The Simulated User Impairment Testing (SUIT) Protocol and Toolbox for Digital Artifacts

Joschua Thomas Simon-Liedtke[1] [0000−0003−4809−1688] and Till Halbach[1] [0000−0002−9566−7013]

Norwegian Computing Center, Gaustadalléen 23A, 0373 Oslo, Norway
{joschua,halbach}@nr.no
https://www.clothes4all.net/?

Abstract. User testing involving people with varying abilities and impairments are two of the key ingredients of universal design that can help to increase the accessibility and usability of an ICT artifact. However, recruiting adequate participants during the design and development process can be difficult, time-consuming and resource intensive. We suggest mitigating these challenges by employing user simulation as additional quality assessment measure throughout the design and development process of a digital ICT artifact.
In the present paper, we present a protocol and a toolbox for Simulated User Impairment Testing (SUIT) that enables non-impaired designers and developers to simulate the perception of people with some of the most common impairments related to cognitive, visual, auditory, vocal, mobility-related and motor-related functioning. We compile a list of possible tools and methods that are readily available or easily obtainable in a production and design setting. Moreover, we complement this list with relevant user examples for each category, as well as examples of the most common barriers. Finally, we point out how the Simulated User Impairment Testing (SUIT) toolbox can be used to add an additional layer of quality assessment during the design and development process of a digital artifact and argue that it cannot replace the user testing part.

Keywords: Universal Design · Accessibility · Simulation · Impairments · Simulations

1 Introduction

Around 15% of the population in the world is living with some sort of impairment [37]. The number of people with impairments in Western societies is expected to increase due to increased life expectancy and ageing [35]. Thus, universal design, defined as the process of designing and implementing systems that are accessible and usable for the widest range of users possible, including people with impairments, has increased its importance in communities and societies around the globe [32,21]. Additionally, it has been shown that increased accessibility and usability of a product or service will benefit non-disabled users [30,29]. Moreover, universal design has become a legal requirement in many countries around the world. The United Nations' Convention on the Rights of Persons with Disabilities has been ratified by many of their member states, and national lawmakers have passed laws for the inclusion of people with impairment1 [16,38]. The Norwegian anti-discrimination law states, for example, that ICT solutions like websites are universally designed [15,16]. The Norwegian lawmakers, and likewise the legislative in other countries like the USA or UK have implemented legal obligations in practice by using a system of audit and certification to a set of guidelines or standards like the Web Accessibility Guidelines (WCAG) [11,42].

User evaluations including people with impairments and varying abilities into the design and development process is one of the key elements of universal design to achieve increased accessibility and usability [4,5]. Even though guidelines are a good starting point to eliminate common technical barriers, user-based evaluations involving real users interacting with an artifact are still necessary to detect and mitigate barriers while making the artifact as a whole more usable [28]. In the ideal case, such user evaluations are conducted with people of different abilities including people with impairments. However, such evaluations involving a wide set of users are often considered to be too time or resource intensive [28]. Thus, supplement methods that can be conducted by singular experts or developers have been proposed including the use of personas, and models or simulations, especially during the early stages of a development process [26,28,2].

We argue that the use of simulation methods can be an effective tool to increase accessibility and usability of an artifact, at the same time as simulations can easily be integrated by designers and developers in the early stages of design process without the need to rely on direct input from external users. On the one hand, the founders of Universal Design published guidelines spanning the spectrum of human abilities in architecture, and individual simulation methods for ICT are mentioned sporadically throughout the literature [33,1,32]. On the other hand, there is no comprehensive framework including practical examples for simulations for the whole spectrum of human abilities applicable in ICT.

In the present paper, we want to close the aforementioned gap by introducing a protocol and a toolbox for Simulated User Impairment Testing (SUIT) spanning the whole spectrum of human abilities relevant for the consumption of, usage of, and interaction with ICT, e.g., cognitive, sensory, voice and movement-related abilities. The SUIT toolbox includes concrete simulation methods that are practically applicable by designers and developers.

In the following sections, we will (I.) review relevant literature about universal design strategies and evaluation methods including guidelines, expert evaluations, personas, and simulations, (II.) introduce the protocol for Simulated User Impairment Testing (SUIT), (III.) list tools and examples for the practical implementation of the SUIT toolbox, (IV.) present a trial study of the SUIT protocol and toolbox evaluating an online artifact used as reference project for the visualization of WCAG 2.0 and 2.1 called Clothes4All, and (V.) discuss some general observations and recommendations collected during our trial study.

2 Background

One strategy to increase accessibility and usability for implementing universal design of a digital artifact is interactive, user-centered design processes with the following iterative phases [24]: (I.) the understanding of users, tasks, and contexts, (II.) the design of the artifact, (III.) the prototyping the artifact, (IV.) the evaluation of the artifact, and (V.) the integration and final implementation of the artifact. A prerequisite of such a design process is inclusive design, which means the inclusion of people with impairments and other marginalized groups as evaluators [24]. Most common evaluation methods include [24,28]: (I.) user testing through for example observations or thinking-aloud techniques; (II.) guidelines, checklists, and automated checking tools; (III.) expert evaluation, (IV.) cognitive walkthroughs, (V.) personas; and (VI.) simulations. These methods are not mutually exclusive. An expert can conduct an evaluation with or without simulation, cognitive walkthroughs can follow checklists, and so on. In an ideal world, user testing would always be the prioritized choice, as it allows detailed insight and first-hand experience by users of the relevant groups. However, user testing can sometimes be impractical, time consuming, resource intensive, and expensive, as well as qualified evaluators with impairments can be difficult to recruit [24,28,19]. Thus, we want to shortly present some of the alternatives that can supplement some user evaluation in some iterations of the design process.

2.1 Guidelines

Accessibility guidelines are tools that help developers to implement websites more easily in the most technically accessible and usable way possible. Guidelines provide best-practice solutions for common challenges and help to avoid common pitfalls. Examples for such guidelines are the Web Accessibility Guidelines (WCAG) that have been proposed in multiple versions [42,43]. WCAG 2.0 have been enshrined in law my the national lawmaker in Norway, while WCAG 2.1 is the standard in the Web Accessibility Directive (WAD) passed by the European Union (EU) [15,9].

