

# Reading disability and document access, a possible approach

**Final proof of concept report**

Pilot Project 09 04 77 25

© European Union, 2021

The Commission's reuse policy is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39 – <https://eur-lex.europa.eu/eli/dec/2011/833/oj>).

Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed, provided that appropriate credit is given and any changes are indicated.

The contents of this publication do not necessarily reflect the position or opinion of the European Commission.

# Contents

|   |          |
|---|----------|
| <b>FINAL REPORT PHASE 2 .....</b>                                   | <b>5</b> |
| 1. INTRODUCTION .....   | 5        |
| 2. SECTION I: SUBSET CREATION.....                                  | 12       |
| 2.1. <i>Subset definition</i> .....                                 | 13       |
| 2.2. <i>Criteria rational</i> .....                                 | 13       |
| 2.3. <i>Methodology and original subset</i> .....                   | 15       |
| 3. SECTION II: SOLUTIONS AND IMPLEMENTATION .....                   | 17       |
| 3.1. <i>Shortlisted solutions</i> .....                             | 18       |
| 3.2. <i>Solutions, processes and implementation</i> .....           | 19       |
| 3.2.1 <i>Proof of Concept RDDA Website</i> .....                    | 20       |
| 3.2.2 <i>AbleDocs</i> .....   | 23       |
| 3.2.3 <i>Accessibility Tagging solution under development</i> ..... | 25       |
| 3.2.4 <i>FACIL'iti</i> .....  | 28       |
| 3.2.5 <i>Siteimprove Accessibility</i> .....                        | 32       |
| 3.2.6 <i>PAC2021</i> .....  | 40       |
| 3.3. <i>Key findings</i> .....                                      | 43       |
| 4. SECTION III: EVALUATION AND ANALYSIS.....                        | 44       |
| 4.1. <i>Scope</i> .....   | 45       |
| 4.2. <i>Methodology</i> .....                                       | 47       |
| 4.2.1 <i>Manual evaluation by EDA</i> .....                         | 48       |
| 4.2.2 <i>Semi-automated evaluation using Siteimprove</i> .....      | 50       |
| 4.2.3 <i>Semi-automated evaluation using PAC2021</i> .....          | 50       |
| 4.3. <i>Analysis</i> .....  | 52       |
| 4.3.1 <i>Language comparison</i> .....                              | 53       |
| 4.3.2 <i>Type comparison</i> .....                                  | 56       |

|  |    |
|--|----|
| 4.3.3 <i>Gain in accessibility</i> .....                         | 57 |
| 4.4. <i>Key findings</i> .....                                   | 74 |
| 5. CONCLUSIONS .....   | 76 |
| 6. LIST OF ACRONYMS .....  | 81 |
| 7. ANNEX .....   | 82 |
| 7.1. <i>Original subset and references for all subsets</i> ..... | 82 |
| 7.2. <i>Generic survey</i> .....                                 | 83 |
| 7.3. <i>FACIL'iti survey</i> .....                               | 87 |

# Reading disability and document access, a possible approach

## Final report Phase 2

### 1. Introduction

The European Parliament (EP) has mandated the Publications Office to carry out the **“Reading disability and document access, a possible approach” pilot project**. As pointed out in the introductory note of Mr Daniele Viotti, former Italian MEP, reading disability<sup>1</sup> is one of the most common neurological conditions that affect European citizens. Public institutions must be accessible and transparent for everyone, including for the most vulnerable people. In recent years, EU institutions made efforts to improve the accessibility of their documents by establishing more user-friendly publications and public websites. However, despite these efforts, access to relevant EU documents/publications remain problematic for people with reading disabilities.

In light of the provisions of “Directive (EU) 2016/2102 of the European Parliament and of the Council on the accessibility of the websites and mobile applications of public sector bodies”<sup>2</sup>, it is important to develop and make available appropriate instruments to ensure that people with reading disabilities have equal access to the documents/publications and websites of EU institutions.

Until 2019, one of the goals of the Publications Office (OP) Strategic Objective (SO 4) “Linked EU information, increased interoperability and federated search” is to enhance the accessibility of existing documents published by the different institutions. Consequently, ensuring that people with reading disabilities can access documents/publications and websites falls within the scope of this Strategic Objective (SO). Therefore, the implementation of this pilot project is entrusted to the OP.C.1 – OP Portal unit which was in charge of SO4.

---

<sup>1</sup> A reading disability can be considered as a condition for which an individual experiences difficulty in reading irrespective of intelligence or visual acuity. Reading disabilities include, for example, dyslexia (learning disorder affecting reading) and alexia (loss of ability to read). For the purposes of the study, the analysis has focused mainly on dyslexia.

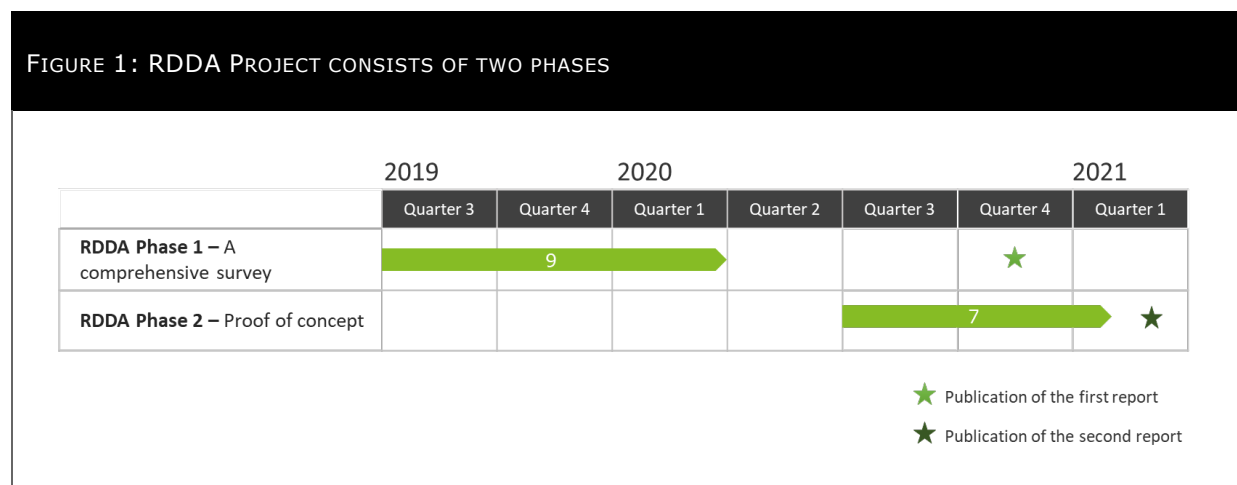
<sup>2</sup> <https://eur-lex.europa.eu/eli/dir/2016/2102/oj>

The project is complementary with the accessibility projects for users with visual impairments managed by the B2/B3 units of the OP.

The 16-month project consists of **two phases**:

- 1) **A comprehensive survey** assessing the targeted audience, the targeted documents/publications and websites, and the relevant tools for enhancing accessibility for users with reading disabilities, mainly dyslexia.
- 2) **Proof of concept (PoC)** by enhancing the accessibility of selected publications and websites.

FIGURE 1 presents the timeline of these phases. The first phase started in July 2019 and ended in March 2020 with the publication of the phase 1 report in December 2020. The second phase started in October 2020 and ended in January 2021 with the publication of the phase 2 report in March 2021.



Deloitte assisted OP in the implementation of phase 2 of the project, in the realisation of the proof of concept.

The whole project tackles the following aspects:

- Define and assess the needs of target user groups.
- Research best practices and existing technologies, including the most recent developments in the area of artificial intelligence.
- Provide guidelines and best practices for producers of information and facilitate accessibility and searchability of “reading disorder-friendly” documents/publications and websites for end-users.
- Provide better access to a selection of official documents/publications and websites for people with reading disabilities by converting them to a suitable form, as defined by the outcomes of the study.
- Estimate the effort and time needed by EU institutions, bodies, and agencies to modify (or replace) existing documents/publications, websites and document production workflows to achieve the targeted level of accessibility.

The aim of the proposed “Reading disability and document access, a possible approach” pilot project is to **make the documents/publications and websites** of the Commission and EU institutions more accessible for people with reading disabilities. It describes state-of-the-art best practices as a basis for recommending how to create a

“reading disorder-friendly” environment for the publications and websites of EU institutions.

It focuses on existing documents/publications (for example PDF and EPUB) and websites (which are mainly in HTML format) published by the European institutions.

First, the project seeks to define the targeted user groups (different levels and nature of disability) and to select a subset of documents/publications and websites of various European institutions in order to assess their accessibility.

Secondly, it identifies the technologies, tools, formats, layout, fonts (e.g., openDyslexia) and best practices that can be applied in an automatic or semi-automatic fashion to existing publications and websites.

A proof of concept based on selected documents/publications and websites demonstrate how the available tools can enhance accessibility.

Based on the project outcomes, the Publications Office proposes recommendations that the institutions could implement with regard to their documents/publications and websites.

### ***Phase 1: A comprehensive study***

The outcome of phase 1 of the RDDA project are presented in a consolidated manner in the *Reading disability and document access, a possible approach* publication<sup>3</sup>.

The first part of the comprehensive study identifies the targeted user groups, explaining the problems and needs of each audience.

The second part of the study assesses the level of accessibility of the documents/publications and websites of the Publications Office, European Commission, and other European institutions for the target groups. It identifies a subset of documents/publications and websites that should be made available, with the first priority being to create a “reading disorder-friendly” environment for the European institutions. The relevant best practices, technologies, tools, and IT Software that could enhance accessibility are described for each type of document/publication and website.

The third part contains an assessment of current technologies, tools, formats, layout, fonts (e.g., openDyslexia), best practices, and IT software that could be used to make existing documents/publications (for example PDF and EPUB) and existing websites (which are mainly in HTML format) more accessible. It describes the state-of-the-art accessibility solutions for users with reading disabilities, and provides full information on

---

<sup>3</sup> Publication Office, Reading disability and document access, a possible approach, Final study report, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-191689283>, [accessed on 18.02.2021]

each tool: where to source it, cost, type of document to which it applies, fully automatic or semi-automatic implementable, and targeted audience.

The project team contacted European institutions, user associations, and civil society bodies to gather the most comprehensive information possible to conduct the study.

Finally, it proposes a roadmap for creating a “reading disorder-friendly” environment for EU institutions. It links the assessment of tools, publications/websites, and targeted user groups in a comprehensive description of how to make the European institutions more accessible. It also proposes a work plan identifying “low-hanging fruits” to be exploited as well as a roadmap with milestones.

The study has been communicated to the selected European institutions that own the websites assessed in the study, to enable them to decide how to best leverage the contents. The outcomes of the study are made public so as to enable all organisations in Europe to benefit from it.

The work carried out in phase 1 defines the scope of work for phase 2. In concrete terms, one of the deliverables of phase 1 is to define technical specifications, guidelines, and a roadmap for identifying the documents/publications and websites that could be converted and made available to the target group.

### ***Phase 2: Running a proof of concept***

The aims of this report are to present in a consolidated manner the outcome of RDDA phase 2. The activities of phase 2 of the project are described in detail in the following paragraph.

Using all the results of phase 1 (study, handbook, and roadmap), a subset of most recent documents/publications and websites is made more accessible using the best practices, technologies, tools and IT software identified by the study. As much as possible the modification of documents/publications and websites is undertaken using automatic or semi-automatic processes. As the budget allows, as many documents/publications and websites are enhanced for accessibility. Finally, an evaluation of the enhanced documents/publications and websites involving users with reading disability is carried out.

### ***Content and structure of the final report***

Based on the activities of phase 2 of the project, the final report is structured as follows:

- **Section I: Subset creation**
  - This section describes the methodology, explains the criteria used to create a subset and it lists the selected documents for remediation and for evaluation.
- **Section II: Solutions and implementations**
  - This section describes the solutions used for remediation and for evaluation and it documents the automated processes of each remediation solution.
- **Section III: Evaluation and analysis**
  - This section describes the manual and automated accessibility evaluation. It presents the list of questions for end users’ evaluation and documents the process of accessibility assessment with semi-automated solutions.



The section presents an analysis of the results with a comparison of the accessibility level and users satisfaction before and after the implementation of the remediation solutions.

- **Conclusion**

- This section summarises the key findings of the proof of concept.

FIGURE 2: STRUCTURE OF THIS REPORT

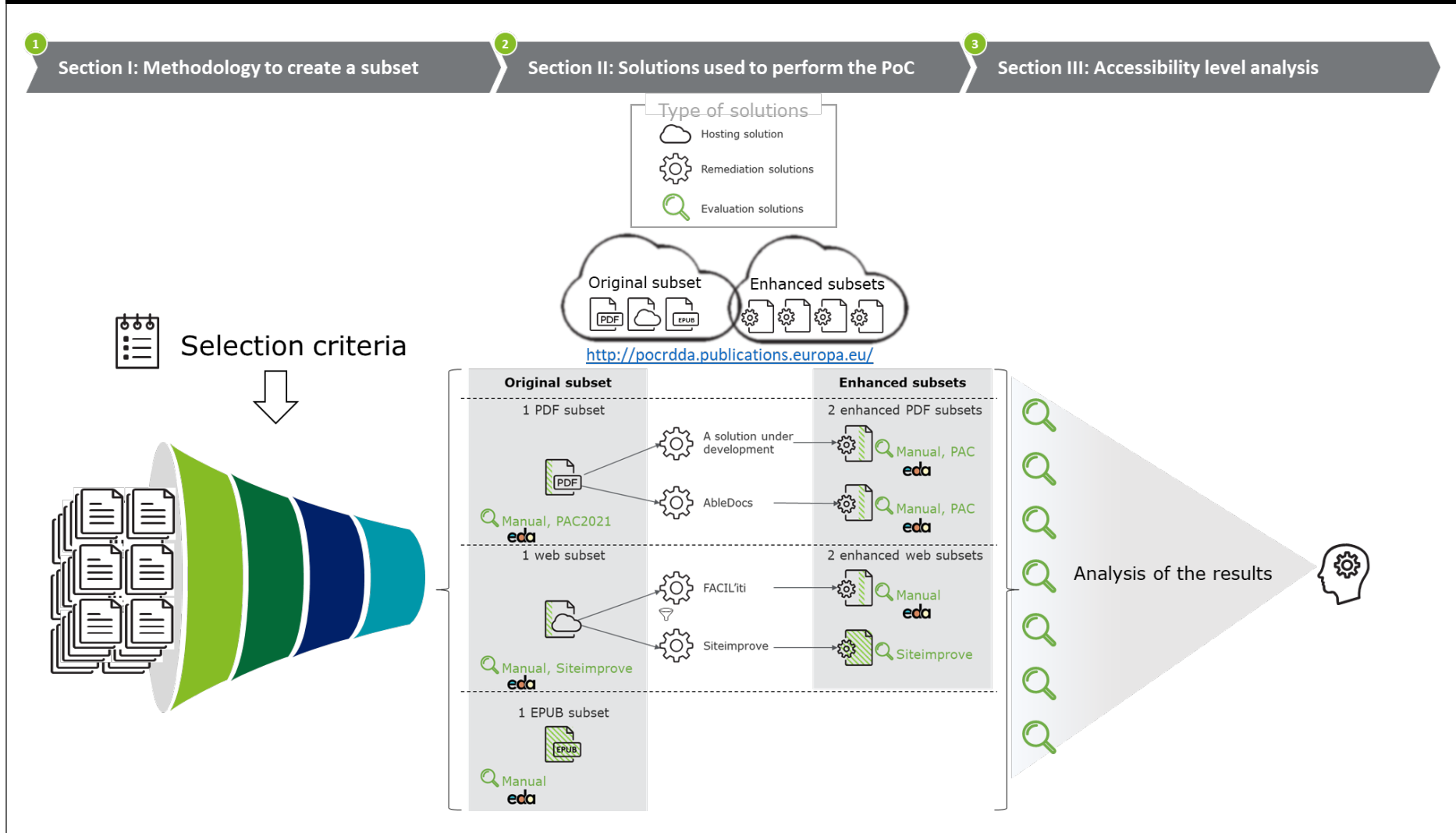


FIGURE 2 presents the structure of this report:

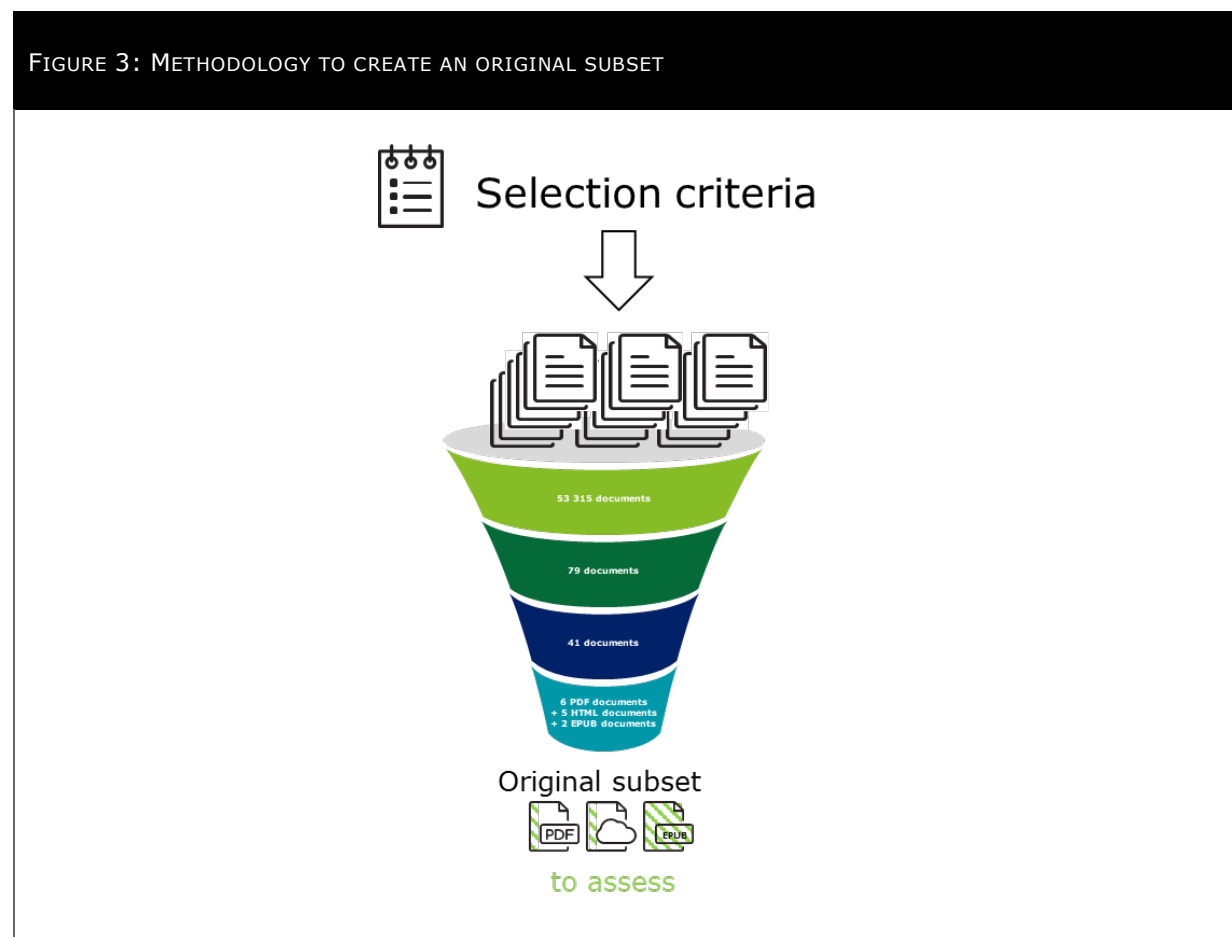
- Section I corresponds to the description of the methodology to create a subset. First, selection criteria are defined and then the methodology is applied to create an original subset. This original subset is made up of PDF, HTML and EPUB documents.
- Section II corresponds to the description of the solutions and their use to perform the PoC. Three types of solutions are used for the PoC:
  - **Hosting solution** for the PoC environment: A [hosting solution](#)<sup>4</sup> has been deployed to host the original and enhanced subsets.
  - **Remediation solutions** to enhance documents: AbleDocs and a solution under development have been selected to remediate the original PDF subset to create two enhanced PDF subsets. FACIL'iti and Siteimprove have been selected to remediate the original web subset to create 2 enhanced web subsets. As Siteimprove is a semi-automated solution, only one page from the original subset has been manually remediated as an example. The EPUB subset is not remediated only evaluated.
  - **Evaluation solutions** to assess the accessibility level of documents: The evaluation of the accessibility level of documents is made manually by the European Dyslexia Association (EDA) and semi-automatically using PAC for PDF and Siteimprove for HTML and PDF. The evaluation is made on the original subset before remediation and on the enhanced subsets after remediation. Since the evaluation process is manual or semi-automatic but not fully automatic, not all documents in the subset will be evaluated, but only a part of them.
- Section III corresponds to the accessibility level analysis. First, a focus is made on the manual evaluation by EDA. Then, the results from EDA and the evaluation solutions are analysed in order to assess the gain in accessibility and the users experience before and after the implementation of the remediation solutions.

---

<sup>4</sup> Hosting solution, <http://pocrdda.publications.europa.eu/>, [accessed on 8.12.2020]

## 2. Section I: Subset creation

The Subset creation section consists in identifying the documents, publications and web pages published by the European Union for remediation and for accessibility level assessment. The purpose of this section is to describe the methodology and explain the criteria used to create a subset of content that includes PDF documents and web pages in multiple languages as shown in FIGURE 3.



Section I of the report is structured as follows:

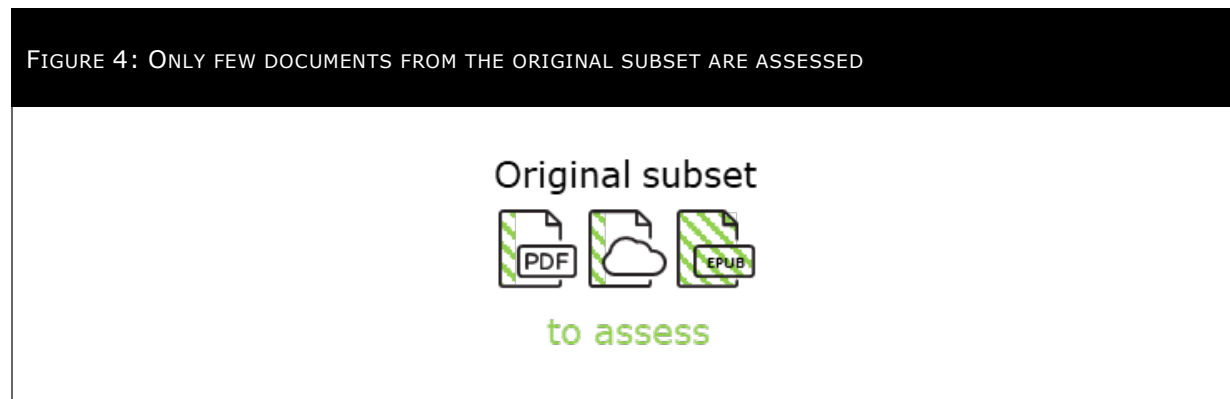
- **Subset definition,**
- **Criteria rational** used to create the original subset of publications to remediate,
- **Methodology and original subset,** methodology implemented to select documents for the original subset to remediate.

Throughout the rest of this document, we shall use the terms:

- Either "subset" or "original subset" to refer to the original list of documents that are selected for remediation.
- Either "enhanced subset" or "remediated subset" to refer to the remediated version of the original document by a remediation solution.

## 2.1. Subset definition

This subsection defines the subset. The original subset is composed of publications that will be enhanced by multiple remediation solutions. As the accessibility assessment of these publications is also made manually, only part of the original subset and remediated subset is assessed. FIGURE 4 shows that only few documents from the original subset are evaluated.



## 2.2. Criteria rational

This subsection lists and explains the criteria used to create the original subset. A summary of the criteria is provided at the end of this subsection.

The OP is an inter-institutional office whose task is to publish the publications of the institutions of the European Union through its website, among others. As [OP Portal](#)<sup>5</sup> covers multiple publications from different Directorate-General (DG) of the EU, all pages from the original subset are selected **from OP Portal**.

Documents and publications published by the Publications Office of the European Union are mostly in PDF format and sometimes in HTML and EPUB format. This is why the subset includes **PDF, HTML and EPUB formats**.

OP Portal hosts more than 50 000 EU publications. Due to limited resources and time, the original subset is limited to **200 pages** including **20 pages to assess, HTML and PDF combined**. The purpose of the criteria described in this section is to ensure the representativeness of these 200 pages. As requested by EDA, **two additional EPUB** documents have been added to the original subset for evaluation but will not be remediated.

---

<sup>5</sup> OP Portal, <https://op.europa.eu/en/home>, [accessed on 31.11.2020]

Documents/publications and websites published by OP may be available in the 24 official languages of the EU. Based on EDA recommendations, **the 5 languages** selected to make up the subset use different alphabets, different level of transparency and together covers more than 50% of native speakers in Europe as shown in TABLE 1.

| <b>TABLE 1: LANGUAGES FOR THE SUBSET</b> |          |                  |   |
|--|----------|------------------|---|
| <b>Languages</b>                         | Alphabet | Transparency     | % of native speakers in Europe <sup>6</sup> |
| <b>Italian</b>                           | Latin    | transparent      | 13%   |
| <b>French</b>                            | Latin    | semi-transparent | 12%   |
| <b>German</b>                            | Latin    | semi-transparent | 18%   |
| <b>English</b>                           | Latin    | non-transparent  | 13%   |
| <b>Greek</b>                             | Greek    | -                | -   |

In order to allow languages comparison in the analysis of the accessibility results, the same content must be **available in all 5 languages**. Thus, all pages from the original subset are in all 5 languages. The number of pages for the original subset is equally distributed per language, so that the original subset includes 40 pages per language (=200/5) to be remediated and 4 to be assessed (=20/5).

