



Making GIFs Accessible

Cole Gleason
cgleason@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania

Amy Pavel
apavel@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania

Himalini Gururaj
hgururaj@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania

Kris M. Kitani
kkitani@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania

Jeffrey P. Bigham
jbigham@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania

Visual animated GIF



Alt Text

Elissa Slater from Big Brother is drinking from a coffee cup, then laughs and spits water everywhere.

Source Audio

[sipping sound]
[spittake sound]
[subtle laughing]
[off screen]: "you guys are talking game"

Audio Description

"Elissa sips water."
[sipping sound]
[spittake sound]
"She spits it out".
[subtle laughing]

Figure 1: The most popular GIF we observed on Twitter was one of a spit-take. We converted GIFs like this into three alternative formats: alternative text, the source audio from the original video, and an audio description recorded over the source audio.

ABSTRACT

Social media platforms feature short animations known as GIFs, but they are inaccessible to people with vision impairments. Unlike static images, GIFs contain action and visual indications of sound, which can be challenging to describe in alternative text descriptions. We examine a large sample of inaccessible GIFs on Twitter to document how they are used and what visual elements they contain. In interviews with 10 blind Twitter users, we discuss what elements of GIF content should be described and their experiences with GIFs online. The participants compared alternative text descriptions with two other alternative audio formats: (i) the original audio from the GIF source video and (ii) a spoken audio description. We recommend that social media platforms automatically include alt text descriptions for popular GIFs (as Twitter has begun to do), and content producers create audio descriptions to ensure everyone has a rich and emotive experience with GIFs online.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility**; **Accessibility systems and tools**.

KEYWORDS

alternative text, GIF, blind, low vision, audio description, social media, image description, accessibility

ACM Reference Format:

Cole Gleason, Amy Pavel, Himalini Gururaj, Kris M. Kitani, and Jeffrey P. Bigham. 2020. Making GIFs Accessible. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*, October 26–28, 2020, Virtual Event, Greece. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3373625.3417027>

1 INTRODUCTION

Social media platforms have become critical broadcast and discussion platforms for conversation online, yet an increase in visual media is making these platforms less accessible to people with vision impairments [13, 27]. Recently, GIFs, silent looping animations, have demonstrated this problem acutely, as they are frequently used and rarely described. GIFs are primarily used on social media to either embody the emotion of the poster or react to another poster's content [34]. If people with vision impairments are unable to understand the visual content of a GIF in a conversation, they miss key channels of emotional tone and information, if not derailing the conversation entirely.

The primary approach to make images accessible is via alternative text [8], which some social networks have recently begun to support for static images [13]. Twitter extended this capability to GIFs on their platform as of January 2020. However, GIFs are more than static images: the visual content over multiple frames often conveys action and contains visual elements that imply sound. Can alternative text adequately describe the emotional tone or meaning that is being visually conveyed? We collected a large sample of popular GIFs on Twitter to examine what kinds of content they contained and how they could be described.

To gather the perspective of blind people on important visual elements to describe, we interviewed 10 Twitter users with vision impairments about their prior experience encountering GIFs. In a second session, they compared three alternative formats for GIFs:



This work is licensed under a Creative Commons Attribution International 4.0 License.

ASSETS '20, October 26–28, 2020, Virtual Event, Greece

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-7103-2/20/10.

<https://doi.org/10.1145/3373625.3417027>

alternative text descriptions, *original source audio* if the GIF was excerpted from a longer video, and spoken *audio descriptions* of action occurring that overlay the source audio. In both interviews, participants stressed that they viewed alternative text as a minimum accessibility requirement. However, depending on both the visual content of the GIF and the original source audio, participants suggested that some audio descriptions presented a more emotive and enjoyable experience of viewing GIFs.

In summary, our contributions are:

- (1) An analysis of GIF usage on Twitter, including how many have alternative text;
- (2) Findings from interviews with 10 Twitter users with vision impairments regarding their past experiences with GIFs; and
- (3) Preferences for accessible alternative formats for GIFs.

In February 2020, few GIFs (0.04%) contained alternative text on Twitter, as the ability to add alternative text to GIFs was new. Therefore, most of our participants had not experienced accessible GIFs on social media, while some participants knew that GIFs were present but undescribed. Based on their experiences with GIFs during our study, many participants were eager to have accessible GIFs on social media – with both alternative text and more expressive audio descriptions.

This work suggests that social media platforms should seek to automatically include alternative text for GIFs on their platforms. In May 2020, Twitter started to include short alternative text descriptions of GIFs taken from their titles on GIF aggregation sites (e.g., GIPHY). They also made it easier for users to add alternative text in general by removing the requirement to enable a special setting, which was noted as holding back alternative text adoption [13]. Based on our investigation of important visual elements in GIFs and discussions with participants, social media platforms should create libraries of descriptive alternative text and automatically include them when users re-use GIFs. Additionally, they should push the accessible experience further by working with content creators to develop rich audio descriptions to convey the emotion in GIFs.

2 RELATED WORK

This research is related to the use of GIFs on social media and existing methods to make images and videos accessible through alternative text and audio descriptions.

2.1 The Use of GIFs on Social Media

The Graphics Interchange Format (GIF, pronounced “jif” [16]) began in 1987 as a format designed to bundle multiple images at a time for later viewing as sequential frames. But the format grew over time with the advent of the World Wide Web and acquired new features: a timed delay between images, transparent backgrounds, and automatic looping of the animation [12]. These features led to widespread use of GIFs on web sites to display animated icons, but the modern emergence of GIFs seen on social media is due to their use on the Tumblr and Reddit platforms.

