The Crowd Work Accessibility Problem

Saiganesh Swaminathan
HCI Institute
Carnegie Mellon University
saiganes@cs.cmu.edu

Kotaro Hara
HCI and Robotics Institutes
Carnegie Mellon University
kotaroh@cmu.edu

Jeffrey P. Bigham
HCI and LT Institutes
Carnegie Mellon University
jbigham@cs.cmu.edu

ABSTRACT

Crowd work is an increasingly prevalent and important kind of work. Because of its flexible nature, crowd work may offer benefits for people with disabilities. Unfortunately, people with disabilities currently lack access to much of this work because the tasks that are posted are often inaccessible. In this paper, we first characterize the accessibility of the tasks posted to a popular crowd marketplace, Amazon Mechanical Turk, by performing manual and automatic checks on 120 tasks from several common types. We then outline research directions that could have positive impact on this problem. Given ongoing and upcoming changes to the world economy and technological progress, we believe it is important to find a way to make sure people with disabilities are able to equally participate in this kind of work.

Keywords

Crowd work; Accessibility

1. INTRODUCTION

People with disabilities do not have equal access to employment. A recent report suggests that people with disabilities are employed at less than half the rate of those without disabilities [7]. Crowdsourcing could offer alternative employment opportunities. In fact, prior work noted that people with disabilities are starting to find crowdsourcing as a source of income [14]. However, some people with disabilities may be segregated due to inaccessible crowd work and limited research has investigated the extent of the accessibility of the crowd work.

In this paper, we conduct an initial study on the accessibility of survey tasks on Amazon Mechanical Turk (AMT). We analyze the survey interfaces' compliance with Web Content Accessibility Guideline 2.0 (WCAG 2.0) Level A. To understand the crowdsourcing tasks' accessibiltiy compliance, we collected 100 survey tasks from AMT and analyzed the survey interfaces' compliance with the accessibiltiy guidelines using a readily available web accessibility checker software.

W4A 2017, April 02-04, 2017, Perth, Western Australia, Australia

DOI: http://dx.doi.org/10.1145/3058555.3058569

We found that only 2 out of the 100 surveys complied with the WCAG 2.0 Level A guidelines. For example, the majority of the surveys' form elements lacked label elements—the problem that prohibits screen reader software to work properly. We also found that 8 out of 25 WCAG 2.0 Level A guidelines were violated, but did not observe accessibility problems such as inaccessible use of color or lack of audio captions. This is likely because we sampled survey tasks; we will discuss this point in the Discussion section.

The contributions of this paper include the first quantitative examination of the survey task accessibility on AMT tasks. Our findings suggest that the majority of the crowdsourcing tasks are not fully compliant with the WCAG 2.0 Level A.

2. RELATED WORK

The accessibility of crowd work is related to both prior work on *(i)* disability and employment, and also *(ii)* on web accessibility, as a result of crowd work being employment that is primarily done on the web.

2.1 Disability and Employment

In the U.S., Title I and Title V of the Americans with Disabilities Act protect people from discrimination on the basis of disability in employment [11]. Despite the comprehensive civil rights legislation for equal employment opportunities, inequalities in employment based on disability remains a problem, and finding and maintaining jobs remain challenging for people with disabilities [2, 7]. In fact, 2015 Annual Disability Statistics Compendium reported that while the employment rate for people without disabilities is 75.4%, that of people with disabilities remained at 34.4% [7].

Crowd work could offer an alternative avenue for employment for people with disabilities; in fact, some are already working on existing platforms [14]. Crowdsourcing eliminates the need to adhere to a fixed schedule, reduces or eliminates the need for in-person social interaction, and enables flexible and remote work to avoid the difficulty of a commute. Despite its potential advantages, however, limited work has investigated the potential advantages and disadvantages of current crowdsourcing platforms. Consequently, we know little about the accessibility of tasks available as crowd work. The goal of this paper, therefore, is to identify accessibility problems in crowdsourcing tasks and articulate potential avenues for addressing these issues.

