UNIVERSITY OF CALIFORNIA,
IRVINE

A historical review of Web Accessibility using WAVE

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Software Engineering

by

Pooja Naresh Bhatia

Thesis Committee:
Professor Sam Malek, Chair
Professor David Redmiles
Associate Professor of Teaching Hadar Ziv

2023

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my thesis advisor and committee chair, Professor Sam Malek, for being my advisor and mentor throughout my Master's degree. His support and guidance throughout my thesis journey are very appreciated. His extensive knowledge in the field of Software Engineering with his invaluable support has helped me complete my thesis. I would also like to thank my committee members Professor David Redmiles and Associate Professor Hadar Ziv for their encouragement, consideration and counsel.

ABSTRACT OF THE THESIS

A historical review of Web Accessibility using WAVE

By

Pooja Naresh Bhatia

Master of Science in Software Engineering

University of California, Irvine, 2023

Professor Sam Malek, Chair

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 in the United States - social network, entertainment, e-commerce, and news media 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. To gain a further deeper understanding, website evaluations were conducted on one website from each of the four categories. The study will aim to provide insights on the current condition of web accessibility and identify areas that still require development by looking at previous data. Researchers, web designers, developers interested in enhancing and improving web accessibility will find this analysis interesting and beneficial. This study is an important step toward raising awareness about a more

accessible online environment and it also offers useful information to developers who are committed to building websites that are more inclusive.

Chapter 1

Introduction

1.1 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 percent of people worldwide live with a disability of some kind or severity [1]. To avoid excluding the people with disabilities populations, 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 web access [2]. 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 [3] as an umbrella term covering the systems and services related to the delivery of assistive products and services [4]. 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's 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 violates 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 [5]. 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 [6]. 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 [7].

The historical review of web accessibility aims at providing insights into the evolution of web accessibility over the period of 10 years. By looking at the data from the past years and leveraging the web archives, the study will target finding how website accessibility has changed or evolved. For the purpose of this study, WAVE subscription API is used in order to collect information about the violations of the website. Ten popular websites in the United States have been selected from four popular categories- social network, entertainment, e-

commerce, and news & media. Further, descriptive statistics analysis has been applied to the data to describe the characteristics and summarize the basic features of the dataset used. Insights about what are the common types of accessibility errors and what is the best year of web accessibility have been discussed. In order to gain a deeper understanding of the category of the website, the website evaluations are conducted on websites which offer real-world examples.

1.2 Purpose of the study

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 been evolved over time. The study provides insights into the accessibility of the popular websites in four different and popular categories to understand the changes in accessibility over time. This will help gain a comprehensive understanding of the history of accessibility issues. Further, it will help identify the current state of web accessibility and whether advancements have been made in this field and identify areas where websites are doing well and where improvements could be made to make the websites more accessible for people with disabilities.

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 will gain insightful information about the areas of improvement needed in the websites and how to make their webpages more accessible and inclusive for all people. The study will also help create awareness of the importance of web accessibility and will encourage web designers & developers to prioritize accessible and inclusive websites.

1.3 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?

Chapter 2

Background & Related Work

2.1 Previous studies on Web Accessibility

In order to gain information about the previous study, a literature search was conducted to gather the knowledge about 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 & 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. Over 96.8% of the homepages were inaccessible and had detected WCAG 2 failures in 2022 [8]. This highlights the need for strong measures to improve the accessibility of the websites.

P. Acosta-Vargas and authors (2016) [9] 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.

Muhammad Akram and Rosnafisah Bt Sulaiman (2017) [10] present 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, G., and authors (2019) [11] evaluate 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's findings state 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 Stephanie Hackett and authors (2005) [12] . This study focuses on

how technological advancements in web design may have affected the accessibility of the webpage for people with disabilities. The authors analyzed a convenience sample of US government websites and a random sample of general websites using the Bobby accessibility tool. The authors came to the conclusion 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 have been done over the years to identify 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 [13] [14] [15] . 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 the website.

2.2 Web Content Accessibility Guidelines (WCAG 2.1)

The Web Content Accessibility Guidelines 2.1 [6] provides the wide range of standards and recommendations to make the webpages more accessible to users with disability. WCAG 2.1 was released in June 2018, followed by the version 2.0 [16] which was release in December 2008. The new version has new success criteria and additions to the conformance section for people with low vision and cognitive impairments.

The WCAG guidelines revolve around four different principles and each of the principle contains guidelines and success criteria for making the websites accessible. The principles of WCAG are mentioned below:

1. Perceivable: It refers to the user's need to be able to perceive the components of user interface and available information. For example, providing text alternatives for the content.

2. Operable: It states that the user interface and the navigational content is operable i.e. the user is able to use the controls and the computer interaction elements. For example, users can operate and navigate through the website using the keyboard-only commands.

3. Understandable: It states that the user is able to easily understand the content of the website. For example, making the content or the text more readable and understandable for people with disabilities.

4. Robust: It means that the website should be robust enough to be interpreted and understood by the wide range of users and assistive technologies. For example, improving the compatibility of the user agents.

WCAG also defines the three levels of conformance and lists the requirements that each of these levels have. Each WCAG principle has multiple guidelines and each guideline has multiple success criterias with the level of conformance. The three conformance levels can be described as:

1. Level A conformance: This is the minimum level of conformity. The requirements are met for level A if the website satisfies all the level A success criteria.

2. Level AA conformance: The requirements are met for level AA if the website satisfies all the level A and level AA success criterias.

3. Level AAA conformance: This is the maximum level of conformity. The requirements are met for level AAA if the website satisfies all the level A, level AA and level AAA success criterias.

WCAG 2.2 is set to be released in April 2023 which will add further guidelines and criteria for users with lower vision, learning disabilities or cognitive health issues.

2.3 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 [17] . 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- WAVE browser extension, WAVE API & testing engine which includes a subscription API and a standalone API, and Accessibility IMpact (AIM) Report. For the purpose of this research, WAVE subscription API was utilized to evaluate the accessibility analysis of the webpage. 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, this enables more accurate evaluation of the issues of webpage.

2.3.1 List of WCAG violations address by WAVE API

WAVE accessibility evaluation tool does not address all the guidelines listed under the WCAG 2.1. The WAVE subscription API addresses the most impactful and relevant WCAG level A and AA success criteria only [18]. The tool does not detect the level AAA conformance level which is the highest level of accessibility that the website can achieve. We should note that addressing level AAA issues is difficult using automated tools and no tool can detect all the accessibility issues with the website, manual testing and automated testing both works optimally together [19]. This section will provide list of issues that are addressed by the WAVE tool. In order to gain better understanding of the types of the errors, the list is further divided into the four WCAG principles. The list also contains multiple success criteria and the conformance level for each guideline. The list also shows which violations/errors are addressed and detected by the WAVE accessibility tool:

1. Perceivable Principle The Table 2.1 provides the list of guidelines under the WCAG Perceivable principle. There are only 3 criteria addressed out or 29 WCAG successes out of which two of the criteria are of level A and the one of them is from level AA.

Guideline	**Success Criteria**	**Conformance Level**	**Addressed by WAVE**
1.1 Text Alternatives	1.1.1 Non-text Content	A	TRUE
1.2 Time-based Media	1.2.1 Audio-only and Video-only (Prerecorded)	A	FALSE
	1.2.2 Captions (Prerecorded)	A	FALSE
	1.2.3 Audio Description or Media Alternative (Prerecorded)	A	FALSE
	1.2.4 Captions (Live)	AA	FALSE
	1.2.5 Audio Description (Prerecorded)	AA	FALSE
	1.2.6 Sign Language (Prerecorded)	AAA	FALSE
	1.2.7 Extended Audio Description (Prerecorded)	AAA	FALSE
	1.2.8 Media Alternative (Prerecorded)	AAA	FALSE
	1.2.9 Audio-only (Live)	AAA	FALSE
1.3 Adaptable	1.3.1 Info and Relationships	A	TRUE
	1.3.2 Meaningful Sequence	A	FALSE
	1.3.3 Sensory Characteristics	A	FALSE
	1.3.4 Orientation	AA	FALSE
	1.3.5 Identify Input Purpose	AA	FALSE
	1.3.6 Identify Purpose	AAA	FALSE
1.4 Distinguishable	1.4.1 Use of Color	A	FALSE
	1.4.2 Audio Control	A	FALSE
	1.4.3 Contrast (Minimum)	AA	TRUE
	1.4.4 Resize text	AA	FALSE
	1.4.5 Images of Text	AA	FALSE
	1.4.6 Contrast (Enhanced)	AAA	FALSE
	1.4.7 Low or No Background Audio	AAA	FALSE
	1.4.8 Visual Presentation	AAA	FALSE
	1.4.9 Images of Text (No Exception)	AAA	FALSE
	1.4.10 Reflow	AA	FALSE
	1.4.11 Non-text Contrast	AA	FALSE
	1.4.12 Text Spacing	AA	FALSE
	1.4.13 Content on Hover or Focus	AA	FALSE

Table 2.1: WCAG Perceivable guidelines with success criteria and conformance level addressed by WAVE

2. Operable Principle The Table 2.2 provides the list of guidelines under the WCAG Operable principle. There are only 7 criteria addressed out or 29 WCAG successes out of which six of the criteria are of level A and the one of them is from level AA.

Guideline	**Success Criteria**	**Conformance Level**	**Addressed by WAVE**
2.1 Keyboard Accessible	2.1.1 Keyboard	A	TRUE
	2.1.2 No Keyboard Trap	A	FALSE
	2.1.3 Keyboard (No Exception)	AAA	FALSE
	2.1.4 Character Key Shortcuts	A	FALSE
2.2Enough Time	2.2.1 Timing Adjustable	A	TRUE
	2.2.2 Pause, Stop, Hide	A	TRUE
	2.2.3 No Timing	AAA	FALSE
	2.2.4 Interruptions	AAA	FALSE
	2.2.5 Re-authenticating	AAA	FALSE
	2.2.6 Timeouts	AAA	FALSE
2.3 Seizures and Physical Reactions	2.3.1 Three Flashes or Below Threshold	A	FALSE
	2.3.2 Three Flashes	AAA	FALSE
	2.3.3 Animation from Interactions	AAA	FALSE
2.4 Navigable	2.4.1 Bypass Blocks	A	TRUE
	2.4.2 Page Titled	A	TRUE
	2.4.3 Focus Order	A	FALSE
	2.4.4 Link Purpose (In Context)	A	TRUE
	2.4.5 Multiple Ways	AA	FALSE
	2.4.6 Headings and Labels	AA	TRUE
	2.4.7 Focus Visible	AA	FALSE
	2.4.8 Location	AAA	FALSE
	2.4.9 Link Purpose (Link Only)	AAA	FALSE
	2.4.10 Section Headings	AAA	FALSE
2.5 Input Modalities	2.5.1 Pointer Gestures	A	FALSE
	2.5.2 Pointer Cancellation	A	FALSE
	2.5.3 Label in Name	A	FALSE
	2.5.4 Motion Actuation	A	FALSE
	2.5.5 Target Size	AAA	FALSE
	2.5.6 Concurrent Input Mechanisms	AAA	FALSE

Table 2.2: WCAG Operable guidelines with success criteria and conformance level addressed by WAVE

3. Understandable Principle The Table 2.3 provides the list of guidelines under the WCAG Operable principle. There are only 2 criteria addressed out or 17 WCAG successes out of which all of the criteria are of level A.

