

Categorization and Comparison of Accessibility Testing Methods for Software Development

Aleksander Bai^{a,1} and Kristin FUGLERUD^a and Rannveig A. SKJERVE^b and Till HALBACH^a

^a*Norwegian Computing Center, Oslo, Norway*

^b*OsloMet, Oslo, Norway*

Abstract. There are many methods for testing accessibility and universal design, ranging from checklists and guidelines to automated testing and finally to human testing with participants from different user groups. It is, however, not straight forward to determine how a testing method relates to impairments and barriers. In this work we have expanded the W3C cognitive barriers from one category to four categories in order to provide a better overview of cognitive barriers and testing methods. We also present an overview of multiple existing accessibility testing methods and what kind of barriers they cover. Finally, we present a recommendation on how to select a collection of accessibility testing methods in order to cover the broadest range of disabilities. Our focus is tools that can empower and ease work processes of software development teams, including both developers and testers.

Keywords. Accessibility, barriers, universal design, testing, methods, software development

1. Introduction

The aim of accessibility testing is to evaluate whether the final solution will be accessible for a wide range of people, including people with various types of impairments, such as limitations in vision, hearing, cognition, movement and speech. Accessibility testing is an important part of a process to achieve universally designed solutions [1].

It is challenging, though, to conduct accessibility testing in an efficient way during the development of software and digital services. Studies have identified several challenges to usability testing, such as costs, both in terms of recruiting participants and evaluating the results [2], difficulties in organizing and integration into work processes [3], and lack of competence and training [4]. It is therefore not uncommon for user testing to be neglected in software development [5], [6]. However, companies state that UD is important, and want more focus on accessibility [7].

To further complicate matters, there are few methods, tools and methodologies that cover all types of impairments and barriers, and it is hard for testers to know what types of impairments and barriers they have covered. To ensure a high degree of accessibility,

¹ Corresponding Author, Aleksander Bai, Norwegian Computing Center, Oslo, Norway; E-mail: aleksander.bai@nr.no

it is necessary to conduct user testing with a variety of users, but having the software team testing for accessibility prior to this gives more efficient user testing [1].

There is a wide range of methods, guidelines, and approaches on testing methods for testing accessibility. In this paper, we give an overview of a number of techniques and actual tools. We focus on testing methods that can be conducted by all members of a software development team, categorized according to which kind of barriers they cover.

The remainder of the paper is organized as follows: After the discussion of related work within accessibility testing and software development in Section 2, we propose a particular grouping of barriers to be used for categorizing the accessibility testing methods in Section 3. After that we present numerous testing methods with their features in Section 4, before we discuss our findings in Section 5. We also discuss limitations of our work in Section 6. Finally, we summarize and highlight future research directions in Section 7.

2. Related work

The strengths and weaknesses of various methods have been assessed in previous work.

In a 2005 study, expert review, screen reader, automated testing and remote user testing were compared to each other [8]. The authors found that methods where multiple experts and multiple tools were combined gave the best results. Another finding was that automated web accessibility testing tools and guidelines alone were inadequate for web designers with little accessibility experience.

The correct comparison of tools is not a trivial task. Brajnik proposed the comparison criteria completeness, correctness, and specificity, developed a detailed comparison process and successfully applied this process to the task of comparing two accessibility evaluation tools [9]. The process includes a statistical analysis of the recorded data points in order to be able to quantify error margins, uncertainties, and confidence intervals.

The same author also proposes the criteria validity, reliability, usefulness, and efficiency for the proper comparison of accessibility testing methods [10]. The criteria are applied to compare the methods "barrier walkthrough" and "conformance test". The former is based on scenarios and personas with impairments, and the latter denotes testing according to standards and/or guidelines such as WCAG. The barrier walkthrough is found to achieve better results than the conformance test method. More detailed results on the same experiment are given in [11].