2.2 Web Accessibility

Accessibility of crowdsourcing tasks is contingent on the web accessibility since the tasks posted on the major platforms like AMT and UpWork require people to use web interfaces. Web Content Accessibility Guidelines (WCAG) published by W3C recommends the online contents to be accessible for everyone [13]. For example, it recommends all non-text contents to have text alternatives (alt text) and all functionality of the contents to be operable through a keyboard interface. Web contents of public web pages are, indeed, mandated to be accessible for everyone by the Section 508 of the Rehabilitation Act [10], and other web pages are recommended to follow general good accessibility practices [12]. Furthermore, numerous technologies have been developed to make web contents more accessible on a client side [3, 4], and developer tools have been made available to simplify the process of making web contents accessible for the users with disabilities (*e.g,.* [5]).

Despite these efforts, however, many web pages remain inaccessible for people with disabilities [6, 9]. In their longitudinal web accessibility study of high-traffic and government web pages, Hanson and Richards reported that many web pages did not adhere to web accessibility criteria even on 2012 [6], though the situation has been improving. This body of work, however, primarily focused on the accessibility of general web page contents and not crowdsourcing task interfaces and limited research exists for accessibility of crowd work

3. STUDY METHOD

In this communication paper, our focus was to study the accessibility of the existing tasks on AMT. We chose AMT because: *(i)* it is one of the most popular crowd work marketplaces; *(ii)* it is one of the few marketplaces in which tasks' details—such as their instructions and Web interfaces—are readily visible to anyone who has an AMT account; and *(iii)* the platform allows requesters to post tasks on web pages that they create themselves, which could cause varying levels of accessibility compliance.

As an initial exploratory study, we focused our investigation to assess the accessibility of survey tasks since they are one of the most prevalent form of work in AMT [8]. To study the tasks' accessibility compliance, we sampled N=100 surveys from AMT and analyzed their accessibility. The process included two main steps: task searching and accessibility analysis. In the first step, we looked for survey tasks by using a search term "survey" in the AMT's task search interface. This allowed us to imitate how a worker would search and browse tasks. We then locally stored the HTML and media files of the first 100 surveys that did not require special worker qualifications. Locally saving the surveys' HTML files was necessary since tasks on AMT appear and disappear frequently as requesters post them and workers accept them. Note, we only stored the first page of each survey task interface; when the first page only included the consent page, we advanced in the task to the first page where we saw the survey interface.

In the accessibility analysis step, we applied an automated accessibility checker to the locally stored survey task interfaces. We used HTML CodeSniffer[1], which allowed us to examine the HTML documents' WCAG 2.0 compliance as well as Section 508 compliance. In this study, we only analyzed the interfaces' compliance with WCAG 2.0 Level A guidelines following the study method of Hanson *et al.*[6].

[1]http://squizlabs.github.io/HTML_CodeSniffer

4. RESULTS

4.1 surveys

Guideline	Errors	# of Surveys
1.1.1 Non-text Content	8	3
1.3.1 Info and Relationships	610	93
2.4.1 Bypass Blocks	5	4
2.4.2 Page Titled	2	1
3.1.1 Language of Page	35	34
3.2.2 On Input	57	56
4.1.1 Parsing	8	3
4.1.2 Name, Role, Value	887	95

Table 1: Categories and frequencies of WCAG 2.0 Level A violations found on the sampled survey tasks (N=100). Of the twenty five categories of the Level A guidelines, we observed that eight categories were violated.

After running the accessibility checker on the surveys, as seen from Table 1, eight out of the twenty five WCAG 2.0 Level A guidelines were violated. Surprisingly, only two of the surveys fully complied with the guidelines. Below, we describe the eight guidelines that were violated in their order of occurrence.

