“Honestly I Never Really Thought About Adding a Description”: Why Highly Engaged Tweets Are Inaccessible

Mallak Alkhathlan, M. L. Tlachac, Lane Harrison, Elke Rundensteiner

“Honestly I Never Really Thought About Adding a Description”: Why Highly Engaged Tweets are Inaccessible

Mallak Alkhathlan[0000−0003−1933−3502], ML Tlachac[0000−0002−6634−678X], Lane Harrison[0000−0003−3029−2799], and Elke Rundensteiner[0000−0001−5375−9254]

Worcester Polytechnic Institute, Worcester MA 01609, USA
{malkhathlan,mltlachac,ltharrison,rundenst}@wpi.edu

Abstract. Alternative (alt) text is vital for visually impaired users to consume digital images with screen readers. When these image descriptions are not incorporated, these users encounter accessibility challenges. In this study, we explore the prevalence and user understanding of alt text in Twitter. First, we assess the availability of alt text by collecting the Twitter Engagement (TWEN) dataset which contains over 1000 high engagement tweets regarding online articles from the most popular Google Keywords. We focused on keywords that create an engagement in Twitter in order to study the possibility of creating priorities of media content that missing alt text then adding descriptions to them by crowdsourcer to help the visually impaired to be equal like others in the social media communities. Our findings reveal approximately 91% of the tweets contained images and videos, less than 1% of the images had alt text. Thus, even highly engaged tweets remain inaccessible to visually impaired individuals. Thus, we designed two guided concepts to raise awareness of high engagement. We then surveyed 100 sighted participants to understand their perception of alt text and evaluate strategies to increase the frequency of alt text for highly engaged content. Our value-based guided concept was well received by the majority of the study participants.

Keywords: Image captions · Image description · Alt text · Accessibility · Visual impairment · Twitter

1 Introduction

Alternative (alt) text, commonly referred to as descriptions, are tags used to describe images, GIFs, diagrams, and illustrations to a non-textual content [50]. Digital images are prevalent, existing on websites, platforms, applications, e-publications, and software. The alt text tag plays a vital role in helping visually impaired users to understand the content. Nowadays assistive technology such as a screen-reader (e.g., Apple Voice-over, Microsoft Narrator, TalkBack) helps people who are visually impaired to render screen reader elements into a speech.

Recent changes in the world including the ease of use and access to mobile phones with cameras have contributed to the heavy dissemination of images on the internet. In 2014, 1.8 billion digital images were uploaded through social media platforms WhatsApp, Facebook, Instagram, Snapchat, and Flickr [33]. This number is only increasing as social media becomes more popular as evidenced by WhatsApp users alone uploading 4.5 billion digital images in 2017 [38].

In 1999, HTML 4.0 specified standards to meet the needs of visually impaired persons. Now authors need to specify an image description in the "alt" field attribute of the image's HTML tag so screen readers are able to access the image [31,40,47]. Yet, image descriptions are lacking on social media. This results in visually impaired users being unable to access images and GIFs, leading to inequality access among Twitter, Facebook, and Instagram social media platforms [36,16,5,64]. Twitter remains one of the more popular social media platforms. In February 2019, Twitter reported 126 million daily users, an increase of 9% from the prior year [46]. Twitter's simple text-based interface makes it more accessible to screen readers than other more popular social media platforms such as Facebook [8]. As such, Twitter is especially popular with visually impaired users. Specifically, a survey of blind Twitter users revealed 88.4% used Twitter for news and 72% used Twitter for entertainment [36].

Unfortunately, images on Twitter are largely inaccessible to the visually impaired. In 2019 [16], researchers highlighted the barrier of missing descriptions. Findings revealed only 0.1% of tweets with images contained image descriptions [16]. According to the World Health Organization, in 2019, the number of visually impaired in the world was approximately 2.2 billion [65]. Twitter media content (i.e. image, GIF) is one of the most important components of the social network. Users interact on Twitter through liking, retweeting, and/or replying to a tweet. This creates more intellectual communication between members of the Twitter community. Twitter is a way to socialize among the visually impaired, though many studies have shown that the visually impaired face difficult challenges to overcome the amount of media contained in Twitter [36,32,67].

As the visually impaired depend on screen readers to consume tweets, media content is inaccessible if it lacks a description. While there is a feature for adding descriptions, it is optional. Despite the improvements to this feature from 2011 to 2020, many sighted users still neglect to add descriptions to media content. Thus, in our work, we propose using content engagement to measure the priority of adding descriptions to multimedia content. This would make crowdsourcing solutions more effective as descriptions would be added to the content with the highest engagement before content with lower engagement. We designed visual icons to notify users that content has high engagement and needs descriptions.

In this study, we first determine the frequency of descriptions in the media content of tweets about a diverse range of highly engaged topics. Then we survey sighed users to understand how to increase description compliance and the perception of our visual icons on tweets with multimedia content. As far as we know, there is no recent tweet data set available to the public that includes highly engaged keywords that causes missing alt text. Our contributions include:

1. Identifying the type of Twitter content that users engage with the most to build a popular data set,
2. Determining the frequency of media content and descriptions in this set of highly engaged tweets, and
3. Surveying sighted users to understand how to increase the description frequency of highly engaged media content on Twitter through crowd-sourcing.

2 Related Work

Our study urges that the community should share the same social media applications interface with the visually impaired users. We believe that the community is able to increase the percentage of visual content descriptions if the way is paved through the visual notification in the application interface like what propose in our research. Our study builds on previous work by Gleason et al. [16] which focuses on images from the most popular Twitter accounts. Unlike Gleason et al., our research includes other forms of media content such as animated GIFs, videos, and URL previews. Additionally, while Gleason et al. focused on tweets from the most popular twitter accounts, we focus on the tweets containing the most popular keywords. This allows us to capture a tweets about highly engaged topics from a diverse range of Twitter users. In first phase of our study, we hypothesize even highly engaged media content are likely to lack text descriptions on Twitter. Our research is further unique in that it covers the difficulties in finding media content with descriptions on Twitter as well as the important role crowdsourcing could play in providing descriptions for visually impaired users.

2.1 Visually Impaired Users and Twitter

Twitter was founded in 2006 and originally provided a rather simplistic text-based interface thus making it popular for people who are blind [8], but the social media platform started featuring embedded photos in tweets on June 2011 [12]. In 2016 [52] Twitter added a description option on Twitter which needed to be adjusted through the settings. Even users who activated this setting to describe an image did not always provide a description. Furthermore, in May 2020 Twitter modified this feature and users can add the description directly without setting [59]. However, this important feature is not well advertised so many users may be unaware of its existence or the value of including descriptions.

