Accessibility Evaluation of Web Systems for People with Visual Impairments: Findings from a Literature Survey

Conference Paper · December 2024

DOI: 10.1145/3702038.3702090

José Cezar de Souza Filho
Polytechnic University of Hauts-de-France

Anna Beatriz Dos Santos Marques
Federal University of Ceará

Ingrid Teixeira Monteiro
Federal University of Ceará

Accessibility Evaluation of Web Systems for People with Visual Impairments: Findings from a Literature Survey

Aline Souza
Federal University of Ceará - UFC
Quixadá, Brazil
alinearaujos@alu.ufc.br

José Cezar de Souza Filho
Univ. Polytechnique Hauts-de-France
LAMIH, UMR CNRS 8201, F-59313
Valenciennes, France
josecezar.juniordesouzafilho@uphf.fr

Carla Bezerra
Federal University of Ceará - UFC
Quixadá, Brazil
carlailane@ufc.br

Victor Anthony Alves
Federal University of Ceará - UFC
Quixadá, Brazil
victorpa@alu.ufc.br

Lara Lima
Federal University of Ceará - UFC
Quixadá, Brazil
laragabrielly@alu.ufc.br

Anna Beatriz Marques
Federal University of Ceará - UFC
Russas, Brazil
beatriz.marques@ufc.br

Ingrid Teixeira Monteiro
Federal University of Ceará - UFC
Quixadá, Brazil
ingrid@ufc.br

ABSTRACT

Accessibility evaluation is essential to determine how accessible a Web system is so that any user can access its content regardless of their limitations. In this context, this work presents a literature survey focused on the accessibility of Web systems, targeting users with visual impairments, emphasizing the importance of including accessibility from the initial phases to the completion of system development. Based on systematic procedures, we seek to provide a current view of the resources available in the technical literature for evaluating Web accessibility, with the primary goal of identifying and bringing together a variety of attributes, methods, metrics, and tools for accessibility improvement of Web systems. We analyzed 1245 papers in the literature; among them, 52 studies demonstrated available resources for evaluating systems by developers. We found that manual evaluation, evaluation with expert users, evaluation with end-users, and heuristic evaluation are the most recurrent accessibility evaluation types. We also cataloged 22 automatic tools, 16 attributes, 15 manual or automatic approaches, and seven heuristics for evaluating accessibility in Web systems. From our findings, it is possible to observe which resources have been used to maintain compliance with accessibility standards, helping professionals to include accessibility in their projects.

CCS CONCEPTS

• **Human-centered computing → Accessibility design and evaluation methods**; • **Information systems → Web applications**; • **General and reference → Evaluation**.

KEYWORDS

web accessibility, evaluation of web systems, visual impairment

ACM Reference Format:
Aline Souza, José Cezar de Souza Filho, Carla Bezerra, Victor Anthony Alves, Lara Lima, Anna Beatriz Marques, and Ingrid Teixeira Monteiro. 2024. Accessibility Evaluation of Web Systems for People with Visual Impairments: Findings from a Literature Survey. In *XXIII Brazilian Symposium on Human Factors in Computing Systems (IHC 2024), October 7–11, 2024, Brasília - DF, Brazil.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3702038.3702090

1 INTRODUCTION

The increasing growth of the large World Wide Web (WWW) information network has been generating observable impacts on society. According to a study in 2019, more than 4.39 billion users have access to the Internet, and this number is growing by around 9% yearly [1]. Nowadays, the Internet provides content that supports its users in getting access to everyday services, for instance, computer-mediated social interaction and communication, education, citizenship, and healthcare, as well as commercial applications, either public or private. Therefore, it becomes necessary that anyone is able to access these contents, including users with disabilities.

The World Health Organization (WHO) estimates in 2021 that 1.3 billion people on the planet, around 16% of the world's population, identify themselves with some disability type [68]. In Brazil, according to a study by the Brazilian Institute of Geography and Statistics (IBGE) [33], the number of people with some type of disability was around 45.62 million (around 23% of the Brazilian population) since 2010, including visual, hearing, motor, or cognitive impairment. On July 6, 2015, the Brazilian Law on the Inclusion of Persons with Disabilities (nº 13.146)[1] was sanctioned, being one of the greatest advances regarding the rights of people with disabilities. The Law is "intended to ensure and promote, under conditions of equality,

[1]https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm

IHC 2024, October 7–11, 2024, Brasília - DF, Brazil

the exercise of fundamental rights and freedoms by people with disabilities, aiming at their social inclusion and citizenship" [20].

Designing interactive systems to ensure the social inclusion of people with disabilities has been a major task for software development companies. Barbosa et al. [12] highlight that considering accessibility in system design allows more people to acquire benefits from the system instead of only a specific group. They define *accessibility* as "a quality criterion related to a user's ability to interact with a system without its interface presenting obstacles to its use or access to its information" [12, p. 35].

Pressman and Maxim [53] point out that system developers face difficulties in choosing what can be used to evaluate their systems, as each evaluation depends, for instance, on the target audience, evaluation goal, and context of use. Silva and Rodrigues [55] carried out an exploratory study with 25 system developers to identify why they did not consider Web accessibility in their projects. The main causes were the lack of supervision by competent bodies and the lack of interest by developers and business owners in studying, implementing, and selling projects with this criterion included.

To improve the development of accessible Web systems, the World Wide Web Consortium (W3C) created the Web Accessibility Initiative (WAI) that seeks to develop strategies, standards, and support resources to help make the Web more accessible to people with disabilities [22]. Among the WAI standards, the Web Content Accessibility Guidelines (WCAG) provide a set of principles, success criteria, techniques, and common flaws, which were identified through the cooperation between organizations and individuals worldwide [63]. The Brazilian National Government created the Electronic Government Accessibility Model (eMAG), a set of recommendations to guide the development and adaptation of the Federal Government's digital content to ensure quality access for the general public [16].

Despite the Web accessibility standards and guidelines, system developers have difficulty implementing and evaluating accessibility in their projects [45]. In some cases, systems are not evaluated or are only evaluated in a summative manner (e.g., after the project ends), which can generate rework for the development team. Hence, it is important to raise awareness among developers and project leaders to include accessibility from the planning activities to the final phase of their projects.

In this paper, we aim to investigate what is currently used to evaluate the quality-in-use with a focus on Web accessibility characteristics. We observed that the research available in the scientific literature addresses accessibility for different audiences and makes an association between one or more types of disabilities. In this way, we carried out a literature survey to gather research focused on accessibility for visual impairment due to the amount of work on this topic. Therefore, we provide an updated view of research papers that evaluate accessibility in interactive systems for visually impaired users. To this end, we cataloged a set of available resources to perform accessibility evaluation in Web systems, thus supporting the quality-in-use improvement of these systems.

2 ACCESSIBILITY IN WEB SYSTEMS

In this Section, we present the concepts of the WCAG and eMAG accessibility guidelines that we used in our survey.

2.1 WCAG

The WCAG guidelines were developed by W3C and first published in 2008. They are primarily defined to support system developers, Web accessibility evaluators, and other practitioners who need a standard to ensure accessible use of Web content. WCAG 2.2 is made up of 13 guidelines categorized into four principles that must be followed [63]: *Perceivable* means that information should be presented to users in a way they can easily perceive them; *Operable* to ensure that user interface components and navigation are functional; *Understandable* describes that information presented in the user interface and its operationalization should be easy to understand; and *Robust* to ensure that available content are interpretable by a wide range of users, including assistive technologies. Table 1 presents the principles and summarized guidelines of WCAG 2.2.

Table 1: WCAG 2.2 Summarized Guidelines [64]

Principle	Guidelines
1. Perceivable	1.1. Provide text alternatives for non-text content. 1.2. Provide subtitles and other alternatives for multimedia. 1.3. Create content that can be presented in different ways, including through assistive technologies, without losing information and structure. 1.4. Make it easier for users to see and hear content.
2. Operable	2.1. Make all functionality available from a keyboard. 2.2. Give users enough time to read and use the content. 2.3. Do not use content that causes seizures or physical reactions. 2.4. Help users navigate and find content. 2.5. Make it easier to use other input devices beyond the keyboard.
3. Understandable	3.1. Make text readable and understandable. 3.2. Make content appear and operate predictably. 3.3. Help users avoid and fix error.
4. Robust	4.1. Maximize compatibility with current and future user tools.

WCAG 2.2 has three levels of compliance: Level A, Level AA, and Level AAA. *Level A* corresponds to the lowest level and ensures minimum accessibility on the Web. *Level AA* covers the requirements of most user groups, addressing the most common access barriers. Finally, *Level AAA* is the highest accessibility standard, requiring compliance with all WCAG accessibility features [4]. There is an initiative to popularize WCAG with easy content in the form of cards to be used by both industry and academia[2].

