

WEB CONTENT ACCESSIBILITY FOR DIGITAL AUDIOVISUAL ASSETS

Author: Walter Forsberg (Consultant), SI-Office of the Chief Information Officer, DPO

Thank You to Our Advisors and Contributors:

***Farleigh Earhart** (Associate General Counsel), SI-Office of the General Counsel*

***Lisa Elbert** (IT Specialist, Web Services), SI-Office of the Chief Information Officer*

***Megan McShea** (Audiovisual Archivist), SI-Archives of American Art*

***Isabel Meyer** (DAMS Branch Manager), SI-Office of the Chief Information Officer*

***Crystal Sanchez** (Video Digital Preservation Spc., DAMS), SI-Office of the Chief Information Officer*

***Jessica Warner** (Sr. Policy & Analysis Program Officer), SI-Office of the Chief Information Officer, DPO*

***Beth Ziebarth** (Director), SI-Accessibility Program*

***Diane Zorich** (Director), SI-Office of the Chief Information Officer, DPO*

Contents

<i>I. Executive Summary</i>	3
<i>II. Purpose</i>	5
<i>III. Background</i>	5
<i>IV. Scope</i>	5
<i>V. W3C Web Content Accessibility Guidelines 2.0</i>	5
WCAG 2.0 PRINCIPLES AND GUIDELINES	6
WCAG 2.0 LEVELS	7
WCAG 2.0 LEVEL AA CHECKLIST (WITH DEFINITIONS) FOR AUDIOVISUAL CONTENT	7
Definitions	9
<i>VI. Existing Accessibility Guidelines and Responsibilities at SI</i>	11
Smithsonian Directive 215 (Adopted May 9, 1994; Updated June 2, 2014)	11
Smithsonian Directive 950 (Adopted April 20, 2012)	12
OCIO, Technical Note: IT-950-TN06, Website Accessibility (Adopted July 7, 2016)	12
<i>VII. Legal Actions Involving Video Captioning and WCAG Compliance</i>	13
<i>VIII. SI Survey of Audiovisual Web Content Accessibility Practices & Policies – Overview</i>	15
<i>IX. Review of WCAG Compliance Plans at Other Institutions</i>	16
<i>X. Recommendations and ‘Pro Tips’</i>	19
General	19
Captioning	19
Compliance	20
Pro Tips	21
APPENDIX A	22
<i>SI Survey of Audiovisual Web Content Accessibility Practices & Policies</i>	22
APPENDIX B	37
<i>Example Style Guide: NMAAHC Audiovisual Captioning Guide</i>	37
APPENDIX C	43
<i>Sample Costs for Creating Captions</i>	43
Machine-generated vs. Human-generated	43
Labor-hours vs. Content-hours	43
Sample Third-party Vendors and Costs	44

I. Executive Summary

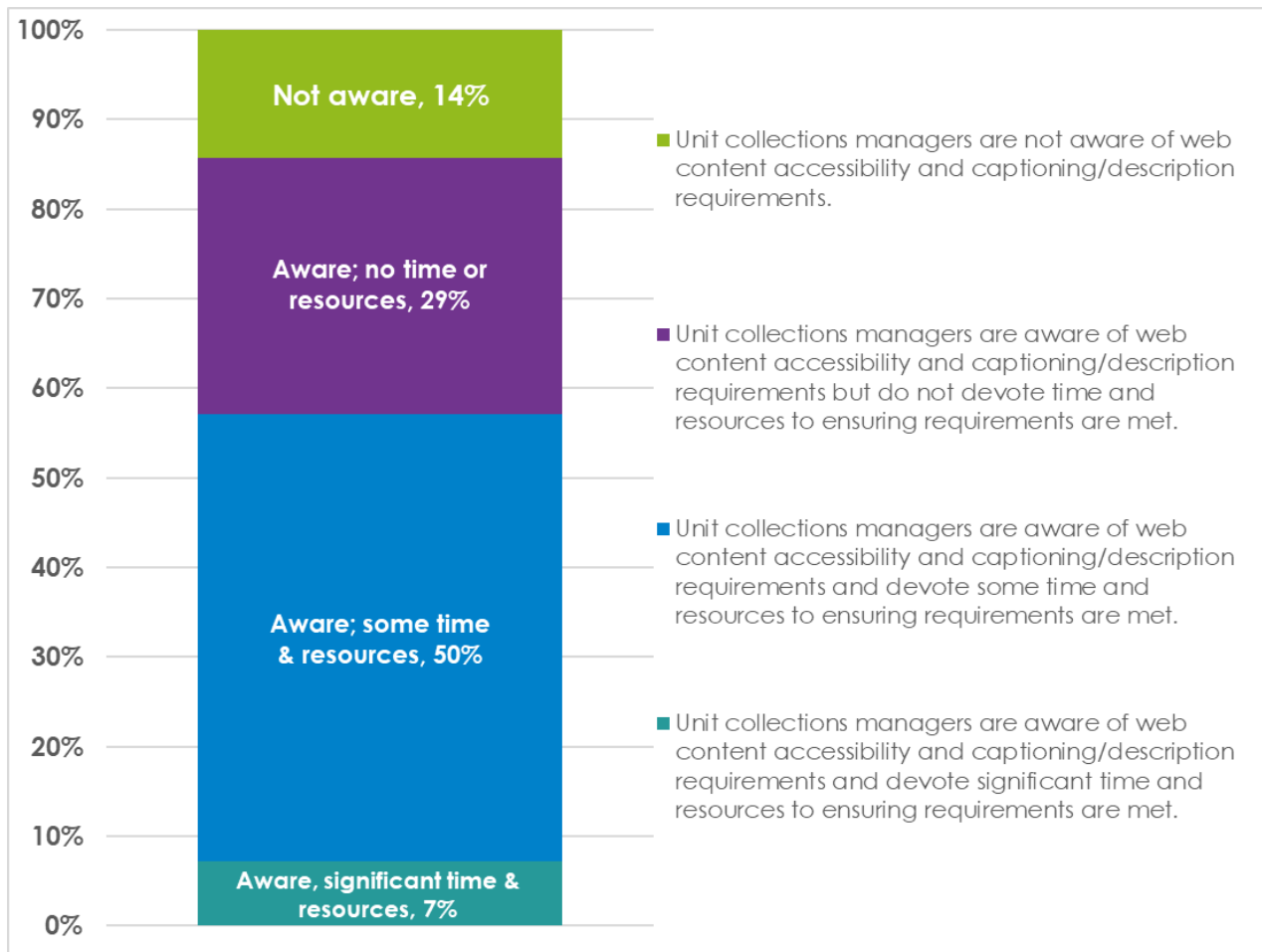
Providing access to Smithsonian Institution collections for people with disabilities, including collections that contain audiovisual materials, is a basic requirement of collections care and our institutional mission. **The Smithsonian has an ethical and legal duty to comply with these requirements reflected in the current policies:**

- [Smithsonian Directive 215, Accessibility for People with Disabilities \(SD 215\)](#)
- [Smithsonian Directive 950, Management of the Smithsonian Web \(SD 950\)](#)
- [Office of Chief Information Officer Technical Note: IT-950-TN06](#)

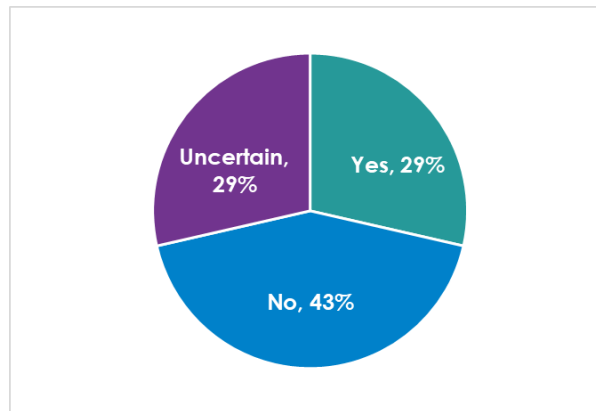
Consistent with federal laws, **Smithsonian Institution policies mandate that such access be ensured for digitized audiovisual works and born-digital audiovisual content streaming over the internet.**

A survey conducted in 2019 among the Smithsonian’s audiovisual collections managers who deal with online audiovisual collections suggests that they are **aware of accessibility requirements but lack formal guidance, training, and resources to make these collections available in accessible formats**, as highlighted in the responses to two of the survey questions below. This may explain why so **few audiovisual collections currently have existing captions, descriptions and/or transcripts.**

Which statement best describes your unit’s policy on web content accessibility and captioning/ description requirements?



Does your unit currently have any dedicated budget lines for creating transcripts/ textual descriptions/ captions for audiovisual content in its collections?



Libraries and academic institutions outside the Smithsonian are also grappling with how to best develop policies and procedures that meet accessibility requirements. Notably, several such institutions have created new departments to adequately meet this challenge.

The last ten years have witnessed significant litigation brought against companies and educational institutions alike for their failure to ensure accessibility to online audiovisual material. Such litigation has proven to be costly for many institutional defendants.

The Smithsonian is committed to ensuring accessibility to digital audiovisual content, yet the logistics and specific responsibilities for doing so are unclear at the grassroots level of collections managers. Notions that accessibility responsibilities are burdensome for collections managers are misguided, as **accessibility protocol are essential activities of collections care**. Several style guides and instructions already exist for meeting accessibility requirements, specifically captioning and audio description. It is a goal of the DPO to further support these efforts by compiling information, providing resources, and making suggestions as to how audiovisual collections managers may continue to address accessibility responsibilities.

To address these needs, the most frequently requested resources by collections managers were:

1. **funding** for outside contracts and/or internal SI or contract staff to **create transcripts/textual descriptions/captions**, and
2. **regulations, training, and guidelines from the Smithsonian on how to create** transcripts/textual descriptions/captions

What Collections Managers Are Saying

“Though we know the requirements for captions and audio descriptions, it is uncertain who should be actually making that happen. As far as I can tell, only one person is doing it, but he doesn't have a great deal of support.”

“Accessibility requirements (captions/textual descriptions) have been implemented for our still image and print materials, and it is something very much on our minds for audiovisual materials, but it has not yet been implemented on our website.”

II. Purpose

This document provides research into current practices surrounding digital audiovisual content, articulation of accessibility standards, as well as guidance regarding web content accessibility requirements for museum, library, and archives unit staff and audiovisual collections managers at the Institution.

III. Background

The documentation, research, analysis, and guidance in this report was developed by the Smithsonian Institution's Digitization Program Office (DPO) in the summer of 2019. The DPO is charged with providing policies that govern digitization activities and digital access and use at the Institution in addition to increasing the quality, quantity, and impact of digitization across the Institution.

IV. Scope

This document applies to all audiovisual assets made accessible on the internet, whether via a Smithsonian Institution web page platform and/or a third-party platform (such as YouTube or Vimeo). This includes born-digital audio and video recordings and audio and video recordings digitized from analog sources. As the 2016-2017 *Pan-Institutional Audiovisual Collections Survey* details, the scale of audiovisual works held across the Institution's collections is massive: some 293,586 analog assets (a number that continues to grow annually) and an uncounted number of digital assets potentially approaching a similar quantity.¹

V. W3C Web Content Accessibility Guidelines 2.0

What does it mean to ensure accessibility to audiovisual content, and how does one achieve fully compliant accessibility?

The World Wide Web Consortium (W3C) provides ample documentation, instruction, and recommendations regarding web content accessibility guidelines (WCAG) via its website www.w3c.org. The WCAG 2.0 standard is the current required level of web content accessibility at the Smithsonian as mandated by *Technical Note: IT-950-TN06*. While a revised level of web content accessibility was published by the W3C on June 5, 2018, entitled WCAG 2.1 (see: <https://www.w3.org/TR/WCAG21/>), the present report focuses on WCAG 2.0 compliance. The first section below presents definitions for WCAG principles and guidelines, as specified by its WCAG 2.0 recommendation published on December 11, 2008.

¹ Smithsonian Institution, *Pan-Institutional Audiovisual Collections Survey: Final Project Report 2016-2017*, (Washington, DC: Smithsonian Institution Archives, 2017), 24, https://siarchives.si.edu/sites/default/files/pdfs/SI_AVSurvey_FinalReport_03282017.pdf. Accessed January 7, 2020.

WCAG 2.0 PRINCIPLES AND GUIDELINES

As defined by the W3C in WCAG 2.0, web accessibility to content consists of four principles:

- 1. Perceivable:** Information and user interface components must be presentable to users in ways they can perceive. This means that the information being presented cannot be invisible to all of the user's senses.
- 2. Operable:** User interface components and navigation must be operable. This means the interface cannot require interaction that a user cannot perform.
- 3. Understandable:** Information and the operation of user interface must be understandable. This means that the content or operation cannot be beyond their understanding.
- 4. Robust:** Content must be robust enough that it can be interpreted reliably by a wide variety of user agents, including assistive technologies. This means that as technologies and user agents evolve, the content should remain accessible.²

To ensure these four principles are met, twelve guidelines organize the WCAG principles and provide guidance on how to achieve them. These guidelines are:

1. Perceivable

- 1.1 Text Alternatives:** Provide text alternatives for any non-text content so that it can be changed into other forms people need, such as large print, Braille, speech, symbols or simpler language.
- 1.2 Time-based Media:** Provide alternatives for time-based media.
- 1.3 Adaptable:** Create content that can be presented in different ways (for example, simpler layout) without losing information or structure.
- 1.4 Distinguishable:** Make it easier for users to see and hear content including separating foreground from background.

2. Operable

- 2.1 Keyboard Accessible:** Make all functionality available from a keyboard.
- 2.2 Enough Time:** Provide users enough time to read and use content.
- 2.3 Seizures and Physical Reactions:** Do not design content in a way that is known to cause seizures.
- 2.4 Navigable:** Provide ways to help users navigate, find content, and determine where they are.

3. Understandable

- 3.1 Readable:** Make text content readable and understandable.
- 3.2 Predictable:** Make web pages appear and operate in predictable ways.
- 3.3 Input Assistance:** Help users avoid and correct mistakes.

² World Wide Web Consortium, *Introduction to Understanding WCAG 2.1*, <https://www.w3.org/WAI/WCAG21/Understanding/intro#understanding-the-four-principles-of-accessibility/>. Accessed January 7, 2020.