Guideline	Success Criteria	Conformance Level	Addressed by WAVE
3.1 Readable	3.1.1Language of Page	A	TRUE
	3.1.2Language of Parts	AA	FALSE
	3.1.3Unusual Words	AAA	FALSE
	3.1.4Abbreviations	AAA	FALSE
	3.1.5Reading Level	AAA	FALSE
	3.1.6Pronunciation	AAA	FALSE
3.2Predictable	3.2.1On Focus	A	FALSE
	3.2.2On Input	A	FALSE
	3.2.3Consistent Navigation	AA	FALSE
	3.2.4Consistent Identification	AA	FALSE
	3.2.5Change on Request	AAA	FALSE
3.3Input Assistance	3.3.1Error Identification	A	FALSE
	3.3.2Labels or Instructions	A	TRUE
	3.3.3Error Suggestion	AA	FALSE
	3.3.4Error Prevention (Legal, Financial, Data)	AA	FALSE
	3.3.5Help	AAA	FALSE
	3.3.6Error Prevention (All)	AAA	FALSE

Table 2.3: WCAG Understandable guidelines with success criteria and conformance level addressed by WAVE

4. Robust Principle The Table 2.4 provides the list of guidelines under the WCAG Operable principle. There are only 1 criteria addressed out or 3 WCAG successes which is of type level A conformance.

Guideline	Success Criteria	Conformance Level	Addressed by WAVE
4.1Compatible	4.1.1Parsing	A	FALSE
	4.1.2Name, Role, Value	A	TRUE
	4.1.3Status Messages	AA	FALSE

Table 2.4: WCAG Robust guidelines with success criteria and conformance level addressed by WAVE

Overall, 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 detecting 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. Whereas 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 will help 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 2.5 shows the combined list of issues that is addressed by WAVE Accessibility Tool.

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	3.1.1Language of Page	A
	3.3 Input Assistance	3.3.2Labels or Instructions	A
4.Robust	4.1 Compatible	4.1.2Name, Role, Value	A

Table 2.5: WCAG guidelines with success criteria and conformance level addressed by WAVE

Chapter 3

Methodology

3.1 Research Design & Methodology

This thesis is a quantitative historical research study which aims to analyze the evolution of the web accessibility changes in the last decade. The figure 3.1 highlights the methodology that is applied for the thesis. The study focuses on four popular website categories and each of the category 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 [20]. 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 so that the data is comparable across different website samples. Finally, the 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, analyze the website category effects on

the web accessibility. Website evaluations are also been performed on selected websites in order to provide deeper understanding and provide concrete examples.

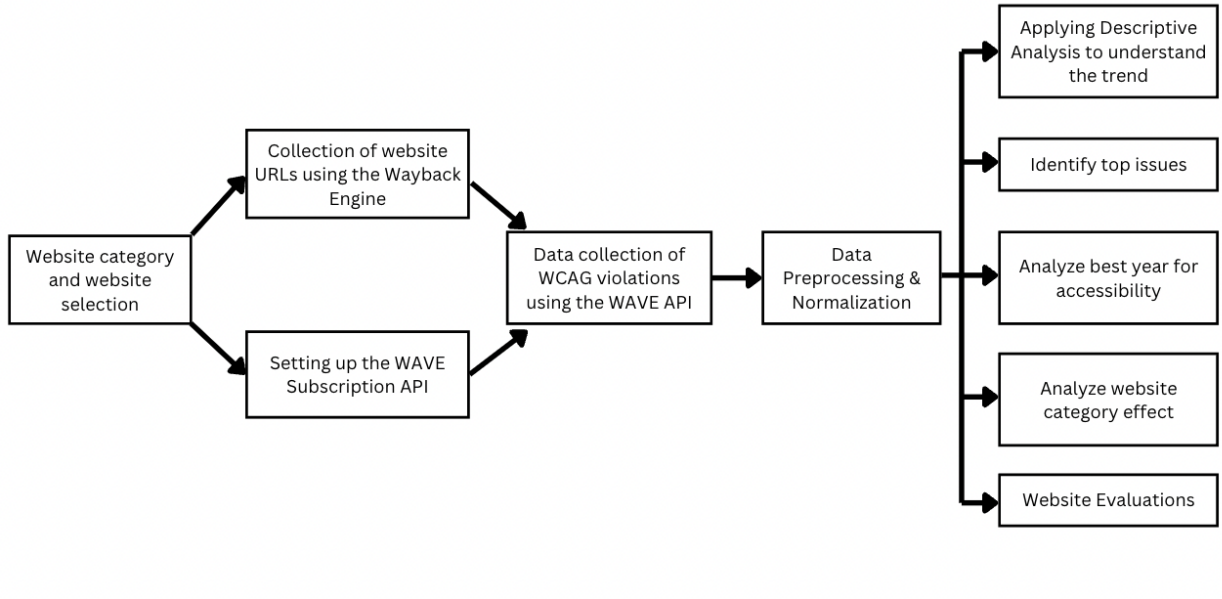


Figure 3.1: Methodology of the research

3.2 Dataset

3.2.1 Data Collection of websites

In order to collect the data on the popular website categories and the popular websites in each category, Semrush.com [21] and SimilarWeb.com [22] was utilized. Semrush and SimilarWeb are digital marketing tool 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 3.1 shows the websites selected for each category as per the traffic.

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://ww.wsj.com/
10	Tagged	https://www.tagged.com/	Costco	https://www.costco.com/	TMZ	https://www.tmz.com/	BBC	https://www.bbc.com/

Table 3.1: Website categories and top 10 websites in each category

The data for the past web content since January 2013 was collected from the web archives using the Wayback Machine. Wayback Machine is a non-profit digital library that archives of 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 in different points of time. There is a search button in the engine in which the current deployed website URL can be entered and the result will appear about the past state of the website. Figure 3.2 shows the snapshot of the Wayback Engine. The data was collected from January 2013 till April 2023 every month and the broken URLs were ignored.

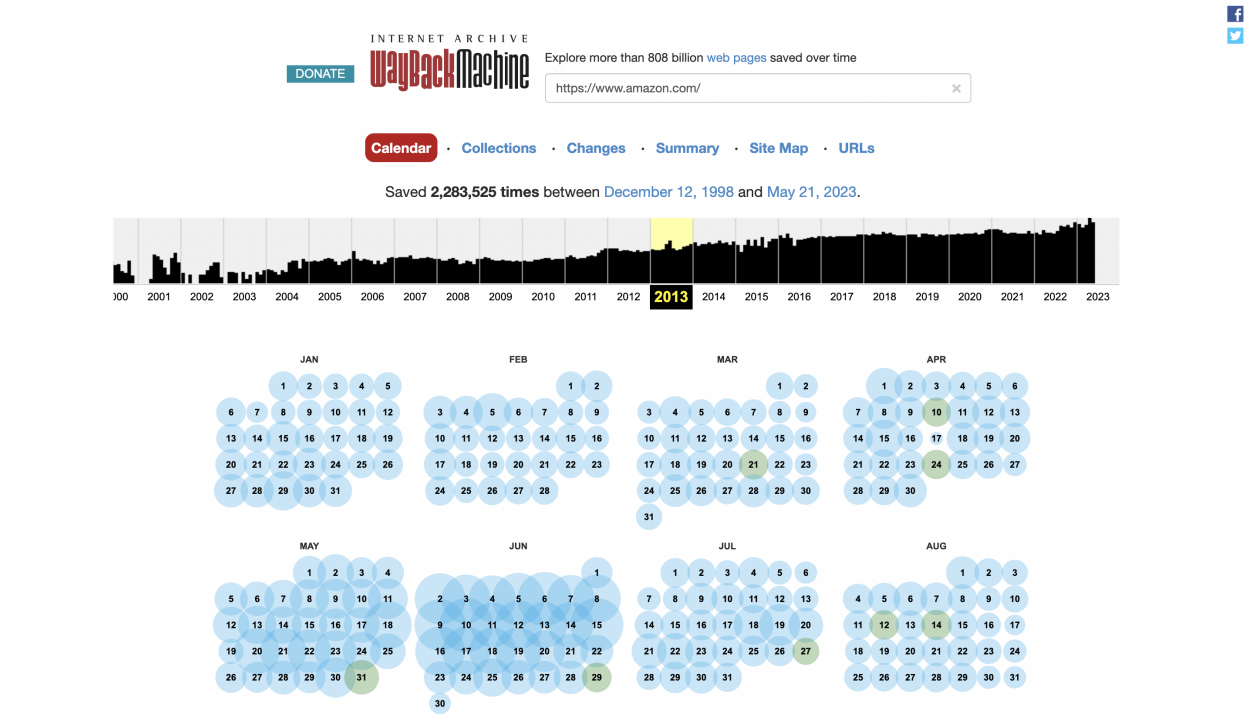


Figure 3.2: Wayback Machine search query result for Amazon website

The websites were eliminated if it falls under any of the following criteria:

1. Information not available for the website since 2013.
2. The website did not exist since 2013.
3. If the website has low number of elements on the homepage i.e. less than 100 elements.
4. If the website has no information available in Wayback Machine.

3.2.2 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 the thesis, the error data is collected which is useful in determining the violations of WCAG found for the webpage. Figure 3.3 is an example of the JSON response by the WAVE API.

```
{
  "status": {
    "success": true,
    "httpstatuscode": 200
  },
  "statistics": {
    "pagetitle": "Example",
    "pageurl": "https://www.example.com/",
    "time": 2.02,
    "creditsremaining": 1388,
    "allitemcount": 65,
    "totalelements": 250,
    "waveurl": "http://wave.webaim.org/report?url=https://www.example.com/"
  },
  "categories": {
    "error": {
      "description": "Errors",
      "count": 2,
      "items": {
        "alt_missing": {
          "id": "alt_missing",
          "description": "Missing alternative text",
          "count": 1
        },
        "link_empty": {
          "id": "link_empty",
          "description": "Empty link",
          "count": 1
        }
      }
    },
    "contrast": {
      "description": "Contrast Errors",
      "count": 43,
      "items": {
        "contrast": {
          "id": "contrast",
          "description": "Very low contrast",
          "count": 43
        }
      }
    },
    "alert": {
      "description": "Alerts",
      "count": 7,
      "items": {
        "h1_missing": {
          "id": "h1_missing",
          "description": "Missing first level heading",
          "count": 1
        },
        "noscript": {
          "id": "noscript"
```

```
        },
        "noscript": {
          "id": "noscript",
          "description": "Noscript element",
          "count": 1
        }
      }
    },
    "feature": {
      "description": "Features",
      "count": 2,
      "items": {
        "alt": {
          "id": "alt",
          "description": "Alternative text",
          "count": 1
        },
        "lang": {
          "id": "lang",
          "description": "Language",
          "count": 1
        }
      }
    },
    "structure": {
      "description": "Structural Elements",
      "count": 5,
      "items": {
        "h2": {
          "id": "h2",
          "description": "Heading level 2",
          "count": 1
        }
      }
    },
    "aria": {
      "description": "ARIA",
      "count": 6,
      "items": {
        "aria_label": {
          "id": "aria_label",
          "description": "ARIA label",
          "count": 3
        },
        "aria_button": {
          "id": "aria_button",
          "description": "ARIA button",
          "count": 3
        }
      }
    }
  }
```


Figure 3.3: WAVE Subscription API example response (Part 2)

3.2.3 Data Normalization

After the data has been collected for all the homepages of the websites, the data normalization is performed on the data in order to make the data standardized. This process will help bring the data to a common scale so that it is easier to make comparison between different websites and categories. In this thesis, the data has been normalized by calculating the error rate of the 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 below formula is used to calculate the percentage of the error rate of a webpage.

$$\text{Error Rate of a webpage} = \frac{\text{Total errors identified by WAVE for a webpage}}{\text{Total elements in a webpage}} * 100$$

The WAVE API JSON response also contains information about the total elements of the webpage along with the error details which makes the error rate more accurate. This normalization of the data helps us in easier comparison across different webpages and identify the trends over the period of 10 years.