2.2 Expert evaluations and personas

Other alternatives for user testing include experts in the field that follow either none, general or detailed guidelines or the usage of Personas [24,26]. Personas are defined as fictional characters representing individual users or user groups [26,7]. Personas were originally rough sketches used in marketing. However, these original Personas had the downsides of not being believable enough, having the characters being communicated poorly, not having provided a clear instruction on how to use the characters, and lacking high-level support within a company [26]. Eventually, Personas evolved to more detailed characters by including statistical data and anecdotes and observations from interviews with target users, having the Personas embedded within the design process and providing clear instructions on why, when, and how to use them [26]. Personas provide a variety of advantages and disadvantages: On the pro side, Personas can create a strong focus on users and work contexts through a fictionalized setting that can create engagement in designers and developers [26]. Moreover, creating Personas can help to make assumptions about the target audience more explicit [26]. Overall, Personas can help to focus the attention on a specific target group by creating a powerful medium for communicating results and improving the understanding of the target group [26]. On the contra side, it can be challenging to compile the right Persona or set of Personas for a specific target group [26]. What is more, designers and developers can be tempted to reuse existing Personas when it would be more appropriate to create new ones [26]. Likewise, overusing personas may replace other valid methods like user-testing, or product evaluation (leading to a so-called "Persona mania") [26]. Moreover, different groups or roles in the design process might have different needs requiring diverging Persona attributes or target audiences [26]. A marketing department, for example, might have different wishes to Personas than for example product development.

2.3 Comparison of user testing, expert evaluation, and personas

Expert evaluations and Persona testing are often considered to be insufficient, as they are not considering users and their needs well enough [28]. Especially persona testing is not considered to be an empirical method and requires decision-making that is likely subject to personal bias by the person interpreting any given persona [26]. For personas with impairments, a tester might base decisions on subjective assumptions that are not based in reality. In contrast, user testing involving users with impairments is often considered to be time or resource intensive, at the same time as it might be difficult to get access to relevant users [28,25]. Furthermore, it can be difficult to set up the testing environment in a way that is accessible for users with impairments [28].

To safe resources and minimize the problem of subjective bias, the testing can be supported by methods that simulate a given impairment. In other words, expert testing that includes simulations can be a feasible alternative to support expert testing or Persona testing. In this paper, we define simulation methods as models that can emulate the perception or experience of people with impairment as accurate as possible to a person without the impairment. Thus, simulations are most relevant where models are possible and economical [24]. Simulation-based testing is helpful in earlier stages of the design process before deploying user testing [3].

2.4 Impairments to consider

The World Health Organization (WHO) published a document defining the whole spectrum of human functions and abilities [41]. Based on this document, we defined categories of human impairments that are relevant for ICT: (I.) Cognition [(i.) Vision, (ii.) Hearing, (iii.) Touch], (II.) Senses, (III.) Voice, (IV.) Movement [(i.) Mobility, (ii.) Motor] [12].

3 The Simulated User Impairment Testing (SUIT) Protocol

The main goal of Simulated User Impairment Testing (SUIT) is to discover possible barriers that people with impairments can encounter when interacting with a digital

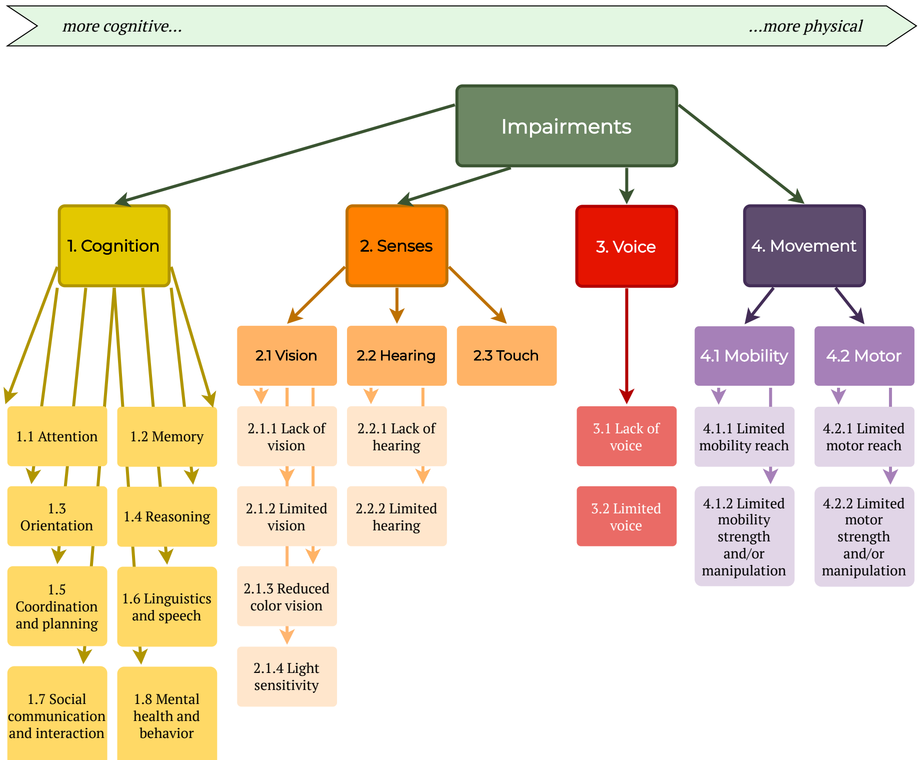


Fig. 1: The impairment categories considered for SUIT.

ICT artifact. In this paper, we are targeting digital artifacts in ICT, a term that encompasses all interactive digital products, systems and applications in the realm of ICT. Typical representations of digital artifacts are, for example, web sites, mobile apps, museum exhibitions, or automats. In other words, we intent to evaluate how accessible and usable the artifact is for people with impairments. A minor goal is to let a non-disabled developer experience an artifact through the eyes and body of a person with impairment. Thus, it is important to emulate the experience of users with impairment as accurate as possible. Another minor goal is to raise awareness around the needs and experiences of people with impairments.

Our Simulation User Impairment Testing (SUIT) Protocol and Toolbox is informed by strategies from the literature, as well as practical experiences from accessibility and usability testing that we have conducted throughout the years. The SUIT protocol is one key element that consists of five major steps:

(1.) The tester defines one or multiple tasks connected to the artifact. Tasks can be related to retrieving information, completing a routine, communication and more.

(2.) The tester defines a metric according to which they intent to evaluate the outcome of the tasks. Such metrics can, for example, measure the completion rate, completion time, accuracy, record the number of observed errors or issues, or simply record qualitative data from observational input.

(3.) The tester defines one or multiple impairments, and chooses an adequate simulation tool to emulate the experience of people with the chosen impairment.
(4.) The tester completes the task(s) with the chosen impairment(s) while recording qualitative observations and relevant data according to the chosen metric.
(5.) The tester evaluates the results from the metric.