4. Robust

4.1 **Compatible:** Maximize compatibility with current and future agents, including assistive technologies.

WCAG 2.0 LEVELS

The WCAG guidelines are arranged into three levels, with corollary evaluation metrics called “Success Criteria,” indicating a scale of conformance among “different groups and different situations”³:

- Level A (lowest level of conformance)
- Level AA
- Level AAA (highest level of conformance)

The *Office of Chief Information Officer Technical Note: IT-950-TN06* requires that:

“Smithsonian Institution websites launched after 07/15/2016, including website refresh projects, shall conform to the W3C WCAG 2.0 Level AA guidelines.

“All Smithsonian websites should strive for Level AA conformance.”⁴

WCAG 2.0 LEVEL AA CHECKLIST (WITH DEFINITIONS) FOR AUDIOVISUAL CONTENT

The W3C has an excellent *Quick Reference* website that walks users through its WCAG 2.0 standard and its Success Criteria: <https://www.w3.org/WAI/WCAG21/quickref/>. This site enables users to toggle between different levels of Success Criteria compliance, WCAG versions, and other filters. This tool includes useful techniques and failures metrics for meeting each guideline. The tool is about as close as one can get to a checklist, including detailed hyperlinked definitions and suggestions for accessibility solutions.

Taking from this tool, what follows is a checklist (with definitions) for conformance with WCAG 2.0, Level AA for audio and video content. (N.B. Level AA compliance encompasses the compliance requirements of level A.) This checklist selects those guidelines specifically applicable to audio and video content using the *Quick Reference*’s filter function. (N.B. Smithsonian audiovisual content managers will need to pay particular attention to the guidelines relating to: Principle 1 - Perceivable.)

1.1.1 Non-text Content (Level A)

All non-text content that is presented to the user has a text alternative that serves the equivalent purpose, except for the situations listed below. If non-text content is **time-based media**, then text alternatives at least provide descriptive identification of the non-text content. (Refer to Quick Reference Guideline 1.2 for additional requirements for media.)

³ World Wide Web Consortium, *Web Content Accessibility Guidelines (WCAG) 2.0*, December 11, 2008, <https://www.w3.org/TR/WCAG20/>. Accessed January 7, 2020.

⁴ Smithsonian Institution, Office of the Chief Information Officer, *Technical Note: IT-950-TN06, Website Accessibility*, (Washington, DC: Smithsonian Institution, July 7, 2016), 3, <https://sinet.sharepoint.com/sites/PRISM2/OCIO/ITPolicies/IT-950-TN06.pdf>. Accessed January 7, 2020.

1.2.1 **Audio-only and Video-only (Prerecorded)** (Level A)

For prerecorded audio-only and prerecorded video-only media, the following are true, except when the audio or video is a media alternative for text and is clearly labeled as such:

- Prerecorded Audio-only: An alternative for time-based media is provided that presents equivalent information for prerecorded audio-only content.
- Prerecorded Video-only: Either an alternative for time-based media or an audio track is provided that presents equivalent information for prerecorded video-only content.

1.2.2 **Captions (Prerecorded)** (Level A)

Captions are provided for all prerecorded audio content in synchronized media, except when the media is a media alternative for text and is clearly labeled as such.

1.2.3 **Audio Description or Media Alternative (Prerecorded)** (Level A)

An alternative for time-based media or audio description of the prerecorded video content is provided for synchronized media, except when the media is a media alternative for text and is clearly labeled as such.

1.2.4 **Captions (Live)** (Level AA)

Captions are provided for all live audio content in synchronized media.

1.2.5 **Audio Description (Prerecorded)** (Level AA)

Audio description is provided for all prerecorded video content in synchronized media.

1.3.3 **Sensory Characteristics** (Level A)

Instructions provided for understanding and operating content do not rely solely on sensory characteristics of components such as shape, size, visual location, orientation, or sound.

1.4.2 **Audio Control** (Level A)

If any audio on a Web page plays automatically for more than 3 seconds, either a mechanism is available to pause or stop the audio, or a mechanism is available to control audio volume independently from the overall system volume level.

2.1.1 **Keyboard** (Level A)

All functionality of the content is operable through a keyboard interface without requiring specific timings for individual keystrokes, except where the underlying function requires input that depends on the path of the user's movement and not just the endpoints. (Note: This does not forbid and should not discourage providing mouse input or other input methods in addition to keyboard operation.)

2.1.2 **No Keyboard Trap** (Level A)

If keyboard focus can be moved to a component of the page using a keyboard interface, then focus can be moved away from that component using only a keyboard interface, and, if it requires more than unmodified arrow or tab keys or other standard exit methods, the user is advised of the method for moving focus away.

2.2.2 Pause, Stop, Hide (Level A)

For moving, blinking, scrolling, or auto-updating information, all of the following are true:

- Moving, blinking, scrolling: For any moving, blinking or scrolling information that (1) starts automatically, (2) lasts more than five seconds, and (3) is presented in parallel with other content, there is a mechanism for the user to pause, stop, or hide it unless the movement, blinking, or scrolling is part of an activity where it is essential; and
- Auto-updating: For any auto-updating information that (1) starts automatically and (2) is presented in parallel with other content, there is a mechanism for the user to pause, stop, or hide it or to control the frequency of the update unless the auto-updating is part of an activity where it is essential.

2.3.1 Three Flashes or Below Threshold (Level A)

Web pages do not contain anything that flashes more than three times in any one second period, or the flash is below the general flash and red flash thresholds.

4.1.2 Name, Role, Value (Level A)

For all user interface components (including but not limited to: form elements, links and components generated by scripts), the name and role can be programmatically determined; states, properties, and values that can be set by the user can be programmatically set; and notification of changes to these items is available to user agents, including assistive technologies.

Definitions

This subsection outlines some of the most common terms regarding web content accessibility for audiovisual works that are used in this report. When appropriate, definitions are quoted directly from the World Wide Web Consortium's (W3C) Web Content Accessibility Guidelines 2.0 *Glossary*, available at: <https://www.w3.org/TR/2006/WD-WCAG20-20060427/appendixA.html>.

Audio Description: "narration added to the soundtrack to describe important visual details that cannot be understood from the main soundtrack alone. Audio descriptions of video provide information about actions, characters, scene changes, and on-screen text. In standard audio description, narration is added during existing pauses in dialogue."⁵ **Audio Description** tracks can take the form of text tracks, separate from a work's original soundtrack and verbalized by the video player, or as a separate audio file containing the verbalized description. Ideally, the **Audio Description** soundtrack can be turned on (or off) as an optional supplemental soundtrack that plays during the interstitial silences of the video's original soundtrack. Often, such functionality will depend on the video player platform, and its designed capabilities. In cases where the amount of description information is too large to convey during these silences, **Extended Audio Description** (WCAG 2.0, Level AAA) may be required wherein the original video is slowed or paused while the **Audio Description** track plays. **Audio Description** is a requirement to fulfill WCAG 2.0, Level AA accessibility success criterion 1.2.5.

⁵ World Wide Web Consortium, *Glossary to Web Content Accessibility Guidelines 2.0, Appendix A Glossary (Normative)*, <https://www.w3.org/TR/2006/WD-WCAG20-20060427/appendixA.html> Accessed January 7, 2020.

Captions: “text presented and synchronized with multimedia to provide not only the speech, but also sound effects and sometimes speaker identification. In some countries, the term **Subtitle** is used to refer to dialogue only and **Captions** is used as the term for dialogue plus sounds and speaker identification.”⁶

Captioning: “the process of converting the audio content of a television broadcast, webcast, film, video, CD-ROM, DVD, live event, or other production into text and displaying the text on a screen or monitor. **Captions** not only display words as the textual equivalent of spoken dialogue or narration, but they also include speaker identification, sound effects, and music description. **Captioning** is critical for students who are deaf or hard of hearing, but it also aids the reading and literacy skills development of many others.”⁷

Closed Captioning: “displays the audio portion of a television program as text on the TV screen, providing a critical link to news, entertainment, and information for individuals who are deaf or hard-of-hearing.”⁸ In the era of analog video production, **closed captioning** was encoded into the video signal on line 21 of the underscan area--typically not visible on consumer cathode ray tube monitors. On-board decoders in consumer cathode ray tube monitors would subsequently decode the closed captions embedded in the signal and display them over the image as white text on rectangular black backgrounds. The open-source “sccyou” tool can be used to extract and convert these embedded closed captions from the video signal, rendering them as a .SCC format file (which can be converted to a .VTT file).⁹

Digital assets: “text, still images, moving images, and sound recordings, research datasets and other types of media originally created in digital format (i.e., born digital) or digitized from another format or state (i.e., a digital surrogate) that are created, stored, or maintained by the Smithsonian.”¹⁰

Hard Captions: captions which are ‘burned in’ to the video image itself, unable to be turned ‘on’ (viewable) or ‘off’ (not viewable). Typically, these **Hard Captions** are created as overlaying text in a video editing program, with the resulting text and video ‘flattened’ into a version that is exported as a video file. Sometimes these types of ‘text-over-image’ are called **Subtitles**; the term **Subtitle**, however, is a term usually reserved for ‘text-over-image’ that displays translation for a different spoken language in the film’s soundtrack or onscreen text, and/or infers a capacity to be turned off. Similarly, **Caption** infers the capacity to be turned off, sometimes garnering them the antonymic term **Soft Captions**.

Subtitles (aka “sub-titles”): See entries for **Captions** and **Hard Captions**.

⁶ *Ibid.*

⁷ Described and Captioned Media Program, *Captioning Key*.
http://www.captioningkey.org/quality_captioning.html. Accessed January 7, 2020.

⁸ Federal Communications Commission, *Consumer Guide: Closed Captioning on Television*,
https://www.fcc.gov/sites/default/files/closed_captioning_on_television.pdf. Accessed January 7, 2020.

⁹ For more information about the “sccyou” closed captions extraction tool, visit the Association of Moving Image Archivists’ GitHub at: <https://github.com/amiaopensource/sccyou>. Accessed January 7, 2020.

¹⁰ Smithsonian Institution, *Smithsonian Directive 609: Digital Asset Access and Use*, (Washington, DC: Smithsonian Institution, July 15, 2011), <https://www.si.edu/content/pdf/about/sd/SD609.pdf>. Accessed January 7, 2020.

Transcripts: text versions of spoken dialogue; popular in oral history disciplines. While they may contain some format of time stamp that ties them to specific moments in a soundtrack, traditional **Transcripts** do not use the same timed-text annotations as **Captions** do. In a basic sense, any time stamps that **Transcripts** may contain are less frequent or tied to specific phrases than those in a **Captions** track or file. The term “timed-text track” is sometimes used as a synonym for **Captions** in web implementations, as is the term “Web Video Text Track” describing the .VTT **Captions** format.

VI. Existing Accessibility Guidelines and Responsibilities at SI

In the course of this research, conversations revealed that some staff within the Institution perceive accessibility protocol as “unfunded mandates,” which place an “undue burden” on collections managers and exhibition staff. Smithsonian policy and extant federal legislation indicate that such views are incorrect. As stated, accessibility protocol are essential component activities of collections care. Nevertheless, these comments can be understood as reflecting the significant resource challenges in providing the guidance and funding necessary to ensure accessibility—challenges which appear unresolved at the level of collections care and management.

This section articulates Smithsonian standards for making its collections accessible to people with disabilities. Commonly cited federal accessibility legislation regarding audiovisual content (such as the Americans with Disabilities Act (ADA), as amended P.L. 110-325, and the Rehabilitation Act, P.L. 93-112) is already covered by several existing Smithsonian policies, including the following:

Smithsonian Directive 215 (Adopted May 9, 1994; Updated June 2, 2014)

Smithsonian Directive 215, Accessibility for People with Disabilities (SD 215) was officially adopted in early 1994, given that:

“a mandate for ‘the increase and diffusion of knowledge’ means little without accessibility to the Smithsonian’s resources. From long-range objectives in charting research to the smallest details of designing exhibits, accommodating staff and visitors with disabilities is a primary goal and responsibility.”¹¹

Broadly, *SD 215* policy covers “all programs held at or by the Smithsonian...”. While *SD 215* does not specifically articulate guidelines for access to Smithsonian audiovisual digital assets online, it does articulate a required Institutional adherence to the Americans With Disabilities Act of 1990 as administered by the United States Department of Justice.

Significantly, *SD 215* states, “All programs, regardless of facility accessibility, will provide effective communication to people with disabilities through their design, supplemental materials, or auxiliary services, such as sign language interpreters, captioning, verbal descriptions or assistive listening devices.”¹²

Responsibility for ensuring accessibility lies with museum and research unit directors. According to *SD 215*, “Unit directors are responsible for ensuring programmatic and facility access to their staff and

¹¹ Smithsonian Institution, *Smithsonian Directive 215, Accessibility for People With Disabilities*, (Washington, DC: Smithsonian Institution, June 2, 2014), <https://airandspace.si.edu/rfp/exhibitions/files/j3-directive-215.pdf>. Accessed January 7, 2020.

¹² *SD 215*, 2.

visitors to the fullest extent possible and practicable, by providing and maintaining accessible facilities, exhibits, services and programs...”¹³

The directive also specifies that “Smithsonian staff” is responsible for “planning, budgeting for, and designing exhibits, programs, and facilities following the guidance in the Smithsonian Guidelines for Accessible Design published by the Accessibility Program.”¹⁴

Smithsonian Directive 950 (Adopted April 20, 2012)

Smithsonian Directive 950 (*SD 950*) outlines policy regarding management of Smithsonian websites and web applications.

Broadly, *SD 950* states “efforts should be made to make websites accessible to all visitors, including those with disabilities...” and “new websites and web applications shall be accessible by visitors with disabilities in accordance with *SD 215, Accessibility for People with Disabilities Policy*. During each major redesign of existing webpages and applications, the appropriate design changes shall be incorporated to make them accessible by visitors with disabilities.”¹⁵

OCIO, Technical Note: IT-950-TN06, Website Accessibility (Adopted July 7, 2016)

Stemming from *SD 950*, the *Office of Chief Information Officer Technical Note: IT-950-TN06* provides additional clarity and detail. It states:

“Smithsonian Institution websites launched after 07/15/2016, including website refresh projects, shall conform to the W3C WCAG 2.0 Level AA guidelines.