Tumblr, a microblogging platform, supported GIF uploads from its inception, leading its community to share a significant number of GIFs that were excerpted from TV shows or movies [12]. Fans used these excerpt GIFs to talk about their favorite characters and moments, while spreading these out-of-context actions and

dialogue (overlaid as a visual caption). Others re-used the visual context from the TV show, but added their own text to give the GIF a new meaning [17]. Reddit popularized the “reaction GIF”, which contain actions or gestures (especially facial expressions) that convey an emotional reaction to a scenario. The original creators of these GIFs may have intended to convey a specific meaning, but interpretations may vary based on the separate understandings of the GIF poster and viewer, their prior knowledge of the source material, and their relationship [19]. GIFs that are shared on most social media platforms and text messaging services today resemble those that spawned on Tumblr and Reddit, and they are typically either act as a response someone else’s post, or as a supplement to text posted by the author to embody an action [34].

The initial uses of GIFs on these two social media platforms demonstrate the two core abilities of GIFs: performance of affect and conveyance of cultural knowledge [25]. They are more engaging than other forms of media due to this and their technical constraints [20]. But these constraints limit GIFs as a visual-only medium, which is a disservice to people with vision impairments who will miss out on emotional tone on social media [14] and be unable to share GIFs themselves.

2.2 Alternative Text on Social Media

Alternative text is a method of attaching a textual description of an image such that a person with a vision impairment can read the description with screen reader software or a Braille display. It has been the standard for making images accessible on the web since 1995 [4]. As social media platforms became popular in recent decades, a rise in user-generated content such as images and other media has not been accompanied by an increase in alternative text usage [27]. For images on Twitter, only 0.1% of images on the platform include alternative text [13]. There have been various proposals to address this issue, including adoption of automatic captioning systems [36], although these are often inaccurate and can be misleading to users [23, 32]. The Twitter A11y project utilized several methods, including automatic descriptions, to attach alternative text to every image a user encountered on Twitter, although not all of the alternative text was high quality [15].

This work expands on this interpretation of alternative text, recognizing that it is important to the accessibility of the web and social media specifically, but audio representations of visual content in GIFs could better serve people with vision impairments. Recent work by Gleason et al. has also explored how to best create alternative text for memes on social media, and whether sound effects could be used instead [14]. Morris et al. [26] also suggests that alt text has not changed much since its inception, and it could support richer representations of visual content, provided both content creators and screen reader software developers could agree on what those were. While well-written alt text [22] is an important first step towards accessible GIFs, GIFs are meant to be emotive and rich content that can quickly convey more than text alone. GIFs allow people to embody physical actions or facial expressions enacted over time (compared to a single moment as in an image), and richer representations for alternative formats may be needed to convey that non-visually.

2.3 Audio description

Like GIFs, longer videos often contain visual content expressed over time. Although videos are not silent like GIFs, they often feature visual content that is inaccessible from the audio track alone. Audio descriptions are the primary method for providing viewers information about this content via a narration track overlaid on top of the video [31]. In the past decade since instating the Twenty-First Century Communications and Video Accessibility Act, audio descriptions have become increasingly common on TV and movies [28, 30], especially with the advent of streaming platforms that add audio descriptions to new content such as Netflix. Audio descriptions are challenging to produce because an author must fit all of the necessary visual content into a limited time provided [29, 35], and are most often professionally produced.

However, audio descriptions are exceedingly rare for online user-generated content for reasons including: lack of video author awareness, challenge of crafting descriptions, and a lack of platform support. Prior work proposed methods to make audio description easier to create including using text-to-speech instead of human narration [21], creating task-specific authoring tools [5, 18, 29], offering methods to add audio descriptions on embedded YouTube videos [1, 18], and hosting audio descriptions [18]. Such tools rely on proactive video authors and third party volunteers, and are challenging to scale. We instead consider the space of GIFs, where we can leverage the resources of centralized GIF creation, and the repetition of the medium in order to make them more accessible.

In our consideration of audio descriptions for GIFs, we analyzed several audio description guidelines often written by and in collaboration with blind authors [2, 3, 10]. While such guidelines primarily offer guidance for long stories, we apply key principles (e.g., describe important visual content, avoid overlapping dialog and key sounds, start general then add detail) in the case of providing audio descriptions for the extremely short medium of GIFs.

3 GIFS ON TWITTER

To explore how GIFs are used on Twitter and what types of content they contain, we used the Twitter API to collect a large, random sample of approximately 108 million tweets continuously from February 26 - March 13, 2020, containing 791,600 GIFs (0.7%). This sample was filtered to remove tweets that Twitter automatically tagged as containing possibly sensitive (*i.e.*, pornographic) material, deleted tweets, retweets, and non-English tweets. After filtering, 303,874 GIFs remained, and only 126 of these (0.04%) contained alternative text. However, the ability to add alternative text to GIFs was launched only 1 month prior to our sample collection, so it may not yet have widespread adoption.

In May 2020, Twitter introduced a feature that automatically included short alt text for GIFs if they were taken from GIF aggregation sites. These are the titles of the GIFs present on these sites, and Twitter added them if users shared a GIF and did not include alternative text themselves. For example, the GIF in Figure 1 had the title “Big Brother Elissa Slater GIF”. While this title includes the name of the person in the GIF and the TV show she appeared on, it fails to describe the visual content of the GIF and the spit-take action occurring. When these titles did describe the action, such as “Oprah Shrug GIF” for Figure 3, it did not include

much detail. Twitter also made it easier for users to add alternative text in general by removing the requirement to enable a setting before seeing the interface to add alternative text. In light of these changes, we collected a smaller sample of 31,000 GIFs in June 2020, and found 47.4% included alternative text with these automatic GIF titles. Because they follow a specific format (short titles ending in “GIF”), we estimate that almost all (99.3%) of the GIF alternative text is automatic titles. Excluding those, 0.3% of GIFs have alternative text likely added by the GIF poster. The remaining analyses in this section are not concerned with the alternative text already on Twitter, and therefore are based on the larger GIF sample.

Prior work has noted two common ways to use GIFs: to supplement your own post or to react to another post [34]. We see this behavior in our large sample as well: 23% of the GIFs were included in original posts and 77% were in reply to other tweets. Notably, of those that were original posts, 89% contained additional text, whereas only 33% of reply GIFs accompanied text. This indicates that someone using a screen reader or Braille display may glean some information from the text content of original tweets with GIFs, supposing the GIF is not the central element. Two-thirds of GIF replies would read as completely blank.