The number of monthly active users on Twitter reached 340 million in 2020 [26]. Tweet engagement is measured through the quantity of likes, retweets, and replies. Engagement and trust are used by influencer marketing to predict the popularity of products [1,61], but many promotions contain visual content without a description [57]. Politics are also frequently discussed on Twitter, in particular during elections [25], and the lack of descriptions impedes the ability of visually impaired users to participate [54]. Regrettably, media content descriptions in tweets are very rare. A study [16] found only 0.1% of 9.22 million tweets with images contained descriptions. Despite the lack of descriptions, Twitter remains popular among visually impaired people who use Twitter to get updates on the latest news, entertainment, and social relations [36]. In addition to images, Twitter also supports videos and GIFs to increase user participation. These media content types allow for expression that static images are unable to capture. However, GIFs are silent animations which make them even more difficult for visually impaired users to consume, creating additional accessibility issues [2]. A survey including 3.7 million accounts found that Tweets with GIFs received 55% greater engagement than those without GIFs [60], demonstrating GIFs are an important tool to increase tweet engagement. Without descriptions, these highly engaged tweets are inaccessible to visually impaired users.

2.2 Methods to Generate Alternative Text

Image-based tweets are highly varied and therefore existing automated image description tools [36,44] are not easily adapted to work on Twitter. In this section we will discuss three strategies for generating content media descriptions.

Human-powered approaches. Human-powered approaches specifically refer to human-in-the-loop labeling provided via crowd-powered systems. Researchers [6] have proposed 13 human-powered access technology design principles inspired by both historical and recent technical advances that have potential to make environments more available for people with disabilities. VizWiz [4] is a free iPhone application that allows visually impaired people take pictures, ask questions, and then receive answers from the Amazon Mechanical Turk (mTurk) crowd-sourcing platform in less than 30 seconds. RegionSpeak [69] is a system that allows blind people to send, receive, and collect more information with crowd-sourcing by providing an easy way to integrate visual details from several images by image stitching which significantly decrease the number of encounters and complete time spent on finding answers. There are also other studies about how to help add image descriptions that rely on human annotators [8,7,42,63].

Automated approaches. Researchers developed and deployed Automatic Alt-Text (AAT) that applies the computer vision technology to identify images and generates alt text on Facebook [68]. There are other approach to automatically generate image descriptions but these are also not intended for visually impaired individuals [14,30,62,53]. Other research focuses on learning how to recognize objects and relationships. This is generally accomplished with trained computer vision models which input an image to generate a related caption [13,28,41].

Hybrid approaches. Caption Crawler is a system that proposes a reverse image search to retrieve existing descriptions on the internet for static images [23]. TweetTalk generates human editing descriptions or tags to save time and financial costs for the recruitment of human crowd workers. This is done through the development and assessment of constructed social media image questions that can instruct humans or AI systems to the invention of captions that contain certain types of details most desired by the visually impaired [44]. Research has also leveraged social media posts to explore the value in alternative text formats and how crowdsourcing could answer new questions about popular imagery [45]. Other approaches often consist of combining generated descriptions or tags with human editing [5,43]. Twitter A11y [17] is browser extension to add alt text on to image tweets accessible, however such a system should not be relied upon as an alternative, as it will likely either misidentify objects or not provide an appropriate level of detail in descriptions, especially for scientific purposes [11].

2.3 UI/UX design awareness on social media

Recently there has been an explosion of data quantity and data sources. For instance, facets of social lives are being increasingly disseminated on social media, essentially allowing for social lives to be transformed into quantifiable information. This increase in data has caused User Interface (UI) and User Experience (UX) design to become increasingly important. In several different contexts, the importance of UX has been studied, especially in the low rate of participatory design awareness. These studies have been used to inform the development of persuasive technologies within the human computer interaction (HCI) community [15,20,37,19,29,66]. The two main ways to portray data is numerically and visually [27,21,39]. Further, research recognizes people historically respond better to evidence when numerical values are involved [39]. Despite enormous research on the accessibility of the Web and on developing standards [22], frameworks and legal standards, awareness of designers [49], and tools to make textbooks accessible [10], access to certain content remains hindered. Unfortunately, creative, dynamic, and adaptable UI/UX design methods that simultaneously benefit both blind users and sighted user have been largely ignored in academic literature. Thus, we explore the use of UX to convey missing alt text using two concepts: (1) numerical guided concept and (2) photo icon guided concept.

3 Twitter Engagement Experiment and Dataset

3.1 Data Collection

In the first phase of our research, we collected the Twitter Engagement (TWEN) data set which consists of data regarding three types of content media on Twitter: image, video, and text. Specifically, this data set is unique as it focuses on Twitter engagement with online articles, thus resulting in a high percent of images and videos. In order to get articles on a variety of popular topics, we choose topics based on the top 100 keywords searched on Google as of June 2018 [48].We have chosen this list of keywords because they might lead us to find the important tweets that people may interact with often then add description to later, In addition this list is not associated with a brand and has removed porn-related keywords. For each of keyword (such as weather, maps, news, donald trump,

#	keywords	URL	Total Twitter Engagement (top 3 URL)	Top 3 URL Tweets (15 tweets)														
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Weather	5000	26385	I	I	I	I	I	I	V	I	I	I	V	V	V	V	V
2	Maps	5000	25151	I	I	V	I	V	V	I	I	I	I	I	I	I	I	I
3	Translate	5000	6354	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
4	Calculator	5000	4184	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
5	Youtube to mp3	116	161	I	I	I	I	I	N	N	N	N	N	I	I	I	I	I
6	speed test	289	1043	N	I	I	I	I	I	I	I	V	V	I	I	B	B	B
7	news	5000	498762	N	V	V	V	V	T	T	T	T	T	I	I	I	I	I
8	thesaurus	375	179	I	T	T	T	I	I	I	B	B	B	I	T	T	T	T
9	poweball	1240	2462	I	T	T	T	B	N	N	N	N	I	I	I	I	I	I
10	donald trump	5000	250554	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I

Fig. 1. Tabular representation of the TWEN data set which indicates the media content of the tweets for the top ten keywords. The tweets are labeled as containing a thumbnail (N), large image (I), video (V), and just text (T). Broken (B) tweets are also noted.

game of thrones, etc..), we use Twitter's premium Tweet search API [55,56], to identify the top three articles with the most Twitter engagement by utilizing premium operators to deliver article engagement based on filtering rules where a queries length are the 100 keywords. We defined the article engagement as the number of times tweets mentioning the article were retweeted, replied, and liked. These articles were from June 22, 2019 through April 4, 2020.