Based on the EDA recommendations, people with reading disability are interested in reading the same content as everyone else. This is why, **the most visited content** is selected for the original subset.

Based on the EDA recommendations, **contents must be understandable by anyone** (ISCED level 2 - International Standard Classification of Education) and **not specialised**. For this reason, all pages from the original subset are not specialised and can be understood by anyone.

To understand the impact of the layout on the remediation and the accessibility evaluation, the pages selected for the original subset have a **diversified layout** (plain text, images, tables, double column, background image, different colours, different fonts, link, footnote ...).

As previously mentioned in the scope of Phase 2: Running a proof of concept (Section 1 Introduction), a subset of **most recent** documents/publications and websites is made more accessible. The pages selected for the original subset include the most recent pages.

---

<sup>6</sup> Commission Européenne, Eurobaromètre, « LES EUROPEENS ET LES LANGUES », 2005, [https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs\\_237.fr.pdf](https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_237.fr.pdf), [accessed on 9.12.2020]

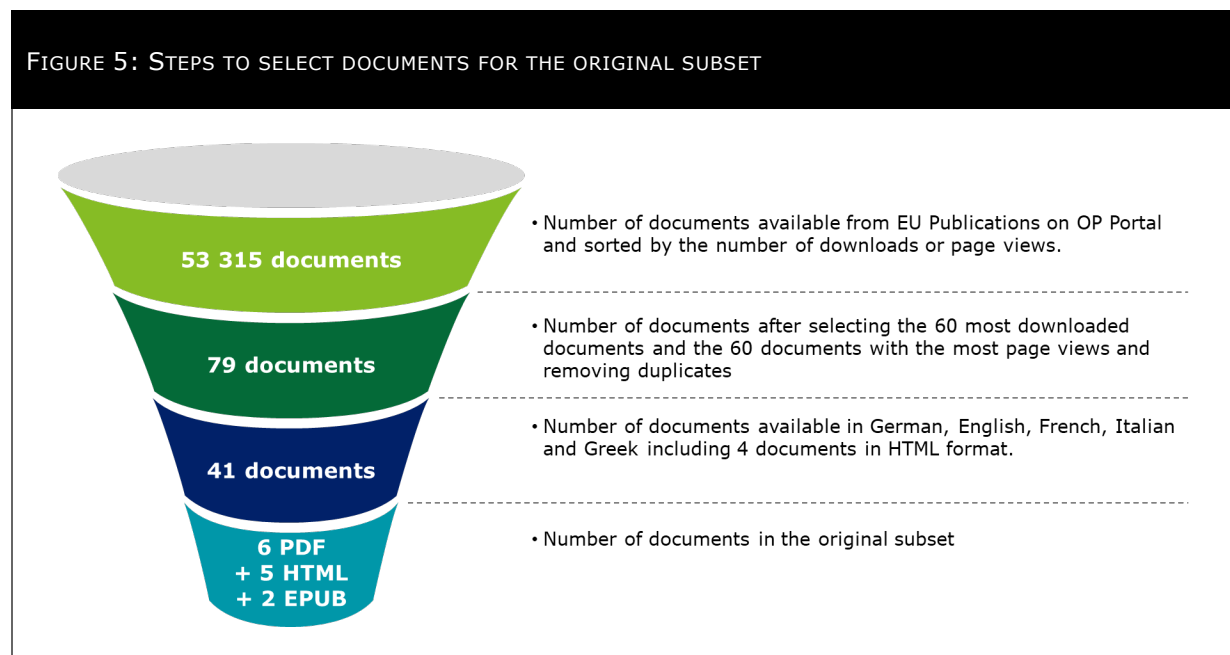
Finally, content that best matches the following criteria has been selected for the original subset:

- Public content from OP Portal,
- Content available in HTML, PDF or EPUB formats,
- Content available in the 5 languages (Italian, French, German, English, Greek)
- Most visited content,
- Understandable by anyone (ISCED level 2),
- Diversified layout (plain text, images, tables, ...),
- Most recent content.

### 2.3. Methodology and original subset

Based on the criteria described on the previous subsection 2.2, the methodology to select pages for the original subset is described in this subsection.

FIGURE 5 presents the different steps to reduce the number of documents and select the 40 pages per language to remediate.



Firstly, the 53 315 documents from EU Publications on OP Portal are sorted by the number of downloads and page views from the statistics of the last 6 months<sup>7</sup>. Then, 79 documents remained after selecting the 60 most downloaded documents and the 60 documents with the most page views and removing duplicates. Only 41 of those

---

<sup>7</sup> Analysis made in October 2020.

documents are available in German, English, French, Italian and Greek including 4 documents in HTML format. The **original subset** is reduced to the following documents in the 5 languages:

- the 6 PDF documents with the most page views,
- the 4 HTML publications,
- 1 additional HTML web page,
- And 2 EPUB documents.

This original subset, as shown in **FIGURE 5**, is the original subset to remediate. As explained in section 2.1, the original subset to assess is only a part of the original subset. The documents to be assessed have been selected as follows:

- 2 PDF documents such as one document is recent (2020) and one is older (2014),
- 2 HTML documents such as one is a web page and one is also a publication assessed in PDF format,
- 2 EPUB documents.

The original subset to remediate is presented in annex 7.1 under the “original subset” sheet. Documents selected for assessment are tagged in column J. Also, section 4.1 presents the scope of the original and enhanced subsets including the documents to assess.

The full list of documents selected for the original subset are listed in the annex 7.1 under “References for all subsets” sheet with a filter “Original” on column B. A [PoC website](#)<sup>8</sup> has been developed in order to host original and enhanced subsets until 2021 under the page “documents”.

The next section focuses on the solutions used for remediation and for accessibility assessment. The remediation process of the original subset is explained for each solution and the semi-automated solutions used for accessibility assessment are presented.

---

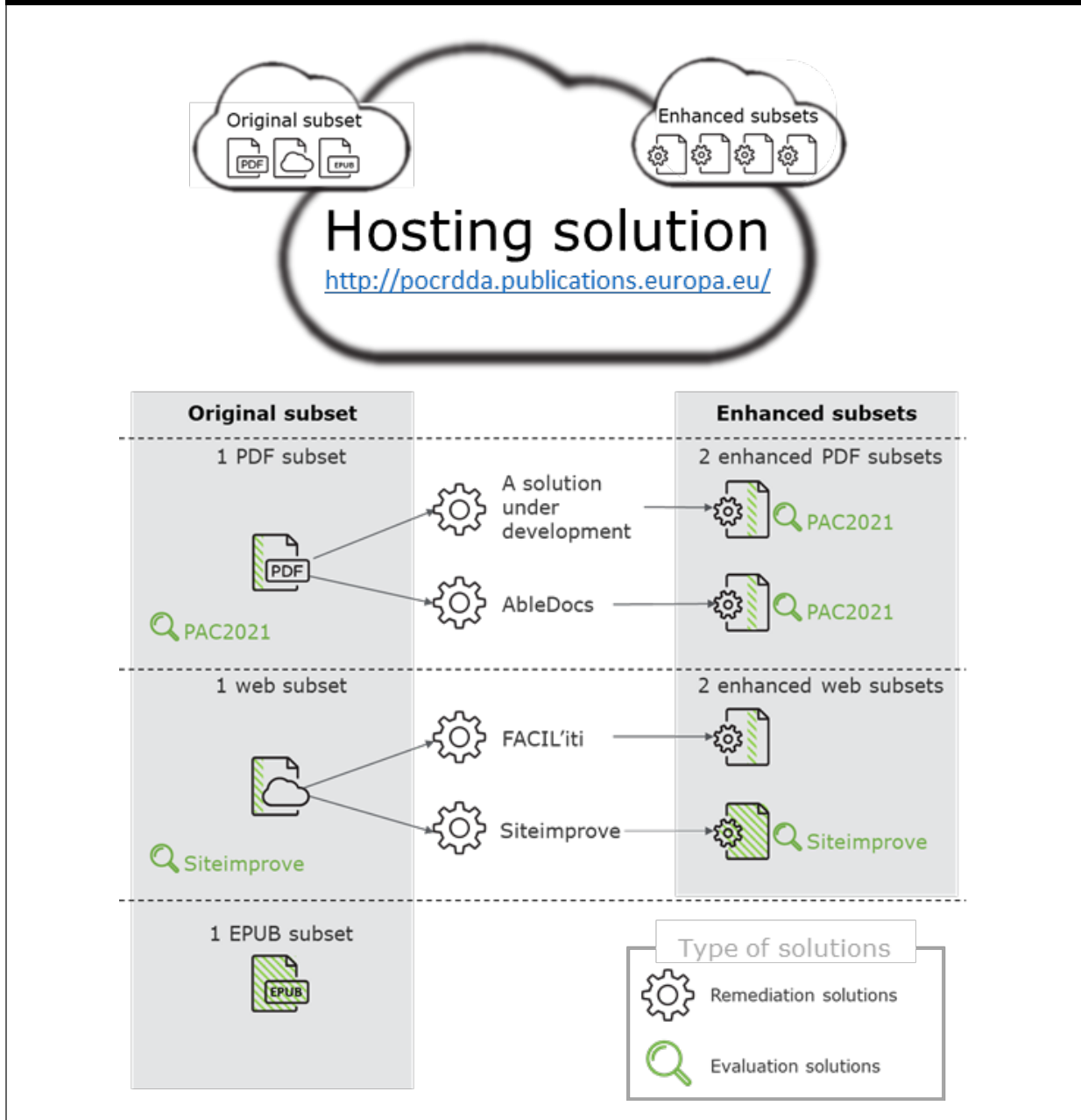
<sup>8</sup> PoC RDDA website, <http://pocrdda.publications.europa.eu/index.html>, [Accessed on 30.11.2020]



### 3. Section II: Solutions and implementation

On the Solutions and Implementations section of the project, a collaboration with the solution providers allowed us to understand the different types of solutions and how to apply them in a Proof of Concept. The purpose of this section is to describe the implemented solutions for hosting, remediation and evaluation, as well as explain the automated processes of each remediation solution, as shown in **FIGURE 6**.

FIGURE 6: SOLUTIONS USED TO PERFORM THE PoC



Section II of the report is structured as follows:

- **Shortlisted solutions** and how they were integrated in the PoC,
- **Solutions processes and implementation** explanation.

### 3.1. Shortlisted solutions

The shortlisted solutions are based on [the annex 9.4 of the RDDA phase 1 final report](#)<sup>9</sup>. The following criteria has been defined to select the solutions to be implemented in the PoC, from the long list of solutions in annex 9.4:

- One solution dealing with the subsets hosting, to provide the PoC environment.
- At least two solutions dealing with the PDF subset, to cover the fully automated remediation of PDF documents.
- At least two solutions dealing with the web subset, to achieve full/semi-automated remediation of web pages.
- At least one solution dealing with the evaluation of PDF, to semi-automatically assess the accessibility level of PDF.
- At least one solution dealing with the evaluation of web pages, to semi-automatically assess the accessibility level of HTML pages.

Based on these criteria, providers and solutions have been selected considering their availability and willingness to participate free of charge.

This subsection lists and summarizes the information regarding the solutions that were analysed for this PoC. The chosen solutions are listed by type (Hosting, Remediation and/or evaluation), subset format (HTML, PDF and/or EPUB) and user<sup>10</sup> (authors and contributors, publishers and webmasters or end-users), as shown in TABLE 2.

| Type               | Subset format       | User       | Solution         | Description  |
|--------------------|---------------------|------------|------------------|--|
| <b>Hosting</b>     | HTML<br>PDF<br>EPUB | End-users  | PoC RDDA Website | Gathers and hosts original and enhanced subsets, providing an interface to facilitate the navigation and subset evaluation.      |
| <b>Remediation</b> | PDF                 | Publishers | AbleDocs         | Remediates and converts original PDF subset to enhanced PDF subset, with fully automated process and PDF/UA standard compliance. |

---

<sup>9</sup> Publications Office, RDDA phase 1 final report, [http://pocrdda.publications.europa.eu/library/REPORTS/tools\\_identification.xlsx](http://pocrdda.publications.europa.eu/library/REPORTS/tools_identification.xlsx), [accessed on 06.01.2021]

<sup>10</sup> Publications Office, Section III: Assessment of the tools of RDDA phase 1 report, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

**TABLE 2: SHORTLISTED SOLUTIONS**

|                                   |             |   |  |   |
|-----------------------------------|-------------|---|--|---|
| <b>Remediation</b>                | PDF         | Publishers  | Accessibility Tagging solution under development | Remediates and converts original PDF subset to enhanced PDF subset, with fully automated process and PDF/UA standard compliance.  |
| <b>Remediation</b>                | HTML        | End-users   | FACIL'iti  | Remediates and converts original web subset to enhanced web subset with fully automated process focusing on dyslexia.   |
| <b>Evaluation and Remediation</b> | HTML<br>PDF | Authors<br>Contributors<br>Publishers<br>Webmasters | Siteimprove<br>Accessibility                     | Evaluates original web subset and automates detailed accessibility report, providing guidelines to manually remediate / enhance web subset (semi-automated process) based on WCAG standard and PDF/UA standard. |
| <b>Evaluation</b>                 | PDF         | Authors<br>Publishers                               | PAC2021  | Supports the manual creation of accessible PDF files and evaluates PDF/UA compliance.   |

### 3.2. Solutions, processes and implementation


This subsection describes the solutions and explains their processes and implementation for remediation and accessibility assessment. First, the hosting solution is described, next remediation solutions and then the evaluation solutions. Each of the following subsection is structured as follows:

- Short description of the solution and the purpose of the solution for the PoC,
- A detailed table providing a general description of the solution, users and provider descriptions of the solution,
- Implementation of the solution.

### 3.2.1 Proof of Concept RDDA Website

To deliver an environment for the PoC, a user interface was designed and hosted in a web server with a [tailored domain](#)<sup>11</sup>. The goal of this PoC website is to enable access to the original and remediated subsets through an easy navigation.

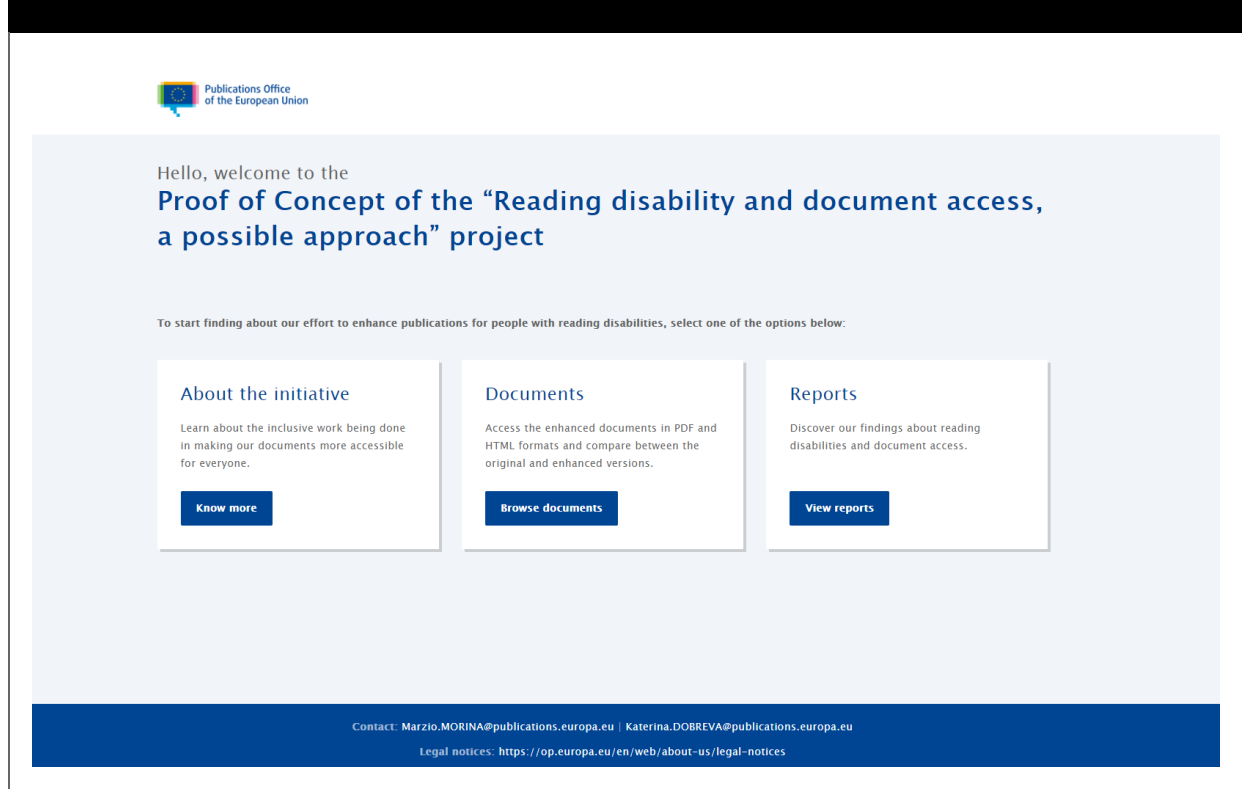
In the following table you can find the website main features:

| <b>TABLE 3: PoC RDDA WEBSITE</b>  |  |
|---|--|
|  |  |
| General description   |  |
| <b>Name of the solution:</b>  | PoC RDDA Website   |
| <b>Description of the solution:</b>   | The website applies the user interface and hosting solution used to perform the PoC RDDA. The solution includes: <ol style="list-style-type: none"> <li>1. User interface design</li> <li>2. Website development (HTML, CSS and JavaScript)</li> <li>3. Web hosting (VPS)</li> </ol> |
| <b>Price:</b>   | N/A  |
| <b>Languages:</b>   | Currently the PoC RDDA Website supports one language (English) for the platform navigation and five languages (German, Greek, English, French and Italian) for the subsets download / view   |
| <b>Alphabets:</b>   | All alphabets can be supported, even Cyrillic, Farsi and Asian character derived languages.  |
| <b>Environment:</b>   | The PoC RDDA website is accessible via any browser and device.   |
| <b>Features:</b>  | The website provides a clear and accessible hierarchical navigation structure, follows the UE official style guide, and uses standard web technologies.  |
| <b>Standards and guidelines:</b>  | The interface design and development follow the UE official style guide and uses standard web technologies with best practices applied using WCAG standard.  |
| <b>Implementation:</b>  | The PoC interface design and website implementation requires two weeks, including internal navigation and original subsets integration. To add the remediated subsets and adjust the navigation, one more week is required.  |
| <b>Demonstration:</b>   | <a href="http://pocrdda.publications.europa.eu/">http://pocrdda.publications.europa.eu/</a>  |
| Users   |  |
| <b>Example of users:</b>  | The PoC RDDA Website is available to the general public, but its purpose is providing access to testers for original and enhanced subsets evaluation. The testers are target end-users   |

<sup>11</sup> PoC RDDA Website URL, <http://pocrdda.publications.europa.eu/>, [available until November 2021]

**TABLE 3: PoC RDDA WEBSITE**

|                                     | with direct or indirect relation with the PoC performance. The PoC RDDA Website will also be used until November 2021 for the dissemination of results.                         |
|-------------------------------------|---|
| Provider                            |   |
| <b>Name of the provider:</b>        | Publications Office of the EU   |
| <b>Description of the provider:</b> | The Publications Office of the European Union was founded in 1969 and is an interinstitutional office whose task is to publish the European Union’s institutions’ publications. |
| <b>Size of the company:</b>         | OP has around 500 employees. The EC has around 32,000 employees including OP and other DGs.   |
| <b>Geographic coverage:</b>         | European Union countries and all internet users in the world that are interested in the subject.  |

**FIGURE 7: SCREENSHOT OF PoC RDDA WEBSITE**

The visual platform used in the PoC was created using standard Web Technologies, such as HTML5, CSS3 and JavaScript.

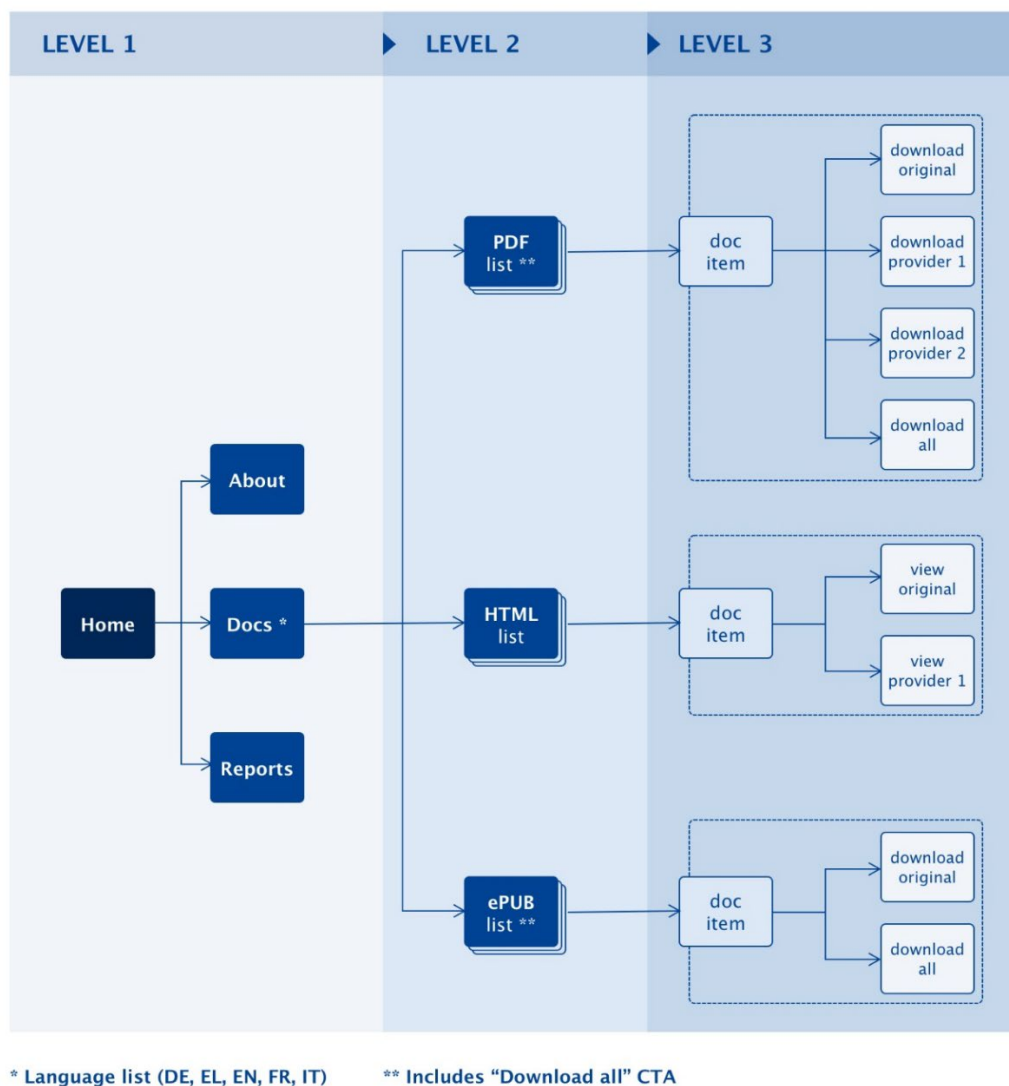
All pages were created manually with best practices in mind, using the same stylesheet to unify the experience and codebase, in an effort to reduce the loading times to a minimum. There is no need for a back-office, user admin or complex database due to the small amount of information to be displayed and overall purpose of the website.

Visually, fonts and colours are inspired by the EU official style guide (FIGURE 7). The UI is designed to favour legibility, contrast, and a clear hierarchical navigation structure. Each document type is clearly identified with an icon that graphically mimics the layout of its correspondent item.

The user flow is comprised of a simple navigation drill-down, where only the home page is not part of a “file system” type navigation (FIGURE 8).

The website “Level 1” is the home page and consists of displaying the main sections for the user journey. “Level 2” relates to the categories / direct content of each section (FIGURE 8 is focusing on the Documents section). The final “Level 3” enables access to the original and enhanced subsets.

FIGURE 8: PoC User Flow



The PoC RDDA Website centralises and enables the other solutions implementation:


- For AbleDocs and a solution under development
  - It provides access to the original PDF subset in a facilitated manner;
  - Integrates the returned remediated PDF subset into the navigation.

- For FACIL’iti
  - It provides URLs for connection with FACIL’iti platform;
  - Integrates the returned tags in the FACIL’iti web subset;
  - Presents link for user setup and a quick guide.
- For Siteimprove Accessibility
  - It provides URLs for scanning and connection with Siteimprove platform;
  - Hosts re-coded PoC Pages and one web page after returned report.

### 3.2.2 AbleDocs

The AbleDocs solution positions itself as a different approach to document accessibility, focused on making content always compliant, accessible, and usable. AbleDocs users / clients are mainly governments, multinational public companies, private companies, and suppliers for the blind. For the proof of concept, fully automated solutions from AbleDocs have been selected to remediate PDF subset.