We suggest including the SUIT protocol throughout the design and development phase as an additional accessibility and usability evaluation, as well as an additional quality assessment step. Moreover, a tester should apply multiple iterations of the SUIT Protocol throughout the design and development process of a digital artifact. It has been shown that a combination of automated checking tools, checklists, expert evaluation, and simulation methods can uncover a high percentage of accessibility and usability issues with respect to their cost [2]. Furthermore, the SUIT protocol is easily compatible with an agile workflow [2]. It can be beneficial to for example have multiple evaluation cycles in which the tester uses automated tools, followed by simulations, checklists, and finally expert evaluations [2]. It is not necessary to apply the complete set of tools represented in the SUIT toolbox at each step of the evaluation. It is possible to rather focus on one or two major impairments during each iteration of the SUIT protocol.

4 The SUIT toolbox

In the following paragraphs, we will describe the various user impairment categories in more detail, including relevant user groups from each subcategory and typical barriers users from those user groups might encounter. Moreover, we list tools, equipment and environments that can be used to carry out simulated user testing. This list includes various assistive devices that users with impairments might employ when interacting with websites online. A system developer or designer should ensure that information on their artifact can be accessed and retrieved both with and without simulating one of the impairments mentioned above (cf. Figure 1).

4.1 Cognitive and mental functions

Cognitive and mental functions are functions of the brain, including both global and specific mental functions, as well as various emotional functions [41]. They can be subcategorized into attention, memory, orientation, reasoning, coordination and planning, linguistics and speech, social communication and interaction, and mental health and behavior.

Attention is defined as all "specific mental functions of focusing on an external stimulus or internal experience for the required period of time" [41] encompassing all functions related to concentration, focus, and motivation. *Examples* of individuals with attention difficulties are people with attention deficit hyperactivity disorder (ADHD) and plain attention deficit disorder (ADD). We are not aware of any typical *assistive technology (AT)* used by this group, but it is worth mentioning that distractions can be reduced by removing distracting devices like mobile phones and services such as social media. Examples of potential *barriers* include situations with complex and over-complicated graphical screen layouts, lengthy text, moving, blinking or flickering content, or background audio [40]. We propose the following *simulation strategies*:

(1.) Use the artifact with an audio source in the background playing at a sufficiently high level.
(2.) Use the artifact with an additional screen on the side, playing for instance a movie or showing a TV program.
(3.) Use the artifact browser extension like Web Disability Simulator (concentration) [20].

Memory is defined as all "specific mental functions of registering and storing information and retrieving it as needed" [41] encompassing both short-term and long-term memory. *Examples* of individuals with memory difficulties are those with short-term or long-term memory loss, dementia, Alzheimer's, and similar [40]. In many cases, ADHD and its variant ADD can have an impact on memory functions as well. We are not aware of any typical *AT* used by this group, but it is worth mentioning that for instance reminder and note taking applications may in general support an individual's memory functions. Examples of potential *barriers* include situations with lengthy text, multiple-step task and problem solving, complex and over-complicated graphical screen layouts, and similar [40]. We propose the following simulation strategies:

(1.) For every instance of use, solve a new task that involves new, unfamiliar parts of the artifact without help and instructions [32].
(2.) Perform multiple-step tasks on the artifact out of order [32].
(3.) In the middle of the solving of a multiple-step task, start over again [32].

Orientation is defined as all "general mental functions of knowing and ascertaining one's relation to self, to others, to time and to one's surroundings" [41]. For digital artifacts, orientation is the ability to orientate to particular targets and to navigate within a given artifact. *Examples* of individuals with orientation difficulties are people with seizure disorders, as well as dementia patients [40]. We are not aware of any typical *AT* used by individuals in this impairment category. Examples of potential *barriers* include situations with complex and over-complicated graphical screen layouts, as well as multiple-step tasks and problem solving. We propose the following simulation strategies:

(1.) For every instance of use, solve a new task that involves new, unfamiliar parts of the given artifact without help and instructions [32].
(2.) Perform multiple-step tasks on the artifact out of order [32].
(3.) Make deliberate mistakes, and try to correct them afterwards [32].
(4.) In the middle of the solving of a multiple-step task, start over again [32].

Reasoning encompasses consciousness function "of the state of awareness and alertness, including the clarity and continuity of the wakeful state" [41], thought functions "related to the ideational component of the mind" [41], and intellectual functions "required to understand and constructively integrate the various mental functions, including all cognitive functions and their development over the lifespan" [41]. For digital artifacts, reasoning is the ability to understand, fathom and explain concepts, ideas, and other abstract aspects of a given artifact, such as an ICT application or system. *Ex-*

amples of individuals with reasoning difficulties are people with learning impairments, individuals with Down's syndrome, and others [40,18]. We are not aware of any typical *AT* used by this group. Examples of potential *barriers* include situations with complex and over-complicated screen layouts, lengthy sentences and unusual terms [40] We propose the following simulation strategies:

(1.) For every instance of use, solve a new task that involves new, unfamiliar parts of the given artifact without help and instructions [32].
(2.) Perform multiple-step tasks on the artifact out of order [32].
(3.) Make deliberate mistakes, and try to correct them afterwards [32].
(4.) In the middle of the solving of a multiple-step task, start over again [32].

Coordination and planning encompasses higher-level cognitive functions "including complex goal-directed behaviors such as decision-making, abstract thinking, planning and carrying out plans, mental flexibility, and deciding which behaviors are appropriate under what circumstances; often called executive functions" [41]. For digital artifacts, coordination and planning describes the ability to plan strategies, make decisions, execute tasks and evaluate outcomes of one's behavior. *Examples* of individuals with coordination and planning difficulties are individuals with Down's syndrome [18]. We are not aware of any typical *AT* used by this group. Examples of potential *barriers* include situations with multiple-step problem solving and change of problem-solving strategies [18]. We propose the following simulation strategies:

(1.) For every instance of use, solve a new task that involves new, unfamiliar parts of the given artifact without help and instructions [32].
(2.) Perform multiple-step tasks on the artifact out of order [32].
(3.) Make deliberate mistakes, and try to correct them afterwards [32].
(4.) In the middle of the solving of a multiple-step task, start over again [32].

Linguistics and speech encompasses "specific mental functions of recognizing and using signs, symbols and other components of a language" [41]. This category has some intersections with Section 4.3. *Examples* of individuals with speech or linguistics difficulties are people with dyslexia, non-native speakers, speakers with strong accents or dialects, shy/quiet speakers, and individuals with Down's syndrome [40,18] Typical *AT* used by this group are text-to-speech converters, speech synthesizer, digital grammar assistants, translation and lookup services, and similar. Examples of potential *barriers* include situations with complex and over-complicated graphical screen layouts, complex sentences and unfamiliar vocabulary, long text without illustrations, moving, blinking or flickering content, or background audio [40] We propose the following simulation strategies:

(1.) Use the website as if you cannot read [32].
(2.) Use uncommon accents or dialects.
(3.) Use browser extensions like Web Disability Simulator (dyslexia, small vocabulary, etc.) [20].