For each article, we collect data about the top five tweets with the highest engagement that reference each selected article's URL, resulting in at most 1500 tweets. We removed ten keywords that contains only one letter such as 'g' and 'f' to avoid matches an exact phrase within the body of a Tweet which may return unexpected results (for example 'f' means Facebook). Of the 1350 extracted tweets, 262 were broken due to the tweets being deleted by the user who posted them or blocked on Twitter. Thus, TWEN is comprised of 1088 tweets. While each tweet has a 280 character limit, Twitter Cards allow for the inclusion of images and videos. Our data set consists of four types of cards. First, a Summary Card includes a thumbnail image of the referenced website, description, and text. Second, a Summary Card with Large Image is very similar to a Summary card except the image is larger than a thumbnail. Third, a Player Card allows for the inclusion of audio, videos, and GIFs. Additionally, our data set contains tweets with only text which we refer to as a Text Card. A tabular view TWEN content is displayed in Figure 1 for the top ten keywords. As seen, there are fifteen tweets, five for each of the articles. We filtered the tweets using operator 'url:' (this refers only to tweets that have URLs that point to media hosted elsewhere). We used 'has:links' and 'has:media' operators to control the API request from the stream. We collect at most 5000 URLs within a tweet for each of the top 100 keywords. Twitter engagement is calculated for the top three article's URLs.

3.2 Findings

Figure 2 contains a summary of TWEN content. Of the 1088 tweets, 902 contain images, 92 contain video, and 94 contain only text. Over 91.4% of the tweets

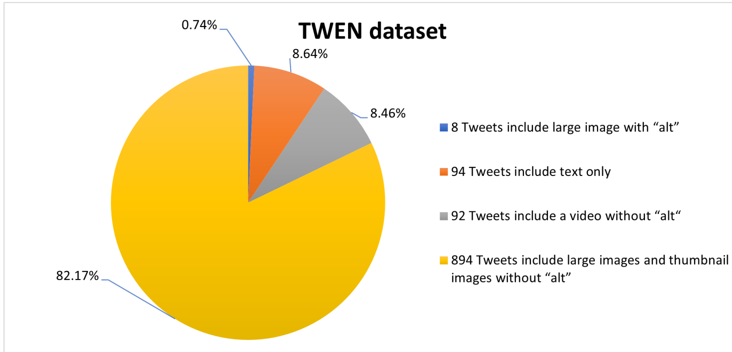

Fig. 2. TWEN shows highly engaged tweets rarely have alt text.

contained content media, demonstrating the prevalence of images and videos in highly engaged tweets about highly engaged articles. In particular, tweets with large images were highly prevalent in TWEN. While videos were less common, Player Cards which support videos were only introduced to Twitter in 2017. None of the videos or thumbnails contained descriptions and only one percent of the 752 large images contained descriptions for visually impaired users. While this alt text can be added to non-textual content on Twitter, it is an optional card property field which is clearly rarely utilized for media content [58].

We also compare the content media of the tweets from each of the article URLs, as seen in Table 1. We collected the five most engaged tweets from each of the three most engaged articles for each keyword. Recall, engagement is the summation of the number of likes, retweets, and replays. Tweets referencing the first URL yielded 279 Large Images, which is higher than the second or third URLs. As with the full data set, tweets with Large Images were the most common media content for all URLs. The most engaged (first) URLs have the highest number of tweets with large images and videos. Our threshold of 100,000 engagements serves to prioritize highly-engaged tweets so that they can reach high description compliance. Based on the data, displayed in Figure 1, we found that the 10 keywords with the highest number of engagements each had more than 100,000 engagements. For example, the keyword 'movies' ranks first with 652,715 engagements. We set our threshold as 100,000 engagements, though this algorithmic threshold could be easily readjusted based on a larger dataset.

TWEN reveals that highly engaged tweets contain plentiful visual content with few descriptions. Large images were the most common type of media content in the tweets. Given the rich visual details in these images, it is most important for them to contain descriptions. Yet, only 8 of the 752 large images contained descriptions. Unfortunately, this poses a major obstacle to Twitter's visually impaired users as they are unable to consume the majority of the most engaged tweets on Twitter. Thus, we consider strategies to help solve the problem of missing descriptions for tweets with high engagement. We propose a publicly visible

Table 1. Distribution of the 1088 TWEN tweets with the article URLs.

Media Content	First URL	Second URL	Third URL
Large Image	279	239	234
Thumbnail	42	60	48
Video	39	26	27
Text Only	32	32	30
Total	392	357	339

icon to notify users to the presence of a highly engaged tweet that needs a description. As there were 10 keywords in TWEN with the engagements of over 100 thousand, we set the threshold value for the icon at 100 thousand engagements. We believe if we are unable reach complete compliance, our threshold serves to prioritize the most highly engaged tweets. The threshold can be readjusted as needed. We design a pseudocode in Algorithm 1 which is an informal high-level description of the operating principle to highlight how we add a flag to tweets that have highly engaged media content. We believe this icon will increase description compliance for the most important media content, thus making highly engaged content more accessible for visibly impaired users on Twitter. In the next experiment, we test the effectiveness of this proposed strategy.

3.3 Contributions and Availability

The TWEN dataset contributes information about the variety and types of imagery being shared on Twitter. Upon publication, we will make the entire tabular representation of TWEN, a subset of which is shown in Figure 1, publicly available at https://osf.io/ksyzx/?view_only=a68dd37bf108424d8006e2d7071e3bdd. Our approach for identifying highly engaged tweets can help inform future research and the design of enhanced automated approaches for captioning.

Algorithm 1 Strategy to identify highly engaged tweets that need a concept flag to encourage descriptions.

```
if the tweet has a media content then
    if media content has a description then
        Don't do anything
    else if the engagement of the tweet ≥ 100,000 then
        Add concept flag to the tweet
    else
        Don't do anything "tweet is not highly engaged"
    end if
else
    Don't do anything "tweet has no media content"
end if
```

4 UX Accessibility Design Experiment

4.1 Motivation and Research Questions

Building from the results of TWEN, we conducted an experiment to explore themes of encouragement and awareness. The purpose of this experiment is to help understand how extensive the problem is that causes missing alt text in media content. This is done by assisting Twitter users in describing tweets that need more attention, i.e., have very large engagement. We also explored strategies to encourage additions of descriptions to media content. Responses were collected with a survey distributed on Twitter. This experiment was designed to answer three main research questions:

- RQ1: To what extent are our findings from the TWEN data set consistent with surveyed participant experience?
- RQ2: Is there a difference in effectiveness between engagement flags using a numeric value versus a photo icon?
- RQ3: What are guidelines for improving the frequency of media content descriptions on social media?

