POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Usability Evaluation Project

Evaluated website: www.unicef.org

Authors: **Emanuele Musto 10691229**

Riccardo Speroni 10671842

Matteo Spreafico 10669138

Manuel Tacca 10707009

Course name: **Hypermedia Applications (Web and Multimedia)**

Academic Year: 2023-2024

Abstract

The aim of this Usability Report is to describe the design, the execution and the results of the usability assessment carried out on the UNICEF website. Specifically, the analysis is performed using the Inspection technique and then the User Testing technique. The first technique involves expert evaluators examining the application and evaluating the compliance with recognized usability principles called heuristics. Such analysis is conducted with reference to the Nielsen heuristics and to a subset of the MILE heuristics. The second technique consists of the data observation and collection of how some representatives of real users interact with the system. Its goal is to discover the actual difficulties encountered by the users when interacting with the website and to obtain systematic feedback on its effectiveness and usability.

Contents

1 | Inspection

Usability Inspection refers to a category of methods used to systematically evaluate the usability of an application, using a team of **expert evaluators** who have knowledge in the field to carry out the examination. In this section, we will focus on heuristics evaluation, which uses a set of well-known and standardized rules to highlight possible problems. Inspectors will typically thoroughly explore the application, evaluating and judging the compliance with the chosen heuristics.

1.1. Inspection Design and General Method

The methodology and the design applied for the inspection part of the project is described in this section. Each evaluator carries an individual, in-depth analysis of the website and the UX-related aspects of the application, by referring to the heuristics defined in the following sections. The focus is on the content layout, the navigation elements, and the overall presentation (look and feel of the application).

The usability evaluation was carried out on all the pages strictly belonging to the www.unicef.org domain. This means that pages such as www.data.unicef.org are not considered in this evaluation, even if they are reachable from the pages of the considered domain. This choice was made due to timing constraints, as the evaluated website already comprises a large number of pages that have revealed some usability issues.

1.2. Heuristics Definition

This section defines the heuristics used to conduct the inspection. Such heuristics are the **Nielsen heuristics** and a subset of the **MiLE heuristics**, which are well-known in the literature. Each heuristic has been assigned a category, according to its focus:

- **Content**. Content heuristics evaluate the quality of the actual information content of the application. This includes text, images, and information quantity.

- **Navigation**. Navigation heuristics evaluate how easily the user can navigate and

go from topics to other related ones.

- **Presentation**. Presentation heuristics refer to the way the information is presented within the application. They evaluate how the topology of information is implemented, whether information is correctly positioned with respect to consistency, hierarchy and aesthetic.

We will also assign each heuristic a code, as shown by the following tables.

Nielsen's heuristics

Code	Category	Heuristic
H1	Navigation	Visibility of system status/bread crumps
H2	Presentation	Match between system and the real world
H3	Navigation	User control and freedom
H4	Presentation	Consistency and standards
H5	Presentation	Error prevention
H6	Presentation	Recognition rather than recall
H7	Navigation	Flexibility and efficiency of use
H8	Presentation	Aesthetic and minimalist design
H9	Presentation	Help users recognize, diagnose and recover from errors
H10	Content	Help and documentation

Table 1.1: Codes for Nielsen's heuristics

MiLE's heuristics

Code	Category	Heuristic
M1	Content	Information overload
M2	Content	Consistency of page content structure
M3	Content	Contextualized information
M4	Content	Content organization (hierarchy)
M5	Navigation	Interaction consistency
M6	Navigation	Group navigation-1
M7	Navigation	Group navigation-2
M8	Navigation	Structural navigation
M9	Navigation	Semantic navigation

M10	Navigation	Landmarks
M11	Presentation	Text lay out
M12	Presentation	Interaction placeholders-semiotics
M13	Presentation	Interaction placeholders-consistency
M14	Presentation	Consistency of visual elements
M15	Presentation	Hierarchy-1
M16	Presentation	Hierarchy-2
M17	Presentation	Spatial allocation-1
M18	Presentation	Spatial allocation-2
M19	Presentation	Consistency of page spatial structure

Table 1.2: Codes for MiLE's heuristics

1.2.1. Nielsen's Heuristics

Jakob Nielsen's heuristics (1994) are a set of rules useful for any interactive system, and define the principles to apply in the development of a website. Below is a brief but precise explanation of each one, along with the category it belongs to.

H1 - Visibility of system status/bread crumbs (*Navigation*)

This heuristic evaluates whether and how the design of the application keeps the user informed of what is going on, at all times and within a reasonable amount of time. Ways of complying with the heuristics are, for instance, the use of bread crumbs or orientation maps.

H2 - Match between system and the real world (*Presentation*)

This heuristic evaluates how the information appears in a natural and logical order, and whether it is conveyed in a language that is familiar to the user.

H3 - User control and freedom (*Navigation*)

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. The user must be able to always be in control of what they are doing, rather than being forced through a lengthy procedure to make changes.

H4 - Consistency and standards (*Presentation*)

Users should not have to wonder whether different words, situations, or actions mean the same thing. The application should use well-known conventions and should be consistent with them in all its pages and contents.

H5 - Error prevention (*Presentation*)

This heuristic evaluates how the application avoids error prone situations and ensures that, before critical paths or decisions are taken, confirmation options are present. Other than making error correction easy, errors should be avoided as much as possible.

H6 - Recognition rather than recall (*Presentation*)

This heuristic evaluates how the application provides visible options and instructions at all times so that the user does not have to remember information from one part of the dialogue to another. Recognition must be properly implemented so that the user always has information and context accessible (i.e. they are not lost).

H7 - Flexibility and efficiency of use (*Navigation*)

This heuristic evaluates how the system provides ways of use for all types of users - both novices and experts - and how the user can tailor their experience and frequent actions on the site.

H8 - Aesthetic and minimalist design (*Presentation*)

This heuristic evaluates how the application provides only the needed information to the user during dialogues, and how its interfaces are clear of elements that do not support its goal: the presentation of unnecessary or rarely needed information can be confusing and can alter the perceived importance of actually relevant aspects.

H9 - Help users recognize, diagnose and recover from errors (*Presentation*)

Error messages should be expressed in plain language (avoiding error codes), precisely indicating the problem and constructively suggesting a solution.

H10 - Help and documentation (*Content*)

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

1.2.2. MiLE's Heuristics

MiLE's heuristics are a larger set of heuristics which focuses on different aspects in a more specific manner. We have used a given, relevant subset of those for the task at hand. Again, below is a brief but precise explanation of each one, along with the category it belongs to.

M1 - Information overload (*Content*)

This heuristic evaluates whether the quantity of information contained in a page is adequate: it shouldn't be too much as it would overwhelm the user, but it shouldn't be too little either as pages need to be useful.

M2 - Consistency of page content structure (*Content*)

This heuristic evaluates whether pages that present topics of the same category have the same type of elements across the site.

M3 - Contextualized information (*Content*)

This heuristic evaluates whether the page contains adequate information to remind the user of where they are in the site, in order to avoid them feeling lost.

M4 - Content organization (hierarchy) (*Content*)

This heuristic evaluates if the hierarchy of the contents and elements of the page is adequate to the relevance they hold: the most important elements should be higher in the topological hierarchy. This can be achieved with spatial disposition of contents, their dimension, or other strategies to convey their relevance or lack thereof.

M5 - Interaction consistency (*Navigation*)

This heuristic evaluates whether pages of the same type have the same navigation links and the same interaction capabilities across the site.

M6 - Group navigation-1 (*Navigation*)

This heuristic evaluates whether it is easy to navigate from, among, and within "groups" of items (e.g. from the group to its members and the other way around, among different groups, and among members of the same group).

M7 - Group navigation-2 (*Navigation*)

This heuristic evaluates whether menus create cognitive overload.

M8 - Structural navigation (*Navigation*)

This heuristic evaluates the ease of navigation between various sections or parts of the same topic.

M9 - Semantic navigation (*Navigation*)

This heuristic evaluates how easily users are able to navigate between related topics, in all directions from one element to a related one.

M10 - Landmarks (*Navigation*)

This heuristic evaluates whether the landmarks (i.e. links available on all the pages of a given website) actually aid in recognizing and reaching the most relevant parts of the website.

M11 - Text lay out (*Presentation*)

This heuristic evaluates how well the text can be read and whether the font size is appropriate.

M12 - Interaction placeholders-semiotics (*Presentation*)

This heuristic evaluates whether interaction elements are easy to understand, and if the symbols or icons used to present them actually convey their functional meaning.

M13 - Interaction placeholders-consistency (*Presentation*)

This heuristic evaluates whether the interaction elements, namely their icon or textual elements, are consistent across the site.

M14 - Consistency of visual elements (*Presentation*)

This heuristic evaluates whether, in pages of the same type, visual elements have the same meaning and functional properties.

M15 - Hierarchy-1 (*Presentation*)

This heuristic evaluates whether the on-screen placement of contents reflects their importance and relevance.

M16 - Hierarchy-2 (*Presentation*)

This heuristic evaluates whether the on-screen placement of visual elements reflects their importance and relevance.

M17 - Spatial allocation-1 (*Presentation*)

This heuristic evaluates whether semantically related elements are close to each other.

M18 - Spatial allocation-2 (*Presentation*)

This heuristic evaluates whether semantically distant elements are placed distant from one another.

M19 - Consistency of page spatial structure (*Presentation*)

This heuristic evaluates whether pages of the same type have the same spatial organization for the various elements that are presented.

1.3.   Metrics Definition

For each presented heuristic, inspectors assign a score which determines how compliant the site is. The metrics used for the inspection are defined by the following table.

Inspection metrics

Score	Meaning
N/A	The heuristic is not applicable
1	The heuristic is **not** satisfied: severe violations have been detected
2	The heuristic is partially satisfied but it is poorly implemented and must be significantly improved
3	The heuristic is partially satisfied but it can be improved
4	The heuristic is almost fully satisfied but there are some imperfections
5	The heuristic is **fully** satisfied: no violation has been detected

Table 1.3: Inspection metrics definition

1.4. Inspection Execution

To carry out the evaluation, each inspector spent a minimum of 4 hours thoroughly exploring the website, carefully analyzing both main and secondary pages. After that each evaluator gave their scores, the inspectors gathered and discussed the individual inspection results. In the following section, aggregated data about the results of the examination are shown and commented, along with the final scores for each heuristic, which have been mediated and agreed taking into account all the involved evaluators.

1.5. Inspection Results

This section reports the evaluations and scores of the heuristics described above, separated by category, firstly with an overview, and lastly with the details of single heuristics, the correlated screenshots and explanations of the identified problems.

1.5.1. Content Heuristics

In this section the evaluated heuristics are all regarding the content category. Following, a visualization and a table summarizing each heuristic and the score attributed and agreed between evaluators.

Figure 1.1: Average and agreed scores by inspectors

Content Heuristics

Code	Heuristic	Score
H10	Help and Documentation	5
M1	Information overload	3
M2	Consistency of page content structure	3
M3	Contextualized information	2
M4	Content organization (hierarchy)	4
Average	**Content**	**3.4**

Table 1.4: Scores for content heuristics

H10 - Help and documentation

Code	Heuristic	Average Score	Agreed Score
H10	Help and documentation	4.5	5

The only documentation provided for this website, and the only one that makes sense to provide, is the FAQ section. Even if it may be not trivial to reach for some people, since the easiest way to get there is through the footer, its content is precise and explains what needs to be known for the user.

Frequently Asked Questions

Find answers to the most frequently asked questions about UNICEF

Available in: English Français Español العربية 中文

- What is UNICEF?
- Where does UNICEF work?
- What does the acronym UNICEF stand for?
- How can I make a donation to UNICEF?
- Where can I get the latest data on issues affecting children?
- Where can I find UNICEF publications?
- Where does UNICEF get its funding?
- How is UNICEF accountable for its work?
- What is the UNICEF Executive Board?
- How can I apply for a job with UNICEF?

Figure 1.2: FAQ Page

M1 - Information Overload

Code	Heuristic	Average Score	Agreed Score
M1	Information Overload	2.25	3

In many pages information is too much, especially in the "what we do" section of the website, which is full of references and other suggested links that make the pages not very concise. Also, the menu is pretty long and verbose.

The actual content of pages is still well organized and for many pages (for example articles) all the information is important.

WHAT WE DO ⌃   RESEARCH AND REPORTS ⌄   STORIES ⌄   ABOUT UNICEF ⌄

Focus areas		How we do it	Global advocacy
All areas	Gender	Partnerships	Action on the climate crisis
Adolescent development	Health	Innovation	Better mental health: #OnMyMind
Child protection	Humanitarian emergencies	Supply and logistics	Child health and survival
Child rights	Immunization	Young people	Children under attack
Children with disabilities	Migrant and refugee children	Humanitarian Action for Children	Tackling the learning crisis
Climate change and environment	Nutrition	Evaluation	
COVID-19 response	Social and behaviour change		
Early childhood development	Social policy		
Education	Water, sanitation and hygiene		

Figure 1.3: Menu Overload

M2 - Consistency of Page Content Structure

Code	Heuristic	Average Score	Agreed Score
M2	Consistency of Page Content Structure	3.5	3

For many pages that belong to the same category, the heuristic is satisfied. There are however many other pages that belong to the same section, that are similar in functionality and content type, but differ one from another, for example pages from the "what we do" section. The thing that creates confusion the most is the fact that the categorization offered by the website (page, appeal, programme...) is not very clear and intuitive, and it's not explained neither, so pages that are considered similar to the user, may not be considered similar for the website.

