

A Historical Review of Web Accessibility Using WAVE

Pooja Naresh Bhatia

bhatiapn@uci.edu

University of California, Irvine

USA

Sam Malek

malek@uci.edu

University of California, Irvine

USA

ABSTRACT

In today's fast-evolving technology, the Internet has become an integral part of our everyday lives. However, for people with disabilities, accessing the online world and navigating through it can be a challenging task. Accessibility on the web ensures that the online world can be used by people with disabilities which is convenient and not complicated to use. The purpose of this study is to offer a thorough knowledge of how web accessibility has changed and evolved over the past ten years. In this historical review, leveraging website archives, patterns and insights in web accessibility over time has been researched. The study collected data on web accessibility violations between January 2013 and April 2023 from 40 homepages of websites across four popular website categories (social network, entertainment, e-commerce, and news media) in the United States using the WAVE subscription API. The data was then further analyzed and insights were found, such as the most common web accessibility issues, the best year for the web accessibility, and the effect of the website category on the levels of web accessibility issues. The study aims to provide insights on the current condition of web accessibility and identify areas that still require development.

KEYWORDS

web accessibility, historical review, accessibility evaluation

ACM Reference Format:

Pooja Naresh Bhatia and Sam Malek. 2024. A Historical Review of Web Accessibility Using WAVE. In *2024 ACM/IEEE Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering (GE@ICSE '24), April 20, 2024, Lisbon, Portugal*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3643785.3648489>

1 INTRODUCTION

The use of the Internet has grown recently and is now a necessity in our daily lives. Even though it has a wide range of uses, people with disabilities still find it difficult to use online resources because of accessibility issues. According to the World Health Organization (WHO), 16% of people worldwide live with a disability of some kind or severity [10]. To avoid excluding the people with disabilities, certain design principles must be used when creating web content. Web accessibility ensures that website designers make their websites simple, easy to use and navigate for people with physical or cognitive impairments to use in order to guarantee inclusive

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GE@ICSE '24, April 20, 2024, Lisbon, Portugal

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0575-5/24/04.

<https://doi.org/10.1145/3643785.3648489>

web access [12]. Special input devices, screen readers, and voice recognition software are just a few of the assistive technologies that have been developed to make it easier for these users to understand online content and navigate through the websites. WHO defines assistive technology [9] as an umbrella term covering the systems and services related to the delivery of assistive products and services [11]. Assistive technologies like screen readers are essential for addressing the needs of users who are blind or visually impaired and making it easier for people to read online content. Using a joystick interface is also extremely advantageous for people with physical disabilities who need alternative methods of controlling computers.

It is crucial to make sure that individuals with disabilities can easily access the information available on the World Wide Web (WWW). Websites that are not accessible can make it difficult for them to navigate and understand the information which can lead to discrimination, and exclusion against them. These practices violate the legal requirement in many countries around the world, including the United States of America. The Americans with Disabilities Act (ADA), passed in 1990, prohibits discrimination against individuals with disabilities [2]. It states that companies or the organizations in the United States must make digital content accessible to people who have disabilities. Several standards and guidelines have also been developed to help web developers create accessible websites, such as Web Content Accessibility Guidelines (WCAG), which provides a set of guidelines for creating accessible web content [1]. Accessibility on the web continues to be a recurrent and significant challenge for many website designers and development teams despite the legal act and requirements in place to take into consideration. This can be a result of various accessibility barriers such as a lack of awareness and training provided to developers, rapid change in technology, or limited resources. Despite the laws and the importance of web accessibility, a majority of the Internet websites still remain not accessible [13].

This paper aims to provide a historical review of web accessibility over the past decade. By looking at the data from the past years and leveraging the web archives, we investigate how website accessibility has changed or evolved. For the purpose of this study, we used WAVE subscription API to collect information about the violations of accessibility guidelines by websites. We applied WAVE to ten popular websites in the United States from four popular categories—social network, entertainment, e-commerce, and news & media. Further, we conducted statistical analysis to understand the characteristics of our dataset and identify any salient patterns in terms of how web accessibility has changed.

The remainder of this paper is organized as follows. Section 2 provides an overview of the background and related work. Section 3 briefly describes our research objectives in this work. Section 4 details our methodology, followed by the discussion of results in

Section 5. Section 6 describes the limitations in our current study and avenues of future research. The paper concludes with a summary of our contributions in Section 7.

2 BACKGROUND & RELATED WORK

2.1 Previous studies on Web Accessibility

In order to gain an understanding of prior work in this area, we conducted a literature search on web accessibility and its evolution over the years. The relevant previous studies and publications were explored from various different research articles, scholarly databases and reputable websites on the internet. To gain vast coverage of the research topic, relevant keywords and phrases were searched. By combining the data from multiple resources, the literature was selected which provides insights about the historical review of web accessibility, performance assessment of websites and web content accessibility guidelines.