4.2 Recruitment and Respondents

To ensure applicable results, we purchased an advertisement on Twitter to recruit Twitter users to take our IRB-approved survey. Specifically, we targeted users with location in the United States and English as their language. We sought responses from people of all genders. The recruited participants should be at least 18 years old, have a Twitter account, and participate voluntarily. We used the Qualtrics survey platform and received 101 valid responses. 36.6% identified as female and 63.4% as male. Participants achieved varied highest levels of education: 3% had less than a high school degree, 15.7% had high school diplomas or equivalents, 12.9% had some college but no degree, 13.9% had Associate's degrees, 41.6% had Bachelor's degrees, 8.9% Master's degree, 1% had Doctoral degrees, and 3% had professional degrees. Users identified varied reasons for using Twitter: news and updates (75.5%), following celebrities and brands (32.7%), entertainment (63.4%), social engagement (48.5%), finding a job (1%), and (5.9%) indicated business (note that some respondents select more than one choice in this question). Further self-reported demographics for the 101 participants are in Table 2.

4.3 Experimental Procedure

There were eight parts to the survey instrument:

1. **Demographic questions:** We presented participants with multiple-choice questions about their Twitter use frequency, age, gender, education, main goals of using Twitter, and length of time they have been using Twitter.

Table 2. Distribution of participant age and frequency/distribution of Twitter use.

Age	%	Frequency of use	%	Duration of use	%
18-24	2%	Multiple times a day	45.5%	< 1 year	2%
25-34	39.6%	Once a day	23.8%	$1-3$ years	18.8%
35-44	35.6%	Multiple times per week	22.7	$3-5$ years	14.8%
45+	22.8%	Monthly	7.9%	> 5 years	64.4%

2. **User experience with Twitter:** We first asked respondents to select the type of media they share most on Twitter. Next, we ask them to rank on a Likert scale the frequency they have *written a description when they tweet a media content (i.e. images, GIFs).* Those who included a description were asked about the source of their descriptions with options including *bot*, *their own words*, and write-in. Lastly, we presented an screenshot to participants and asked them if *they think image in the bounding box has a description?*
3. **High priority of the engagement:** In the prior experiment, we found only 1% of the highly engaged tweets with media content contained descriptions. As such, this section of questions is to determine if the level of engagement influences users willingness to act as a crowdsourcer by adding a description. We first provide a definition of engagement (like + retweet + replay). We then ask participants to select what media content they would *choose to write a description for* with options varying the level of engagement.
4. **Concept effectiveness:** Respondents were asked to rate on Likert scales the *effectiveness of the numerical value concept* and the *effectiveness the photo guided concept* in comparison to the default concept used by Twitter.
5. **Concept Interest:** Respondents were asked to rate our numerical value and photo guided concepts as well as the Twitter default concept on Likert scales. Options ranged from being *interested* to *disinterested*.
6. **Overall awareness**: We asked respondents to indicate if they are aware of the recent Twitter announcement on May 27, 2020 explaining how descriptions can be added without having enter settings. Participants were also asked to rate on Likert scale their willingness to turn a *screen reader on to see whether or not the image has a description.* We educate the participants on how a screen reader works and why it is used in case they are not aware of this type of technology. As twitter currently hides media descriptions, we also asked participants to rate on Likert scale if they agree with the statement: *Twitter can improve on how it displays invisible descriptions.*
7. **Overall preference**: We asked respondents to select if they preferred the Twitter default, numerical value, or photo guided concept.
8. **Request for further information**: Participants who selected at least *sometimes* regarding their description frequency in part two were automatically prompted to answer the open-ended question: *Why did they add a description to the media content (i.e. images, GIFs)?* Those who selected *Never* were prompted to answer the open-ended question: *Why did they not add a description when they tweeted media content (i.e. images, GIFs)?*

Fig. 3. Numerical value guided concept: the screenshot featuring the numerical value guided concept.

Fig. 4. Photo guided concept: the screenshot featuring the photo guided concept.

Fig. 5. Twitter default concept: displays the screenshot with Twitter's default concept.

During the survey, participants were asked questions related to our design concepts which we refer to as *guided concepts*. The related questions regarded a news feed on the Twitter interface. Participants compared our numerical value guided concept and photo guided concept with the default concept currently used by Twitter. For this study, we chose a tweet with $likes = 9.9k$, $retweets = 2.7k$ and $replies = 49$. As seen in Figures 3-5, the tweet also contains both text and an image missing a description. Our hypothesis is that many users will fail to distinguish between text at the top of the tweet and the description.

For all three concepts, we first asked participants to familiarize themselves with the concept. The screenshot featuring the numerical value guided concept is displayed in Figure 3. The tweet includes a small icon on the top right-hand side of the image to indicate very high engagement. The screenshot featuring the photo guided concept is displayed in Figure 4. The tweet includes a depiction of an eye on the top right-hand side of the image to indicate very high engagement. Lastly, Figure 5 displays the screenshot with Twitter's default concept.

4.4 Analysis

We use qualitative and quantitative analysis in our survey. We focus on three independent variables: 1) *numerical value guided concept*, 2) *photo guided concept*, and 3) *default Twitter concept*. Currently, Twitter hides the descriptions.

We also focus on a variety of different dependent variables. During the second part of the survey, we collected responses to four questions regarding user experience with Twitter. Participants could respond to *"I often tweet <type of content>* with content types 1 = *text*, 2 = *image*, 3 = *GIF/video*, and 4 = *poll*. A Likert scale ranging from 1 = *always* to 5 = *never* was used to gauge *"if they have ever written a description when they tweet a media content"*. Participants could respond to *"what type of descriptions have been used to write a description"* with 1 = *bot*, 2 = *my own words*, and 3 = *other*. For the question regarding description existence, participants could respond to *"<type of answer> the image has a description"* with answer types 1 = *I think*, 2 = *I don't think*, 3 = *I think there is no way to know if*, or 4 = *other*. We will refer to these questions about **user experience with Twitter** as UE1, UE2, UE3, and UE4.

In part three, responses included *very highly engaged*, *very low engaged*, *both*, and a write-in option. We will refer to the **high priority of the engagement** as HP1. In part four we collected three Likert scale ratings on the effectiveness of the numerical value, photo guided, and Twitter default concepts. We refer to these five-point scales regarding **concept effectiveness** as EC1, EC2, and EC3. Likewise, in part five, we collect three Likert scale ratings regarding interest in the numerical value, photo guided, and Twitter default concepts. We refer to these five-point scales regarding **concept interest** as ICO1, ICO2, and ICO3.