🖹 Page

Convention on the Rights of the Child

For every child, every right.

Figure 1.4: Page in "What we do" section

⊡ Programme

Children with disabilities

Every child has the right to live in an inclusive world.

Figure 1.5: Programme in "What we do" section

M3 - Contextualized Information

Code	Heuristic	Average Score	Agreed Score
M3	Contextualized Information	2.25	2

This heuristic is only partially satisfied, in some pages only there is a breadcrumb that indicates in what part of the website the user is. Apart from the pages with the bread crumbs, the only way to tell which part of the website the user is navigating, is to figure it out by reading the text and make a guess. Sometimes it's enough because the titles are coherent with the navigation bar links, some other times it's not enough.

Page with breadcrumbs in "Stories" section

Page without breadcrumbs in "Stories" section

M4 - Content Organization (hierarchy)

Code	Heuristic	Average Score	Agreed Score
M4	Content Organization (hierarchy)	4	4

The heuristic is overall satisfied, the elements inside pages are organized hierarchically and the most relevant content is emphasized. The only flaw is the dimension of the images, that are way too big and force users to scroll down in order to see more relevant content.

1.5.2. Navigation Heuristics

In this section the evaluated heuristics are all regarding the navigation category. Following, a visualization and a table summarizing each heuristic and the score attributed and agreed between evaluators.
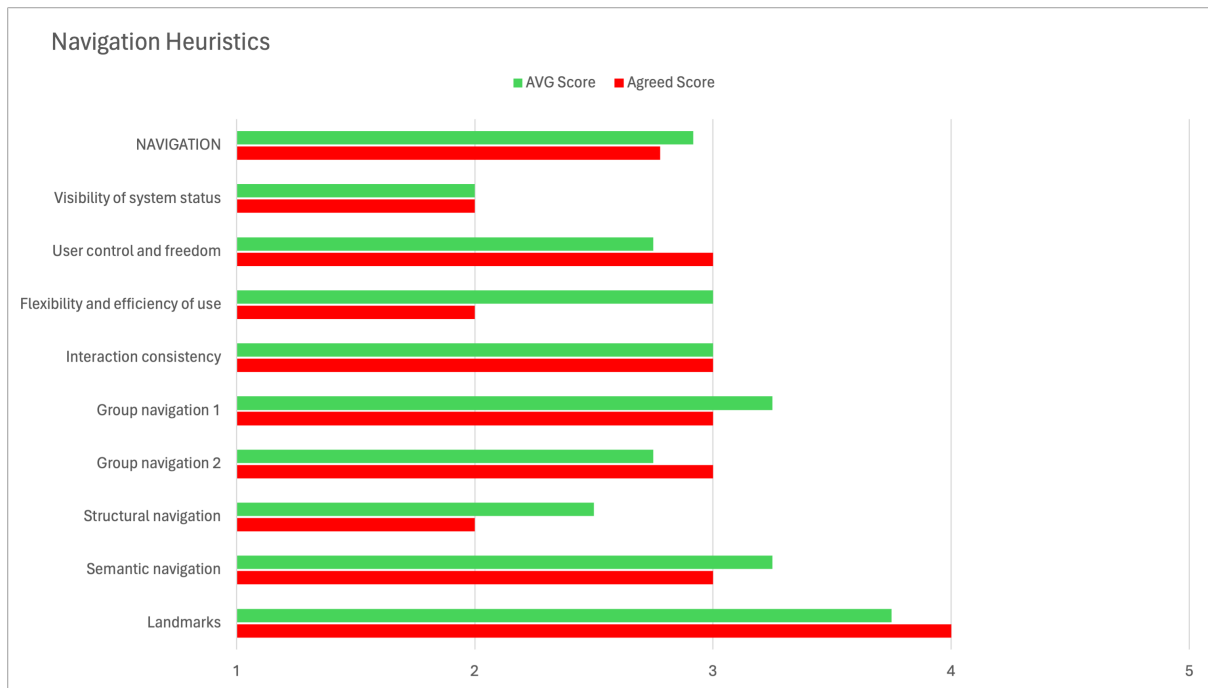
Figure 1.6: Average and agreed scores by inspectors

Navigation Heuristics

Code	Heuristic	Score
H1	Visibility of system status	2
H3	User control and freedom	3
H7	Flexibility and efficiency of use	2
M5	Interaction consistency	3
M6	Group navigation 1	3
M7	Group navigation 2	3
M8	Structural navigation	2
M9	Semantic navigation	3
M10	Landmarks	4
Average	**Navigation**	**2.78**

Table 1.5: Scores for navigation heuristics

H1 - Visibility of System Status

Code	Heuristic	Average Score	Agreed Score
H1	Visibility of System Status	2	2

Breadcrumbs are the main tool used by the website to show the system status, as there are not other status-based functionalities. However they are present only in some pages, in many other pages the system does not provide information about his status. Moreover, breadcrumbs are not that much intuitive, especially with pages located deep in the navigation tree. For newcomers, also, it's difficult to understand if the page belongs to the main website, or to a sub-section of it

Example of the careers sub-section of the website

H3 - User Control and Freedom

Code	Heuristic	Average Score	Agreed Score
H3	User Control and Freedom	2.75	3

Regarding the functionalities that the website offers, which are only a few (for example donation form, research function), there is always the possibility to undo, but not redo (even if it's less important considering the functionalities). A demonstration of redo lack is the donation form, after writing the mail field, going back to the amount of money choice will result in the mail field being cleared.

Regarding instead the navigation, going from a subsection of the website back to the main page is not intuitive and easy to carry out.

Donation Form

Example of the careers sub-section of the website

H7 - Flexibility and Efficiency of Use

Code	Heuristic	Average Score	Agreed Score
H7	Flexibility and Efficiency of Use	3	2

The only landmark that is mostly working is the "return to home" one, that is available and useful for both novice and advanced users. There are many examples of pages that

change the behaviour of that landmark, and have a different way of return back to the home page. At last, there are even some pages that don't allow to go back at all.

Regarding advanced users only, if the goal is to reach a certain sub-section of the website, or some specific parts, there is no fast and intuitive way of reaching them.

Figure 1.7: Supply sub-section: different landmark effect, "Visit Unicef Global" to go back to the home page

Another sub-section, not possible to return to the home page at all

M5 - Interaction consistency

Code	Heuristic	Average Score	Agreed Score
M5	Interaction consistency	3	3

For the most part the interaction for the pages is the same: there are always suggested topics to review, links to donate and share. There are however some exceptions, for example additional forms for donation (not present in every page of similar structure), links to jump across the page (present only in some pages), related topic tags (present only in some pages).

Figure 1.8: Additional donation form for Sudan's cause

Figure 1.9: Links to jump across the page
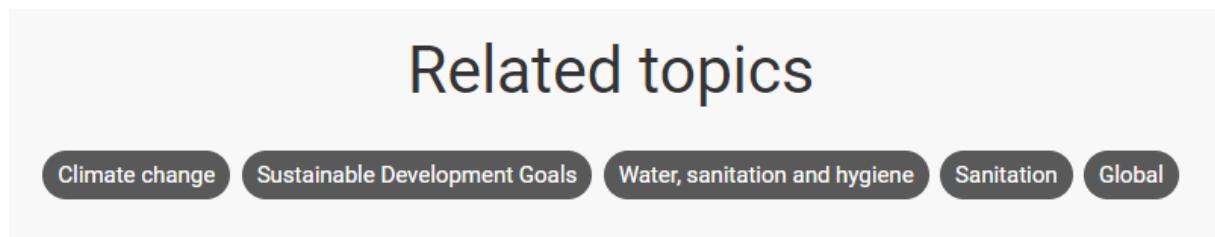
Figure 1.10: Related topic tags

M6 - Group Navigation 1

Code	Heuristic	Average Score	Agreed Score
M6	Group Navigation 1	3.25	3

Regarding the main drop down menu, the heuristic is satisfied, it is easy to navigate across it.

Regarding groups of pages, in those without breadcrumbs, it's difficult to navigate from one to another, and to do so the user needs to go through the same path that lead to that page. For example, from the "what we do page" to the "children with disabilities" page.

Another aspect to consider is the fact that going from an article or page, to the next, requires users to scroll up to the top of the page, and then choose manually the next page from the menus. It would have been easier to have a "prev/next" button, and also a "go to the top" button.

M7 - Group Navigation 2

Code	Heuristic	Average Score	Agreed Score
M7	Group Navigation 2	2.75	3

The heuristic is partially satisfied, because in some parts of the menu, the links are a bit too many, but most importantly sometimes they are useless because they lead to the same page.
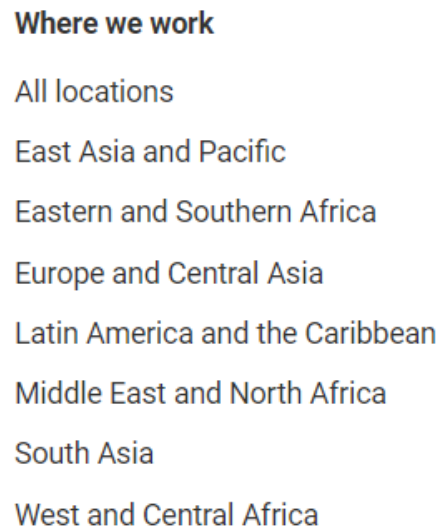
Where we work

All locations

East Asia and Pacific

Eastern and Southern Africa

Europe and Central Asia

Latin America and the Caribbean

Middle East and North Africa

South Asia

West and Central Africa

Figure 1.11: Example of links that lead to the same page

Stories and explainers

Stories of impact

Issue explainers

Photography
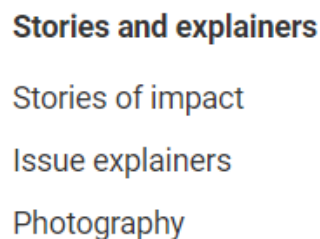
Figure 1.12: Another example of links that lead to the same page

M8 - Structural Navigation

Code	Heuristic	Average Score	Agreed Score
M8	Structural Navigation	2.5	2

In general the pages are big, especially because of images that take up much space. The consequence is that the user must scroll down a lot in order to visit each page. For this reason it could have been useful to have functions that help the user navigate with ease through the pages, for example with a "go to the top" or "go to the bottom" buttons.

There are pages that provide jump links at the top that redirects to some relevant points in the page, but these pages are only a few.

Example of page with big images

Figure 1.13: Example of page with jump links

Example of similar page with no jump links

M9 - Semantic Navigation

Code	Heuristic	Average Score	Agreed Score
M9	Semantic Navigation	3.25	3

This heuristic is satisfied only mainly in the "what we do" section, thanks to the breadcrumbs, that allow the users to navigate from one topic to a related one.

There also also a very few pages that include, at the bottom, some tags with related topics, but it is a unidirectional navigation, it's not possible from one topic to search for all the articles related to it, unless the user takes advantage of workarounds like the research function.
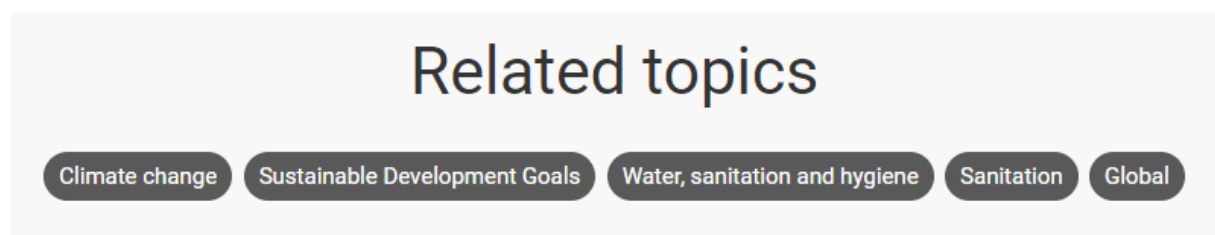
Figure 1.14: Related topics tags

M10 - Landmarks

Code	Heuristic	Average Score	Agreed Score
M10	Landmarks	3.75	4

Landmarks are almost always present, sometimes they are too little and misplaced, making it difficult to see and use them.