**TABLE 4: ABLEDOCS**

|  |   |
|--|---|
| General description  |   |
| <b>Name of the solution:</b>   | AbleDocs  |
| <b>Description of the solution:</b>  | <p>Depending on the document type and remediation service, AbleDocs can apply the following solutions:</p> <ul style="list-style-type: none"> <li>• ADService is an on-demand remediation service that returns accessible and fully compliant documents.</li> <li>• ADGateway allows clients to transmit content directly into their proprietary production environment with embedded accessibility testing tools.</li> <li>• ADStream is a high-volume/high availability scaled solution for fully automated remediation of transactional documents, such as statements and standard types of content.</li> <li>• ADForms provides forms services and solutions supporting AcroForm, XFA-LiveCycle, AEM, and HTML type.</li> <li>• ADPublish is a custom workflow process to produce different formats from a single source, fully compliant.</li> <li>• ADScan is a website crawling tool that tests files under PDF/UA standard. It helps to understand which documents are compliant and helps to design a document accessibility strategy.</li> <li>• ADLegacy is for old content that remains on websites. It provides on-demand 24 hours turnaround to every PDF.</li> </ul> |
| <b>Price:</b>  | <p>The cost depends on the volume and complexity of the files and the strategy defined.</p> <p>The price can range from 0.004 cents per page to €5 per page including manual remediation. AbleDocs recommends assessing the operating environment, the client capabilities, and the file types to find the right solution for a cost sustainable document accessibility strategy.</p>   |

| <b>TABLE 4: ABLEDOCS</b>            |  |
|-------------------------------------|--|
| <b>Languages:</b>                   | Today, 38 languages are supported by the solutions. The gateway interface is in English, French, Danish, Spanish, or German. It is easy for AbleDocs to add new languages.   |
| <b>Alphabets:</b>                   | All alphabets can be supported, even Cyrillic, Farsi and Asian character derived languages.  |
| <b>Environment:</b>                 | ADGateway is accessible via any browser. ADStream can be installed on Windows, Linux, Unix, web based or on-premises.  |
| <b>Features:</b>                    | Everything produced by AbleDocs using automated and manual remediation is PDF/UA compliant. Some documents can be fully automated when the layout and the structure are similar (e.g., legal documents).   |
| <b>Standards and guidelines:</b>    | AbleDocs complies with PDF/UA, which does not conflict with WCAG2.1 AA. ADScan checks every component of PDF/UA as well as colour contrast (which is not part of PDF/UA) but will not modify the layout.   |
| <b>Implementation:</b>              | ADStream requires 2 weeks of on boarding to configure the files and processing environment to be accessed. Other AbleDocs solutions are implementable on a same day basis.   |
| <b>Demonstration:</b>               | <a href="https://www.youtube.com/watch?v=xQNCJPOQzEk">https://www.youtube.com/watch?v=xQNCJPOQzEk</a>  |
| <b>Users</b>                        |  |
| <b>Example of users:</b>            | AbleDocs has around 2,200 clients at every level of government in Canada, US, UK, Denmark, Switzerland, Germany, France, and Australia. Clients include, for example, governments, multinational public companies, private companies, single graphic designers, freelancers, or suppliers for the blind. |
| <b>Provider</b>                     |  |
| <b>Name of the provider:</b>        | AbleDocs   |
| <b>Description of the provider:</b> | AbleDocs has leading experts in PDF accessibility around the world. It was founded in 2019 with the merger of four accessibility companies.  |
| <b>Size of the company:</b>         | AbleDocs has around 130 employees.   |
| <b>Geographic coverage:</b>         | AbleDocs operates globally. It has four US offices and four EU offices.  |

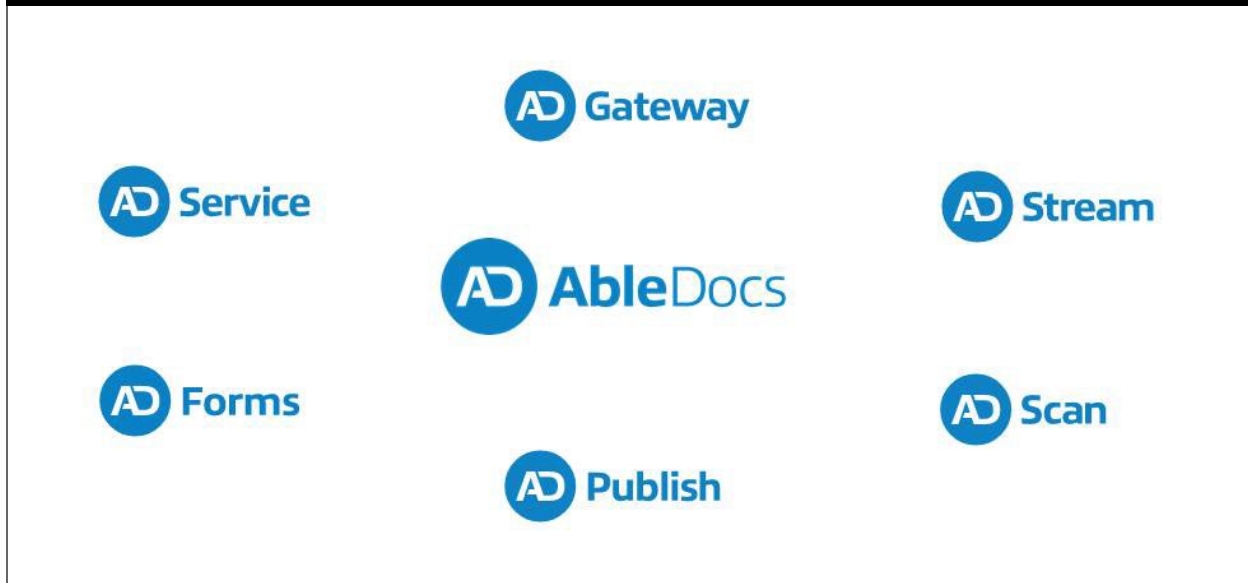
AbleDocs [accessibility management](#)<sup>12</sup> approaches the remediation solution as an ecosystem of documents - different types of organizations deal with different document types at different times. Document accessibility needs rangeability, and on that basis, AbleDocs offers 6 distinct solutions, as shown in FIGURE 9.

---

<sup>12</sup> PDF Association, "5 minutes with AbleDocs" (Electronic Document Conference 2019), <https://www.youtube.com/watch?v=xQNCJPOQzEk>, [accessed on 10.12.2020]



FIGURE 9: ABLEDOCS SOLUTION ECOSYSTEM



To perform the PoC and automated remediation of the PDF subset, AbleDocs applied the following solutions:

- AService
  - On-demand remediation service that returns accessible and fully compliant documents.
- ADStream
  - Built-in Engine to automatically remediate high volume documents with repeated design, not template content.

The AbleDocs remediation solution starts with the configuration of one sample file from the PDF subset. From that initial file setup, all of the PDF subset is automatically processed and remediated (tagging), regardless of document language (AbleDocs supports multi-language).

After the automated tagging, supported by the ADStream engine, the remediated subset goes through quality control review, both automatically and manually. For the "alternative-text" accessibility requirement AbleDocs remediation is always manual.

### 3.2.3 Accessibility Tagging solution under development

The purpose of the the solution which is under development is to automatically tag PDFs to lessen the amount of manual remediation needed on the path to PDF/UA compliance.

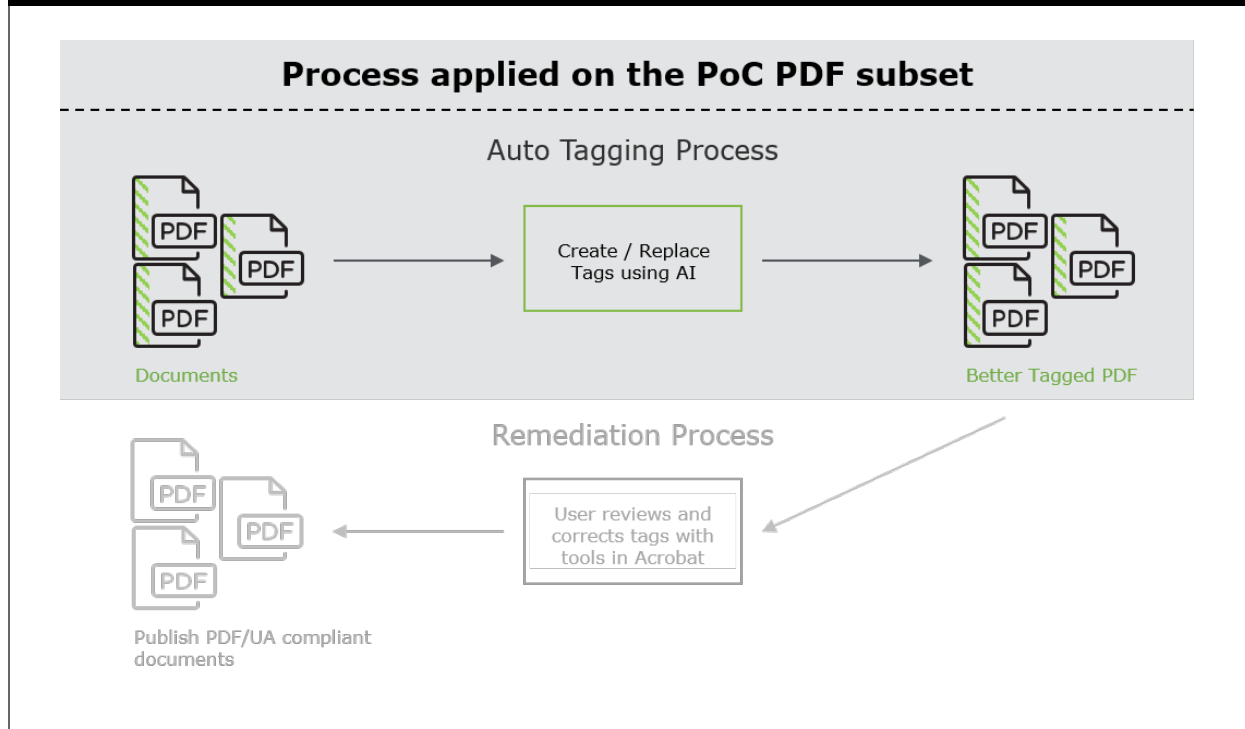
Using artificial intelligence and machine learning technology, new features are being developed and applied (pre-beta version) in automated accessibility tagging.

**TABLE 5: ACCESSIBILITY TAGGING SOLUTION UNDER DEVELOPMENT**

| N/A                                 |  |
|-------------------------------------|--|
| General description                 |  |
| <b>Name of the solution:</b>        | Not identified. The solution is currently under development  |
| <b>Description of the solution:</b> | Artificial intelligence (AI) and machine learning technology are being developed and applied (beta version) in automated accessibility tagging.  |
| <b>Price:</b>                       | Accessibility Tagging is under development and not yet available.  |
| <b>Languages:</b>                   | Under development for European languages and many additional languages.  |
| <b>Alphabets:</b>                   | All alphabets can be supported.  |
| <b>Environment:</b>                 | It is supported on desktop and mobile, using most of the operating systems.  |
| <b>Features:</b>                    | <p>There are tools that:</p> <ul style="list-style-type: none"> <li>• Use machine Learning / Artificial Intelligence to ensure better processes automations;</li> <li>• Have different functionalities to tag and fix those tags within the document. After the initial automated tagging, the remaining remediation work is done manually to allow users to adjust the tag;</li> <li>• Auto-generate the form fields to facilitate the creation of accessible forms;</li> <li>• Organise pages;</li> <li>• Run OCR and tag PDF</li> </ul> |
| <b>Standards and guidelines:</b>    | The solution is based on WCAG 2.1 A & AA and EN 301 549 standards.   |
| <b>Implementation:</b>              | The accessibility features are easy to implement. Specific training and knowledge is beneficial for understanding how to create accessible PDFs. Web seminars and help documentation are available online to learn.  |
| <b>Demonstration:</b>               | The developer can readily provide demonstrations to anyone, anywhere at any time. It also provides a two-day training course.  |
| Users                               |  |
| <b>Example of users:</b>            | Its clients are in the public sector, educational institutions, companies and also, in the independent remediation houses.   |
| Provider                            |  |
| <b>Name of the provider:</b>        | Not identified for this under development solution.  |
| <b>Description of the provider:</b> | N/A  |
| <b>Size of the company:</b>         | N/A  |
| <b>Geographic coverage:</b>         | N/A  |

The under development solution consists in an “auto tagging process” (via accessibility API) before a “remediation process” (involving user review and correction with tools such as Acrobat) as shown in FIGURE 10.

FIGURE 10: ACCESSIBILITY TAGGING SOLUTION UNDER DEVELOPMENT: END-TO-END PROCESS



As commonly recognized, AI is not yet perfect, revealing some shortcomings such as:

- Alternative text currently cannot be automatically generated for images;
- Not all document types (forms, for instance) can be converted yet;
- Complex tables with nested headers or non-traditional formats are hard to remediate:
  - AI does not recognize cell properties, so these must be created manually.

Auto-remediated PDFs must always be manually checked and remediated for full PDF/UA compliance.

However, using AI for text-only documents and less complex layout structures provides 100% compliant remediation.

On images with captions, AI cannot place a caption tag but instead creates and separates the caption with a paragraph tag (<p>), which makes it an easier task for screen readers and it's a step towards a better automated remediation.


To perform an automated remediation of the PDF subset, the solution applied the AI auto tagging process, and reduced the manual remediation left to be done.

Overall, we can see the improved results in using the AI service accessibility tagging.

### 3.2.4 FACIL'iti

FACIL'iti is a web service that improves the accessibility of a website or a web application. The FACIL'iti solution is the intermediary between the website of the client's company and the user which allows the latter to adapt the site's interface to its own needs. For the proof of concept, FACIL'iti is used to remediate HTML subset.

**TABLE 6: FACIL'ITI**

|  <b>FACIL'iti</b><br>VOTRE SITE + FACILE D'ACCÈS |   |
|---|---|
| General description   |   |
| <b>Name of the solution:</b>  | FACIL'iti   |
| <b>Description of the solution:</b>   | FACIL'iti is a web plugin that modifies the text (of client's websites) to the user's customised preferences in terms of text size, font selection, letter groups (highlights these), space between letters, words, and lines. This solution is not open source.  |
| <b>Price:</b>   | The price is based on a monthly subscription fee according to website traffic, ranging from €350 to €900. A discount is available where access to several websites is required.   |
| <b>Languages:</b>   | FACIL'iti supports content in many languages, notably Cyrillic and Greek. However, the FACIL'iti interface is available in English, Spanish, French, German, Italian, Japanese, Russian, Dutch, Korean, Chinese, and Greek.   |
| <b>Alphabets</b>  | Latin, Cyrillic, Greek and Kana alphabets are supported by the tool. The interface is available in Latin, Cyrillic and Kana alphabets.  |
| <b>Environment:</b>   | The solution is implemented on the server side (license).   |
| <b>Features:</b>  | <ul style="list-style-type: none"> <li>• FACIL'iti does not work like a traditional zoom because it can result in obstructed text and field; instead, the changes occur on the CSS file layer.</li> <li>• FACIL'iti uses open source fonts that do not require a license.</li> <li>• Text-to-speech is not part of FACIL'iti's features. However, the tool allows access to the alternative text for OCR tools.</li> <li>• FACIL'iti works by leveraging the use of TAGS. FACIL'iti covers everything in the HTML file.</li> <li>• Flash content will not work with FACIL'iti.</li> </ul> |
| <b>Standards and guidelines:</b>  | FACIL'iti focuses on the user side of WACG guidelines, providing a service that addresses several disabilities, in particular, visual (such as colour blindness), movement (such as Parkinson's disease) and cognitive (such as dyslexia) disabilities. It focuses on user's needs and not on regulations in place. Therefore, the use of FACIL'iti does not make a website fully compliant with regulations. FACIL'iti has co-developed its solution with relevant associations for each pathology and worked with thousands of users, testing, and refining their product.              |
| <b>Implementation:</b>  | FACIL'iti's integration level depends on the website's accessibility/WCAG compliance. FACIL'iti does not provide accessibility services for improving a website's code. FACIL'iti is easy and fast to setup; the process takes about 10 minutes.  |
| <b>Demonstration:</b>   | <a href="https://www.facil-iti.com/">https://www.facil-iti.com/</a><br><a href="https://youtu.be/BrbhjQPEfAw">https://youtu.be/BrbhjQPEfAw</a>  |

**TABLE 6: FACIL'ITI**

| Users                               |   |
|-------------------------------------|---|
| <b>Example of users:</b>            | Hundreds of entities use FACIL'iti including New York City, Tokyo, Kyoto and Yokohama, as well as MPOD, Paris Tourist Office, TF1 and Total. The number of FACIL'iti users is close to 500,000.   |
| Provider                            |   |
| <b>Name of the provider:</b>        | FACIL'iti   |
| <b>Description of the provider:</b> | FACIL'iti has been on the market since 2015 with ongoing development. The current roadmap for FACIL'iti development focuses on delivering services aimed at different disabilities, including Alzheimer's disease. According to FACIL'iti, over 90% of users interviewed find the tool to be useful to some extent. |
| <b>Size of the company:</b>         | FACIL'iti has 16 employees in France and 5 in Japan.  |
| <b>Geographic coverage:</b>         | France, Japan, United States, Canada, amongst others.   |

The FACIL'iti solution includes:

- Connection to the website of the client's company via a small computer code (tag);
- The possibility to choose the display settings to adapt the website (to magnify the texts, darken the backgrounds, highlight certain words, etc) via a configuration interface.

Many different entities use FACIL'iti, in order to enhance end-users experience (with different disabilities) on their website.

The automated remediation process includes:

- Website connection to FACIL'iti;
- Implement FACIL'iti tag on desired pages.

The FACIL'iti solution is cookie based and needs cookies to work. These technical cookies are never used to collect the user browsing data and are never passed on to third parties.

The data used by the FACIL'iti solution, as well as the security and commitments made in order to implement its services, are detailed in FACIL'iti [personal data and security policy](#)<sup>13</sup>.

---

<sup>13</sup> FACIL'iti "personal data ad security policy", <https://ws.facil-iti.com/data-protection-and-security.html> [accessed on 12/12/2020]

Regarding the PoC RDDA Website, the focus of the FACIL'iti solution is dyslexia. The FACIL'iti tag was only implemented on the web HTML subset. After the solution implementation, the user journey for a non-registered profile in FACIL'iti is:

1. The user arrives to a "select language" list of links to original and remediated (tagged) web subset HTML pages;
2. The user selects a remediated version and language;
3. An overlay appears with an explanation about FACIL'iti, the user notices a link to "Setup a FACIL'iti profile" and clicks on it (**FIGURE 11**);
4. The user accesses the platform and creates a profile with preferences (**FIGURE 12**), to then automatically return to the previous / original web page;
5. The user accesses and views the selected remediated and language version of the web subset.

**FIGURE 11: PoC RDDA WEBSITE FACIL'ITI ACCESS**

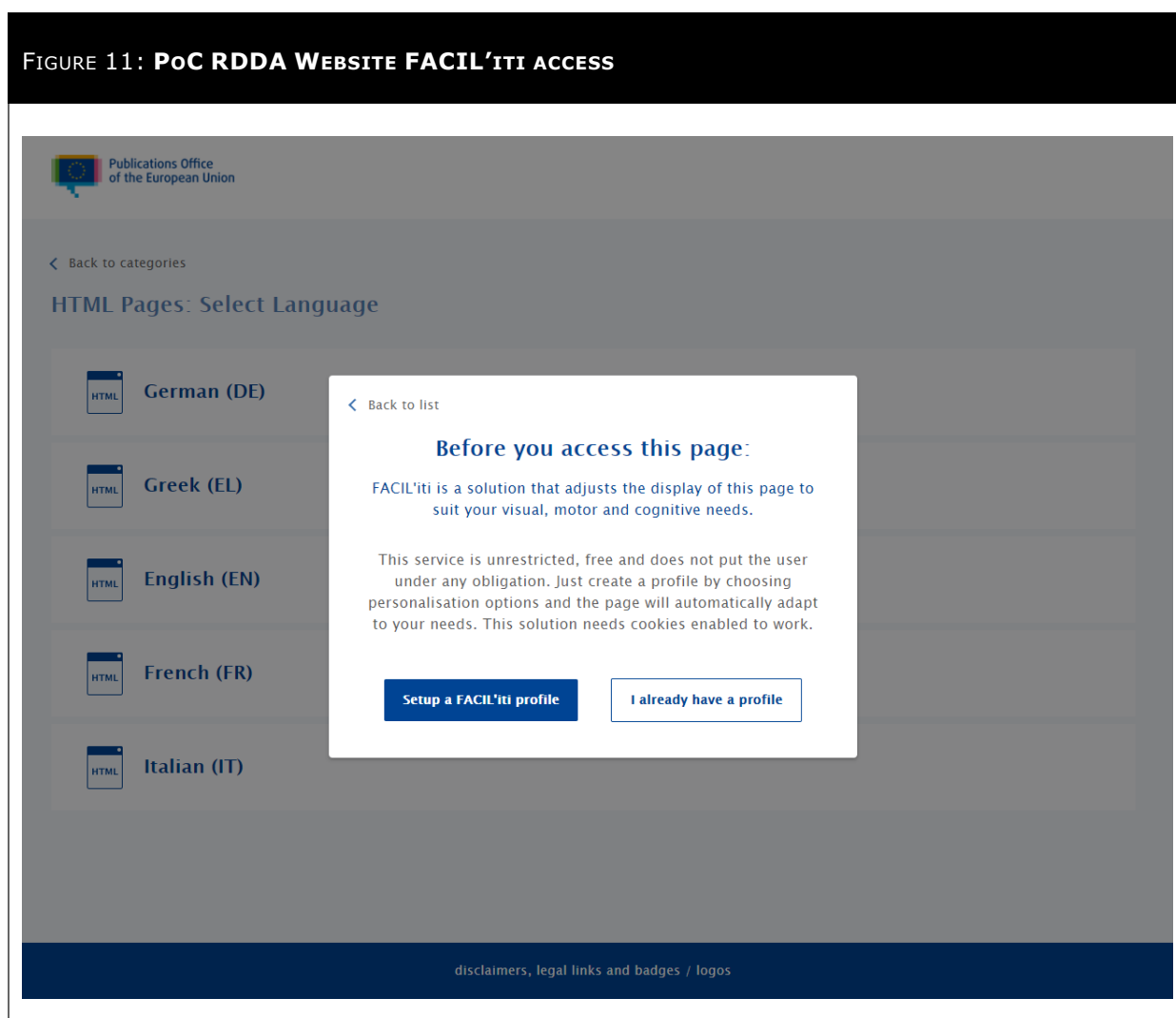


FIGURE 13 shows how to create a personalized dyslexia filter in FACIL'iti.

And, the user journey for a previously registered profile in FACIL'iti is:

1. The user gets access to a "select language" list of links to original and remediated (tagged) web subset HTML pages;
2. The user selects a remediated version and language;

3. An overlay appears with an explanation about FACIL'iti, the user notices a "I already have a profile" and clicks on it;
4. The user accesses and views the selected remediated and language version of the web subset.

FIGURE 12: CREATE A PROFILE IN FACIL'ITI PLATFORM

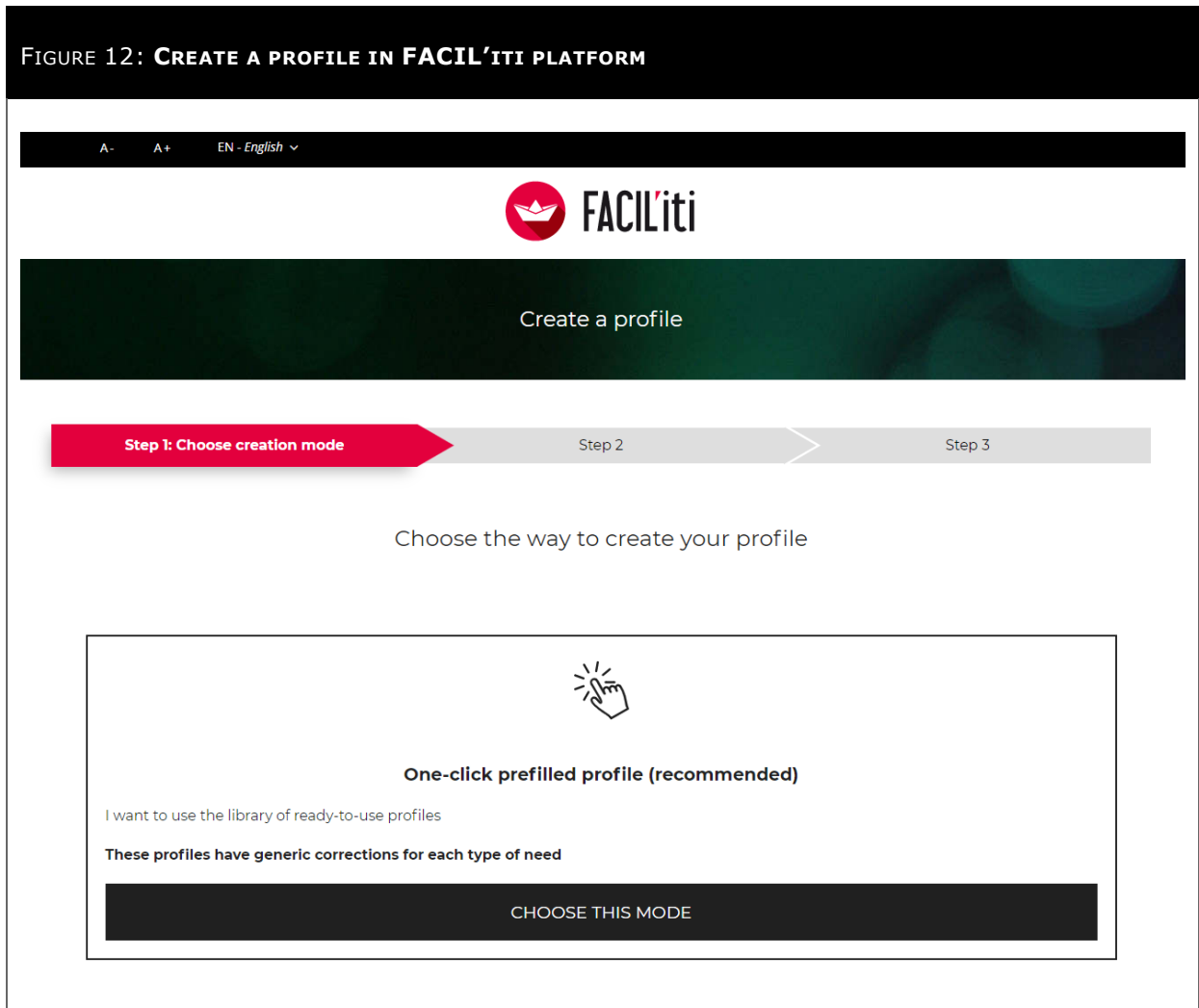


FIGURE 13: HOW TO CREATE A FACIL'ITI PROFILE

The image is a screenshot of the FACIL'iti website's profile creation process, enclosed in a red dashed border. At the top center is the FACIL'iti logo, a red circle with a white book icon, followed by the text "FACIL'iti". Below the logo is the heading "HOW TO CREATE A PERSONALIZED DYSLEXIA FILTER ?".