In part six, we collected one quantitative reply and two Likert scales. Options regarding awareness of Twitter's announcement about descriptions were *aware* and *not aware*. Participants rated their likelihood of screen reader use and the potential Twitter improvement on five-point scales. We refer to these **overall awareness** replies as OA1, OA2, and OA3. In part seven, participants selected

Table 3. WN: type of participants responses about why not to add a description.

Category	Description
Time consuming	statement explaining how adding a description takes time and effort
Type of follower	statement about personal followers (i.e. friends/ family) who are not vision impaired
Misunderstood the necessity of alt text	statement about how media content can speak for itself without the need to add more explanation
Not expecting blind or low vision (BLV) audience on Twitter	statement expressing surprise about the existence of users who are blind and might follow them
How to add alt text?	statement about the location of the option to add description and how it is used

Table 4. WA: the type of participants responses about why to add a description.

Category	Description
Accessible	statement about how adding descriptions would enable blind users to understand concept of tweets
Post humors	statement regarding how adding descriptions often makes media content more fun
Easy	statement about how easy it is to write a description
Deep meaning	statement about the amount of details when delivering a vague image
More info	statement about how adding descriptions would give clarifications
Out of topic	statement that is not related to the description

whether they preferred the *default Twitter concept*, *numerical value guided concept*, or *photo guided concept*. We refer to **overall preference** as OP1.

In last part of the survey, we collected open ended answers regarding motivation for adding and not adding descriptions. We performed qualitative analysis by conducting an inductive and deductive reasoning method that performed axial coding [51]. We used inductive logic to identify the reasons for missing image description and deductive reasoning to attribute missing descriptions to a particular concept which allows us to apply it to the theory. We then leverage a semantic analysis technique [18] to identify and code text segments according to the parent codes. According to Braun and Clark [9], when searching for themes one should "not [be] looking for anything beyond what a participant has said or written". Thus, our semantic analysis involves dividing the data into subsets and assigning a unique child code to a parent theme code among all participants' responses. For motivating **description exclusion**, child codes include the statements in Table 3, referred to as WN. For motivating **description inclusion**, child codes include the statements in Table 4, referred to as WA.

4.5 Findings

User Experience with Twitter. *UE1:* 75.2% of participants reported that the majority of their tweets contained only text. 18.8%, 5.0%, and 1.0% reported they most tweeted images, videos, and polls, respectively. *UE2:* 43.6% of participants reported never writing a description while only 1.0% reported always writing descriptions. The remaining participants either wrote descriptions sometimes (31.7%) or half of the time (13.9%). Thus, most participants fail to add a description to most of their media content, even when it is a known option. *UE3:* 98.0% of participants reported using *their own words* when including a description. 2.0% of participants reported using *other* when including a description. None reported using a *bot*. *UE4:* When shown the image, 53.5% of participants reported they thought there was a description while 39.6% of participants reported they thought there was no description. Only 6.9% of participants correctly indicated that there was no way for them to know. This confirms that many people do not know that the description is not visible on the tweet. Currently, the only ways to view descriptions are screen readers or browsers that allow viewing of the HTML source code. We also found a statistically significant relationship between among female and male participants. 71.4% of female participants reported that there was no way to know if there is a description, a view shared by only 28.6% of male participants. To determine the statistical significance, we used Chi-Squared test with p-value = 0.01 and effect size (Cramér's V) = 0.280.

High Priority of the Engagement. *HP1:* About half (53.5%) of participants reported they would add descriptions to the content of highly engaged tweets. A third (32.7%) reported they would want to add descriptions to both tweets with high and low engagement. 1% reported other and the remaining 12.9% would add descriptions to only lowly engaged tweets.

Concept Effectiveness. *EC1-EC3*, respondents rated the perceived effectiveness of the numerical guided *(EC1)*, photo guided *(EC2)*, and Twitter default *(EC3)* concepts. The five-point Likert scale responses were normally distributed. The concept effectiveness for the numerical guided *EC1* (M=3.25, SD=1.09) was rated higher than the photo guided *EC2* (M=2.95, SD=1.15): t(100)=2.243, p = 0.027. The effect size using Cohen's d to indicate a standardized difference between two means is 0.264, suggesting a small effect. The concept effectiveness was also rated higher for numerical guided *EC1* (M=3.25, SD=1.09) compared to Twitter default *EC3* (M=2.75, SD=1.10): t(100)=2.6, p = 0.026. Cohen's d for this comparison is 0.451, suggesting a medium effect. We do not refer to these differences as statistically significant, however, as a Bonferonni correction for two comparisons would yield an adjusted threshold of $p < 0.05/2$ or 0.025. We instead refer to these results as weak but inconclusive evidence that the numerical guided concept *(EC1)* is more effective than the photo guided concept *(EC2)* and the default concept *(EC3)*. The results suggest that future experiments in this area may benefit from larger sample sizes or expanded Likert ranges (e.g. a 7- or 9-point scale) to provide more opportunity for participants to relatively order the alternatives.

Concept Interest. *ICO1-ICO3*, respondents rated both the numerical value guided *(ICO1)* than the photo value guided *(ICO2)* and the Twitter default *(ICO3)* concepts. The five-point Likert scale responses were normally distributed. The numerical guided *ICO1* (M=3.31, SD=1.18) was rated higher than photo value guided *ICO2* (M=3.02, SD=1.11): t(100)=2.26, p = 0.025. The effect sizes using Cohen's d to indicate a standardized difference between two means is 0.25, suggesting a small effect. The numerical guided *ICO1* (M=3.31, SD=1.18) was also rated higher than the Twitter default *ICO3* (M=2.69, SD=1.18): t(100)=2.61, p = 0.01. The effect sizes using Cohen's d to indicate a standardized difference between two means is 0.52, suggesting a medium effect. Applying the same Bonferroni correction, the latter result favoring numerical guided *ICO1* over the Twitter default *ICO3* would be considered strong statistical evidence. There was little difference in the rated interest between *ICO2* and *ICO3*.

Table 5. WN: category distribution of participants who are not adding descriptions.

Category	**%**	**P**	**Example quote from participant** *P*
Time consuming	9.1%	P99	*"To tweet with the added image description and text takes extra time."*
Type of followers	18.2%	P77	*"None of the people who follow me are visually impaired, so I don't think it's necessary to put in the effort."*
Misunderstood the necessity of alt text	20.5%	P37	*"The ones that i use are specific, they generally speak for themselves when i do tweet media so there is no need to add a description."*
Not expecting BLV audience on Twitter	38.6%	P29	*"Honestly I never really thought about adding a description for those that have vision issues."*
How to add alt text?	13.6%	P33	*"I'd have no idea how to do it. I do not know how to expand on that, sorry. This will have to be enough."*

Table 6. WA: category distribution of participants who are adding descriptions.