Figure 1.15: "Visit Unicef Global" landmark positioned at the top right part of the page, and small

1.5.3.   Presentation Heuristics

In this section the evaluated heuristics are all regarding the presentation category. Following, a visualization and a table summarizing each heuristic and the score attributed and agreed between evaluators.
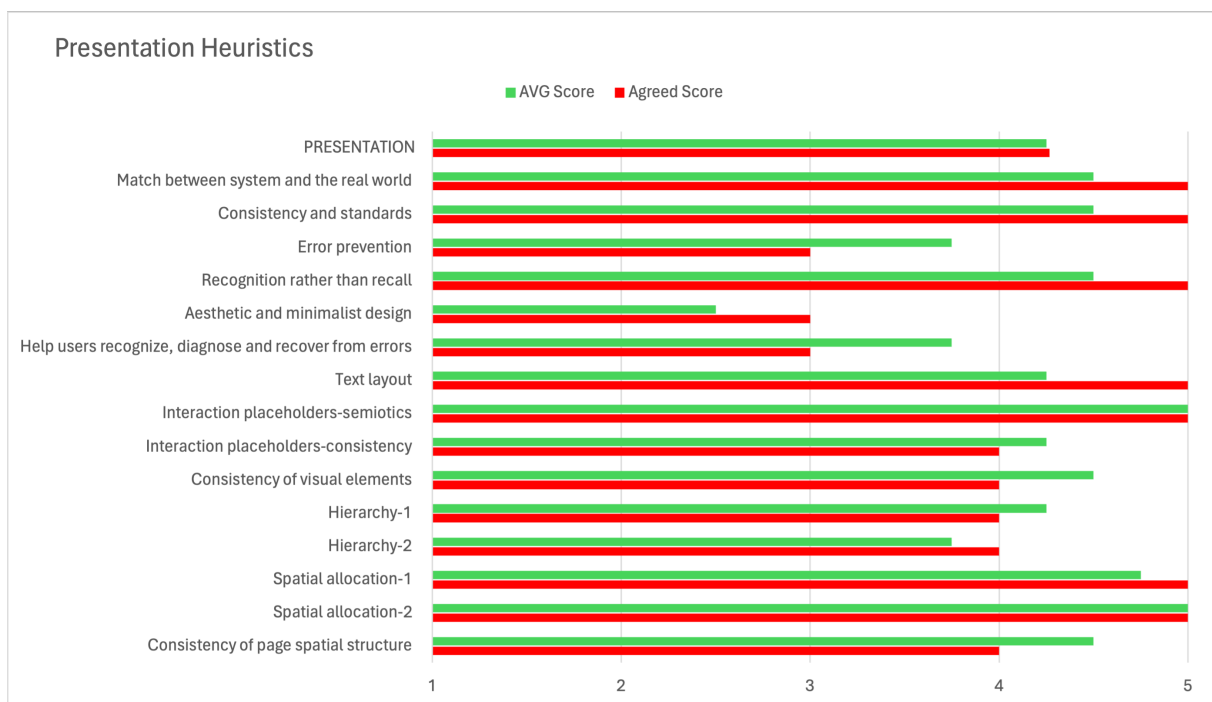

Figure 1.16: Average and agreed scores by inspectors

Presentation Heuristics

Code	Heuristic	Score
H2	Match between system and the real world	5
H4	Consistency and standards	5

H5	Error prevention	3
H6	Recognition rather than recall	5
H8	Aesthetic and minimalist design	3
H9	Help users recognize, diagnose and recover from errors	3
M11	Text layout	5
M12	Interaction placeholders-semiotics	5
M13	Interaction placeholders-consistency	4
M14	Consistency of visual elements	4
M15	Hierarchy 1	4
M16	Hierarchy 2	4
M17	Spatial allocation 1	5
M18	Spatial allocation 2	5
M19	Consistency of page spatial structure	4
Average	**Presentation**	**4.27**

Table 1.6: Scores for presentation heuristics

H2 - Match between system and the real world

Code	Heuristic	Average Score	Agreed Score
H2	Match between system and the real world	4.5	5

The website uses a language that is close to the users. Each navigation label is clear and icons, where used, are effective. The inspectors have appreciated that, in the donation page, appropriate icons and text show the donor what UNICEF can buy thanks to their act.
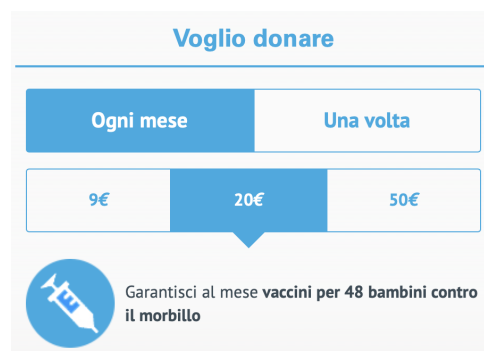
Figure 1.17: Donation Page (Partial, in Italian)

H4 - Consistency and standards

Code	Heuristic	Average Score	Agreed Score
H4	Consistency and standards	4.5	5

The website's user interface uses standards conventions that are consistent across the site (such as the "search" and "share" icons).

H5 - Error prevention

Code	Heuristic	Average Score	Agreed Score
H5	Error prevention	3.75	3

The website's form quality is fluctuating. For instance, in the first step of the donation wizard, if the user enters the string "1'000", the amount is validated (and rejected) only after pushing the "donate" button, and no information on the reason of the rejection is provided. Furthermore, when a user wants to join a mailing list, they can do it via two different forms, which are significantly different in terms of validation: most notably, in the "popup" one that appears as you reach the mailing list page, no validation is performed.

Figure 1.18: Donation Page Error (Partial, in Italian)

H6 - Recognition rather than recall

Code	Heuristic	Average Score	Agreed Score
H6	Recognition rather than recall	4.5	5

There are not many places where it is crucial that the website reminds something to the user, since it is content-intensive. Still, in the donation page, the user is recalled at every step of the amount of money they are donating and of the type of donation (one-time/monthly).
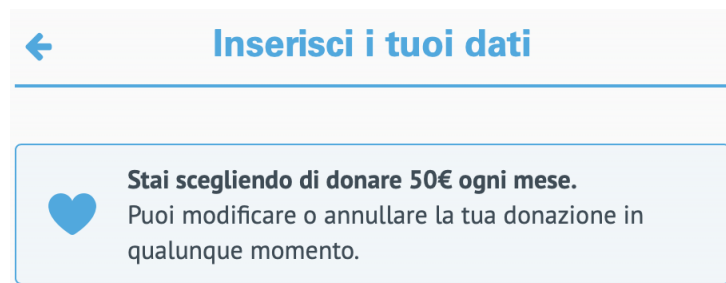

Figure 1.19: Donation Page Recall (Partial, in Italian)

H8 - Aesthetic and minimalist design

Code	Heuristic	Average Score	Agreed Score
H8	Aesthetic and minimalist design	2.5	3

Pages tend to have too much information, and the presence of tens of links in each of them make the navigation very prone to distraction. The aesthetic is overall good: the only things worth noticing are the excessive dimension of some images and of the report containers.

Page with a big initial image and with big containers for reports at the end

H9 - Help users recognize, diagnose and recover from errors

Code	Heuristic	Average Score	Agreed Score
H9	Help users recognize, diagnose and recover from errors	3.75	3

Apart from the already mentioned fluctuating form quality, it is possible to reach from the website pages that are no longer present. In such scenarios, the "404 error" text is shown, even if with some CSS. Even though the following text expands a bit on the error, the inspectors think that it is not a good idea to show such a text in a website that is primarily used by people with little to no expertise in Computer Science.

404 error: page not found

Sorry, we can't find the page you are looking for. As we continue to keep our site up to date, this page may have been moved or updated.

BACK TO OUR HOME PAGE

Figure 1.20: 404 Error

M11 - Text lay out

Code	Heuristic	Average Score	Agreed Score
M11	Text lay out	4.25	5

The text is always readable and has an appropriate size with respect to its importance, hence there is no violation to point out.

M12 - Interaction placeholders-semiotics

Code	Heuristic	Average Score	Agreed Score
M12	Interaction placeholders-semiotics	5	5

The interaction elements are easy to understand, hence there is no violation to point out.

M13 - Interaction placeholders-consistency

Code	Heuristic	Average Score	Agreed Score
M13	Interaction placeholders-consistency	4.25	4

This heuristic is generally well-implemented, however there are a few imperfections to point out. First of all, the button that takes the user to the donation page comes some-

times with different text (i.e. "donate", "donate now", or "become a donor"). Furthermore, the "donate" button is sometimes of different colors, which can generate confusion, and the color usually used for the donation button has been used elsewhere (*here*) with a different meaning ("take action").
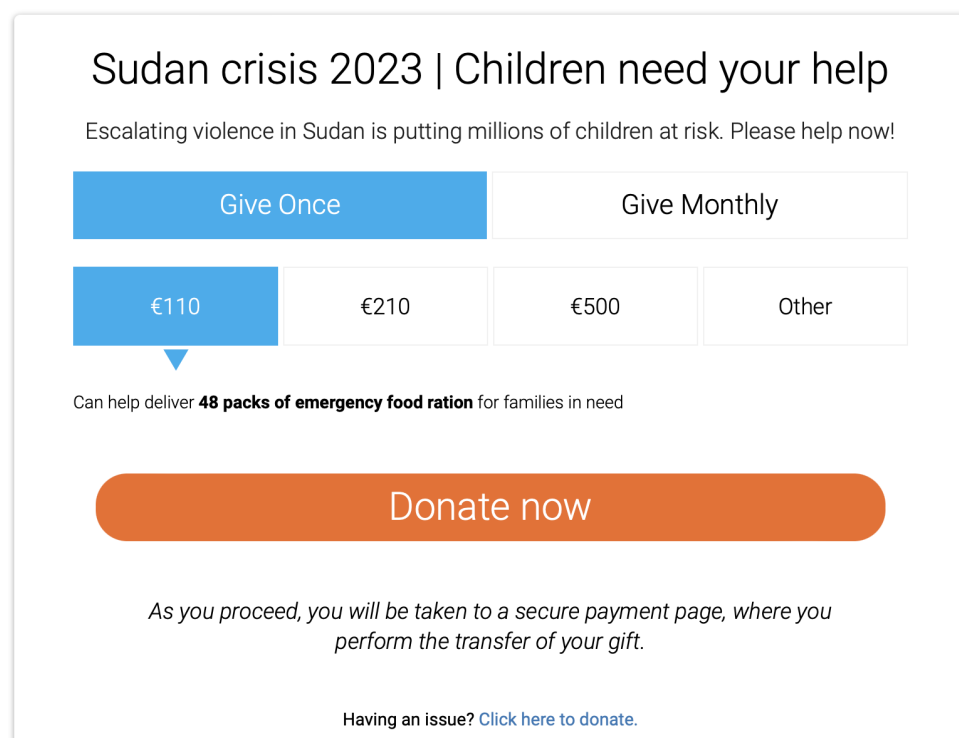

Figure 1.21: Most Common Donation Button

Figure 1.22: Sudan Donation Button

M14 - Consistency of visual elements

Code	Heuristic	Average Score	Agreed Score
M14	Consistency of visual elements	4.5	4

Generally, the website's visual elements are consistent, except for the example discussed in the previous heuristic's comment.

M15 - Hierarchy 1

Code	Heuristic	Average Score	Agreed Score
M15	Hierarchy 1	4.25	4

This heuristic is generally well-implemented. However, there are a few things worth pointing out: first of all, in the "What we do" sections, the photos are so big that whatever comes after could be regarded as secondary content, while it's the actual content. We understand that the big pictures could have been the result of design choices aimed at building empathy, but they often feel like an obstacle to overcome in order to use the website. Moreover, at the following *link*, one could expect to read about the people behind UNICEF first, but the actual first thing that one can read is about what they do (but there's already a huge section of the website dedicated to such matter).

M16 - Hierarchy 2

Code	Heuristic	Average Score	Agreed Score
M16	Hierarchy 2	3.75	4

This heuristic is generally well-implemented, with the notable exception of the big images (which we have already discussed in the previous paragraph).

M17 - Spatial allocation 1

Code	Heuristic	Average Score	Agreed Score
M17	Spatial allocation 1	4.75	5

There are no violations to report with respect to this heuristic.

M18 - Spatial allocation 2

Code	Heuristic	Average Score	Agreed Score
M18	Spatial allocation 2	5	5

There are no violations to report with respect to this heuristic.

M19 - Consistency of page spatial structure

Code	Heuristic	Average Score	Agreed Score
M19	Consistency of page spatial structure	4.5	4

This heuristic is generally well-implemented, but there's an imperfection to highlight: when browsing UNICEF's annual reports, the 2022's report is significantly different from other years' reports.