**STEP 1**  
Click on this button :  
A red button with the text "CREATE A PERSONALISED PROFILE" and a mouse cursor icon pointing to it.

**STEP 2**  
Choose « Personalized profile »  
A screenshot of a web interface showing a "Personalized profile" section with a gear icon. Below it, there is a text input field with "CHOOSE THIS MODE" and a mouse cursor icon pointing to it.

**STEP 3**  
Enter a profil name and validate :  
A form with a text input field containing "Dyslexia". Below the field, there is a checkbox labeled "I WANT TO PROTECT MY PROFILE WITH A PASSWORD (OPTIONAL)". At the bottom of the form is a black button with the text "VALIDATE YOUR PROFILE" and a mouse cursor icon pointing to it.

**STEP 4**  
Scroll down to choose « Identify Letters » and click on « see options ».  
(you can also choose « space letters and words » option below)  
A screenshot of a web interface showing a section titled "Identify letters, groups of letters or words (recommended for dyslexia)". There is a mouse cursor icon pointing to a black button with the text "SEE OPTIONS".

**STEP 5**  
Enter the letters that are confusing and activate :  
A form with a text input field containing "dubmrl". Below the field, there is a black button with the text "ACTIVATE" and a mouse cursor icon pointing to it.

At the bottom right of the screenshot, there is a red thumbs-up icon and the text "That's ok ! You can now return on the website".

### 3.2.5 Siteimprove Accessibility

The purpose of the Siteimprove solution is to analyse and report web pages and media accessibility compliance, detailing potential issues and fixes. For that, most Siteimprove users have roles that include web management or editorial responsibilities. They tend to be part of web, communications, or marketing teams. For the proof of concept, Siteimprove accessibility is used to evaluate the accessibility level of HTML pages before and after remediation.



**TABLE 7: SITEIMPROVE ACCESSIBILITY**

**Siteimprove**

**General description**

|                                     |  |
|-------------------------------------|--|
| <b>Name of the solution:</b>        | Siteimprove accessibility  |
| <b>Description of the solution:</b> | <p>Siteimprove provides users with an easy-to-use platform for website optimisation. One part of the broader platform is Siteimprove Accessibility, which helps users discover and prioritise issues. It takes control of WCAG 2.1 barriers which prevent people from accessing content on the web and provides users with guidance and practical recommendations on how to address issues.</p> <p>In addition to the platform, Siteimprove also provides manual testing services and educational resources that help reduce compliance risks.</p> <p>Besides accessibility, the platform offers content quality, SEO, Google Ads, Analytics, website performance and Data Privacy solutions.</p> <p>The solution is not open source, but Siteimprove is in the process of moving to a new accessibility-checking engine (Alfa) which will be partially open source:<br/> <a href="https://github.com/Siteimprove/alfa">https://github.com/Siteimprove/alfa</a>.</p> |
| <b>Price:</b>                       | <p>Pricing is available on request and depends on the number of modules of the platform that users want to purchase. Siteimprove Accessibility is licensed by number of pages and includes unlimited users and simple on-boarding training for small teams.</p>  |
| <b>Languages:</b>                   | <p>The Siteimprove platform supports the following languages: English, Swedish, German, Norwegian Bokmål, French, Spanish, Italian, Dutch, Portuguese, Finnish and Japanese. It does not have any known language limitations regarding accessibility testing. Any language dependent checks are verified by the users.</p>   |
| <b>Alphabets:</b>                   | <p>The Siteimprove platform supports Latin, Greek and Cyrillic alphabets, but is not able to support symbol alphabets.</p>   |
| <b>Environment:</b>                 | <p>Siteimprove’s Accessibility solution is cloud-based. It scans and tests public websites, intranets, and PDFs.</p>   |
| <b>Features:</b>                    | <p>Features include:</p> <ul style="list-style-type: none"> <li>• WCAG 2.1 accessibility testing of websites and PDFs.</li> <li>• Easy task coordination and resource allocation (difficulty ratings).</li> <li>• Guiding individuals of all skill sets to carry out accessibility testing (potential issues, guided testing).</li> <li>• Supporting users in understanding what matters most to their visitors (analytics/issue coverage metrics).</li> <li>• Ensuring maximum remediation impact with the least effort (page sections, focus on most exposed/at risk content).</li> <li>• Remediation support (pass and fail examples for failing content).</li> <li>• Peer-reviewed, fully documented and transparent checks (W3C ACT accessibility checks).</li> <li>• Harmonisation of accessibility testing standards (ACT format).</li> </ul>   |

**TABLE 7: SITEIMPROVE ACCESSIBILITY**

|                                     |   |
|-------------------------------------|---|
| <b>Standards and guidelines:</b>    | Siteimprove follows WCAG 2.1 and Accessibility Conformance Testing (ACT) Rules Format for both websites and PDFs.   |
| <b>Implementation:</b>              | <p>Siteimprove Accessibility is a hosted plug and play service. The Siteimprove platform automatically crawls user websites once an account has been created. The solution then breaks down accessibility issues that are discovered as part of the crawl into prioritised, manageable tasks, providing guidance and practical recommendations along the way. This makes it straightforward even for non-experts to start working towards making their websites accessible.</p> <p>Additionally, Siteimprove Client Success Managers will support new users through the implementation and on-boarding phase.</p>   |
| <b>Demonstration:</b>               | <a href="https://www.youtube.com/watch?v=LeRAK3xc51Y">https://www.youtube.com/watch?v=LeRAK3xc51Y</a> <sup>14</sup>   |
| <b>Users</b>                        |   |
| <b>Example of users:</b>            | Siteimprove has approximately 70,000 users in both the public and private sectors. Many users have roles with web management or editorial responsibilities. They tend to be part of web, communications, or marketing teams. Siteimprove has a very strong foundation in the public sector with a lot of government organisations and educational institutions amongst its clients.   |
| <b>Provider</b>                     |   |
| <b>Name of the provider:</b>        | Siteimprove   |
| <b>Description of the provider:</b> | <p>Siteimprove was founded in 2003. It offers the world’s most comprehensive digital optimisation platform, providing its clients with actionable insights to improve their visitors’ digital experience. It supports organisations in driving digital accessibility, improving content quality, understanding and improving their visitors’ journey and optimising their digital marketing activities. Siteimprove empowers its clients to focus on initiatives that drive the biggest impact on their visitors’ digital experience.</p> <p>Siteimprove’s solution has been developed with accessibility in mind and most of all features are accessible. Siteimprove are continuously optimising what they do and how they do it and are committed to creating intuitive, barrier-free products that everybody can use, regardless of their disabilities.</p> <p>Siteimprove partners with the European Commission, the W3C and the European Disability Forum to improve digital accessibility.</p> <p>Siteimprove also provides manual testing services and broader educational content on accessibility such as the <a href="#">“All-in-One</a></p> |

---

<sup>14</sup> Siteimprove, “Siteimprove DCI™ Score: Accessibility, SEO, and Quality Assurance” <https://www.youtube.com/watch?v=LeRAK3xc51Y>, [accessed on 04.01.2021]

**TABLE 7: SITEIMPROVE ACCESSIBILITY**

|                             |   |
|-----------------------------|---|
|                             | <a href="#">Accessibility Handbook</a> <sup>15</sup> and digital academy courses to help train teams on accessibility and other subject matters. It also provides the free Siteimprove Chrome Extension which is used to test a single instance (page). |
| <b>Size of the company:</b> | Siteimprove has 575 employees.  |
| <b>Geographic coverage:</b> | Siteimprove covers North America, Europe, Australia/New Zealand and Japan.  |

Siteimprove is a hosted plug and play service automatically crawls user websites once an account has been created and URLs are provided. The solution breaks down accessibility issues that are discovered as part of the crawl into prioritised and manageable tasks, providing guidance and practical recommendations along the way.

Siteimprove built-in engine uses a crawler technology that scans the URL and collects the HTML and CSS. After collecting this information, it utilises a service ([alfa](#)<sup>16</sup> checker engine) to read through the HTML and CSS to provide the results. The scan also includes accessibility testing on PDF documents available on the website.

These results are then provided back into the platform and presented in a user-friendly approach (FIGURE 14), by:

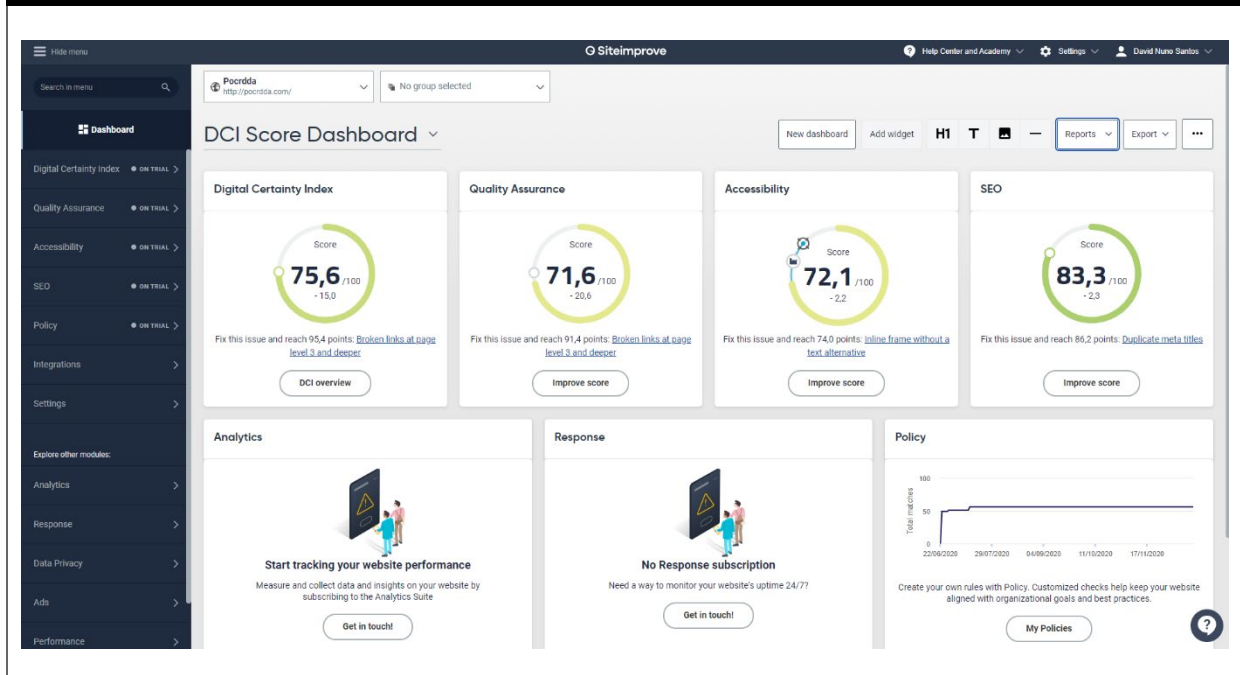
- Recreating the HTML and CSS;
- Pinpointing / highlighting the problems;
- Utilising and crossing with the common acceptance format of ACT (Accessibility Conformance Testing);
- Providing them back again, in an easier and digestible format inside the tool.

---

<sup>15</sup> Siteimprove, "The All-in-One Digital Accessibility E-Book", <https://hello.siteimprove.com/en/ebook/The-All-in-One-Digital-Accessibility-eBook/download>, [accessed on 04.01.2021]

<sup>16</sup> Alfa Engine Checker, <https://github.com/Siteimprove/alfa> [accessed on 06.01.2021]

FIGURE 14: SITEIMPROVE PLATFORM



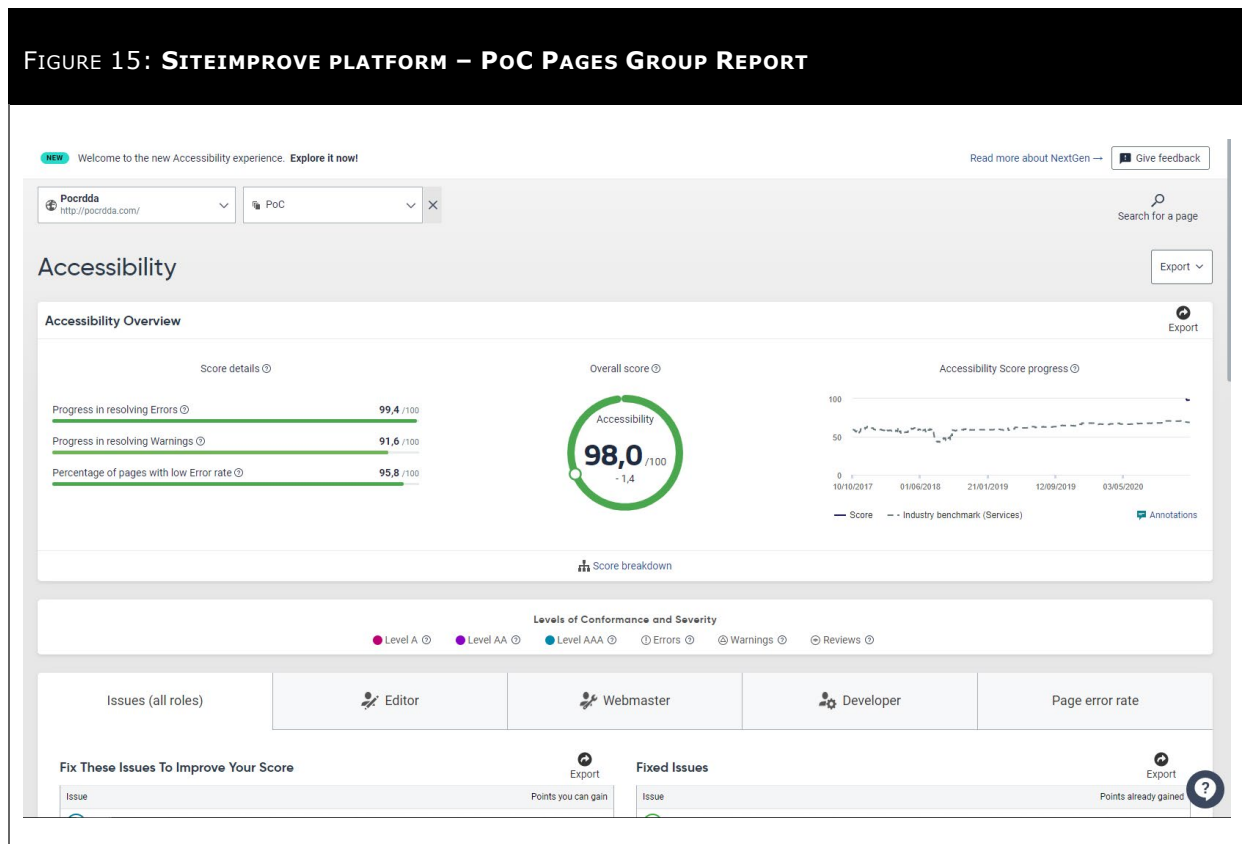
The overall page score presented by the Siteimprove solution is called the "Digital Certainty Index (DCI)" which is a measure of a page's performance regarding the following main indicators:

- **Quality Assurance:** How credible and usable the user-facing aspects of a site are?
  - The overall QA Score is measured by a site's content quality, content freshness, security, and the user experience.
- **Accessibility:** How easy it is to navigate and read a page regarding best practices common to user disabilities? This indicator considers the following aspects:
  - Text alternatives: Detection of existent text alternatives for non-text content.
  - Adaptability: Usage of html styling when CSS should be used; Presence of non-unique element IDs; Properly applied HTML5 and WAI-ARIA; Landmarks (i.e. <nav> or <section> html tags).
  - Distinguishable: Use of colour to clearly identify elements like hyperlinks; Colour contrast; Images with text in them.
  - Timing: Measure of the time a page takes to load, hypothetical time-limits for a user to read the content.
  - Navigable: Detection of an option to skip repeated content; Well-structured content with section and structure headings.
  - Compatible: Detect correct element parsing (i.e. presence of unique IDs), name, role, value (missing required html attributes, WAI-ARIA roles correspond to the intended function of an element.).
- **SEO:** How well the user-facing and technical aspects of your site contribute to search engine optimization and, ultimately, to higher rankings and organic traffic?

In the proof of concept, only the accessibility indicator is used.

After the Siteimprove automatic evaluation of the PoC website and web subset, it's possible to distinguish groups of pages (e.g.: PoC web pages VS original/enhanced subsets) with their respective reports and accessibility score, as shown in **FIGURE 15**.

**FIGURE 15: SITEIMPROVE PLATFORM – PoC PAGES GROUP REPORT**



Some of the Accessibility indicator features include:

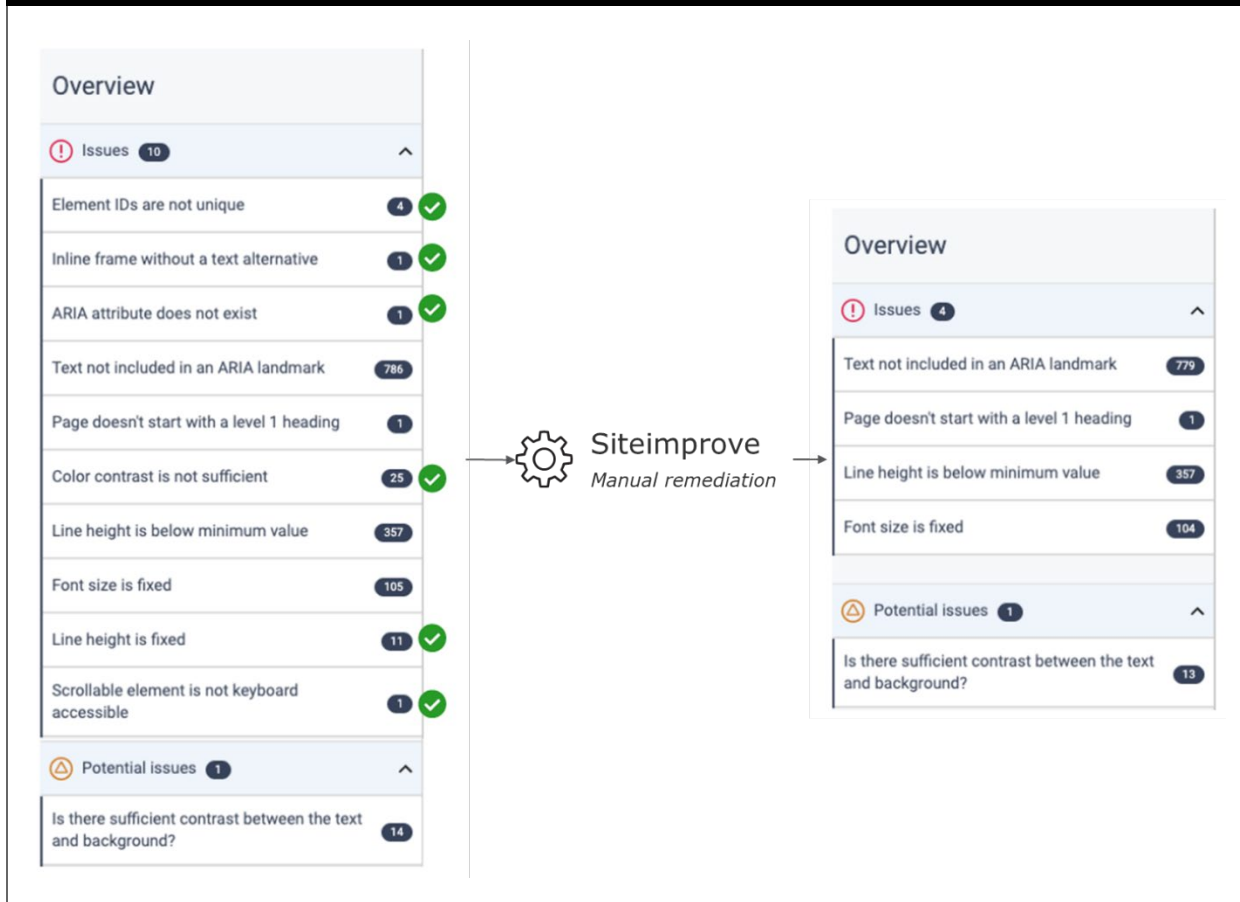
- **Multi-site management**, allowing more than one website / tailored domain for assessment;
- **Site page grouping**, making it possible to create groups of pages with different “match types”;
- **Site target compliance**, which enables setting the compliance level target (A, AA or AAA) and provides a score goal;
- **Accessibility score**, which provides a score by page groups VS the entire website and enables separated reports;
- **Issues filtering**, allowing to automatically filter issues according to the site’s target and manage a workflow of displaying and analysing issues, dividing them in potential and solved issues;
- **Reports**, allowing sending, scheduling and managing reports;
- **Export** of dashboards, graphics, reports, and issues’ lists.

The most common issues based on the original HTML subset are listed in the table below.

| <b>TABLE 8: MOST COMMON ISSUES BASED ON THE ORIGINAL HTML SUBSET</b> |                                 |  |
|--|---------------------------------|--|
| <b>Category of issue</b>   | <b>Type of issue + Severity</b> | <b>Description</b>   |
| <b>Content writing</b>   | Issue - error                   | Missing or repeated text alternatives  |
| <b>Content writing</b>   | Issue - error                   | Form fields with missing labels  |
| <b>Content writing</b>   | Issue - error                   | Empty html tags  |
| <b>Content writing</b>   | Issue - error                   | Missing page titles  |
| <b>Content writing</b>   | Issue - error                   | Similar text inside different links  |
| <b>Content writing</b>   | Potential Issue - warning       | Different links in the same context suspected to:<br>- Be duplicate because of very similar text                       |
| <b>Content writing</b>   | Potential Issue - warning       | Different links in the same context suspected to:<br>- Have the same target URL because of similar text and appearance |
| <b>Development</b>   | Issue - error                   | Empty elements   |
| <b>Development</b>   | Issue - error                   | Page language not defined  |
| <b>Development</b>   | Issue - error                   | Lack of keyboard accessibility (ex.: horizontally scrollable elements)   |
| <b>Development</b>   | Issue - error                   | ARIA roles and tags not setup properly   |
| <b>Development</b>   | Issue - error                   | Clipped/hidden text on browser (user) zoom   |
| <b>Development</b>   | Issue - error                   | Fixed font-size (should be relative to the user's browser setting, usually at 16px)                                    |
| <b>Development</b>   | Issue - error                   | Fixed line-height (should be relative to the user's browser setting, usually at 1.5 minimum)                           |
| <b>Visual Design</b>   | Issue - error                   | Insufficient colour contrast   |
| <b>Visual Design</b>   | Issue - error                   | Line-height below minimum value (not enough spacing between lines)   |
| <b>Visual Design</b>   | Issue - error                   | Text with uneven spacing (relative to line-height values and line-breaks)  |
| <b>Visual Design</b>   | Potential Issue - warning       | Possible insufficient contrast between text and background   |

As an example of Siteimprove tools usage, some manual remediations are applied for a single web page named "The EU in 12 lessons" in English. The figure below shows the issues and potential issues, as it is shown on Siteimprove, before and after some manual remediations.