Social communication and interaction functions are defined as "general mental functions [. . .] required to understand and constructively integrate the mental functions that lead to the formation of the interpersonal skills needed to establish reciprocal social interactions, in terms of both meaning and purpose" [41]. For digital artifacts, this social communication and interaction encompasses the ability to ask for support or help from, interact with, or take contact with others. *Examples* of individuals with social communication and interaction difficulties are Autism spectrum disorder (ASD) [40]. We are not aware of any typical *AT* used by this group. Examples of potential *barriers* include situations with complex and over-complicated graphical screen layouts, complex sentences and unfamiliar vocabulary, long text without illustrations, moving, blinking or flickering content, or background audio [40] We are not aware of any simulation strategies at the moment.

Mental health and behavior encompasses emotional functions "related to the feeling and affective components of the processes of the mind" [41]. temperament and personality functions "of constitutional disposition of the individual to react in a particular way to situations, including the set of mental characteristics that makes the individual distinct from others" [41]. *Examples* of individuals with mental health and behavior difficulties are anxiety, stress, delirium, depression, paranoia, schizophrenia [40]. We are unaware of any typical *AT* used by this group. Examples of potential *barriers* include situations with complex and over-complicated graphical screen layouts, complex sentences and unfamiliar vocabulary, long text without illustrations, moving, blinking or flickering content, or background audio [40] We propose the following simulation strategies:

(1.) For every instance of use, solve a new task that involves new, unfamiliar parts of the given artifact without help and instructions [32].
(2.) Perform multiple-step tasks on the artifact out of order [32].
(3.) Make deliberate mistakes, and try to correct them afterwards [32].
(4.) In the middle of the solving of a multiple-step task, start over again [32].
(5.) Let yourself get distracted by a radio, or TV, or let yourself be interrupted by a colleague. [32].
(6.) Use a browser extension like the Funkify disability simulator [10].

4.2 Senses

In this section, we discuss the human senses relevant for digital artifacts including seeing, hearing, and touching [41].

Vision addresses all functions related "to sensing the presence of light and sensing the form, size, shape and color of the visual stimuli" [41]. We can subcategorize vision impairments into lack of vision, low vision, reduced color vision, and light sensitivity

Examples of individuals with a lack of vision are those who are blind or deafblind [40]. Typical *AT* used by this group are screen readers, refreshable braille displays, magnifiers, and similar. Examples of potential *barriers* include situations where no text alternative is given for visual elements, missing text or audio alternative for videos, or browsing without full keyboard support [40]. We propose the following simulation strategies:

(1.) Use the artifact with screen reader (and turned-off monitor): Jaws, NVDA, VoiceOver, TalkBack, Voice Assistant and others [23,31].
(2.) Use the artifact with a keyboard only.
(3.) Use the artifact with a non-graphical browser.

Examples of individuals with limited or low vision are those with reduced focus or acuity (i.e., increased blurriness), tunnel vision, central field loss, clouded vision, as well as those affected by cataract, glaucoma, age-related macular degeneration, retinitis pigmentosa, stroke, diabetes retinopathy, and others [40]. Typical *AT* used by this group are magnifiers (both optical and digital), screen readers, and separate color schemes. Examples of potential *barriers* include situations where it is impossible to resize text or images sufficiently, change contrast, or use custom color schemes [40]. We propose the following simulation strategies:

(1.) Increase the distance between eye and monitor, or reduce the browser zoom.
(2.) View the artifact through a tube.
(3.) View the artifact with one eye closed or with peripheral vision only [32].
(4.) View the artifact with a single or multiple layers of Cambridge Simulation Glasses [39].
(5.) Use a browser extension: NoCoffee, Web Disability Simulator, SEE, A11Y, and others [17,20,27,34].

Examples of individuals with reduced color vision are those affected by color vision deficiencies, including deuteranomaly, and protanomaly [40]. Typical *AT* used by this group are special color schemes and frequency band switching glasses. Examples of potential *barriers* include situations where differentiation between two elements is only possible by color alone, or layout ordered by color alone [40]. We propose the following simulation strategies:

(1.) Use low-brightness glasses.
(2.) Use a browser extension: NoCoffee, Web Disability Simulator, SEE, A11Y, and others [17,20,27,34].
(3.) Use of special user style sheets.

Examples of individuals with light sensitivity are people with seizure disorders, including epilepsy and migraines, and people with conditions like retina detachment and meningitis [40]. Typical *AT* used by this group are light-filtering glasses and shields, as well as special color schemes like dark mode. Examples of potential *barriers* include constellations with high-brightness screens, viewing screens in combination with stray light or strong reflections, and visual flickering at certain frequencies and with particular patterns [40]. We propose the following simulation strategies:

(1.) View the artifact with maximum screen brightness.
(2.) View the artifact with a direct strong light source from above, below or from the sides.

Hearing encompasses all functions related "to sensing the presence of sounds and discriminating the location, pitch, loudness and quality of sounds" [41]. We can subcategorize hearing impairments into lack of hearing, and limited hearing.

Lack of hearing is synonymous with deafness [40]. Typical *AT* used by this group are cochlear implants, and transcript and caption services, as well as speech-to-text converters and services [6]. Examples of potential *barriers* include situations where applications rely on the voice alone, such as calling services, and a lack of sign language support [40]. We propose the following simulation strategies:

(1.) Use a muted speaker or headset output.

(2.) Use high-fidelity noise cancellation headsets.

Examples of individuals with a limited hearing function are people with hearing aids, implant-operated, people affected by Tinnitus, as well as many elder individuals and people with Down's syndrome [14,40]. Typical *AT* used by this group are various hearing aids including conversation amplifiers and teleloop systems, various assistive listening and alerting devices, speech-to-text and transcription services, and media playing and volume controls [22,40]). Examples of potential *barriers* include situations where audio content is presented without captions or transcripts, situations with a high level of background noise, and services which lack media playing and volume controls. We propose the following simulation strategies:

(1.) Use reduced speaker or headset volume for media and sound playback.
(2.) Listen to the audio source with background noise.
(3.) Use high-fidelity noise cancellation headsets (with or without additional background noise), or ear plugs.