3.1 Determining unique GIFs

When online memes use repeated visual elements, it becomes easier to make them accessible as portions of alternative text can be reused between images [14]. We were interested to see if GIFs were often reused, and if so, how many unique GIFs might need to be described. We analyzed the first frame from each GIF to output a perceptual image hash [6]. To verify this method, 10 instances of 25 GIFs were manually examined to ensure they correspond to the same GIF, excluding minor changes due to compression or resolution differences. It is possible that some GIFs could be incorrectly marked as unique if they had significantly different first frames, but the likelihood of this is small as many were shared from aggregator websites and contain the same set of frames. The total number of unique GIFs that were tweeted at least once is 127,916 (42%), and the remaining GIFs were repeated. Several (187) of these GIFs exceeded 100 uses, and the remainder form a long tail of usage distribution (Figure 2). This suggests that accessible formats could be reused for the most popular GIFs.

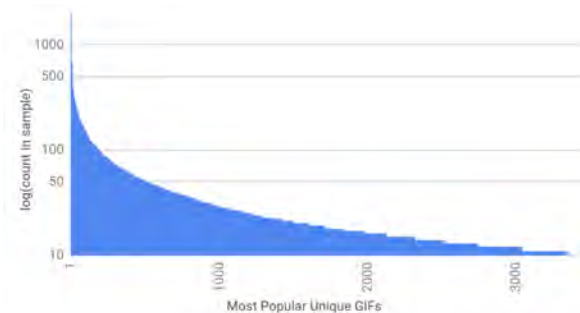


Figure 2: Histogram of the all of the most popular GIFs in our sample (used at least 10 times). The y-axis shows how often each unique GIF was used, on a logarithmic scale.

3.2 Visual Elements of GIFs

We randomly sampled 97 of the most popular 1,000 GIFs to understand the kind of visual content they contained. Popularity was determined by the number of unique times a GIF was shared, not retweeted or liked. To first identify the important visual elements of GIFs, two members of the research team iteratively coded three small, separate random samples of GIFs (30 at a time) to describe them textually and add open codes. They met frequently to discuss their codes, which were based on elements relied on to textually describe the focus of this GIF (e.g., number of characters, text captions, is the character performing an action) and other elements of composition that differed between GIFs (e.g., live-action versus animation, shot length). Once the kinds of visual elements were agreed upon, the researchers then proceeded to code the 97 popular GIFs to describe the frequency of various visual elements, which are reported below:

Original or Excerpt: 87 of the GIFs were excerpted from a longer video, while 10 seemed to be created just to share as a GIF.

Animated or Live-Action: 77 GIFs were live-action content, while 20 were animated. All 10 of the original GIFs mentioned above were animated.

How many characters?: 75 of the GIFs contained only 1 person or character, whereas 14 displayed 2 or more characters, and 8 contained none.

If there is text, is it dialogue?: 14 GIFs contained text, and 7 of these were lines of dialogue from the original source. The others were either overlaid by the GIF author or original GIFs that displayed text only.

Are there visual indications of sound?: 37 GIFs contained some visual indication of sound, with 11 being dialogue, 11 vocalizations that were not speech, and 18 sound effects (e.g., clapping). A GIF could contain more than one indication of sound.

Is the character(s) face important?: 85 of the GIFs had at least one face present, and we identified 58 of them as being important visual context (*i.e.*, the face was the focus).

Is the character performing an action?: 53 GIFs contained the character performing some action or gesture, including clapping, walking, dancing, etc.

Camera Angle Shot: 36 of the GIFs were close-up shots of a person's face, 36 were medium-length shots of someone's torso and head, and 16 were full-body shots of someone from a distance.

This analysis gives us a good understanding of the kind of visual content that might need to be described in a GIF. Most are excerpted from longer, live-action videos and contain characters. About a third contain visual indications of sound, meaning many gestures or actions may be non-verbal. In around 60% of the GIFs, a character's face is the focal point, indicating facial expressions will be critical for understanding GIF content.

4 FORMATIVE INTERVIEWS

This analysis of a large sample of tweets gave us insight into the quantitative nature of GIFs on Twitter, but we desired a qualitative perspective from people with vision impairments to assess the

impact on accessibility. To do this, we conducted a formative study with 10 Twitter users who had a vision impairment. The participants were equally split between men and women, and they averaged 36.2 years old (min = 20, max = 52). Only one participant (P1) used their vision to access content on Twitter, but she often used a screen reader as her level of vision can fluctuate. All participants had used Twitter for at least 5 years, except P1 who used it for 3 years. More detailed demographics are available in Table 1.

In our formative interview, we asked participants about encountering GIFs on Twitter or other social networks, showed them examples of alternative text that we wrote for 10 GIFs, and solicited their feedback on what information to include in accessible GIFs. The interview questions are available in Appendix A.1.

4.1 Prior Experience with GIFs

We asked our participants about their prior experiences encountering GIFs on social media or the web as a whole. Five of the participants stated they commonly encounter GIFs online, and the others either saw them sporadically or not at all. Three participants used the TWBlue client to access Twitter [9] which does not notify the user when they encounter a tweet with a GIF included, so those three participants were not very aware of GIFs. Two participants who frequently encountered GIFs noted that it was typically in replies to other tweets or in comments for posts on Facebook. Five participants stated that when they encounter GIFs, they are not sure if they are missing content that is important to the conversation, while five participants stated they mostly ignore GIFs because they are inaccessible. Four participants had experiences where the use of inaccessible GIFs interrupted conversation, with P1 relating how it interrupted an interpersonal relationship:

About three years ago I was talking to this guy who only reacted in reaction GIFs, and I could never tell what emotion they were feeling about a particular question. [...] I think he assumed that there was a lot more accessibility available for GIFs than there actually was. Because I couldn't see almost anything he was sending me and I ended up just like, 'You know what? We're done. We're not talking.' – P1

The participants stated that they did not usually share GIFs because interfaces to select GIFs on their mobile phones or social network applications did not provide enough information about the GIF to choose one.