The comparison of methods and tools, however, depends heavily on the used accessibility model, of which multiple definitions exist [12]. Moreover, the named work gives a taxonomy of accessibility evaluation methods and discusses the evaluation of accessibility methods with regard to the author's own definition of accessibility model, as well as the importance of context. It also lists the pros and cons of the methods "conformance review", "subjective assessment", "screening techniques", "barrier walkthrough", and "user testing"; however, on a rather high abstraction level.

The accessibility evaluation method "barrier walkthrough" is discussed in detail in [13]. The results show that there should be three to five experienced experts in order to find all problems on a webpage in a reliable manner. In contrast to this, up to 14 non-experienced reviewers are needed to achieve the same quality level, as their assessments diverge considerably.

Other work focuses on often neglected aspects of accessibility testing methods, namely their inherent costs and benefits [14]. The authors argue that both differ considerably across testing methods, and that both aspects thus have to be taken into account for choosing the appropriate method. Bai have investigated how to prioritize testing methods from a data perspective, by looking at how many issues they find compared to the cost of using the tools [15]. An evaluation of different accessibility testing methods and which methods uncover what type of problems has also been studied by [16].

Specifying that "automated testing" or "semi-automated testing" has been used as evaluation method, however, is simply not enough, as discussed in [17]. Tools, even though they share the same purpose, will likely have differing quality in terms of correctness, completeness, robustness, time consumption, and other parameters. Therefore, it is not only of importance to choose the correct accessibility evaluation method, but also to choose an appropriate tool to be used within a particular method. Also, there are multiple ways in how to apply such tools, and variants in application order [18].

3. Barrier groups

Working with accessibility is a negotiation between diversity and practicality. On the one hand standards and categories are useful tools when advocating for, and working with accessibility. On the other hand they can become too broad or too generalizing, which reduces the diversity in ways of functioning for people in the defined categories [19].

It is very challenging to group or categorize impairments properly, because they are often overlapping in one or several aspects. There are numerous attempts at grouping various impairments, and no clear consensus on the optimal approach, neither in information technology nor in medicine. Impairment categorization is also highly dependent on context.

From a use of ICT solutions perspective, it can make more sense to categorize barriers rather than specific diagnoses, such as dementia or dyslexia, since barriers in ICT are created by a lack of accommodation for different ways of functioning rather than lack of accommodation for a specific diagnosis.

If we are going to describe, evaluate and discuss testing methods for accessibility, we need to have consistent categorization of the various function barriers. We therefore consider function barrier categories from a testing point of view, and not from a medical point of view. We also focus on testing digital solutions and not physical environments, and this will also impact on our categorization.

The W3C group has categorized disabilities into five groups [20], even though they point out that the list is not intended to be an exhaustive listing of all disabilities and barriers. Their focus is web accessibility barriers that people commonly experience:

- Auditory
- Cognitive, learning, and neurological
- Physical
- Speech
- Visual

The W3C definition includes mild or moderate hearing loss in one or both ears (“hard of hearing”) to substantial and uncorrectable hearing loss in both ears (“deafness”) in the auditory category. Some people with auditory disabilities can hear sounds but sometimes not sufficiently to understand all speech, especially when there is background noise. For the cognitive, learning, and neurological disabilities category W3C includes neurodiversity and neurological disorders, as well as behavioral and mental health disorders that are not necessarily neurological.

The physical (also called “motor disabilities”) group include weakness and limitations of muscular control (such as involuntary movements including tremors, lack of coordination, or paralysis), limitations of sensation, joint disorders (such as arthritis), pain that impedes movement, and missing limbs. Difficulty in producing speech that is recognizable by others or by voice recognition software is included in the speech disabilities category. Finally the visual disabilities category includes mild or moderate vision loss in one or both eyes (“low vision”) to substantial and uncorrectable vision loss in both eyes (“blindness”).

The disabled world organization uses eight disability categories [21], and their labelling has a medical perspective:

- Mobility and Physical Impairments
- Spinal Cord Disability:
- Head Injuries - Brain Disability
- Vision Disability
- Hearing Disability
- Cognitive or Learning Disabilities
- Psychological Disorders
- Invisible Disabilities

The categorization of the disabled world has much overlap with W3C, but since their main focus is on rehabilitation it make sense to categorize disabilities into groups that can have the same treatment and rehabilitation.