“All Smithsonian websites should strive for Level AA conformance.”¹⁶

The technical note “applies to all units, employees, contractors, consultants, and volunteers who own or manage Smithsonian websites and/or web applications, both publically [sic] -accessible and internal-only, whether hosted by the Institution within its data center or externally on behalf of the Institution. All personnel must follow its contents and the procedures specified herein to maintain adherence to *SD-950 “Management of the Smithsonian Web and SD-215 Accessibility for People with Disabilities”*.”¹⁷

Technical Note: IT-950-TN06 articulates a broad and detailed range of parties responsible for ensuring WCAG 2.0 Level AA compliance. Specifically, *Technical Note: IT-950-TN06* mandates, “Units must incorporate accessibility validation into their on-going website maintenance processes,” and at the points of creation and regular maintenance, website owners, sponsors, managers, and developers must “ensure their websites are accessible to people with disabilities.”¹⁸

¹³ *SD 215*, 3.

¹⁴ *SD 215*, 3.

¹⁵ Smithsonian Institution, *Smithsonian Directive 950, Management of the Smithsonian Web*, (Washington, DC: Smithsonian Institution, April 20, 2012), 21, <http://prism2.si.edu/SIorganization/OCFO/OPMB/SD/SD950.pdf>. Accessed January 7, 2020.

¹⁶ *Technical Note: IT-950-TN06*, 3.

¹⁷ *Ibid.*, 2.

¹⁸ *Ibid.*, 3, 1.

VII. Legal Actions Involving Video Captioning and WCAG Compliance

There is a score of ongoing, lower-court litigation surrounding captioning and audio description requirements for internet-posted video content. As *Forbes* magazine describes it, “ADA Title III litigation has become a cottage industry, with federal claims increasing from 57 in 2015 to 814 in 2017.”¹⁹ Court rulings and corollary obligations for video content providers often conflict. In the absence of ADA regulations specific to websites, determinations on the matter will rest with future higher-court decisions. Nevertheless, several private companies and educational institutions have been the subject of lengthy and expensive litigation for failing to caption and/or provide audio description for video content. A common premise for plaintiff litigation is to allege violation of the Americans With Disabilities Act (ADA), Title III.

Below are brief summaries of several relevant legal actions in the United States.

- *NAD, et al. v. Netflix* (2010-2012)
 - In 2010, the National Association of the Deaf (NAD), Western Massachusetts Association of the Deaf and Hearing Impaired, and Lee Nettles filed suit in the United States District Court for the District of Massachusetts Western Division against online video service provider Netflix, alleging violation of Title III of the ADA for failing to provide textual captions for online video content. A consent decree reached in October 2012 saw Netflix agree to commit to create and provide captioning for all its online streaming video content by 2014, and to reimburse NAD’s legal fees of \$755,000.00, along with other practical compliance and financial concessions.²⁰
 - In light of the *NAD, et al. v. Netflix* consent decree, Netflix subsequently entered into a notable additional settlement with the American Council of the Blind in 2016, agreeing to provide Audio Description (as required by WCAG 2.0, Level AA, Success Criterion 1.2.5) for a range of its online content.²¹
- *Cullen v. Netflix* (2011-2015)
 - In 2011, private citizen Donald Cullen filed a class action lawsuit against Netflix for failure to caption online video content. In a decision running counter to the Massachusetts court decision *NAD, et al. v. Netflix*, the U.S Court of Appeals for the Ninth Circuit in San Francisco ruled that the ADA did not apply to Netflix because “Netflix’s services are not connected to any ‘actual, physical place[]’”.²² Critically, the

¹⁹ Glenn G. Lamm, “Ninth Circuit Decision Underscores Need For Clarity On ADA’s Application In Cyberspace,” in *Forbes.com*, (January 31, 2019), <https://www.forbes.com/sites/wlf/2019/01/31/ninth-circuit-decision-underscores-need-for-clarity-on-adas-application-in-cyberspace/#78c463ab51dd>. Accessed January 7, 2020.

²⁰ National Association of the Deaf, et al. v. Netflix, Civil Action No. 11-30168-MAP, (October 9, 2012), <https://dredf.org/captioning/netflix-consent-decree-10-10-12.pdf>. Accessed January 7, 2020.

²¹ Todd Spangler, “Netflix to Expand Audio Descriptions for Blind Subscribers,” *Variety*, (April 14, 2016), <https://variety.com/2016/digital/news/netflix-audio-descriptions-blind-settlement-1201753569/>. Accessed January 7, 2020.

²² *Cullen v. Netflix*, No. 13-15092, D.C. No. 5:11-cv-01199-EJD, Memorandum, (2015), 2, <https://d3bsvxk93brmko.cloudfront.net/datastore/memoranda/2015/04/01/13-15092.pdf>. Accessed January 7, 2020.

court deemed its ruling as “not for publication,” suggesting it not be held up as quotable precedent in other cases outside its judicial purview (which includes much of the Western United States).²³

- *NAD, et al. v. Harvard University* (2015-present)
 - In 2015, the NAD alleged that Harvard University violated Title III of the ADA for failing to provide textual captions for online video content. Litigation is ongoing. Core issues of this case include: whether Harvard, as an educational institution, is liable for its own content posted on third-party websites (such as YouTube, and SoundCloud); what constitutes an “undue burden” vis-à-vis captioning requirements; and the relevance of the 1996 Communications Decency Act—a statute initially intended to regulate internet pornography, which holds that internet service providers do not qualify as “publishers” of content.²⁴ A Consent Decree reached in November 2019 stipulates, among several agreed actions, that Harvard create, support, and monitor a “cure process” by which the public can request creation of captions for online audiovisual materials that have none, as well as request corrections to captions that contain “material errors.”²⁵ In addition to consenting to henceforth provide captions for its online audiovisual content, Harvard agreed to cover \$1,575,000 in legal fees for the NAD.
- *U.S. Department of Justice’s Letter of Findings to University of California at Berkeley* (2016)
 - In a ‘Letter of Findings’ to UC Berkeley in 2016 the DOJ summarized ongoing investigation claiming that the educational institution’s thousands of freely-available online video courses failed to provide captioning when posted online via third-party platforms YouTube and iTunes U.²⁶ As a result, UC Berkeley decided to take down roughly 20,000 videos, no longer making them publicly available.²⁷ The Letter of Findings specifically highlights UC Berkeley’s failure to adhere to WCAG 2.0, Level AA success criteria.²⁸

²³ Joe Mullin, “9th Circuit rules Netflix isn’t subject to disability law,” *arstechnica.com*, (April 2, 2015), <https://arstechnica.com/tech-policy/2015/04/9th-circuit-rules-netflix-isnt-subject-to-disability-law/>. Accessed January 7, 2020.

²⁴ Seyfarth Shaw, LLC, “Four-Year Court Battle Between Deaf Advocates and Harvard Over Closed Captioning of Videos Proceeds to Discovery With Some Limitations,” (April 5, 2019), <https://www.adatitleiii.com/2019/04/four-year-court-battle-between-deaf-advocates-and-harvard-over-closed-captioning-of-videos-proceeds-to-discovery-with-some-limitations/>. Accessed January 7, 2020.

²⁵ Consent Decree, *NAD v. Harvard University* (No. 3:15-cv-30023-KAR, District Court of Massachusetts, November 8, 2019), <https://harvardcaptioningsettlement.files.wordpress.com/2019/12/nad-harvard-consent-decree.pdf>. Accessed January 7, 2020.

²⁶ U.S. Department of Justice, *The United States’ Findings and Conclusions Based on its Investigation Under Title II of the Americans with Disabilities Act of the University of California at Berkeley*, DJ No. 204-11-309, (August 30, 2016), https://www.ada.gov/briefs/uc_berkeley_lof.pdf. Accessed January 7, 2020.

²⁷ Douglas Ernst, “Berkeley removing 20K free videos after DOJ ruling, closed-captioning complaint,” *Washington Times*, (March 7, 2017), <https://www.washingtontimes.com/news/2017/mar/7/berkeley-removing-20k-free-educational-videos-afte/>. Accessed January 7, 2020.

²⁸ U.S. Department of Justice [Ibid.], 6-7.

VIII. SI Survey of Audiovisual Web Content Accessibility Practices & Policies – Overview

As part of the research for this document, in April 2019 the DPO undertook a survey of current practices among a group of audiovisual collections stakeholders across the Institution. Questions were created by SI-DPO consultant contractor Walter Forsberg and SI-DPO Senior Policy & Analysis Program Officer Jessica Warner, in consultation with SI-OCIO DAMS Video and Digital Preservation Specialist Crystal Sanchez and SI Accessibility Program Director Beth Ziebarth.

Sixteen (16) core audiovisual stakeholders were asked to complete the survey, fourteen (14) of which responded (an 87.5% response rate). For a full report of the survey results including charts, see APPENDIX A.

Core audiovisual stakeholders from the thirteen (13) Institution units listed below responded to the initial survey. (Different staff from one unit—the National Museum of the American Indian—completed the survey twice.)

- Archives of American Art (AAA)
- Center for Folklife and Cultural Heritage (CFCH)
- Freer-Sackler Galleries (FSG)
- National Air and Space Museum (NASM)
- National Museum of African American History and Culture (NMAAHC)
- National Museum of American History (NMAH)
- National Museum of the American Indian (NMAI)
- National Museum of Natural History (NMNH)
- Office of the Chief Information Officer (OCIO)
- Smithsonian American Art Museum (SAAM)
- Smithsonian Institution Archives (SIA)
- Smithsonian Institution Libraries (SIL)
- Smithsonian Tropical Research Institute (STRI)

Over two-thirds of respondents make some Smithsonian audiovisual digital assets available on the internet [Question 2], yet over 90% of respondents do not have any articulated or specific guidelines as to how to create transcripts, textual descriptions, or captions for said content [Question 16]. Unfamiliarity with captioning protocol was echoed by the fact that only two respondents attested to having ever received any training or specific how-to guidance on creating transcripts, textual descriptions, or captions [Question 8]. Yet, respondents appear overwhelmingly aware that their online collections are subject to accessibility requirements [Questions 19, 22, and 23]. **The need for guidance as to how to be compliant seems to be the overwhelming core take-away from the survey**, and respondents suggested that they would traditionally expect such guidance from either the SI AV Archivists Interest Group and/or OCIO [Question 24].

While they may not necessarily have guidance on making captions for Smithsonian collection, stakeholders appear to be quite familiar with the phenomenon of captions for audiovisual content. All survey respondents reported having viewed videos online via streaming players capable of displaying captions, and all but two respondents reported having viewed broadcast television with traditional broadcast television Closed Captions [Questions 3 and 4]. Experience with screen-reading software—frequently used with text-based, non-audiovisual materials—was notably less prevalent among

respondents, with just over one-quarter of respondents reporting experience using such software [Question 5].

Exactly what would constitute an accessibility strategy for audiovisual collections seems unclear to respondents at the present time. **Staffing for accessibility purposes appears to be a significant need** as only three respondents could attest to having specific staff or contractors dedicated to this purpose [Question 7]. Almost two-thirds of respondents were uncertain if their unit had any existing on-site hearing or visual impairment accommodations for watching or viewing audiovisual content [Question 10]. And nearly half of respondents were uncertain if accessibility provisions like captions and transcripts were being created for social media and other non-archival audiovisual material at their unit [Question 18].

Inevitably, budgetary matters surrounding the cost of ensuring accessibility emerged as an issue. Less than one-third of respondents have **dedicated budget lines** for creating transcriptions and captioning [Question 12], and fewer than **one-quarter make budgetary provisions for these costs when embarking on a media digitization or reformatting project** [Question 14]. This may explain why the vast majority of respondents reported that **less than 20% of their audiovisual collections currently have existing captions and/or transcripts** [Question 20]. When they do create captions and transcriptions, over **two-thirds of respondents use third-party vendors** for this purpose [Question 13]. Increased exposure to handy resources for creating new captions and/or extracting existing ones from NTSC video could possibly augment these numbers [Questions 15 and 17].

IX. Review of WCAG Compliance Plans at Other Institutions

As part of this research, several responses were obtained from other libraries and academic institutions regarding their policies and protocols vis-à-vis web content accessibility for audiovisual collections. What follows are selected responses from ongoing solicitations, which have been anonymized upon request. (N.B. Emphasis added via bold text by this paper's authors and contributors.)

- **XXXXXX University** comments:

"Accessibility for audiovisual content is something we've been thinking a lot about lately. There's a working group focused on presentation of AV material that has been looking at that specifically - I'm not on that group, but they did recently issue their recommendations. I'm not sure that I can share that document at this point, since it's only been made available internally, but the key points are:

- That transcription and captioning for audio and video be provided to researchers upon request, in a format and virtual environment that fully complies with WCAG 2.0 accessibility requirements.
- That audio description for time-based media with visual components be provided to researchers upon request, in a format and virtual environment that fully complies with WCAG 2.0 accessibility requirements and follows best practices for verbal description.
- As a pilot, XXXXXX University Libraries should allocate \$30,000 per year to meet patron-generated transcription, captioning, and audio description requests.
- That the XXXXXX University viewer interface be compatible with common assistive technologies such as screen readers, and navigable via keyboard, so that researchers who use those tools can navigate it.

- That accessibility consultants, XXXXXX University legal counsel, and users (XXXXXX-UL staff and XXXXXX-UL's broader community of users) provide feedback during the development of the above features where necessary."

- **University of XXXX Libraries** comments:

"We hope to pilot in-house speech-to-text with additional human labor to correct transcripts in the next year, or so. For now, we are just trying to budget for captions (created by a vendor) in our reformatting budgets."

- **XXXXXX Public Library** comments:

"Our Web and Mobile Content Accessibility Policy sets WCAG 2.0, Level AA as our accessibility standard. At the same time this policy was adopted, a **Digital Accessibility Coordinator role (FTE) was created** to help implement this policy. We plan to adapt the policy to stay current to WCAG versioning. Since the Policy was adopted, we have not released updates to our core website that do not meet these standards. We are working through secondary, internal, and third-party built websites. We are also renovating our central website with an accessibility first mindset. This should make accessible many pages and apps published prior to our policy.