Name, Role, Value (4.1.2): The majority of the survey tasks (N=95) violated the guideline. The problems were caused by the HTML elements of the surveys not providing ways to programmatically access values in input elements. For example, a drop-down menu (*i.e.,* <select> element) in Figure 1a had neither a corresponding <label> element nor element attributes (*e.g.,* aria-label). This prohibits a screen reader from identifying the role of this pull-down menu, which in turn negatively affects people with visual impairments to perform the survey tasks.

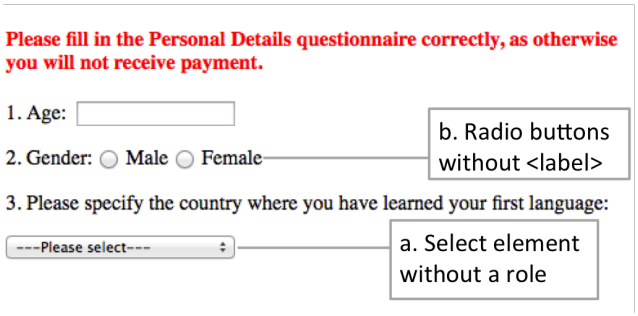

Figure 1: A snippet of survey where input elements lack proper labels (violate WCAG 1.3.1 and 4.1.2)

Info and Relationships (1.3.1): We observed that 93 surveys violated the guildeline on *Info and Relationships.*

These errors were caused by form elements that were missing corresponding labels. For example, the text description in Figure 1b ("Gender") was not associated with the radio button elements with "for" attribute. While sighted people could associate the description with the radio buttons with their visual proximity, lack of the explicit association prohibits the screen reader to identify the label for the radio buttons. This negatively affects the experience of people with visual impairments who use screen readers. Another benefit of using labels is that the users can click on the label itself to set focus to the form element as this increases the target area. This is useful to people with motor disabilities, particularly when selecting small checkboxes and radio buttons.

On Input (3.2.2): More than half of the surveys (N=56) violated this guideline. The problems were caused by the surveys that had forms with no submit buttons. In such forms when there are more than one input field, implicit submission through keyboard would not be possible *i.e,* hitting enter key would not trigger form submission. It affects users with motor impairments and visual impairments who primarily rely on keyboard instead of mouse for interacting with web forms.

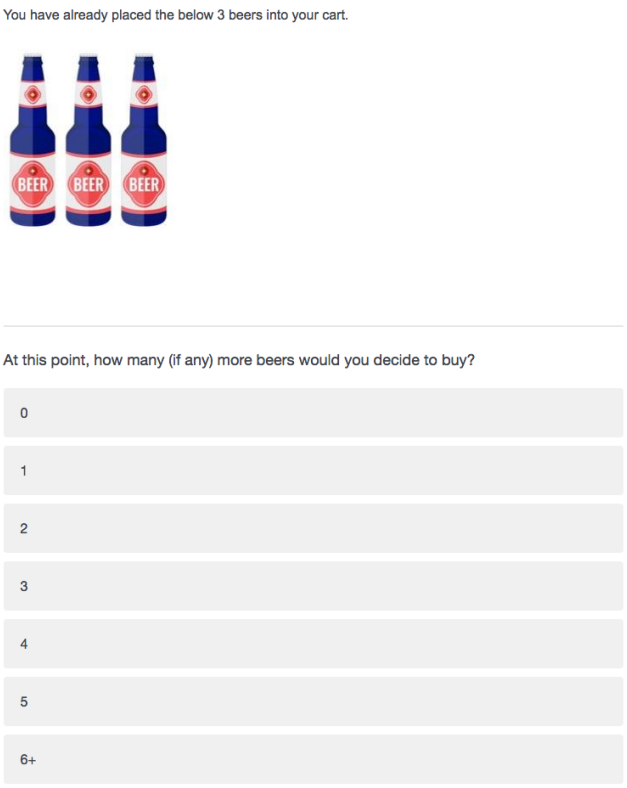

Figure 2: Example of task violating WCAG 1.1.1, not providing alt-text for image

Language of the page (3.1.1): In the one third of the survey documents (N=34), no language attribute was specified in the <html> tag. The lack of language could influence the technologies like screen reader as the information is used to switch language pronunciation and recognize the language specific syntax. The lack of properly working speech synthesis could negatively impact people who are blind since they might miss out on understanding the page without the correct information. Furthermore, this may create missed opportunities for work as there are an increasing number of language-specific tasks on AMT [1].