WHO has created an International Classification of Functioning, Disability and Health (ICF) that categorizes disability into multi-dimensional concepts [22]. The four main domains in ICF are: 1) Body function 2) Body structure 3) Activities and Participation and 4) Environmental Factors. The body function domains is further broken down into:

- Mental functions
- Sensory functions and pain
- Voice and speech functions
- Functions of the cardiovascular, haematological, immunological and respiratory systems
- Functions of the digestive, metabolic, endocrine systems
- Genitourinary and reproductive functions
- Neuromusculoskeletal and movement-related functions
- Functions of the skin and related structures

The categorization of WHO is very broad and complex and can cover a large range of applications. However, their main focus is a medical perspective, and they do not have a good classification scheme for ICT interactions.

Another interesting categorization is based on the 1996/97 Disability Follow-up Survey in Great Britain [23]. This categorization is most relevant for product design. It is developed to support estimating the number of people who are excluded from particular aspects of product use, and to help product designers making informed design decisions. The categories are vision, hearing, cognition, mobility, dexterity and reach.

It is worth noting that there are separate categories for mobility, dexterity and reach. Mobility covers walking, steps and balancing, while dexterity is about grasping and fine-finger manipulation, in both the dominant and non-dominant hand. We believe that it is the latter that has the highest relevance to software usage, and this is covered in the third W3C category "Physical".

Our goal is to find the most common and non-overlapping categorization of function barriers that users experience when using ICT solutions, and we believe that the W3C categorization is the more useful from a software perspective. However, their second category that includes cognitive, learning and neurological barriers can be expanded since it's a very large and complex group as defined by W3C. Researchers have pointed out that W3C WCAG 1.0 and 2.0 have limitations when it comes to addressing cognitive disabilities [24]. W3C acknowledges that this is still a challenge in the newly published WCAG 2.1 [25], by stating that it makes "some accommodation for learning disabilities and cognitive limitations". Also, it could be argued that software in general puts high demands on the cognitive abilities of the user because it usually contains a lot of information.

Another argument for expanding the cognitive or Learning Disability category is because it is a large number of people that experience cognitive functional barriers, and it make sense to break the group into smaller sizes. It is difficult to find accurate number on how many people that experience the different functional barriers, but several studies indicate that cognitive disabilities are more common than the other types of disabilities, such as vision and hearing. Statistics from the US [26] for 2016 explains that while 2.4% of the population have some vision disabilities and 3.5% of the population have hearing disabilities, a whole 4.8% of the population have some form of cognitive disability. This does not, of course, directly relate to functional barriers for ICT solutions, but it gives a strong indication that the cognitive category should be broken down into smaller categories. A final argument for breaking down the cognition category, is that ICT solutions places very high and particular demands on the cognitive functions because it is very information intensive.

We have therefore expanded the W3C categories with some of the mental functions from ICF that are relevant from the perspective of testing of ICT barriers. This includes attention and memory functions which are merged into one category, thought and higher-level cognitive functions which are merged into one category, and finally language and calculation which are also merged into one category. Finally we end up with three cognitive categories instead of one as shown in Table 1:

The merging of functions into categories are based on similarities of barriers that are created. For example a page that requires five successive, correct, operations creates barriers because it requires attention that can last for all five operations *and* a memory of several earlier operations. Hence, attention and memory is grouped in one category.

Table 1. Function barrier categories.

Abbrv.	Name	Description
Aud	Auditory	Barriers created by content that is accessible by audio only. This is related to variations in hearing.