- **Guidelines: We have an evolving set of guidelines for compliance with WCAG.** These guidelines have been created **for a range of audiences including: permanent and contracted team members in Engineering, UX, Product, Scrum, Marketing and Communications, and Procurement.** We also have materials to support staff across the organization in content creation on our own and third-party platforms. These guidelines are **supported by outreach and training by the Digital Accessibility Coordinator.**
- Training: **Staff training is frequent** and happens throughout library units. They are adapted to suit the needs and familiarity levels of targeted staff. Sessions range from **structured introductions to working sessions.**
- Procurement: We have **incorporated VPAT** (Voluntary Product Accessibility Template) documentation into our development process and are increasingly collecting these from vendors. We have **introduced language requiring vendors to comply with WCAG 2.0 into our standard contracts.** During the RFP process **we analyze vendors through a set of standard questions** aimed to gauge accessibility of their products and, when appropriate, **conduct accessibility audits** of the product.
- Original Content: With WCAG 2.0 as our standard, departments creating and publishing original content are responsible for covering the cost of text description services for organization created materials created now and in the future. We have **contracted with a preferred vendor** and departments creating content work directly with that vendor. Staff also create text alternatives in house. Departments have started using CART (Communications Access Realtime Translation) on Livestream and captions and transcripts for prerecorded materials. We are working to expand this to all media. As we start to incorporate video description, we anticipate this may require some more work to secure funding and adjust budgets. Departments may be more selective in what material they publish to be sure that what does go on the web hits all of our standards. We have not yet settled on how to manage the cost of providing text alternatives for already published materials that date back years. We anticipate they will be covered by

a central fund rather than the department that originally authored the materials. We may also seek grant and donor funding to cover the costs of a bulk remediation.

- **Collection Materials:** In terms of collection materials, as a way to make content accessible but control costs we are currently **considering an “On demand” service**, modeled on how we handle our print collections. For our print research collections, **we offer training on use of technologies on personal devices as well as library provided technologies** like personal reading machines or CCTVs with speech; this technology can recognize text and read print aloud. Patrons also have access to Bookshare. These **services are funded by both centralized funds for accessible collections and general collection development**. When there is no existing accessible format or these technologies can’t remove barriers, we **provide a digitization and reformatting on demand process. Funded mostly through the digital department**, staff digitizes and reformats materials into accessible formats including machine readable digital text, which can be magnified, and otherwise manipulated, and synthesized speech audio books delivered digitally and on cartridges for NLS Talking Book Players. The library also has an extensive circulating collection: included in the library’s services and collection development policies are attention to accessible formats of our circulating collection. This includes collecting print, large print, Braille, and talking books of popular titles in digital and print format.
- **Audiovisual:** We want to improve the accessibility of circulating AV materials. Staff **are being trained on inclusive public programming**, including showing AV materials with captions and audio description. Again, we work to make sure collections developed specifically for programming include materials with these text alternatives. We also provide CART and ASL (American Sign Language) services, when requested, for public programs. These services are paid for out of a central fund.”
- **XXX University Libraries** comments:

“XXX University’s policy on accessibility states that all university websites published after November 1, 2016 are required to be compliant with WCAG 2.0 AA, and that older sites are expected to upgrade over time. XXXU’s **Communications office has developed guidelines for streaming video**. However, these guidelines are focused around promotional videos and instructor/university-created instructional videos, not archival materials. We don’t currently have any specific accessibility guidelines for archival video and audio and do not have a specific budget or planning process underway for any large-scale accessibility remediation of these materials.”
- **XXXXXX University Libraries** comments:

“As for WCAG, a Smithsonian white paper is a great idea and I'm really glad to learn that this research is underway. We've been chatting with Dave Rice about this a lot and, internally at XXXXXX University Libraries, we just started conversations for XXXXXX University Libraries' compliance. **There's a new department established within the university** to manage this and we have an early May [2019] meeting to discuss how Special Collections content will be impacted and consider **workflows and budget**. IT Accessibility team as well as the Libraries U/X and XXXXXX University TV are all involved.”

- **Vendor X** comments regarding their new video platform service:

“To respond to your questions, we are **currently using .VTT and OHMS XML as import formats** to populate the transcript (which automatically populates captions in our video player) and index data in our proprietary video player’s back end. While it's not out of the question that we will publish guidance on WCAG in the future, it's currently not a planned effort. I can say that we are considering WCAG in the development of our video player and will continue to be as responsive/compliant as we are able in the continued evolution of the platform. We know that this is a big question and issue for many of our colleagues and clients. I can also say that under any circumstance, our video player will unquestionably be advantageous in this regard relative to other platforms and options.”

X. Recommendations and ‘Pro Tips’

General

Work to build a groundswell for undertaking and ensuring accessibility for audiovisual works online. It is certain that others in your unit, especially Social Media and Information Technology (IT) teams, are facing similar challenges.

- Schedule a conversation with your unit’s IT and/or web services lead. Inquire as to extant (or absent) IT department protocol for ensuring accessibility to audiovisual content (e.g. livestreams, event webcasting).
- Inquire as to IT budget lines for accessibility services such as captioning and how collections materials might qualify for such funds. Keep in mind that per SD 950, a unit’s accessibility responsibilities belong in no small part to its web managers. Funding from general unit exhibition, visitor services, and/or other sources may be available outside regular audiovisual collections care and digitization budget lines.
- Talk to your colleagues in the OCIO, DPO, Web Services, and DAMS teams about how best to approach ensuring accessibility and obtaining funding to do so. Squeaky wheels get grease.
- Inquire as to a specific staff member responsible for ensuring accessibility compliance at your unit. If a specific staff member is not designated, consider suggesting that accessibility compliance be officially established within a specific staff member’s roles and responsibilities.
- If a specific staff member is not designated as responsible for ensuring accessibility compliance at your unit, consider creating a “Digital Accessibility Coordinator” position within your unit. Remember that per SD 215, responsibility for ensuring accessibility lies with museum and research unit directors. Ensuring accessibility to digital collections has a direct impact on the ability to make them available via your unit’s websites and internet portals.

Captioning

While neither Smithsonian policy nor the W3C provide specific guidelines for exactly how captions should be created stylistically, below are some key considerations to guide audiovisual collections managers through this process.

- Build captioning and description costs into production and digitization budgets so that content can be easily made available online in compliance with Smithsonian accessibility policies, including WCAG 2.0, Guideline 1.2.2 Captions (Prerecorded).
- Outsourcing for the creation of video captions is usually less expensive than paying for Smithsonian staff to create them.

- ‘Proof-watch’ and review captions once they are completed to ensure accuracy and adherence to preferred styles. Auto-transcription tools (e.g. YouTube) are not 100% accurate.
- Consider training a Cataloger in captioning so that any ‘proof-watching’ or review processes can also provide a ‘double-dip’ opportunity to catalog materials.
- Follow a grammar and style guide, such as William Strunk Jr. and E. B. White’s *Elements of Style*, for grammatical guidance and stylistic unity. (See also APPENDIX B NMAAHC Audiovisual Captioning Guide.)
- Identify speakers. When assigning speaker identifications for dialogue, focus on a speaker’s diegetic role and/or identity in lieu of gendered or racial assumptions.
- Use block brackets for any non-diegetic, non-verbal, caption descriptions. This can ensure that WCAG 2.0, Guidelines 1.1.1 for Non-text content, 1.2.1 Audio-only and Video-only (Prerecorded), and 1.2.3 Audio Description or Media Alternative (Prerecorded) are met.
- Scripted dialogue appears in fully formed sentences; documentary dialogue, less so. Make reading easy for your audience without eradicating speech styles.
- Captions require adequate on-screen display time to be read by their intended audience. Test this by reading each caption aloud.
- Song titles alone don’t provide much information for captions audiences. Transcribe the lyrics.
- Use descriptive indicators for non-verbal plot and on-screen gestures (e.g. [telephone rings]).
- Conserve caption screen space whenever possible (e.g. numerals).
- Consider creating captions for all title cards, credits, and other on-screen text to accommodate screen-reading software, and/or Audio Description tracks.
- For oral histories, timed-text captions tracks will prove more useful for the dual purposes of 1) on-screen captions and 2) keyword-searchable interview transcripts. If traditional oral history transcriptions are provided for in an oral history production budget, consider creating timed-text captions instead, as they can serve both purposes.

Compliance

Ensuring accessibility is everyone’s responsibility, and the spectrum of responsible parties outlined in OCIO’s *Technical Note: IT-950-TN06* makes this abundantly clear. Within any large institution, however, it’s not always evident where ‘the buck stops.’ While the Office of the General Counsel has suggested that the present report cannot provide specific compliance evaluations for analyzing Smithsonian websites, below are some of the most common challenges encountered in meeting WCAG 2.0 Level AA for audiovisual digital assets.

For more in-depth and interactive techniques for ensuring compliance with WCAG 2.0, consult the W3C’s *Quick Reference* guide: <https://www.w3.org/WAI/WCAG21/quickref/>.

Pro Tips

- **Problem:** Captions contain typographic errors, perhaps due to automatic transcription.
Solution: Proof-watching and review can ensure accuracy and legibility for captions.
- **Problem:** No captions and/or descriptions are present.
Solution: Make inclusion of accurate captions and/or a description sidecar file (such as a .VTT) part of your unit's Submission Information Package (SIP) and workflow whenever uploading audiovisual digital assets to the Smithsonian DAMS and/or your access website.
- **Problem:** Captions formatting is problematic across different operating systems and text-reader software.
Solution: Always use UTF-8 encoding and try to build the Web Video Text Tracks (.VTT) format as the base of your captioning workflows.
- **Problem:** You encounter comments such as, "How do I make captions?", or the like, from a co-worker.
Solution: Articulate your unit's audiovisual accessibility protocol, style guide, and/or captioning workflows in a document. Working together with your unit's IT staff to develop a suite of quality documentation will prove invaluable to staff less familiar with the concepts of accessibility for audiovisual content. It will also demonstrate a good-faith effort in case of a potential event of alleged accessibility compliance violation.

APPENDIX A

SI Survey of Audiovisual Web Content Accessibility Practices & Policies

The following charts and tables provide detailed results from the April 2019 survey of current captioning, description, and digital audiovisual content accessibility practices and policies at the Smithsonian.

Question 1. What Smithsonian Institution unit do you most closely work with/for?

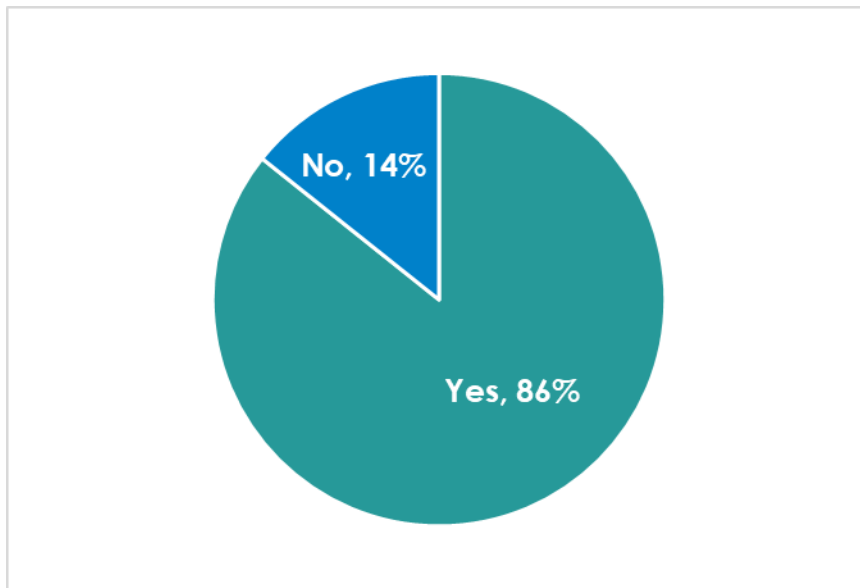
Unit	Number of Responses
AAA	1
CFCH	1
FSG	1
NASM	1
NMAAHC	1
NMAH	1
NMAI	2
NMNH-Anthro	1
OCIO	1
SAAM	1
SIA	1
SIL	1
STRI	1
Total	14

Question 2. Describe your unit's approach to making audio visual collections available on the internet (check all that apply)?

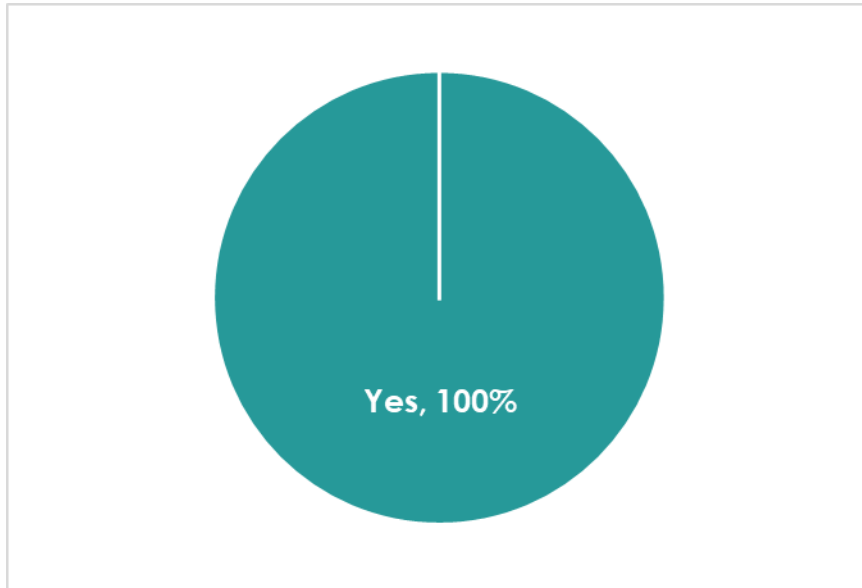
- Unit makes some audiovisual collections available online via YouTube.
- Unit makes some audiovisual collections available online via another platform (not YouTube).
- Unit does not make audiovisual collections available online.