Over the past ten years, web accessibility has come a long way and made advancements by increasing awareness, setting regulations and laws, and availability of automated accessibility compliance tools. Several studies have also been conducted in the past two decades to evaluate the state of web accessibility which highlights the challenges and opportunities for improvement. The WebAIM Million, which is one of the notable studies analyzes the accessibility of the top one million web pages every year since 2019. It found that over 96.8% of the homepages were inaccessible [6]. This highlights the need for strong measures to improve the accessibility of the websites.

Acosta-Vargas and colleagues (2016) [14] presented a study assessing the accessibility of the websites of 20 different universities worldwide, using the Web Content Accessibility Guidelines 2.0 (WCAG 2.0) published by the World Wide Web Consortium (W3C) and the Website Accessibility Conformance Evaluation Methodology (WCAG-EM). The study states that universities with higher academic rankings have more accessible websites than the lower prestige institutes. The article helps identify the importance of web accessibility and provides the recommendations to the universities for compliance with accessibility guidelines.

Akram and Sulaiman (2017) [16] presented a systematic literature review. They analyzed Saudi Arabia's government and academic websites to determine if they were accessible to those with disabilities. They discovered that many of these websites did not adhere to appropriate accessibility standards. This means that people with disabilities may not be able to use them effectively.

Agrawal and colleagues (2019) [15] evaluated the website usability and accessibility for the airline industry in India. TAW automated accessibility checker tool was used to get the data of the violations of the website. The study found that none of the Indian airlines websites adhere to the WCAG 2.0 accessibility standards. This study helps us in understanding the necessity of improving the accessibility of airline websites.

The study on a retrospective look at website accessibility over the period from 1997 to 2002 was performed by Hackett and colleagues (2005) [18]. This study focuses on how technological advancements in web design may have affected the accessibility of several webpages for people with disabilities. The authors analyzed a sample of US government websites and a random sample of general websites

using the Bobby accessibility tool. The authors came to the conclusion that random websites became inaccessible as complexity increased, whereas US government sites remain accessible even when there was an increase in the complexity of the webpage.

Various studies efforts over the years have identified accessibility problems of the websites in different categories. The studies have evaluated the accessibility problem and identified the various causes of inaccessible content on the web [17] [21] [19]. They used a variety of techniques including manual as well as automated testing and a combination of both. There is a need to further evaluate the evolution of web accessibility in the recent years and identify the areas of improvement. With the advancement of technology in the recent years, it is important to identify what has changed and what needs attention. These developments may have given the web designers and developers new techniques and opportunities, but it is important to see if the advancement has raised new challenges or issues with the regard to accessibility of websites.

2.2 WAVE Subscription API

The WAVE Web Accessibility evaluation tool is one of the popular automated web accessibility checker tools. This tool enables web designers and developers to identify the issues on their website that do not conform to the WCAG guidelines [7]. The WAVE tool was developed by WEBAIM (Web Accessibility in Mind) and it was launched in 2001.

WAVE provides three different services for identifying the issues with the website accessibility: (1) WAVE browser extension, (2) WAVE API & testing engine which includes a subscription API and a standalone API, and (3) Accessibility IMpact (AIM) Report. For the purpose of this research, WAVE subscription API was utilized to evaluate the accessibility analysis of websites. It is a paid service that provides automated analysis of the web content. The WAVE API uses the API engine to evaluate the website after all the scripts have been applied and CSS has been updated, which enables an accurate evaluation of the issues on websites.

WAVE addresses 13 success criteria out of 78 which is approximately 16.66% of the WCAG violations. WAVE accessibility evaluation tool does not address all of the WCAG issues. However, it does detect the issues that are useful and are most common accessibility issues. WAVE addressed issues which can be automatically detected by analyzing the HTML web content of the website such as text alternatives and color contrast. It does not effectively detect issues that are related to learning or cognitive disabilities or which require human intervention such as issues related to audio or video content. It is essential to keep in mind that the combination of automated and manual testing is typically necessary to identify accessibility issues properly. Also, out of 13 success criteria, 11 of them are of conformance type Level A and only 2 are of Level AA. Table 1 shows the combined list of issues that is addressed by WAVE Accessibility Tool.