3.3 Data Analysis

3.3.1 Examination of the trend of web accessibility violations over the past ten years (RQ1)

To answer this research question, the study will determine whether the websites have advanced to meet the Web Content Accessibility Guidelines (WCAG) by finding the WCAG

violations of each website. The WAVE Subscription API is used to extract the accessibility errors of the websites to further identify the pattern in the number of violations and decide whether there has been a general improvement over time. Descriptive statistical analysis is performed on the normalized data in order to gain understanding of the data and describe the main features of the dataset.

3.3.2 Identification of the most common errors that violate the WCAG guidelines (RQ2)

This question will be studied by examining the most frequent web accessibility errors found on the websites that are found using the WAVE subscription API. This will be found by analyzing a large sample of websites to determine the most common errors that prevent the websites in meeting the WCAG guidelines. This will allow us to identify the areas that need improvement in order to make websites more inclusive and accessible.

3.3.3 Examination of how web accessibility differs depending on the type of website category (RQ3)

The study will investigate the impact of the category of the websites on their accessibility. The websites are classified into four categories- social network, entertainment, e-commerce and news & media. The study will determine which category has the highest number of violations and provide insights using the descriptive statistics of the data. Insights will be obtained by deriving mean error rate, standard deviation and minimum & maximum points in the data in order to understand the data for central tendency and variability.

Chapter 4

Results

4.1 Trend of Web Accessibility over the past decade

The overall trend of violations of Web Accessibility has shown to have decreased over ten years. Figure 4.1 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 4.1 shows the overview of the average error rate each year.

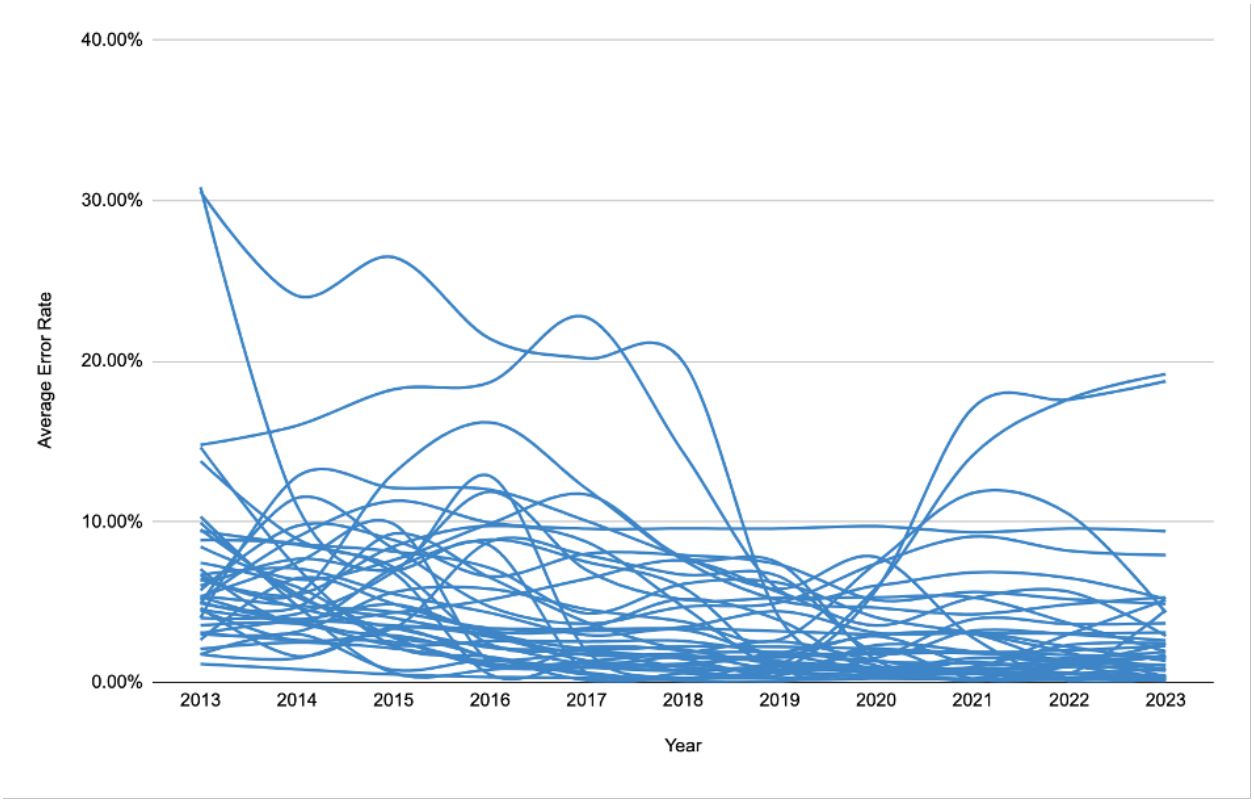


Figure 4.1: Decrease in Web Accessibility issues from January 2013 till April 2023

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 4.1: 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 con-

tinued to decrease from the year 2017 at a slower rate. 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 addressed by the WAVE API. This may have been an effect of the update in the Section 508 laws and regulations published in January 2017 [23]. The US board published a ruling that had major changes that focus on industry alignment which incorporates the WCAG 2.0 in the webpage design beginning January 2018. Additionally, WCAG 2.1 was released in June 2018 which provides more detailed information and examples about the success criteria. This may have raised awareness among the web designers and developers about the importance of the accessibility of the website. We should also note that there may also be other factors that might have an effect on the decrease in the violations such as technological change or increased awareness.

Best year: The best year can be thought of as two different types. To determine the lowest percentage of web accessibility violations, the best year will be identified as the one that has the lowest percentage of violations across all websites. On the other hand, to show improvement over time, the best year would be determined as the largest decrease in number of violations compared to the previous years. The year 2020 and 2019 can be considered 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 years of around 1.12%. This indicates that over time there have been efforts in addressing the issues and the developers are becoming more aware about it. In spite of that, still, improvement needs to be made in order to have webpages that are easier to use for people with disabilities.

4.2 Most common errors that violate the WCAG guidelines

In this section, we will talk about the most common accessibility issues that are addressed by WAVE. Table 4.2 shows the 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%.

No	Error	Average Error Rate
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 4.2: Average error rate of different type of issues addressed by WAVE

Some accessibility issues are easier to resolve than others. For example, “empty heading” is easier to fix than “broken ARIA menu”. Some errors can be fixed quickly by running the automated tool and manually address it, on the other end, some errors require specialized knowledge and a deeper understanding of the design.

These common errors can serve as a starting point for web designers to make their websites more inclusive. It can also be noted that many of the errors can be fixed just by mentioning the alternative texts for the images and empty labels.

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

This section will cover the impact of website categories on web accessibility and this study will identify the category with the highest number of violations and provide insights using the descriptive statistics of the data.

Table 4.3 shows the categories of websites with the highest to lowest average error rates over the decade. Social Networking websites have the highest average error rate whereas e-commerce website’s average is the least. In this section, we will discuss nuanced details about the web accessibility of websites for each category.

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 4.3: Average error rate for website categories

4.3.1 Social Network

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. According to the findings of the data, it was observed that social networking websites have the highest average error rate among the four popular 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.

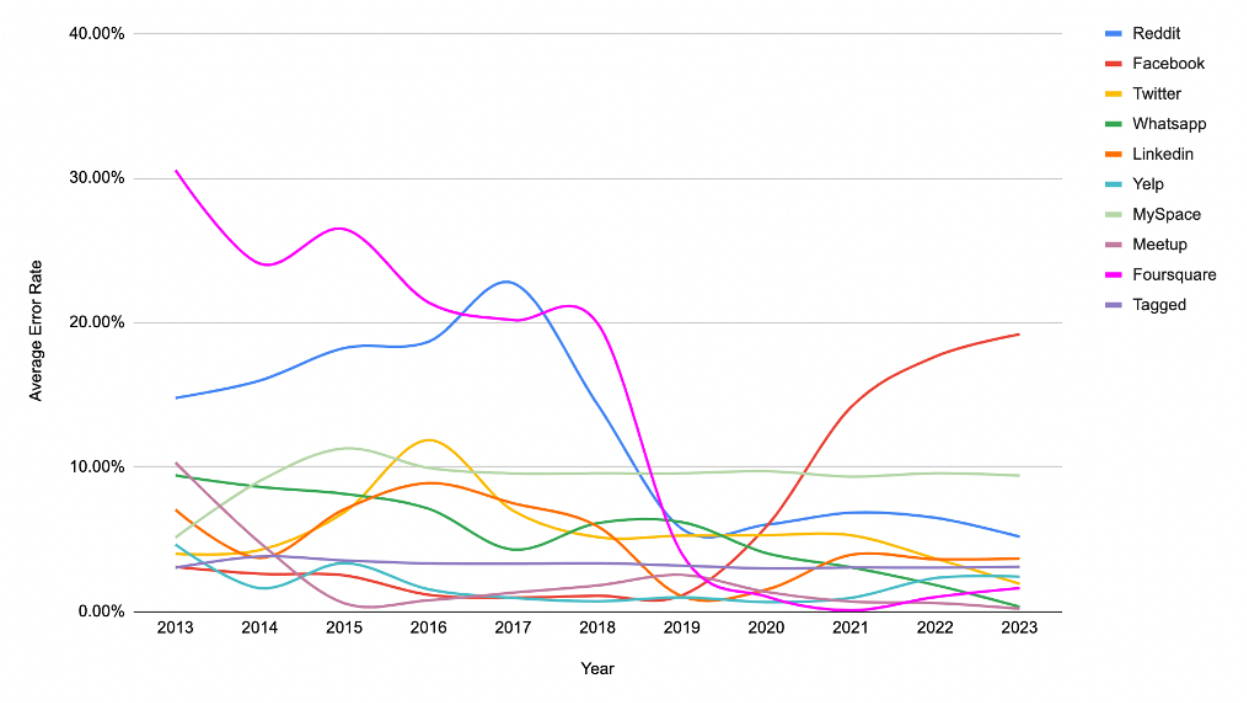


Figure 4.2: Average Error rate for Social Networking websites over the years

Year	**Average Error Rate**
2013	9.20%
2014	7.86%
2015	8.81%
2016	8.47%
2017	7.78%
2018	6.80%
2019	3.97%
2020	3.86%
2021	4.74%
2022	4.99%
2023	4.71%

Table 4.4: Yearly average error rate of Social Networking websites

Website	**Average Error Rate**	**Standard Deviation**
Foursquare	14.45%	11.82%
Reddit	12.73%	6.70%
MySpace	9.28%	2.29%
Twitter	5.74%	3.15%
Whatsapp	5.71%	2.70%
Facebook	5.48%	6.74%
Linkedin	4.99%	3.20%
Tagged	3.26%	0.34%
Meetup	2.41%	3.68%
Yelp	1.80%	1.51%

Table 4.5: Average error rate Standard Deviation of Social Networking websites

Figure 4.2 is a line graph that shows the overall trend of accessibility violations of ten different websites in the social networking category and Table 4.4 and 4.5 provides the numerical data about the statistics of the websites. We can make various observations on the trend that is shown:

1. Overall, there has been a decrease in web accessibility violations at a slower pace over the years which indicates improvement in accessible pages. There has been a decrease of approximately 4.49% in reported issues.