Touch is defined as all functions related to "sensing surfaces and their texture or quality" [41]. *Examples* of individuals with decreased touch sensation are those affected by numbness (hypoesthesia) caused by illnesses and injuries. We are not aware of any typical *AT* used by this group. Examples of potential *barriers* include situations where alerts are only signaled through vibration, and where temperature is only communicated through heat radiation. We propose the following simulation strategies:

(1.) Use gloves or mittens to decrease sensitivity [32].
(2.) Use the artifact with wet or oily fingers [32].

4.3 Voice

encompasses all functions related to producing sounds and speech [41]. Vocal functions can be sub-categorized into lack of voice, and limited voice.

Examples of individuals with limited voice and lack of voice include people with muteness, stuttering, cluttering, dysarthria, speech sound disorder, individuals affected by stroke, foreign language speakers, and those speaking language dialects [32,40]. Typical *AT* used by this group are augmentative and alternative communication / AAC devices, picture boards, and text-to-speech converters [22]. Examples of potential *barriers* include situations where device control relies on interaction by voice alone, systems that rely on phone lines as only means of communication, home assistants, as well as noisy environments [40]. We propose the following simulation strategies:

(1.) Use without voice, or muted / lowered microphone on headset.
(2.) Use speech distortion.
(3.) High background noise (for interference with voice input).

Examples of limited voice include stuttering, cluttering, dysarthria, speech sound disorder, and similar, stroke-affected, people using a electrolarynx, non-native language speakers, people with a low speaking voice, and those speaking with strong dialects [40,32]. Typical *AT* used by this group are augmentative and alternative communication or AAC devices, picture boards, and text-to-speech converters [22]. Examples of potential *barriers* include situations where device control relies on interaction by voice alone, systems that rely on phone lines as only means of communication, home assistants, as well as noisy environments [40]. We propose the following simulation strategies:

(1.) Use a voice distorter.
(2.) Use the artifact while lowering the microphone input.
(3.) Use uncommon accents or dialects.
(4.) Use the website with loud background noises.

4.4 Movement

This section addresses all human functions related to movement and mobility including functions of joints, bones, reflexes and muscles [41].

Mobility or lower body functions encompasses all functions and impairments related to the lower body structures (legs, feet, pelvis, etc.) and their related actions (standing, walking, getting up, or sitting down) [36,35]. Mobility functions can subcategorized into limited mobility reach, and limited mobility strength and/or manipulation.

Examples of individuals with limited mobility reach are wheelchair users, amputees, and people suffering from arthritis [32]. Typical *AT* used by this group are wheelchair, crutches, and protheses of various fidelity [32]. Examples of potential *barriers* are stairs, stepping thresholds, narrow aisles, and so on. We propose the following simulation strategies:

(1.) Use the artifact in a wheelchair (manual and/or electrical).
(2.) Use the artifact while using a stroller or walker.
(3.) Use the artifact while using crutches.
(4.) Use the artifact while resting on one leg.

Examples of individuals with limited mobility strength and/or manipulation are people of extreme size or weight, those suffering from Parkinson's disease or multiple sclerosis, stroke or similar, as well as some affected by cerebral palsy [32]. Typical *AT* used by this group are crutches [32]. Examples of potential *barriers* include situations where one has to rise from a seated position or stand upright, games with a stepping mat, and exhibits that need to be controlled by a foot or feet [32]. We propose the following simulation strategies:

(1.) Use the artifact with crutches.
(2.) Use the artifact while resting on one leg.
(3.) Use the artifact with balancing shoes or B-shoes.

Motor or upper body functions are defined as all functions and impairments related to the upper body structures (arms, hands, neck, back, etc.) and their related actions (lifting, grasping, or pushing/pulling) [36,35]. Motor functions can be sub-categorized into limited motor reach, and limited motor strength and/or manipulation.

Examples of individuals with limited motor reach are people with amputations, those suffering from arthritis, carpal tunnel syndrome, and others, as well as conditions related to cerebral palsy [32]. Typical *AT* used by this group are head-operated pointing devices, mouth stick, on-screen keyboard with trackball, joystick, or switch-operated by foot, shoulder, or head [40]. Examples of potential *barriers* include situations that require two arms or two hands for interaction, and situations with the need to push a button or lever or to turn a wheel [32]. We propose the following simulation strategies:

(1.) Use the artifact with the hand tied to body through rubber band, or elbow fixated against the body [32].
(2.) Use the artifact with a rope with weight around wrist [32].
(3.) Use the artifact with a single hand or single arm [32].
(4.) Use a Browser Extension: Web Disability Simulator [20].

Examples of individuals with limited motor strength and/or manipulation are those suffering from Parkinson's disease, multiple sclerosis, stroke, people with muscular dystrophy, tremor and spasms, as well as repetitive stress injury / RSI, and many affected by cerebral palsy [32,40]. Typical *AT* used by this group are ergonomic keyboard or mouse, voice recognition, eye tracking, and exoskeleton [40]. Examples of potential *barriers* for these users include situations that require precise input on screen or keyboard, repeated input over time, no full or limited keyboard support, insufficient time limits, faulty navigation mechanisms and page functions, and the simple need to push a button or turn a wheel [32,40]. We propose the following simulation strategies:

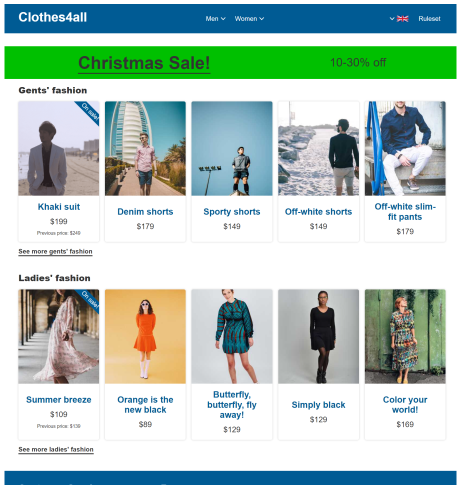
(1.) Use the artifact with plastic or simulation gloves (such as Cambridge Simulation Glasses, with various material thicknesses, increased push resistance, larger area of touch, and similar [39]).
(2.) Use the artifact with mittens or fists.
(3.) Use the artifact with single (unfamiliar) finger of the dominant and non-dominant hand [32].