2022 annual report

2021 annual report

1.6. Inspection Final Comment

Among the three main aspects that were evaluated, the presentation aspect is the one that received less criticism: indeed, the website is visually captivating and, generally, visually consistent. However, concerning the other two aspects (content and navigation), the Inspection technique has highlighted several problems which should be prioritized in a future redesign phase. More specifically, the huge amount of content present on almost all of the pages, along with the huge amount of links that come with it, may lead to distraction and make the site considerably less usable. Plus, the bad implementation of bread crumbs makes the navigation more difficult, impacting the effectiveness and efficiency of users. Lastly, as explained in the previous sections, landmarks are sometimes inconsistent and the different donation functions are not well-integrated, worsening the frustration of the end user.

2 | User Testing

User Testing is another usability evaluation technique. It is based on recruiting a relevant sample of candidates - that represent the actual possible users of the website - and observing them interact with it. Candidates are usually assigned tasks, which are completed under the supervision of an expert observer ("moderator"), who will gather data during the procedure. Said data is meant to highlight possible issues and flaws not found during the inspection phase, as direct interaction with the end user is involved.

2.1. User Testing Design

2.1.1. Candidates Definition and Recruitment

For the scope of the study, a sample of a total of 20 candidates has been recruited. Since a maximum of 12 people is needed in order to make the collected data converge to the "real" value, we decided to recruit our 20 candidates by considering two groups: 10 candidates should have technical skills with respect to Computer Science, whereas the other 10 candidates should not. Concerning both the candidate sets, the gender ratio is 1:1. In this way, we want to evaluate if the expertise in Computer Science can influence the fruition of a website that should instead be easy to use for all categories of users.

All recruited people are young people of ages 18-25. This demographic was chosen as young people are more likely to empathise with UNICEF's mission, and even if they may not represent the most valuable economic agents, they surely are more likely to make donations through a website than older people, who are more likely to use dedicated telephone numbers following a TV advertisement. All candidates were direct connections of the experts and were not promised any compensation.

2.1.2. Task Definition

After that each expert came up with a relevant set of tasks, they gathered and decided a reasonable subset of them that best suited the study that they wanted to perform. Hence,

they were chosen based on generally relevant aspects of the website, with special attention to areas that had already been deemed as problematic during the inspection phase. Plus, for each task, we defined a reasonable, expected time limit by which users should be able to perform the task. This choice was made because it is not best practice to keep the testers actively engaged for more than one hour (considering also the time needed to explain the user testing activity and the submission of the forms). The following list explains each task, along with a brief motivation for the choice and the expected time limit.

1. **You are a university student who wants to donate 15€ monthly to UNICEF, as you are interested in their cause. Proceed to do so. (1 min)** This task is perhaps the most important one, as the experts believe that the main reason why the UNICEF's website has been built is to collect as many donations as possible. Hence, we want to be sure that such a task can be easily performed by everyone.

2. **You are particularly interested about the emergency in Yemen, so you decided to donate 100€ specifically for that cause. Proceed to do so. (5 min)** This task is similar to the previous one, but the experts believe that it could be significantly harder due to how the website handles donations that are specific for a certain emergency.

3. **You are wondering what UNICEF stands for, thus you decided to look it up under the Frequently Asked Questions (FAQs) section. Find the answer to your question. (3 min)** During the inspection, an expert came up with the possibility that the FAQs section may not be trivial to find for non-expert users. Plus, there are various different ways to find it. With this task, we want to be sure that every user can eventually be able to read the FAQs, since they comprise valuable information about the most important things related to UNICEF.

4. **You would like to participate actively to the UNICEF's cause. Find the page in which you can browse through the available jobs. (4 min)** Apart from donations, the experts believe that this task is among the most crucial ones. An organization like UNICEF needs to attract as many people as they can in order to perform their helping action.

5. **You heard about malnourished pregnant or breast-feeding women's issue on the news. Find where the website mentions ten key actions to improve adolescent girls' and women's nutrition. (7 min)** With this task, we want to test how such a content-intensive website can be easily navigated to find valuable information. Again, there are many ways to reach the required content, and the

experts want to see which paths the users will partake.

6. **You are doing a school research about how water issues impact the life of children. Find all the contents about water published from 01 February 2024 to 20 March 2024. (5 min)** With this task, we want to assess how easy to use the search function of the website is. Since the website is content-intensive, we believe that it should be a powerful function that every user can easily notice and use.

7. **During your last class, your professor told you about the 17 sustainable development goals. In particular, you want to gain more information about the "zero hunger" goal. Search for a downloadable document which talks about how to implement it. (6 min)** This task aims at understanding whether users can easily download more specific information only if they really need it.

2.1.3. Evaluation Variables Definition

In order to gather data from the user testing process, the moderators agreed on some variables for the tasks at hand. Such variables can be divided into quantitative variables and qualitative variables.

Quantitative Variables

These variables can be analyzed either during or after the actual testing phase, and yield data that can be represented through numbers and aggregation/visualization. The chosen variables for the study are:

- **Effectiveness.** This variable indicates how effectively and with how much ease a user completed a task. The evaluation of this variable is done by the observer right after the execution of each task, and is immediately recorded. The metrics used are explained by the following table.

Effectiveness metrics

Score	Meaning
0	The user gave up on completing the task
1	The user was not able to complete the task within a reasonable, established time limit

| 2 | The user was eventually able to complete the task but needed slight assistance |
| 3 | The user completed the task on their own and within the established time limit |

Table 2.1: Effectiveness metrics definition

- **Efficiency (time on task).** During the execution of each task, time was tracked and recorded. Tasks that take the user too long to perform represent and highlight possible flaws.

- **Wandering time.** During the execution of each task, wandering time was also tracked and recorded. High wandering times may either mean that the website suffers from information overload, or that there are so many links that the user needs time to understand where to go next.

- **Search bar utilization.** This variable was evaluated during the testing phase by the moderator, who counted the number of times the user tried using the search bar in order to complete the task.

- **Number of errors and wrong paths.** This variable was evaluated during the testing phase by the moderator, who counted the number of times the user strayed from a path that would bring to the destination.

- **Perceived task difficulty.** This variable was evaluated after the execution of each task: users were asked to fill out a form in which they could express how difficult thee task was perceived. The metrics used are explained by the following table.

Perceived task difficulty metrics

Score	Meaning
1	Extremely easy
2	Somewhat easy
3	Neutral
4	Somewhat difficult
5	Extremely difficult

Table 2.2: Perceived task difficulty metrics definition

Qualitative variables

Contrary to the previous category, qualitative variables represent aspects of the user experience that cannot be quantified. During the whole process of testing, users were encouraged to make comments about their impressions and feelings towards the site. Forms offered after the testing also contained spaces to leave comments on (both concerning the whole website and for each individual task). Finally, each tester also compiled the System Usability Scale (SUS) form, which is a well-known questionnaire in the literature. Relevant comments were recorded by observers, and were then used to evaluate the following variables:

- **Disorientation.** This variable aims at estimating the level of ease with which users navigate through the site, and how well the paths intended to reach specific information are followed. Data regarding this variable were gathered by considering user comments and expressions, as well as the forms that were offered.

- **Satisfaction.** Measuring the level of satisfaction that end users have with respect to the product under evaluation is of extreme importance. This variable reflects whether the site met the users' expectations and if it caused frustration. Once again, users' comments, both spoken and written in the forms were taken into account.

2.2. User Testing Execution

2.2.1. Before the Test

Before its execution, the test was clearly explained to each user. In particular, it was pointed out that the evaluation concerns the website and not their performance, and the users were reassured that they could leave at any point during the test. Plus, users were invited to think aloud for the whole duration of the test, in order to allows the moderators to better track their emotions and their satisfaction while using the application.

If the test was conducted in presence, the moderators gave each user a laptop with a browser (either Safari or Google Chrome) and two different tabs opened: the tasks form and UNICEF website's home page. If the test was conducted remotely, the user had to prepare their own laptop with a suitable browser, and had to leave both the webcam on and the microphone on. Then, they opened the given task form and started to share their screen for the entire duration of the test.

2.2.2.   During the Test

The user reads aloud the first task and, when they are ready, they go to the homepage of the website and begin to complete the task. At that point, the evaluators start the chronometer, and examine the user's actions and reactions. At the end of each task, the chronometer is stopped, and the user is asked to fill out the form for that specific task, in order to collect the perceived difficulty of the task and a brief comment. This process was repeated for each task. Notice that wandering time was collected with the same chronometer by using the "lap" functionality.

2.2.3.   After the Test

After the users had finished solving the tasks, they were asked to answer the System Usability Scale form to collect their opinion on the website and they were also given the possibility to leave a general comment on their overall experience with the product.

2.3.   User Testing Results and Discussion

In this section, the results obtained from User Testing will be reported, focusing on allowing a clear understanding of the collected data through a precise and visual way. From now on, in the following tables, we will distinguish between skilled users and not skilled users by coloring the user identifier with violet for the first group (skilled), and with magenta for the second one (not skilled).

2.3.1.   Effectiveness

Effectiveness is defined as the degree of completion with which each user carried out the tasks. As previously discussed in the Evaluation Variables section, the degree of completion was caught by a score from 0 to 3. The overall effectiveness score of the system is 2.486/3, approximately. In the following table and in the following visualizations, the complete effectiveness results are reported.

User Effectiveness

User	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	1	3	3	1	0	2	3
2	3	3	3	3	2	1	3
3	3	3	3	3	3	3	1

4	3	1	3	3	2	3	3
5	3	3	3	3	3	0	3
6	3	3	3	3	3	2	3
7	3	3	3	3	3	2	3
8	3	3	3	3	2	2	3
9	3	3	3	3	2	2	2
10	3	3	3	3	2	2	2
11	1	0	3	3	0	2	3
12	3	1	3	3	1	1	3
13	3	3	3	3	0	1	3
14	3	1	3	3	0	3	3
15	3	3	3	3	1	3	3
16	3	3	3	3	3	1	3
17	3	3	3	3	1	1	3
18	3	2	3	3	2	3	3
19	3	3	3	3	1	2	3
20	3	3	3	3	2	1	3
tot. avg	2.8	2.5	3	2.9	1.65	1.85	2.8
skilled avg	2.8	2.8	3	2.8	2.2	1.9	2.6
not skilled avg	2.8	2.2	3	3	1.1	1.8	3

Table 2.3: User effectiveness
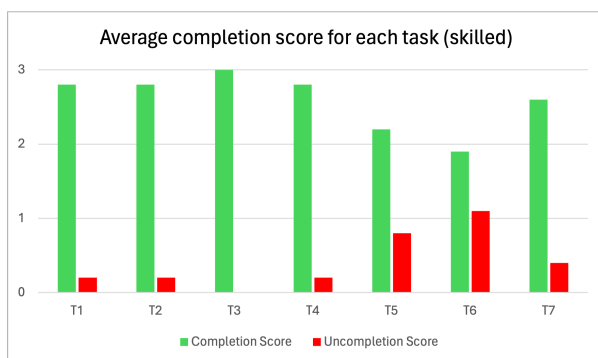
Figure 2.1: Effectiveness for skilled users
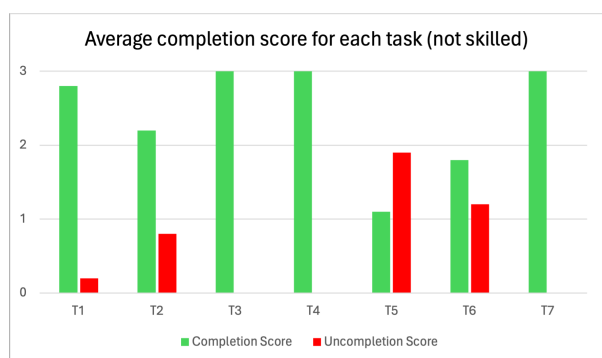
Figure 2.2: Effectiveness for non skilled users

The task in which testers showed a minor completion rate is Task 5 (ten key actions

to improve adolescent girls' and women's nutrition). This task was designed in order to understand the ease of navigation into such a content-intensive web site, and the obtained result suggests that it was not easy for the testers to find a specific information. Some testers tried to reach the goal of the task by using the search feature provided, but all of them had an hard time in finding valuable information with that tool, showing frustration and annoyance. This highlights the poor quality of implementation of the search feature, which should be a key aspect for deeply content-related systems. Other users tried to navigate in a more standard way, using the menu and going deeper into the pages of the site by following hyperlinks. This strategy showed generally more effectiveness, even though some users found it difficult to manage to complete the task. To opt for the latter strategy was a more common choice for skilled testers, who obtained a better completion score, as it is possible to see from the two charts reported above.

Task 6 also showed a small degree of completion by the testers, as the vast majority of them expected to be able to filter for date from the section of the site called "publications by topic". When they noticed it was not possible, they often found themselves wandering across the section, not knowing exactly what to do to carry on the task and sometimes even resigning. This has resulted in an average score of 1.85/3 for Task 6, once again revealing the lack of ease in navigation and of missing consistency in the way the system provides features for searching specific pieces of content.

Even though Tasks 5 and 6 provided a smaller effectiveness score, it is still important to notice that, for the other tasks, most of the users were able to complete them.