A&M	Attention and memory	Barriers created by content containing complex or long-time operation sequences or invasive sensory content. Relates to variations in attention and concentration. Operations that require sustaining, shifting, dividing and sharing attention and concentration are included in this category, as well as long- and short term memory based operations.
HLL	Higher level logic	Barriers created by complex content containing complex comprehension or problem-solving tasks. Relates to variations in control and content of thought, decision-making, abstract thinking, planning and carrying out plans.
L&N	Language and numbers	Barriers created by dense textual or numerical content, difficult terms, use of inaccessible representation such as fonts, and/or a lack of alternative styling mechanisms. Relates to variations in recognition and use of terms, signs, symbols and other components of language, and approximation and manipulation of mathematical symbols and processes.
Phy	Physical	Barriers created by content that contains one-mode-only operations, such as navigation by mouse only. Relates to variations of muscular control or strength, sensation, joint flexibility, pain related to movement, physical variations of body.
Spe	Speech	Barriers created by content that requires operation by speech such as voice recognition software. Relates to variations in speech production.
Vis	Visual	Barriers created by content that is accessible by visual means only. Relates to variations in vision.

Even though we provide good arguments for why we think our barrier categories are better than W3C, it needs to be verified by how well the categories covers separate and unique barriers. We will discuss this further in Section 5.

4. Testing methods

Based on the barrier categories in Table 1 we wanted to investigate to what extent different accessibility testing methods cover each of the function barrier categories. This will give a better picture of what kind of testing methods that should be used together to cover as many barriers as possible. Existing studies show that a triangulation of methods gives better results than using a single method [1], but there has not been any research that have investigate what kind of barriers that accessibility testing method are actually covering.

There are several methods for testing accessibility in software development; ranging from fully automated tools to manual checklists and user testing [27]. There are also different approaches of grouping these methods, but no clear standard from a software development perspective. In terms of software development, how much time it takes to conduct a test is vital since accessibility testing is often neglected because of constrained resources [28]. We have therefore categorized methods based on their cost in terms of how long they take to complete on average and how different they are. The different method types are listed in Table 2. There are studies that have tried to categorize cost [16], and we have based our work on a similar definition with low, medium, high. Low typically means that a testing method can test a single webpage in less than 5 minutes, while medium between 5 and 30 minutes depending on how well the tester knows the method. High means more than 30 minutes for a single webpage.

Table 2. Method types.

Abbrv.	Cost	Description
Auto	Low	A tool or program that will automatically assess an ICT solution and provide a report of all its findings. A popular example is the Wave automatic checker.

Check	Medium	A checklist or guideline for assessing barriers, like the WCAG standard.
Sim	Low	Simulation using either wearables or use of tools that simulate a barrier. A popular example is the Cambridge Simulation Glasses.
AT	Medium	Use of assistive technology that is normally used by a person to overcome or help with a barrier. The most typical example is a screenreader.
Exp	High	Walkthrough methods that typically requires an expert or someone proficient with the method. Walkthrough includes usability inspection methods like heuristic evaluation and personas testing.

We have evaluated four typical accessibility testing tools for each method type, and an overview can be seen in Table 3. The evaluation of coverage was performed based on the description of barriers in each category. For example, for the category *L&N* a method that includes checks of accessibility in textual and/or numerical content received an x.

Many of the automatic checkers had only one rule or test that within the attention and memory (*A&M*) barrier category. However, it was at most one single test, and we believe it would be unreasonable to make a check for the *A&M* category based on a single test. Each automatic tool has a set of rules, or algorithms that tests single accessibility aspects. A tool was deemed to include a check for a category if rules covering barriers from that category is present in the tools rule set. Only rules that specifically identifies an error were included. For example in the case of HTMLCodeSniffer, the rule with the description "Ensure that the image submit button's alt text identifies the purpose of the button" was excluded because it is a *possible* error, but the tools does not test the content of the alt text. On the other hand, a rule with the description "Image submit button missing an alt attribute" was included because it tests whether the alt attribute is present. HTMLCodeSniffer received and x in the *Vis* column in Table 3 because there were several rules checking the accessibility of visual content.