FIGURE 16: ISSUES AND POTENTIAL ISSUES BEFORE AND AFTER SOME MANUAL REMEDIATIONS USING SITEIMPROVE



The table below shows some of the remediations that have been tested based on Siteimprove recommendations.

| TABLE 9: HTML REMEDIATIONS TEST |                          |   |   |                          |
|---------------------------------|--------------------------|---|---|--------------------------|
| Category of issue               | Type of issue + Severity | Description                             | Remediation   | Status after remediation |
| Development                     | Issue - error            | Element IDs are not unique              | Searched and fixed duplicate IDs  | Fixed                    |
| Development                     | Issue - error            | Inline frame without a text alternative | Added an "aria-label" attribute   | Fixed                    |
| Development                     | Issue - error            | ARIA attribute does not exist           | Fixed aria attribute:<br><br>From: "<nav id="main-nav" aria-role="navigation" class="navbar-fixed-top">"<br><br>To: "<nav id="main-nav" role="navigation" aria-label="navigation" | Fixed                    |

**TABLE 9: HTML REMEDIATIONS TEST**

|                        |                           |  |  |   |
|------------------------|---------------------------|--|--|---|
|                        |                           |  | class="navbar-fixed-top">  |   |
| <b>Content writing</b> | Issue - error             | Text not included in an ARIA landmark                      | This issue needs a deeper understanding of the text to ensure correct usage of aria landmarks ("region" or "role" attributes for content regions or html elements purpose, respectively) | Not fixed - Further review needed from the content team   |
| <b>Development</b>     | Issue - error             | Page does not start with a level 1 heading                 | Added a hidden h1 heading with the page title  | Not fixed - error persists (probably because it is hidden via CSS)  |
| <b>Visual Design</b>   | Issue - error             | Colour contrast is not sufficient                          | Darkened colours using WebAIM contrast checker ( <a href="https://webaim.org/resources/contrastchecker/">https://webaim.org/resources/contrastchecker/</a> )                             | Fixed   |
| <b>Visual Design</b>   | Issue - error             | Line height is below minimum value                         | Added 1.5 value for line-height, where possible.<br><br>Some line-height values were used for visual purposes (non-text elements) and should be further considered/reviewed.             | Not fixed: Further review needed from the design team   |
| <b>Visual Design</b>   | Issue - error             | Font size is fixed   | This depends greatly on the design + the default font-size configured on the users' browser. Further review is required.   | Not fixed: Further review needed from the design team   |
| <b>Visual Design</b>   | Issue - error             | Line height is fixed                                       | Fixed using a relative line-height, where possible.  | Fixed   |
| <b>Development</b>     | Issue - error             | Scrollable element is not keyboard accessible              | Fixed adding tabindex="0" to the iframe html tag.  | Fixed   |
| <b>Visual Design</b>   | Potential Issue - warning | Possible insufficient contrast between text and background | Needs review from the design team  | Partially fixed: Further review needed from the design team, but in this particular case this potential issue can be dismissed / marked as fixed. |

### 3.2.6 PAC2021

The freeware program PAC provides a fast way to test the accessibility of PDF files. PAC is suitable for experts and end users conducting accessibility evaluations. For the proof of concept, PAC2021 is used to evaluate the accessibility level of PDF pages before and after remediation.



**TABLE 10: PAC2021**

# PDF/UA FOUNDATION

## General description

|                                     |  |
|-------------------------------------|--|
| <b>Name of the solution:</b>        | PAC (PDF Accessibility Checker) 2021   |
| <b>Description of the solution:</b> | The PDF Accessibility Checker (PAC) 2021 is a freeware provided by the not-for-profit "PDF/UA Foundation" for PDF accessibility testing.   |
| <b>Price:</b>                       | Free   |
| <b>Languages:</b>                   | Currently, the PAC2021 software is displayed in three languages (English, French and German).  |
| <b>Alphabets:</b>                   |  |
| <b>Environment:</b>                 | PAC2021 is supported on desktop using Windows operating systems.   |
| <b>Features:</b>                    | <p>PAC2021 features include:</p> <ul style="list-style-type: none"> <li>• a PDF/UA-check with all machine-testable PDF/UA compliance checkpoints.</li> <li>• a WCAG 2.1 AA check with all machine-testable WCAG 2.1 AA compliance checkpoints.</li> <li>• a screen reader preview that visually evaluates the logical document structure.</li> <li>• a view of the tag tree with all properties and attributes statistics concerning the document structure</li> </ul> |
| <b>Implementation:</b>              | The users must download and execute the program on their computer. The application can be run locally and does not require an installer.   |
| <b>Demonstration:</b>               | <a href="https://www.youtube.com/watch?v=qgZEi99xBKo">https://www.youtube.com/watch?v=qgZEi99xBKo</a> (in German)  |
| <b>Users</b>                        |  |
| <b>Example of users:</b>            | PAC2021 can be used by the general public, as well as experts conducting accessibility assessments.  |
| <b>Provider</b>                     |  |
| <b>Name of the provider:</b>        | PDF/UA Foundation  |
| <b>Description of the provider:</b> | <p>Founded in 2020 to promote digital accessibility for people with print disabilities, the PDF/UA Foundation is today the centre of competence for PDF accessibility around the world.</p> <p>"PDF/UA Foundation" sees itself as a mediator between the user group of people with disabilities and information providers from the public and private sectors.</p>   |
| <b>Size of the company:</b>         | "PDF/UA Foundation" is a relatively small company with around 10 volunteer members.  |
| <b>Geographic coverage:</b>         | Global   |

The PAC2021 (PDF Accessibility Checker, version 4, by the PDF/UA Foundation is a free software, an automatic evaluation tool for PDF accessibility testing which supports and facilitates the development of high-quality accessible PDF documents and forms with the following features:

- PDF user accessibility test with all machine-testable checkpoints of the Matterhorn protocol<sup>17</sup>;
- PAC 2021 now includes all relevant WCAG 2.1 AA checks that apply to PDF;
- Screen reader preview that visually evaluates the logical document structure;
- View of the Tag tree with all the properties and attributes;
- Statistics concerning the document structure.

With PAC2021 ISO 14289-1 (PDF/UA) validator it is possible to evaluate the PDF/UA compliance of accessible PDFs and it is suitable for experts and end users conducting accessibility evaluations. It requires installation by users on their computer, with the following dependencies:

- Windows XP, Vista and Windows 7 or later;
- Adobe Reader Version 8 or later;
- Mozilla Firefox 3 or later, Internet Explorer 6 or later, Google Chrome;
- Microsoft .NET Framework 2.0 SP" or later.

There are several accessibility assessment checks (Matterhorn protocol) implied in the tool workflow, such as:

1. If the document is marked as containing tags;
2. If a title has been set for the PDF document;
3. If a language has been assigned to the PDF document;
4. If the security settings of the PDF document allow assistive technology to access the tag structure;
5. If all pages of the PDF document are configured so that when using the Tab key, the performed jumps follow the tag structure;
6. If the document has a consistent heading structure;
7. If the document contains bookmarks;
8. If all font characters within tagged text blocks can be converted into distinct Unicode characters;
9. If all contents of the PDF document are tagged;
10. If the document has a logical reading order;
11. If all tagged non-text elements contain an alternative text;
12. If all tags and rolls within the PDF document correspond to the ISO standard and if they have been used correctly;
13. If the visual presentation of tagged text has a sufficient contrast ratio according to the WCAG 2.0 standard;
14. If the visual spaces in a PDF document are also contained within the content and the tag tree.

---

<sup>17</sup> Matterhorn Protocol - a list of all the possible ways to fail PDF/UA, as referenced by PDF Association - <https://www.pdfa.org/resource/the-matterhorn-protocol-1-02/>, [accessed on 18.12.2020]

### 3.3. Key findings

The main findings gathered throughout the solutions implementation processes transcribe to the following conclusions:

#### ***KF1) Tags are the foundation of accessible PDF documents***

Tags are the foundation of accessible PDF documents, delivering the structural information, which enables assistive technology, such as screen readers, to:

- Properly recognize different text elements, such as headings, lists, graphics and tables;
- Read the document following the correct reading order.

#### ***KF2) Currently, the "automated only" remediation processes still do not allow a full PDF/UA compliance***

Many authoring systems can produce tagged PDF documents. Unfortunately, however, the quality of those automatically generated tagged documents is often insufficient, which makes additional visual as well as technical checks inevitable. AbleDocs and a solution under development are focusing efforts in improving their capability to produce automatically tagged PDF documents. Currently, the "automated only" remediation processes still do not allow a full PDF/UA compliance.

#### ***KF3) User experience vs compliance with standard***

Some remediation solutions, such as FACIL'iti, focus on the user experience and do not aim to be fully compliant with standards.

#### ***KF4) The evaluation solutions are based on standards***

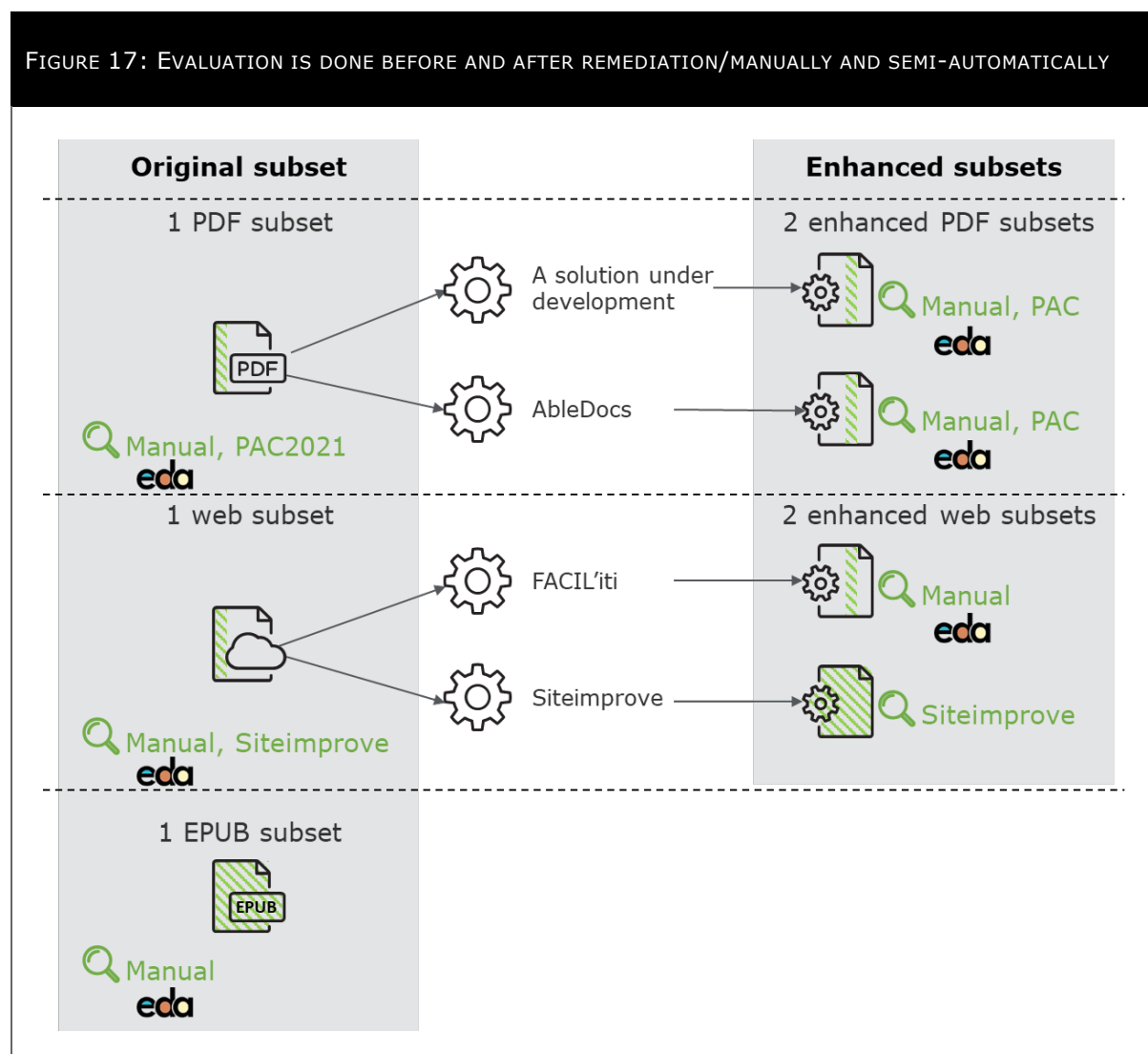
The evaluation solutions are based on standards, such as WCAG for web evaluation with Siteimprove and PDF/UA for PDF evaluation with PAC2021.

## 4. Section III: Evaluation and analysis

This section Evaluation and analysis consists in the presentation of the evaluation, the assessment and the analysis of the results before and after the implementation of the solutions.

| TABLE 11: EVALUATION VS ASSESSMENT VS REMEDIATION |   |
|---|---|
| Evaluation  | The evaluation is done manually by EDA  |
| Assessment  | The assessment is done semi-automatically using Siteimprove for HTML pages et PAC2021 for PDF documents.                              |
| Remediation                                       | The remediation is done automatically by AbleDocs and a solution under development for PDF documents and by FACIL'iti for HTML pages. |

As shown on **FIGURE 17**, the assessment and the evaluation is done before and after remediation. Also the evaluation is done manually by EDA and the assessment is done semi-automatically using Siteimprove for HTML pages and PAC2021 for PDF documents (TABLE 11). For the Proof of Concept, AbleDocs and a solution under development have been asked to remediate documents using only automated solutions in order to evaluate the possibility to implement such remediation to a large scale of documents.



Section III of the report is structured as follows:

- **Scope** of the documents to assess and evaluate, as only a part of the original and remediated subsets is assessed;
- **Methodology** followed to collect, prepare and process the evaluated level of accessibility from EDA or using Siteimprove or PAC2021;
- **Analysis** of the results following different axes such as:
  - Language comparison;
  - Type comparison;
  - Gain in accessibility.

#### 4.1. Scope

This section presents the scope of documents that have been evaluated by EDA and using Siteimprove and PAC2021. As a reminder, the evaluation subsets are reduced compared to the original and remediated subsets.

FIGURE 18 shows the documents of the original subset and the documents for which the level of accessibility is assessed.

- The original subset includes **two EPUB documents** named "The ABC of EU law" and "The EU in 2019" in the five languages (English, French, German, Greek and Italian).
  - *The accessibility level of these two documents is evaluated by EDA.*
- The original subset also includes **five HTML documents** named "Europe in 12 lessons", "The ABC of EU law", "The European Union What it is and what it does", "Travelling in Europe 2020" and "Who we are" in the five languages.
  - *Only the accessibility level of "The European Union What it is and what it does" and "Who we are" is evaluated by EDA and assessed using Siteimprove.*
- The original subset also includes **six PDF documents** named "EU & me", "Europe in 12 lessons", "Let's explore Europe!", "The ABC of EU law", "The European Union What it is and what it does", and "Travelling in Europe 2020" in the five languages.
  - *Only the accessibility level of "Europe in 12 lessons" and "The European Union What it is and what it does" is evaluated by EDA and assessed using PAC2021.*

FIGURE 18: DOCUMENTS OF THE ORIGINAL SUBSET

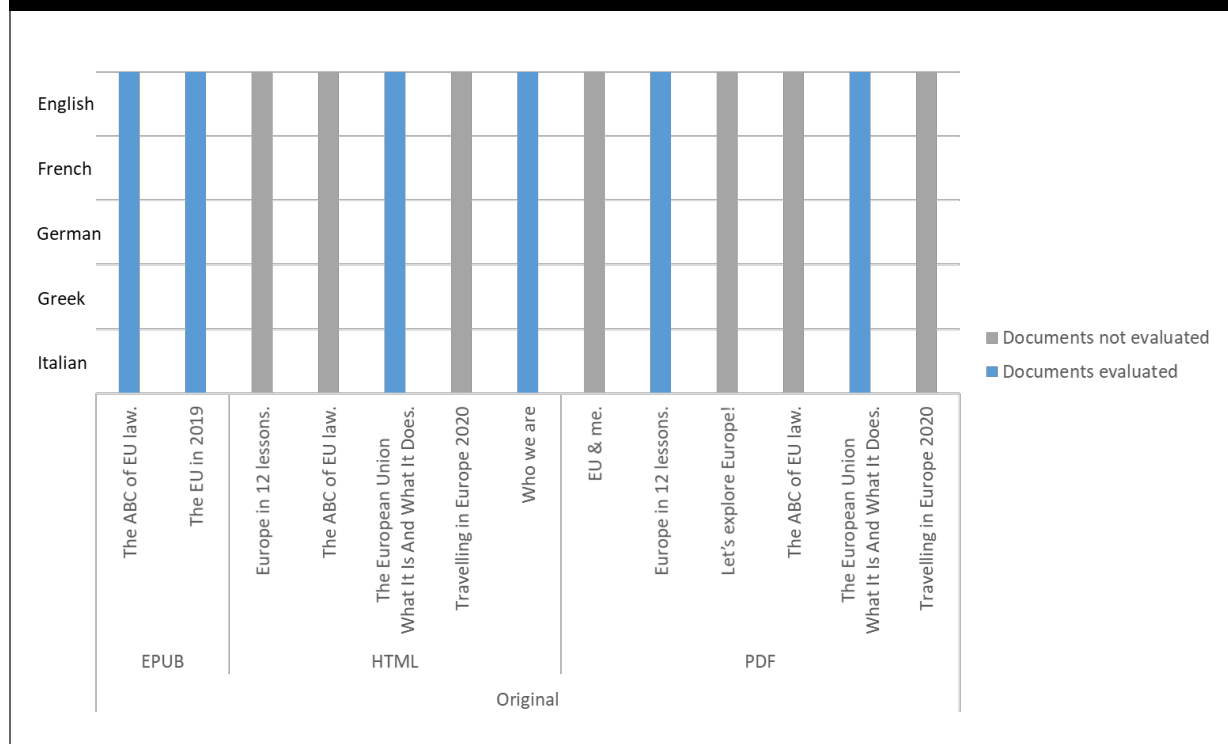
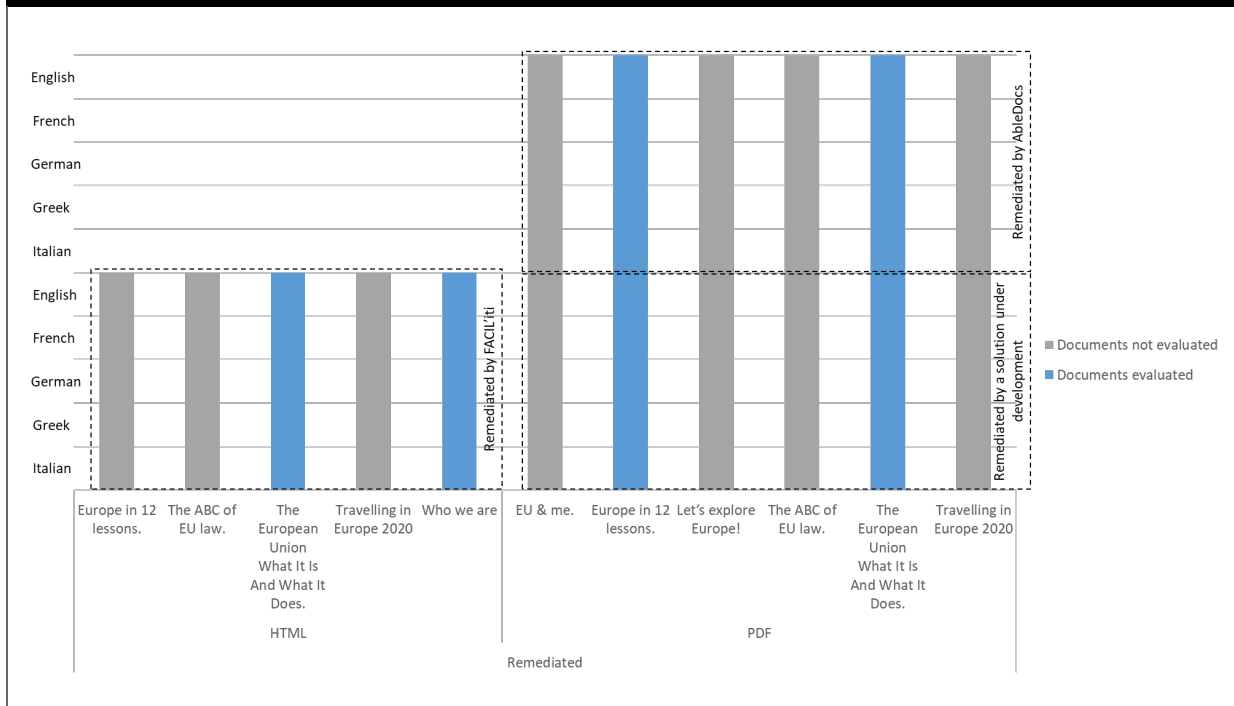


FIGURE 19 shows the documents of the remediated subset and the documents for which the level of accessibility is assessed.

- The remediated subset includes the **five HTML documents** of the original subset in the five languages (English, French, German, Greek and Italian) which are remediated by FACIL'iti.
  - Only the accessibility level of "The European Union What it is and what it does" and "Who we are" is evaluated by EDA.
- The remediated subset also includes the **six PDF documents** of the original subset in the five languages which are remediated by AbleDocs and a solution under development.
  - Only the accessibility level of "Europe in 12 lessons" and "The European Union What it is and what it does" is evaluated by EDA and assessed using PAC2021

FIGURE 19: DOCUMENTS OF THE REMEDIATED SUBSET



Finally, the **original subset** includes 2 EPUB documents, 5 HTML documents and 6 PDF documents in the five languages (English, French, German, Greek and Italian). For each type of documents, two are selected for accessibility assessment. The **remediated subset** includes 5 HTML documents remediated by FACIL'iti and 6 PDF documents remediated by AbleDocs and a solution under development in the five languages. In the context of this Proof of Concept, providers of solutions have been asked to provide only automated remediation. The documents from the remediated subset for which the level of accessibility is assessed, are the same as the original subset.

For more information on the process to select the original subset and the documents for which the accessibility level is assessed, please refer to the section I. In addition, the documents to assess are listed in Annex 7.1 and tagged as document to be assessed.

In conclusion, only 6 documents (2 EPUB, 2 HTML and 2 PDF) are evaluated due to the lack of time and human resources for manual evaluation and semi-automatic assessment. The results presented in this report should be read with care as only 6 documents out of 53 315 documents from OP Portal are evaluated and assessed.

## 4.2. Methodology

First, this subsection explains how the manual evaluation is done by EDA. It presents the surveys sent to a panel of users including people with dyslexia and experts in this domain. It also explains how the raw results are checked and processed before to be analysed.

Second and third, this section explains how the semi-automated evaluations is done using Siteimprove for HTML pages and PAC2021 for PDF documents.

#### 4.2.1 Manual evaluation by EDA

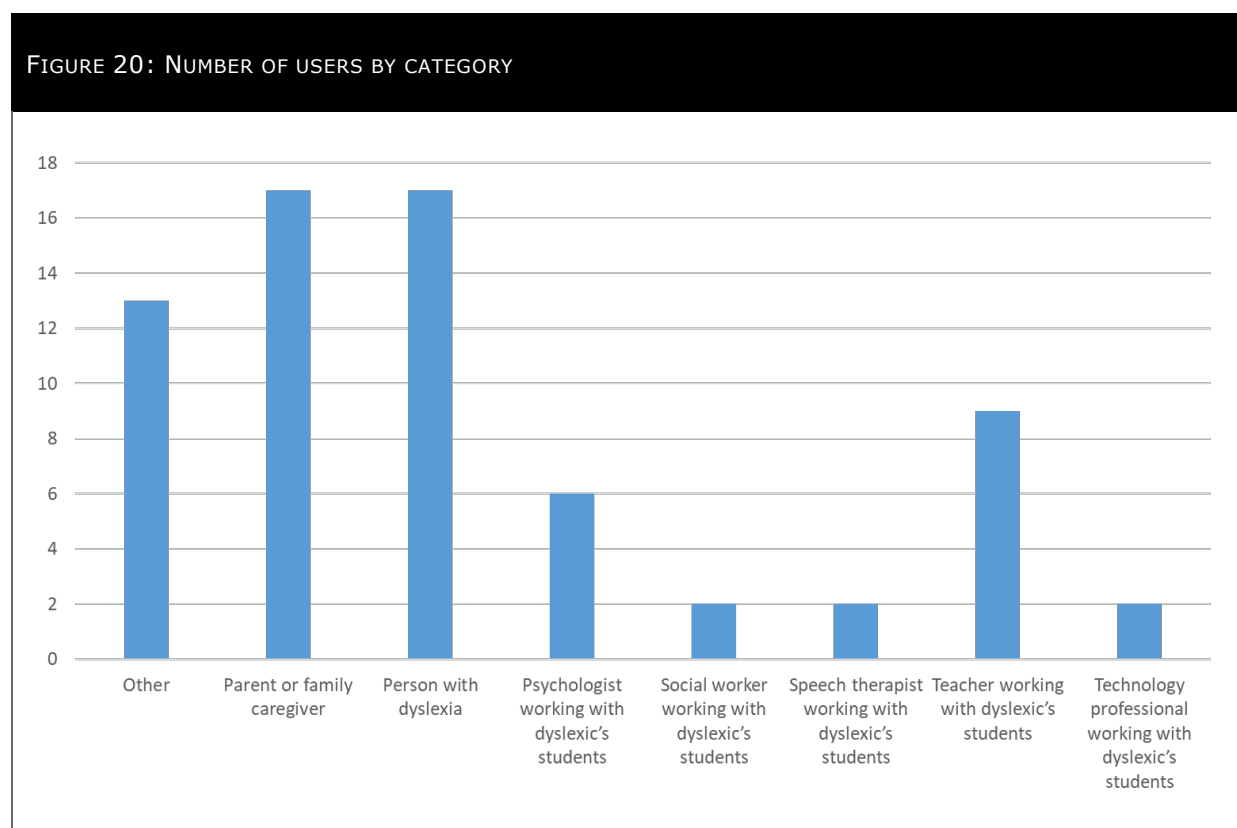
The European Dyslexia Association is an umbrella non-profit organisation with members in 24 countries in and around EU. In this way, EDA is the best organisation to collect data in multiple countries. EDA was in charge to select users for the evaluation, create a method of assessment using a survey and manage the assessment with national organisations and their users.