2.3.2. Efficiency

Efficiency is defined as the amount of time required for each user to carry out the tasks. This time is defined as the time that elapses between the opening of the homepage and the instant in which the user completes the activity. If the user was not able to finish within a specific time limit, we considered the threshold we set uniquely for each task as the efficiency time we used to compute the averages.

In the following table, the measured times are reported in seconds, along with the average time spent for each task.

Time on Task

User	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	60	238	177	240	420	263	63

2	25	55	111	29	304	300	198
3	60	63	24	62	118	47	360
4	22	300	12	16	276	40	313
5	18	28	44	64	125	276	112
6	25	72	35	30	73	270	90
7	21	29	33	37	138	233	81
8	26	127	30	14	179	159	21
9	12	48	21	15	307	241	297
10	20	62	51	207	114	178	144
11	60	170	146	111	354	300	236
12	33	300	50	22	420	300	130
13	28	264	60	34	420	300	22
14	12	300	12	60	420	236	140
15	60	253	46	17	420	45	19
16	25	154	50	24	124	300	195
17	13	85	65	27	420	300	44
18	26	228	25	90	398	145	146
19	25	280	58	35	420	300	172
20	30	100	52	25	314	300	46
tot. avg	30.05	157.8	55.1	57.5	288.2	266.65	141.45
skilled avg	28.9	102.2	54.4	71.4	205.4	200.7	167.9
not skilled avg	31.2	213.4	57.4	44.5	371	252.6	115

Table 2.4: Time on task (in seconds)
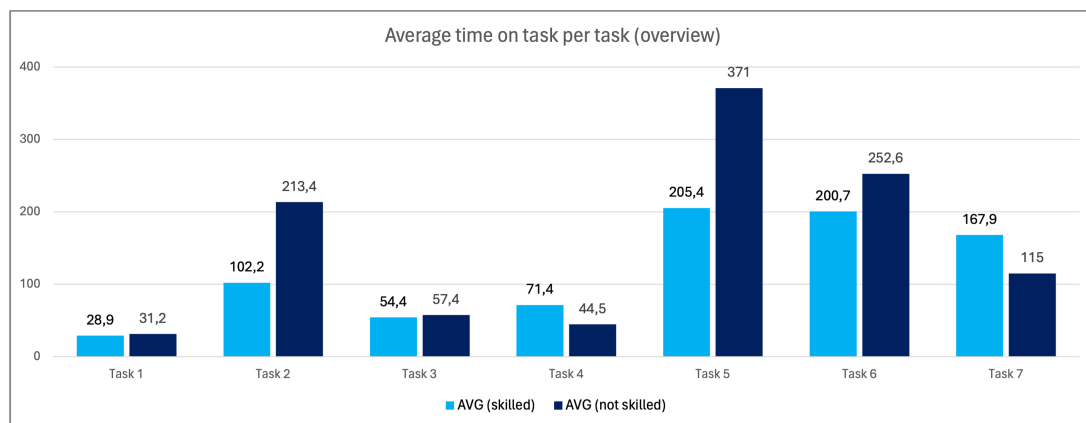
Figure 2.3: Time spent by users on each task

The average time spent for each task is pretty consistent with the efficiency ratings: the tasks 5 and 6 are clearly the two most time consuming.

On the other hand, it is important to notice that Task 2 (Yemen donation) has a low efficiency too: from the results obtained from User Testing, we noticed that many testers expected to find in the "Donation" page the possibility of directing their charity to all the specific causes UNICEF takes care about, including Yemen emergency, but this was not the case. It was necessary for them to reach the Yemen page in order to donate there, and this resulted in a substantial loss of time in the completion of the task and therefore a significant loss in efficiency. In figure 2.4 and 2.5 it catches an eye how T2 completion was more efficient for skilled users, it indicates that their technical experience helped them to find a workaround for the issue described above.

In Task 4 (search jobs in UNICEF), there are some users that took a significantly higher amount of time to carry out the task with respect to the average. That is because they found hard to notice the button "search jobs" in the "Take action" page, because of the chromatic similarity of the button with its background, so they scrolled down the page, even questioning if they took the right path, as they did not understand which would be the correct next step. This allowed us to also notice that in the "Take Action" section, it is hard to distinguish among the variety of possibility a person has to collaborate with UNICEF (e.g. jobs, volunteering, internship, partnership etc.).

Even taking into account the above comments and observation, we can anyway say that the overall task completion time was inline with what we expected. Still, it is advisable to be careful to not sacrifice the usability aspects due to aesthetics reasons, and also to be more consistent in terms of feature update: the testers were really annoyed by the fact that some emergencies were more highlighted than others.
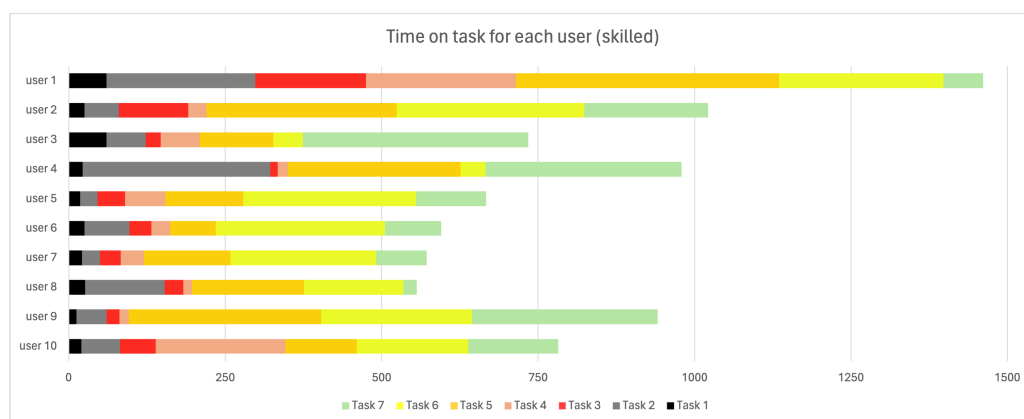
Figure 2.4: Time spent for each task by skilled users
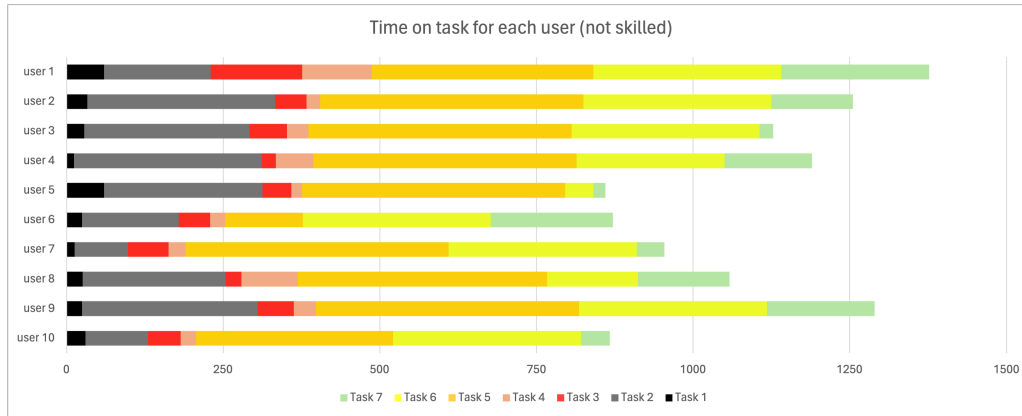
Figure 2.5: Time spent for each task by non skilled users

2.3.3. Wandering time

Wandering time, as the name suggests, represents the amount of time users spent wandering around the site, without a clear idea of where to go in order to complete the task, gathering information about the web site structure and content. The wandering times are expressed, in seconds, in the following table.

Wandering Time

User	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	40	128	97	187	390	186	22
2	8	15	45	14	228	242	118
3	33	37	1	10	34	57	280
4	2	260	1	2	192	5	245
5	1	3	11	21	56	238	59
6	0	39	18	12	20	210	10
7	0	13	15	15	118	200	40
8	0	112	7	0	125	144	0
9	0	20	0	0	252	221	170
10	0	40	47	140	80	151	98
11	60	170	60	28	196	240	187
12	5	251	45	0	288	250	27
13	5	210	60	0	322	226	0
14	0	210	0	30	240	215	102
15	50	191	5	0	322	0	0

16	0	30	10	0	20	120	30
17	0	25	35	10	240	180	0
18	3	70	5	30	240	40	45
19	0	180	40	20	360	180	90
20	10	40	20	5	180	210	10
avg	10.85	102.2	26.1	26.2	195.15	165.75	76.65

Table 2.5: Wandering time (in seconds)

The outcome of this measurement is really similar to what we have already discussed in the Efficiency section above. Tasks 2, 5 and 6 evidenced the most the presence of information overload: when the users where required to find specific pieces of information, they often lost themselves within the pages of the web site, trying to understand which was the correct path that would lead them to the task completion.

2.3.4. Search bar utilization

Search bar utilization is a metrics registered by the moderators and indicates, for each task, the number of times each user decided to use the search feature provided by the system. The results are reported in the following table and in the following visualization.

Search bar usage

User	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	0	0	0	0	0	1	0
2	0	0	0	0	0	1	1
3	0	0	0	0	1	6	6
4	0	1	0	0	3	1	3
5	0	0	1	0	1	0	1
6	0	0	0	0	0	1	0
7	0	0	0	0	0	1	0
8	0	0	0	0	0	1	0
9	0	0	0	0	0	1	1
10	0	0	0	0	0	1	0
11	0	0	1	0	0	1	4
12	0	1	0	0	1	1	1
13	0	2	0	0	3	1	0

14	0	3	0	0	4	1	1
15	0	1	0	0	7	1	0
16	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0
18	0	1	0	0	3	1	2
19	0	0	0	0	2	1	1
20	0	0	0	0	0	0	0
tot	0	9	2	0	25	21	21
skilled	0	1	1	0	5	14	12
not skilled	0	8	1	0	20	7	9

Table 2.6: Search bar usage
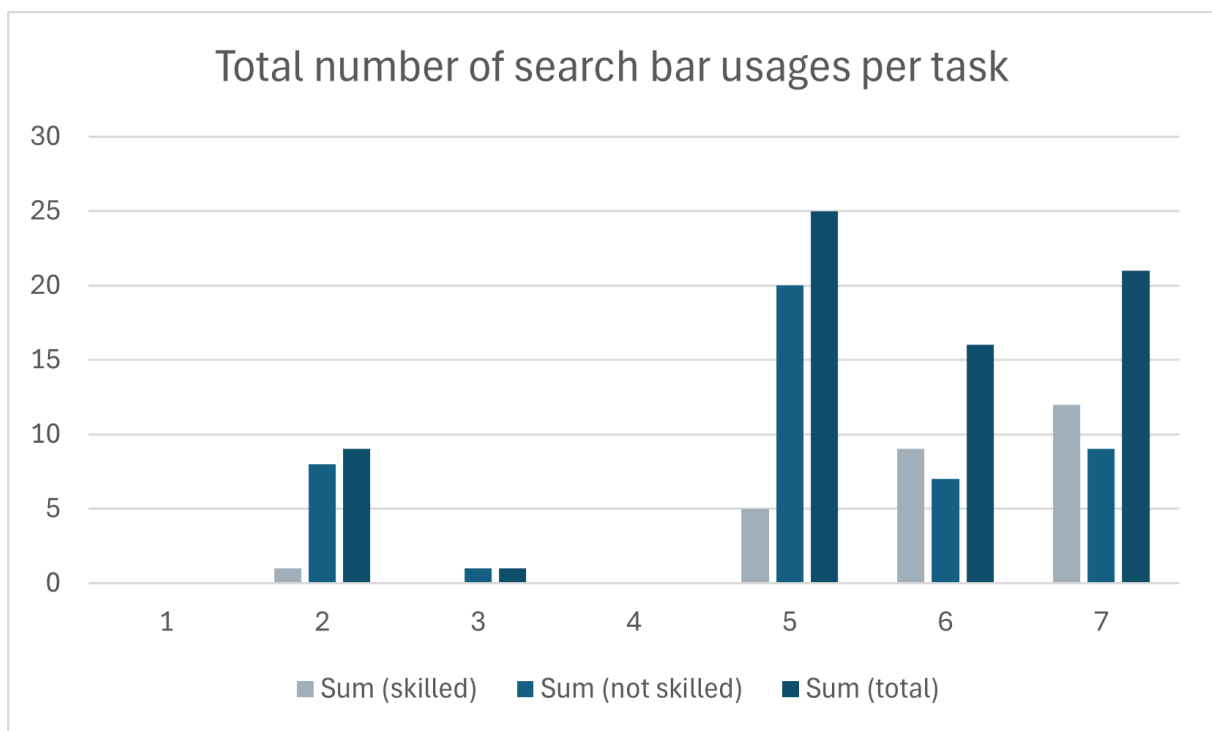
Figure 2.6: Total number of search bar usages filtered by user category

We can notice how in general not skilled users decided to opt for the usage of the tool more than skilled ones.