The checklists were deemed to include a check if one of the checkpoints in the list covered barriers from the category. For example, a checklist containing checks of auditory content received an x in the *Aud* column of the table. Simulation tools were evaluated based on the kinds of barrier it simulated, for example a tool simulating dyslexia received an x in the *L&N* column. Expert testing methods assumes that the expert designing the test is aware of barriers from all categories and includes them in the test design.

Table 3. Accessibility testing methods.

Name	Type	Aud	A&M	HLL	L&N	Phy	Spe	Vis
Android Accessibility Scanner	Auto	-	-	-	-	X	-	X
aXe	Auto	X	-	-	-	X	-	X
WebAIM Wave	Auto	X	-	-	-	X	-	X
HTMLCodeSniffer	Auto	-	-	-	-	X	-	X
WCAG 2.0	Check	X	X	-	X	X	-	X
Electronic and information technology accessibility standards	Check	X	X	-	X	X	-	X
a11y Web Accessibility Checklist	Check	X	-	-	-	-	-	X
WebAim Accessibility Checklist	Check	X	X	-	X	X	-	X
Cambridge Simulation glasses	Sim	-	-	-	-	-	-	X

WebAim Low Vision Simulation	Sim	-	-	-	-	-	-	X
Dyslexia Simulation	Sim	-	-	-	X	-	-	-
Funkify Disability Simulator	Sim	-	X	-	X	X	-	X
Switch Access for Android	AT	-	-	-	-	X	-	-
VoiceOver	AT	-	-	-	-	-	-	X
NVDA	AT	-	-	-	-	-	-	X
High contrast mode	AT	-	-	-	-	-	-	X
Barrier walkthrough	Exp	X	X	X	X	X	X	X
Personas testing	Exp	X	X	X	X	X	X	X
Cognitive walkthrough	Exp	-	X	X	X	-	-	-
Heuristic evaluation	Exp	-	X	X	X	-	-	-

As can be seen from Table 3 the number of categories that are covered varies quite a lot between the different method types. This is expected, since the time and resources required to use the different method types also varies a lot, and also because some of the methods have a focus on particular functions, such as vision in the case of Cambridge Simulation glasses.

5. Discussion

Our categorization of function barriers allows us to investigate which barriers that accessibility testing methods cover. We have shown that it is practical to expand the W3C groups into several cognitive barriers, since this makes it clear that many barriers are inadequately covered by testing methods.

We have marked some of the testing methods with a check for attention and memory *A&M*, but very few have tests or guidelines for memory. In most cases it is the tests for attention that gave a check for this barrier category. This might indicate that we should have memory as a separate barrier category. In the US alone, an estimated 5.4 million have mild cognitive impairments (MCI) not counting dementia [29]. A typical barrier for people with MCI is loss of short-term memory, which support the indication that memory barriers is experienced by many people. Complicated webshops and checkout processes are obvious candidates where testing methods needs to check for memory issues. There were also very few methods that have any tests for barriers in understanding and handling numbers in the *L&N* category, and more tests for number barriers should be developed and implemented.

Also more testing methods for higher level logic (*HLL*) barriers and speech (*Spe*) is required, since very few methods consider that category. This is important since complex pages like webshops and information webpages often make use of higher level logic. Government webpages that explains what kind of tax reduction you are entitled to is a typical example.

Visual (*Vis*) and physical (*Phy*) barriers are the categories that are best covered by all testing methods. Ideally all testing methods should cover all barriers, but this is not required as long as the testers make sure to select a group of methods that covers as many barriers as possible.

The methods of the expert type (*Exp*) are the only testing methods that show very good coverage for all barriers. However, that comes with a cost since they require extensive training and takes longer to conduct than the other method types [30]. This does not mean that they should not be used of course, but we believe it is likely that they will not be selected by software teams with limited time and resources. At best they are maybe used once throughout the project, and probably towards the end before user testing. It also takes time to master these expert methods, as Brajnik have shown that non-experiences do not achieve the same quality level as experienced experts [13]. This means that other method types need to improve their coverage of barriers.