2. Homepage of the Foursquare website has the highest average error rate of 14.45% and the highest variability in the data with 11.82% deviation. If we observe the graph, we can see that Foursquare's website was almost 30% inaccessible in the year 2013. Since 2013, there has been an improvement and at present only approximately 2% of the website is not accessible.

3. Reddit's homepage is the second highest with an average error rate of 12.73%. In 2017, the Reddit website was the most inaccessible with around 23% inaccessibility compared to other years. However, overall, there has been an improvement.

4. MySpace has shown little to no improvement over the decade. With a standard deviation of 2.29%, MySpace's violations have been consistent.

5. Facebook's error rate has increased in the past decade. In the initial years, it was as low as below 4%, however, it has now reached to 20% inaccessibility of the homepage. Since 2019, there has been a considerable increase in the violations.

6. LinkedIn, Tagged, Whatsapp, Meetup and Twitter's error rate has been fluctuating with an overall decrease in the issues.

7. Yelp had the lowest average error rate with only approximately 1.8% of the website inaccessible.

4.3.2 Entertainment

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.

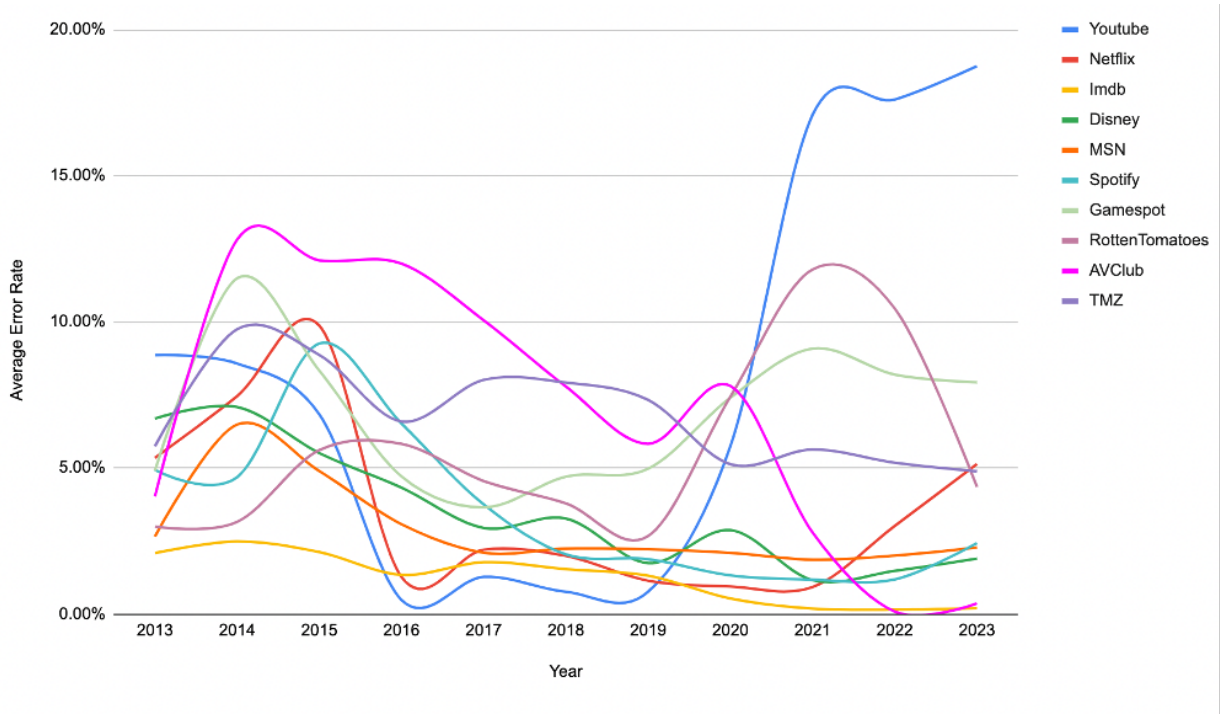


Figure 4.3: Average Error rate for Entertainment websites over the years

Year	**Average Error Rate**
2013	4.82%
2014	7.40%
2015	7.34%
2016	4.61%
2017	4.03%
2018	3.60%
2019	2.98%
2020	4.13%
2021	5.17%
2022	4.93%
2023	4.82%

Table 4.6: Yearly average error rate of Entertainment websites

Website	**Average Error Rate**	**Standard Deviation**
AVClub	7.29%	4.56%
Youtube	7.18%	7.28%
TMZ	6.94%	1.81%
Gamespot	6.78%	3.15%
RottenTomatoes	5.78%	3.36%
Disney	3.65%	2.30%
Spotify	3.63%	3.03%
Netflix	3.47%	3.69%
MSN	2.94%	2.74%
Imdb	1.31%	0.84%

Table 4.7: Average error rate Standard Deviation of Entertainment websites

Figure 4.3 is a line graph that shows the overall trend of accessibility violations of ten different websites in the entertainment category and Table 4.6 and 4.7 provides the numerical data about the statistics of the websites. We can make various observations on the trend that is shown:

1. When comparing the average error rate for the years 2013 and 2023, there has been little to no improvement in the accessibility of entertainment website homepages.

2. AV Club has the highest average error rate of 7.29%. However, AVClub has been showing a decreasing trend overall.

3. YouTube being the second highest overall had very minimal errors in the year 2016-2019. Since 2019, it has been drastically increasing with approximately 18% of the webpage inaccessible today.

4. The homepage of TMZ has 6.94% of inaccessible elements and its standard deviation is only 1.81%. This indicates that the website has been consistent with the errors overall.

5. Gamespot and Rotten Tomatoes have shown fluctuations with little improvement.

6. Disney, MSN and Spotify have shown a decrease in accessibility issues.

7. IMDB had the lowest overall error rate with around 1.31

8. Netflix showed an improvement in the initial years, however, the error rate has been increasing slightly since 2021.

4.3.3 News & Media

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 constant change in the content.

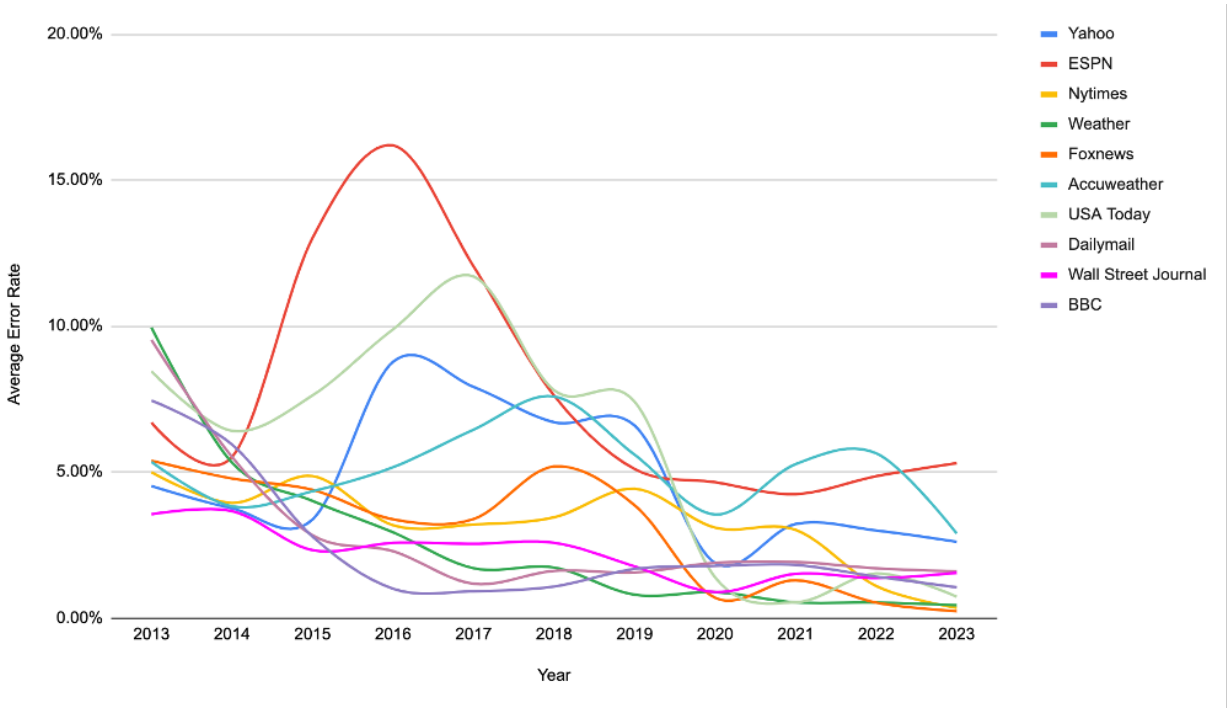


Figure 4.4: Average Error rate for News Media websites over the years

Year	**Average Error Rate**
2013	6.58%
2014	4.86%
2015	4.95%
2016	5.53%
2017	5.10%
2018	4.53%
2019	3.88%
2020	2.07%
2021	2.33%
2022	2.16%
2023	1.67%

Table 4.8: Yearly average error rate of News Media websites

Website	**Average Error Rate**	**Standard Deviation**
ESPN	7.90%	4.59%
USA Today	6.09%	4.29%
Accuweather	5.20%	1.84%
Yahoo	4.89%	2.45%
Nytimes	3.42%	1.73%
Foxnews	3.19%	1.97%
Dailymail	2.95%	2.72%
Weather	2.76%	3.15%
BBC	2.53%	2.22%
Wall Street Journal	2.25%	0.96%

Table 4.9: Average error rate Standard Deviation of News Media websites

Figure 4.4 is a line graph that shows the overall trend of accessibility violations of ten different websites in the News & Media category and Table 4.8 and 4.9 provides the numerical data about the statistics of the websites. We can make various observations on the trend that is shown:

1. Overall, News & Media homepages have shown a decrease in the error rate. From 2013 to 2013 there has been a decrease of approximately 4.91% in the average error rate.
2. Homepage of ESPN has the highest average error rate of 7.90%. In 2016, around 17% of the webpage was inaccessible, which is the highest among all the websites. Since 2016, there has been a decrease overall.
3. Since 2017, overall there has been improvement in the accessibility of the webpages.
4. USA Today, Accuweather and Yahoo's average error rate has been fluctuating. However, there has been an improvement in recent years.

5. Nytimes, Foxnews, DailyMail, Wall Street Journal and BBC have shown overall decreasing and lower error rates.

4.3.4 E-commerce

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 an appealing user interface.