There are some exceptions in Task 6 (find articles about water by filtering for date) and Task 7 (find "zero hunger" downloadable). For task 6, that is mostly because it was necessary to use it in order to complete the task, and most of the skilled testers managed

to have the correct intuition. Regarding task 7, the difference causes an interesting effect, letting more not skilled users being able to complete the task, while lowering the completion score for skilled users: this is a clear trend reversal with reference to the other 6 tasks, where technical testers did generally better.
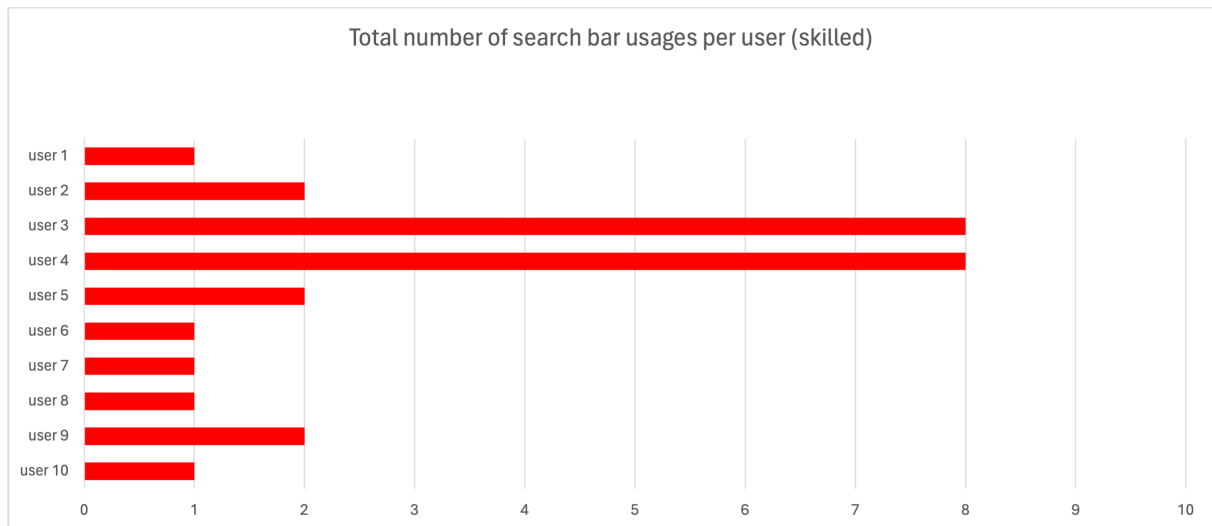
Figure 2.7: Total number of search bar usages per skilled user
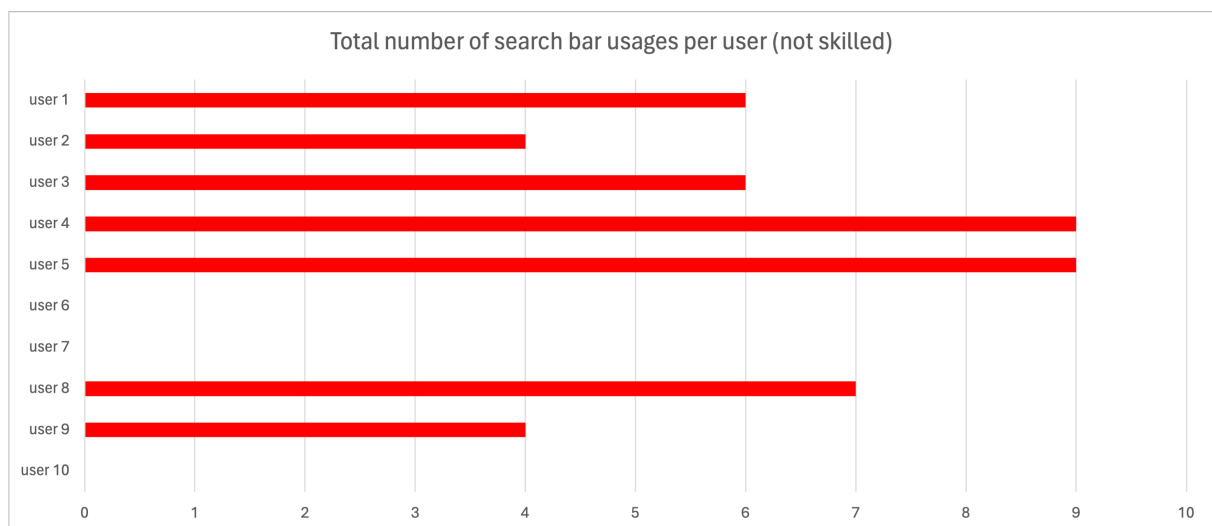
Figure 2.8: Total number of search bar usages per not skilled user

2.3.5. Number of errors and wrong paths

The "Number of errors" variable measures the number of wrong actions or paths taken by the user while browsing the website. Here we can observe the values recorded during the execution of the test.

Number of errors and wrong paths

User	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	1	1	0	1	4	2	0
2	0	0	0	0	3	4	2
3	0	0	0	0	0	0	8
4	0	3	0	0	2	0	3
5	0	0	0	0	2	4	1
6	0	1	0	0	0	3	1
7	0	0	0	0	3	2	1
8	0	1	0	0	4	1	0
9	0	1	0	0	5	4	4
10	0	1	0	3	0	1	2
11	1	2	2	1	4	2	0
12	0	2	0	0	4	1	1
13	0	2	0	0	9	2	0
14	0	8	0	1	10	2	4
15	0	1	0	0	9	0	0
16	0	3	0	0	1	3	2
17	0	2	0	0	4	3	0
18	0	4	0	1	4	1	2
19	0	2	0	0	3	2	1
20	0	1	0	0	3	4	0
tot	2	35	2	7	74	41	32
skilled	1	8	0	4	23	21	22
not skilled	1	27	2	3	51	20	10

Table 2.7: Number of errors for each task
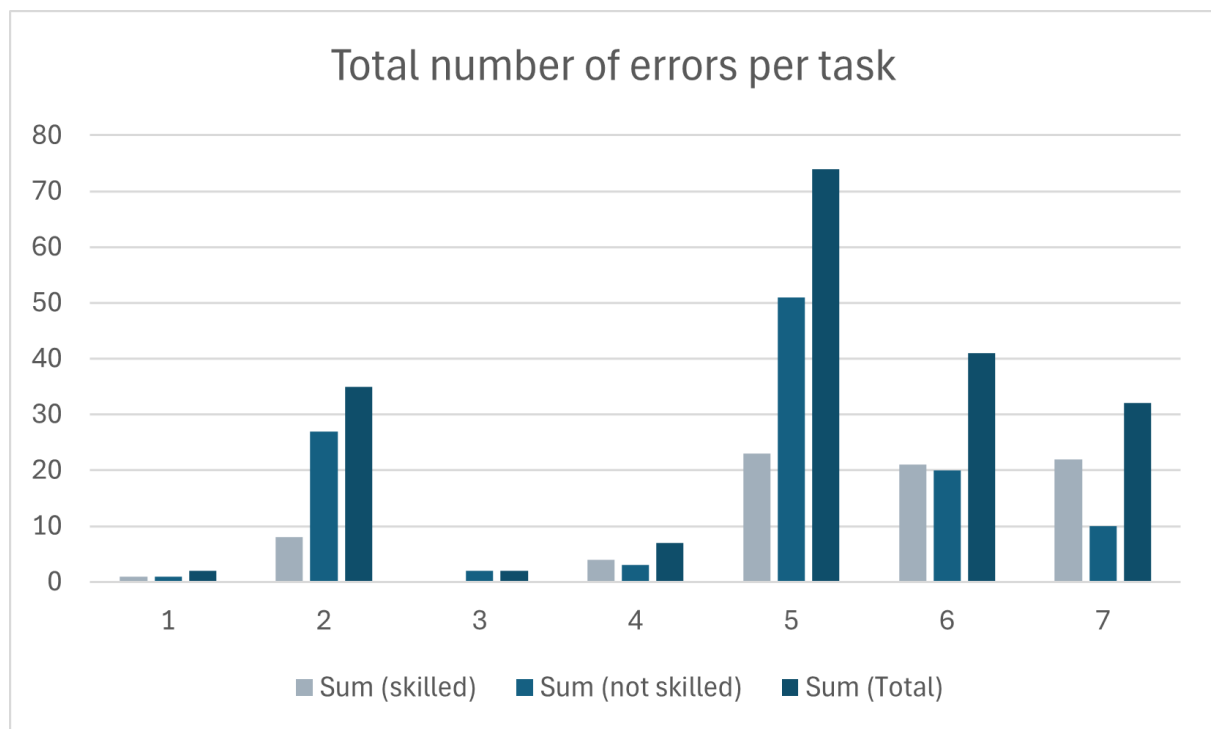
Figure 2.9: Number of errors filtered by user category

Again, the most problematic tasks are 5 and 6. This clearly suggests the difficulty users found in reaching the pages required to complete these two tasks.

Also, task 2 (Yemen donation) and task 7 (Find "zero hunger" downloadable document) results present a significant amount of errors from the testers: this evidences that they have not always found the navigation through the system intuitive.

Since we already discussed the obstacles for task 2, it is now important to focus on task 7: some users expected to find the "17 sustainable development goals" under the section "What we do" of the main menu, while it is under "About UNICEF". Moreover, it was pretty common that, even after managing to find the "17 sustainable development goals", a tester clicked on wrong hyperlinks, and had to spend some time finding a way back to where they were, causing an increased number of errors.

On the chart in Figure 2.9 note how on T6 and T7 skilled testers have more errors than not skilled ones. As it was already said, it is mostly because of the search feature usage.

2.3.6. Perceived task difficulty

The perceived task difficulty is a variable that derives from an evaluation by each user after the execution of each task. It is a measurement on a scale from 1 (extremely easy)

to 5 (extremely difficult) used to gain information on how the task was perceived.

As it was already described in Section 2.1.1, during user testing we divided the 20 testers in two groups: skilled users and not skilled users. In order to gain more insights, we will compare how these two groups perceived the tasks, and comment the results.

In the following chart, we compare the evaluations of the two groups.
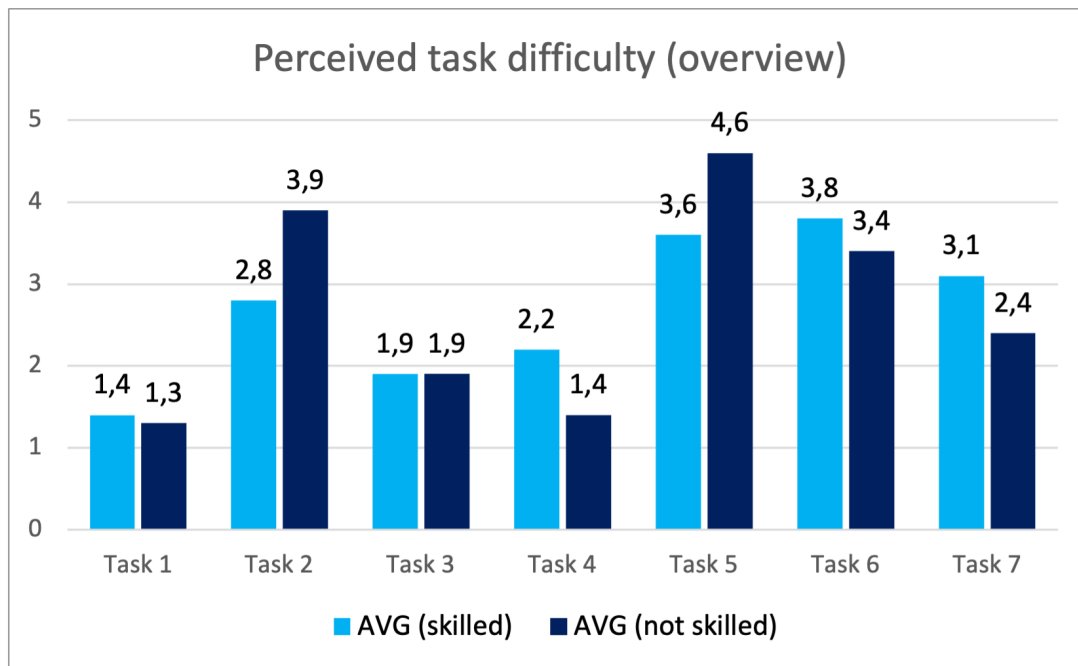
Figure 2.10: Overall perceived task difficulty filtered by user category

The first thing that catches the eye is the difference in evaluation of task 2 and task 5. This is due to the fact that skilled testers have proven to be better at managing path errors and not to be discouraged from it, leading them to an higher completion rate and therefore to a lower perceived task difficulty.