##### Users

The users are adults with dyslexia in EU member states and/or persons working daily to provide them accommodation. Finally, the number of users is shown in **FIGURE 20** based on the following categories he/she belongs to:

- Parent or family caregiver (17 users),
- Person with dyslexia (17 users),
- Psychologist working with dyslexic's students (6 users),
- Social worker working with dyslexic's students (2 users),
- Speech therapist working with dyslexic's students (2 users),
- Teacher working with dyslexic's students (9 users),
- Technology professional working with dyslexic's students (2 users),
- Others (13 users).

**FIGURE 20: NUMBER OF USERS BY CATEGORY**



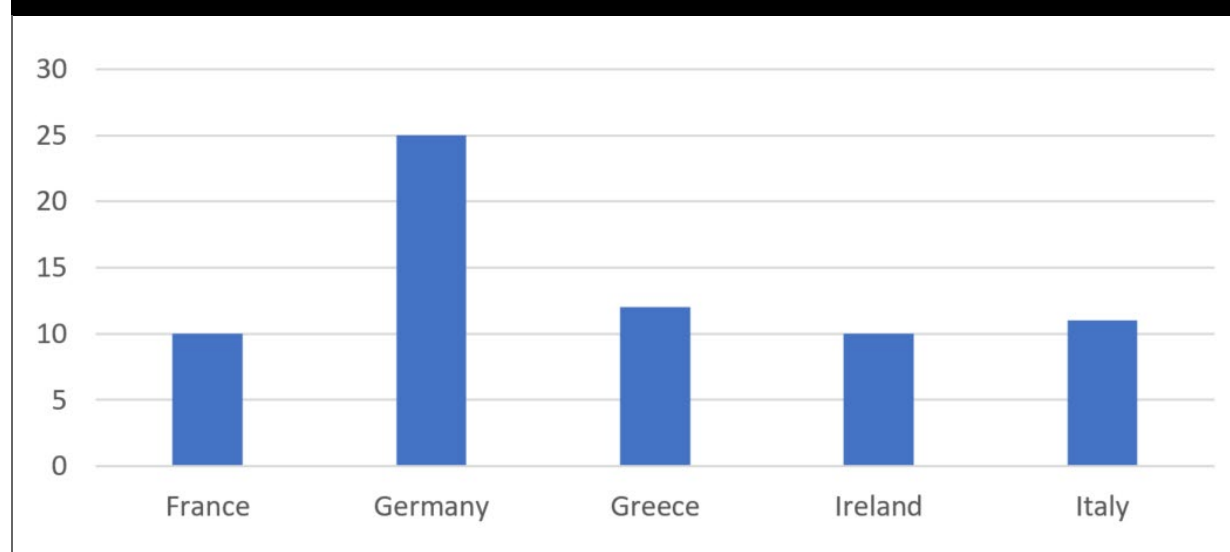


EDA has been supported by the following national member organisations in order to gather users that read the different languages selected in this proof of concept:

- Fédération Française des DYS, for French documents;
- Bundesverband Legasthenie, for German documents;
- Ελληνική Ένωση Δυσλεξίας, for Greek documents;
- Dyslexia Association of Ireland, for English documents;
- Associazione Italiana Dislessia, for Italian documents.

As shown in FIGURE 21, 10 users answer the surveys in French (France), 25 users in German (Germany), 12 users in Greek (Greece), 10 users in English (Ireland) and 11 users in Italian (Italy).

FIGURE 21: NUMBER OF USERS BY COUNTRY



## Surveys

Two surveys have been built:

1. **A generic survey**, to evaluate the accessibility level of a document. This survey is used on the original subsets, including PDF, HTML and EPUB documents and also on remediated PDF documents as the remediation provided by AbleDocs and a solution under development is static and can be compared.
2. **A FACIL'iti survey**, to evaluate the tool and the accessibility level of remediated documents using FACIL'iti. This survey is only used on FACIL'iti remediated documents because the tool is based on user preferences so the display of the document will be different for each user and the questions designed for the generic survey are not any more relevant and comparable.

The questions of the generic survey are presented in Annex 7.2 and the questions of the FACIL'iti survey are presented in the Annex 7.3. Each user was asked to fill in the survey for each document to assess.

## Preparation of the data

Users filled in by themselves the surveys online. The quality of the answers was checked before starting the analysis. This involved sorting and cleaning the answers based on the following criteria:

- Only answers provided using a computer are taken into account (more than 98% of the answers);
- Only the first answer is taken into account when a user answers the survey for the same document multiple times;
- Only answers with full completion of the survey are taken into account;
- Only answers from native speakers are considered.

Finally, 529 answers for the generic survey are considered in the following analysis and 85 answers for the FACIL'iti survey.

To analyse and compare user experiences about reading original and remediated documents, a **user experience score** is calculated for each document. The answers of each question of the **generic survey** are numbered from 0 to 3 where 0 is not good (very difficult to read, not suitable at all, ...) and 3 is great (very easy to read, very suitable, ...). The "non-applicable" answers and the comments questions are not taken into consideration for the calculation of the user experience score. For each assessment/filled-in survey, an assessment score is calculated by averaging the responses of all the questions in the survey that are then set out of 100 points. As for one document, the survey has been filled-in multiple times by different users, the user experience score per document corresponds to the average of the assessment scores.


### 4.2.2 Semi-automated evaluation using Siteimprove

In the scope of the proof of concept, Siteimprove is used only to analyse the accessibility level of HTML pages from the original subset. For each document, Siteimprove provides an accessibility score out of 100. This accessibility score is based on WCAG level (A, AA, AAA) conformance, WAI-ARIA authoring practices and accessibility best practices. This accessibility score will be used for the analysis of accessibility of HTML documents.

### 4.2.3 Semi-automated evaluation using PAC2021


PAC2021 is used to analyse the accessibility level of PDF pages from the original and enhanced subset. The tool checks a PDF document and generates a report as shown on FIGURE 22.


FIGURE 22: PAC2021 TEST REPORT EXAMPLE



# PDF Accessibility Checker 2021

PRERELEASE Version: 21.0.0.0

  
[About](#)




Title  
**Die Europäische Union: Was sie ist und was sie tut**

Filename  
**OR-DE-P-02.pdf**

|           |             |           |             |
|-----------|-------------|-----------|-------------|
| Language  | Tags        | Pages     | Size        |
| <b>de</b> | <b>6149</b> | <b>64</b> | <b>8 MB</b> |


PDF/UA


WCAG





✘ **This PDF file is not PDF/UA compliant.**

| Checkpoint   | Passed | Warned | Failed |
|--|--------|--------|--------|
| <span style="color: green; font-weight: bold;">✔</span> PDF Syntax           | 13178  | 0      | 0      |
| <span style="color: green; font-weight: bold;">✔</span> Fonts                | 24     | 0      | 0      |
| <span style="color: red; font-weight: bold;">✘</span> Content                | 263577 | 0      | 1579   |
| <span style="color: gray; font-weight: bold;">⊘</span> Embedded Files        | 0      | 0      | 0      |
| <span style="color: green; font-weight: bold;">✔</span> Natural Language     | 128023 | 0      | 0      |
| <span style="color: red; font-weight: bold;">✘</span> Structure Elements     | 1192   | 0      | 35     |
| <span style="color: yellow; font-weight: bold;">⚠</span> Structure Tree      | 5772   | 80     | 0      |
| <span style="color: green; font-weight: bold;">✔</span> Role Mapping         | 6224   | 0      | 0      |
| <span style="color: red; font-weight: bold;">✘</span> Alternate Descriptions | 12208  | 0      | 657    |
| <span style="color: red; font-weight: bold;">✘</span> Metadata               | 2      | 0      | 1      |
| <span style="color: red; font-weight: bold;">✘</span> Document Settings      | 2      | 0      | 53     |


Results in Detail


Logical Structure


Screen Reader Preview


Document Statistics

In this example, the PDF file is not PDF/UA compliant. The report shows the different checkpoints that have been passed, warned or failed. PAC2021 evaluates the accessibility of PDF files according to [ISO/DIN-Standard 14289-1 \(PDF/UA\)](#)<sup>18</sup> and includes the following checkpoints:

- Basic requirements:
  - PDF syntax,
  - Fonts,
  - Content,
  - Embedded Files,
  - Natural Language,
- Logical structure:
  - Structure elements,
  - Structure tree,
  - Role Mapping,
  - Alternative Descriptions,
- Metadata and Settings:
  - Metadata,
  - Document settings.

As Siteimprove, an accessibility score out of 100 is calculated based on passed, warned or failed checkpoints. The accessibility score corresponds to the average of passed checkpoints percentage. In the example of FIGURE 22, the accessibility score is 86/100. This accessibility score will be used for the analysis of the accessibility of PDF documents.

### 4.3. Analysis

This section analyses the results and the accessibility level based on three axes in order to answer the questions mentioned in TABLE 12.

| TABLE 12: QUESTIONS TO ANSWER PER AXES |  |
|--|--|
| Axes                                   | Questions  |
| Language comparison                    | 1. Do the type of language (non – semi - transparency) and alphabet (non - Latin) impact reading?<br>2. Do the type of language and alphabet impact remediation? |
| Type comparison                        | 3. Does the type of document (PDF or HTML) impact reading?   |
| Gain in accessibility                  | 4. Does the remediation improve reading?   |

---

<sup>18</sup> Standard ISO 14289-1 (PDF/UA), <https://www.iso.org/standard/64599.html>, [accessed on 14.01.2021]

### 4.3.1 Language comparison

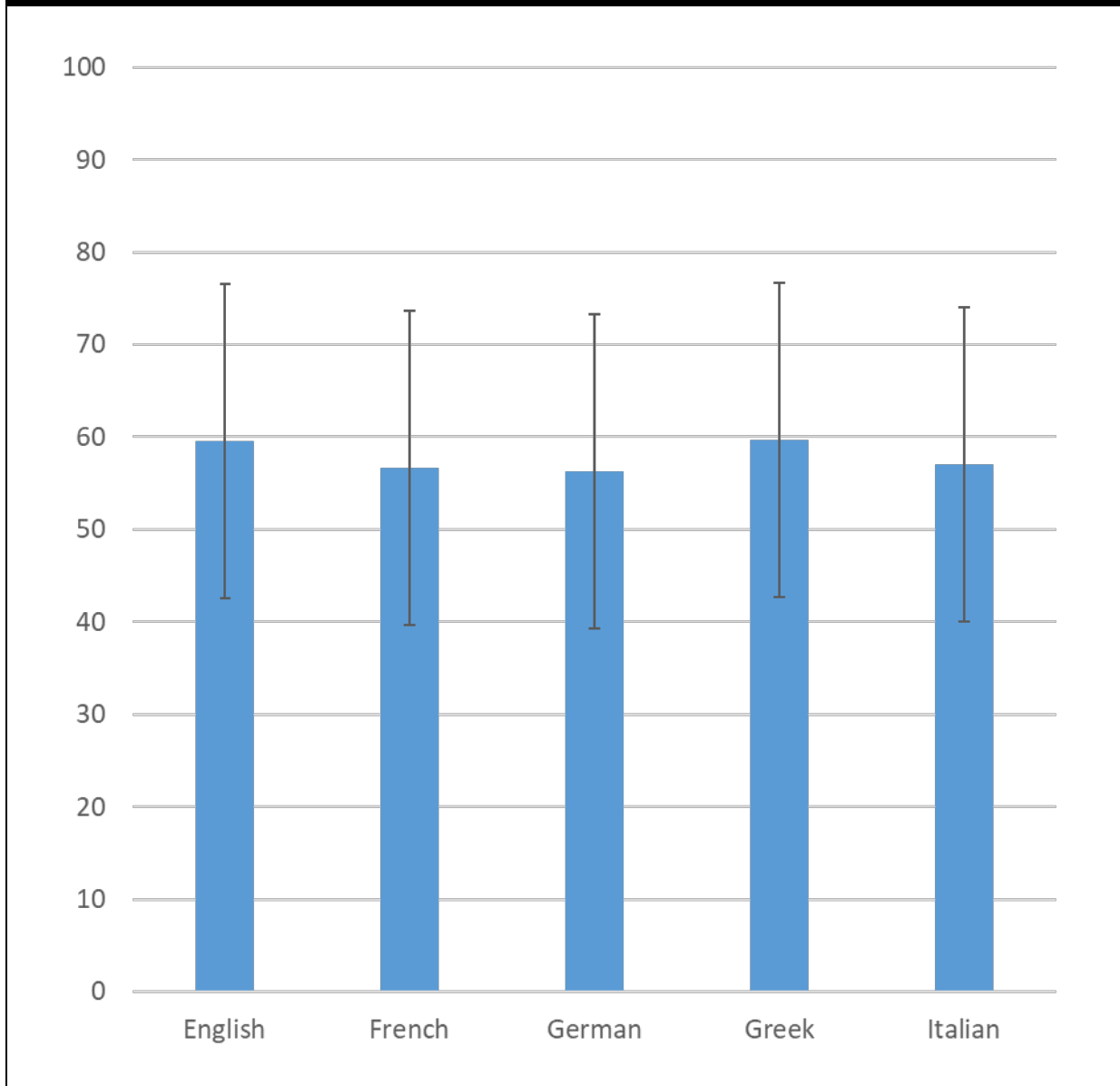
This section is structured to answer to these two questions:

- Do the type of language (non – semi - transparency) and alphabet (non - Latin) impact reading?
- Do the type of language and alphabet impact remediation?

Do the type of language (non – semi - transparency) and alphabet (non - Latin) impact reading?

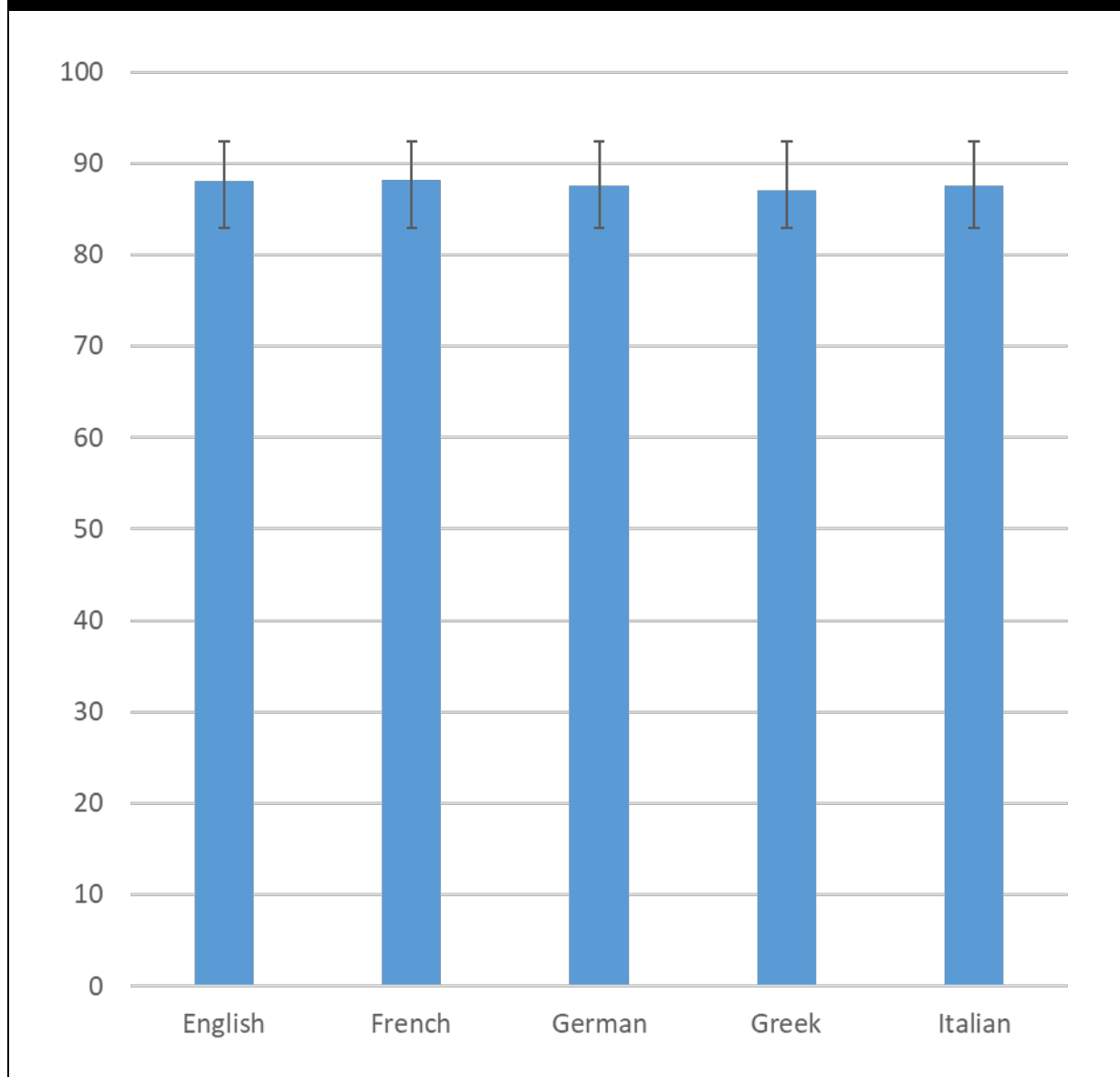
FIGURE 23 shows the average of user experience score per language with standard deviation. The average of user experience score varies between 56,2 for German documents and 59,6 for Greek documents which is insignificant for the collected data.

FIGURE 23: AVERAGE OF USER EXPERIENCE SCORE PER LANGUAGE WITH STANDARD DEVIATION



In the same way, FIGURE 24 shows the average of accessibility score per language using the semi-automatic solutions, Siteimprove and PAC2021. The accessibility score varies between 87 for Greek and 88,2 for French which is insignificant.

FIGURE 24: AVERAGE OF ACCESSIBILITY SCORE PER LANGUAGE WITH STANDARD DEVIATION

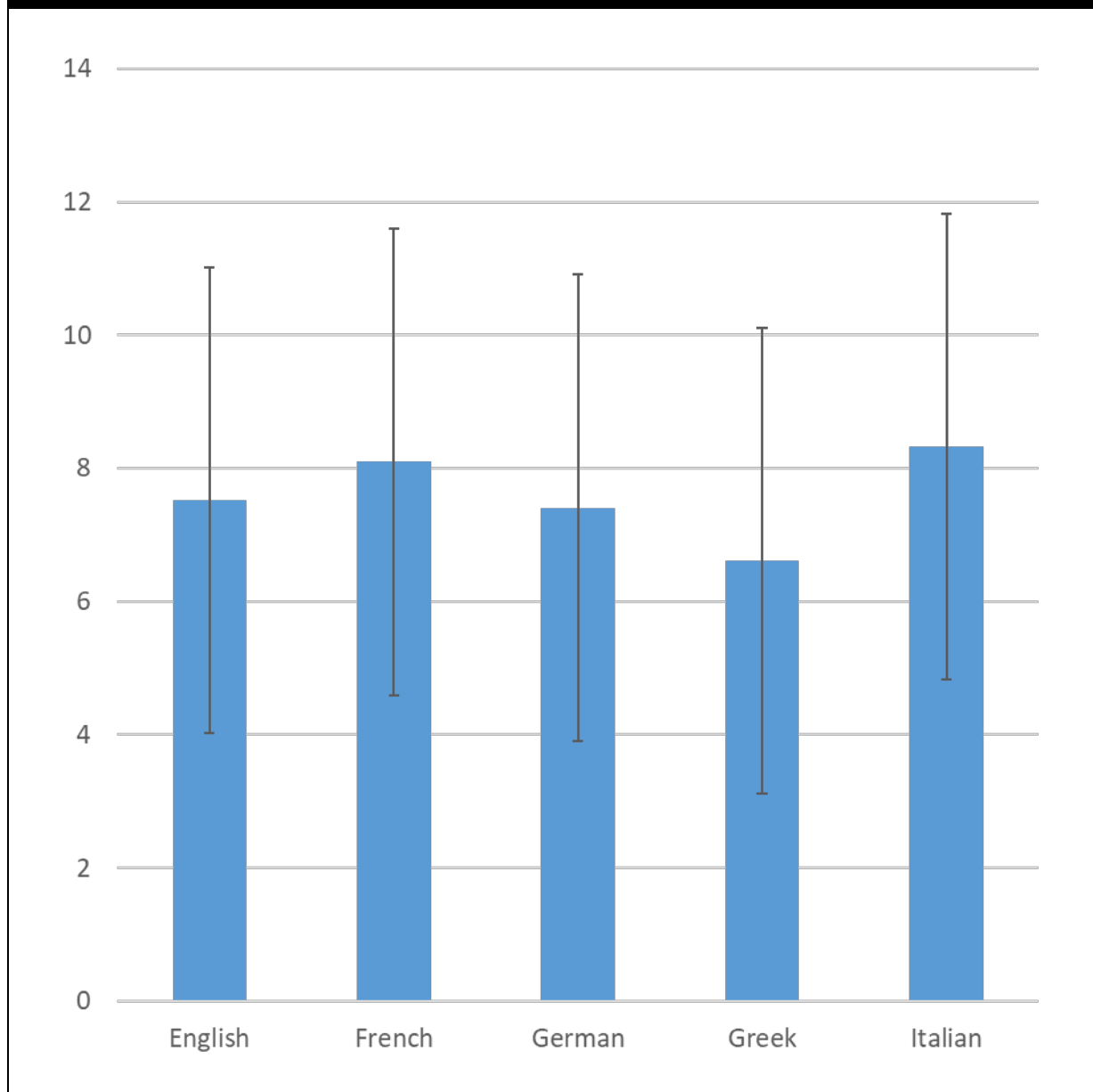


Finally, based on the data collected, it appears that the type of language (non – semi - transparency) or alphabet (non - Latin) do not impact the reading or the accessibility score.

Do the type of language and alphabet impact remediation?

The difference/delta of the accessibility score before and after remediation is calculated on PDF subset. FIGURE 25 shows the average delta of accessibility score per language. It shows that the average delta per language varies between 6,6 for Greek documents and 8,3 for Italian documents which is insignificant.

FIGURE 25: AVERAGE OF DELTA OF ACCESSIBILITY SCORE PER LANGUAGE WITH STANDARD DEVIATION



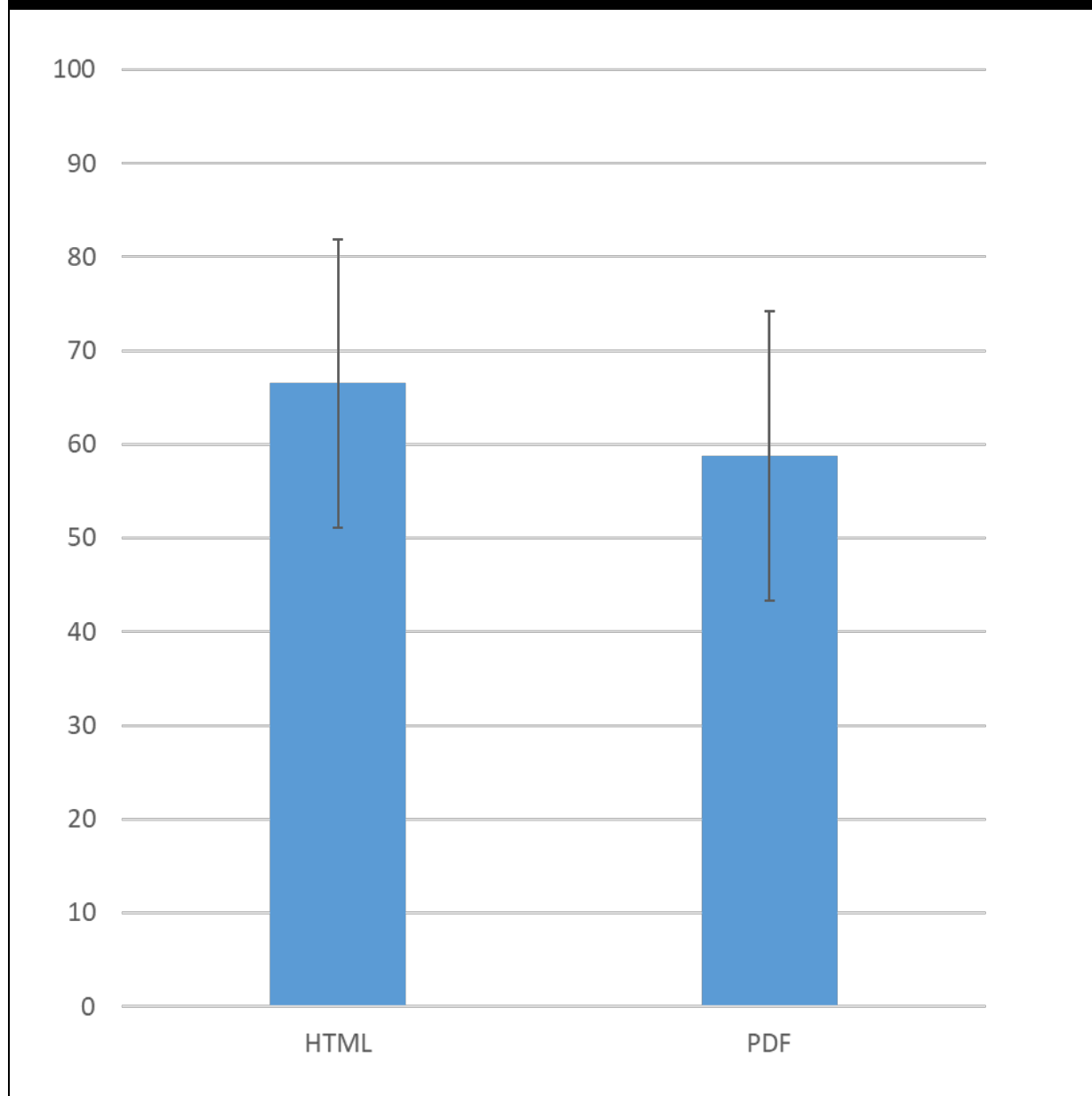
Finally, based on the data collected, it appears that the type of language (non – semi - transparency) or alphabet (non - Latin) do not impact the remediation.