In terms of workarounds, 4 participants explicitly stated they used the surrounding textual content, if available, to guess at what a GIF might contain. P1 was the only participant who reported using external software, such as Microsoft Seeing AI [24], to describe GIFs. Four other participants said it was too much work, as the GIF might not be very interesting and they must take a screenshot to extract a single frame from the GIF to get a description. Of course, this is unlikely to fully describe a GIF, as they contain action over multiple frames. P1 recounted this:

My brother [said] "Hey, watch, this garden hose turned into a snake!" So we had to do it frame by frame so I could figure out what was going on. – P1

Three participants said friends would verbally describe GIFs they wanted to share in person, or send text descriptions in online

Table 1: The demographics of the participants who engaged in both online interviews, including age, gender, level of vision, years at the designated level of vision, years using Twitter, other social networks used, and screen reader software used.

ID	Age	Gender	Level of vision	Level of vision years	Years on Twitter	Other social media	Screen readers
P1	23	F	Low-vision	Since age 3	3 years	Instagram, Pinterest, Facebook Messenger	NVDA, VoiceOver, Select to Speak
P2	25	F	Light perception	Since birth	5.5 years	Facebook	VoiceOver, NVDA, JAWS
P3	39	F	Light perception	Since birth	12 years	Facebook	VoiceOver
P4	33	M	Totally blind	Since birth	13 years	None	NVDA, JAWS
P5	41	M	Totally blind	Since age 26	13 years	Facebook, Instagram	NVDA, VoiceOver, Talk-back
P6	54	M	Totally blind	Since age 1	11 years	None	NVDA, Voiceover
P7	46	M	Totally blind	Since birth	13 years	Facebook, LinkedIn	JAWS, NVDA, Narrator, VoiceOver
P8	29	F	Totally blind	Since age 17	7 years	Facebook, LinkedIn	JAWS, NVDA, VoiceOver
P9	20	M	Light perception	Since birth	7 years	Facebook, Youtube	NVDA, VoiceOver
P10	52	F	Totally blind	Since birth	10 years	None	JAWS, NVDA, Narrator

messages, but this was infrequent. Six participants had seen people online describe a GIF posted by someone else at least once, but P5 noted that asking others to describe this content either in person or online has high social barriers:

Oh, you know, I don't want to wear out my welcome. It's socially awkward. But at the same time, I feel like I need some access to that culture. – P5

4.2 Information to include in GIFs

To elicit feedback on what information to include in GIFs, we prepared alt text for 10 GIFs and read each to the participants during the formative interview. The GIFs were selected by sampling 100 random GIFs and manually choosing 10 that roughly spanned the visual elements presented in Section 3.2. This formed a diverse sample to elicit discussion about important information.

After each GIF, we asked what elements of the alt text participants thought were important and which they they might remove. We attempted to include a lot of information in the alternative text descriptions, so that participants were aware of the majority of the visual elements. The alt text and GIFs are available as supplemental material.

All participants noted that the most important elements of the GIF descriptions were: the people or characters present and what they actions they are taking. If there was not a definite character or person in the GIF, then the focus should be described. All participants wanted to know what text said, if it was present. When text is present, care should be taken to distinguish if it is dialogue from the GIF source video or not. One GIF was a clip from Saturday Night Live with unrelated text overlaid, similar to an image macro meme [11], and participants were unsure if the text was dialogue from the SNL skit.

If a GIF was taken from a movie or TV show, participants wanted to know information about the character, actor/actress, and the work they appeared in. There was some disagreement between participants about which of these three was important to include. Three participants thought the character was most important as the action or dialogue might be more closely linked with the character.

Others mentioned that different actors can play the same character (as in a GIF for The Batman), and sighted people viewing GIFs may recognize the actor or actress even if they never saw the film or show. P8 suggested:

So you've got 'Princess Diaries', 'Princess Mia', and 'Anne Hathaway', right? Having two out of those three I think is probably good. – P8

Participants wanted most of the information to be present, but also alternative text to be concise. When alternative text mentioned the clothing of the character or person in the GIF, most participants were not personally interested but were reluctant to suggest removal in case others may be interested. Some participants already knew pieces of information in GIFs (e.g., P2 and P8 were aware Michael Jordan played for the Chicago Bulls), but thought others might benefit from it. The only information that the majority of participants felt comfortable suggesting to remove was information about overall GIF coloring such as "It is very dark and red" or redundant information that appeared elsewhere in the description. In one case, the alternative text described Michael Jordan performing a "reverse one-handed dunk" and included a more lengthy description of the same action. Participants wanted one or the other to make the GIF more concise. Four participants stated that length was not their primary concern, and that the description needs to be proportional to the amount of action occurring:

It's long, unfortunately. I know you want to keep these brief, but I think sometimes for sake of being complete, it just takes as long as it takes. – P10

4.3 Stated preferences for accessible formats

Both before and after hearing example alternative text for GIFs, we asked participants about their thoughts on how to make GIFs accessible. Before hearing the alt text descriptions, almost all participants suggested that GIFs be accompanied by alternative text on sites like Twitter. Specifically, participants wanted Twitter to make it easier to add alternative text to GIFs on mobile devices and make users more aware of alternative text. P5 suggested that alternative text be turned on by default for everyone, something that has been



Figure 3: A popular reaction GIF of Oprah Winfrey shrugging. She turns to look to the camera, glances to the side, stares at the camera, then shrugs with her palms up.

suggested in prior work [13]. Three participants wondered if GIFs could be automatically captioned as they were used to from applications like Microsoft Seeing AI. P1 wanted human-authored descriptions to be added to all of the GIFs that Twitter and others offer in their GIF libraries:

Just put alt text in across all the GIF libraries, because I feel like other users aren't going to take the time to know what alt text is or how to write it. – P1

After experiencing the alt text descriptions for 10 GIFs and recognizing that many were extracted from other videos, seven participants brought up audio formats as another possibility. Two participants suggested that the source audio by itself would not have enough context, but five suggested that audio descriptions could be recorded or extracted from the original video if it was described. However, all participants were confident they still wanted alternative text for GIFs as a minimum accessibility requirement. Alt text is quicker and less disruptive as it can be read in the screen reader's voice and speed. It is also universally accessible to people who browse social media with a Braille display. P10 said she sometimes struggles to hear audio descriptions over background noise and music. So these participants noted that they would like to have source audio and preferably audio descriptions if available, but that alternative text always needs to be there to fall back on situationally or for more context.