On the other hand, task 7 was slightly better perceived by not skilled users. This is because most of skilled users decided to use the search feature in order to find the "zero hunger" downloadable and this brought a certain degree of perplexity, resulting in a higher perceived difficulty.

2.3.7. Disorientation

Disorientation estimates the level of ease with which users navigate through the web site and how well path intended to reach a certain information are actually correct.

The form that was asked to compile gives us a general idea for this qualitative metric. In the following charts, the average scores given by not skilled and skilled users are reported. Green, yellow and red indicate respectively good opinions, neutral opinions and bad opinions.
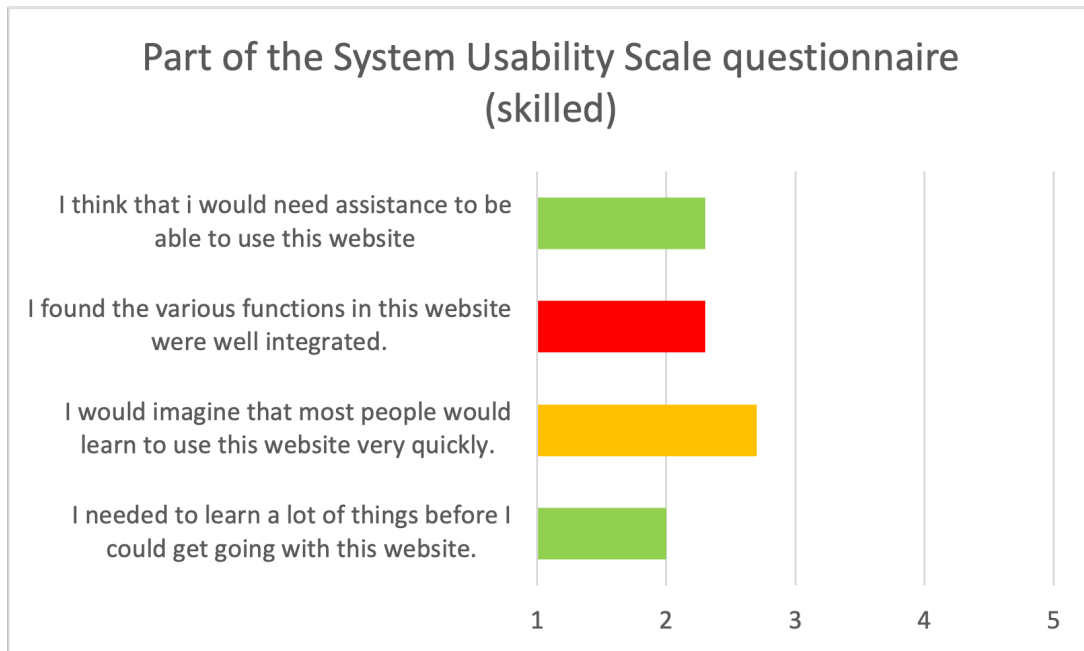
Figure 2.11: Skilled users' opinion on disorientation
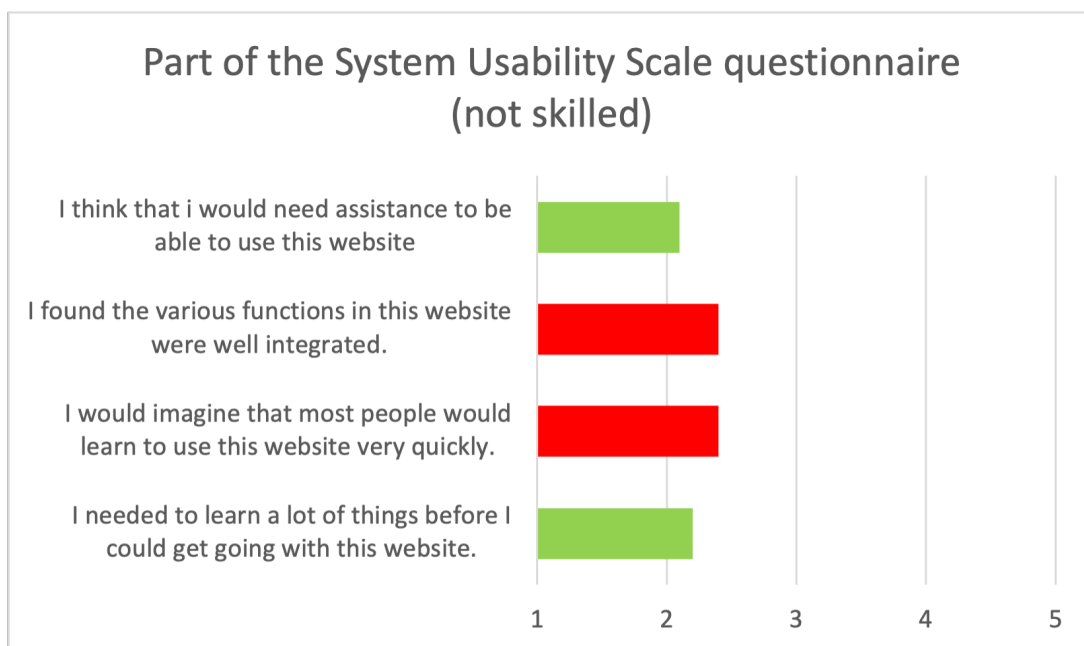
Figure 2.12: Not skilled users' opinion on disorientation

As it is possible to notice, the difference in ability of the users did not impact significantly the overall experience in the system, which was perceived by the most of the testers as not too much annoying and disorientating, even though they recognized the need for some practice before being able to effectively navigate through the site, and also that some functions and features were not really useful.

2.3.8. Satisfaction

Satisfaction is an estimate of the level of satisfaction that users have with respect to the product they were to use. In the following charts, the average scores given by not skilled and skilled users are reported.
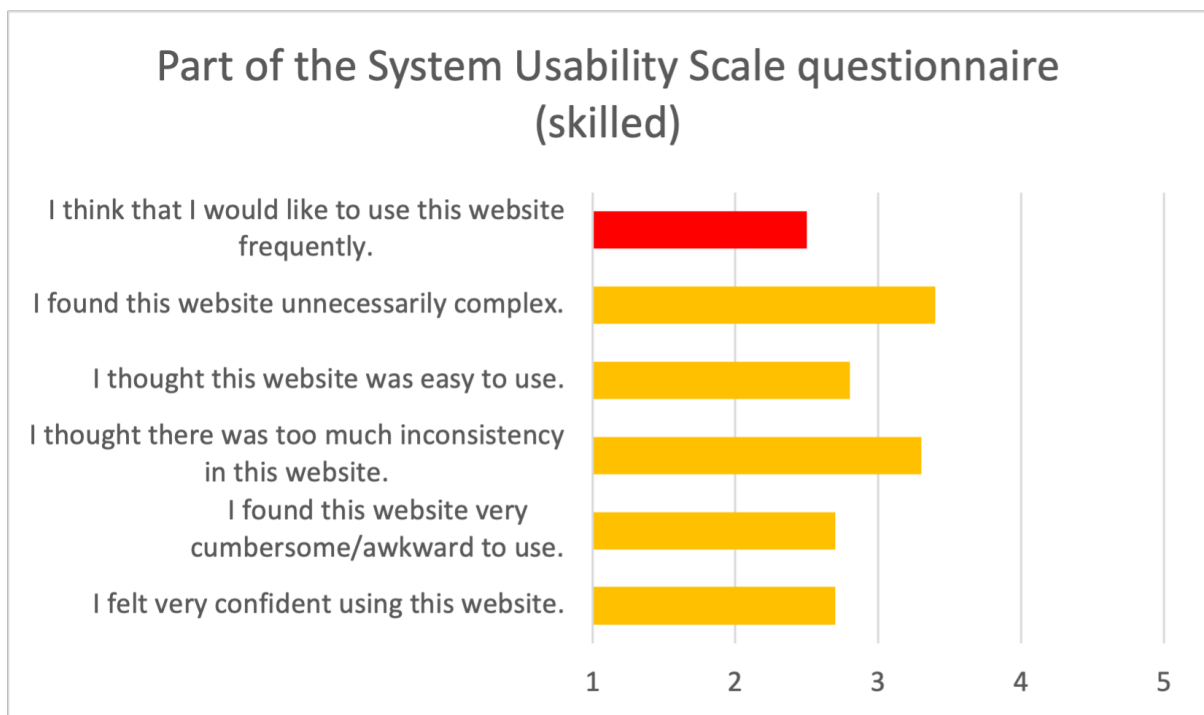
Figure 2.13: Skilled users' opinion on satisfaction
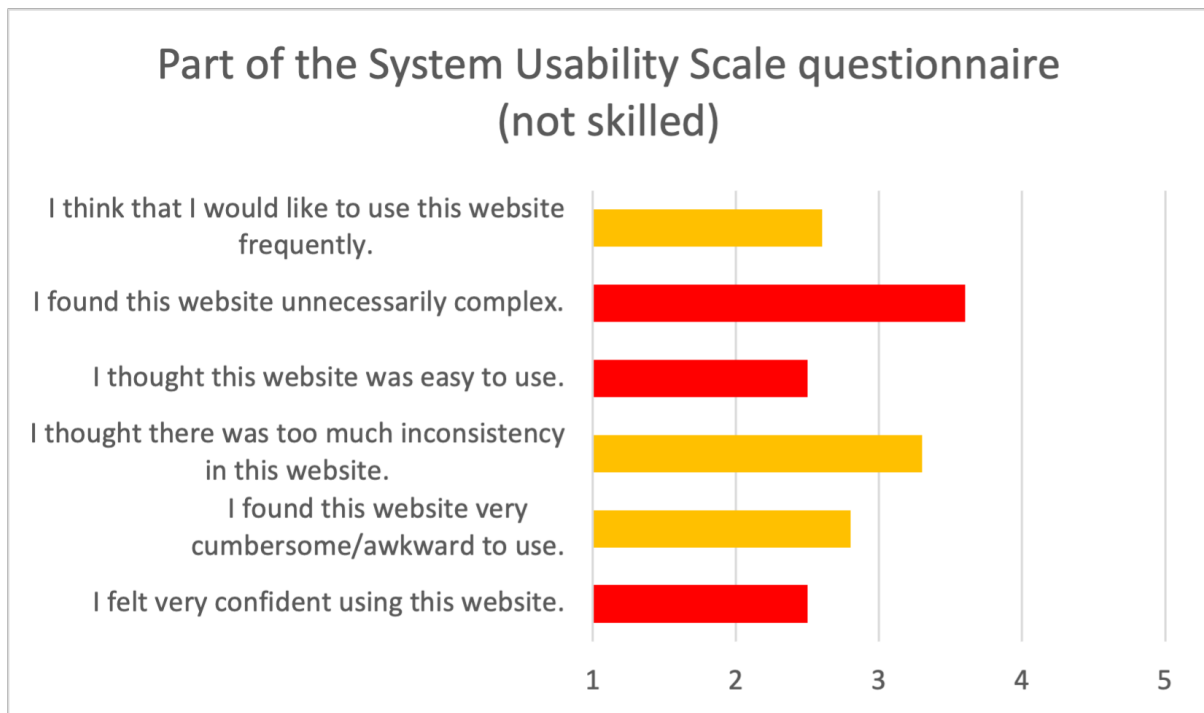
Figure 2.14: Not skilled users' opinion on satisfaction

The overall sentiment is not particularly good: users did not find it particularly easy to use, nor they felt confident during the task execution.

Still, it is to be said that skilled users where more condescending in their evaluation, giving scores that express a partial satisfaction with their usage. Not skilled users were generally more harsh, particularly relatively to their confidence while using the website.

3 | Conclusions

This section discusses the overall conclusions drawn from the study, from both the perspective of the Inspection and User Testing evaluation. Some suggestions for redesign in order to increase the usability of the site are also provided.

3.1.   Final Comment

Considering both the Inspection and User Testing results, we can finally give a general comment about the website's usability.

First of all, we have noticed that there have been no significant differences between the results of the users with technical skills (skilled) and the ones without technical skills (non-skilled). In other words, having knowledge about how websites are built did not help in any way in completing the tasks.

Generally, the concerns that came up with the Inspection were confirmed by the User Testing: the website is full of content and sometimes it is not easy to navigate through it. In each page, there is a huge quantity of links that easily distracts the user. In such a content-intensive website, the search function should be a powerful tool that allows to effectively retrieve a specific piece of information upon request. However, especially thanks to the User Testing, we noticed that only a part of the website content is reachable from the search bar, and the filters are not immediately visible, hence we think that its implementation does not comply with high usability standards. Plus, both the expert users and the recruited testers agree that the multiple donation functions are not well-integrated, since there are multiple forms (not always easy to reach), each with their own specific beneficiaries, to perform the same kind of interaction. This causes in the end user a sense of frustration, and may also impact their willingness to further interact with the website.

Concerning the visual design, the Inspection did not highlight any severe violation, and neither did the User Testing. Generally, the visual design is readable and consistent. Plus, many design choices effectively support the content and create empathy, with the

ultimate goal of raising awareness and collecting donations.