#### 4.3.2 Type comparison

This subsection tries to understand if the type of document (PDF or HTML) impact reading. FIGURE 26 shows the average of user experience score per type of document on the document named “The European Union. What it is and what it does.” of the original subset. It shows that the average of user experience score for the HTML document is 66,5 and for the PDF document 58,7 which is insignificant. Based on the results of this panel of users, it is not possible to draw a complete conclusion.



FIGURE 26: AVERAGE OF USER EXPERIENCE SCORE PER TYPE OF DOCUMENT ON THE DOCUMENT NAMED "THE EUROPEAN UNION. WHAT IT IS AND WHAT IT DOES." OF THE ORIGINAL SUBSET.

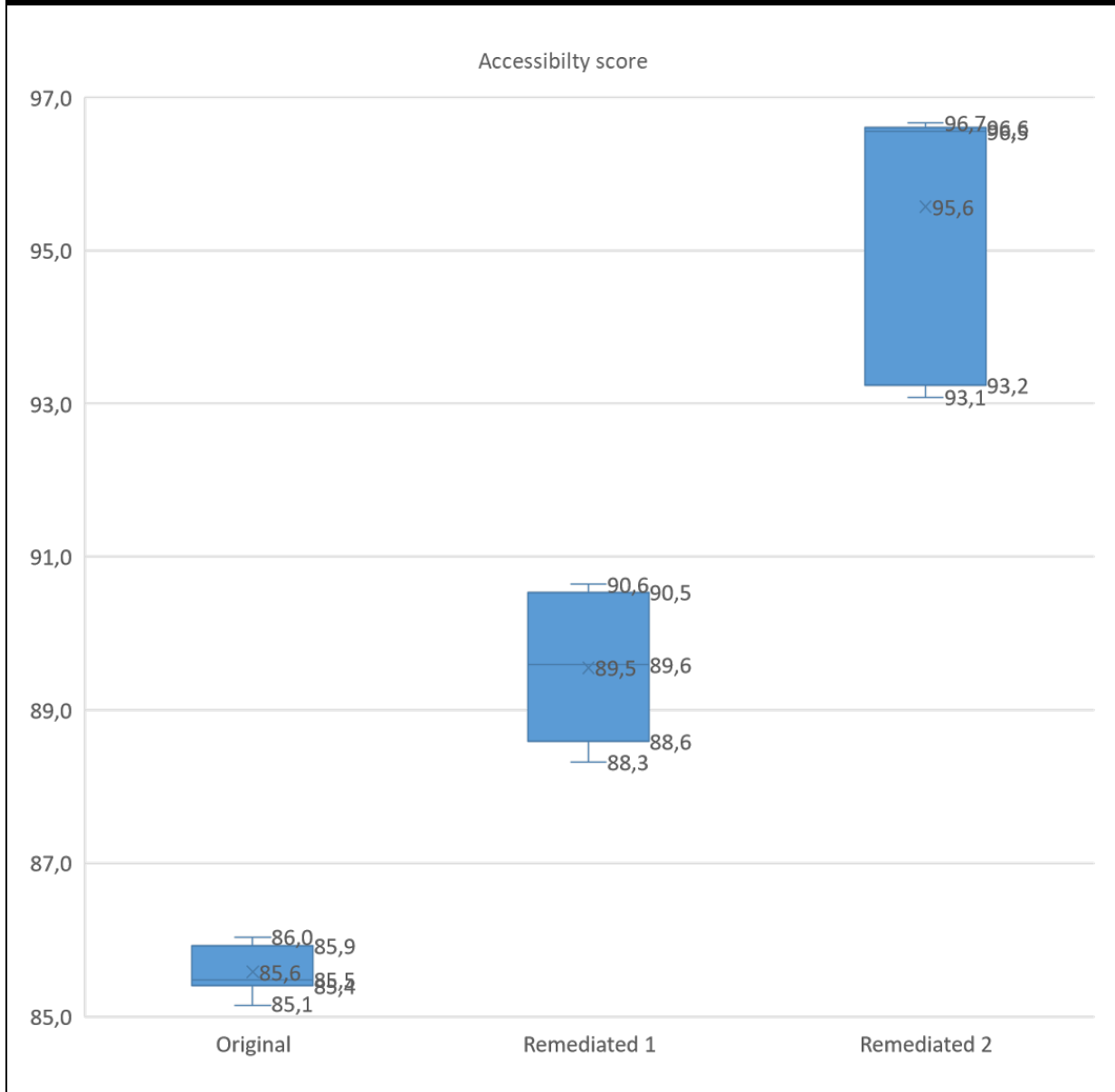


### 4.3.3 Gain in accessibility

This section analyses the gain in accessibility, by means of remediation, and user satisfaction.

FIGURE 27 shows the accessibility score box plot for the original and remediated PDF subsets. It shows that the accessibility score based on PDF/UA increased when using fully automated solutions. The accessibility scores of remediated documents by fully automated solutions does not comply 100% with PDF/UA and manual remediation might be needed to increase the accessibility score and to probably comply with PDF/UA standard.

FIGURE 27: ACCESSIBILITY SCORE BOX-PLOT PER SUBSET FOR PDF DOCUMENTS



The question "Q30. Did using FACIL'iti make the document easier to read?" with its answers (yes, no, unsure) from FACIL'iti survey allows to understand if the user experience is increased or not using FACIL'iti solution. FIGURE 28 shows that most of the users that answered the survey (64,7%) find the document easier to read using FACIL'iti while 35,3% do not find it easier.

FIGURE 28: PERCENTAGE OF NO, UNSURE AND YES ANSWERS TO Q30 OF FACIL'ITI SURVEY

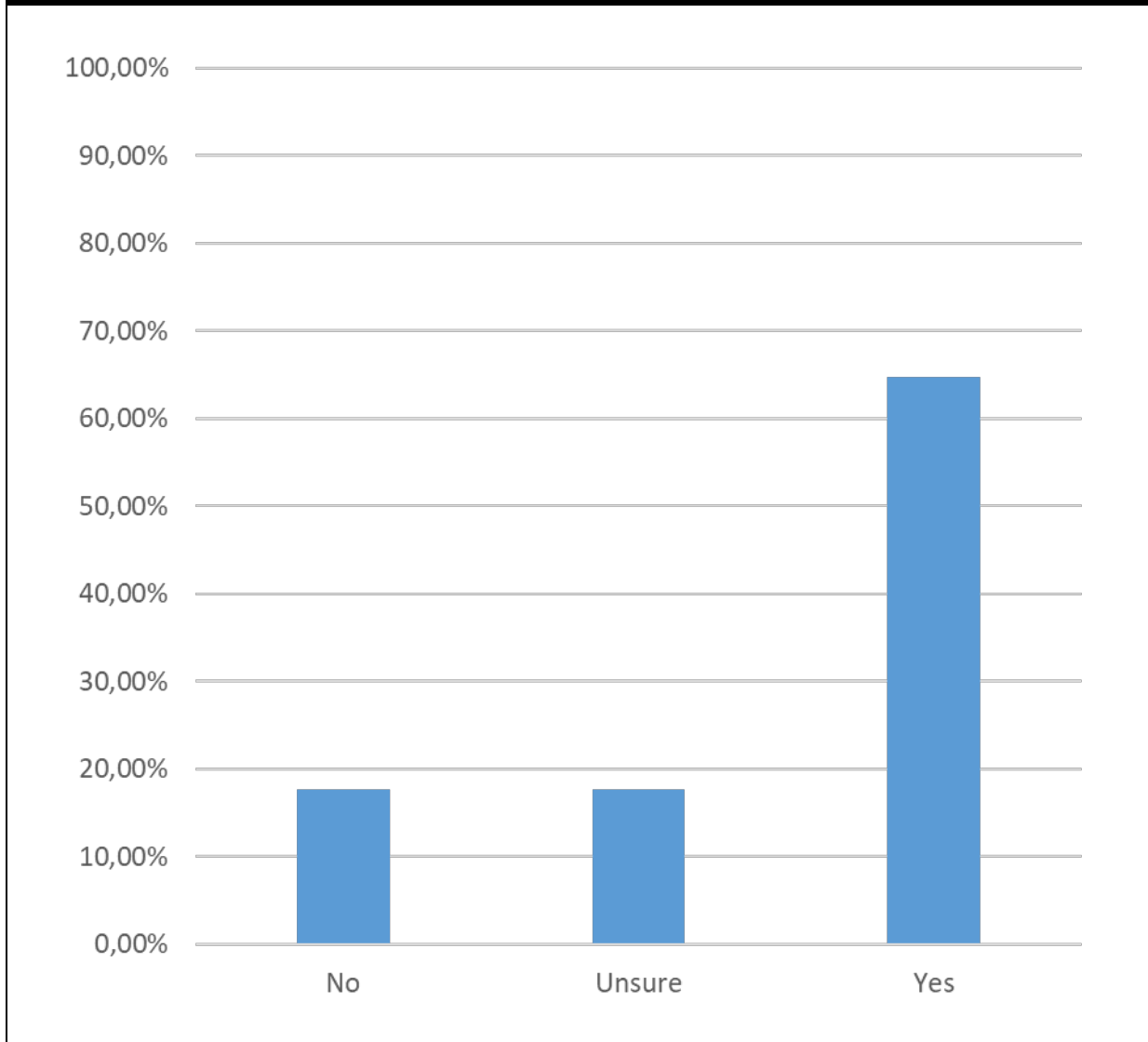
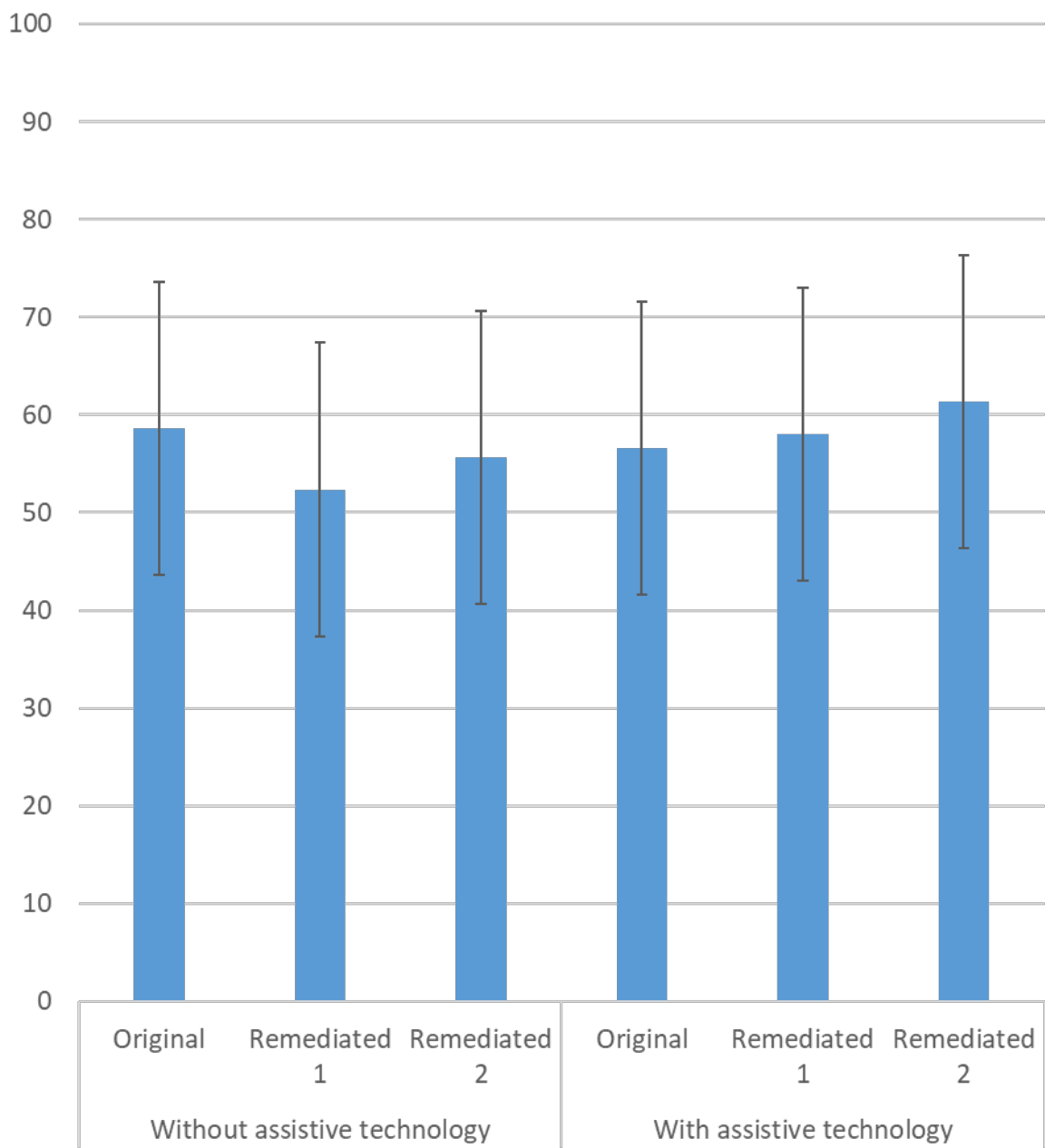


FIGURE 29 shows the average of user experience score per subset for PDF documents. The average varies between 58,6 for the original subset and 52,3 for one of the remediated subsets without assistive technology and the average varies between 61,3 for one of the remediated subsets and 56,5 for the original subset with assistive technology. In both cases without and with assistive technology, the difference is insignificant for the collected data. The remediation from AbleDocs and a solution under development applies universal design meaning that the layout of the document does not change. Indeed, the remediation mainly concerns the inclusion of tags which may be useful for most of the reading assistive technologies. This might explain the very slight increase on the results with the use of assistive technologies and the feeling of a non-improvement for user without assistive technology.

FIGURE 29: AVERAGE USER EXPERIENCE SCORE PER SUBSET FOR PDF DOCUMENTS WITH STANDARD DEVIATION



Also only 8% of the responses are provided using an assistive technology. As confirmed by EDA, the use of assistive technology is not so common for people with reading disability who usually need tailored text (cf. *Section 2.2.1 Proposed user groups segmentation* from RDDA phase 1 report<sup>19</sup>).

As the selected solution providers for PDF remediation (AbleDocs and a solution under development) provide environmental remediation (i.e. mainly for the use of assistive technology), a comparison between before and after remediation, using PAC2021 (subsection Manual analysis with PAC2021) and screen readers (subsection Manual analysis with screen readers), is carried out manually on a very limited scope of two pages. The objective is to evaluate and identify how user experience using a screen reader can be improved by AbleDocs and a solution under development.

The pages number 3 and 62 from the document entitled "The European Union what it is and what it does" are selected because they have different layouts:

- Page 3, as shown on FIGURE 30, corresponds to the table of content with multiple columns, a heading and a list of links, referring to the different parts of the document.
- Page 62, as shown on FIGURE 31, corresponds to a content page with two columns of text, a graphic and links to other pages.

---

<sup>19</sup> Publication Office, RDDA phase 1 report, Section 2.2.1 Proposed user groups segmentation, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

# Contents

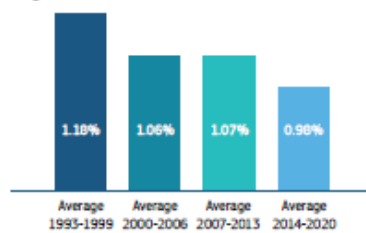
|   |    |
|---|----|
| The European Union:<br>What it is and what it does.....         | 5  |
| 1. The European Union in brief.....                             | 7  |
| 2. What the European Union does .....                           | 13 |
| Climate action and the<br>European Green Deal.....              | 14 |
| Economy, finance<br>and the euro.....                           | 15 |
| Migration and asylum.....                                       | 16 |
| Borders and security.....                                       | 17 |
| Business and industry.....                                      | 18 |
| Single market.....  | 19 |
| Digital economy<br>and society.....                             | 20 |
| Employment<br>and social affairs.....                           | 21 |
| Education<br>and training.....                                  | 22 |
| Research<br>and innovation.....                                 | 23 |
| Regional policy.....  | 24 |
| Transport.....  | 25 |
| Agriculture and<br>rural development.....                       | 26 |
| Maritime affairs<br>and fisheries.....                          | 27 |
| Environment.....  | 28 |
| Energy.....   | 29 |
| Foreign affairs<br>and security policy.....                     | 30 |
| EU enlargement.....   | 31 |
| European neighbourhood<br>policy.....                           | 32 |
| Trade.....  | 33 |
| International cooperation<br>and development.....               | 34 |
| Humanitarian aid<br>and civil protection.....                   | 35 |
| Justice and<br>fundamental rights.....                          | 36 |
| Public health.....  | 37 |
| Food safety.....  | 38 |
| Consumers.....  | 39 |
| Banking and<br>financial services.....                          | 40 |
| Competition.....  | 41 |
| Taxation.....   | 42 |
| Customs.....  | 43 |
| Culture and media.....  | 44 |
| Youth.....  | 45 |
| Sport.....  | 46 |
| Budget.....   | 47 |
| Fraud prevention.....   | 48 |
| 3. How the European Union makes decisions and takes action..... | 51 |

The Commission is responsible for managing and implementing the EU budget and the policies and programmes adopted by the Parliament and the Council. Most of the actual implementation and spending is done by national and local authorities but the Commission is responsible for supervising it. The Commission handles the budget under the watchful eye of the *European Court of Auditors*. Both institutions aim to ensure sound financial management.

Every year, following a recommendation by the Council, the European Parliament decides whether to give its final approval, or 'discharge', on the way the Commission has implemented the EU budget. This procedure ensures full accountability and transparency, and when granted the discharge leads to the formal closure of the accounts for a given year.

The European Court of Auditors is the independent external audit institution of the European Union. It checks that the EU's income has been received correctly, that its expenditure has been incurred in a legal and regular manner and that financial management has been sound. It performs its tasks independently from the other EU institutions and governments.

The size of the EU budget as a percentage of gross national income



Date: EU budget payment ceiling expressed as a percentage of EU GNI  
Source: European Commission

**A modern budget for an EU that protects, empowers and defends**

In May 2018 the Commission presented its proposals for a modern, long-term budget for the 2021-2027 period. The proposals are a realistic response to an exceptionally challenging context: technological and demographic change, migration, climate change and scarce resources, unemployment and security threats are compounded by geopolitical instability.

The Commission has proposed to boost funding in areas where the EU can contribute most effectively. This will be done by expanding and modernising existing programmes with a proven track record, and by creating new tailor-made programmes in areas where a fresh approach is needed to help the EU deliver on its ambitions. For example:

- investing in innovation and the digital economy;
- creating training and employment opportunities for young people;
- continuing the EU's work on a comprehensive approach to migration and border management;
- building up the EU's capacity in security and defence;
- strengthening the EU's external action and investing in climate action and environmental protection;
- strengthening the economic and monetary union.

See also the following pages in Section 2: [Economy, finance and the euro](#); [Banking and financial services](#); [Taxation](#); [Budget](#); and [Fraud prevention](#).

## Manual analysis with PAC2021

Firstly, the analysis is made with PAC2021 which offers a feature called "screen reader preview" that displays the structure of the imported documents. It shows the headings, paragraphs, figures and other elements of the document in coloured boxes which provides an overview of the document's hierarchy, and allows a first comparison between the original and the remediated documents. The structure of this subsection Manual analysis with PAC2021 is as follows:

- Page 3 comparison between the original document and AbleDocs,
- Page 3 comparison between the original document and a solution under development,
- Page 62 comparison between the original document and the remediated documents.

These comparisons focus on four main aspects:

- Headings,
- The structure of the document,
- Section titles,
- Figures.

### *Page 3 comparison between the original document and AbleDocs*

Using this PAC2021 feature on the page 3, three components vary between the original document and the one remediated by AbleDocs (FIGURE 32 and FIGURE 33):

- **Headings:** considered as paragraphs in the original document, they are correctly identified by "H" (Heading) tags in the one remediated by AbleDocs.
- **The structure of the table of content:** the structures of the two documents are slightly different, the same tags are used (TOC: Table of Content, TOCI: Table of Content Item) but with a different hierarchy.
- **Section titles:** titles to other pages are marked as link in the AbleDocs document, while titles in the original document are marked as reference and link. The dots separating section titles and page numbers are correctly detected in the remediated version and not in the original version.



FIGURE 32: PAC 2021 SCREEN READER PREVIEW EXTRACT ON PAGE 3 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" ORIGINAL DOCUMENT

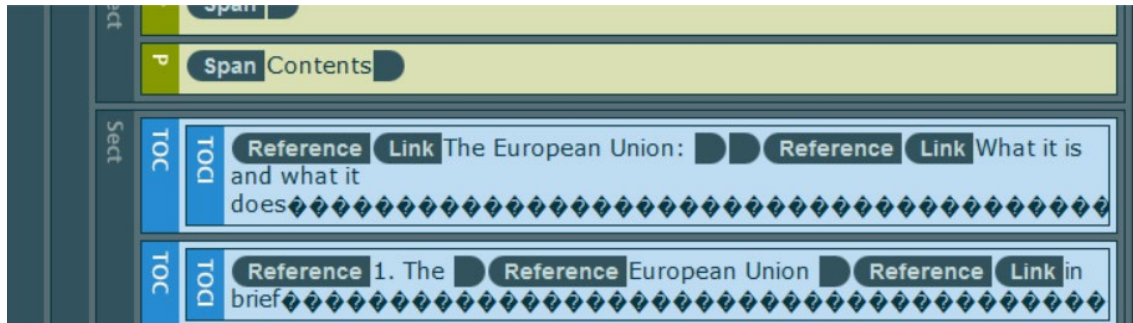


FIGURE 33: PAC 2021 SCREEN READER PREVIEW EXTRACT ON PAGE 3 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" REMEDIATED WITH ABLEDOCS



Page 3 comparison between the original document and a solution under development

Regarding the document remediated by the solution under development (screen readers FIGURE 34), modifications on the structure of the table of content and section titles are noticeable:

- **The structure of the table of content:** in the original document, the structure is composed of TOC and TOCI tags, whereas in the remediated document it is composed of L (List), LI (List Item), LBody (List Body) and Table. Furthermore, this structure is not consistent throughout the page, the tags L and LI change to TOC and TOCI in the middle of the page.
- **Section titles:** the titles in the remediated document are tagged as reference and not as link and reference. The dots used as decorators and separators for the title and page number are no longer present in the subtitles.

FIGURE 34: PAC 2021 SCREEN READER PREVIEW EXTRACT ON PAGE 3 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" REMEDIATED WITH THE SOLUTION UNDER DEVELOPMENT

|       |                               |       |  |                                  |                    |
|-------|-------------------------------|-------|--|----------------------------------|--------------------|
| p     | Contents                      |       |  |                                  |                    |
| p     | Reference The European Union: |       |  |                                  |                    |
| L     | LI                            | LBody | Reference What it is and what it does.....5            |                                  |                    |
| L     | LI                            | LBody | Reference 1. The European Union in brief.....7         |                                  |                    |
| L     | LI                            | LBody | Span Reference 2. What the European Union does .....13 |                                  |                    |
| Table | TR                            | TH    | p  | Reference Climate action and the | Reference European |
| Table | TR                            | TD    | p  | Spai Reference Economy, finance  |                    |
| Table | TR                            | TD    | p  | Spai Reference and the euro      | 15                 |

FIGURE 35 shows an extract of PAC2021 screen reader preview of the page 62 of the original document. The remediations made on this page by AbleDocs (FIGURE 36) and the solution under development (FIGURE 37), are mainly on the headings and on the figure at the end of the page:

- **Heading:** the remediation performed on the document by AbleDocs and the solution under development improve the headings tagging.
- **Figure:** Compare to the original document, where the figure is in plain text, AbleDocs converts the figure into a two-column table. The first column contains the caption and the second, the value associated with the caption. The solution under development convert the figure in an image.