3 RESEARCH OBJECTIVES

The purpose of the research on the historical analysis of web accessibility is to look back over the past 10-year period to observe how web accessibility has evolved over time. The study provides insights into the accessibility of the popular websites in four different and

WCAG Principle	Guideline	Success Criteria	Conformance Level
1. Perceivable	1.1 Text Alternatives	1.1.1 Non-text Content	A
	1.3 Adaptable	1.3.1 Info and Relationships	A
	1.4 Distinguishable	1.4.3 Contrast (Minimum)	AA
2. Operable	2.1 Keyboard Accessible	2.1.1 Keyboard	A
	2.2 Enough Time	2.2.1 Timing Adjustable	A
		2.2.2 Pause, Stop, Hide	A
	2.4 Navigable	2.4.1 Bypass Blocks	A
		2.4.2 Page Titled	A
		2.4.4 Link Purpose (In Context)	A
		2.4.6 Headings and Labels	AA
3. Understandable	3.1 Readable	3.1.1 Language of Page	A
	3.3 Input Assistance	3.3.2 Labels or Instructions	A
4. Robust	4.1 Compatible	4.1.2 Name, Role, Value	A

Table 1: WCAG guidelines with success criteria and conformance level addressed by WAVE

popular categories to understand the changes in accessibility over time. This helps us gain an understanding of the trends in accessibility issues affecting websites. It also helps us identify the current state of web accessibility, including areas in which advancements have been made, and where further improvements are needed. Furthermore, the study aims to be a helpful and insightful resource for the individuals who are responsible for the creation of the websites and online content. Web designers and developers also gain insightful information about the areas of improvement needed in the websites and how to make their website more accessible and inclusive for all people. The study also helps create awareness of the importance of web accessibility and encourages web designers & developers to prioritize accessible and inclusive websites.

Specifically, we study the following three research questions:

- **RQ1:** Have the websites evolved to meet the Web Content Accessibility Guidelines (WCAG)?
- **RQ2:** What are the most common errors in Web Accessibility across all the websites?
- **RQ3:** How does web accessibility differ among various types of website categories?

4 METHODOLOGY

4.1 Research Design & Methodology

Figure 1 shows the methodology that is followed in this study. The study focuses on four popular website categories and each of the categories has ten different websites that are selected. After the websites have been selected, the data for the website URL is collected using the web archives from the Wayback Machine [8]. Another prerequisite is to setup the WAVE subscription API in order for us to collect the WCAG violation data. The website URL information that is collected from the archives is further used to make the API calls as an input and the WCAG violations that are addressed by the WAVE API are collected as the result. After collecting the data, it is normalized in order to scale the data to a standard range, such that the data is comparable across different website samples. Finally, several data analysis techniques are applied to the normalized data in order to understand the evolution, identify the common issues of web accessibility, identify the best year for accessibility, and analyze the website category effects on the web accessibility.

4.2 Dataset

4.2.1 *Data Collection of websites.* In order to collect the data on the popular website categories and the popular websites in each category, “Semrush.com” [3] and “SimilarWeb.com” [4] were utilized. Semrush and SimilarWeb are digital marketing tools that provide insights about the Internet traffic and the rankings of the websites. Based on the traffic information of the websites on these tools, four popular website categories were selected manually. These four categories are Social Network, E-commerce, Entertainment and News & Media. Table 2 displays the websites selected for each category based on their traffic.

The data for the past web content starting from January 2013 was collected from the web archives using the Wayback Machine. Wayback Machine is a non-profit digital library that archives the Internet websites. It captures the webpage snapshots periodically and stores them in the archive, which allows users to take a look at a single website at different points of time. It provides a search engine that takes the current deployed website URL and returns a list of past states of the website at that URL. The data was collected from January 2013 till April 2023 for every month and the broken URLs were ignored.

The websites were eliminated if they fell under any of the following criteria:

- (1) Information was not available for the website since 2013.
- (2) An instance of the website did not exist every month since 2013.
- (3) If the website has low number of HTML elements on the homepage, i.e., less than 100 elements.
- (4) If the website has no information available in Wayback Machine.

4.2.2 *WCAG violation data collection from WAVE Subscription API.* The WAVE Subscription API takes the website URL as an input, which was collected from the Wayback Machine, and provides the information about the analysis performed by the automated tool. The API call returns the data in the JSON format. This JSON data contains information about the statistics of the webpage such as total HTML elements, item count and it also provides data about the errors, alerts, features, structure and aria. For the purpose of our