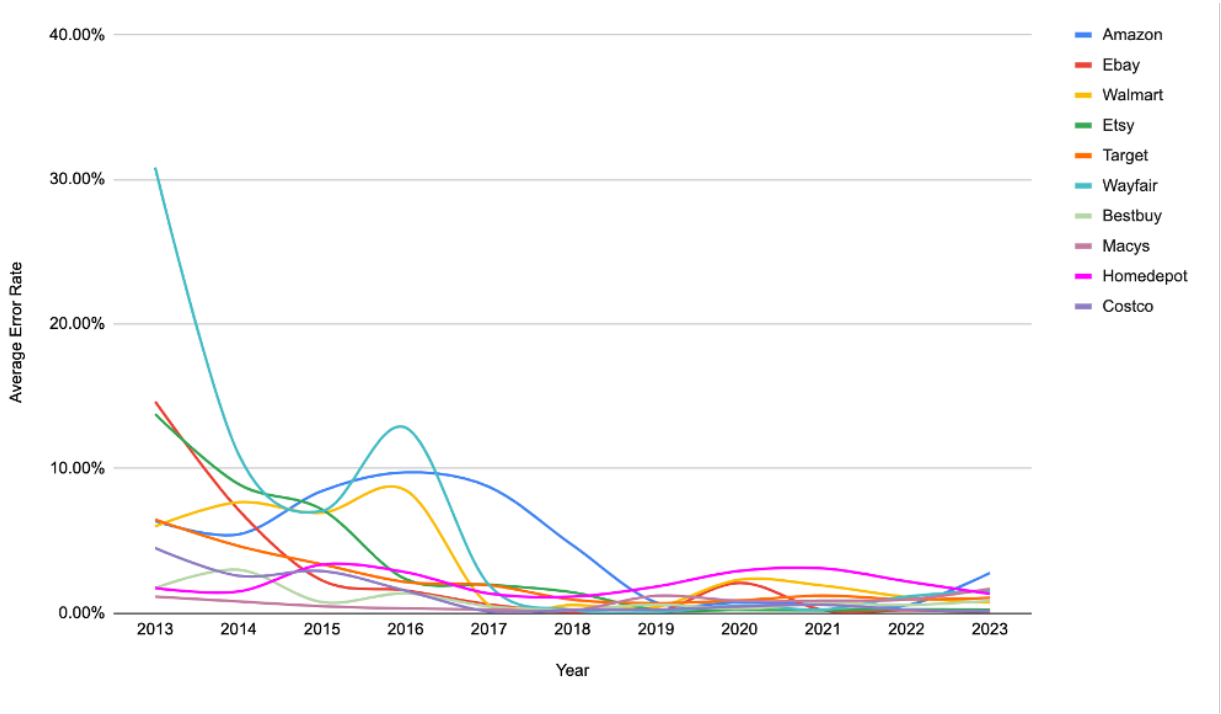


Figure 4.5: Average Error rate for E-commerce websites over the years

Year	**Average Error Rate**
2013	8.73%
2014	5.28%
2015	4.29%
2016	4.33%
2017	1.81%
2018	0.98%
2019	0.62%
2020	1.14%
2021	0.97%
2022	0.82%
2023	1.05%

Table 4.10: Yearly average error rate of E-commerce websites

Website	**Average Error Rate**	**Standard Deviation**
Wayfair	6.43%	10.01%
Amazon	4.62%	4.13%
Etsy	3.55%	4.67%
Walmart	3.52%	3.75%
Ebay	2.80%	4.57%
Target	2.29%	2.02%
Homedepot	2.18%	1.10%
Costco	1.31%	1.60%
Bestbuy	0.98%	1.05%
Macys	0.75%	0.49%

Table 4.11: Average error rate Standard Deviation of E-commerce websites

Figure 4.5 is a line graph that shows the overall trend of accessibility violations of ten different websites in the E-commerce category and Table 4.10 and 4.11 provides the numerical data about the statistics of the websites. We can make various observations on the trend that is shown:

1. E-commerce websites have shown the largest decrease in average error rate from 2013 to 2023 which is around 7.68%. This indicates improvements and efforts have been made towards the accessibility of the websites.

2. Wayfair has the highest average error rate with around 6.43% inaccessibility. Wayfair's homepage was 30% inaccessible in the year 2013, which is the highest among all the websites in the category. This has now reduced to around 2% of inaccessible elements which indicates a lot of improvement over the decade.

3. After the year 2018, the overall average error rate has been consistently less than 3%.

4. Amazon and Walmart show an increase in errors in 2016, which dropped since the year 2017.

5. Macy's average error rate was the lowest among all the websites with around 0.75%.

4.4 Website Evaluations

4.4.1 YouTube (Entertainment)

YouTube is one of the popular and largest streaming platforms which allows people to share and view video content. Over the decade, YouTube has undergone various changes including the accessibility features, user interface and technology. Figure 4.6 shows the overall trend

of YouTube web accessibility violations over the past decade along with the average error rate each year.

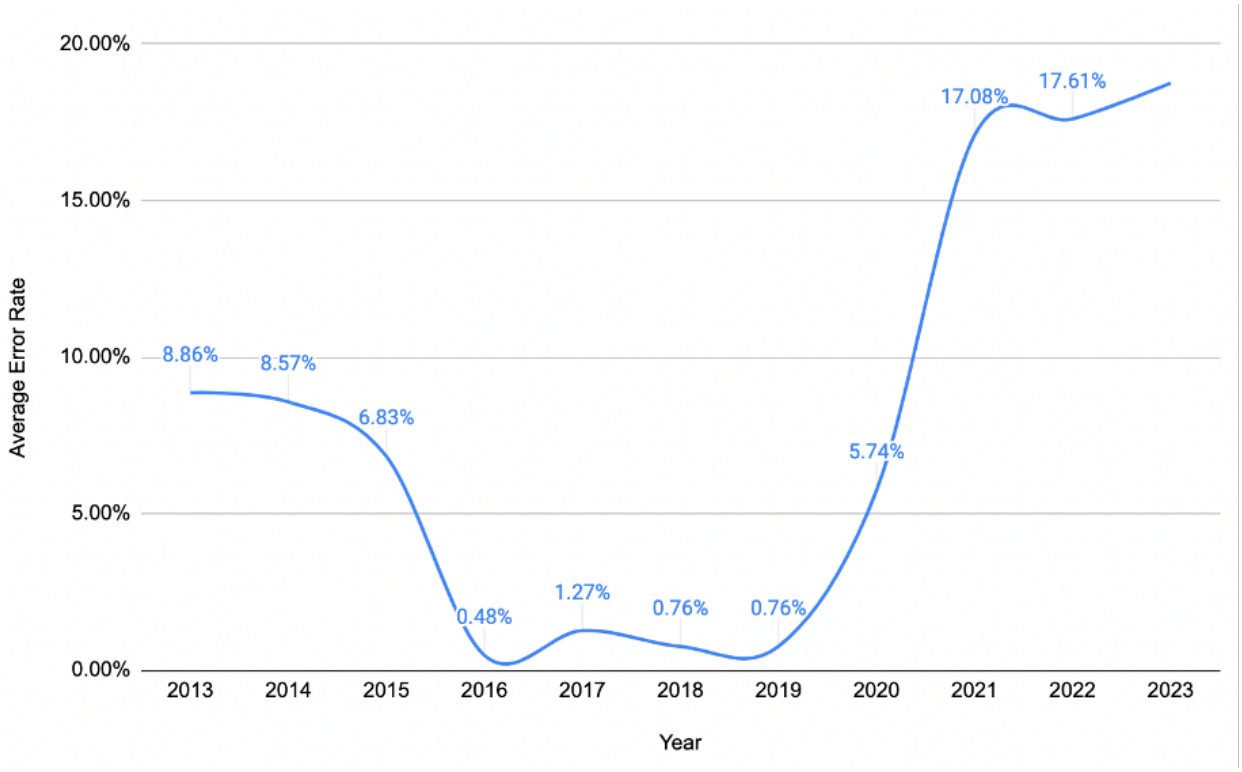


Figure 4.6: YouTube Web Accessibility violations over the years

YouTube also underwent major user interface changes since 2013 which is highlighted in the Figure 4.7. These three figures show the three major UI transformation and redesign for YouTube. This also introduced new elements and alteration of the current functionalities.

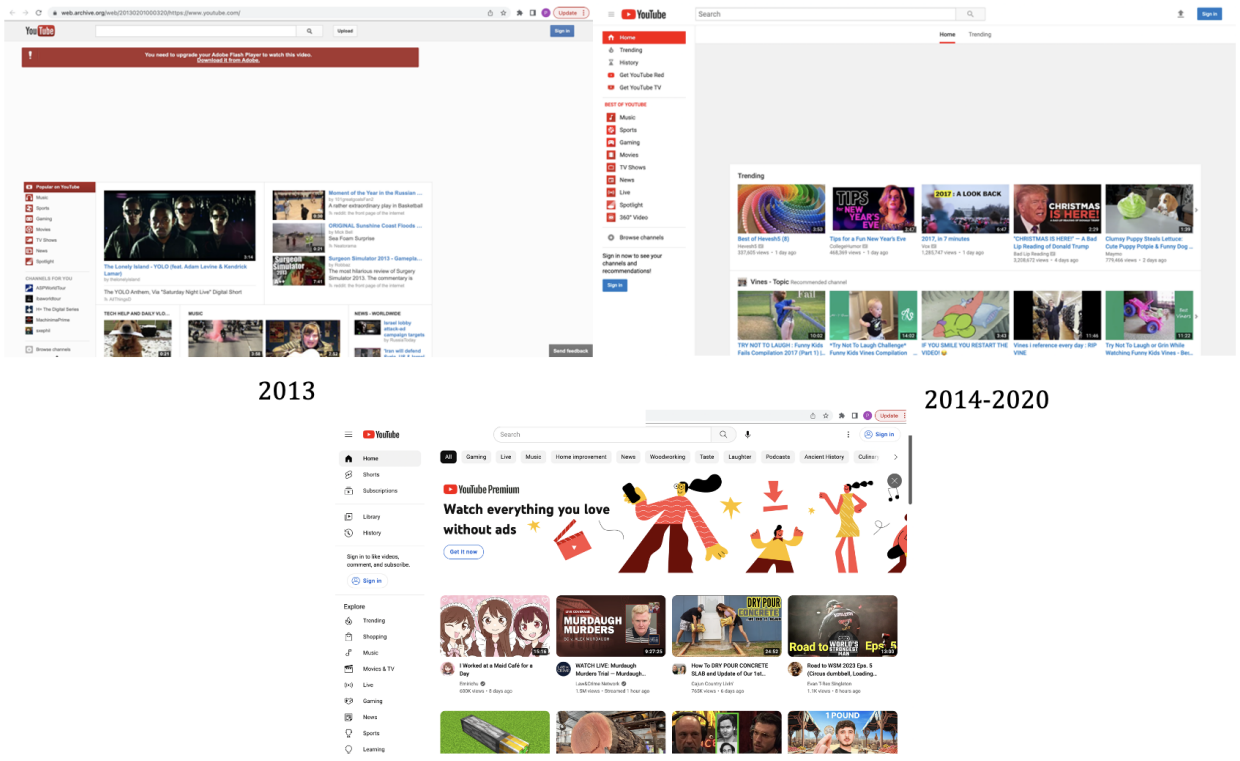


Figure 4.7: YouTube User Interface

Overall trend: According to the data collected, YouTube's average error rate has been fluctuating over the decade. In 2013, the average error rate for YouTube was 8.86% which was then decreased to 0.76% in 2019. This indicates that from year 2013 to 2019 there was overall decrease in the web accessibility violations. However, we can notice that from the year 2020, the average error rate of YouTube website increased which corresponds with the UI redesign in mid 2020. After which the average error rate has increased considerably and it has reached to average error rate of 18.75% in 2023. This suggests that one factor that could have contributed to the YouTube's lower error rate in the early years was simplicity in the design and user interface. The website had relatively straightforward UI in the early years which prioritized the video content over the additional functionalities and the features. The simplicity of the UI and lesser features made the website easier to navigate and access the content using the assistive technology. The complexity of the website has grown over time which may have affected the accessibility of YouTube.

Most common errors: There were 11 different types of errors identified by WAVE API for the YouTube website over the decade which are listed in table 4.12. 'Linked image missing alternative text' was the most common error found with an average error rate of approximately 3.01%. Missing alternative text in images can make it impossible for the people who are blind or have low vision to understand the context. The second most common error was 'Very low contrast' with 1.92% average error rate which can make it difficult to understand the textual information and distinguish between the graphical elements for users with visual impairments.