2.2 eMAG

The eMAG model's first version was published in 2005 to support the development of websites/portals and the adaptation of accessible digital content of the Brazilian Federal Government, thus ensuring information access to any user, regardless of their physical-motor abilities as well as the cultural and social context to which they belong [16]. The latest version (3.1) of eMAG was published in April 2014. It provides recommendations to guide the easy implementation of virtual accessibility in a standardized way and consistent with Brazilian needs. These recommendations were defined based on the WCAG guidelines (Section 2.1) through research carried out within the scope of the virtual accessibility project, in which people with disabilities participated [16]. The eMAG model provides a total of 45 accessibility recommendations categorized into six sections according to the implementation needs, as follows:

- **Markup**: Codes programmed to indicate the structure and presentation of elements in electronic documents;

[2]https://guia-wcag.com/en/, a set of toolkits also available in Portuguese at http://acessibilidadetoolkit.com/.

- **Behavior (Document Object Model – DOM)**: Tree representation of elements in HTML/XML documents for dynamic manipulation;
- **Content/Information**: Data and resources provided in electronic documents to convey a message;
- **Presentation/Design**: Visual and aesthetic aspects that affect the appearance of an electronic document;
- **Multimedia**: Content elements, such as audio and video, that enrich the user experience;
- **Form**: Set of interactive fields in an electronic document for inserting and manipulating data.

The eMAG model supports system developers in creating accessible websites with different user groups in mind since its process and recommendations make it possible to develop a website that meets the needs of a target audience with some disability [16]. The eMAG principles and recommendations are presented in Table 2.

Table 2: eMAG Sections and Recommendations [16]

Section	Recommendations
1. Markup	1.1. Respect Web standards. 1.2. Organize HTML code logically and semantically. 1.3. Correctly use header levels. 1.4. Order reading and tabulation in a logical and intuitive manner. 1.5. Provide anchors to go directly to a block of content. 1.6. Do not use tables for diagramming. 1.7. Separate adjacent links. 1.8. Divide the information areas. 1.9. Do not open new instances without the user's request.
2. Behavior (DOM)	2.1. Make all page functions available via keyboard. 2.2. Ensure that programmable objects are accessible. 2.3. Do not create pages with frequent automatic updates. 2.4. Do not use automatic page redirection. 2.5. Provide an alternative to modify the session time limit. 2.6. Do not include situations with screen flashing. 2.7. Ensure user control over content changes.
3. Content/Information	3.1. Identify the primary language of the page. 3.2. Inform language change in the content. 3.3. Provide a descriptive and informative title for the page. 3.4. Inform the user about their location on the page. 3.5. Describe links clearly and succinctly. 3.6. Provide text alternatives for website images. 3.7. Use image maps in an accessible way. 3.8. Make documents available in accessible formats. 3.9. Use titles and summaries appropriately in tables. 3.10. Associate data cells with header cells. 3.11. Ensure reading and understanding of information. 3.12. Explain acronyms, abbreviations, and unusual words.
4. Presentation/Design	4.1. Minimum contrast value between background and foreground. 4.2. Do not use only color or other sensory characteristics to differentiate elements. 4.3. Allow resizing without loss of functionality. 4.4. The focused element should be visually evident.
5. Multimedia	5.1. Provide an alternative to video. 5.2. Provide an alternative to audio. 5.3. Offer audio description for pre-recorded video. 5.4. Provide audio control for sound. 5.5. Provide animation control.
6. Form	6.1. Text alternative for form image buttons. 6.2. Associate tags with their fields. 6.3. Establish a logical navigation order. 6.4. Do not automatically cause context changes. 6.5. Provide instructions for data input. 6.6. Identify and describe data input errors and confirm the information sending. 6.7. Group form fields. 6.8. Use specific security strategies instead of CAPTCHA.

3 RELATED WORK

Several recent research studies in the literature have used different methodologies to address digital accessibility. Certain studies focus specifically on users with visual impairments, whereas others cover a broader scope.

Bi et al. [14] applied a qualitative and quantitative approach to investigate the challenges and benefits of incorporating accessibility into software design and development. To do this, they gathered data from 15 interviews and 365 responses from a personal opinion survey in 26 countries, spanning the five global continents, to understand the practitioners' perceptions regarding accessible design and development in practice. The authors presented empirical evidence of practitioners' opinions, highlighting the key obstacles they face regarding accessibility. In a practical and complementary way, AlMeraj et al. [5] evaluated the current state of accessibility in higher education institutions' websites, using a quantitative approach to evaluate the websites regarding their compliance with the WCAG 2.0 standards. They evaluated 41 higher education courses' homepages and universities' landing pages using the following measurement tools: AChecker, Total Validator, WAVE, and ARIA for HTML/CSS. Both studies point out challenges and recommendations regarding accessibility for software design and development.

Ara et al. [8] and Paiva et al. [51] conducted Systematic Literature Reviews (SLRs) on the development of accessible software. Ara et al. [8] identified requirements, challenges, engineering techniques, ontologies, frameworks, Application Programming Interfaces (APIs), algorithms, and testing tools for different levels of satisfaction focusing on visual impairments, whereas they highlighted a research gap related to other disabilities. They provided an updated view of processes, methods, techniques, and tools to support the accessibility development by Web designers and developers. Paiva et al. [51] reviewed 94 papers published between 2011 and 2019 to investigate how accessibility is adopted across the distinct software life cycle phases. Most papers focus on the design and testing phases, and some papers have addressed accessibility since the beginning of the process. Both studies highlight the current focus on accessibility is predominantly on partial or total visual impairment.

Rosa and Valentim [54] carried out a systematic mapping to identify Accessibility, Usability, and User Experience (UX) technologies used during software design for people with visual impairments. They mapped 23 different technologies to this topic and some research gaps, such as (i) most design technologies focus only on one concept, Accessibility or Usability or UX; (ii) most technologies do not specify which level of visual impairment they are intended for; and (iii) when the technology specifies the type of disability, they are designed only for the work in question, not providing a standard of replicability. Furthermore, most of the technologies identified are specific to a type of platform, not allowing the technology to be used to assist in designing applications for other platforms.

Silveira et al. [57] presented a systematic mapping of scientific literature to identify the attributes and measures regarding accessibility focusing on people with visual impairments. This mapping provides a consolidation of the existing knowledge and identifies research opportunities. The authors highlight that, due to the amount of research on the topic, they needed to refine the study scope to focus on visual impairments since this type of disability has a vast body of knowledge and is the most common among people with some type of disability.

Our study, similar to Silveira et al. [57] and Rosa and Valentim [54], focuses on accessibility for users with visual impairments. However, we go beyond evaluation measures and metrics to identify different types of resources to support accessibility evaluation.

We applied a methodology similar to Ara et al. [8] and Paiva et al. [51] to select and analyze papers regarding accessibility in a broader period (from 2009 to 2023). However, our study investigates how accessibility for visually impaired users is addressed in Web systems, focusing on selecting and analyzing papers regarding this application domain. Similar to AlMeraj et al. [5], we also discuss challenges and recommendations regarding accessibility for software design and development. However, while AlMeraj et al. [5] focuses on accessibility evaluation in higher educational institutions' websites through automatic measurement tools, we aim to conduct a structured literature survey to identify techniques, methods, measures, and tools to support accessibility evaluation in Web systems.

4 SURVEY METHODOLOGY

In this literature survey, we aim to carry out a broad and descriptive investigation with a focus on identifying studies that address accessibility research in Web systems for people with visual impairments. The main difference between our survey in relation to traditional Systematic Literature Reviews (SLRs) and Systematic Literature Mappings (SLM) lies in the more rigid structure of SLRs and SLMs; we use a lighter review structure. To this end, we defined our main research question: "What has been used for accessibility evaluation of Web systems focusing on visually impaired users?".

The survey methodology consisted of the planning and execution stages. When planning, we base ourselves on a research question and create our search string to apply to selected libraries. Still, at this stage, we defined the inclusion and exclusion criteria to be applied to papers extracted from libraries. We also define an extraction form. The execution consisted of extracting papers from libraries according to the search string and applying 2 filters with experts in the field. Duplicate papers were removed, and we applied the first filter to the title and abstract of the papers using the inclusion and exclusion criteria. We applied the second filter to read the full text using inclusion and exclusion criteria and obtained a set of final papers to extract data and obtain survey findings.

Considering our research goal and that we followed a lighter review structure instead of a systematic strategy, we decided to not apply a quality assessment for the selected papers, since it is out of our survey's scope to apply comparison criteria between the findings found in the papers or to exclude papers according to a quality threshold. These are procedures applied only for standard systematic literature reviews.

4.1 Planning

We defined a search string to identify papers in digital libraries. The string was organized and structured in order to gather secondary studies (either systematic literature reviews or systematic mappings) and also primary studies, searching for papers that address methods and/or measures for evaluating accessibility in Web systems, quality-in-use improvement, accessibility guidelines adoption, and inclusiveness of visually impaired users, among other studies included according to the keywords.

Therefore, we searched for research papers in the scientific literature that consider a set of keywords and their synonyms, structured through the following search string:

(("software accessibility" OR "web accessibility" OR "usability of web" OR "website usability" OR "accessibility of web" OR "adaptability of web sites") AND ("visually handicapped" OR "visually impaired" OR "blind" OR "visual disability" OR "low vision") AND ("attribute" OR "property" OR "metric" OR "measure" OR "measurement" OR "evaluation" OR "assessment")).