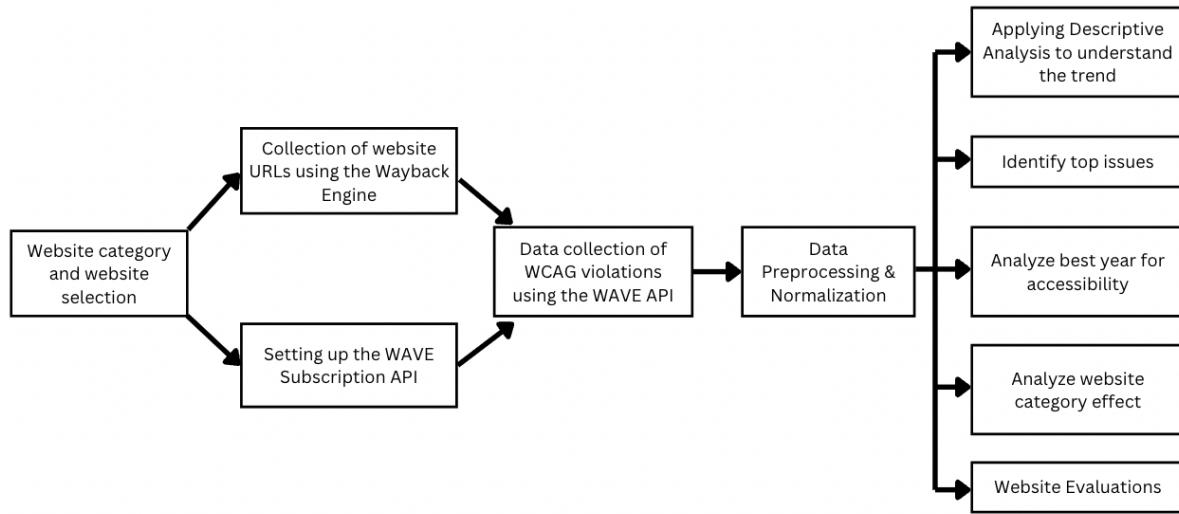


Figure 1: Methodology of the research

Category	Social Network		E-commerce		Entertainment		News & Media	
	Website	URL	Website	URL	Website	URL	Website	URL
1	Reddit	https://www.reddit.com/	Amazon	https://www.amazon.com/	Youtube	https://www.youtube.com/	Yahoo	https://www.yahoo.com/
2	Facebook	https://www.facebook.com/	Ebay	https://www.ebay.com/	Netflix	https://www.netflix.com/	ESPN	https://www.espn.com/
3	Twitter	https://www.twitter.com/	Walmart	https://www.walmart.com/	Imdb	https://www.imdb.com/	Nytime	https://www.nytimes.com/
4	Whatsapp	https://www.whatsapp.com/	Etsy	https://www.etsy.com/	Disney	https://www.disney.com/	Weather	https://weather.com/
5	LinkedIn	https://www.linkedin.com/	Target	https://www.target.com/	MSN	https://www.msn.com/	Foxnews	https://www.foxnews.com/
6	Yelp	https://www.yelp.com/	Wayfair	https://www.wayfair.com/	Spotify	https://www.spotify.com/	Accuweather	https://www.accuweather.com/
7	Myspace	https://myspace.com/	Bestbuy	https://www.bestbuy.com/	Gamespot	https://www.gamespot.com/	USA Today	http://www.usatoday.com/
8	Meetup	https://www.meetup.com/	Macys	https://www.macys.com/	RottenTomatoes	https://www.rottentomatoes.com/	Dailymail	https://www.dailymail.co.uk/
9	Foursquare	https://www.foursquare.com/	Homedepot	https://www.homedepot.com/	AVClub	https://www.avclub.com/	WSJ	https://www.wsj.com/
10	Tagged	https://www.tagged.com/	Costco	https://www.costco.com/	TMZ	https://www.tmz.com/	BBC	https://www.bbc.com/

Table 2: Website categories and top 10 websites in each category

study, we mainly collected error data that is useful in determining the violations of WCAG.

4.2.3 Data Normalization. After the data has been collected for all of the websites, the data normalization is performed on the data in order to make the data standardized. This process helps bring the data to a common scale, such that it is easier to make comparison between different websites and categories. In this study, the data has been normalized by calculating the *error rate* of each webpage. For calculating the error rate, the total number of errors that were identified by the WAVE API and the total elements of the webpage are required. The following formula is used to calculate the percentage of the error rate of a webpage:

$$\text{error rate of a webpage} = \frac{\text{total errors for a webpage}}{\text{total elements in a webpage}} \times 100$$

The WAVE API JSON response contains information about the total elements of the webpage along with the error details, which allows us to accurately compute the error rate per webpage. This

normalization of the data helps us perform meaningful comparisons across different webpages and identify the trends over the period of 10 years.

5 RESULTS

5.1 Trend of Web Accessibility over the past decade

The overall violations of Web Accessibility appears to be decreasing in the past ten years. Figure 2 shows the line graph that highlights the overall decrease in the accessibility violations identified by WAVE Subscription API for the 40 different homepages of the selected websites. Table 3 shows the overview of the average error rate each year.