YouTube	Other platform	None
X	X	
X	X	
		X
X		
X		
X		
		X
X	X	
X	X	
X		
X		
X	X	
		X
	X	
10	6	3

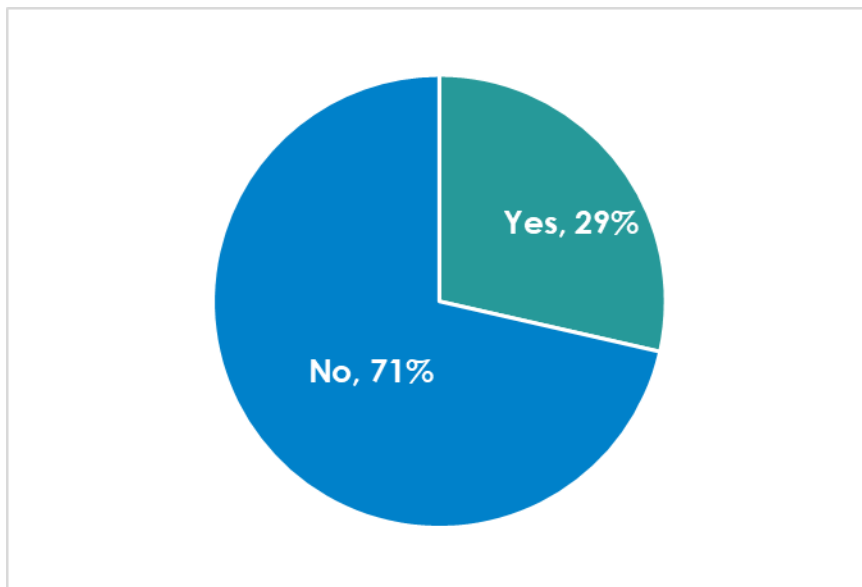
Question 3. Have you ever viewed broadcast television with Closed Captions?



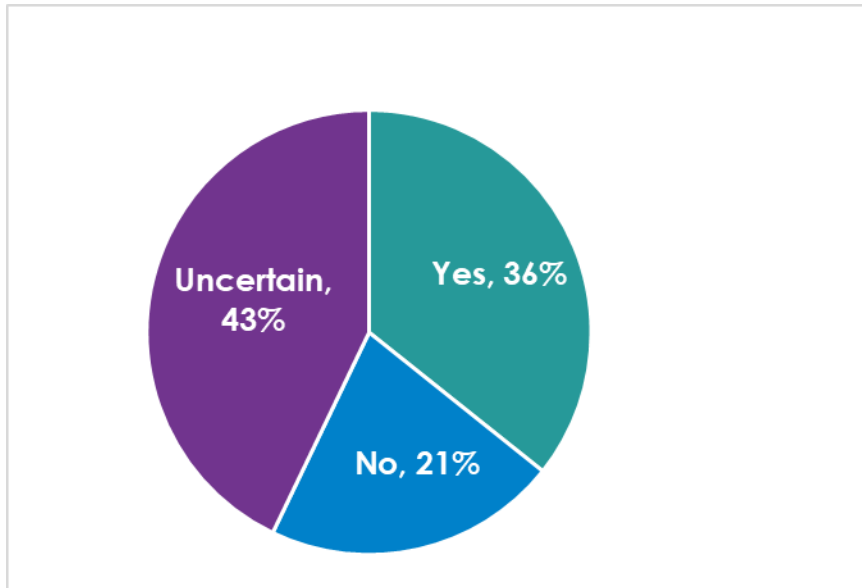
Question 4. Have you ever viewed online/ internet video content with Closed Captions?



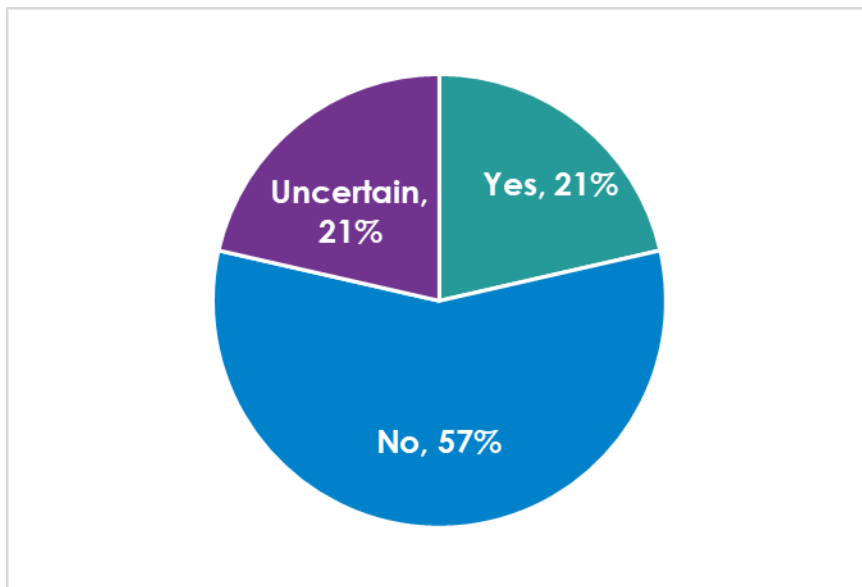
Question 5. Have you ever used screen-reader software?



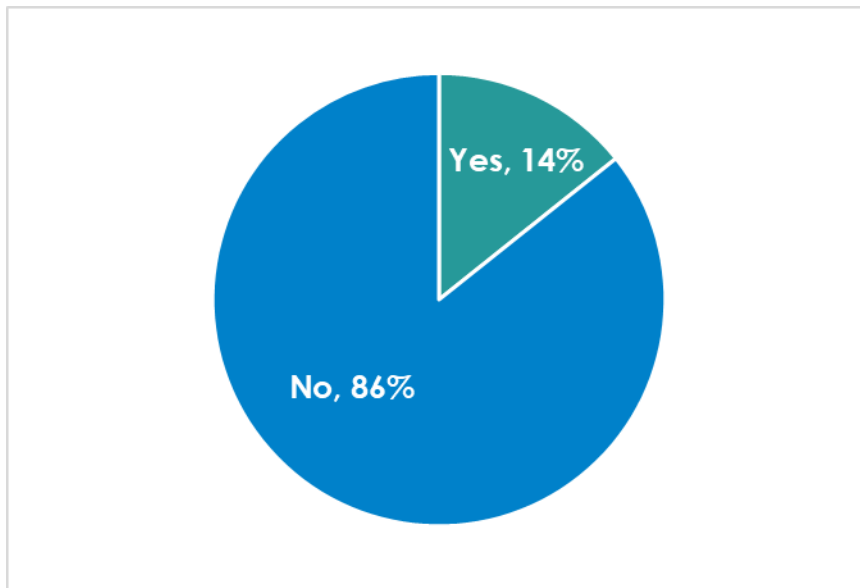
Question 6. Is facilitating accessibility to audiovisual collections content for hearing- or visually-impaired patrons an articulated job responsibility for you, or a colleague in your unit?



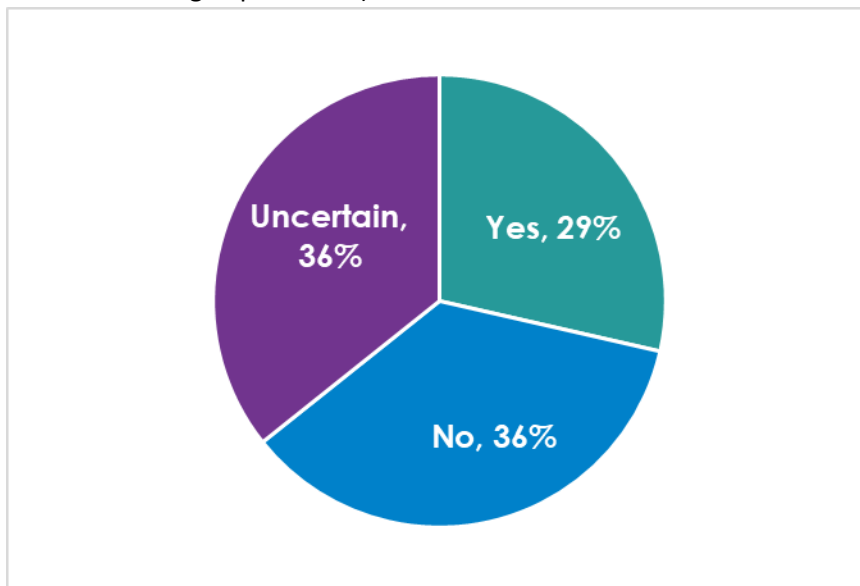
Question 7. Does your unit have staff or contractors dedicated to accessibility as it relates to transcripts/textual descriptions/captions?



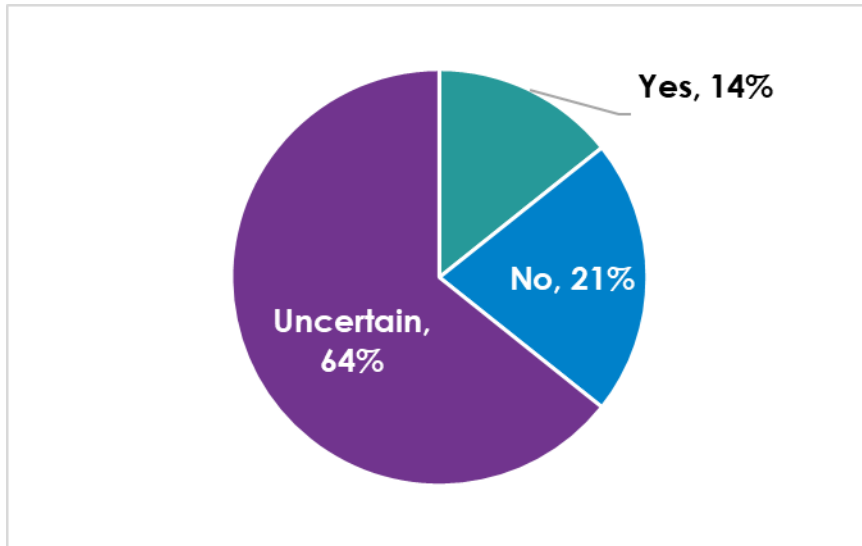
Question 8. Have you ever received any professional training, workshop education, or specific how-to guidance on creating transcripts/textual descriptions/captions?



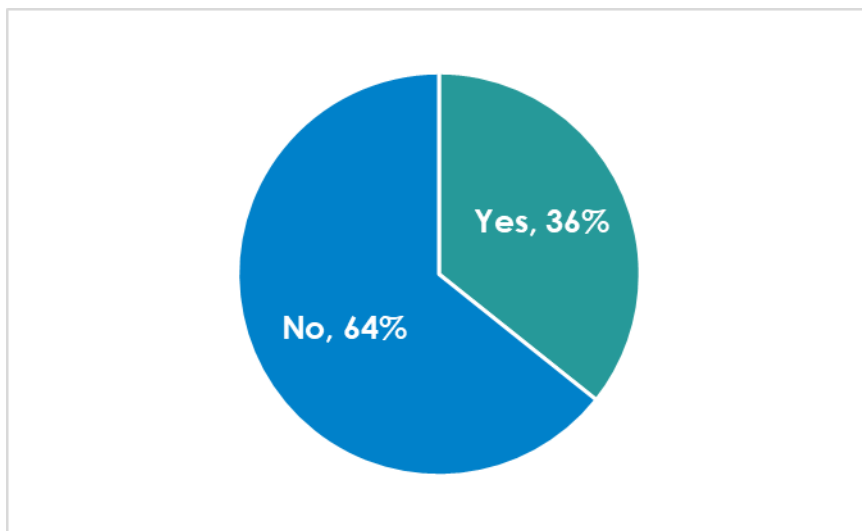
Question 9. Does your unit employ any collections staff with perceptual accessibility impairments (i.e. visual or hearing impairments)?



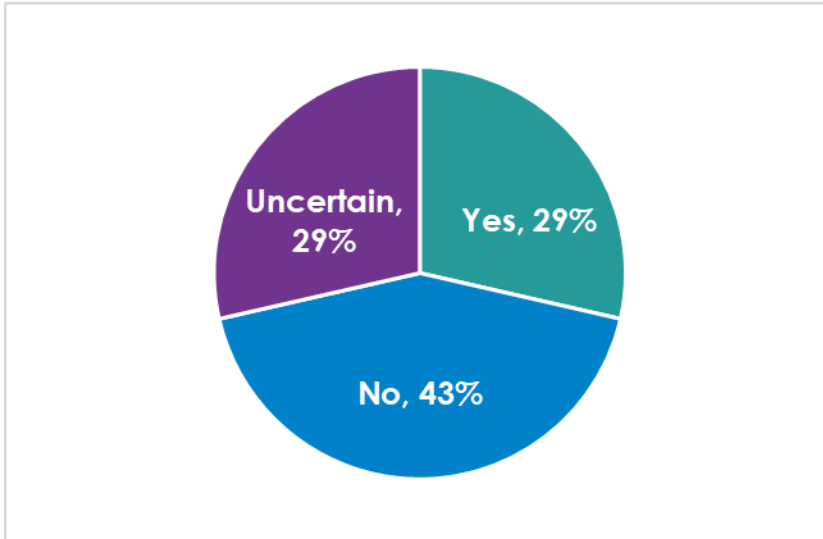
Question 10. Does your unit have any on-site hearing or visual impairment accommodations for viewing or listening to audiovisual content in its collections?



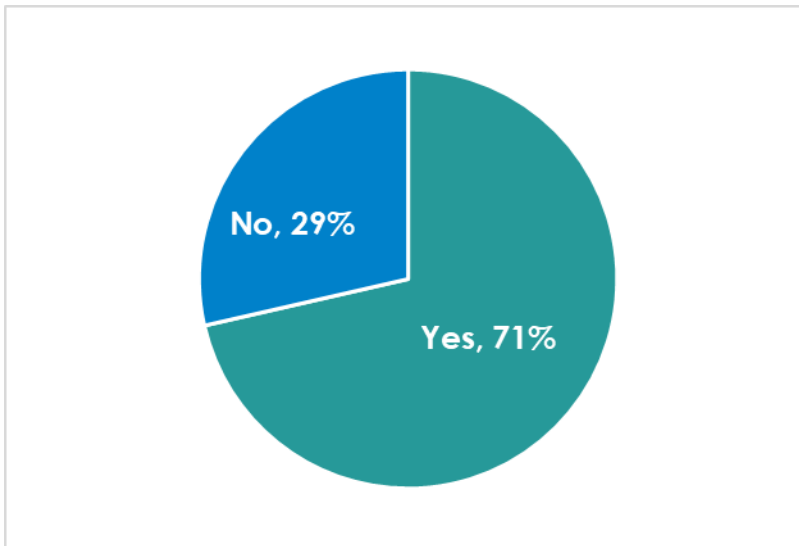
Question 11. Does your unit make transcripts/textual descriptions/captions for audiovisual content available and downloadable on its website as a text document (ie. as .DOC, .PDF, .VTT, .SRT, or other type text files), separate from their functionality as part of an audio or video player?