Previously it has been recommended that when software teams are testing for accessibility they should make sure to include at least three different testing methods [1], and make sure to include methods that are supplementing each other in terms of barriers they are covering. Table 3 clearly illustrates the need for combining various methods in order to cover as many barriers as possible. While the testers should also have the cost of utilizing the various methods in mind, the table can be an aid when deciding upon the best combination of methods. The cheaper methods may be conducted more often than the expensive. However, the testers must then be aware of which barriers they are actually testing for.

6. Limitations

We have not verified that the accessibility testing methods actually cover the barriers that they claim to test for. With other words, we are trusting the documentation for each of the methods.

Also, the category "Cost" for method types is set only from an experience point of view, with further influences from other research. The proper estimation and exact quantification of cost is outside the scope of this work.

7. Conclusion

In this work, we have discussed how barriers can be categorized based on existing classifications. We argue for splitting the cognition category into several smaller categories, since this is a large and diverse group. From a testing point of view, it can be beneficial to subdivide the cognitive group into categories that can be tested by different testing methods. The same approach could also be applied to the other barriers, but we have focused on the cognitive barrier in this paper, since this category seems to be weakly covered in the literature.

We have further compared numerous accessibility testing methods for software development. The methods have been categorized into five method types. For each method type we have evaluated four testing methods and shown what barriers they cover. An analysis of some of the most common accessibility testing methods, and to what extent they cover the various barrier categories, has also been presented. The matrix showing the functional barrier categories and testing methods gives a quick overview of which testing method to choose from for a particular barrier. Our goal is that this overview of accessibility testing methods can be used by software development teams when they need to decide what kind of method they should use.

Our work further indicate that software teams should prioritize to educate and train several members of the team in different expert methods, since these are the only methods that cover all barriers.

More studies into other testing methods should be done to provide a better overview of accessibility testing methods and which barriers they cover. It is also quite clear that more testing methods are needed for the cognition domain, and in particular for barriers related to memory, numbers, and higher-level logic. They should be developed and implemented for several types of methods, such as checklists and walkthroughs. Finally, more attention should be given to how to categorize barriers within the cognition domain. We believe it not beneficial to group all cognitive barriers together in one group and have shown some of the consequences in this paper by illustrating the lack of tests for cognitive barriers.

Acknowledgment

This work has been supported by the ITIK project funded by the Norwegian Directorate for Children, Youth and Family Affairs in the UnIKT program, grant number 62127.

References

- [1] K. S. Fuglerud, "Inclusive design of ICT: The challenge of diversity," University of Oslo, Faculty of Humanities, 2014.
- [2] L. C. Cheng and M. Mustafa, "A Reference to Usability Inspection Methods," in *International Colloquium of Art and Design Education Research (i-CADER 2014)*, 2015, pp. 407–419.
- [3] M. Baez and F. Casati, "Agile Development for Vulnerable Populations: Lessons Learned and Recommendations," in *Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Society*, 2018, pp. 33–36.
- [4] C. Putnam, M. Dahman, E. Rose, J. Cheng, and G. Bradford, "Best practices for teaching accessibility in university classrooms: cultivating awareness, understanding, and appreciation for diverse users," *ACM Trans. Access. Comput.*, vol. 8, no. 4, p. 13, 2016.
- [5] J. Nielsen and S. Gilutz, *Usability return on investment*. Nielsen Norman Group, 2003.
- [6] B. Bygstad, G. Ghinea, and E. Brevik, "Software Development Methods and Usability: Perspectives from a Survey in the Software Industry in Norway," *Interact. Comput.*, vol. 20, no. 3, pp. 375–385, 2008.
- [7] Larry Goldberg, "Teaching Accessibility: A Call to Action from the Tech Industry." 2015.
- [8] J. Mankoff, H. Fait, and T. Tran, "Is your web page accessible?: a comparative study of methods for assessing web page accessibility for the blind," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2005, pp. 41–50.
- [9] G. Brajnik, "Comparing accessibility evaluation tools: a method for tool effectiveness," *Univers. Access Inf. Soc.*, vol. 3, no. 3–4, pp. 252–263, 2004.
- [10] G. Brajnik, "Web accessibility testing: when the method is the culprit," in *International Conference on Computers for Handicapped Persons*, 2006, pp. 156–163.
- [11] G. Brajnik, "A comparative test of web accessibility evaluation methods," in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, 2008, pp. 113–120.
- [12] G. Brajnik, "Beyond Conformance: The Role of Accessibility Evaluation Methods," *Proceedings of the 2008 international workshops on Web Information Systems Engineering*. Springer-Verlag, Auckland, New Zealand, 2008.
- [13] G. Brajnik, Y. Yesilada, and S. Harper, "The expertise effect on web accessibility evaluation methods," *Human-Computer Interact.*, vol. 26, no. 3, pp. 246–283, 2011.
- [14] T. H. Røssvoll and K. S. Fuglerud, "Best practice for efficient development of inclusive ICT," *HCI International 2013*. Springer, Las Vegas, Nevada, USA, 2013.
- [15] A. Bai, H. C. Mork, and V. Stray, "A Cost-Benefit Analysis of Accessibility Testing in Agile Software Development Results from a Multiple Case Study," *Int. J. Adv. Softw. Vol. 10, Number 1*