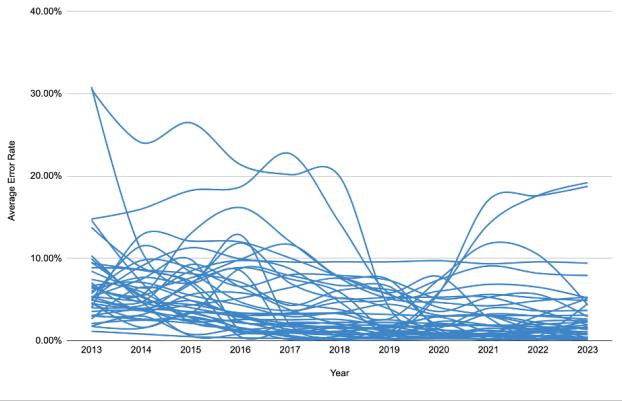


Figure 2: Decrease in Web Accessibility issues from January 2013 till April 2023

Year	Average Error Rate
2013	7.33%
2014	6.35%
2015	6.35%
2016	5.74%
2017	4.68%
2018	3.98%
2019	2.86%
2020	2.80%
2021	3.30%
2022	3.23%
2023	3.06%

Table 3: Average accessibility error rate since January 2013 till April 2023

The average error rate was 7.33% in the year 2013 which has decreased to 3.06% in the year 2023. That is a decrease of approximately 4.27% over the decade. The error rate has continued to decrease from the year 2017 at a slower pace. This indicates that web accessibility has been given attention and the websites are improving in terms of inclusivity. Overall, the data shows that efforts are being made in order to make the websites more accessible.

Years 2019 and 2020 have less than 3% of violations that are detected by the WAVE API. This may have been an effect of the update in the Section 508 of the Rehabilitation Act in 2017[5], which addresses the accessibility of electronic and information technology for people with disabilities. The law requires federal agencies' electronic and information technology, including websites, to be accessible to all individuals, including those with disabilities. Several high-profile lawsuits in 2017-2018 time frame further brought attention to the issues of accessibility, including a court ruling that handed victory to a blind man who sued Domino's pizza over site accessibility [20]. Additionally, WCAG 2.1 was released in June 2018 which provides more detailed information and examples about the success criteria for a truly accessible site. This may have raised awareness among the web designers and developers about the importance of the accessibility of the website. It is hard to know for

certain what has triggered this trend. There may also be other factors that might explain the decrease in the violations, such as growing popularity of certain frameworks such as React for construction of websites. Additional research is needed to understand the underlying reasons.

The best year for accessibility can be considered from two perspectives. To determine the lowest percentage of web accessibility violations, the best year can be identified as the one that has the lowest rate of violations across all websites. On the other hand, to show improvement over time, the best year would be determined as the largest drop in rate of violations compared to the prior year. The year 2020 is the best year in terms of the least average error rate of WCAG violations. The best year can also be considered as the year 2019 in terms of improvement over time because it had the largest decrease in the issues when compared with the previous year of around 1.12%. This indicates that over time there have been efforts in addressing the issues and the developers are paying more attention to accessibility. In spite of that, still, improvements need to be made in order to have webpages that are easier to use for people with disabilities.

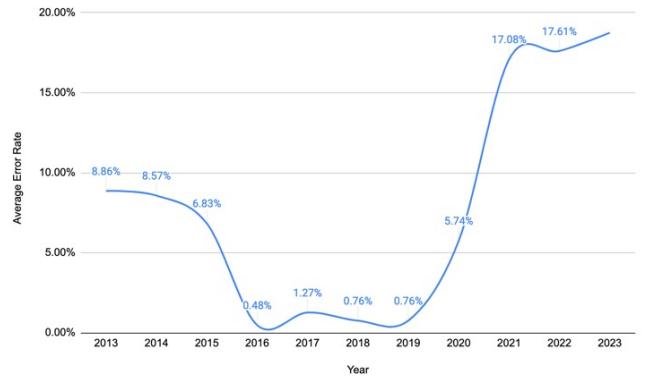


Figure 3: YouTube web accessibility violations over the years.