The protocol we defined for this survey considered four digital libraries: IEEE Xplore[3], ScienceDirect[4], Springer[5], and ACM Digital Library[6]. From them, the ScienceDirect search engine limits the database search to eight connectors, which consecutively reduces the number of keywords. Therefore, we needed to tailor our string to carry out the search on this digital library. Hence, the search string applied on ScienceDirect was the following:

(("web accessibility" OR "usability") AND ("visually impaired" OR "blind" OR "visual disability") AND ("metric" OR "measure" OR "evaluation" OR "assessment")).

We adopted Parsifal[7] as a supporting tool throughout the survey process. Parsifal is a robust tool built based on the work by Kitchenham and Charters [38], supporting the following phases: planning, conduction through identification and selection of papers, data extraction, and interpretation of primary or even secondary studies available on digital libraries to answer research questions regarding a particular research topic or thematic area. To support the selection process, we defined a set of criteria (Table 3) to screen research papers unrelated to our goal and research question.

Table 3: Inclusion and Exclusion Criteria

Type	Criteria	ID
Inclusion	The paper addresses techniques, methods, or tools for evaluating or testing accessibility in Web systems.	CI1
	The paper presents measures for accessibility evaluation in Web systems.	CI2
	The paper discusses specific characteristics related to accessibility for visually impaired users.	CI3
Exclusion	Doctoral or master's thesis.	CE1
	The paper is a duplicated record or presents the same study already presented in other papers.	CE2
	The paper focuses on mobile devices (unless it also addresses Web accessibility).	CE3
	Short papers (up to three pages) since they usually do not provide a substantial contribution.	CE4
	Non-scientific papers or non-peer-reviewed papers; the paper does not meet any inclusion criteria.	CE5
	The paper's full text is not available.	CE6
	The paper was not written in English/Portuguese.	CE7

We applied two filters for screening papers. The first filter was based on the analysis of each paper's title, abstract, and keywords. Then, the accepted papers were included in the second filter, in which we analyzed each paper's full text according to the established inclusion and exclusion criteria. We also verified whether

[3]https://ieeexplore.ieee.org/
[4]https://www.sciencedirect.com/
[5]https://link.springer.com/
[6]https://dl.acm.org/
[7]https://parsif.al/

the papers meet our research question. Finally, we carried out data extraction and analysis, in which the information found was categorized, for instance, into tools, techniques, methods, and measures to gather all information in an organized way.

During data extraction, we applied an extraction form. According to Keele et al. [37], it should be designed to collect the necessary data to address the research question and the survey selection criteria. For this study, we designed a data extraction form organized into seven fields as follows: a) paper's title; b) paper's authors; c) year of publication; d) digital library; e) keywords; f) type of contribution, which includes, for instance, types of methods, techniques, and tools; and g) the contribution content, such as a technique's description and measures applied.

The information was analyzed to understand the type of contribution made and the context of use of each resource. In cases where the information is not well explained, we analyzed these excerpts, and when necessary, they were rewritten. In the normalization stage, we combined the information that was repeated in different works so that they appeared only once in the final version. After extracting the contributions and the examples obtained from each paper, we were able to synthesize a set of available resources found in our literature survey, including techniques, methods, measures, and tools applied for evaluating accessibility in Web systems.

After analysing the information, this study was submitted to peer review by three experts: an undergraduate student with 2 years' market experience in testing and accessibility; a PhD student with 6 years' experience in Human-Computer Interaction (HCI); and a researcher with over 15 years' experience in testing and software engineering. They assessed the study's coherence and contribution to research. This set of resources was considered more complete in view of the works present in the literature on this domain, as it brings together data from papers available in the literature from 2009 to 2023, collecting resources for evaluating systems involving users with visual impairments. Thus, aggregating the recommendations proposed by different works.

4.2 Execution

The application of the search string in digital libraries was carried out on February 22, 2023, when 1245 papers were found, of which 77 were from IEEEXplore, 123 from ScienceDirect, 446 from Springer, and 599 from ACM Digital Library. All work was recorded and organized in the Parsifal tool to support data management.

When starting the study, we removed all duplicate publications according to exclusion criterion CE2, considering that the same paper can be published in different libraries. 13 papers were removed from the list. Then, we start with the first filter (Table 4), using the other exclusion criteria with the review of titles, abstracts, and keywords. In this step, 1,149 papers were excluded, and 83 that met the established criteria were included.

Table 4 – Number of Papers in the First Filter by Library

Digital Library	# Included	# Excluded	# Duplicated	Total
IEEE Xplore	10	66	1	77
ScienceDirect	9	105	9	123
Springer	25	421		446
ACM DL	39	557	3	599
Total	83	1149	13	1245

In the second filter (Table 5), we evaluated the papers' full texts according to the defined exclusion and inclusion criteria. Our objective was to include only the studies that allow us to answer the research question previously defined. To do this, we focus on the inclusion criteria CI1, CI2, and CI3. In this step, we carried out a complete reading of the 83 papers, identifying relevant information for the survey. As a result, we excluded 31 papers and included 52 papers in our final set of papers, which are aligned with the survey's goal.

Table 5: Number of Papers in the Second Filter by Library

Digital Library	# Included	# Excluded	Total
IEEE Xplore	6	4	10
ScienceDirect	6	3	9
Springer	19	6	25
ACM DL	24	15	39
Total	52	31	83

After the second filter, each paper was explored individually. This approach allowed us to ensure that their content presents pertinent information. In this step, 52 papers were read completely, and their information was extracted. Table 6 lists the final set of papers. Figure 1 shows a visual representation of the execution phase scenario.

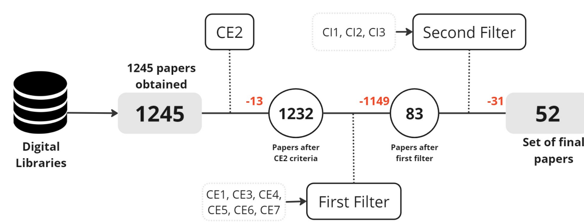

Figure 1: Filtering of papers during the execution stage

4.3 Ethical Concerns

Resolution 510/2016 from the National Health Council of Brazil [17] states that "research carried out exclusively with scientific texts to review the scientific literature" is exempt from evaluation by the Research Ethics Committees and the National Research Ethics Committee. This research conducted a literature survey and, therefore, does not involve experimentation with human beings. Hence, this research protocol was not submitted for ethical appraisal by an ethics committee in the country.

5 FINDINGS

In the scientific literature, several studies can be used to understand how to carry out Web accessibility evaluation in order to develop and adapt evaluation methods to improve Web systems development processes. Each resource follows a way of carrying out the evaluation depending on the defined goal. The evaluators can conduct the evaluation in different phases of the system, either during development or after its completion. In this study, we cataloged the main accessibility evaluation resources in the scientific literature to