No	**Error**	**Approximate Average Error Rate**
1	Linked image missing alternative text	3.01%
2	Very low contrast	1.92%
3	Empty heading	0.69%
4	Empty button	0.58%
5	Spacer image missing alternative text	0.32%
6	Missing alternative text	0.24%
7	Empty form label	0.13%
8	Empty link	0.12%
9	Broken ARIA menu	0.12%
10	Image map missing alternative text	0.00%
11	Missing form label	0.00%

Table 4.12: Unique violations of YouTube

Best Year: 2016 was the best year for web accessibility on YouTube with a lowest average error rate of only 0.48%. It is also the year in terms of the improvement over time and with the largest decrease in the number of violations when comparing it with the previous year.

2016 may be identified as the year in which the increased efforts were shown by YouTube to improve accessibility.

4.4.2 Facebook (Social Network)

Facebook is one of the largest social media platforms where users can create profile, interact with each other, upload videos and images. Over the decade, Facebook's homepage has had changes in the accessibility with websites being 3.09% accessible in 2013 to now with over 19.19% inaccessible elements. This evaluation will identify what might be the cause of these violations and identify the type of error Facebook's homepage has. Figure 4.8 shows the trend of accessibility for Facebook over time showing the average error rate for each year.

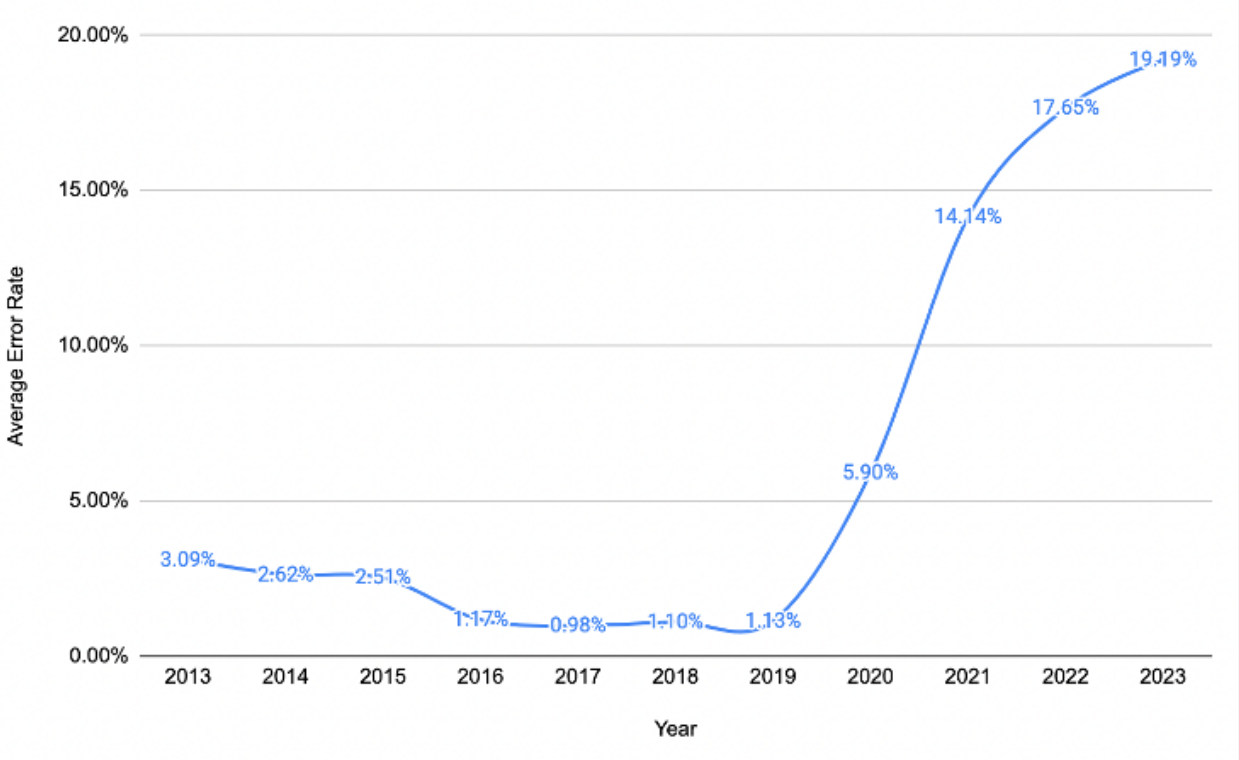


Figure 4.8: Facebook Web Accessibility violations over the years

Facebook also underwent minimal changes in the user interface which is shown in the figure

4.9. It highlights the major changes over the past decade.

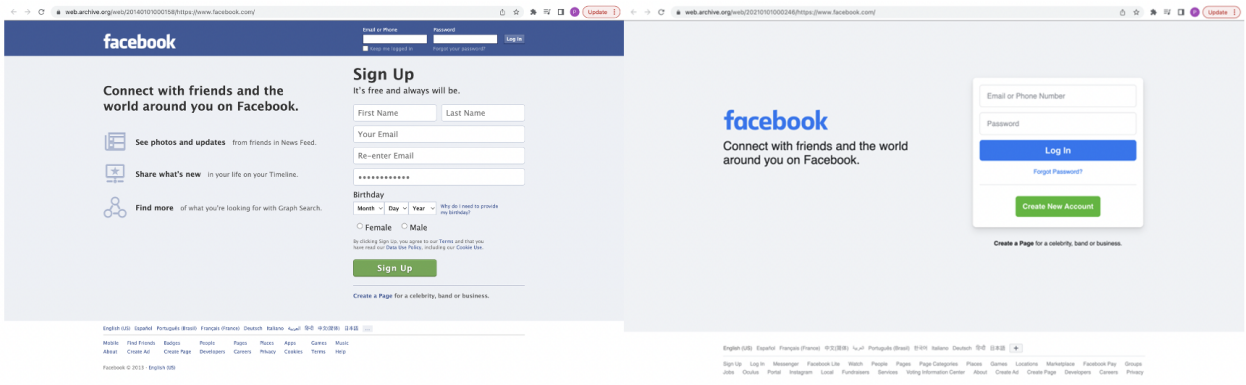


Figure 4.9: Facebook User Interface

Overall trend: According to the data that was found, Facebook had maintained the accessibility of the homepage from 2013 to 2019 and showed an overall decrease in the violations. However, there was a drastic change between the years 2019 and 2023 with an increase of over 18.06% accessibility issues. Figure 4.9 shows the changes that were made in the year 2020 that corresponds with the increment of the errors. Changes in the user interface might have been one of the factors due to which there was an increase in the accessibility issues. Even though the UI looks straightforward and simple, there is 19.19% of inaccessibility. In the next paragraph, we discuss the types of errors that it has and identify the root cause.

Most common errors: Table 4.13 shows the different types of errors that the homepage of Facebook has. The most common error identified is "Very low contrast" with an average error rate of 5.06%. Figure 4.10 shows the "Very low contrast" average error rate each year. If the low contrast errors will be resolved, the website will show an overall improvement in the accessibility for people with low vision. The other errors are less frequent, however, many errors are easier to resolve just by adding alternative text and missing information.

No	Error	Approximate Average Error Rate
1	Very low contrast	5.06%
2	Empty form label	0.16%
3	Missing alternative text	0.14%
4	Empty link	0.08%
5	Broken ARIA menu	0.04%

Table 4.13: Unique violations of Facebook

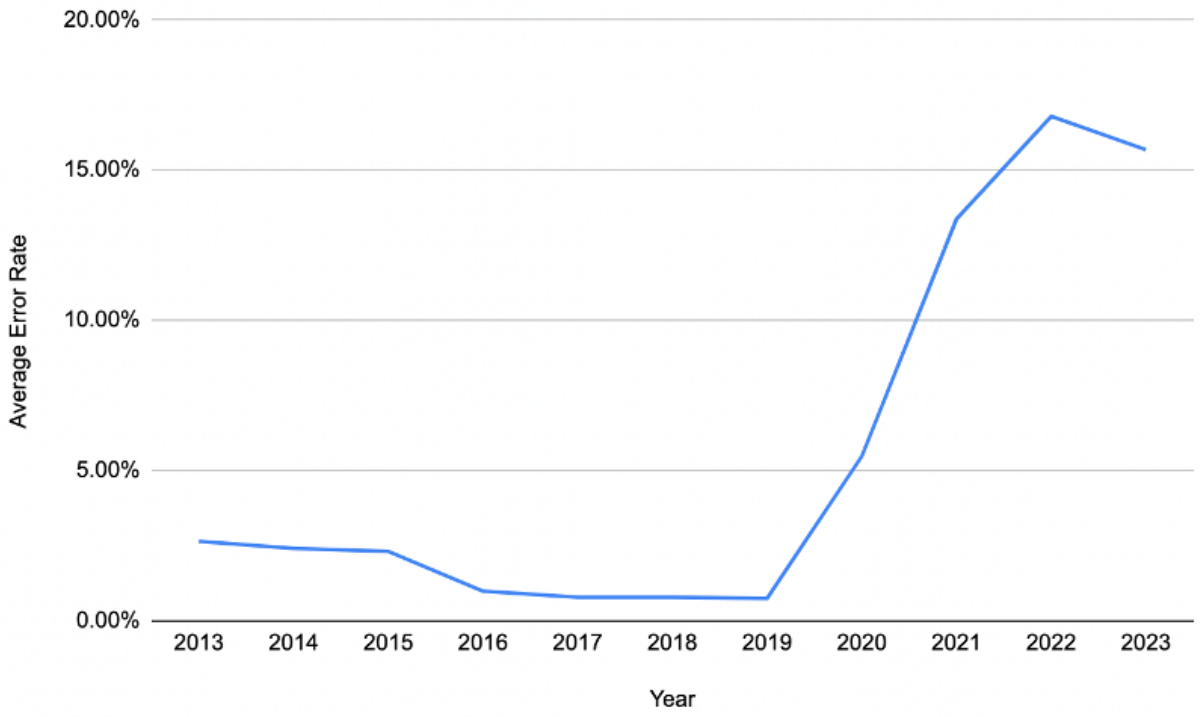


Figure 4.10: "Very low contrast" average error rate

Best Year: The period from 2016 and 2019 can be considered the best year for Facebook's accessibility as it had less than 1.2% of accessibility issues.

4.4.3 Wayfair (E-commerce)

Wayfair is one of the most popular E-commerce websites which sells furniture equipment and home goods online. As shown in table 4.12, Wayfair has the highest average error rate with fluctuating accessibility over the decade. This evaluation will help us identify the issues that the homepage has and what might have affected it. Figure 4.11 shows the trend of accessibility for Wayfair over time showing the average error rate for each year.