Although we observed an overall decreasing trend in web accessibility issues, in a few cases we noticed significant spikes that were different from the overall trend. For instance, Figure 3 shows the the overall trend for YouTube web accessibility violations over the past decade. A further investigation of YouTube revealed that the drastic changes in error rate coincide with the major changes to its user interface. YouTube has gone through three major user interface changes since 2013, as shown in Figure 4. These changes have introduced new elements and features that clearly impacted the platform's accessibility. Notably, we notice that after the UI redesign in mid 2020s, the average error rate of YouTube website has increased sharply and reached an average error rate of 18.75% in 2023. This suggests that one factor that could have contributed to YouTube's lower error rate in the early years was simplicity in the design of its user interface. The website had relatively straightforward UI in the early years that mainly revolved around the video content. The simplicity of the UI made the website easier to navigate using assistive technologies (e.g., screen readers). The features and functionalities available on the website have grown dramatically

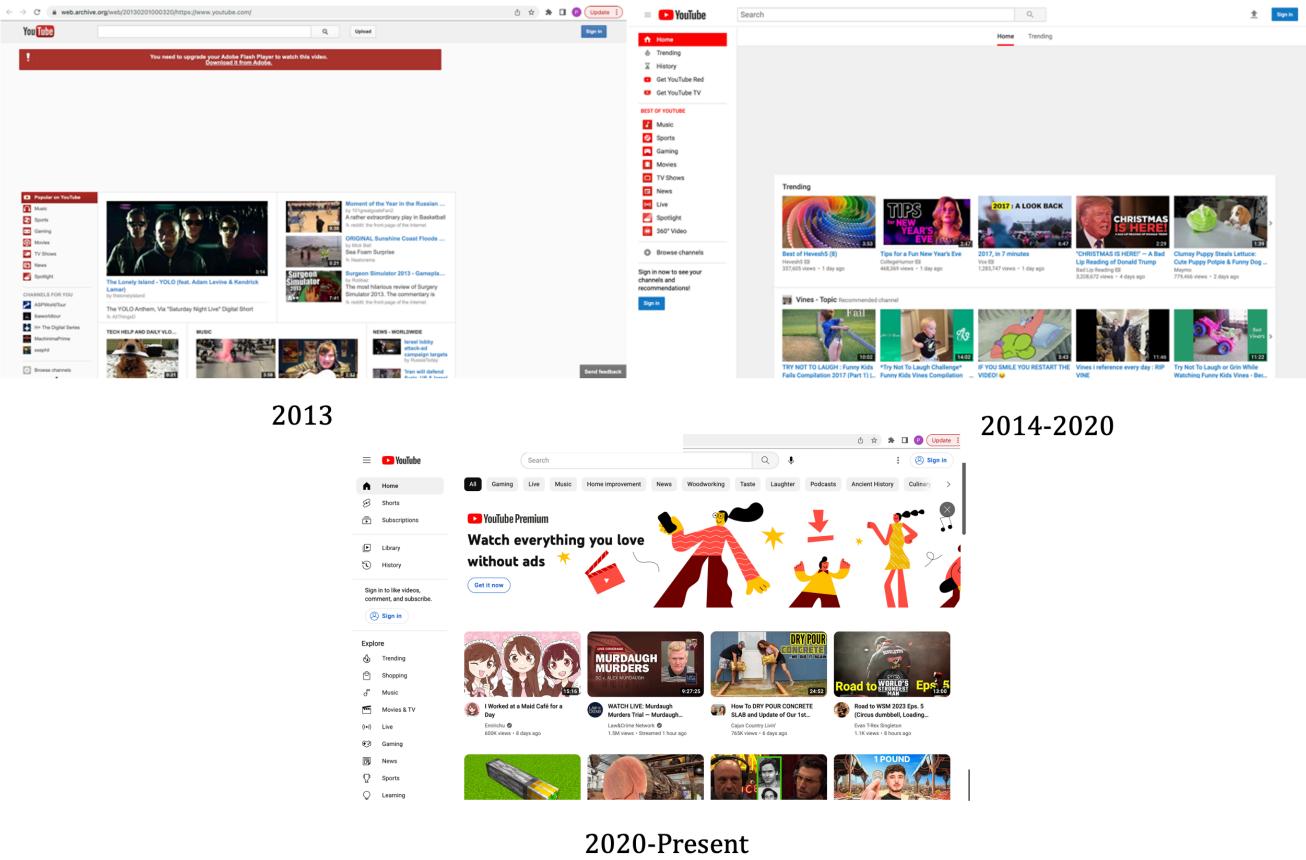


Figure 4: YouTube user interface over the years

over time, which appear to have negatively affected the accessibility of the website.

5.2 Most common errors that violate the WCAG guidelines

In this section, we describe the most common accessibility issues that are affecting the websites. Table 4 shows 23 different types of errors that the WAVE API is able to detect in a webpage along with their average error rate among all the websites. The most common error was “Very low contrast” with an approximate average error rate of 2.78%. This can make it difficult for people with visual impairments to understand the content of the webpage. The second most common error is “Missing alternative text” with approximately around 0.53%.