Table 6: Final Set of Papers Selected for this Literature Survey

ID	Title	Reference
S1	Hera-FFX: A Firefox Add-on for Semi-Automatic Web Accessibility Evaluation	[29]
S2	Assistive Tools to Reach European Accessibility Web Standards	[21]
S3	Evaluating Groupware Accessibility	[29]
S4	The Complementary Role of Two Evaluation Methods in the Usability and Accessibility Evaluation of a Non-Standard System	[2]
S5	Accessibility by Demonstration: Enabling End Users to Guide Developers to Web Accessibility Solutions	[15]
S6	A web accessibility assessment on the Texas public school system	[46]
S7	Development and Trial of an Educational Tool to Support the Accessibility Evaluation Process	[10]
S8	Analyzing Web Accessibility in Finnish Higher Education	[32]
S9	Automatic web accessibility metrics: Where we are and where we can go	[60]
S10	Using Acceptance Tests to Validate Accessibility Requirements in RIA	[65]
S11	Guidelines Are Only Half of the Story: Accessibility Problems Encountered by Blind Users on the Web	[52]
S12	Guidelines, Icons and Marketable Skills: An Accessibility Evaluation of 100 Web Development Company Homepages	[31]
S13	Evaluating the Accessibility of Web Applications	[28]
S14	Accessibility evaluation improvement using Case Based Reasoning	[9]
S15	Is Moodle Accessible for Visually Impaired People?	[18]
S16	Benchmarking Web Accessibility Evaluation Tools: Measuring the Harm of Sole Reliance on Automated Tests	[61]
S17	The Differences in Accessibility of TV and Desktop Web Applications from the Perspective of Automated Evaluation	[23]
S18	Investigating Accessibility on Web-Based Maps	[47]
S19	The Model of Accessibility to Electronic Government: Applicability in DATAPREV	[27]
S20	Heuristic evaluation of the visual accessibility of the moodle Virtual Learning Environment	[24]
S21	Social Media and Accessibility	[67]
S22	Active Learning for Web Accessibility Evaluation	[71]
S23	WAEM: A Web Accessibility Evaluation Metric Based on Partial User Experience Order	[59]
S24	The accessibility of Cyprus Islands' Higher Education Institution Websites	[35]
S25	Accessibility of Turkish university Web sites	[39]
S26	Developing and Evaluating a Thai Website Accessibility Checker	[6]
S27	Crowdsourcing-Based Web Accessibility Evaluation with Golden Maximum Likelihood Inference	[58]
S28	Software Accessibility for Visually Impaired People: A Systematic Mapping Study	[57]
S29	Peruvian Public Universities and the Accessibility of Their Websites	[13]
S30	A Heuristic Method to Evaluate Web Accessibility for Users With Low Vision	[1]
S31	Types of Problems Encountered by Automated Tool Accessibility Assessments, Expert Inspections and User Testing: A Systematic Literature Mapping	[56]
S32	Are urban public libraries websites accessible to Americans with Disabilities?	[44]
S33	Web Navegate Visual Comfort : Inclusival and Egalitarian Opportunities for Visual Impairment Users	[42]
S34	Preliminary Evaluation of Interactive Search Engine Interface for Visually Impaired Users	[7]
S35	Web Accessibility Testing for Singapore Government E-Services	[41]
S36	Web Accessibility Evolution in the Brazilian Government	[50]
S37	Web accessibility of healthcare Web sites of Korean government and public agencies: a user test for persons with visual impairment	[69]
S38	Accessibility Issues in Indonesian E-Commerce Portals: Issues and Recommendations for Business Improvement and Growth	[11]
S39	What Makes Videos Accessible to Blind and Visually Impaired People?	[43]
S40	Accessibility, usability, and security evaluation of Hungarian government websites	[25]
S41	Accessibility of university websites worldwide: a systematic literature review	[19]
S42	The accessibility of state occupational safety and health consultation websites	[70]
S43	Evaluating the accessibility of higher education institution websites in the State of Kuwait: empirical evidence	[5]
S44	Refining the Acceasy Framework from an accessibility assessment	[30]
S45	Web accessibility investigation and identification of major issues of higher education websites with statistical measures: A case study of college websites	[36]
S46	An Analysis of Web Content Accessibility of Municipality Websites for People with Disabilities in Norway: Web Accessibility of Norwegian Municipality Websites	[34]
S47	Does the law matter? An empirical study on the accessibility of Finnish higher education institutions' web pages	[40]
S48	Improving accessibility of CMS-based websites using automated methods	[26]
S49	Accessibility of COVID-19 Websites of Asian Countries: An Evaluation Using Automated Tools	[49]
S50	Effect of Potential Issues Flagged by Automated Tools on Web Accessibility Evaluation Results: A Case Study on University Department Websites	[3]
S51	Web Accessibility: An Evaluation of CCSC Central Plains Participants' University Home Pages	[66]
S52	Accessibility engineering in web evaluation process: a systematic literature review	[8]

gather information to support the evaluator in choosing the most appropriate resource(s) in each case.

By analyzing the final set of papers included in this survey, a valuable set of resources was identified to evaluate the accessibility of Web systems for users with visual impairments. The main types of contributions can be classified into Attributes, Tools, Heuristics, Measures, Methods, Metrics, Types of Evaluations, and Types of Tests, which we obtained from various sources. We describe them in detail below.

5.1 Related Attributes and Measures

According to Montagud et al. [48], a quality attribute is a measurable physical or abstract property of an artifact produced during product development. Most attributes have measures that help verify whether they are being met, such as the Presence of Alternative Texts (textual equivalents) attribute refers to the "*alt*" tags, which provide alternative information for images. These tags are essential to ensure that, in situations such as slow loading or for users who rely on screen readers, the image content is described [57].

According to Silveira et al. [57], measures represent a structured set of parameters that enable an objective evaluation of the product's quality attributes. It allows quantifying and comparing the presence or absence of such attributes, allowing a detailed analysis of how each manifests itself. Table 7 exemplifies the attributes and measures for people with visual impairments mapped in the technical literature.

Two of the studies analyzed presented accessibility attributes and measures. The study S28 [57] presented a systematic mapping study and identified fourteen attributes related to 24 accessibility measures for this target users. The study S42 [70] used manual tests to complement the automated tests, in which a manual evaluation was conducted on some of the critical areas that the automated tools had problems verifying.

Table 7 – Attributes and Measures for Visually Impaired Users Identified in the Studies

ID	Attribute	Measure	Study ID
1	Affection	Measure was not mentioned.	**S28**
2	Audibility	Existence of textual equivalents, Redundant text, and Separated characters by spaces.	**S28**
3	Control	Measure was not mentioned.	**S28**
4	Shortcut availability	Measure was not mentioned.	**S28**
5	Efficiency	Measure was not mentioned.	**S28**
6	Understanding Learning/Capacity	Degree of understanding, Number of violations of the principle of understanding for each task.	**S28**
7	Ease of operation	Degree of operability related to the use of: Keyboard, Time for execution, Navigation complexity, Existence of anchors, Number of violations of the principle of operability in each task.	**S28**
8	Navigability	Reach time, Existence of headers or links to main content, Rate of accessible links on the page, Ratio of header tags, navigation skip links, FORM tags and TABLE tags structure.	**S28**
9	Perception	Degree of content perception, Number of violations of the principle of perception in each task.	**S28**
10	Presence of Alternative Texts	Percentage of Alternative Texts Found.	**S28, S42**
11	User productivity	Execution time for each task.	**S28**
12	Presence of Useful Form Labels	Percentage of Useful Form Labels.	**S42**
13	Image Redundancy	Percentage of redundant images.	**S28**
14	User Satisfaction	Degree of satisfaction	**S28**
15	Skip to Content	Presence of links to skip to content.	**S42**
16	Utility of Links	Utility of Text Links	**S28, S42**

The S42 study [70] explains that the "*alt*" attributes are considered useless in certain situations: if they were absent, by incorrectly marking it as empty (*alt=""*) or containing full or partial filenames such as *alt="logo.jpg"* or *alt="image343. png"*, vague descriptions such as *alt="image"* or *alt="logo"* hinder access to content for users with disabilities. We intend to emphasize the use of this mechanism to evaluate the percentage of alternative texts found in relation to the images presented and determine the accessibility of a Web page in relation to accessible images, thus guaranteeing access to the content by people with visual impairments.

Attributes and Measures are still rarely used, as we observed in this study. Only two studies emphasized the search for attributes and measures. However, identifying and defining these attributes and their respective measures can bring significant results not only to researchers but especially to users with visual impairments.

5.2 Tools

We found 22 different tools used to perform automatic evaluation on Web systems. Some studies used more than one resource, either a method, technique, tool, and/or type of evaluation, to determine the system's accessibility. It is important to note that these tools alone may not identify all problems in a system.

Automatic tools perform automatic checks against accessibility guidelines and standards on Web systems. They are useful for quickly checking and identifying areas that need attention to ensure accessibility for all users, especially those with disabilities. Table 8 presents the automatic tools mapped in the scientific literature and the respective studies that cite their use.

5.3 Set of Heuristics

Nielsen's heuristics provide a set of general principles widely applied to identify usability problems in interactive systems [7]. They are also widely used to identify areas to improve in the user interface design and user experience. Heuristic Evaluation is a method conducted by experts who apply a set of pre-defined heuristics to evaluate the usability of an interactive system. The process involves experts examining the system for heuristics violations and highlighting areas that require corrections. Table 9 presents the heuristics found regarding accessibility in the scientific literature.

Table 8 – Types of Automatic Tools Identified in the Studies

ID	Tool	Study ID
1	AccessMonitor	S18 and S19
2	AChecker	S12, S14, S16, S18, S25, S29, S30, S31, S41, S42, S43, S49, and S51
3	ASESWeb	S36
4	CynthiaSays	S18
5	Deque	S16
6	Examinator	S29 and S30
7	Fujitsu	S3
8	HERA-FFX	S1
9	HTML/CSS/ARIA	S43
10	MAUVE++	S31
11	QualWeb	S13 and S17
12	Siteimprove Accessibility Checker	S47 and S50
13	SortSite	S16 and S31
14	TAW (Web Accessibility Testing)	S8, S16, S18, S24, S29, S30, S41, S45, S46 and S51
15	Tenon	S31 and S41
16	TVT (Total Validator Test)	S38
17	W3C CSS	S30
18	W3C Markup	S30
19	WaaT7	S17
20	WatchFire Bobby Software (Discontinued)	S6
21	WAVE	S5, S24, S31, S32, S40, S41, S43, S46, S47
22	Web Accessibility Checker	S17

Table 9 – Heuristics for Visually Impaired Users Identified in the Studies

ID	Heuristic	Reference
1	Presence of speech	S39
2	Informative language	S39
3	Infrequent visual changes	S39
4	Simple visual content	S39
5	Description of visual objects	S39
6	On-screen text description	S39
7	Few visual references	S39

The study S34 [7] used Nielsen's heuristics to evaluate an interactive search engine called InteractSE, a Google search interface for users with visual impairments that aims to minimize the presentation text of search results that need to be read by the screen reader. The study S39 [43] carried out a study on accessible videos for blind and visually impaired users since the users interviewed in the study reported that the search for accessible videos was carried out using a time-consuming approach via trial and error. Furthermore, users with disabilities reported some methods used to characterize accessible and inaccessible videos. With this data and based on the study's findings, the authors formulated seven heuristics and seven corresponding metrics to measure video accessibility.