5 PERCEPTIONS OF ACCESSIBLE ALTERNATIVES

Based on the formative interviews with participants, we developed some sample accessible alternatives for GIFs and asked participants to examine them in a second 30-minute session as a means to understand their perceptions of these alternative formats.

5.1 Materials

We determined that there were three likely formats for alternative representation of GIFs that could be more inclusive: alternative text, original source audio, and audio descriptions. Alternative text drew on existing best practices for describing images online and audio descriptions were based on best practices for accessible movies or TV shows. We also experimented with only the source audio, bringing in the audio context from the original source material if the GIF was excerpted from other media.

From our prior sample of 97 popular GIFs, we chose a representative 15 (Figure 4) that covered different aspects of their composition (e.g., facial expressions, action, source material). 13 were excerpted from longer videos, and two contained dialogue with text. One had

additional text overlaid, and another was just a GIF of text. The chosen GIFs, alternative text, and audio files for the below alternative formats are all included as supplemental material.

5.1.1 Alternative Text. Alternative text was a natural choice for an accessible alternative format for GIFs, as it is the existing standard for making images accessible online, and GIFs on the web and social media may already include alternative text (although this is uncommon on most social media sites). Most of our participants would prefer alternative text descriptions to make GIFs accessible as a minimum requirement, and expect sites like Twitter to support their inclusion. We composed alternative text descriptions for all 15 of the popular GIFs we selected. Based on prior conversations with participants, we ensured the GIF described the person or characters, actions occurring, and setting of the GIF (if important). If the GIF was from a known television property (which many were), we varied which descriptions included the character's first name, last name, and TV show name, as a way to provoke more discussion on the topic. Our composed alternative text averaged 15.9 words (min = 10, max = 20). An example for Figure 3 is "Oprah Winfrey turns to look straight at the camera, shifts her eyes sideways and then back to center, then shrugs."

5.1.2 Source Audio. For GIFs that are excerpted from TV shows, we hypothesized that some GIFs could be accessible with the source audio alone, as if a video clip had been shared instead of the GIF. To evaluate this, we found the original source audio for as many of the 15 GIFs as possible. Two of the GIFs were not excerpted from a video, and we were unable to find the source audio for another three GIFs, as they did not contain enough identifying information or the video had since been deleted. We trimmed the recovered audio for the remaining 10 GIFs to be representative of the visual content. However, some of the source audio has additional dialogue that was not in the original visual GIF. Our source audio files were on average 5.0 seconds long (min = 2.0, max = 9.7). An example for Figure 3 is audio of someone talking off screen, saying "I always look back at that and say, you know, when I feel like I'm hungry"

5.1.3 Audio Description. Finally, our conversations with participants revealed that audio descriptions might be a viable way to make GIFs accessible, as it is a common method to describe longer videos. GIFs, as a sequence of frames, are a format somewhat in between a static image and a video. Therefore, as audio descriptions often describe action and accompany sound, we developed short audio descriptions for each GIF with source audio. One drawback with audio descriptions is that there is often very little space to add the description audio between music, sound effects, and other

dialogue in the original video. We did not attempt to ensure that the entirety of the alternative text fit into the audio description, and instead focused on brevity and conveying the most important information according to audio description guidelines [7]. We also sometimes extended the amount of source audio to allow the audio descriptions to fit, but we were careful to ensure this did not give additional context that was outside the scope of the original visual GIF. Our audio descriptions for the 10 GIFs with source audio averaged 7.9 words (min = 3.0, max = 16.0) and 5.4 seconds (min = 2.0, max = 9.7). An example for Figure 3 is a narration track over the original audio with the script “Oprah looks at us, to the side, and back at us, shrugging with her palms up.”

5.2 Procedure

All of our participants from the formative interview returned for a second 30-minute session in which they listened to the alternative formats for the 15 GIFs. Participants were engaged over an online voice call using Zoom, and they were compensated \$20 via an Amazon or Paypal gift card.

Because of discussions in the formative interview about how alternative text was a critical minimum requirement for accessibility, all participants heard the formats in the order of: alt text, source audio, and audio description. After hearing all available formats for a GIF example, a member of the research team asked the following questions:

- (1) How would you (or someone else) use that GIF on social media?
- (2) (If multiple formats:) Which format did you prefer and why?

The first question ensured the participant felt confident in the meaning of the GIF, and that the understood meaning from the accessible alternative was similar to the meaning interpreted visually. The second question on format preference elicited whether a particular format excelled or failed for a specific GIF, as the content in the GIF or source audio affected which format participants preferred. After listening to all examples, participants answered questions (listed in Appendix A.2) about their overall format preferences.

5.3 Study Scope and Limitations

The purpose of the second session was for the participants to experience the source audio and audio description formats alongside the format they heard in the formative interview (alt text). This would help them compare the formats and provide qualitative feedback about the preferred format and included information.

Our participants highlighted in the formative interviews that alternative text was critical, so we chose to explicitly highlight the comparison in the second session as a preference, not a mutually exclusive choice. Because of this, we did not randomize the ordering of the formats, as someone listening to the formats would likely always hear alternative text first. Therefore, we do not make statistical claims about the stated preferences of the participants. As participants only heard 15 GIFs, it is possible that participants might develop different preferences with exposure to more GIFs of different content.