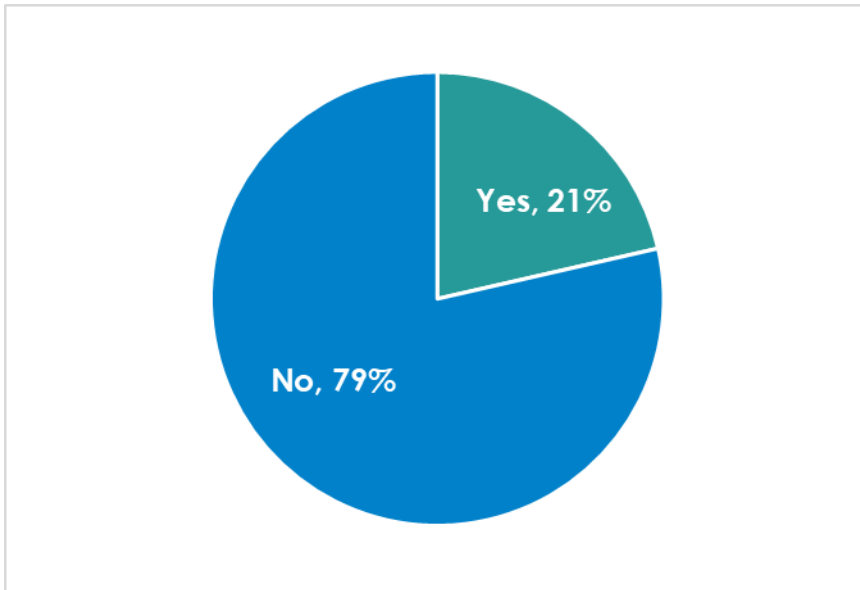
Question 12. Does your unit currently have any dedicated budget lines for creating transcripts/textual descriptions/captions for audiovisual content in its collections?



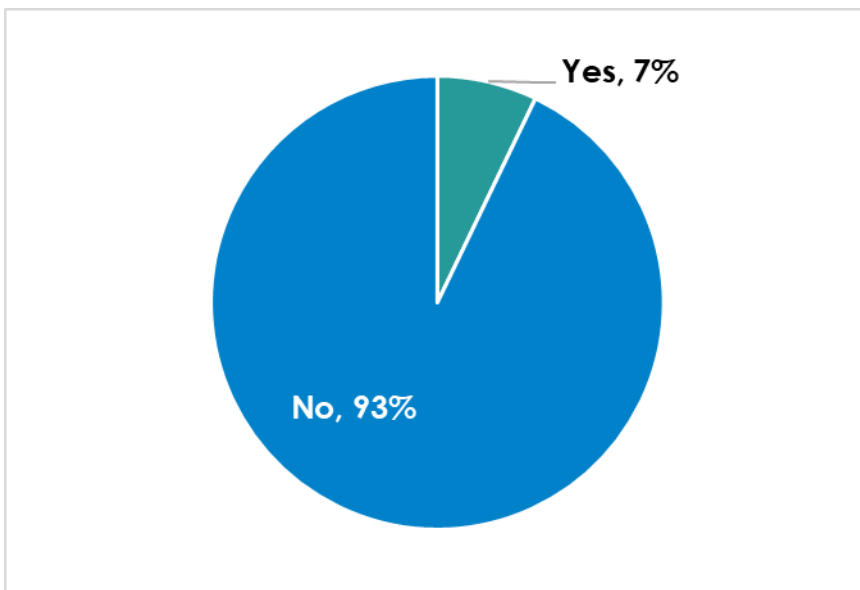
Question 13. Does your unit currently use outside third-party vendors for creating transcripts/textual descriptions/captions for audiovisual content in its collections?



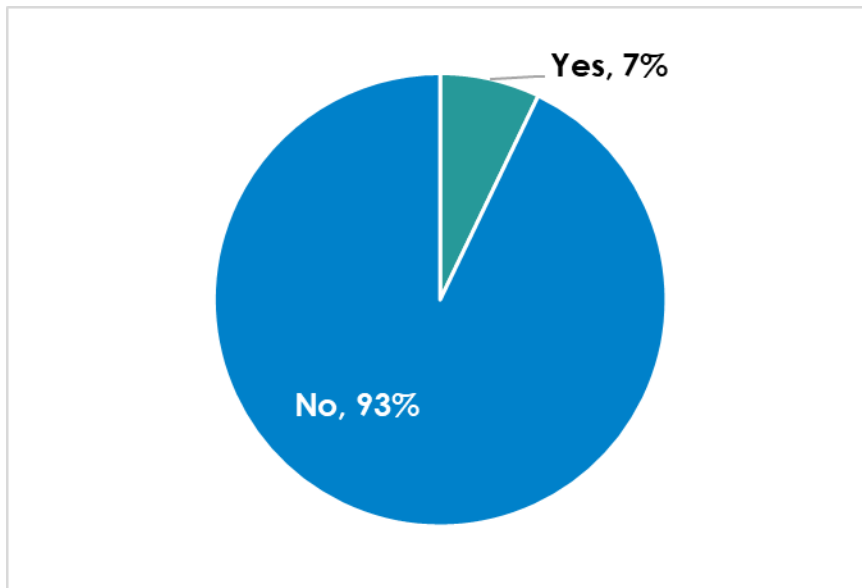
Question 14. When embarking on a media digitization project does your unit currently budget for the creation of transcripts/textual descriptions/captions for resulting digitized audiovisual content?



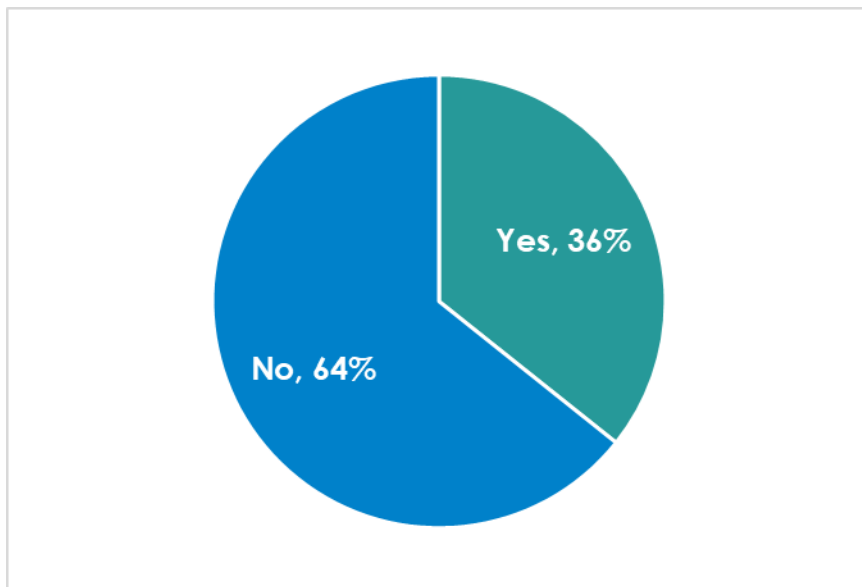
Question 15. When digitizing analog videotapes, does your unit have a workflow to check for the existence of and/or extract line-21 (CEA-608) closed captions?



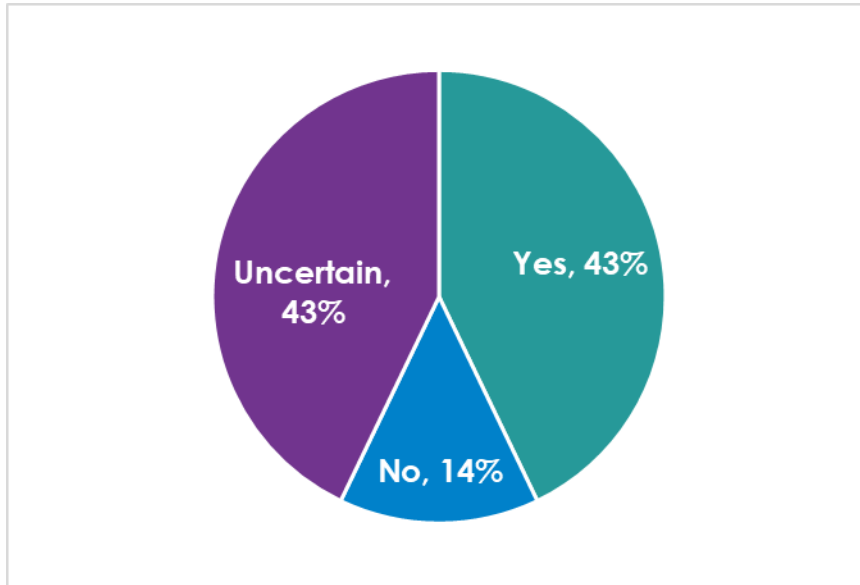
Question 16. Does your unit have articulated and/or specific guidelines for the creation of transcripts/textual descriptions/captions (for audiovisual content or otherwise)?



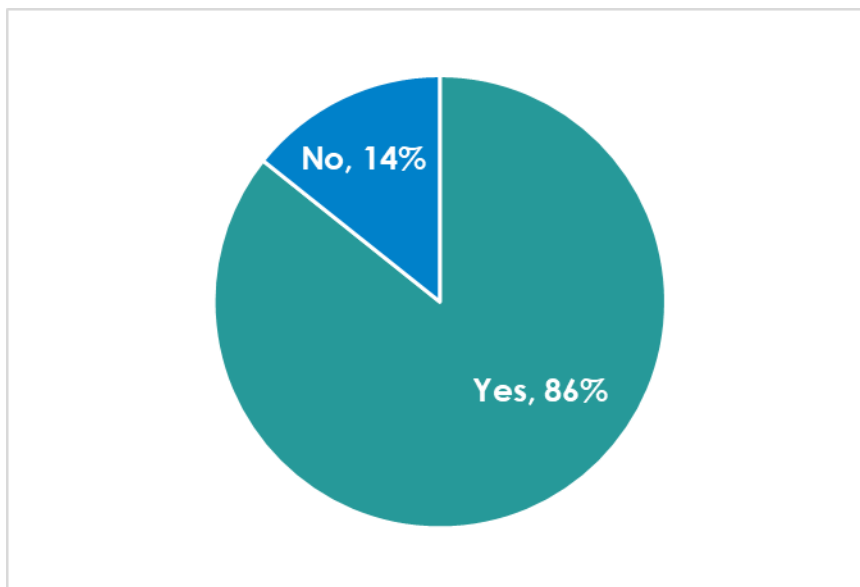
Question 17. Are you aware of any software or tools at your unit available for creating transcripts/textual descriptions/captions for audiovisual content?



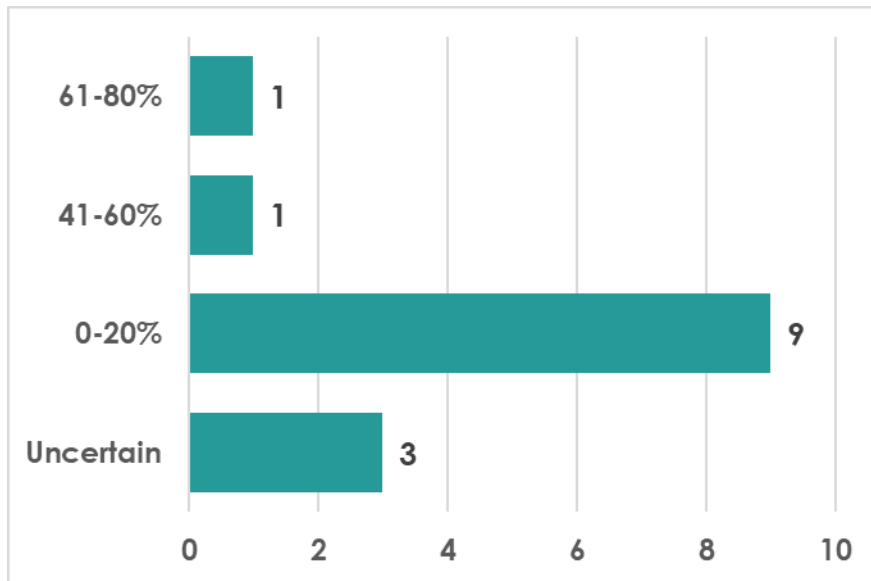
Question 18. Does your unit currently create transcripts/textual descriptions/captions for newly-produced video content, social media video content, webcast content, oral histories, or other video productions (i.e. non-archival video content)?



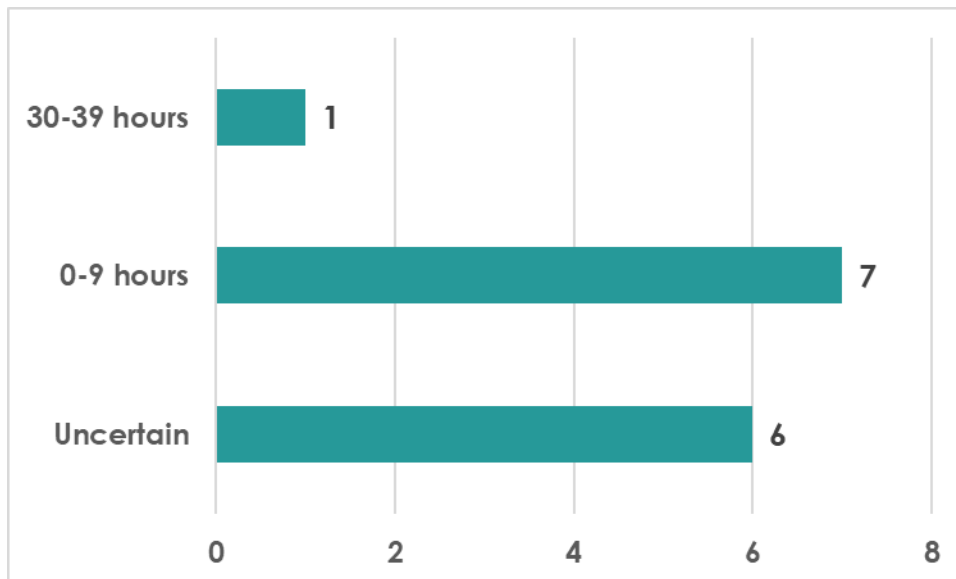
Question 19. Are you aware that federal agencies publishing any content (images, audiovisual) online are legally obliged to provide transcripts/textual descriptions/captions?