5 Pilot study of SUIT on Clothes4All

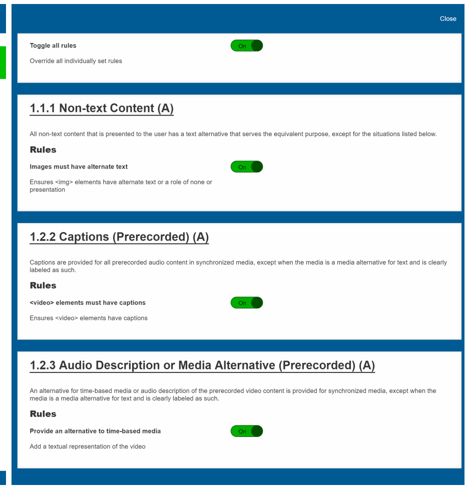
We tested the protocol on an artifact we developed, called Clothes4All, which is website that can be used to study different aspects of web accessibility. The artifact is a fictive online shop that sells clothes for all, and can be accessed under `www.clothes4all.net` (cf. Figure 2a). The website shows how the WCAG 2.1 can be implemented in practice, at the same time as it gives the user the option to toggle on/off conformance to various WCAG success criteria (cf. Figure 2b). By toggling WCAG success criteria on/off, the user can investigate how a website with high conformance to WCAG can compare to a website with low conformance to WCAG, at the same time as experienced developers and designers can inspect the source code of the website for best-practice examples.

When designing the pilot study of SUIT on Clothes4All, we defined a set of tasks relevant to the content of the website:

> (1.) Scenario 1: Obtaining information about the products. You want to know more about the various offers and deals.
>   (a.) What is the product for women with the lowest price currently on sale? How much do you save compared to previously?



(a) The homepage of Clothes4All, a fictive online shop selling clothes that fit all.



(b) Examples for some of the WCAG 2.1 rules that users can toggle on/off.

> (b.) Is the leather coat for men on sale?
> (c.) Are the slim-fit denim jeans for women on sale?
> (d.) …
>
> (2.) Scenario 2: Buying men's clothing. You want to buy some men's clothing. Complete the following tasks. Please start with an empty shopping basket.
> (a.) How many categories of men's fashion do you find?
> (b.) How many fashion items for men are in the following categories:
> i. Suits.
> ii. Blazers.
> iii. Shirts.
> (c.) …
>
> (3.) …
>
> *Examples of possible tasks for Clothes4All*

Then, we defined a metric according to which the tasks should be evaluated:

> (1.) Could you solve the task? Yes / No (If you cannot solve the task within XX minutes, you should mark the case as unsolvable.)
> (a.) If No: What prevented you from completing the task? What barriers did you encounter?
>
> (2.) How much time did you use for solving the task? Measure the so-called time-on-task, e.g. from the beginning of solving the task until either it is completed or abandoned.

> (3.) Please write down some descriptions, explanations, or comments around the process.
> (4.) How satisfied are you with the process? Pick a number on a scale from 1 to 10, where 1 is least satisfied, and 10 is most satisfied. Set a number on how you experienced the process. Think of the variety of feelings that may have been involved, e.g. joy, frustration, like, dislike, curiosity, surprise, or confusion.
>
> *The SUIT metric for Clothes4All*

Finally, we customized the SUIT protocol (cf. section 3) for the Clothes4All website:

> (1.) Choose a single or multiple user impairments that you want to simulate. Please consult Section 2.4 for a list of impairments.
> (2.) Pick the appropriate tools, e.g. assistive technologies (ATs) or equipment, from the SUIT toolbox. A collection of tools for any particular impairment is included in Section 4.
> (3.) Turn off one or several of the associated success criterion switches in the settings / rule set of Clothes4All.
> (4.) Now pick one task of your choice from the provided list. Try to solve the task once with the relevant success criterion switches turned off as described in the previous step (this equals the first trial), and then with all switches turned on (equals the second trial). Write down your observations according to the metric defined before.
>
> *The SUIT protocol for Clothes4All*

During the test trial, we went through all categories and picked one simulation tool for each category. We then recorded the data according to the results. The goal of this trial run was to investigate the feasibility and practicability of the SUIT protocol and toolbox, rather than uncovering barriers on the website. After all, the website was designed in a way that it offered one version that was more accessible and usable than the others. Thus, we knew where we could expect to find barriers and challenges. In other words, the purpose of the trial was not to discover barriers themselves, but to investigate whether or not the protocol and the tools where appropriate to discover barriers and challenges on the artifact.

6 Discussion

To begin with, we could conduct the trial successfully simulating each of the categories mentioned in the toolbox above (cf. Section 4). However, performance and relevance differed for some of the categories. Simulations of sensory impairments performed quite well, and sensory tools could easily be obtained. There is, for example, a variety of models and tools that predict various visual impairments like color vision deficiencies and others. Thus, we could choose among multiple tools, plug-ins and products to simulate the experience of people with sensory impairments. The same can be said about auditory impairment simulation tools. Likewise, movement-related impairments simulation tools were easily available and applicable in a regular office setting. At the same time, the motor-related simulation tools were more relevant for our digital artifact than the mobility-related tools. The reason for this is the fact that the user interacts mostly with their upper body extremities, e.g., the arms and hands, rather than their lower body extremities. In contrast, tools representing cognitive impairments had limited availability, and questionable accuracy. Cognitive process can be very subjective, and models are not readily available. Some of the tools or methods for simulation described in this paper can be very vague or inaccurate. Thus, the results should be interpreted carefully.

Moreover, the relevance for some of the tools depend strongly on the type of digital ICT artifact that is investigated. We were testing on a website artifact that mainly stimulates visual and auditory senses while relaying on input from upper body extremities. At the same time, the tasks required significant cognitive interaction. Lower body and touch-related functions were not as relevant for the investigated artifact. This could be different for other digital ICT artifacts like museum exhibitions, where users are encouraged to watch, touch, interact with, walk into, lift, or manipulate the artifact [13]. Testing for wheelchair accessibility is for example more relevant for a museum exhibition that a user can walk into than it is for a website. We want to evaluate the relevance of the SUIT protocol and toolbox for other areas by testing them on other digital ICT artifacts like museum exhibits in future research.

One of the strengths of the SUIT toolbox is that it can help to raise awareness and increase visibility for people with impairments and their needs. While using a screen reader, a tool that is very common among people with low-vision or blindness, we became aware of how missing labels and alt-texts can significantly decrease the user experience of a given digital ICT artifact. Likewise, retrieving auditory information from a video without subtitles while the sound was turned off was obviously impossible. We argue that designers and developers using the simulation tools will more easily understand the importance of accessibility and usability guidelines for their products. Likewise, the tools can increase the acceptance and dissemination of accessibility and usability guidelines, checklists, and best-practice examples.

One challenge of the SUIT toolbox was the accuracy of the categorization. The list of simulation methods in Section 4 does not represent a complete set of simulation methods for the experiences that people with impairments may encounter. In the same way, the (sub)categories do not represent all possible impairments. The (sub)categories can be organized differently depending on a given project or user needs. For online artifacts, for example, mobility barriers are less relevant than for museum exhibits [13]. It is very common that the categorization of human impairments differ throughout the literature [28,44,13,8].