As the same 10 participants were present in both the formative interviews and evaluation of alternative representations, our findings cannot represent all users with vision impairments. Longer

term evaluations with larger cohorts may be necessary to solidify or confirm these results.

5.4 Findings

Six participants were confident in the meaning of all but 1-2 of the 15 total GIFs, and their descriptions were similar to a visual interpretation of the same GIF. Three participants reported that they were unsure how to use at least 3 of the GIFs, often the GIFs with the least context present in the source audio or unclear visual expressions. P4 was unsure how to use 8 of the 15 GIFs or what they meant. Further numbers are reported by GIF in Table 2.

The GIFs that presented the most confusion sometimes had sound that could be interpreted multiple ways or was hard to discern, such as the spit-take clip from Big Brother (Figure 1). In the source audio for this clip, another contestant starts talking right after the on-screen Elissa Slater performs a spit-take. Participants were confused who was speaking, and if the spit-take was meant to imply laughing or indignation. A GIF of Oprah shrugging (Figure 3) was confusing to participants because the action was entirely visual, yet another woman is speaking in the source audio, leading to additional context that is not important to the visual GIF.

Subtle character actions proved difficult to describe. A GIF of the character Stringer Bell from the show The Wire involves subtle facial expressions like a “side eye”. Participants were not sure what this gesture implied. P8 suggested that nonverbal gestures that are well-known may require the description author to editorialize more to describe the meaning.

Table 2: Participant understanding and format preference for each GIF. From left to right the columns are: GIF source (Figure 4), the number of participants who understood that GIF, the number who preferred Alt Text (AT), Source Audio (SA), and Audio Descriptions (AD). Note: * P5 and P9 always responded that they would prefer to use alt text in combination with other formats. Their preference for alt text is represented by the (+X) notation in this column, and they are also counted among the other format they preferred.

GIF Source	Understood	AT*	SA	AD
Spongebob	9	4 (+2)	1	5
Big Bang Theory	9	5 (+2)	2	3
Judge Judy	10	N/A	N/A	N/A
The Office - No!	8	1 (+2)	4	5
Brooklyn 99	10	1 (+2)	3	6
The Wire	5	9 (+1)	1	0
Utah Jazz	5	N/A	N/A	N/A
Big Brother	6	4 (+2)	2	4
Full House	9	4 (+2)	1	5
Original GIF (Text)	9	N/A	N/A	N/A
Obama’s Address	10	4 (+2)	3	3
Ryan Gosling	8	N/A	N/A	N/A
The Office - Party	10	2 (+2)	1	7
Original GIF (Cats)	10	N/A	N/A	N/A
Oprah’s Next Chapter	5	4 (+2)	3	3



Figure 4: The first frame of all 15 GIFs we used in our second session. Their source is annotated below each GIF.

5.4.1 Format Preference. Overall, six of the 10 participants stated they preferred the audio description format as the best way to experience GIFs, with the caveat that most participants expected alternative text to be present as a fallback option if the audio description was hard to understand or they were not able to listen to audio files at the moment. Three participants preferred to use alt text, and P9 preferred to use a combination of the alternative text and source audio to understand the GIF content.

Source audio by itself was viewed as the most inaccessible format, as it often did not describe the action in the scene or was too noisy to pick apart distinct sounds in the clip due to background music, dialogue, or laugh tracks. For example, the audio for a GIF of character Stringer Bell from *The Wire* had a mostly silent audio track, as he sits in silence while the GIF focuses on his expression. Eight of the 10 participants stated source audio was their least favorite format, while P1 and P9 disliked audio descriptions:

I don't like the audio descriptions because at that point I would have already looked at the alt text to know what was going to happen. So I would be more paying attention to the [source] audio. – P9

5.4.2 Seeking out GIF conversations. All participants said they were unlikely to specifically seek out conversations that contained accessible GIFs, but most would be more engaged when they encountered them their existing social media accounts. P8 noted:

One of the biggest bummers is if I'm reading through social media and [...] the post is accessible, and then I'm reading the comments and it'll say like, "comment with a GIF". I'm like, "Damn, that really sucks". – P8

While we focused the conversation on large social media networks, P2 and P3 both mentioned they would engage more with content posted on their workplace communication platforms (*i.e.*, Microsoft Teams and Slack) as GIFs are common there. P5 wondered if means of making GIFs accessible could be extended to short videos on Instagram or TikTok, as they were interested in trying

out those platforms that remain mostly inaccessible non-visually. P8 echoed this about TikTok more negatively:

That whole app is not even accessible. I've given up on trying new social medias. – P8

6 DISCUSSION

In both sessions, our participants made it clear that alternative text must always be available for GIFs on social media. Alt text is what people are familiar with on the web, it works well with screen reader software, and can be customized to be read in a preferred voice or speed. It does not vary in volume, and can be skimmed quickly, as well as being universally accessible to a user with a Braille display. The kinds of visual information present in GIFs that is needed to write alternative text is similar to that of images, although a user must also describe action occurring over time. While a still image might be described as "Oprah shrugs", the GIF in Figure 3 may include several distinct actions to describe such as "Oprah turning to look at the camera, shrugging with her palms up in the air, giving a sly smile, and turning back to the speaker". Participants reported a tension between an "objective" account of visual action versus a shorter description that gives a subjective interpretation like "Oprah shrugs as if to say 'I told you so'". They discussed that the description length should be proportional to the amount of action occurring, and this tension is more clear when listening to audio descriptions, where space is very limited.