The use of heuristics is recommended to perform quick evaluations of a system. They can be used in the initial stages of the design or development process to identify accessibility problems and immediate solutions. Furthermore, heuristics are useful in evaluating existing interfaces, helping to identify possible improvements and corrections. Combining heuristics with other evaluation methods is also recommended to obtain more complete and accurate results.

5.4 Methods

Evaluation methods refer to approaches, strategies, and procedures that can be used to examine and verify, for instance, the effectiveness of a given system, product, or service. The methods may vary in different approaches, whether manual or automated. Table 10 lists the methods mapped in the literature.

Table 10 – Evaluation Methods Identified in the Studies

ID	Evaluation Method	Reference
1	Replacement of CSS/SCSS class	S48
2	Replacement of MVC-based extension	S48
3	Replacement of HTML output	S48
4	Data input verification	S48
5	Semi-supervised active prediction machine learning	S22
6	Modification of the Barrier Walkthrough heuristic method	S30
7	Manual evaluation through assistive technology with users and experts in the field	S52
8	Questionnaire-based evaluation for visually impaired individuals through various data analysis techniques	S52
9	Questionnaire-based evaluation to discover navigation strategies of individuals with low vision causing accessibility barriers	S52
10	Automatic evaluation system to identify the most effective validator for accessibility testing	S52
11	Statistical analysis of data to validate the reliability of the questionnaire result	S52
12	Quantitative data analysis using statistical analysis methods	S52
13	Manual evaluation criteria for assessing the accessibility of Australian private and governmental websites in relation to DDA(Disabilities discrimination act) standards	S52
14	User evaluation of Adobe's online design platform tool with the assistance of mixed-panel data analysis	S52
15	Statistical evaluation for website quality analysis	S52

The study S48 [26] details that there are relatively few automated solutions to simplify the correction of problems. They suggested four methods to solve accessibility problems on websites based on content management systems:

- **CSS/SCSS class replacement method:** The goal is to automatically correct CSS/SCSS codes that do not meet WCAG requirements. Formatting restrictions are set for each element on the page, overriding inappropriate user settings and ensuring accessibility. This is possible due to the priority of CSS rules: formatting restrictions defined for the element take priority over user settings. The advantage of this solution is that if multiple websites use the same structure (same CSS styles), then the repair rule can be easily applied to all of those sites.
- **Model-View-Controller (MVC)-based extension replacement method:** This method addresses accessibility issues such as function descriptions for Web elements, fixing missing or cluttered titles, and improving header cells in tables, among other errors. The process begins with a check of the website by an automatic tool, after which a study of the source code of the section with error is carried out to identify which extension is responsible for the inappropriate formatting for accessibility. After identifying the extension and its version, the original file responsible for rendering is copied to the appropriate folder. This so-called instantiation operation can be automatically executed in the Content Management System (CMS) administration. After this operation, the newly created files must be modified to fix the accessibility issues marked by the automatic scan tool. After the necessary fixes are completed, the online checker should be run again to check the improved *interface*. The process should be repeated until the page is error-free.
- **HTML output replacement method**: If the scanning engine detects an accessibility issue, it will be changed immediately. If a vulnerability issue is detected, it can be fixed with a replacement extension without modifying the original content according to previously defined rules. Each rule must have a pair, the faulty source code, and the appropriate source code. Regular expressions can also be used to make a substitution rule flexible. Using this method can improve almost any problem accessibility errors detected by modifying the site's source code.
- **Data entry check**: The objective is to warn during content creation if the content uses formatting that does not adequately meet WCAG recommendations.

Using the methods described allows Web systems to comply with WCAG guidelines and helps improve their accessibility. In current content management systems, the actual content is usually inserted using so-called "What You See Is What You Get" (WYSIWYG) editors. After starting the plugin among the examples mentioned in the study, it can detect the error when the WYSIWYG editor is used. It suggests a solution and then moves on to the next error. Detected errors are created in advance (errors and their corresponding rules), and new cases can be added via JavaScript code.

The study S30 [1] used a modification of the Barrier Walkthrough (BW) method proposed by Giorgio Brajnik considering the WCAG. The modification consisted of including persistence to determine the severity of an accessibility barrier. The method was applied to 40 websites, including those from 30 universities in Latin America, according to ranking Webometrics, and 10 websites among the most visited, according to ranking Alexa. This method comprises 10 phases: (1) select the website, (2) select the type of users, (3) identify users' goals and scenarios, (4) explore interaction mechanisms, (5) list the barriers according to the users, (6) apply the UX Check tool, (7) evaluate the website with the BW method, (8) record data, (9) analyze results, and (10) suggest recommendations. This method allows for measuring the accessibility of websites and obtaining sample data for analysis.

The study S22 [71] emphasizes the use of artificial intelligence in evaluating accessibility on the Web and proposes a new method for evaluating the accessibility of websites, consisting of five steps: 1) active prediction semi-supervised machine learning is used; 2) the team of evaluators selects the most informative pages or sites to be evaluated for active learning; 3) evaluate selected pages to obtain their corresponding accessibility results using automatic tools; 4) consider the evaluated pages as training data to obtain prediction models; 5) finally, the accessibility results of the remaining pages are predicted using machine learning.

The study S52 [8] is a systematic literature review that analyzed several works and mapped numerous accessibility evaluation methods, such as: (1) manual evaluation using assistive technology with users and domain experts; (2) questionnaire-based evaluation for people with visual impairments through various data analysis techniques; (3) questionnaire-based evaluation to discover the navigation strategies of people with low vision that cause accessibility barriers; (4) automatic evaluation system to identify the most effective validator for accessibility testing; (5) statistical analysis of data to validate the reliability of the questionnaire result; (6) quantitative data analysis using statistical analysis methods; (7) manual evaluation criteria for evaluating the accessibility of Australian private and government websites against Disability Discrimination Act (DDA) standards; (8) user evaluation of Adobe's online design platform tool with the help of mixed panel data analysis; (9) statistical evaluation to analyze the quality of websites.

Different methods to improve accessibility are proposed by studies, and the choice of a specific method depends on a combination of factors that guide the evaluation. These factors include 1) the evaluation's goal, such as identifying problems, validating the use of standards, or analyzing the needs of the system's users; 2) the target audience, considering their needs, skills, and context of use; and 3) project needs, where more complex projects may require more detailed and comprehensive evaluation methods.

5.5 Metrics

In general, evaluation metrics can be described as the readability measurement tool of the website content to make website content universally accessible. According to the Vigo and Brajnik [60] classification algorithm, different metrics can evaluate different system characteristics. Table 11 lists the metrics regarding visual impairment mapped in the scientific literature and the respective studies that apply them.

Table 11 – Metrics Identified in the Studies

ID	Metric	Reference
1	Web Accessibility Barriers (WAB)	S9
2	Page Measure (PM)	S9
3	Web Accessibility Quantitative (WAQM)	S9
4	Web Accessibility Experience Metric (WAEM)	S23
5	Percentage of time without speech	S39
6	Percentage of low lexical density speech	S39
7	Shot change rate	S39
8	Number of visual entities detected per minute	S39
9	Percentage of visual entities not in speech	S39
10	Number of on-screen text detected not in speech per minute	S39
11	Number of unresolved reference words per minute	S39

Vigo and Brajnik [60] carried out an empirical study to evaluate the validity and adequacy of metrics. A structure was proposed to evaluate existing metrics. Using this structure, they show the results obtained when they applied the seven existing automatic accessibility metrics. The aim is that metrics can cover validity, reliability, sensitivity, adequacy, and complexity in the context of four scenarios in which they can be used. Seven metrics were experimentally analyzed on 1,543 pages. Among the metrics evaluated, three of them showed better results, as follows:

- **Quantitative Web Accessibility Metric (WAQM):** analyzes several parameters, such as the presence and correct implementation of accessible elements, the use of correct attributes in images, and the structuring and semantics of the code.
- **Page Measure (PM):** examines factors that affect page accessibility, such as the presence and quality of alternative text in images, content readability, navigation, and screen reader comprehension of the site and assistive technologies.
- **Web Accessibility Barriers (WAB) Metric:** identifies specific points at which the page may be inaccessible or pose challenges for user groups with disabilities.

According to the study S23 [59], the WAEM metric can combine better the results of accessibility evaluation with the user experience of people with disabilities, aligning the evaluation metric with the Partial User Experience Order (PUEXO). This metric considers aspects such as ease of navigation, clarity/readability of content, effectiveness of navigation tools, and other elements that can affect usability for people with disabilities.

In the study S39 [43], seven quantitative metrics were instantiated to evaluate the accessibility of videos, as follows: 1) percentage of duration without speech, 2) percentage of low lexical density speech, 3) rate of shot changes, 4) number of visual entities detected per minute, 5) percentage of visual entities that are not in speech, 6) number of text detected on the screen that is not in speech per minute, and 7) number of unresolved reference words per minute. Thus, accessibility is considered essential to videos that, through video accessibility metrics, allow blind and visually impaired people to quickly find videos of interest.