Moreover, we intended to integrate the SUIT simulation toolbox into a meaningful testing routine for accessibility and usability following clearly defined procedures. As described in Section 3, designers and developers should define concrete tasks and their desired outcome beforehand. The outcome should be measurable according to observable metrics like pass/fail, success rate, or completion time. By defining observable metrics, testing results will become comparable over time, as well as allowing developers and designers to compare and weigh various options and solutions against each other. These comparisons will not only help development within any ongoing project but have the potential of being transferable to other future or previous projects as well.

Finally, the SUIT simulation toolbox is not intended to replace thorough user testing including people with impairments. It is intended as additional quality assessment during the production and development process that has the potential of mitigating some of the most basic barriers that can arise during the development of a project. Thus, it is still necessary to conduct multiple iterative cycles with user testing involving people with impairments to increase accessibility and usability of the digital ICT artifact. One reason for this is that simulation methods will never be able to capture the complete experience of people with impairments. People that use simulation tools for a limited time span, will experience digital ICT artifacts differently than people who must live with the limitations of an impairment on a daily basis. People with impairments have, for example, strategies and mechanisms to cope with and mitigate some of the most common challenges and barriers of typical ICT artifacts. These strategies might be unknown to non-disabled people who can develop alternative coping strategies that might be irrelevant for people with impairments. In other words, user testing with people with impairment is still the most important necessity to increase accessibility and usability of a digital ICT artifact for all.

7 Conclusion

In this paper, we present a protocol and toolbox for Simulated User Impairment Testing (SUIT) that allows designers and developers to experience the perspectives of users with impairments during the development of a digital ICT artifact through simulation tools. By using simulation tools, designers and developers can assess their digital ICT artifact through the eyes and bodies of a person with impairments and varying abilities. With The tools of the Simulated User Impairment Testing (SUIT) toolbox, designers and developers can simulate the experience of people with cognitive, visual, auditory, vocal and movement-related impairments. We show how these tools can be used to evaluate the design process of a digital ICT artifact using the SUIT protocol. The observations from our prototype trial show that the tools and methods can be easily implemented and integrated into an agile design and development workflow. We show that the SUIT toolbox can meet the economical and practical demands and restrictions imposed by the industry. At the same time, we observed that sensory impairments, related to vision, hearing, and touch, as well as physical impairments could be easily simulated. However, there were only a few reliable simulations for cognitive impairments. Finally, we argue that this kind of simulation approach can raise awareness for and acceptance of the needs and experiences of people with impairments. In future research, we see the need to extend the simulation strategies specifically for cognitive impairments. In the present paper, we could only test the SUIT protocol and toolbox on websites, which was within the scope defined in our research project called Clothes4All. However, other digital ICT artifacts like museum exhibits can have more focus on other human functions like mobility or motor than websites. Thus, we want to test the SUIT protocol and toolbox on other digital ICT artifacts like museum exhibits in future research as well.

Acknowledgement

This work was partly supported by the UnIKT program of the Norwegian Directorate for Children, Youth and Family Affairs.

References

1. Bai, A., Fuglerud, K.S., Skjerve, R.A., Halbach, T.: Categorization and Comparison of Accessibility Testing Methods for Software Development. Studies in Health Technology and Informatics **256**, 821–831 (Jan 2018), https://europepmc.org/article/med/30371447
2. Bai, A., Mork, H.C., Stray, V.: A cost-benefit evaluation of accessibility testing in agile software development. ICSEA 2016 p. 75 (2016)
3. Bai, A., Skjerve, R., Halbach, T., Fuglerud, K.: Evaluating accessibility testing in automated software build processes. Norsk IKT-konferanse for forskning og utdanning (Nov 2019), https://ojs.bibsys.no/index.php/NIK/article/view/646
4. Begnum, M.E.N.: Universal Design of ICT: A Historical Journey from Specialized Adaptations Towards Designing for Diversity. In: Antona, M., Stephanidis, C. (eds.) Universal Access in Human-Computer Interaction. Design Approaches and Supporting Technologies. pp. 3–18. Lecture Notes in Computer Science, Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-49282-3_1
5. Bonacin, R., Dos Reis, J.C., Baranauskas, M.C.C.: Universal Participatory Design: Achievements and Challenges. SBC Journal on Interactive Systems **10**(1), 2–16 (2019)
6. Clerk Center, Gallaudet University: Assistive Technology. https://clerccenter.gallaudet.edu/national-resources/info/info-to-go/assistive-technology.html (2017), last checked on February 24th, 2022.
7. Cooper, A.: The inmates are running the asylum. In: Software-Ergonomie'99. Springer (1999)
8. European Telecommunications Standards Institute (ETSI): EN 301 549 v3.1.1 (2019-06): Accessibility requirements for ICT products and services (Jun 2019)
9. European Union (EU): Web Accessibility Directive (WAD) - Directive (EU) 2016/2102. https://eur-lex.europa.eu/eli/dir/2016/2102/oj (Oct 2016), last checked on February 23rd, 2021.
10. Funkfiy: Funkify. https://www.funkify.org/ (2022), https://www.funkify.org/, last checked on February 24th, 2022.
11. Giannoumis, G.A.: Implementing Web Accessibility Policy. Case Studies of the United Kingdom, Norway, and the United States. Ph.D. thesis, The University of Bergen (Jun 2019), http://hdl.handle.net/1956/21054
12. Halbach, T., Simon-Liedtke, J.T.: Categories of User Impairment . In: IADIS International Conference on Interfaces and Human Computer Interaction 2021 (part of MCCSIS 2021) (2021), https://nr.brage.unit.no/nr-xmlui/handle/11250/2766973
13. Halbach, T., Tjøstheim, I.: Towards Reliable Accessibility Assessments of Science Center Exhibits. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) Human-Computer Interaction – INTERACT 2019. pp. 33–41. Lecture Notes in Computer