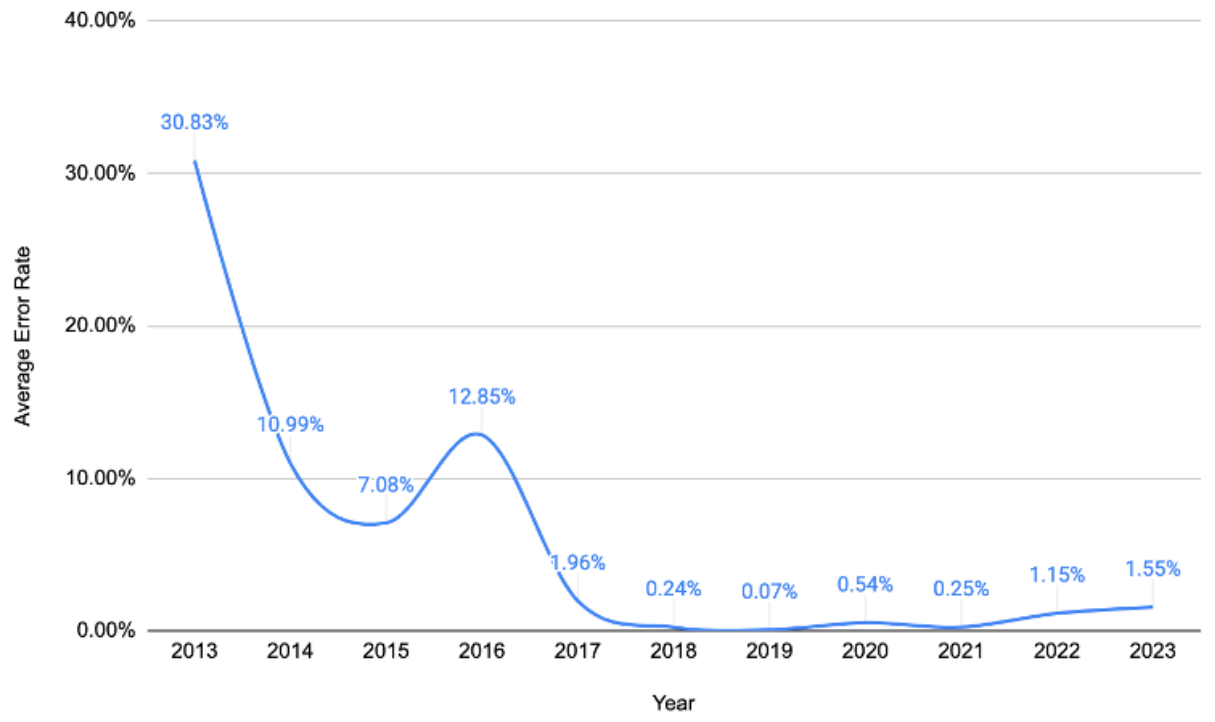


Figure 4.11: Wayfair Web Accessibility violations over the years

Along with changes in the accessibility features over the decade, Wayfair's homepage also underwent various changes in the user interface which is shown in the figure 4.12. It highlights the medium to major changes over the years.

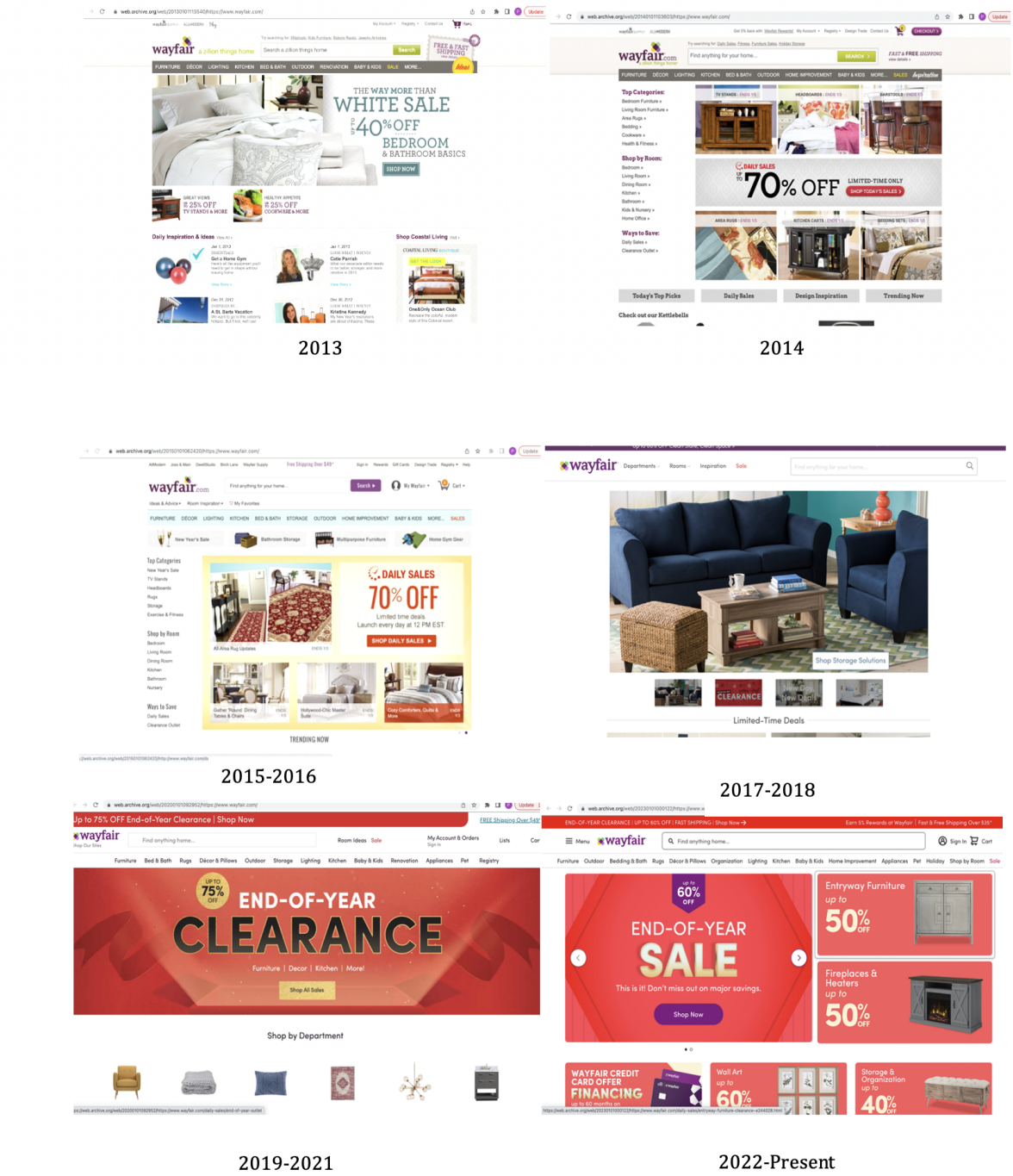


Figure 4.12: Wayfair User Interface

Overall trend: Wayfair has shown an improvement in terms of the WCAG violations on their homepage. In 2013, 30% of the webpage was inaccessible which mostly comprised of low contrast errors. By the year 2023, the webpage has become more accessible with only

1.55% of inaccessible elements. Figure 4.12 also highlights the changes in Wayfair's UI which shows that it underwent multiple UI changes over the past decade which might be the reason of the fluctuating trend between the years 2013 and 2018. Additionally, Wayfair's user UI was complex with on an average of 1788 elements on the homepage which might have been another factor that made it difficult to maintain the accessibility.

Most common errors: Table 4.14 shows the violations that were reported by WAVE API for the homepage of Wayfair. There are 11 different types of errors and the most common error is the "Very Low Contrast" with an average error rate that comprises of the major chunk of the errors. Attention was given to the contrast errors to make the user interface accessible for low vision individuals, which was the reason of decrease in accessibility violations. In January 2013, 463 contrast errors were found which was resolved by April 2023 with 0 contrast errors.

No	**Error**	**Approximate Average Error Rate**
1	Very low contrast	5.09%
2	Missing alternative text	0.41%
3	Linked image missing alternative text	0.30%
4	Empty link	0.28%
5	Spacer image missing alternative text	0.10%
6	Broken ARIA menu	0.08%
7	Missing form label	0.08%
8	Empty heading	0.06%
9	Document language missing	0.02%
10	Empty button	0.02%
11	Empty form label	0.00%

Table 4.14: Unique violations of Wayfair

Best year: The period from 2018 to 2021 can be considered the best years for Wayfair with as less as 0.55% accessibility issues for this year. In terms of improvement when compared to the previous years, year 2014 had shown a decrease of approximately 19.84% of errors when compared to 2013.

4.4.4 ESPN (News & Media)

ESPN is a sports platform that provides sports highlights, news and scores online. ESPN's homepage has the highest average error rate with lot of fluctuations in the violations over the years. Figure 4.13 shows the average error rate from 2013 to 2023 for ESPN.

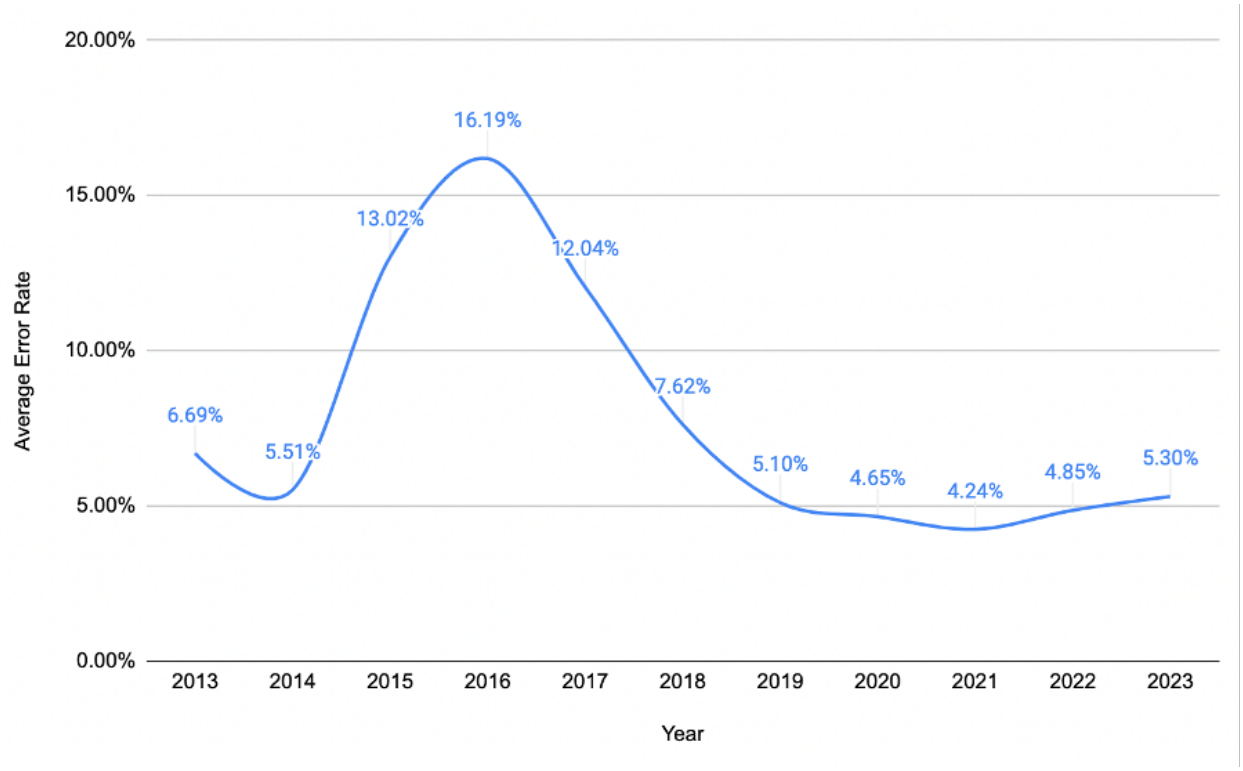
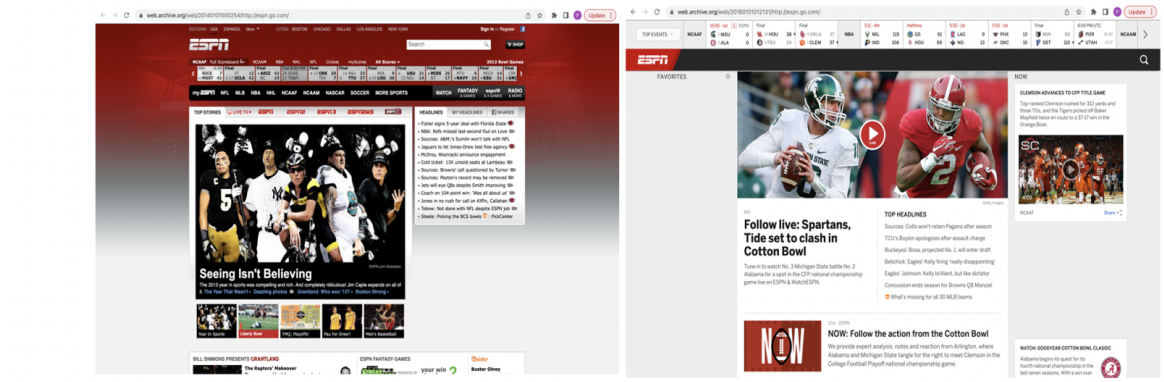


Figure 4.13: Wayfair User Interface

ESPN also underwent user interface changes over the decade with three major changes shown in the figure 4.14.