Different metrics for evaluating accessibility were found in our survey. To choose a metric, the following aspects must be analyzed: 1) evaluation's goal: identify problems, validate the use of standards, analyze the experience of system users; 2) target audience: needs and experiences; and 3) metric approach: whether the aim of the analysis improves accessibility in the project, whether the metric is effective and relevant. After analyzing the options, it should be possible to select the metric or set of metrics that contributes to improving the system.

5.6 Types of Evaluation

To ensure compliance of quality characteristics, different types of evaluation can be used to evaluate a system, the evaluation must follow existing recommendations or a specific plan of what needs to be evaluated. Among the most mentioned types of evaluation are: Manual Evaluation, Evaluation with Expert Users, Evaluation with End Users, and Heuristic Evaluation. Analyzing the works selected in the 2nd filter, we observed that most of the works applied different types of accessibility analysis, complementing their research through some of the evaluations mentioned.

- **Manual evaluation**: Manual evaluation consists of analyzing a selected page or system; users are invited to perform specific tasks on a given system so that after this, the tasks are thoroughly analyzed, and any problems identified are documented so that the development team can seek solutions appropriate for improving the system.
- **Evaluation with experts**: The expert evaluation involves inviting qualified professionals familiar with different frequently developed systems and asking them to evaluate their usability, performance, and functionalities. These experts can easily identify errors and offer valuable insights to improve the system.
- **Evaluation with end-users**: Evaluation with end users represents direct interaction with users from the target audience, in which users are invited to use the system and perform specific tasks. In contrast, their interactions are observed

and recorded. As they are the target audience, they provide a more complete understanding of failures, helping to understand needs and detecting practical problems that were not previously noticed.

- **Heuristic evaluation**: Heuristic evaluation is an evaluation method based on a set of established principles used as a reference to identify usability problems when evaluating user interfaces. In this study, the use of heuristic evaluation was identified in two papers that, together with other resources, evaluate the accessibility of Web systems.
- **Automatic evaluation**: Automatic tools use algorithms and perform checks to examine content's compliance with established accessibility guidelines and can identify common problems such as lack of alternative text in images, lack of adequate contrast, inappropriate header structures, and other elements that can make it difficult to read. navigation. Tools help identify apparent issues, but these tools have limitations and cannot capture all accessibility flaws in the system.

Each system evaluation method has different objectives and approaches. The choice must be analyzed according to the feedback you want after evaluating the system. Table 12 exemplifies the attributes/measures for visual impairment mapped in the technical literature and the respective studies they used.

Table 12 – Types of Accessibility Evaluation Identified in the Studies

ID	Types of evaluation	Study ID
1	Manual evaluation	S1, S11, S12, S14, S15, S21, S36, S37, S41, S42 and S50
2	Evaluation with experts	S15, S19, S26, S31, S33, S34, S43 and S44
3	Evaluation with end-users	S11, S26, S31, S33 and S37
4	Heuristic evaluation	S20, S30 and S34
5	Automatic evaluation	S1, S3, S5, S6, S8, S12, S13, S14, S16, S17, S18, S19, S24, S25, S29, S30, S31, S32, S36, S38, S40, S41, S42, S43, S45, S46, S47, S49, S50 and S51

5.7 Types of Test

This section describes the studies associated with the purpose of testing to validate the accessibility of Web platforms. Tests are specific methods used to verify different scenarios and functionalities of a software system, each type of test has different objectives and must be applied at different stages of software development. The types of accessibility tests identified are displayed in Table 13.

Table 13 – Types of Tests Identified in the Studies

ID	Types of tests	Reference
1	Validation test	S25
2	Acceptance test	S10
3	Mobile device compatibility test	S25
4	Browser graphical test	S25
5	Text-only browser test	S25

In S10 [65], the authors describes the development of an approach that is based on acceptance tests that allow evaluating the accessibility of applications. This approach supports the software development cycle by providing developers with an automated tool that checks all aspects of the system, from client-side to server-side implementation. Because it is implemented as an executable automated test, this solution integrates seamlessly into the Continuous Integration software development process, ideal for constantly improving web applications.

In the S25 study [39], the authors applied a combination of methods to evaluate the level of accessibility of the home pages of the websites of 10 of the best universities in the world, employing a series of different evaluation methods and collecting data from multiple sources, to achieve a more accurate and comprehensive analysis of accessibility levels. The strategies described below were used to collect data and evaluate functionality and accessibility.

- **Browser graphic test**: the existence of "alt" tags, changes in text size and navigation capacity were verified;
- **Text-only browser test**: the Lynx text browser was used to check that all information on the page, visible in graphical browsers, remained accessible and understandable in text format.
- **Automated Tests**: Three automated tools were used to check accessibility features.
- **Validation tests**: checking the correct use of HTML and CSS for each university's home page using the W3C validation service with the Validator and Jigsaw tools.
- **Mobile-friendly tests**: checking whether websites render as well on mobile devices as they should on other platforms with the help of Google's compatibility testing tool called mobile-friendly.

The type of testing to be applied during software construction depends on the life cycle phase the software is in, as there is no way, for example, to carry out tests on the interface if the system is at the beginning of the coding process. It is necessary to plan the tests to ensure that they are carried out at the appropriate time.

6 DISCUSSION

In the literature, there is a growing concern about making the web more inclusive and accessible for all users [62]. The reviewed work highlights the importance of ensuring compliance with accessibility guidelines such as WCAG and considering different user groups' specific experiences and needs, including the system's target audience, in the evaluation. We observe the evolution of accessibility guidelines and standards, such as WCAG and eMAG, and the increasing adoption of these guidelines from the initial phases of design and development instead of considering the inclusion of accessibility only after the system is finalized. This implies that the literature is currently focused on continuously improving accessibility, reducing digital barriers, and increasing compliance with legal regulations, as well as fostering technological innovation and highlighting corporate social responsibility, contributing to a more equitable and accessible web for all. In addition, considering accessibility from the earliest stages of development results in solutions that are inclusive from the start. However, this practice is only recently being adopted, and many web systems that didn't begin their development with accessibility in mind still remain inaccessible.

In the evaluation and accessibility test of Web systems, the resources most frequently cited in the identified studies are manual evaluation and the use of automatic tools. Manual evaluation is performed by experts who directly review the system, using checklists to verify compliance with accessibility guidelines. Among the most mentioned automatic tools are *AChecker*, which checks compliance with WCAG and eMAG guidelines; the TAW; and WAVE, which evaluates the accessibility of websites, identifying problems and offering suggestions for improvements. Each verification tool and method has different approaches and can identify different accessibility issues, emphasizing certain aspects or focusing on varying compliance criteria. Therefore, it is recommended to use multiple accessibility checking features when evaluating Web systems, especially since no single method can guarantee complete accuracy.

Taking into account the increased use of automatic tools in accessibility evaluation, the HCI industry needs to train specialists who are trained to manage and operate these tools. Thus, in parallel to empirical and human experience, these professionals must have technical expertise in process automation and the use of accessibility tools, integrating technical and practical skills to improve the effectiveness and accuracy of evaluations. Furthermore, the need to combine different accessibility evaluation methods has several implications for the HCI industry and web accessibility. It demands the training of professionals in a variety of techniques, increasingly driving the specialization of professionals and encouraging multidisciplinary collaboration between designers, developers, and accessibility specialists.

Accessibility for visually impaired users involves adapting content so that screen readers can easily navigate and locate information. Good practices include the adequacy of titles and subtitles, access and operation via keyboard, ensuring good contrast, and the consistency and organization of information. These practices aim to make web systems more inclusive, allowing visually impaired users to access, understand and interact with online content effectively and independently. The characteristics can be analyzed in Table 9, which contains the heuristics related to this topic. However, although the studies analyzed suggested a variety of resources and tools available for accessibility for visually impaired people, the need for combined evaluation approaches, along with technical expertise in automated tools and processes, may be delaying this implementation.

7 THREATS TO VALIDITY

Some possible threats to the results' validity were verified for this literature survey. We discuss them as follows.

The main research question, entitled "What has been used for accessibility evaluation of Web systems focusing on visually impaired users?" covers several aspects, such as techniques, methods, metrics, and tools, among other resources used to evaluate the accessibility of Web systems. However, this generic approach may have limited the results of this survey. Given the diversity of research existing on this topic, our study has considered various aspects and areas of study; however, it may also have neglected some topics due to the broad scope of the research question. A more specific approach could have allowed a more comprehensive analysis of the different aspects related to evaluating the accessibility of Web systems for users with visual impairments.

The paper selection may have been influenced by the availability of information, which may affect the representativeness of the results on what has been used to evaluate the accessibility of Web systems for visually impaired users. This influence may result in a sample that does not fully reflect the diversity of existing studies, limiting the generalization of results. Consequently, the conclusions derived from the data collected may have been influenced by the authors' biases in interpreting these data. The interpretation may have introduced biases that impact the results.