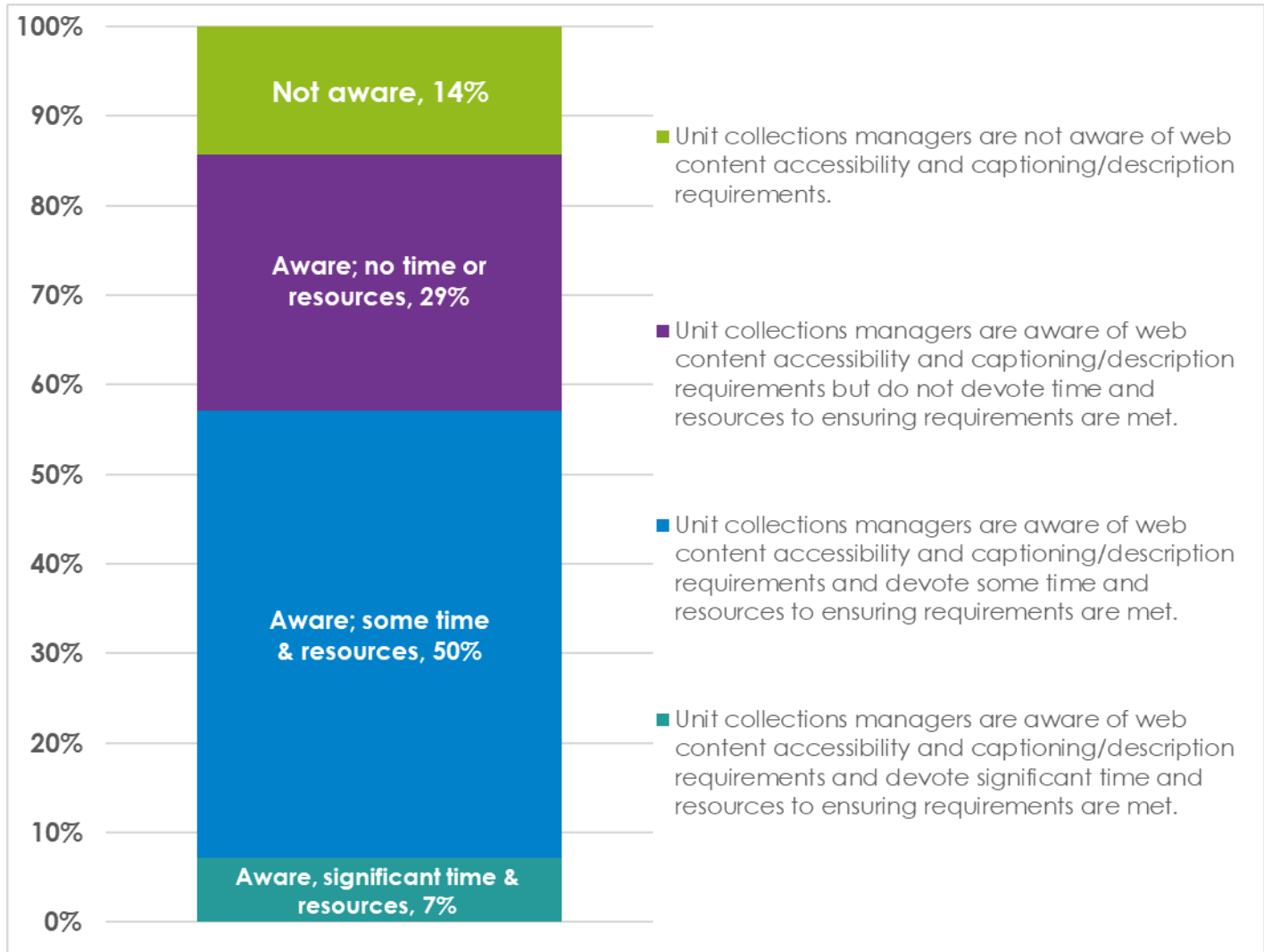
Question 20. What percentage of your unit's audiovisual collections have transcripts/textual descriptions/captions already created for them?



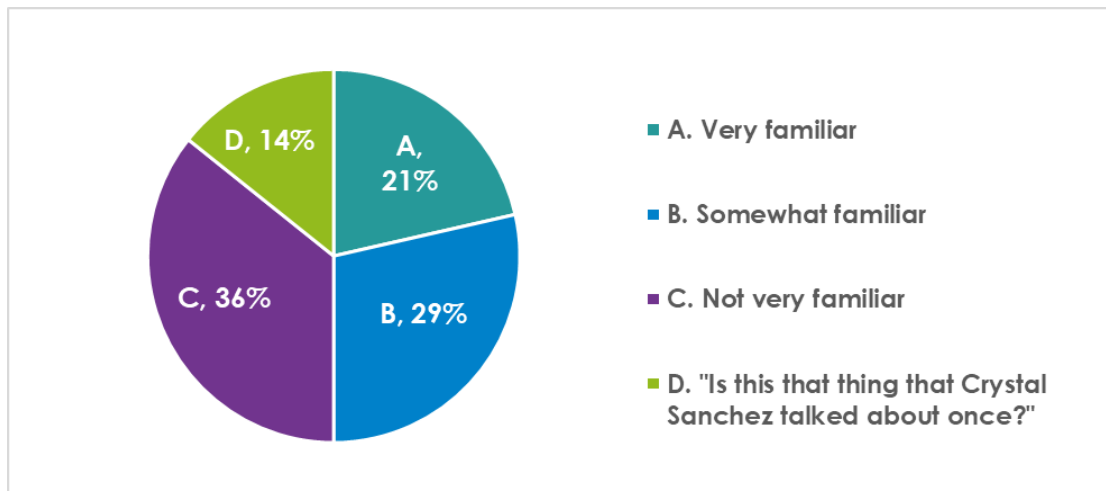
Question 21. On average how much time a week do staff in your unit spend working on web content accessibility, including making transcripts/textual descriptions/captions for audiovisual collections?



Question 22. Which statement best describes your unit's policy on web content accessibility and captioning/description requirements?



Question 23. How familiar are you personally with Web Content Accessibility Guidelines 2.0?



Question 24. Which of the following organizational areas would you say has provided the most guidance with regards to articulating web content accessibility protocol for collections material? (Check multiple boxes if appropriate.)

OCIO Staff	AVAIL	Unit IT Team	Unit Web Team	Unit Oral History Dept	Other	None
					X	
X	X					
X	X					
	X		X			
X						
			X			
						X
	X				X	
X						
	X					
	X			X		
X		X	X		X	
X		X				
X	X					
7	7	2	3	1	3	1

Question 25. If you replied "Other" to Question #24, please indicate your source(s) for web content accessibility guidance.

- Media Archivist
- Unit Media Staff
- SI Access Dept & Webmasters Meeting

Question 26. What resources would you be most interested in?

- Funding for outside contracts to create transcripts/textual descriptions/captions.
- Funding for internal SI staff/contractors to create transcripts/textual descriptions/captions.
- Regulations and/or guidelines from Smithsonian on "how to" create transcripts/textual descriptions/captions.
- Training on "how to" create transcripts/textual descriptions/captions.

Outside Contracts	Internal SI Staff	Regulations/ guidelines	Training
X	X	X	X
X		X	X
X	X	X	X
	X	X	X
	X	X	X
X	X	X	X
X		X	X
X	X		
X			
X	X	X	X
	X		X
X	X	X	X
		X	
X	X	X	X
10	10	11	11

GENERAL FEEDBACK: If you would like to provide additional feedback, directly respond to, or elaborate on any specific survey question, a text field is provided below.

- *Thank you!*
- *As an added use case, we have a lot of paper transcripts of recordings in our collections and have discussed providing access to digitized paper transcripts as a placeholder for captions. Also, we have significantly higher quantities of audio than video/film in our collections and so it would be important to use for training/guidelines to explicitly include audio-only accessibility solutions.*
- *Though we know the requirements for captions, it is uncertain who should be actually making that happen. As far as I can tell, only one person is doing it, but he doesn't have a great deal of support.*
- *Accessibility requirements (captions/textual descriptions) have been implemented for our still image and print materials, and it is something very much on our minds for audiovisual materials, but it has not yet been implemented on our website.*
- *It will be helpful to have an executive summary to share findings with unit stakeholders in order to raise awareness of our collective obligations. Also, it would be helpful to know what the CIMC and the members of the CIS-IRM allocations sub-committee think about potential projects with captioning components.*

APPENDIX B

Example Style Guide: NMAAHC Audiovisual Captioning Guide



Draft Version 0.8 (November 2019)

By: Walter Forsberg, Web Content Accessibility Specialist (Contractor);
Blake McDowell, Media Archivist and Conservator

I. INTRODUCTION

Transcripts, captions, sub-titles, and other textual alternatives to audio and video appear in a variety of file formats and textual styles. This guide provides some general principles and elementary considerations about implementing captions for audiovisual content, and outlines information on captioning and transcription style and file formats.

II. ELEMENTARY CONSIDERATIONS FOR CAPTIONING

II.1 Follow William Strunk Jr. and E. B. White's *Elements of Style* for grammatical guidance.

Looking for guidance on whether, or not, to use a comma for the abbreviation for “junior,” when it follows the name “William Strunk” (hint: do not)? Curious about that possessive singular apostrophe, when referencing Frederick Douglass (hint: it’s Douglass’s)? Just ask Strunk and White! They literally (“often incorrectly used in support of exaggeration or violent metaphor”) wrote the book on these matters.

II.2 Identify speakers. When assigning speaker identifications for dialogue, focus on a speaker’s diegetic role and/or identity, in lieu of gendered or racial assumptions.

Speaker identifications such as a character name (in fiction works) or actual names (in documentary works) are accurate and generous ways of identifying speakers of dialogue. Performing a modicum of pre-captioning research can help identify these speaker identifications, immensely. Sometimes accuracy in such identifications is not always possible or self-evident. In such circumstances, attempt descriptive identifiers that enable agency and avoid possibly inaccurate or offensive mischaracterizations.

For example: Use “Undertaker” or “Executive assistant,” instead of “Man” or “Woman.” Use “Newspaper seller” instead of “Paper Boy.” Use “Paramour,” instead of “Unidentified buff dude” or “Attractive nameless dame.” Use “Coiffed person” instead of “Becky with the good hair.” As a last resort, employ descriptive screen geographies for identification, such as, “Person, at left,” or format your captions so that they employ caption placement beneath the appropriate speaker.

Re-identify speakers after another speaks, or another descriptive indicator is employed. This will help captions audiences keep track of who’s saying what.

II.3 Use block brackets for any non-diegetic, non-verbal, caption descriptions.

Follow named/described speaker identifications with a colon—e.g. “[Bobby Seale:] You don't fight racism with racism, the best way to fight racism is with solidarity.” Don't use a colon for descriptive indicators—e.g. the sound of a shotgun firing would be “[Shotgun firing]” not “[Shotgun: firing]”.

II.4 Scripted dialogue appears in fully-formed sentences; documentary dialogue, less so. Make reading easy for your audience, without eradicating speech styles.

Because it is scripted, dialogue in fiction works will often appear spoken in full sentences. However, real-life humans rarely, like, uh, speak in such a manner. Nevertheless, captions-readers will have an easier time if captions appear without every excruciating “um,” “er,” and half-finished spoken thought. If disruptive speech patterns are essential to content, use a descriptor in the Speaker Identification block bracket to precede transcribed dialogue—for example, “[Preston Lay, Jr., stuttering:] I don't know,” instead of “[Preston Lay, Jr.:] I, I, I, I, d... d... d... don't know.”

As the DCMF's *Captioning Key* reminds us: “all captioning should include as much of the original language as possible; words or phrases which may be unfamiliar to the audience should not be replaced with simple synonyms. However, editing the original transcription may be necessary to provide time for the caption to be completely read and for it to be in synchronization with the audio.”²⁹

II.5 Captions require adequate on-screen display time to be read by their intended audience. Test them by reading each, aloud.

A good rule of thumb to follow is that each on-screen set of captions requires a minimum of one second of screen time. Decent captioning tools like YouTube and MacCaption indicate increments of a second to ensure that your captions audience has time to read all transcribed text. Verify that any potential captioning vendor follows this rule. Some older, legacy captions and sub-title file formats have limitations on the number of characters per line of text. For example, the “Scenarist Closed Caption,” or .SCC, format cannot encode more than 32 characters per line, so try to avoid captions lines that are excessively long.

II.6 Song titles, alone, don't provide much information for captions audiences.

Transcribe the lyrics.

While James Brown was hardly the world's most outspoken feminist, consider the redemptive second clause of chorus for his 1966 song, *It's a Man's, Man's, Man's World*: “But, it would be nothing, nothing, without a woman or a girl.” Captions audiences will take-away quite different meaning from a full transcription of the song's lyrics, than if the title alone is presented in block brackets. Consider using quotation marks for the song lyrics. N.B. Copyright sensitivity may impede adherence to this consideration for more risk-averse entities.

²⁹ Described and Captioned Media Program, *Captioning Key*. Accessed at: http://www.captioningkey.org/quality_captioning.html.

II.7 Use descriptive indicators for non-verbal diegesis and gestures.

Visually impaired captions audiences using screen reader software will glean more from your captions if you provide descriptive indicators in block brackets to convey what’s happening on-screen when dialogue isn’t being spoken. Baseline industry standards such as “[Laughter.]” or “[Coughs.]” are fine, however additional descriptive indicators such as “[Jamal trips on shoelace.]” or “[Chokes on food.]” can provide additionally informative context.