Category	**%**	**P**	**Example quote from participant** *P*
Accessible	10.53%	P80	*"I would want them to have the same info I do."*
Post humors	7.0%	P64	*"help to convey the message, usually a parody."*
Easy	1.8%	P12	*"Because its quick to do and easy."*
Deep meaning	19.3%	P73	*"I think that it is necessary in some cases to give some depth or understanding to the image."*
More info	40.4%	P97	*"I think I would do it if people did not understand the context or they can't otherwise read it."*
Out of topic	21.5%	P55	*"I always add GIFS."*

Overall Awareness. *OA1:* Only 16.8% of participants were aware of Twitter's announcement about descriptions. *OA2:* 6.9% of participants reported that they would definitely turn on the screen reader and 20.8% reported that they probably would. 23.8% of participants reported they would definitely not turn on the screen reader and 25.7% reported that they probably would not. The remaining 22.8% were neutral. Almost half of respondents were unwilling to turn on a screen reader while less than a third of participants were willing to turn on a screen reader. As such, We conclude that effective strategies must make descriptions visible to the public. *OA3:* 28.7% and 43.6% of participants strongly and somewhat agreed respectivelythat Twitter can improve how it displays descriptions. 24.8% were neutral, 3.0% somewhat disagreed, and none strongly disagreed.

Overall Preference. *OP1:* Only 15.8% of participants reported that they preferred Twitter's default concept. 54.5% preferred the numerical value guided concept and 29.7% preferred the photo guided concept. This confirms the participants wanted to know if they can help others by writing a description.

Request for Further Information. *WN:* We found participants who are not adding descriptions can be distributed into five categories, as displayed in Table 5. *WA:* We found participants who are adding descriptions can be distributed into six categories as displayed in Table 6.

5 Discussion

One purpose of this study is to understand how to best integrate the BLV population in social media. Making information accessible to all is an important concept in justice and equality [3]. Currently, fairness [24] and artificial intelligence [34,35] research do not meet the needs of the visually impaired community. Artificial intelligence is not yet capable of accurately describing images. Further, there is a mistrust in image descriptions with prior research [23] indicating visually impaired users desire multiple descriptions to confirm accuracy. Our study reveals new design opportunities for improving UX accessibility technologies by enabling crowdsourcing to involved in adding description. In this study, we found that informing users about the need for descriptions will be useful in urging society to invent multiple ways to adding accurate descriptions to media content on social media. For instance, displaying a guided concept on tweets that have received high engagements and need a description will enable willing members of the community to add descriptions of the image. This could have a significant positive effect to the BLV community by providing multiple descriptions for the most pertinent visual content on social media.

5.1 Implication of TWEN

Through our collection of TWEN, we demonstrate a technique to identify highly engaged tweets across a wide variety of popular topics. This technique could be leveraged in the future and the resulting engagement count inform the threshold for adding a guided concept to highly engaged images missing descriptions. The TWEN data set revealed the need for such a guided concept as less than

1% of the images contained descriptions. The 8 descriptions in TWEN were for the keywords “trump” and “mail”. The evidence about keywords corresponding to high engagement is valuable. This knowledge could be used to identify which keywords are correlated with high circulation, thus assisting developers to identify content that may need crowdsourced annotators to provide descriptions. Further, we suggest that Twitter should allow users to add description to the media content when the link's website has Twitter Cards enabled.

5.2 UX Accessibility Design Experiment Research Questions

To answer RQ1, our findings from the second experiment are consistent with our findings from the first experiment. Only 8 of the 902 images in TWEN contained a description. This is not surprising given how many of the surveyed participants did not know what constituted a description. Only 6.9% of participants knew there was no way to know if the image shown to them contained a description (in UE4). Additionally, only 1% of participants always added a description to their posted media content (in UE2). Our participants were mostly avid Twitter users with 45.5% using Twitter more than once a day and 64.6% of them having been a Twitter user for more than 5 years. Yet, only 16.8% knew of the change announced by Twitter regarding the ability to add a description without the need to modify the settings (in OA1).

To answer RQ2, participants preferred engagement flags that used numerical values. 72.3% indicated Twitter's display of descriptions could be improved (in OA3). We presented two alternatives to Twitter's default concept: a numerical value guided concept and a photo guided concept. All three are displayed in Figures 3-5. Participants found the numerical value guided concept more effective (in EC1) and more interesting (in ICO1) than the photo guided (in EC2 and ICO2) and default (in EC3 and ICO3) concepts.

To answer RQ3, we found most participants were amendable to writing descriptions despite only 1% of them always doing so (in UE2). The lack of descriptions seems to stem partially from lack of awareness about descriptions. For instance, only 6.9% of participants knew they could not determine the existence of a description when viewing a tweet containing an image (in UE4). This is confirmed by over a third of participants having replies to why they did not include descriptions (WN) that fell into *misunderstanding* and *How we add alt text?* categories. Another 38.6% of participants were *not expecting BLV to be an audience on Twitter*. Thus it is important to educate social media users on description purpose, how to add descriptions, and the prevalence of visually impaired users. Further, promoting awareness of high engagement may increase willingness to add descriptions. While only a third of participants indicated willingness to always add descriptions, a majority indicated wiliness to add descriptions only to highly engaged media content (in HP1). The 27.3% of open-ended responses (WN) in the *type of followers* and *time consuming* categories also support the need for targeted and effortless solutions. Thus, we must recommend strategies such as our well-received numeric value guided concept for highlighting highly engaged media content without a description.

5.3 Limitations and Future Work

In this study we demonstrated that icons can be used to indicate highly-engaged tweets and encourage users to add descriptions. As there is no international standard icon to indicate BLV in social media platforms, future work involves designing a multicultural icon that could be understood by many people around the world and exploring the pros and cons of different icons. Future icon experiments may include differing formats, shapes, sizes, placements, and colors. However, the icon design is beyond the scope of this paper. Our proposed guided concepts need to be further tested to determine effectiveness, therefore our future work will include more tests to determine effectiveness of guided concepts. As the participants in our study were already are primed to know what the context and meaning of the engagement number, future work involves testing the icons on different populations who have not been asked prior questions about image descriptions and engagement. As it is outside the scope of this research, future research could also involve solutions to the logistical concerns of using crowdsourcing to add descriptions. Notably, there are decorative images that may not require descriptions. Further, some complex images may require domain knowledge to provide a description, so a solution would ideally allow for the people who contribute to writing descriptions specify topics of interest and expertise.