Once alt text is present for GIFs on social networks, the majority of participants were interested in additional modalities to describe GIFs, as the audio descriptions or (in 2 cases) source audio can give a more rich, emotive experience. Just as sighted people utilize GIFs to embody actions or expressions in supplement to text, people with vision impairments should have that option with audio GIFs. A caveat here is that the original GIF author may not have considered the audio content when designing the visual GIF. Thus, the source audio could be useless (with purely visual actions like a shrug) or

be discordant with the visual meaning (e.g., Big Brother contestant talking over the spit-take). Not all GIFs may benefit from the inclusion of source audio alone or with an audio description, but those that are centered on dialogue or vocalizations would benefit from these additional formats. Additionally, our participants disagreed about the information that should be included in the audio descriptions, as they all heard the longer alternative text and knew what was excluded from the briefer audio description. For now, we would recommend audio description best practices to decide what information to include, but future research could explore modular audio descriptions that allow people with vision impairments to choose what information is most important to them. As alternative text should always be present and contain all of the information, this decision is not as important as it is for longer media where audio descriptions are the only accessible format.

When talking about making images accessible on social media, research largely focuses on automatic solutions to scale the problem [23, 32, 36] or human-written descriptions [13]. Automatic approaches can scale human-written descriptions for viral memes that change, as long as the visual content remains the same [14]. Like image memes, GIFs are often used repeatedly online, so information about the origin of a GIF may help convey meaning. Current efforts to document a meme's origin and spread on sites like Know Your Meme rarely describe the visual content, as they assume a sighted audience. Future work may investigate integrating this information along descriptions of the visual content to better convey their thematic meaning.

Recent GIFs seem less likely to be modified and remixed compared to memes, as many excerpted from TV shows are produced and distributed by the television networks [33]. If TV production and network companies are producing this content, they could make it accessible before distributing it to GIF aggregation website or smartphone keyboard applications. In fact, content produced for broadcast TV may already have produced audio descriptions that are sufficient for the excerpted GIFs, depending on the script. For user-generated GIFs that are not made accessible by their creator, third-party volunteers or crowd workers could generate alt text or audio description templates similar to the proposed solution for memes.

This work has primarily focused on the consumption of GIFs on social media posts, but half of participants said they would like to share GIFs if accessible formats were available. The addition of alternative text or audio descriptions to GIF search engines would aid people with vision impairments in selecting the perfect GIF. Further research may need to explore accessible tooling to assist people with vision impairments in the creation of new GIFs, such as excerpting video clips. We focused on GIFs specifically, as these were common on social networks like Twitter, Facebook, or Reddit. But one participant mentioned they would like to see an extension of this work to short videos, such as those popularized on Vine or Tik Tok. As those videos prominently feature audio, audio descriptions seem like a promising solution, but may need additional tooling to support creation by all social media users.

7 CONCLUSION

GIFs are a common and expressive way to display emotional reactions or embody physical actions on social media. In the words of P10, they are “supporting actors” for posts, but as a visual medium, they are inaccessible to people with vision impairments. In this work, we examined just how prevalent GIFs are, and how many were made accessible by Twitter posters (0.04-0.3%). In formative interviews with blind participants, we discussed prior accessibility issues with GIFs online, leading to the development of three accessible alternative formats. This led to a second interview with probes of the accessible alternative formats. Our participants stressed the importance of alt text as a minimum requirement, but also enjoyed the expressiveness of audio descriptions when it fit the GIF well. We recommend that platforms continue efforts to include alternative text with GIFs on their site, and consider formats such as audio descriptions for the most popular GIFs.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. (#DGE-1745016). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] [n.d.]. 3PlayMedia. <https://www.3playmedia.com/>
- [2] [n.d.]. American Council of the Blind, Audio Description Project, Guidelines for Audio Describers. <https://www.acb.org/adp/guidelines.html>.
- [3] N. Reivers A. Remael and G. Vercauteren. [n.d.]. Pictures painted in Words: ADLab Audio Description Guidelines. <https://dcmp.org/learn/captioningkey/624>.
- [4] Tim Berners-Lee and Dan Connolly. 1995. *Hypertext Markup Language - 2.0*. RFC 1866. Internet Engineering Task Force. 1–77 pages. <https://doi.org/10.17487/RFC1866>
- [5] Carmen J Branje and Deborah I Fels. 2012. Livedescribe: can amateur describers create high-quality audio description? *Journal of Visual Impairment & Blindness* 106, 3 (2012), 154–165.
- [6] Johannes Buchner. 2020. A Python Perceptual Image Hashing Module. <https://github.com/JohannesBuchner/imagehash> [Online; accessed 6-May-2020].
- [7] ADI AD Guidelines Committee. 2003. Guidelines for Audio Description. <https://www.acb.org/adp/guidelines.html> [Online; accessed 6-May-2020].
- [8] W3 Consortium. 2018. Web Content Accessibility Guidelines (WCAG) 2.1. <https://www.w3.org/TR/WCAG21/>
- [9] Manuel Cortez. 2019. TWBlue. <https://twblue.es/> [Online; accessed 6-May-2020].
- [10] DCMF. [n.d.]. Description Key. <https://dcmp.org/learn/captioningkey/624>.
- [11] Marta Dynel. 2016. “I has seen Image Macros!” Advice Animals memes as visual-verbal jokes. *International Journal of Communication* 10 (2016), 29.
- [12] Jason Eppink. 2014. A brief history of the GIF (so far). *Journal of Visual Culture* 13, 3 (2014), 298–306. <https://doi.org/10.1177/1470412914553365>
- [13] Cole Gleason, Patrick Carrington, Cameron Cassidy, Meredith Ringel Morris, Kris M. Kitani, and Jeffrey P. Bigham. 2019. “It’s Almost Like They’re Trying to Hide It”: How User-Provided Image Descriptions Have Failed to Make Twitter Accessible. In *The World Wide Web Conference* (San Francisco, CA, USA) (WWW ’19). ACM, New York, NY, USA, 549–559. <https://doi.org/10.1145/3308558.3313605>
- [14] Cole Gleason, Amy Pavel, Xingyu Liu, Patrick Carrington, Lydia B. Chilton, and Jeffrey P. Bigham. 2019. Making Memes Accessible. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) (ASSETS ’19). Association for Computing Machinery, New York, NY, USA, 367–376. <https://doi.org/10.1145/3308561.3353792>
- [15] Cole Gleason, Amy Pavel, Emma McNamee, Christina Low, Patrick Carrington, Kris M. Kitani, and Jeffrey P. Bigham. 2020. Twitter A11y: A Browser Extension to Make Twitter Images Accessible. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI ’20). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376728>
- [16] Doug Gross. 2013. It’s settled! Creator tells us how to pronounce ‘GIF’. <https://www.cnn.com/2013/05/22/tech/web/pronounce-gif/index.html> [Online; accessed