Bypass Blocks (2.4.1): We observed 4 of the 100 surveys that violated this guideline. These errors are caused by surveys that had empty titles in the iframes of the survey page. This problem affects blind and visually impaired users as they rely on screen readers to read the titles of iframes to get a quick overview of contents in them. They can use the overview information to then skip blocks of content or jump between texts in page. By addressing this issue, we can make it easier for screen reader users to quickly scan surveys to decide if they want to work on them.

Non-text content (2.4.1): We observed 3 surveys of the 100 that violated this guideline. The errors were relating to the perceptibility of images in a survey, *i.e.,* not providing any alt text to images. It creates an issue for people with visual impairments since the text provides an alternative route for them to understand the content in the images. For example in Figure 2, the tasks shows the user an image of three bottles to illustrate the scenario, a simple text like "blue bottles with beer label" would help increase the experience of blind users working on the task.

Parsing (4.1.1): We found that 3 surveys violated the guideline 4.1.1. The guideline 4.1.1 relates to errors occurred during the parsing of HTML content by screen readers. For example, HTML pages with incomplete tags could cause parsing error and make screen readers misinterpret the logical order of content. Hence fixing such errors could help in blind users understanding the tasks and the content in them in right way and ultimately resulting in better data quality for the requesters or task providers.

Page Titled (2.4.2): We found 1 survey where 2.4.2 was violated. This guideline provides criterion to make sure that all web pages have titles, which helps users get a quick overview about the page. This can help blind users to gather a quick overview about the type of survey or task being presented when they work.

5. DISCUSSION

We explored the accessibility of the survey tasks on AMT. Our results indicated that only 2% of the survey tasks on AMT complied with WCAG 2.0 Level A. Among the potential AMT workers with disabilities, we observed that visually impaired people are to be most affected. The eight types of the violations that we observed could negatively affect this population (*i.e.,* 98% of the surveys were potentially inaccessible). While smaller in the magnitude, 56% of the surveys were potentially inaccessible for people with motor impairments due to the *On Input* (3.2.2) violation. The accessibility issues we found on the survey tasks would have minimal impact on people with other types of disabilities.

It is likely that other types of crowdsourcing tasks have different accessibility problems. To get an initial sense of the kinds of accessibility problems found in task types other than surveys, we manually sampled (N=6) image transcription and audio transcription tasks. Many of the problems observed were similar to those we observed with surveys. For instance, only one of the 12 tasks properly labeled the form fields workers were expected to use. Seven had problems with dynamic content that were not coded to be accessible

to screen reader and screen magnifier users. Five could not be completed only with the keyboard.

Of course, some tasks are inherently inaccessible to some workers, *i.e.*, audio transcription is inaccessible to deaf workers, and image transcription is inaccessible to blind workers. Yet, it is interesting that these tasks often included problems that would prevent a worker with a disability from completing a task whose core was accessible to them. For instance, two of the audio transcription tasks included vital information in images, which might prevent a blind user from completing it. In one, all of the instructions were presented in an image without alternative text, and in the other all of the controls for controlling audio playback and submitting the task were contained within images without alternative text. The image transcription tasks could have been completed by a deaf worker, but three of the tasks required the worker to be able to also use the mouse (could not be completed using the keyboard only). Future work should further investigate the accessibility of various types of tasks.