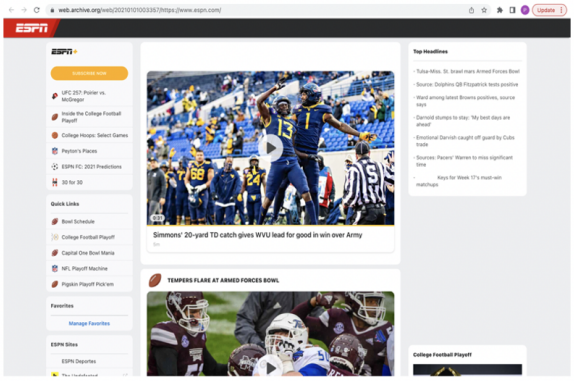


Figure 4.14: ESPN User Interface

Overall trend: Based on the data that is gathered, it is observed that there has been fluctuations year to year with the highest error rate of 16.19% in the year 2016. There was a drastic rise of accessibility issues from the year 2014 to 2016. Since 2016, error rate has been falling gradually with a small rise after 2021. ESPN had underwent changes in the UI along with decrease in accessibility issues as well. Attention was given to the accessibility features of the website along with making the UI more intuitive and responsive. ESPN also has complex UI structure with on an average 2670 elements on the homepage, this might also be the reason for the highest average error rate.

Most common errors: ESPN's homepage reported 13 different types of errors that are shown

in the table 4.15. Like Facebook and Wayfair, "very low contrast" was the most common error with an average of 3.84% error rate. In 2013, 5.83% of the website was inaccessible and in 2023, it is 3.15% of inaccessibility due to the low color contrast. The average error rate for the very low contrast error has been shown in the figure 4.15. The overall trend shows a decrease in this type of error, however, it is very steadily decreasing. Efforts still needs to be taken in order to resolve this issue. Many of the other errors found on ESPN's homepage can be fixed by providing alternative texts and missing information.

No	**Error**	**Approximate Average Error Rate**
1	Very low contrast	3.84%
2	Missing alternative text	1.86%
3	Empty link	1.21%
4	Linked image missing alternative text	0.67%
5	Empty button	0.15%
6	Missing form label	0.12%
7	Empty heading	0.02%
8	Language missing or invalid	0.02%
9	Image button missing alternative text	0.01%
10	Image button missing alternative text	0.01%
11	Broken ARIA reference	0.00%
12	Image map area missing alternative text	0.00%
13	Empty form label	0.00%

Table 4.15: Unique violations of ESPN

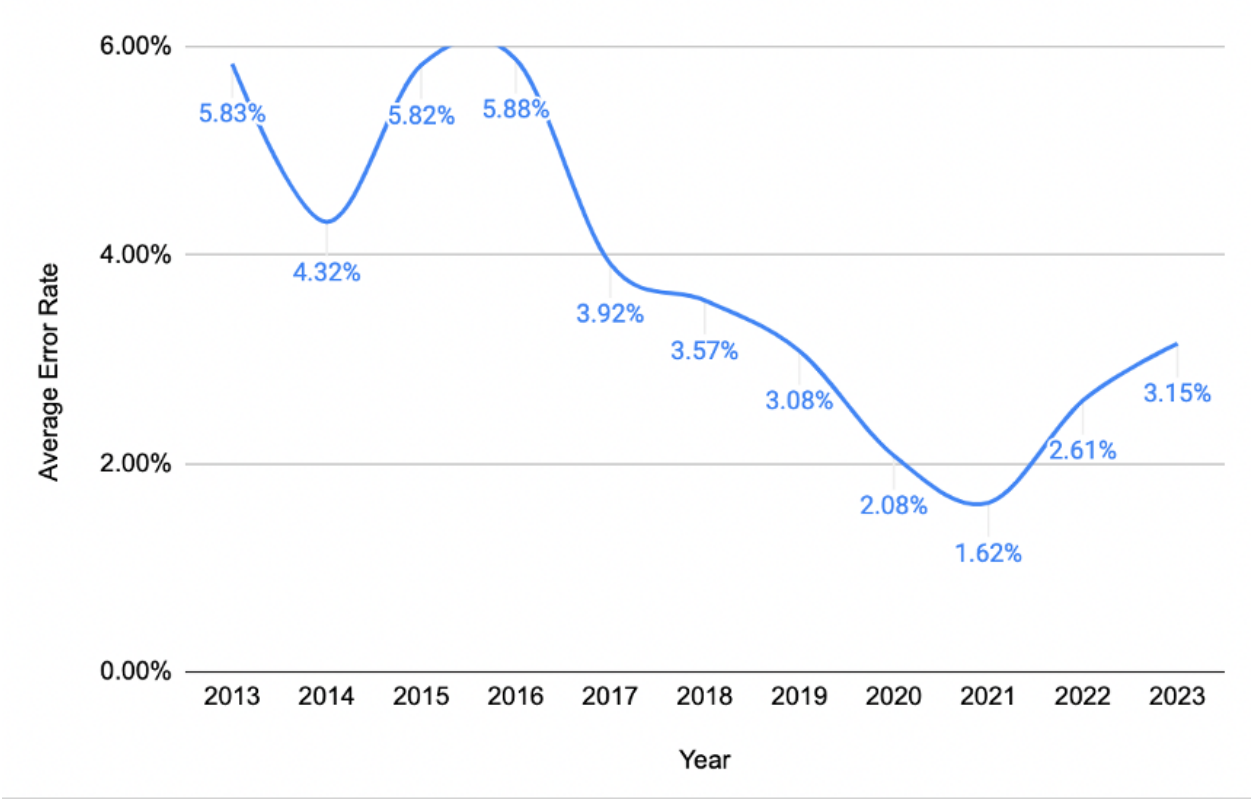


Figure 4.15: ESPN "very low contrast" average error rate

Best year: The best year for ESPN in terms of the lowest errors found was 2021 with a rate of 4.24%. In terms of improvement when compared to the previous years, the best year is 2018 with a decrease of 4.42% error rate compared to 2017.

Chapter 5

Limitations & Future Direction

5.1 Limitations

This thesis does provide insights into the overall trend of 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 the data about the WCAG violations of the website. However, manual testing is always required along with automated testing because many errors are not identified by the automated tool which requires 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 have not been addressed, only the most impactful and relevant issues have been reported.

3. Website category not aligned with homepage content: Homepages of the website may not purely be as their categorization. For example, WhatsApp's homepage is not

purely a social network, whereas Reddit's homepage is a social network.

5.2 Future Direction

The following recommendations can be made for the future work of this thesis:

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 the users who actually experience these issues can provide a more nuanced and detailed understanding of the web accessibility issues that people with disabilities face.

3. Impact of technology: New technologies have been developed in the past decade to make the user interface more appealing and responsive. Examination and comparison between the technology and the WCAG violations can help understand the affect of the emerging technologies on web accessibility.

Chapter 6

Conclusion

In conclusion, this thesis presents historical research along with a comprehensive investigation of the evolution of web accessibility over the course of the previous ten years. This study conducts a study of 40 different websites across four website categories that are widely known. In order to acquire the WCAG violations data, we made use of the WAVE subscription API and the Wayback Machine. Over the course of the past ten years, there has been evidence that indicates 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 a webpage. The study looked into the most common accessibility issues on the web as well as the year that was the most successful for accessibility. In addition to this, the thesis presents an examination of the impact that the different website categories have on the number of web accessibility violations. Website evaluations on websites such as YouTube, Facebook, Wayfair, and ESPN help those reading understand the accessibility aspects of these websites in greater depth, as well as how these companies have responded to accessibility violations over a period of time. 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 & supports attempts aimed at enhancing everyone's

access to and inclusion on the internet.

Bibliography

[1] World Health Organization. https://www.who.int/news-room/fact-sheets/detail/disability-and-health. Accessed March 2023.

[2] World Wide Web Consortium. https://www.w3.org/WAI/fundamentals/accessibility-intro/. Accessed March 2023.

[3] Web Assistive Technologies. http://sites.stedwards.edu/accessibility/web-assistive-technologies/. Accessed March 2023.

[4] World Health Organization. https://www.who.int/news-room/fact-sheets/detail/assistive-technology. Accessed March 2023.

[5] National Network. https://adata.org/learn-about-ada. Accessed March 2023.

[6] Web Content Accessibility Guidelines (WCAG) 2.1. https://www.w3.org/TR/WCAG21/. Accessed March 2023.

[7] Hayfa Abu Addous, Mohd Zalisham, and Nurlida Basir. Web accessibility challenges. *International Journal of Advanced Computer Science and Applications*, 7, 2016.

[8] The WebAIM Million. https://webaim.org/projects/million/2022. Accessed March 2023.

[9] Patricia Acosta-Vargas, Sergio Luján-Mora, and Luis Salvador-Ullauri. Evaluation of the web accessibility of higher-education websites. pages 1–6, 2016.

[10] Muhammad Akram and Rosnafisah Bt Sulaiman. 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.

[11] Gaurav Agrawal, Dinesh Kumar, Mukesh Singh, and Dhaval Dani. Evaluating accessibility and usability of airline websites. 1045, 2019.

[12] Stephanie Hackett, Bambang Parmanto, and Xiaoming Zeng. A retrospective look at website accessibility over time. *Behaviour & Information Technology*, 24(6):407–417, 2005.

[13] Stephanie Hackett and Bambang Parmanto. A longitudinal evaluation of accessibility: Higher education web sites. *Internet Research*, 15:281–294, 2005.

[14] Osama Sohaib and Kyeong Kang. Assessing web content accessibility of e-commerce websites for people with disabilities. 2016.

[15] Abid Ismail and K. S. Kuppusamy. Accessibility of indian universities' homepages: An exploratory study. *Journal of King Saud University - Computer and Information Sciences*, 30(2):268–278, 2018.

[16] Web Content Accessibility Guidelines (WCAG) 2.0. `https://www.w3.org/TR/WCAG20/`. Accessed March 2023.

[17] WAVE Subscription API. `https://wave.webaim.org/api/`. Accessed March 2023.

[18] WCAG-WAVE Mapping. `https://docs.google.com/spreadsheets/d/1oSyK4QiyHf1zx-xrc4P9M7efYVv0vohNxVWjeHan8Iw/edit#gid=902071383`. Accessed March 2023.

[19] Helen Petrie, Fraser Hamilton, and Neil King. Tension, what tension?: Website accessibility and visual design, 2004.

[20] Wayback Machine. `https://web.archive.org/`. Accessed March 2023.

[21] Semrush. `https://www.semrush.com/`. Accessed March 2023.

[22] SimilarWeb. `https://www.similarweb.com/`. Accessed March 2023.

[23] The Section 508 Update. `https://www.section508.gov/blog/accessibility-news-the-section-508-Update/`. Accessed March 2023.