- 21-July-2020].
- [17] Ödül Akyapi Gürsimsek. 2016. Animated GIFs as vernacular graphic design: producing Tumblr blogs. *Visual Communication* 15, 3 (2016), 329–349. <https://doi.org/10.1177/1470357216645481>
 - [18] The Smith-Kettlewell Eye Research Institute. [n.d.]. YouDescribe. <https://youdescribe.org/>
 - [19] Jialun "Aaron" Jiang, Casey Fiesler, and Jed R Brubaker. 2019. "The Perfect One": Understanding Communication Practices and Challenges with Animated GIFs. *arXiv* 2, CSCW (2019), 1–20. <https://doi.org/10.1145/3274349> arXiv:1903.09623
 - [20] Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, Juan Pablo Hourcade, Saeideh Bakhshi, David A Shamma, Lyndon Kennedy, Yale Song, Paloma de Juan, and Joseph 'Jofish' Kaye. 2016. Fast, Cheap, and Good: Why Animated GIFs Engage Us. (2016), 575–586. <https://doi.org/10.1145/2858036.2858532>
 - [21] Masatomo Kobayashi, Trisha O'Connell, Bryan Gould, Hironobu Takagi, and Chieko Asakawa. 2010. Are synthesized video descriptions acceptable?. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. 163–170.
 - [22] Veronica Lewis. 2019. How to Write Alt Text for GIFs (on Veronica With Four Eyes). <https://veroniiica.com/2020/02/03/how-to-write-alt-text-for-gifs/> [Online; accessed 6-May-2020].
 - [23] Haley MacLeod, Cynthia L. Bennett, Meredith Ringel Morris, and Edward Cutrell. 2017. Understanding Blind People's Experiences with Computer-Generated Captions of Social Media Images. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 5988–5999. <https://doi.org/10.1145/3025453.3025814>
 - [24] Microsoft. 2017. Seeing AI Talking camera app for those with a visual impairment. <https://www.microsoft.com/en-us/seeing-ai/>
 - [25] Kate M Miltner and Tim Highfield. 2017. Never Gonna GIF You Up: Analyzing the Cultural Significance of the Animated GIF. *Social Media + Society* 3, 3 (2017), 205630511772522. <https://doi.org/10.1177/2056305117725223>
 - [26] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. 2018. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). ACM, New York, NY, USA, Article 59, 11 pages. <https://doi.org/10.1145/3173574.3173633>
 - [27] Meredith Ringel Morris, Annuska Zolyomi, Catherine Yao, Sina Bahram, Jeffrey P. Bigham, and Shaun K. Kane. 2016. "With most of it being pictures now, I rarely use it". In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. 5506–5516. <https://doi.org/10.1145/2858036.2858116>
 - [28] Jaclyn Packer, Katie Vizenor, and Joshua A Miele. 2015. An overview of video description: history, benefits, and guidelines. *Journal of Visual Impairment & Blindness* 109, 2 (2015), 83–93.
 - [29] Amy Pavel, Gabriel Reyes, and Jeffrey P. Bigham. 2020. Rescribe: Authoring and Automatically Editing Audio Descriptions. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software & Technology (UIST 2020)*.
 - [30] Audio Description Project. [n.d.]. Master AD List. <https://acb.org/adp/masterad.html>
 - [31] Audio Description Project. [n.d.]. What is Audio Description? <https://acb.org/adp/ad.html>
 - [32] Abigale Stangl, Meredith Ringel Morris, and Danna Gurari. 2020. "Person, Shoes, Tree. Is the Person Naked?" What People with Vision Impairments Want in Image Descriptions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376404>
 - [33] Louis Staples. 2019. Are Memes Ruining Television? <https://www.gq.com/story/are-memes-ruining-television> [Online; accessed 6-May-2020].
 - [34] Jackson Tolins and Patrawat Samermit. 2016. GIFs as Embodied Enactments in Text-Mediated Conversation. *Research on Language and Social Interaction* 49, 2 (2016), 75–91. <https://doi.org/10.1080/08351813.2016.1164391>
 - [35] Kelly Warren. [n.d.]. A day in the life of an audio describer. <https://dcmp.org/learn/277-a-day-in-the-life-of-an-audio-describer>. [Online; accessed 6-May-2020].
 - [36] Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic Alt-text: Computer-generated Image Descriptions for Blind Users on a Social Network Service. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (*CSCW '17*). ACM, New York, NY, USA, 1180–1192. <https://doi.org/10.1145/2998181.2998364>

A APPENDIX: INTERVIEW QUESTIONS

A.1 Session 1

- (1) Collection of demographic information.
- (2) How often do you encounter GIFs? In what contexts?
- (3) Do you remember the last time you encountered a GIF? What cues do you use to interpret it?

- (4) Have you encountered GIFs elsewhere on the web? Are they accessible there?
- (5) Have you encountered GIFs where people add informal alt text (in the original post or in the comments)?
- (6) Has not being able to access the visual content of a GIF prevented you from understanding something in the past?
- (7) Have you had any experience of someone helping you access a GIF? What was the context?
- (8) What would you do to make GIFs accessible?

A.2 Session 2

- (1) Which format did you most prefer? Why?
- (2) Which format did you least prefer? Why?
- (3) Given a tool that could provide all three formats for popular GIFs, which do you think you would enable at least some of the time? Why?
- (4) Given accessible alternatives for popular GIFs, do you imagine you would seek out more conversations that contain GIFs? If so, where would you look?