Our findings suggest that effort in improving crowd work accessibility is needed. We believe three approaches should be pursued: (i) encourage requesters to make crowd work more accessible, (ii) design technologies to recommend accessible crowd work for people with disabilities, and (iii) automatically fix accessibility problems on crowd work interfaces. First, future work should explore ways to encourage requesters to make their tasks more accessible. In-depth analysis on the economic impact of segregating people with disabilities could influence behavior of requesters. Potential research questions include, for instance, *"what % of potential crowd workers are currently segregated due to inaccessible crowd work, and how much productivity an organization can gain by harnessing this labor force?"* While it is unlikely to influence the individual requesters' behaviors, it would influence larger organizations that heavily rely on crowdsourcing.

Second, future work should design and develop a task recommendation engine that takes into account of accessibility for recommendation criteria. For instance, we could implement the engine into a web browser plug-in. The simplest form of the recommendation engine would filter out tasks with many accessibility errors, thereby never suggesting inaccessible tasks to people with disability. We also envision the engine to let the user specify their levels of disabilities for personalizing task accessibility.

Third, fixing inaccessible tasks on AMT could be another approach for making crowdsourcing more accessible. We imagine (semi)-automated technologies can be implemented as a two-stage process. In the first stage, accessibiltiy problems on crowd task interfaces could be automatically identified by using the accessibiltiy checker. Then, we could address the identified problem via crowdsourcing and/or automated methods. For example, if a <label> tag is missing for a form element, we could inject an appropriate HTML element with crowdsourcing or machine learning.

6. CONCLUSION

In this paper, we explored the accessibility of survey tasks posted on AMT. To study the tasks accessibility compliance, we sampled N=100 surveys from AMT and analyzed their accessibility using an automated accessibility checker. We identified only 2% of the 100 survey tasks we sampled from AMT fully complied with WCAG 2.0 Level A. We identified that eight out of the twenty five WCAG 2.0 Level A guidelines were violated. Among 98% of the surveys that were potentially inaccessible, all surveys had issues for people with visual impairments and 56% of the surveys affected people with motor impairments.

7. ACKNOWLEDGMENT

The contents of this paper were funded by the National Science Foundation and the National Institute on Disability, Independent Living, and Rehabilitation Research.

8. REFERENCES

[1] Amazon. DARPA Language Challenge. https://requester.mturk.com/case_studies/cs/darpa.

[2] M. Ameri, L. Schur, M. Adya, S. Bentley, P. McKay, and D. Kruse. The disability employment puzzle: A field experiment on employer hiring behavior. Technical report, National Bureau of Economic Research, 2015.

[3] C. Asakawa and H. Takagi. *Transcoding*, pages 231–260. Springer London, London, 2008.

[4] J. P. Bigham. Making the web easier to see with opportunistic accessibility improvement. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 117–122, New York, NY, USA, 2014. ACM.

[5] Google. Accessibility Developer Tools.

[6] V. L. Hanson and J. T. Richards. Progress on website accessibility? *ACM Trans. Web*, 7(1):2:1–2:30, Mar. 2013.

[7] A. J. Houtenville, D. L. Brucker, and E. A. Lauer. Annual compendium of disability statistics: 2015. *Durham, NH: University of New Hampshire, Institute on Disability*, 2016.

[8] P. G. Ipeirotis. Analyzing the amazon mechanical turk marketplace. *XRDS*, 17(2):16–21, Dec. 2010.

[9] J. T. Richards, K. Montague, and V. L. Hanson. Web accessibility as a side effect. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '12, pages 79–86, New York, NY, USA, 2012. ACM.

[10] United States Department of Justice. Section 508 of the Rehabilitation Act (29 U.S.C. 794d). Technical report, 1998.

[11] C. R. D. United States Department of Justice. Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328. Technical report, 1990.

[12] W3C. Introduction to Web Accessibility, 1994.

[13] W3C. Web Content Accessibility Guidelines (WCAG) Overview, 2011.

[14] K. Zyskowski, M. R. Morris, J. P. Bigham, M. L. Gray, and S. K. Kane. Accessible Crowdwork? Understanding the Value in and Challenge of Microtask Employment for People with Disabilities. In *Proc. of CSCW 2015*, 2015.