As with music lyrics, musical styles, tonalities, instrumentation, and other sonic qualities can convey much to captions audiences. Consider adding qualifying description, such as, “[Upbeat orchestral swing music],” to a descriptive caption instead of the listless and perfunctory, “[MUSIC].”

II.8 Conserve caption screen space whenever possible.

Notwithstanding recommended adherence to grammatical and style principles espoused by Strunk and White’s *The Elements of Style*, minimizing on-screen characters by using abbreviations will be useful. Shorter captions make for speedier and easier reading. In lieu of “8 o’clock” (9 characters) consider using, “8:00” (4 characters), or, something even shorter--“8.” After an initial speaker identification of “[Louise Thompson Patterson:],” consider subsequently using her last name only, “[Patterson:],” or a further abbreviation, such as, “[LTP:].” Using “OK” instead of “okay” will prove more historically accurate for this abbreviation of “Oll Korrekt,” and will also save caption screen space. An ampersand takes up two fewer characters than does “and.”

II.9 Create captions for all title cards, credits, and other on-screen text.

Screen-reading software generally won’t be able to read such text-based on-screen information unless it is transcribed.

III. HOUSE STYLE

Notably, the W3C’s Web Content Accessibility Guidelines are not accompanied by any ‘style guide,’ per se, on how captions should appear. Because so many different published approaches to transcription are feasible, published or otherwise, it is a good idea to develop a ‘House Style’ outlining specific tendencies, preferences, and other favored approaches.

III.1 House Style: Examples

Many excellent style guides for transcriptions already exist--particularly available from oral history programs at academic institutions.

- The Smithsonian Institution’s Archives of American Art published a revised version of their *Oral History Program Style Guide* in May 2019, available here: <https://www.aaa.si.edu/documentation/oral-history-program-style-guide>.
- Columbia University published a revised version of its *Oral History Transcription Style Guide* in August 2018, available here: <https://www.incite.columbia.edu/publications-old/2019/3/13/oral-history-transcription-style-guide>.
- Baylor University’s Institute for Oral History published a revised version of its *Style Guide: A Quick Reference for Editing Oral History Transcripts* in March 2018, available here: <https://www.baylor.edu/oralhistory/doc.php/14142.pdf>.

For guidance with specific regards to web content accessible captions for audiovisual content, both commercial video platform Netflix and the Department of Education-funded organization, Described and Captioned Media Program (DCMP), have published excellent detailed guidelines for captioning and describing audiovisual content.

- The DCMP's *Captioning Key* is considered a 'best practice' guide, available here: http://www.captioningkey.org/about_c.html.
- Netflix's *Timed Text Style Guide* is a useful resource, resulting from actual federal litigation, available here: <https://partnerhelp.netflixstudios.com/hc/en-us/articles/215758617-Timed-Text-Style-Guide-General-Requirements>.

House styles and unit-specific style guides can be said to be ever-evolving. In this regard, NMAAHC is no exception. Below are a few house style rules-of-thumb encountered in its first year of captioning audiovisual content

III.2 House Style: Textual Appearance

- **Caption placement:** Not necessary. Optional caption placement acceptable, especially when accurate speaker identifications are not possible.
- **Font:** No specific font style requirements; YouTube's captioning software defaults to downloads in Monaco size 14. Font size should be uniform throughout captions transcript and should not be crazy (hint: 10-point, 12-point, or 14-point seem reasonable).
- **Languages (non-English):** Unless you are a native speaker, thereof, do not translate. Instead, use a block bracketed descriptive indicator, such as: "[Kathleen Cleaver, speaking in French]."
- **Line appearance:** No more than two lines of captions should appear on-screen at one time. Each caption screen should last a minimum of one second. No specific character-per-line limitation, but captions should appear as uniform in width as possible. The shorter the better and more readable for captions audiences.
- **Multiple speakers:** When captioned dialogue for two speakers appears on-screen at the same time, employ block bracketed speaker identifications for each. When a group voices dialogue, indicate so using block bracketed speaker identifications. For example, "[Choir, in unison:] Amen."

III.3 House Style: Specific Words and Mannerisms

- **African American:** No hyphen, as per past museum usage.
- **black:** As used to describe "black people," or "African Americans," generally not capitalized. It could be however past museum usage has not capitalized it. (Likewise, **white** and **colored** are not capitalized terms, as per past museum usage.)
- **Negro:** Always capitalized, as per W. E. B. DuBois's famous footnote: "I shall, moreover, capitalize the word, because I believe that eight million Americans are entitled to a capital letter."³⁰
- **The Stepin Fetchit paradigm:** Captioning dialogue for caricatured racist stereotypes can be complicated. While

³⁰ W. E. Burghardt DuBois, "The Philadelphia Negro: A Social Study" (New York: Lippincott, 1899), 1. For more on this topic, see: Donald L. Grant and Mildred Bricker Grant, "Some Notes on the Capital 'N,'" in *Phylon*, Vol. 36, No. 4 (4th Qtr., 1975), pp. 435-443.

array of offensive behavior and deliberately ignorant verbal dialogue (what Donald Bogle calls, “arch-coon”).³¹ When captioning such dialogue, consider using properly-spelled and punctuated words in lieu of attempting a verbatim colloquial textual appearance. Individual captioning treatments will likely be subjective in this instance. A descriptive note in block-brackets preceding dialogue may be used to infer caricatured speech patterns.

IV. CAPTIONS FILE FORMATS

NMAAHC employs the .VTT (‘Web Video Text Track,’ or ‘WebVTT’) file format for captions and descriptions because the OCIO Media Asset Delivery System (MADS) player is based on J-Player, which mandates use of the WebVTT format. WebVTT is also the specified captions format for HTML5 players, so in addition to its formal simplicity it would appear to hold promise as a future-proof format.

Many other captions, transcription, and subtitle file formats exist. Facility in translation from one file format to another is dependent on which formats one is translating to, and from. Many captioning software such as the YouTube Creator Studio tool, Amazon Transcribe, Closed Caption Creator, and MacCaption can automate such translations, however human-verification and a complete proof-watching is always recommended. Such human-verification and proofing can ensure that subject-specialized knowledge and context form part of the captions, enabling a more egalitarian text alternative to verbal speech and cues.

Below is a list of several file formats, common for captions and subtitles. While identifying a captions or subtitles file format is most easily accomplished by looking at its three character extension, included are some helpful examples of what each’s textual formatting looks like.

- **SCC:** “Scenarist Closed Caption” file. Typically used in analog broadcast workflows to represent line 21 closed captions, as some software can automatically generate an SCC file from a video source containing line 21 closed captions. Also widely used in early iTunes, iPhone, and iPod content. Format protocol is limited to 32 text characters per line of captions/subtitles. SCC files are double-spaced with interleaved blank lines, and consist of a time-stamp in HH:MM:SS;FRAMES and two-byte hexadecimal works, each separated from each other using spaces.

Example of SCC formatting (via www.theneitherworld.com):

```
01:02:53:14    94ae 94ae 9420 9420 947a 947a 97a2 97a2 a820 68ef f26e
```

```
01:02:55:14    942c 942c
```

³¹ Donald Bogle, *Toms, Coons, Mulattoes, Mammies, & Bucks: An Interpretive History of Blacks in American Films*, fourth edition, (New York: Bloomsbury, 2013), 41.

- **SRT:** “SubRip Subtitle” file. A simple and common captions and subtitles sidecar format, particularly among collector and torrent communities. Common format used with Matroska-wrapped digital video files. Supported by video formats such as DivX, and many DVD-ripping software such as SubRip and Mac the Ripper. SRT files consist of four parts, all textual.

Example of SRT formatting:

1
01:04:23,301 --> 01:04:27,102
What we've got here
is failure to communicate.

Explanation of SRT formatting:

Caption/Subtitle sequence number
HH:MM:SS,MIL
Caption/Subtitle text
Caption/Subtitle text

2
01:04:27,103 --> 01:04:30,202
Some men
you just can't reach.

Caption/Subtitle sequence number
HH:MM:SS,MIL
Caption/Subtitle text
Caption/Subtitle text

- **STL:** “Spruce Subtitle File.” Primarily used with DVD Studio Pro software. Not textually editable.
- **WebVTT:** “Web Video Text Tracks” file. A W3C standard developed as a simple, purely textual captioning file format, specifically for internet-based videos. WebVTT is the required caption format for HTML5 browser video.

Example of WebVTT formatting:

WEBVTT
Kind: captions
Language: en

00:00:59.760 --> 00:01:00.480
[Elsie Bellwood:] Hey, Cookie.

00:01:00.480 --> 00:01:01.100
[Piano Player Cookie:] Uh-huh?

00:01:01.100 --> 00:01:02.760
[Elsie:] Will you play "Beautiful Baby" for me?

APPENDIX C

Sample Costs for Creating Captions

Disclaimer: This section is not an endorsement of any of the following third-party vendors.

Machine-generated vs. Human-generated

Many options exist for creating “text alternatives” for audiovisual content, i.e., captions for transcribed dialogue and textual description of non-audio content, plot, action, etc. Creation of such text alternatives can be undertaken in-house by Smithsonian staff and contractors or outsourced to a third-party vendor. In both cases options exist to employ automated speech-recognition software (in-house, via free software such as that offered by YouTube or via various third-party proprietary software aka “machine-generated”) or to employ human transcription labor (aka “human-generated”).

It is important to note the significant imperfection inherent in machine-generated speech-recognition technologies. Such technologies have yet to reach a level of accuracy that is reliable—both vis-à-vis accuracy of transcribed dialogue and in terms of grammatically correct sentence separations, line breaks, punctuation, capitalization, speaker identification, etc. In fact, even when proofed by humans at a third-party vendor, machine-generated captions are likely to involve significant errors and oversights whenever subject-specialized content is involved. Speaker identification, for example, is a recurring challenge: a computer or a third-party vendor may not correctly identify the speaker in a voiceover by Malcolm X. Subject accuracy remains a recurring challenge: a computer or a third-party vendor may not undertake the additional effort to confirm the spelling of a geographic location, proper noun, or other element critical to a full understanding of the content. Proper grammar persists as a glaring challenge: a computer or third-party vendor may not necessarily follow the regulations of a specific style guide or the basics of elementary grammar (e.g. correct usage of “their” vs. “there”).

Even when these inherent speech-to-text software issues are overcome, failure to properly regulate line breaks and basic separations of transcribed dialogue over multiple lines of text can result in clunky, open-ended, incongruous text alternatives that will negatively impact the captions and audio descriptions when read by screen-reading accessibility software. Such ‘ugly’ and inelegant line breaks in captions also fail to accurately capture the style of speech they are intended to convey.

For these reasons it is recommended that all captions and text alternatives be proofread/proof-watched by a human—preferably by a subject specialist in the relevant content.

Labor-hours vs. Content-hours

When estimating resource allocations for the creation of captions and text alternatives, it is useful to consider the ratio between time required to create (or correct) captions and text alternatives— (“labor-hours”) and the time or duration of the content itself (“content- hours”). Several vendors base their rates on labor hours, others on content hours. Depending on the type of content, this difference can result in a significant cost variation. For example, a dialogue-heavy feature film involving fast-talking characters may require far more labor-hours to transcribe and caption ‘from

scratch' than an oral history with a slow-talking interviewee. For human-generated captioning 'from scratch,' one vendor suggests a labor-hour to content-hour ratio of between 5:1 and 10:1 times the length of video content.¹ Another vendor suggests the ratio to be between 6:1 and 8:1 times the length of video content.² The ratio of labor-hour to content-hour for proofreading/proof-watching machine-generated captions may be significantly smaller, may be equal, or (in the case of ugly and inelegant line breaks) may be more significant.

Sample Third-party Vendors and Costs

There are innumerable third-party captioning and transcription vendors in the marketplace, with a small number specifically offering services designed to comply with web content accessibility targets. As specified in *SD 215*, Smithsonian staff is responsible for planning and budgeting for accessibility protocol such as captions and descriptions, and it is recommended that any digitization or digital video project incorporate accurate cost estimates for these. Below is a chart comparing captioning costs and options from several vendors contracted by Smithsonian units in the past.

¹ Sofia Enamorado and 3 Play Media, "How Long Does It Take to Manually Caption Videos?," (June 3, 2019). See: <https://www.3playmedia.com/%202018/12/20/long-take-manually-caption-videos/> Accessed January 7, 2020.

² Michael Sesling and Audio Transcription Center, email correspondence with Walter Forsberg (October 31, 2017).

Smithsonian Institution - Web Content Accessibility for Digital Audiovisual Assets

Vendor	Website	Service Offered	Accuracy Claim	File Submission/ Retrieval	Cost Basis for Captions (Turnaround time)	Payment Options	Notes
3Play Media	https://www.3playmedia.com/	Machine-generated, human-verified transcriptions, captions, audio description, and text alternatives. Non-English translation and transcription available.	99%	Via portal.	\$2.50/content minute (10 days); \$3.00/per content minute (4 days); \$3.75/ content minute (2 days); \$4.50/content minute (1 day); \$5.50/content minute (8 hours); \$8.50/content minute (2 hours).	Credit card.	Pricing and turnaround dependent on content duration.
Audio Transcription Center	https://audiotranscriptioncenter.com/	Human-generated, human-verified transcriptions, captions, audio description, and text alternatives.	99%	Via email attachment, or Dropbox.	\$30/content hour (\$2.00/ content minute).	Credit card or purchase order.	Does not charge rush fees.
Caption Sync	https://www.automaticsync.com/captionsync/	Human-generated, human-verified transcriptions, captions, audio description, and text alternatives. Non-English translation and transcription available.	No claim.	Via portal.	\$2.45/content minute (4 days); \$2.49/content minute (2 days); \$3.15/ content minute (1 day); \$3.75/content minute (8 hours).	Credit card.	
REV.com	www.rev.com	Machine-generated, human-verified transcriptions, captions, audio description, and text alternatives. Non-English translation and transcription available.	99% (for audio files that are clearly audible).	Via portal.	\$1 per minute of content (1 day).	Credit card.	Pricing and turnaround dependent on content duration.
WGBH - Media Access Group	https://www.wgbh.org/foundation/what-we-do/media-access-group	Human-generated, human-verified transcriptions, captions, audio description, and text alternatives.	No claim.	Via portal.	\$7.00/content minute, with \$70 minimum (unspecified turnaround).	Credit card or purchase order.	