6 Conclusion

In this study we collected the TWEN data set about highly engaged tweets and a surveyed sighted users regarding their experience with descriptions. Over 90% of TWEN included images, suggesting highly engaged tweets frequently have images. However, less than 1% of these images contained descriptions, posing accessibility issues for visually impaired users. The subsequent survey revealed the importance of awareness in increasing description frequency as many participants were unaware of the purpose of descriptions. Despite this, many participants seemed willing to help by providing descriptions for highly engaged content if properly alerted of the need, such as through our well received numerical value guided concept. Thus, if guided concepts are used to draw attention and a method is provided to standardize description writing efforts, adding descriptions to highly engaged content could become more common. Our findings offer a tangible guide regarding what type of media content is available on Twitter and a strategy to increase description frequency.

Acknowledgment

We thank Imam Abdulrahman Bin Faisal University (IAU) and Saudi Arabian Cultural Mission to the USA (SACM) for financially supporting Mallak Alkhathlan. We thank the US Department of Education P200A180088: GAANN grant and the WPI Data Science Department for financially supporting ML Tlachac. We thank Brittany Lewis and the WPI DAISY lab for their advice and support.

References

1. Arora, A., Bansal, S., Kandpal, C., Aswani, R., Dwivedi, Y.: Measuring social media influencer index-insights from facebook, twitter and instagram. Journal of Retailing and Consumer Services **49**, 86–101 (2019)
2. Aslam, S.: Twitter by the numbers: Stats, demographics fun facts (2020), https://www.omnicoreagency.com/twitter-statistics/
3. Bennett, C.L., Keyes, O.: What is the point of fairness? disability, ai and the complexity of justice. arXiv preprint arXiv:1908.01024 (2019)
4. Bigham, J.P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R.C., Miller, R., Tatarowicz, A., White, B., White, S., et al.: Vizwiz: nearly real-time answers to visual questions. In: Proceedings of the 23nd annual ACM symposium on User interface software and technology. pp. 333–342 (2010)
5. Bigham, J.P., Kaminsky, R.S., Ladner, R.E., Danielsson, O.M., Hempton, G.L.: Webinsight: making web images accessible. In: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility. pp. 181–188 (2006)
6. Bigham, J.P., Ladner, R.E., Borodin, Y.: The design of human-powered access technology. In: The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility. pp. 3–10 (2011)
7. Brady, E., Morris, M.R., Bigham, J.P.: Gauging receptiveness to social microvolunteering. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 1055–1064 (2015)
8. Brady, E.L., Zhong, Y., Morris, M.R., Bigham, J.P.: Investigating the appropriateness of social network question asking as a resource for blind users. In: Proceedings of the 2013 conference on Computer supported cooperative work. pp. 1225–1236 (2013)
9. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative research in psychology **3**(2), 77–101 (2006)
10. Center, D.: Image sorting tool (2014), http://diagramcenter.org/decision-tree.html/
11. Chiarella, D., Yarbrough, J., Jackson, C.A.L.: Using alt text to make science twitter more accessible for people with visual impairments. Nature Communications **11**(1), 1–3 (2020)
12. Dorsey, J.: search+photos (2011), https://blog.twitter.com/en_us/a/2011/searchphotos.html
13. Elzer, S., Schwartz, E., Carberry, S., Chester, D., Demir, S., Wu, P.: A browser extension for providing visually impaired users access to the content of bar charts on the web. In: WEBIST (2). pp. 59–66. Citeseer (2007)
14. Fang, H., Gupta, S., Iandola, F., Srivastava, R.K., Deng, L., Dollar, P., Gao, J., He, X., Mitchell, M., Platt, J.C., Lawrence Zitnick, C., Zweig, G.: From captions to visual concepts and back. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2015)
15. Georges, V., Courtemanche, F., Senecal, S., Baccino, T., Fredette, M., Leger, P.M.: Ux heatmaps: mapping user experience on visual interfaces. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 4850–4860 (2016)
16. Gleason, C., Carrington, P., Cassidy, C., Morris, M.R., Kitani, K.M., Bigham, J.P.: "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. pp. 549–559 (2019)

17. Gleason, C., Pavel, A., McCamey, E., Low, C., Carrington, P., Kitani, K.M., Bigham, J.P.: Twitter a11y: A browser extension to make twitter images accessible. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2020)
18. Goddard, C.: Semantic analysis: A practical introduction. Oxford University Press (2011)
19. Gray, C.M., Kou, Y.: Ux practitioners' engagement with intermediate-level knowledge. In: Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems. pp. 13–17 (2017)
20. Gray, C.M., Kou, Y., Battles, B., Hoggatt, J., Toombs, A.L.: The dark (patterns) side of ux design. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–14 (2018)
21. Grosser, B.: What do metrics want? how quantification prescribes social interaction on facebook. Computational Culture (4) (2014)
22. Guidelines, W.C.A.: How to meet wcag (quick reference) (2019), https://www.w3.org/WAI/WCAG21/quickref/
23. Guinness, D., Cutrell, E., Morris, M.R.: Caption crawler: Enabling reusable alternative text descriptions using reverse image search. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–11 (2018)
24. Guo, A., Kamar, E., Vaughan, J.W., Wallach, H., Morris, M.R.: Toward fairness in ai for people with disabilities sbg@ a research roadmap. ACM SIGACCESS Accessibility and Computing (125), 1–1 (2020)
25. Julian Ausserhofer, A.M.: National politics on twitter: Structures and topics of a networked public sphere. Information, Communication Society **16**(3), 291–314 (2013). https://doi.org/https://doi.org/10.1080/1369118X.2012.756050
26. Kemp, S.: Digital 2020: 3.8 billion people use social media (2020), https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media
27. Kennedy, H., Hill, R.L.: The feeling of numbers: Emotions in everyday engagements with data and their visualisation. Sociology **52**(4), 830–848 (2018)
28. Keysers, D., Renn, M., Breuel, T.M.: Improving accessibility of html documents bygenerating image-tags in a proxy. In: Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility. pp. 249–250 (2007)
29. Kitson, A., Buie, E., Stepanova, E.R., Chirico, A., Riecke, B.E., Gaggioli, A.: Transformative experience design: designing with interactive technologies to support transformative experiences. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–5 (2019)
30. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: European conference on computer vision. pp. 740–755. Springer (2014)
31. Loiacono, E.T., Romano Jr, N.C., McCoy, S.: The state of corporate website accessibility. Communications of the ACM **52**(9), 128–132 (2009)
32. MacLeod, H., Bennett, C.L., Morris, M.R., Cutrell, E.: 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. pp. 5988–5999 (2017)
33. Meeker, M.: Internet trends 2014 (2014), https://cryptome.org/2014/05/internet-trends-2014.pdf
34. Meredith Whittaker, Meryl Alper, C.L.B.S.H.L.K.M.M.M.R.M.J.R.E.R.M.S.S.M.W.: Disability, bias, and ai (2019), https://ainowinstitute.org/disabilitybiasai-2019.pdf
35. Morris, M.R.: Ai and accessibility: A discussion of ethical considerations. arXiv preprint arXiv:1908.08939 (2019)