Science, Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-29381-9_3

14. Hørselshemmedes Landsforbund (HLF): Om HLF. https://www.hlf.no/kontakt-oss/om-hlf/ (2022), last checked on February 24th, 2022.
15. Kommunal- og moderniseringsdepartementet: Forskrift om universell utforming av informasjons- og kommunikasjonsteknologiske (ikt)-løsninger (for-2013-06-21-732). https://lovdata.no/dokument/SF/forskrift/2013-06-21-732 (Jun 2013), last checked on 2020-09-17.
16. Kulturdepartementet: Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven, lov-2017-06-16-51). https://lovdata.no/dokument/NL/lov/2017-06-16-51 (Jun 2017), https://lovdata.no/dokument/NL/lov/2017-06-16-51, last checked on 2020-09-17.
17. Leventhal, A.: NoCoffee – Vision Simulator for Chrome . https://accessgarage.wordpress.com/2013/02/09/458/ (Feb 2013), last checked on February 24th, 2022.
18. Malt, E.A., Dahl, R.C., Haugsand, T.M., Ulvestad, I.H., Emilsen, N.M., Hansen, B., Cardenas, Y.E.G., Skøld, R.O., Thorsen, A.T.B., Davidsen, E.M.M.: Health and disease in adults with Down syndrome. Tidsskrift for Den norske legeforening (Feb 2013). https://doi.org/10.4045/tidsskr.12.0390, https://tidsskriftet.no/en/2013/02/health-and-disease-adults-down-syndrome
19. McFadden, E., Hager, D.R., Elie, C.J., Blackwell, J.M.: Remote Usability Evaluation: Overview and Case Studies. International Journal of Human–Computer Interaction **14**(3-4), 489–502 (Sep 2002). https://doi.org/10.1080/10447318.2002.9669131
20. Metamatrix: Web Disability Simulator - Chrome Nettmarked. https://chrome.google.com/webstore/detail/web-disability-simulator/olioanlbgbpmdlgjnnampnnlohigkjla (2019), last checked on February 24th, 2022.
21. Miljøverndepartementet: T-1468 B/E Universell utforming (Nov 2007), https://www.regjeringen.no/no/dokumenter/t-1468-universell-utforming/id493083/
22. National Institute on Deafness and Other Communication Disorders (NIDCD): Assistive Devices for People with Hearing, Voice, Speech, or Language Disorders. https://www.nidcd.nih.gov/health/assistive-devices-people-hearing-voice-speech-or-language-disorders (Nov 2019), last checked November 9th, 2020.
23. Norges Blindeforbund: Skjermleser. https://www.blindeforbundet.no/hjelpemidler-og-produkter/skjermleser (Mar 2021), last checked on February 24th, 2022.
24. Petrie, H., Bevan, N.: The Evaluation of Accessibility, Usability, and User Experience. The universal access handbook **1**, 1–16 (2009)
25. Power, C., Petrie, H.: Working With Participants. In: Yesilada, Y., Harper, S. (eds.) Web Accessibility, pp. 153–168. Human–Computer Interaction Series, Springer, London (2019). https://doi.org/10.1007/978-1-4471-7440-0_9
26. Pruitt, J., Grudin, J.: Personas: practice and theory. In: Proceedings of the 2003 conference on Designing for user experiences. pp. 1–15. DUX '03, Association for Computing Machinery, San Francisco, California (Jun 2003). https://doi.org/10.1145/997078.997089
27. Q42: SEE. https://github.com/Q42/SEE. (Jun 2014), last checked on February 24th, 2022.
28. Sauer, J., Sonderegger, A., Schmutz, S.: Usability, user experience and accessibility: towards an integrative model. Ergonomics **63**(10), 1207–1220 (Oct 2020).

https://doi.org/10.1080/00140139.2020.1774080, https://doi.org/10.1080/00140139.2020.1774080
29. Schmutz, S., Sonderegger, A., Sauer, J.: Implementing Recommendations From Web Accessibility Guidelines: Would They Also Provide Benefits to Nondisabled Users. Human Factors **58**(4), 611–629 (Jun 2016). https://doi.org/10.1177/0018720816640962
30. Schmutz, S., Sonderegger, A., Sauer, J.: Effects of accessible website design on nondisabled users: age and device as moderating factors. Ergonomics **61**(5), 697–709 (May 2018). https://doi.org/10.1080/00140139.2017.1405080
31. Statlig spesialpedagogisk tjeneste (Statped): Skjermleserne TalkBack og Voice Assistent for Android. https://www.statped.no/laringsressurser/syn/opplaring-i-mobil-og-nettbrett/android-for-synshemmede/skjermleserne-talkback-og-voice-assistent-for-android/ (2019), last checked on February 24th, 2022.
32. Story, M.F., Mueller, J.L., Mace, R.L.: The Universal Design File: Designing for People of All Ages and Abilities. Revised Edition. Center for Universal Design, NC State University, Box 8613, Raleigh, NC 27695-8613 ($24) (1998), https://eric.ed.gov/?id=ED460554
33. Stray, V., Bai, A., Sverdrup, N., Mork, H.: Empowering agile project members with accessibility testing tools: a case study. In: International Conference on Agile Software Development. pp. 86–101. Springer, Cham (2019)
34. The A11Y Project: The A11Y Project. https://www.a11yproject.com/ (2021), https://a11yproject.com/, last checked on February 24th, 2022.
35. The US Census Bureau: Americans with Disabilities: 2010 (Jul 2012), https://www.census.gov/library/publications/2012/demo/p70-131.html
36. The US Census Bureau: Americans with Disabilities: 2014 (Nov 2018), https://www.census.gov/library/publications/2018/demo/p70-152.html
37. United Nations High Commissioner for Refugees (UNHCR): Disability: Prevalence and Impact - Syrian Arab Republic. Humanitarian Needs Assessment Programme (HNAP) (Sep 2019), a Nationwide Household Survey Using Washington Group Methodology
38. United Nations (UN): Convention on the rights of persons with disabilities (crpd) (2006), https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html
39. University of Cambridge: Cambridge Simulation Glasses. http://www.inclusivedesigntoolkit.com/csg/csg.html (2017), last checked on February 24th, 2022.
40. W3C Web Accessibility Initiative (WAI): Diverse Abilities and Barriers (Oct 2020), https://www.w3.org/WAI/people-use-web/abilities-barriers/, last checked on October 15th, 2020.
41. World Health Organization (WHO): International Classification of Functioning, Disability and Health (ICF) (May 2001), https://www.who.int/classifications/icf/en/, last checked 2020-11-10.
42. World Wide Web Consortium (W3C): Web content accessibility guidelines (wcag) 2.0. https://www.w3.org/TR/WCAG20/ (Dec 2008), last checked on 2020-09-17.
43. World Wide Web Consortium (W3C): Web content accessibility guidelines (wcag) 2.1. https://www.w3.org/TR/WCAG21/ (Jun 2018), last checked on 2020-09-17.
44. World Wide Web Consortium (W3C): Web content accessibility guidelines (wcag) 2.2 [working draft]. https://www.w3.org/TR/WCAG22/ (Aug 2020), last checked on 2020-09-17.