Some accessibility issues are easier to resolve than others. Some errors can be fixed quickly by running the automated detection tool and manually fixing it (e.g., missing alternative text, empty heading), while some errors require specialized knowledge and a deeper understanding of the design choices (e.g., how the webpage design determines the navigation order for a keyboard user). Regardless,

these errors can serve as a good starting point for web designers to make their websites more inclusive.

5.3 Effect of type of website category on the levels of web accessibility violations

We now discuss the impact of website categories on web accessibility. Table 5 shows the categories of websites with the highest to lowest average error rates over the past decade.

No	Category	Overall Average Error Rate
1	Social Network	6.59%
2	Entertainment	4.90%
3	News Media	4.12%
4	E-commerce	2.84%

Table 5: Average error rate for website categories

Internet users all across the world utilize social networking sites almost regularly. It has become an essential part for users all around the world, where people can interact with each other, share photos & videos, create profiles and much more. We found that social networking websites have the highest average error rate among

No	Error	Average Rate	Error
1	Very low contrast	2.78%	
2	Missing alternative text	0.53%	
3	Linked image missing alternative text	0.38%	
4	Empty link	0.33%	
5	Missing form label	0.18%	
6	Empty button	0.10%	
7	Empty heading	0.09%	
8	Document language missing	0.09%	
9	Spacer image missing alternative text	0.06%	
10	Empty form label	0.02%	
11	Broken ARIA menu	0.02%	
12	Broken ARIA reference	0.02%	
13	Multiple form labels	0.01%	
14	Image map missing alternative text	0.01%	
15	Broken skip link	0.01%	
16	Image button missing alternative text	0.00%	
17	Empty table header	0.00%	
18	Image map area missing alternative text	0.00%	
19	Missing or uninformative page title	0.00%	
20	Page refreshes or redirects	0.00%	
21	Blinking content	0.00%	
22	Invalid longdesc	0.00%	
23	Marquee	0.00%	

Table 4: Average error rate of different type of issues addressed by WAVE

the four popular website categories. This can be due to the nature of these social websites, which is dynamic, and large amount of content is updated and changed constantly.

Another popular category is Entertainment which refers to online content and streaming platforms. This category has the second highest overall average error rate with around 4.90% of inaccessibility. It can be due to the reason that entertainment websites have content that falls under multimedia such as videos, streaming and images. Also, entertainment websites may tend to make the user interface appealing and prioritize the experience of the user rather than the functionality.

News & Media is another popular category that provides news and information of politics, celebrities, sports and much more. This category was the third highest in inaccessibility with an average error rate of around 4.12%. The reason for an inaccessible webpage can be that it can have breaking news that needs to be delivered urgently or continuous changes to the content.

E-commerce websites offer users and companies to buy and sell their goods and products online. This website category has performed best among all the categories with an average error rate of 2.84%. This can be because the e-commerce web user interface is consistent and has a streamlined layout. Also, e-commerce websites may have a focus on the products and the services rather than an appealing user interface.

6 DISCUSSION

6.1 Limitations

This study provides certain insights into the overall trends in web accessibility. However, the authors recognize the limitations of the study:

- (1) Automated testing: The study used an automated accessibility evaluation tool, WAVE, to collect WCAG violations of websites. However, manual testing is always required along with automated testing, because many errors are not identified by automated tools and require human intervention and judgement.
- (2) Limited WCAG guidelines addressed: WAVE only addresses 13 success criteria out of 78 which is approximately 16.66% of the WCAG violations. The issues that are not in the scope of the WAVE API are not addressed in our study, and only the most impactful and relevant issues have been reported.
- (3) Website category not aligned with homepage content: Webpages may belong to more than one category. For example, WhatsApp's homepage is not purely a social network and has features that may fall under other categories. However, in our categorization of web pages, we only considered a webpage to belong to a single (primary) category.

6.2 Future Directions

Several avenues of future work exist for continuation of this study:

- (1) Expansion of the dataset: This research focuses on only four popular categories. Other website categories can be examined. Also, expanding the number of websites in each category can help in providing a more comprehensive view of web accessibility over the years.
- (2) Conducting user study: This study uses an automated tool in order to get the results of the WCAG violations of the homepages of the websites. Conducting manual testing of the websites by disabled users who actually experience these issues can provide a more nuanced and detailed understanding of the web accessibility issues.
- (3) Impact of technology: New technologies, particularly in the form of JavaScript frameworks, have been developed in the past decade to make the user interface more appealing and responsive. Examination and comparison between these technologies and the WCAG violations can help us understand the impact of the emerging technologies on web accessibility.