- 2, 2017, 2017.
- [16] A. Bai, H. C. Mork, T. Schulz, and K. S. Fuglerud, "Evaluation of accessibility testing methods. which methods uncover what type of problems?," *Stud. Health Technol. Inform.*, vol. 229, pp. 506–516, 2016.
- [17] T. Halbach and W. Lyszkiewicz, "Accessibility Checkers for the Web: How Reliable are they, actually?," in *Proceedings of 14th International Conference on {WWW/Internet} ({ICWI})*, 2015.
- [18] T. Halbach and K. S. Fuglerud, "On Assessing the Costs and Benefits of Universal Design of ICT," in *Studies in Health Technology and Informatics*, 2016, pp. 662–672.
- [19] H. Persson, H. Åhman, A. A. Yngling, and J. Gulliksen, "Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects," *Univers. Access Inf. Soc.*, vol. 14, no. 4, pp. 505–526, 2015.
- [20] W3C, "W3 Diversity of web users." <https://www.w3.org/WAI/intro/people-use-web/diversity>.
- [21] Disabled World, "Disability: Definiton, Types and Models." <https://www.disabled-world.com/disability/types/>.
- [22] World Health Organization, "International Classification of Functioning, Disability and Health (ICF)." <http://apps.who.int/classifications/icfbrowser/>.
- [23] S. D. Waller, M. D. Bradley, P. M. Langdon, and P. J. Clarkson, "Visualising the number of people who cannot perform tasks related to product interactions," *Univers. Access Inf. Soc.*, vol. 12, no. 3, pp. 263–278, 2013.
- [24] J. Borg, A. Lantz, and J. Gulliksen, "Accessibility to electronic communication for people with cognitive disabilities: a systematic search and review of empirical evidence," *Univers. Access Inf. Soc.*, vol. 14, no. 4, 2014.
- [25] W3C, "Web Content Accessibility Guidelines (WCAG) 2.1." .
- [26] Cornell University, "Disability Status Report United States." 2016.
- [27] N. Ghasemifard, M. Shamsi, A. R. R. Kenari, and V. Ahmadi, "A new view at usability test methods of interfaces for human computer interaction," *Glob. J. Comput. Sci. Technol.*, 2015.
- [28] J. Nielsen, "Return on investment for usability," *Jakob Nielsen's Alertbox, January*, vol. 7, 2003.
- [29] B. L. Plassman et al., "Prevalence of cognitive impairment without dementia in the United States," *Ann. Intern. Med.*, vol. 148, no. 6, pp. 427–434, 2008.
- [30] C. Wharton, J. Bradford, R. Jeffries, and M. Franzke, "Applying cognitive walkthroughs to more complex user interfaces: experiences, issues, and recommendations," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1992, pp. 381–388.