we were able to further understand through focus group discussions and to address through alterations to the interface.

Even with the custom fields and what we thought were relatively clear labels of “Synopsis,” “Description,” and “Audio Description,” our screenreader users had difficulty identifying the context for specific images within the larger collection. As one male participant—older than 65 and adventitiously blind—said, “I had no idea what I was really looking at, and was frustrating, because they weren’t in any kind of logical order.” Based on such feedback, we added a new custom field to the collection, called “About This Collection,” which linked to contextualizing information. Responding to this adjustment, the same reviewer reported, “All of a sudden, when I looked back at it for the second time, they were in a logical order that I could understand, and each picture built on the one before it. It was much better and much easier for me to work with. I could follow what was going on. ... It was always the same kind of information. It was consistent. I knew where I was going to find everything.”

Accessibility for people who are DeafBlind, blind, or low vision thereby was shown not only to be about having AD available somewhere on the screen. AD also needs to be prioritized in the interface hierarchy, in consistent ways, to allow screenreader users to easily find it and use it. There needs to be readily available context clues that situate the media in ways that help to make sense of the material as a part of a collection, rather than in a vacuum, and if a sighted user of the material can determine how the image fits into the larger whole at a glance, a user without strong sight should also be able to use audible means to make the same grounding of the information.

After our changes were made based on feedback, including the addition of the “About This Collection” field, all our reviewers reported satisfaction with the delivery methods in CONTENTdm. For example, a woman, older than 65, and adventitiously blind, said during the focus group that interface design to her is just as important as the content because if she can’t find it, then it might as well not exist. In the first version of the CONTENTdm project, she acknowledged troubles using it and understanding the collection, but after the changes were made, she said, “I didn’t have to figure that one out. It was very easy.” Another reviewer, a 55-64-year-old male, who is adventitiously blind, added, “If it’s a chore for (blind people), then they’re just not going to put the time and effort into learning the interface.”

Conclusion

We would have liked to pursue a more elegant solution to the problem of integrating UniD with CONTENTdm through an API. Having access to a read-write API would have greatly simplified our process. However, difficulty in execution is not the excuse that history accepts. Our test users reported that the AD greatly increased their understanding of the collection and its contents. That was the ultimate goal. We hope, by sharing our example, you can do this, too, with your institution's collections and make the world a more accessible place.

Notes

1. Determining the number of people who are DeafBlind, blind, or low vision in the United States depends on how you count these people; most related organizations have different ways that sometimes create radically different numbers. We are using the number stated by the National Federation of the Blind (<https://nfb.org/resources/blindness-statistics> (<https://nfb.org/resources/blindness-statistics>)), but for comparison, the American Foundation for the Blind determined the number to be 50.18 million adult Americans in 2022 (<https://www.afb.org/research-and-initiatives/statistics> (<https://www.afb.org/research-and-initiatives/statistics>)).
2. Audio Description research in this area includes B. Oppegaard and M. Rabby, "Inclusive Measures: Establishing Audio Description Tactics That Impact Social Inclusion," in A. Lancaster and C. King (eds.), *Amplifying Voices in UX: Balancing Design and User Needs in Technical Communication* (Albany, NY: SUNY Press, 2024); T. Peters and L. Bell, "Audio Description Adds Value to Digital Images," *Computers in Libraries* 26, no. 4 (2006): 26–28; C. Lyons and T. Peters, "Audio Description Illinois," Workshop outline (Springfield, IL: Audio Description Illinois, 2008) (<https://web.archive.org/web/20160403094702/http://www.alsaudioillinois.net/workshops.cfm> (<https://web.archive.org/web/20160403094702/http://www.alsaudioillinois.net/workshops.cfm>)); and K. Lonbom, "Listening to Images: Exploring Alternate Access to a Digital Collection," in C. Cool and K. B. Ng (eds.), *Recent Developments in the Design, Construction, and Evaluation of Digital Libraries: Case Studies* (New York, NY: Information Science Reference, 2013).
3. Gayle Schechter noted in 2019 that cultural heritage collections often lack consistent metadata such that—for a busy professional—some description is better than none at all. Perhaps acknowledging the lack of time and resources in libraries/archives, the Society of American Archivists has called for the use of alt text for images but makes little reference to longform or audio description.
4. A separate column could be written just about the upcoming legal reckoning, based on the Department of Justice ruling (<https://www.ada.gov/resources/2024-03-08-web-rule/> (<https://www.ada.gov/resources/2024-03-08-web-rule/>)), which calls out as source material WCAG 2.1 (<https://www.w3.org/TR/WCAG21/> (<https://www.w3.org/TR/WCAG21/>)) and Section 508 (<https://www.access-board.gov/about/law/ra.html#section-508-federal-electronic-and-information-technology>).
5. Audio Description research in this area includes D. Bergin and B. Oppegaard, "Automating Media Accessibility: A Approach for Analyzing Audio Description Across Generative Artificial Intelligence Algorithms," *Technical Communication Quarterly* 38, no. (2024) 2, <https://doi.org/10.1080/10572252.2024.2372771> (<https://doi.org/10.1080/10572252.2024.2372771>).