7 CONCLUSION

We conducted an investigation of the evolution of web accessibility over the course of past decade. To that end, we evaluated 40 different websites across four popular categories for WCAG violation. In order to acquire the WCAG violations data, we used the WAVE subscription API and the Wayback Machine. Over the course of the past ten years, we found that there has been an improvement in the accessibility of websites. However, more effort is still required to reduce the number of WCAG violations found on webpages. The study looked into the most common accessibility issues on the web as well as the year that was the most successful from an accessibility perspective. In addition to this, the study presents an examination

of the impact that the different website categories have on the number of web accessibility violations. The need of addressing accessibility concerns in order to build an inclusive environment online is emphasized by this research. This study raises awareness among web designers and encourages and supports attempts aimed at enhancing everyone's access to and inclusion on the Internet.

Our research artifacts are publicly available on the following GitHub repository: <https://github.com/poojabhatia9599/historical-web-accessibility-data>

ACKNOWLEDGMENTS

This work was supported in part by award numbers 2211790 and 2106306 from the National Science Foundation. We are grateful to Fable for granting us access to its platform, which allowed us to conduct extensive user studies. We appreciate the anonymous reviewers of this paper for their detailed feedback, which helped us improve the work.

REFERENCES

- [1] [n. d.]. Web Content Accessibility Guidelines (WCAG) 2.1. <https://www.w3.org/TR/WCAG21/>. Accessed March 2023.
- [2] 2023. National Network. <https://adata.org/learn-about-ada>. Accessed March 2023.
- [3] 2023. Semrush. <https://www.semrush.com/>. Accessed March 2023.
- [4] 2023. SimilarWeb. <https://www.similarweb.com/>. Accessed March 2023.
- [5] 2023. The Section 508 Update. <https://www.section508.gov/blog/accessibility-news-the-section-508-Update/>. Accessed March 2023.
- [6] 2023. The WebAIM Million. <https://webaim.org/projects/million/2022>. Accessed March 2023.
- [7] 2023. WAVE Subscription API. <https://wave.webaim.org/api/>. Accessed March 2023.
- [8] 2023. Wayback Machine. <https://web.archive.org/>. Accessed March 2023.
- [9] 2023. Web Assistive Technologies. <http://sites.stedwards.edu/accessibility/web-assistive-technologies/>. Accessed March 2023.
- [10] 2023. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>. Accessed March 2023.
- [11] 2023. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/assistive-technology>. Accessed March 2023.
- [12] 2023. World Wide Web Consortium. <https://www.w3.org/WAI/fundamentals/accessibility-intro/>. Accessed March 2023.
- [13] Hayfa Abu Addous, Mohd Zalisham, and Nurilda Basir. 2016. Web Accessibility Challenges. *International Journal of Advanced Computer Science and Applications* 7 (2016). <https://doi.org/10.14569/IJACSA.2016.071023>
- [14] Patricia Acosta-Vargas, Sergio Luján-Mora, and Luis Salvador-Ullauri. 2016. Evaluation of the web accessibility of higher-education websites. (2016), 1–6. <https://doi.org/10.1109/ITHET.2016.7760703>
- [15] Gaurav Agrawal, Dinesh Kumar, Mukesh Singh, and Dhaval Dani. 2019. Evaluating Accessibility and Usability of Airline Websites. 1045 (2019). https://doi.org/10.1007/978-981-13-9939-8_35
- [16] Muhammad Akram and Rosnafisah Bt Sulaiman. 2017. A Systematic Literature Review to Determine the Web Accessibility Issues in Saudi Arabian University and Government Websites for Disable People. *International Journal of Advanced Computer Science and Applications (IJACSA)* 8, 6 (2017). <https://doi.org/10.14569/IJACSA.2017.080642>
- [17] Stephanie Hackett and Bambang Parmanto. 2005. A longitudinal evaluation of accessibility: Higher education web sites. *Internet Research* 15 (2005), 281–294. <https://doi.org/10.1108/10662240510602690>
- [18] Stephanie Hackett, Bambang Parmanto, and Xiaoming Zeng. 2005. A retrospective look at website accessibility over time. *Behaviour & Information Technology* 24, 6 (2005), 407–417. <https://doi.org/10.1080/01449290500066661>
- [19] Abid Ismail and K. S. Kuppusamy. 2018. Accessibility of Indian universities' homepages: An exploratory study. *Journal of King Saud University - Computer and Information Sciences* 30, 2 (2018), 268–278. <https://doi.org/10.1016/j.jksuci.2016.06.006>
- [20] Jonahtan Lazar. 2022. Ableist Language in the Web Accessibility Case of Robles v. Domino's Pizza LLC: When "Everyone" Doesn't Include People with Disability. 2 (2022). <https://doi.org/10.51357/id.vi2.197>
- [21] Osama Sohaib and Kyeong Kang. 2016. Assessing Web Content Accessibility of E-Commerce Websites for People with Disabilities. (2016).