The search string definition may have resulted in the exclusion of studies relevant to the research. The absence of a relevant term or an excess of terms can distort the results, leading to the inappropriate inclusion or exclusion of studies, which can compromise the survey's validity. Furthermore, the choice of keywords in the search string may not adequately represent the diversity of approaches and practices for evaluating Web accessibility for people with visual impairments, limiting the generalization of results. To mitigate this threat, we defined the search string based on the key terms of the research question, which were previously reviewed by two researchers working with software engineering and human-computer interaction, respectively. However, despite this mitigation, there is still the possibility that some relevant studies were not included due to limitations inherent in the search string definition process.

8 FINAL REMARKS

In this study, we carried out a literature survey regarding accessibility evaluation of Web systems for people with visual impairments. We selected 52 papers that evaluated accessibility in existing systems, such as government websites and higher education institutes, focusing on people with visual or multiple disabilities.

As findings, we identified studies related to quality improvement of accessibility in Web systems for visually impaired users, from tools for checking compliance to testing methods with real users. The various examples presented verify compliance with the main WCAG guidelines and eMAG standards to make systems more accessible. In the selected papers, we found the following frequently mentioned evaluation resources: a total of 16 accessibility-related attributes and their respective measures related to visual impairment. The most common type of evaluation was manual evaluation, mentioned in 11 occurrences, followed by evaluation by experts with eight references, evaluation with end users mentioned six times, and Heuristic Evaluation mentioned three times. Furthermore, we mapped 22 automatic tools, among which the AChecker, TAW, and WAVE tools stand out. We also found a set of seven heuristics regarding accessibility. Furthermore, we mapped 11 accessibility metrics, 15 evaluation methods, and five specific types of tests for accessibility evaluation.

The findings of our study provide a centralized source of available resources used for accessibility evaluation in Web systems to support researchers and accessibility evaluators in selecting a resource from those cited according to their study/evaluation's goal and needs or to obtain an overview of the current state of accessibility evaluation in scientific literature. In so doing, our study contributes to the scientific community interested in Web accessibility and to software professionals who can consider the mapped resources in the development of more accessible interactive systems. As future perspectives, we suggest adapting this study to take into account other types of disabilities, thus offering valuable insights from a wider range of users with diverse needs and characteristics.

Also, a possible future work is to investigate the types of evaluation of accessibility for blind people in the context of digital games.

REFERENCES

[1] Patricia Acosta-Vargas, Luis Antonio Salvador-Ullauri, and Sergio Luján-Mora. 2019. A heuristic method to evaluate web accessibility for users with low vision. *IEEE Access* 7 (2019), 125634–125648.

[2] Funmi Adebesin, Paula Kotzé, and Helene Gelderblom. 2010. The complementary role of two evaluation methods in the usability and accessibility evaluation of a non-standard system. In *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. 1–11.

[3] Georgios Akritidis and Christos Katsanos. 2021. Effect of Potential Issues Flagged by Automated Tools on Web Accessibility Evaluation Results: A Case Study on University Department Websites. In *Proceedings of the 25th Pan-Hellenic Conference on Informatics*. 113–117.

[4] Nancy Alajarmeh. 2021. The extent of mobile accessibility coverage in WCAG 2.1: Sufficiency of success criteria and appropriateness of relevant conformance levels pertaining to accessibility problems encountered by users who are visually impaired. *Universal Access in the Information Society* (2021), 1–26.

[5] Zainab AlMeraj, Fatima Boujarwah, Dari Alhuwail, and Rumana Qadri. 2021. Evaluating the accessibility of higher education institution websites in the State of Kuwait: empirical evidence. *Universal Access in the Information Society* 20, 1 (2021), 121–138.

[6] Kewalin Angkananon, Mike Wald, and Piyabud Ploadaksorn. 2017. Developing and evaluating a Thai website accessibility checker. In *HCI International 2017–Posters' Extended Abstracts: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 19*. Springer, 3–10.

[7] Aboubakr Aqle, Kamran Khowaja, and Dena Al-Thani. 2020. Preliminary evaluation of interactive search engine interface for visually impaired users. *IEEE Access* 8 (2020), 45061–45070.

[8] Jinat Ara, Cecilia Sik-Lanyi, and Arpad Kelemen. 2023. Accessibility engineering in web evaluation process: a systematic literature review. *Universal Access in the Information Society* (2023), 1–34.

[9] Cecilia Avila, Silvia Baldiris, Ramon Fabregat, and Juan Carlos Guevara. 2012. Accessibility evaluation improvement using case based reasoning. In *2012 Frontiers in Education Conference Proceedings*. IEEE, 1–6.

[10] Christopher Bailey and Elaine Pearson. 2011. Development and trial of an educational tool to support the accessibility evaluation process. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*. 1–10.

[11] Raghuram Balaga. 2020. Accessibility Issues in Indonesian E-Commerce Portals: Issues and Recommendations for Business Improvement and Growth. In *Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2020, Tiruchirappalli, India, December 18–19, 2020, Proceedings, Part II*. Springer, 581–591.

[12] Simone Diniz Junqueira Barbosa, Bruno Santana da Silva, Milene Selbach Silveira, Isabela Gasparini, Ticianne Darin, and Gabriel Diniz Junqueira Barbosa. 2021. *Interação Humano-Computador e Experiência do Usuário*. Autopublicação.

[13] Fanny Dolores Benites Alfaro and Claudia María Del Pilar Zapata Del Río. 2018. Peruvian public universities and the accessibility of their websites. In *Design, User Experience, and Usability: Users, Contexts and Case Studies: 7th International Conference, DUXU 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part III 7*. Springer, 589–607.

[14] Tingting Bi, Xin Xia, David Lo, John Grundy, Thomas Zimmermann, and Denae Ford. 2022. Accessibility in Software Practice: A Practitioner's Perspective. *ACM Trans. Softw. Eng. Methodol.* 31, 4, Article 66 (jul 2022), 26 pages. https://doi.org/10.1145/3503508

[15] Jeffrey P Bigham, Jeremy T Brudvik, and Bernie Zhang. 2010. Accessibility by demonstration: enabling end users to guide developers to web accessibility solutions. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. 35–42.

[16] Brasil. 2014. eMAG - Modelo de Acessibilidade em Governo Eletrônico. *Portal do Governo Brasileiro* (2014). https://www.gov.br/governodigital/pt-br/acessibilidade-digital/eMAGv31.pdf

[17] Brasil. 2016. Resolução nº 510, de 07 de abril de 2016. *Diário Oficial da República Federativa do Brasil* (2016). Retrieved August 16, 2024 from https://www.gov.br/conselho-nacional-de-saude/pt-br/acesso-a-informacao/legislacao/resolucoes/2016/resolucao-no-510.pdf Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana. (In Portuguese).

[18] Rocío Calvo, Ana Iglesias, and Lourdes Moreno. 2012. Is Moodle accessible for visually impaired people?. In *Web Information Systems and Technologies: 7th International Conference, WEBIST 2011, Noordwijkerhout, The Netherlands, May 6-9, 2011, Revised Selected Papers 7*. Springer, 207–220.

[19] Milton Campoverde-Molina, Sergio Luján-Mora, and Llorenç Valverde. 2023. Accessibility of university websites worldwide: a systematic literature review. *Universal Access in the Information Society* 22, 1 (2023), 133–168.

[20] Casa Civil. 2015. Lei Nº 13.146, de 6 de julho 2015. *Institui a lei brasileira de inclusão da pessoa com deficiência (estatuto da pessoa com deficiência). Brasília* (2015).

[21] Sonia Colas, Nicolas Monmarché, and Mohamed Slimane. 2009. Assistive Tools to Reach European Accessibility Web Standards. In *Universal Access in Human-Computer Interaction. Applications and Services: 5th International Conference, UAHCI 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings, Part III 5*. Springer, 654–663.

[22] World Wide Web Consortium et al. 2008. Web content accessibility guidelines (WCAG) 2.0. (2008).

[23] Daniel Costa, Luís Carriço, and Carlos Duarte. 2015. The differences in accessibility of tv and desktop web applications from the perspective of automated evaluation. *Procedia Computer Science* 67 (2015), 388–396.

[24] Douglas Costa, Heitor Costa, and Paulo A Parreira. 2016. Heuristic evaluation of the visual accessibility of the Moodle Virtual Learning Environment. In *2016 XLII Latin American Computing Conference (CLEI)*. IEEE, 1–9.

[25] Balázs Csontos and István Heckl. 2021. Accessibility, usability, and security evaluation of Hungarian government websites. *Universal Access in the Information Society* 20 (2021), 139–156.

[26] Balázs Csontos and István Heckl. 2022. Improving accessibility of CMS-based websites using automated methods. *Universal Access in the Information Society* 21, 2 (2022), 491–505.

[27] Isabel Maria de Souza, Cristiano Maciel, and Claudia Cappelli. 2016. The model of accessibility to electronic government: applicability in dataprev. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research*. 287–292.

[28] Nádia Fernandes, Daniel Costa, Carlos Duarte, and Luís Carriço. 2012. Evaluating the accessibility of web applications. *Procedia Computer Science* 14 (2012), 28–35.

[29] José Luis Fuertes Castro, Ricardo González, Emmanuelle Gutiérrez, and Loïc Martínez Normand. 2009. Hera-FFX: a Firefox add-on for semi-automatic web accessibility evaluation. (2009).