36. Morris, M.R., Zolyomi, A., Yao, C., Bahram, S., Bigham, J.P., Kane, S.K.: " with most of it being pictures now, i rarely use it" understanding twitter's evolving accessibility to blind users. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 5506–5516 (2016)
37. Obrist, M., Wurhofer, D., Beck, E., Karahasanovic, A., Tscheligi, M.: User experience (ux) patterns for audio-visual networked applications: inspirations for design. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. pp. 343–352 (2010)
38. Online, B.: Whatsapp users share 55 billion texts, 4.5 billion photos, 1 billion videos daily (2017), https://www.businesstoday.in/technology/news/whatsapp-users-share-texts--photos-videos-daily/story/257230.html
39. Porter, T.M.: Trust in numbers: The pursuit of objectivity in science and public life. Princeton University Press (1996)
40. Power, C., Freire, A., Petrie, H., Swallow, D.: Guidelines are only half of the story: accessibility problems encountered by blind users on the web. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 433–442 (2012)
41. Ramnath, K., Baker, S., Vanderwende, L., El-Saban, M., Sinha, S.N., Kannan, A., Hassan, N., Galley, M., Yang, Y., Ramanan, D., et al.: Autocaption: Automatic caption generation for personal photos. In: IEEE Winter Conference on Applications of Computer Vision. pp. 1050–1057. IEEE (2014)
42. Rodríguez Vázquez, S.: Measuring the impact of automated evaluation tools on alternative text quality: a web translation study. In: Proceedings of the 13th Web for All Conference. pp. 1–10 (2016)
43. Rowe, N.C.: Marie-4: A high-recall, self-improving web crawler that finds images using captions. IEEE Intelligent Systems **17**(4), 8–14 (2002)
44. Salisbury, E., Kamar, E., Morris, M.R.: Toward scalable social alt text: Conversational crowdsourcing as a tool for refining vision-to-language technology for the blind. In: 5th AAAI Conference on Human Computation and Crowdsourcing (2017)
45. Salisbury, E., Kamar, E., Morris, M.R.: Evaluating and complementing vision-to-language technology for people who are blind with conversational crowdsourcing. In: IJCAI. pp. 5349–5353 (2018)
46. Shaban, H.: Twitter reveals its daily active user numbers for the first time (2019), https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-first-time/
47. Shi, Y.: E-government web site accessibility in australia and china: A longitudinal study. Social Science Computer Review **24**(3), 378–385 (2006)
48. Siegemedia: The 100 most popular google keywords (mar 2020), https://www.siegemedia.com/seo/most-popular-keywords
49. Spyridonis, F., Daylamani-Zad, D.: A serious game to improve engagement with web accessibility guidelines. Behaviour & Information Technology **39**(4), 1–19 (2020). https://doi.org/https://doi.org/10.1080/0144929X.2019.1711453
50. Stangl, A., Morris, M.R., Gurari, D.: " 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. pp. 1–13 (2020)
51. Strauss, A., Corbin, J.: Basics of qualitative research techniques. Sage publications Thousand Oaks, CA (1998)
52. todd: Accessible images for everyone (2016), https://blog.twitter.com/en_us/a/2016/accessible-images-for-everyone.html

53. Tran, K., He, X., Zhang, L., Sun, J., Carapcea, C., Thrasher, C., Buehler, C., Sienkiewicz, C.: Rich image captioning in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2016)
54. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M.: Predicting elections with twitter: What 140 characters reveal about political sentiment. In: Fourth international AAAI conference on weblogs and social media (2010)
55. Twitter: Rules and filtering: Premium (2020), https://developer.twitter.com/en/docs/twitter-api/premium/rules-and-filtering/operators-by-product
56. Twitter: using-premium-operators (2020), https://developer.twitter.com/en/docs/twitter-api/premium/rules-and-filtering/using-premium-operators
57. Twitter, I.: Promoted tweet (2020), https://business.twitter.com/en/help/campaign-setup/advertiser-card-specifications.html
58. Twitter, I.: the suggested minimum properties for cards (2020), https://developer.twitter.com/en/docs/twitter-for-websites/cards/overview/summary-card-with-large-image
59. @TwitterA11y: Adding descriptions to images (2020), https://twitter.com/TwitterA11y/status/1265689579371323392
60. @TwitterBusiness: Fun fact: A study of twitter accounts (2020), https://twitter.com/TwitterBusiness/status/1070423034467540992?s=20
61. Vilenchik, D.: The million tweets fallacy: Activity and feedback are uncorrelated. In: Twelfth International AAAI Conference on Web and Social Media (2018)
62. Vinyals, O., Toshev, A., Bengio, S., Erhan, D.: Show and tell: A neural image caption generator. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3156–3164 (2015)
63. Von Ahn, L., Ginosar, S., Kedia, M., Liu, R., Blum, M.: Improving accessibility of the web with a computer game. In: Proceedings of the SIGCHI conference on Human Factors in computing systems. pp. 79–82 (2006)
64. Voykinska, V., Azenkot, S., Wu, S., Leshed, G.: How blind people interact with visual content on social networking services. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. pp. 1584–1595 (2016)
65. (WHO), W.H.O.: Blindness and vision impairment (2020), https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment
66. Wood, G., Long, K., Feltwell, T., Rowland, S., Brooker, P., Mahoney, J., Vines, J., Barnett, J., Lawson, S.: Rethinking engagement with online news through social and visual co-annotation. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2018)
67. Wu, S., Adamic, L.A.: Visually impaired users on an online social network. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 3133–3142 (2014)
68. Wu, S., Wieland, J., Farivar, O., Schiller, J.: 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. pp. 1180–1192 (2017)
69. Zhong, Y., Lasecki, W.S., Brady, E., Bigham, J.P.: Regionspeak: Quick comprehensive spatial descriptions of complex images for blind users. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 2353–2362 (2015)