3.2. Suggestions for Redesign

This website is clearly overwhelmed with content, but we can hardly find a way to cut something out. In this sense, we suggest that a redesign activity should focus on improving navigation and consistency, by doing the following:

1. **Search bar improvement.** In such a content-intensive website, the search bar should be the easiest and fastest way to get a specific piece of information that may be fitted among images and other pieces of text, several clicks away from the home page. In this sense, we suggest the following:

   - The search bar should be more visible;

   - The user should be able to filter for date, topic, and content type (article, downloadable document, report, etc.) as soon as they click on the search function;

   - All the content of the website (also the text on the website's pages) should be reachable from the search bar.

2. **Bread crumbs improvement.** Bread crumbs have been implemented only in specific sections of the website, but they should be extended to all the sections, since they represent an easy way to keep the user informed about the current "depth" of the navigation and also an easy way to go back to a specific "depth" to partake another navigation path.

3. **Landmarks consistency improvement.** In some sections of the website, such as the one related to career opportunities, the main landmark (i.e. the one with UNICEF's logo) no longer takes the user to the home page, but takes the user back to the career section. This silent change of functionality easily causes disorientation and frustration to the user, who is forced either to notice a very little clickable text that takes them back to the home page, or to use browser functions to go back. We suggest that the landmarks should not change what they do during the navigation. If the user wants to go back to the career section and not to the home page, a correct implementation of bread crumbs may help in this sense.

4. **Donation consistency improvement.** Donations are undoubtedly one of the main reasons why UNICEF's website has been created. However, as the usability evaluation techniques clearly highlighted, there's no reason to keep many forms,

located in many different places, for making donations to different beneficiaries. Instead, all donations should be made in only one place and with only one form, so that the user always knows exactly where to go when they want to donate, and the user should be given the possibility to choose a specific beneficiary in that page.

A | Annexes

A.1. Inspection Individual Results

The following tables report the individual evaluations performed by all the inspectors who participated in this project (individual comments are not reported since they have been merged into a more significant comment in the Inspection chapter). They will be presented divided by category, along with some aggregated data. Each inspector is identified by the first letter of their name and their surname.

A.1.1. Content Heuristics

Content Heuristics

Code	Agreed Score	EM	RS	MS	MT
H10	5	5	5	3	5
M1	3	1	1	5	2
M2	3	4	5	1	4
M3	2	2	1	3	3
M4	4	4	4	5	3
Average	**3.4**	**3.2**	**3.2**	**3.4**	**3.4**

Table A.1: Individual scores for content heuristics

A.1.2. Navigation Heuristics

Navigation Heuristics

Code	Agreed Score	EM	RS	MS	MT
H1	2	2	2	2	2
H3	3	3	2	3	3

H7	2	2	4	3	3
M5	3	3	2	3	4
M6	3	5	3	2	3
M7	3	3	1	5	2
M8	2	3	2	2	3
M9	3	3	3	3	4
M10	4	2	5	4	4
Average	**2.78**	**2.89**	**2.67**	**3**	**3.11**

Table A.2: Individual scores for navigation heuristics

A.1.3. Presentation Heuristics

Presentation Heuristics

Code	Agreed Score	EM	RS	MS	MT
H2	5	4	5	5	4
H4	5	5	4	5	4
H5	3	4	3	4	4
H6	5	3	5	5	5
H8	3	1	2	4	3
H9	3	4	4	3	4
M11	5	5	2	5	5
M12	5	5	5	5	5
M13	4	4	5	5	3
M14	4	4	5	5	4
M15	4	4	4	5	4
M16	4	2	4	5	4
M17	5	5	5	5	4
M18	5	5	5	5	5
M19	4	4	5	5	4
Average	**4.27**	**3.93**	**4.2**	**4.73**	**4.13**

Table A.3: Individual scores for presentation heuristics

A.2.   User Testing Individual Results

The following tables report the results of the user testings performed by all the inspectors who participated in this project (additional comments were reported but are not shown here since they have been used to craft a more significant comment in the User Testing chapter). They will be presented divided by inspector.

A.2.1.   User Testing by Emanuele Musto (testers with no technical skills)

Tester 1

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	1	1m (time-out)	1m	1	0
2	0	2m 50s	2m 50s	2	0
3	3	2m 26s	1m 00s	2	1
4	3	1m 51s	28s	1	0
5	0	5m 54s	3m 16s	4	0
6	2	5m	4m 0s	2	1
7	3	3m 56s	3m 7s	0	4

Table A.4: User Testing by Emanuele Musto - Tester 1

Tester 2

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	33s	5s	0	0
2	1	5m (time-out)	4m 11s	2	1
3	3	50s	45s	0	0
4	3	22s	0s	0	0
5	1	7m (time-out)	4m 48s	4	1
6	1	5m (time-out)	4m 10s	1	1
7	3	2m 10s	27s	1	1

Table A.5: User Testing by Emanuele Musto - Tester 2

Tester 3

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	28s	5s	0	0
2	3	4m 24s	3m 30s	2	2
3	3	1m 0s	1m 0s	0	0
4	3	34s	0s	0	0
5	0	7m (time-out)	5m 22s	9	3
6	1	5m (time-out)	3m 46s	2	1
7	3	22s	0s	0	0

Table A.6: User Testing by Emanuele Musto - Tester 3

Tester 4

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	12s	0s	0	0
2	1	5m (time-out)	3m 30s	8	3
3	3	22s	0s	0	0
4	3	1m 0s	30s	1	0
5	0	7m (time-out)	4m	10	4
6	3	3m 56s	3m 35s	2	1
7	3	2m 20s	1m 42s	4	1

Table A.7: User Testing by Emanuele Musto - Tester 4

Tester 5

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar

1	3	1m	50s	0	0
2	3	4m 13s	3m 11s	1	1
3	3	46s	5s	0	0
4	3	17s	0s	0	0
5	1	7m (time-out)	5m 22s	9	7
6	3	45s	0s	0	1
7	3	19s	0s	0	0

Table A.8: User Testing by Emanuele Musto - Tester 5

A.2.2. User Testing by Riccardo Speroni (testers with technical skills)

Tester 1

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	1	1m (time-out)	40s	1	0
2	3	3m 58s	2m 8s	1	0
3	3	2m 57s	1m 37s	0	0
4	1	4m (time-out)	3m 7s	1	0
5	0	7m (time-out)	6m 30s	4	0
6	2	4m 23s	3m 6s	2	1
7	3	1m 3s	22s	0	0

Table A.9: User Testing by Riccardo Speroni - Tester 1

Tester 2

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	25s	8s	0	0
2	3	55s	15s	0	0
3	3	1m 51s	45s	0	0
4	3	29s	14s	0	0

5	2	5m 4s	3m 48s	3	0
6	1	5m (time-out)	4m 2s	4	1
7	3	3m 18s	1m 58s	2	1

Table A.10: User Testing by Riccardo Speroni - Tester 2

Tester 3

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	1m	33s	0	0
2	3	1m 3s	37s	0	0
3	3	24s	1s	0	0
4	3	1m 2s	10s	0	0
5	3	1m 58s	34s	0	1
6	3	47s	5s	0	1
7	1	6m (time-out)	4m 40s	8	6

Table A.11: User Testing by Riccardo Speroni - Tester 3

Tester 4

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	22s	2s	0	0
2	1	5m (time-out)	4m 20s	3	1
3	3	12s	1s	0	0
4	3	16s	2s	0	0
5	2	4m 36s	3m 12s	2	3
6	3	40s	5s	0	1
7	3	5m 13s	4m 5s	3	3

Table A.12: User Testing by Riccardo Speroni - Tester 4

Tester 5

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	18s	1s	0	0
2	3	28s	3s	0	0
3	3	44s	11s	0	0
4	3	1m 4s	21s	0	0
5	3	2m 5s	56s	2	1
6	0	4m 36s	3m 58s	4	0
7	3	1m 52s	59s	1	1

Table A.13: User Testing by Riccardo Speroni - Tester 5

A.2.3. User Testing by Matteo Spreafico (testers with no technical skills)

Tester 1

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	25s	0s	0	0
2	3	2m 34s	30s	3	0
3	3	50s	10s	0	0
4	3	24s	0s	0	0
5	3	2m 04s	20s	1	0
6	1	5m (time-out)	2m	3	0
7	3	3m 15s	30s	2	0

Table A.14: User Testing by Matteo Spreafico - Tester 1

Tester 2

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	13s	0s	0	0
2	3	1m 25s	25s	2	0

3	3	1m 05s	35s	0	0
4	3	27s	10s	0	0
5	1	7m (time-out)	4m	4	0
6	1	5m (time-out)	3m	3	0
7	3	44s	0s	0	0

Table A.15: User Testing by Matteo Spreafico - Tester 2

Tester 3

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	26s	3s	0	0
2	2	3m 48s	1m 10s	4	1
3	3	25s	5s	0	0
4	3	1m 30s	30s	1	0
5	2	6m 38s	4m	4	3
6	3	2m 25s	40s	1	1
7	3	2m 26s	45s	2	2

Table A.16: User Testing by Matteo Spreafico - Tester 3

Tester 4

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	25s	0s	0	0
2	3	4m 40s	3m	2	0
3	3	58s	40s	0	0
4	3	35s	20s	0	0
5	1	7m (time-out)	6m	3	2
6	2	5m	3m	2	1
7	3	2m 52s	1m 30s	1	1

Table A.17: User Testing by Matteo Spreafico - Tester 4

Tester 5

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	30s	10s	0	0
2	3	1m 40s	40s	1	0
3	3	52s	20s	0	0
4	3	25s	5s	0	0
5	2	5m 14s	3m	3	0
6	1	5m (time-out)	3m 30s	4	0
7	3	46s	10s	0	0

Table A.18: User Testing by Matteo Spreafico - Tester 5

A.2.4. User Testing by Manuel Tacca (testers with technical skills)

Tester 1

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	25s	0s	0	0
2	3	1m 12s	39s	1	0
3	3	35s	18s	0	0
4	3	30s	12s	0	0
5	3	1m 13s	20s	0	0
6	2	4m 30s	3m 30s	3	1
7	3	1m 30s	10s	1	0

Table A.19: User Testing by Manuel Tacca - Tester 1

Tester 2

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	21s	0s	0	0
2	3	29s	13s	0	0

3	3	33s	15s	0	0
4	3	37s	15s	0	0
5	3	2m 18s	1m 58s	3	0
6	2	3m 53s	3m 20s	2	1
7	3	1m 21s	40s	1	0

Table A.20: User Testing by Manuel Tacca - Tester 2

Tester 3

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	26s	0s	0	0
2	3	2m 07s	1m 52s	1	0
3	3	30s	7s	0	0
4	3	14s	0s	0	0
5	2	2m 59s	2m 05s	4	0
6	2	2m 39s	2m 24	1	1
7	3	21s	0s	0	0

Table A.21: User Testing by Manuel Tacca - Tester 3

Tester 4

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	12s	0s	0	0
2	3	48s	20s	1	0
3	3	21s	0s	0	0
4	3	15s	0s	0	0
5	2	5m 7s	4m 12s	5	0
6	2	4m 1s	3m 41s	4	1
7	2	4m 57s	2m 50s	4	1

Table A.22: User Testing by Manuel Tacca - Tester 4

Tester 5

Task No.	Score	Total Time	Wandering Time	#Errors/ Wrong Paths	#Uses of Search Bar
1	3	20s	0s	0	0
2	3	1m 2s	40s	1	0
3	3	57s	47s	0	0
4	3	3m 27s	2m 20s	3	0
5	2	1m 54s	1m 20s	0	0
6	2	2m 58s	2m 31s	1	1
7	2	2m 24s	1m 38s	2	0

Table A.23: User Testing by Manuel Tacca - Tester 5

A.3. Division of Labour and Gathering of Data

Everyone involved in this report participated in every aspect of the project: everybody performed an individual inspection, proposed a set of tasks, performed user testing with 5 testers, and analyzed the collected data. However, in order to be more efficient, we decided to split the remaining work as follows:

- Emanuele Musto and Matteo Spreafico have written the most of this report, but have let Riccardo Speroni and Manuel Tacca enter their own data and write some parts of it (more specifically, Riccardo Speroni produced all the visualizations of section 2.3, and Manuel Tacca wrote all their textual comments).

- Riccardo Speroni and Manuel Tacca have rephrased the selected tasks (that were in a sort of "draft" form) in a more appropriate way and have built the two forms that were used to collect data in the User Testing phase. The forms have been built using the *Microsoft Forms* tool.

Notably, Emanuele Musto provided the other inspectors with an Excel sheet with which inspectors could easily enter the data they observed during the user testings (such data have been reported in the previous tables), and Matteo Spreafico adapted a selected LaTeX thesis template of Politecnico di Milano to this project's needs.

List of Figures

List of Tables