[30] Renan Soares Germano, Tereza A Nobre de Sousa, and Maria Amelia Eliseo. 2021. Refining the Acceasy Framework from an accessibility assessment. In *2021 16th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 1–6.

[31] Teresa D Gilbertson and Colin HC Machin. 2012. Guidelines, icons and marketable skills: an accessibility evaluation of 100 web development company homepages. In *Proceedings of the international cross-disciplinary conference on web accessibility*. 1–4.

[32] Behzad Joze Hashemian. 2011. Analyzing web accessibility in Finnish higher education. *ACM Sigaccess Accessibility and Computing* 101 (2011), 8–16.

[33] IBGE IBGE. 2010. Censo demográfico 2010. *IBGE: Insituto Brasileiro de Geografia e* (2010), 34–35.

[34] Yavuz Inal, Deepti Mishra, and Anne Britt Torkildsby. 2022. An Analysis of Web Content Accessibility of Municipality Websites for People with Disabilities in Norway: Web Accessibility of Norwegian Municipality Websites. In *Nordic Human-Computer Interaction Conference*. 1–12.

[35] Erkut İ İşeri, Kaan Uyar, and Ümit İlhan. 2017. The accessibility of Cyprus Islands' higher education institution websites. *Procedia computer science* 120 (2017), 967–974.

[36] Abid Ismail and KS Kuppusamy. 2022. Web accessibility investigation and identification of major issues of higher education websites with statistical measures: A case study of college websites. *Journal of King Saud University-Computer and Information Sciences* 34, 3 (2022), 901–911.

[37] Staffs Keele et al. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report. Technical report, ver. 2.3 ebse technical report. ebse.

[38] Barbara Kitchenham and Stuart Charters. 2007. Guidelines for performing systematic literature reviews in software engineering. (2007).

[39] Serhat Kurt. 2017. Accessibility of Turkish university Web sites. *Universal Access in the Information Society* 16 (2017), 505–515.

[40] Merja Laamanen, Tarja Ladonlahti, Hannu Puupponen, and Tommi Kärkkäinen. 2022. Does the law matter? An empirical study on the accessibility of Finnish higher education institutions' web pages. *Universal Access in the Information Society* (2022), 1–17.

[41] Zui Young Lim, Jia Min Chua, Kaiting Yang, Wei Shin Tan, and Yinn Chai. 2020. Web accessibility testing for singapore government e-services. In *Proceedings of the 17th International Web for All Conference*. 1–5.

[42] Gabriel EF Lima, Natália N Vieira, Renato Campos Mauro, Fábio Paschoal Júnior, and Rafael Castaneda Ribeiro. 2020. Web Navegate Visual Comfort: Inclusival and Egalitarian Opportunities for Visual Impairment Users. In *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 1–6.

[43] Xingyu Liu, Patrick Carrington, Xiang 'Anthony' Chen, and Amy Pavel. 2021. What makes videos accessible to blind and visually impaired people?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[44] Yan Quan Liu, Arlene Bielefield, and Peter McKay. 2019. Are urban public libraries websites accessible to Americans with disabilities? *Universal Access in the Information Society* 18 (2019), 191–206.

[45] Beatriz Martins and Carlos Duarte. 2023. A large-scale web accessibility analysis considering technology adoption. *Universal Access in the Information Society* (2023), 1–16.

[46] Shane May and Qi Zhu. 2010. A web accessibility assessment on the Texas public school system. *Universal Access in the Information Society* 9 (2010), 87–96.

[47] Jonathas Leontino Medina, Maria Istela Cagnin, and Débora Maria Barroso Paiva. 2015. Investigating accessibility on web-based maps. *ACM SIGAPP Applied Computing Review* 15, 2 (2015), 17–26.

[48] Sonia Montagud, Silvia Abrahão, and Emilio Insfran. 2012. A systematic review of quality attributes and measures for software product lines. *Software Quality Journal* 20 (2012), 425–486.

[49] Telcia Niom and Frank Lin. 2022. Accessibility of COVID-19 Websites of Asian Countries: An Evaluation Using Automated Tools. *SN Computer Science* 3, 6 (2022), 498.

[50] Alberto Dumont Alves Oliveira, Andre Pimenta Freire, and Marcelo Medeiros Eler. 2020. Web accessibility evolution in the Brazilian government. In *XVI Brazilian Symposium on Information Systems*. 1–8.

[51] Débora Maria Barroso Paiva, André Pimenta Freire, and Renata Pontin de Mattos Fortes. 2021. Accessibility and Software Engineering Processes: A Systematic Literature Review. *Journal of Systems and Software* 171 (2021), 110819. https://doi.org/10.1016/j.jss.2020.110819

[52] Christopher Power, André Freire, Helen Petrie, and David Swallow. 2012. Guidelines are only half of the story: accessibility problems encountered by blind users on the web. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 433–442.

[53] Roger S Pressman and Bruce R Maxim. 2021. *Engenharia de software-9*. McGraw Hill Brasil.

[54] João Ricardo dos S. Rosa and Natasha Malveira C. Valentim. 2020. Accessibility, usability and user experience design for visually impaired people: a systematic mapping study. In *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems* (Diamantina, Brazil) *(IHC '20)*. Association for Computing Machinery, New York, NY, USA, Article 5, 10 pages. https://doi.org/10.1145/3424953.3426626

[55] Bruno Gomes da Silva and Kamila Rios H Rodrigues. 2018. Accessibility challenges in web systems implementation. In *Anais do IX Workshop sobre Aspectos da Interação Humano-Computador para a Web Social*. SBC, 105–116.

[56] Carlos Alberto Silva, Arthur FBA de Oliveira, Delvani Antônio Mateus, Heitor Augustus Xavier Costa, and André Pimenta Freire. 2019. Types of problems encountered by automated tool accessibility assessments, expert inspections and user testing: a systematic literature mapping. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*. 1–11.

[57] Barbara Cristina A Silveira, Thiago Silva-de Souza, and Ana Regina C da Rocha. 2018. Software Accessibility for Visually Impaired People: a systematic mapping study. In *Proceedings of the XVII Brazilian Symposium on Software Quality*. 190–199.

[58] Shuyi Song, Jiajun Bu, Andreas Artmeier, Keyue Shi, Ye Wang, Zhi Yu, and Can Wang. 2018. Crowdsourcing-based Web accessibility evaluation with golden maximum likelihood inference. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–21.

[59] Shuyi Song, Can Wang, Liangcheng Li, Zhi Yu, Xiao Lin, and Jiajun Bu. 2017. WAEM: a web accessibility evaluation metric based on partial user experience order. In *Proceedings of the 14th International Web for All Conference*. 1–4.

[60] Markel Vigo and Giorgio Brajnik. 2011. Automatic web accessibility metrics: Where we are and where we can go. *Interacting with computers* 23, 2 (2011), 137–155.

[61] Markel Vigo, Justin Brown, and Vivienne Conway. 2013. Benchmarking web accessibility evaluation tools: measuring the harm of sole reliance on automated tests. In *Proceedings of the 10th international cross-disciplinary conference on web accessibility*. 1–10.

[62] Beat Vollenwyder, Serge Petralito, Glena H Iten, Florian Brühlmann, Klaus Opwis, and Elisa D Mekler. 2023. How compliance with web accessibility standards shapes the experiences of users with and without disabilities. *International Journal of Human-Computer Studies* 170 (2023), 102956.

[63] WAI. 2022. *WCAG 2 Overview*. Retrieved June 08, 2022 from https://www.w3.org/WAI/standards-guidelines/wcag/

[64] WAI. 2023. *WCAG 2.1 at a Glance*. Retrieved Nov 20, 2023 from https://www.w3.org/WAI/standards-guidelines/wcag/glance/

[65] Willian Massami Watanabe, Renata PM Fortes, and Ana Luiza Dias. 2012. Using acceptance tests to validate accessibility requirements in ria. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*. 1–10.

[66] Michael Whitney and Stephen Dannelly. 2022. Web accessibility: an evaluation of CCSC central plains participants' university home pages. *Journal of Computing Sciences in Colleges* 37, 6 (2022), 46–53.

[67] Gian Wild. 2016. Social media and accessibility. In *HCI in Business, Government, and Organizations: eCommerce and Innovation: Third International Conference, HCIBGO 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016, Proceedings, Part I 3*. Springer, 129–140.

[68] World Health Organization. 2023. *HTML*: Disability. Retrieved Nov 10, 2023 from https://www.who.int/health-topics/disability

[69] Yong Jeong Yi. 2020. Web accessibility of healthcare Web sites of Korean government and public agencies: a user test for persons with visual impairment. *Universal Access in the Information Society* 19, 1 (2020), 41–56.

[70] Norman E Youngblood, Muhammet Fehmi Capanoglu, and Richard Sesek. 2021. The accessibility of state occupational safety and health consultation websites. *Universal Access in the Information Society* 20 (2021), 85–92.

[71] Mengni Zhang, Can Wang, Zhi Yu, Chao Shen, and Jiajun Bu. 2017. Active learning for web accessibility evaluation. In *Proceedings of the 14th International Web for All Conference*. 1–9.