FIGURE 35: PAC 2021 EXTRACT SCREEN READER PREVIEW ON PAGE 62 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" ORIGINAL DOCUMENT

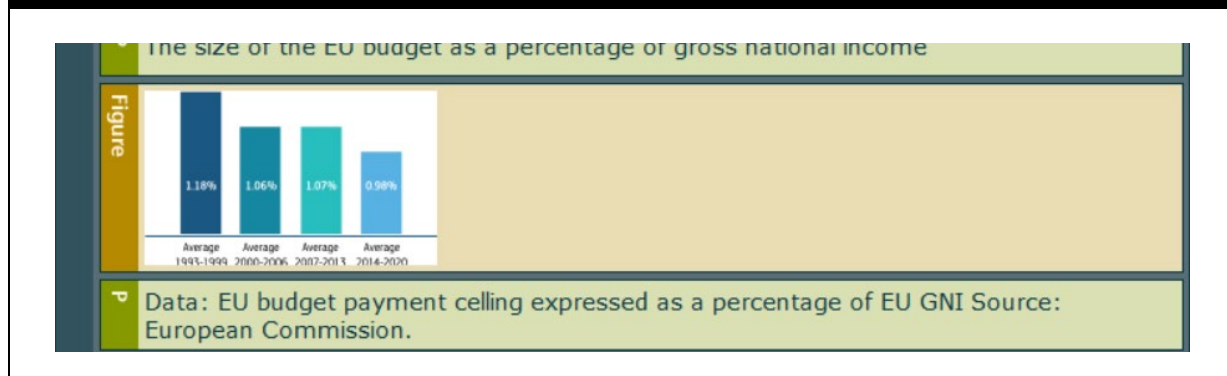
correctly, that its expenditure has been incurred in a legal and regular manner and that financial management has been sound. It performs its tasks independently from the other EU institutions and governments.

|   |      |   |      |   |      |                          |
|---|------|---|------|---|------|--------------------------|
| P | Sect | P | Span | The size of the EU budget as a percentage | Span | of gross national income |
|   | Sect | P | Span | 0.98%                                     |      |                          |
|   | Sect | P | Span | 1.07%                                     |      |                          |

FIGURE 36: PAC 2021 EXTRACT SCREEN READER PREVIEW ON PAGE 62 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" REMEDIATED WITH ABLEDOCS

|     |        |    |  |      |         |           |    |       |
|-----|--------|----|--|------|---------|-----------|----|-------|
| Div | Capiti | P  | The size of the EU budget as a percentage of gross national income |      |         |           |    |       |
|     | Table  | TR | TH   | Span | Average | 1993-1999 | TD | 1.18% |
|     |        | TR | TH   | Span | Average | 2000-2006 | TD | 1.06% |
|     |        | TR | TH   | Span | Average | 2007-2013 | TD | 1.07% |
|     |        | TR | TH   | Span | Average | 2014-2020 | TD | 0.98% |

FIGURE 37: PAC 2021 EXTRACT SCREEN READER PREVIEW ON PAGE 62 OF "THE EUROPEAN UNION WHAT IT IS AND WHAT IS DOES" REMEDIATED WITH THE SOLUTION UNDER DEVELOPMENT



Finally, the manual analysis with PAC 2021 on one example shows that the remediations made by AbleDocs improve the tagging of headings and table of contents, the structure of the document and figure.

#### Manual analysis with screen readers

Secondly, an evaluation of the remediations effects on screen readers is carried out. Screen readers are software applications that attempt to convey visual information via non-visual means, like text-to-speech and sound icons. The analysis is made with two screen readers Acrobat Reader and Nvaccess software:

- Acrobat Reader is one of the most used software for reading pdf, with or without screen readers, magnifiers, and braille printers. In the scope of this experience, the Acrobat Reader "Read out loud" version 21.1.20135.42 is used on Windows 10.
- Nvaccess<sup>20</sup> is a free and open-source software that allows people with blindness and visual impairment, to access and interact with the Windows operating system and many third-party applications. In the scope of this experience, the Nvaccess version 2020.4 is used on Windows 10.

The observations can vary using different versions of the screen readers and operating system.

TABLE 13 presents examples of observations of page 3 and page 62 of the document entitled "The European Union what it is and what is does" using the screen reader feature of Acrobat Reader and Nvaccess. A first listening session is carried out on the original document to identified reading inconsistencies and observations. Then, listening sessions are carried out on remediated documents to find out if inconsistencies and

---

<sup>20</sup> Nvaccess website, <https://www.nvaccess.org/> , [accessed on 26.03.2021]

observations have been corrected. The observations are grouped by category (pause in reading, reading page numbers, reading references, reading of punctuation as decoration and figure reading), by screen reader and for each remediated document.

TABLE 13: OBSERVATIONS USING SCREEN READER FEATURE OF ADOBE READER AND NVACCESS

| Category                              | Original document   | AbleDocs  | Solution under development   | Acrobat reader | Nvaccess |
|---------------------------------------|---|---|--|----------------|----------|
| <p>Pause in reading</p> <p>Page 3</p> | <p>Pause during reading are not at the right time, for example:</p> <ol style="list-style-type: none"> <li>1. Additional pauses when not needed, for example « Climate action and the 'Pause' European Green Deal ». (Acrobat)</li> <li>2. Missing pauses after page numbers, for example, no pause after page number 5 and 7 and read 51 and 52 (Acrobat and Nvaccess).</li> <li>3. Some pauses are missing in reading, leading the reader to interpret them as a single line. E.g. "Regional policy 24 'Missing pause' Transport 25" (Nvaccess).</li> </ol> | <p>The pause problem is no longer present in this version (Acrobat and Nvaccess).</p> | <p>Some of the pauses during reading have been corrected but not all of them.</p> <ol style="list-style-type: none"> <li>1. Not corrected: additional pauses are heard (Acrobat and Nvaccess).</li> <li>2. Corrected: Missing pauses are no longer present (Acrobat and Nvaccess).</li> <li>3. Corrected: Missing pauses are no longer present (Nvaccess).</li> <li>4. An unnecessary pause is taken by the reader between subheading and page number. Leading the reader to read the page number and the following subtitle in the same sentence. For example: "Regional policy 'Pause' 24 Transport 'Pause' 25 ..." (Nvaccess).</li> </ol> | <p>✓</p>       | <p>✓</p> |

**TABLE 13: OBSERVATIONS USING SCREEN READER FEATURE OF ACRBAT READER AND NVACCESS**

|   |  |  |  |          |          |
|---|--|--|--|----------|----------|
| <p>Reading page numbers</p> <p>Page 3</p>                 | <p>The reading of page numbers is split, e.g., 15 is read "one five"</p>   | <p>The page numbers are correctly read.</p>                                      | <p>The page numbers are correctly read.</p>                                      | <p>✓</p> |          |
| <p>Reading references</p> <p>Page 3 and 62</p>            | <p>References are read once.</p>   | <p>The reference titles are read twice.</p>                                      | <p>The reference titles are read twice.</p>                                      |          | <p>✓</p> |
| <p>Reading of punctuation as decoration</p> <p>Page 3</p> | <p>The dots separating the title of a section and its page number are read by Nvaccess, e.g. "The European Union in brief ..... 7" becomes "The European Union in brief 10 dots, 7".</p> | <p>The problem of reading the dots in the table of content is still present.</p> | <p>The problem of reading the dots in the table of content is still present.</p> |          | <p>✓</p> |

**TABLE 13: OBSERVATIONS USING SCREEN READER FEATURE OF ACRBAT READER AND NVACCESS**

|                                   |   |   |  |          |          |
|-----------------------------------|---|---|--|----------|----------|
| <p>Figure reading<br/>page 62</p> | <p>The reading engine does not make the connection between the information contained in the figure and its caption. The data is listed one after the other without context.</p> | <p>Acrobat is able to read the caption and the value associated in the figure. The reading is consistent and allows a better understanding of the information.</p> <p>Nvaccess also makes it possible to read the structure of the table, allowing better control of the reading (rhythm, repetition) while maintaining the integrity of the information.</p> | <p>The caption and the data of the figure are read independently, as in the case of reading by text. This method is therefore not optimal for understanding.</p> | <p>✓</p> | <p>✓</p> |
|-----------------------------------|---|---|--|----------|----------|



Screen readers are useful for people with audio needs and image comprehension needs (cf. *Section 2.2.1 Proposed user groups segmentation* from RDDA phase 1 report<sup>21</sup>). The remediations performed by AbleDocs and the solution under development, have improved the reading of the documents using Acrobat reader and Nvaccess screen readers by correcting issues such as pauses in reading and reading page numbers.

Furthermore, the manual analysis with screen readers shows that two different screen readers read documents in different ways and do not necessarily encounter the same problems. Finally, the design of both the user tools (screen readers) and the PDF document environment (tagging, alt text, ...) are important in order to provide a good user experience for people with audio needs.

---

21 Publication Office, RDDA phase 1 report, Section 2.2.1 Proposed user groups segmentation, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

#### 4.4. Key findings

This section summarises the key findings of the analysis of the accessibility and user experience score. The key findings refer to 1) the study seems to not systematically correlate gain in accessibility and user experience, 2) the type of language and alphabet does not appear to impact reading and remediation, 3) automated PDF remediation allows to be more compliant with PDF/UA standard, 4) Automated PDF remediation does not allow a document to become fully compliant with PDF/US standard and 5) Format solutions for publishers and webmasters may have less impact on readers than solutions for end users.

##### ***KF1) The study seems to not systematically correlate gain in accessibility and user experience***

Within the scope of this proof of concept, results tend to show that tools based on standards such as AbleDocs and a solution under development based on PDF/UA and Siteimprove based on WCAG standard, allow to better comply with the standard and thus to improve accessibility. But, in the same scope of this study, this did not imply a systematic improvement of the user experience. Some solutions, such as FACIL'iti, seem to be less focused on standards but more on providing a better user experience.

##### ***KF2) The type of language (non – semi - transparency) and alphabet (non - Latin) do not appear to impact reading and remediation***

As shown in this section, the accessibility score, the delta of remediation of the accessibility and the user experience score is almost similar in English, French, German, Greek and Italian. It shows that the type of language (non – semi - transparency) and alphabet (non - Latin) seems to not impact the reading and the remediation.

##### ***KF3) Automated PDF remediation allows to be more compliant with PDF/UA standard***

Based on the data available in this study the analysis could lead to the following conclusion, using fully automated remediation provided by AbleDocs and a solution under development on PDF subset allows to increase the accessibility score and be more compliant with PDF/UA standard.

##### ***KF4) Automated PDF remediation does not allow a document to become fully compliant with PDF/UA standard***

The results of this study tend to show that using fully automated remediation provided by AbleDocs and a solution under development on PDF subset allows to increase the accessibility score but does not comply 100% with PDF/UA standard. The automated solutions are more and more efficient, but manual remediation is still needed, especially to provide alternative description.

##### ***KF5) Format solutions for publishers and webmasters may have less impact on readers than solution for end users***

In RDDA phase 1 report<sup>22</sup>, solutions are classified by users of the solution as follows:

- Authors and contributors, who write the content,
  - Guidelines solutions,
  - Editing solutions,
- Publishers and webmasters, who publish documents and web pages,
  - Evaluation solutions,
  - Formats solutions,
- End-users, who read and access documents or web pages,
  - Fonts solutions,
  - Speech solutions,
  - Image solutions,
  - Paper based solutions,
  - Navigation agent solutions.

Within the framework of this proof of concept, mainly format and automatic solutions (AbleDocs and a solution under development) have been selected to be applicable on a large scale in the institutions of the European Union. Analysis of the results seems to show that, while these solutions increase compliance with the PDF/UA standard, they do not improve the user experience as much as a font solution (FACIL'iti) seems to do. This may be explained by the fact that these solutions do not make direct changes in the visualisation of the document (layout). Only users using assistive technologies may see a difference in the tags, hierarchy and organisation of the document but, even if this creates one document for all users, then respecting the "universal design principle", it seems user prefer to be able to adapt the visualisation to their needs, rather than use specific assistive technology.

Finally, based on the data available in this study, it may be interesting to add functionalities on existing solutions (document readers or websites for example) allowing an inclusive design and thus adapting the layout, navigation and/or other elements to the specific needs of the user. A feature that could be universal since it is accessible to all users but allowing an inclusive design and customisation according to the user's needs (e.g. using FACIL'iti for HTML pages).

---

<sup>22</sup> Publication Office, RDDA phase 1 report, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

## 5. Conclusions

Accessibility is becoming an increasingly important topic, with both public and private sectors working on ways of ensuring equality of rights and access to documents and information. The general information provided by the European institutions should be accessible to all citizens, regardless of age or disabilities. Today, more and more solutions are being developed to increase the accessibility of digital documents and websites. RDDA phase 1 report<sup>23</sup> lists some of these solutions and makes a theoretical analysis of some of them. This report complements the previous analysis with a proof of concept. Some of the solutions identified have been implemented and run to understand the impact on accessibility and user experience.

This report shows the results of the second phase of the “Reading disability and document access, a possible approach” pilot project. Phase 2 involved running a proof of concept to:

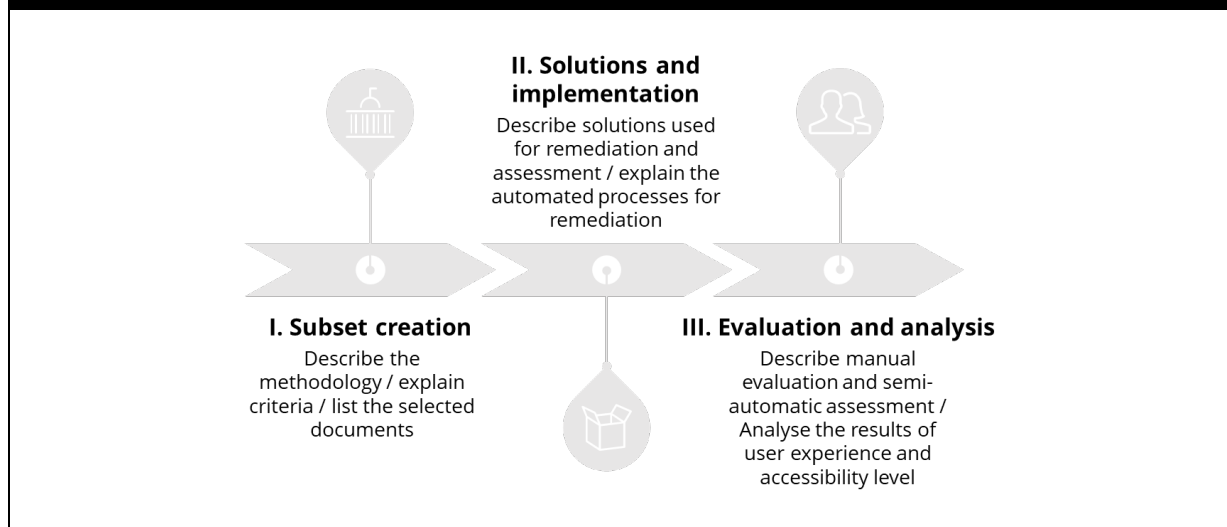
- Make a subset of documents/publications more accessible using solutions identified in RDDA phase 1;
- Assess and evaluate the enhanced documents/publications and websites involving users with reading disability.

The structure of RDDA phase 2 report is described in FIGURE 38.

---

<sup>23</sup> Publication Office, RDDA phase 1 report, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

FIGURE 38: STRUCTURE OF RDDA PHASE 2 REPORT



The conclusions section of the report presents a summary, as well as key findings and outcomes from each section. Future work to enhance documents and web content is also presented.

### **Section I: Subset creation**

The purpose of the first section of the report is to create the original subset. It presents the criteria rationale and the methodology to select the documents and publications of the original subset. The content that best matches the following criteria has been selected for the original subset:

- Public content from OP Portal,
- Content available in HTML, PDF or EPUB formats,
- Content available in the 5 languages (Italian, French, German, English, Greek)
- Most visited content,
- Understandable by anyone (ISCED level 2),
- Diversified layout (plain text, images, tables, ...),
- Most recent content.

Finally, from 53 315 documents available on OP Portal, the original subset includes 6 PDF documents, 5 HTML pages and 2 EPUB documents.

### **Section II: Solutions and implementation**

The purpose of the second section is to describe the solutions and their implementation. Three types of solutions are used in this report:

- **Hosting solution**, allowing to have one platform with the original and enhanced subsets, to install FACIL'iti and to enable dissemination.
- **Remediation solutions**, allowing to enhance the original subset.
- **Evaluation solutions**, allowing to assess the level of accessibility of documents.

The **hosting solution** is built specifically for the PoC and is available at <http://pocrdda.publications.europa.eu/> until end 2021.

The **remediation solutions** are selected from the shortlist of solutions of the RDDA phase 1 report<sup>24</sup> such as:

- At least two solutions dealing with the PDF subset are selected, to cover the fully automated remediation of PDF documents,
- At least two solutions dealing with the web subset are selected, to achieve full/semi-automated remediation of web pages.

So based on these criteria, providers and solutions have been selected considering their availability and willingness to participate free of charge. AbleDocs and a solution under development are selected to remediate PDF documents using only fully automated solutions by mainly tagging the document. FACIL'iti and Siteimprove solutions are selected to remediate HTML pages.

The **evaluation solutions** are selected such as:

- At least one solution dealing with the evaluation of PDF, to semi-automatically evaluate the accessibility level of PDF,
- At least one solution dealing with the evaluation of web pages, to semi-automatically evaluate the accessibility level of HTML pages.

In this PoC, PAC2021 solution is used to assess the accessibility level of PDF documents and Siteimprove is used to assess the accessibility level of HTML documents.

The **key findings** of the implementation of solutions are as follows:

- Tags are the foundation of accessible PDF documents.
- Currently, the "automated only" remediation processes still do not allow a full PDF/UA compliance.
- AbleDocs and a solution under development are focusing efforts in improving their capability to produce automatically tagged PDF documents.
- Some remediation solutions, such as FACIL'iti, focus on the user experience and do not aim to be fully compliant with standards.
- The evaluation solutions are based on standards, such as WCAG for web evaluation with Siteimprove and PDF/UA for PDF evaluation with PAC2021.

### ***Section III: Evaluation and analysis***

---

<sup>24</sup> Publication Office, RDDA phase 1 report, <https://op.europa.eu/en/publication-detail/-/publication/801de94b-27c0-11eb-9d7e-01aa75ed71a1/language-en/format-PDF/source-183027858>, [accessed on 04.01.2021]

The purpose of the third and last section is to understand how the remediation may impact the accessibility. To do so, the accessibility of documents is evaluated manually by EDA and assessed semi-automatically using PAC2021 and Siteimprove solutions.

The scope of evaluated and assessed documents is reduced to 6 documents (2 EPUB, 2 HTML and 2 PDF) due to the lack of time and human resources for manual evaluation and semi-automatic assessment. The results presented in this report should be read with care as only 6 documents out of 53 315 documents from OP Portal are evaluated and assessed.

The key findings of the analysis of the accessibility level and user experience is listed below:

- Within the scope of this proof of concept, results tend to show that tools based on standards such as AbleDocs and a solution under development based on PDF/UA and Siteimprove based on WCAG standard, allow to better comply with the standard and thus to improve accessibility. But, in the same scope of this study, this did not imply a systematic improvement of the user experience. Some solutions, such as FACIL'iti, seem to be less focused on standards but more on providing a better user experience.
- The type of language (non – semi - transparency) and alphabet (non - Latin) do not appear to impact reading and remediation.
- Automated PDF remediation allows to be more compliant with PDF/UA standard.
- Within the framework of this proof of concept, mainly format and automatic solutions (AbleDocs and a solution under development) have been selected to be applicable on a large scale in the institutions of the European Union. Analysis of the results seems to show that, while these solutions increase compliance with the PDF/UA standard, they do not improve the user experience as much as a font solution (FACIL'iti) seems to do. This may be explained by the fact that these solutions do not make direct changes in the visualisation of the document (layout). Only users using assistive technologies may see a difference in the tags, hierarchy and organisation of the document but, even if this creates one document for all users, then respecting the “universal design principle”, it seems user prefer to be able to adapt the visualisation to their needs, rather than use specific assistive technology.
- Finally, based on the data available in this study, it may be interesting to add functionalities on existing solutions (document readers or websites for example) allowing an inclusive design and thus adapting the layout, navigation and/or other elements to the specific needs of the user. A feature that could be universal since it is accessible to all users but allowing an inclusive design and customisation according to the user's needs (e.g. using FACIL'iti for HTML pages).

### ***Future work***

Future work that may lead to improve levels of accessibility includes the following suggestions:

- Use available solutions to increase the compliance with standards or to improve user experience.
- Increase awareness of authors, contributors, publishers, and webmasters to create accessible documents and publications, as fully automated remediation solutions do not exist yet, and manual remediation is needed.

- Increase awareness of your readers by providing them a list of solutions and functionalities that they can use to improve user experience.

Improving the levels of accessibility remains a challenge and cannot be achieved overnight. However, the results and suggestions included in the two reports of RDDA project can be a first step in that direction. The guidelines and the reports could be used as a framework by European institutions and incentivise other organisations and private companies to improve the level of accessibility of their online content.



## 6. List of acronyms

This section presents the list of acronyms alphabetically arranged.

| <b>TABLE 14: LIST OF ACRONYMS</b> |   |
|-----------------------------------|---|
| <b>Acronyms</b>                   | <b>Definition</b>   |
| <b>ACT</b>                        | Accessibility Conformance Testing                           |
| <b>AI</b>                         | Artificial Intelligence                                     |
| <b>API</b>                        | Application Programming Interface                           |
| <b>CSS</b>                        | Cascading Styles Sheet                                      |
| <b>CTA</b>                        | Call To Action  |
| <b>DCI</b>                        | Digital Certainty Index                                     |
| <b>DG</b>                         | Directorate-General   |
| <b>EC</b>                         | European Commission   |
| <b>EDA</b>                        | European Dyslexia Association                               |
| <b>EP</b>                         | European Parliament   |
| <b>EPUB</b>                       | electronic publication                                      |
| <b>EU</b>                         | European Union  |
| <b>HTML</b>                       | HyperText Markup Language                                   |
| <b>ISCED</b>                      | International Standard Classification of Education          |
| <b>IT</b>                         | Information Technology                                      |
| <b>L</b>                          | List  |
| <b>LI</b>                         | List Item   |
| <b>LBody</b>                      | List Body   |
| <b>ML</b>                         | Machine Learning  |
| <b>OP</b>                         | Publications Office   |
| <b>OS</b>                         | Operating System  |
| <b>PAC</b>                        | PDF Accessibility Checker                                   |
| <b>PDF</b>                        | Portable Document Format                                    |
| <b>PDF/UA</b>                     | PDF/Universal Accessibility                                 |
| <b>PoC</b>                        | Proof of Concept  |
| <b>RDDA</b>                       | Reading Disability and Document Access, a possible approach |
| <b>SEO</b>                        | Search engine optimization                                  |
| <b>SO</b>                         | Strategic Objective   |
| <b>TOC</b>                        | Table of Content  |
| <b>TOCI</b>                       | Table of Content Item                                       |
| <b>URL</b>                        | Uniform Resource Locator                                    |
| <b>VPS</b>                        | Virtual Private Server                                      |
| <b>WAI</b>                        | Web Accessibility Initiative                                |
| <b>WCAG</b>                       | Web Content Accessibility Guidelines                        |

## 7. Annex

### 7.1. Original subset and references for all subsets

The original subset and all references for all subsets are presented in the attached Excel file named "OP\_RDDA\_Phase2\_Annex\_list of selected documents.xlsx". This file has 2 sheets representing:

- The "original subset" for one language and presenting for each document:
  - The type of document (PDF, HTML, EPUB),
  - The title of the document,
  - The date of publication,
  - The link to the document on OP Portal,
  - The number of page views,
  - The number of downloads,
  - The number of pages,
  - The tag if the document is to be remediated,
  - The tag if the document is to be assessed.
- The "references for all subsets" presenting for each document:
  - The reference of the document,
  - The subset to which the document belongs,
  - The language of the document,
  - The type of document (PDF, HTML, EPUB),
  - The title in English of the document,
  - The link to the document on OP Portal,
  - The link to the document on PoC environment,
  - The tag if the document is to be assessed,
  - The accessibility score with Siteimprove,
  - The accessibility score calculated based on PAC2021.

## 7.2. Generic survey

The generic survey is used on the original subsets, including PDF, HTML and EPUB documents and on remediated PDF documents by AbleDocs and a solution under development. This survey evaluates the accessibility level of a document and includes the following questions:

Q1. What is your personal ID?

Q2. The document being assessed is:

Q3. Which device are you doing this assessment on?

- Computer
- Tablet
- Mobile phone

Q4. Which internet browser are you currently using?

- Internet explorer
- Microsoft Edge
- Chrome
- Safari
- Firefox
- Android browser
- iOS Browser
- Other
- Not applicable

Q5. Are you using any personal assistive technology?

- Yes, if yes, please name the tool
- No

### **Section A**

*In this section, we are assessing if the design of the text is adapted to your needs.*

Q6. What is your first impression of the readability of this document as a whole?

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q7. Would you say that the font (shape/design) of the main text is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q8. Would you say that the font of the titles is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q9. Would you say that the font of the picture's caption is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q10. Any other comment on the font(s) in this document? :

Q11. Would you say that the size of the font of the main text is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q12. Would you say that the size of the font of the titles is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q13. Would you say that the size of the font of the picture's caption is ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q14. Would you say that the spacing between the letters of the main text is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q15. Would you say that the spacing between the letters of the titles is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q16. Would you say that the spacing between the letters of the picture's caption is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q17. Would you say that the spacing between the words of the main text is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q18. Would you say that the spacing between the words of the titles is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q19. Would you say that the spacing between the words of the picture's caption is

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q20. Would you say that the space between the lines in the main text is ...

- Very suitable
- Suitable
- Not suitable
- Not suitable at all
- Not applicable

Q21. The contrast and colours of the letters and the background make the document ...

- Very easy to read
- Easy to read
- Difficult to read
- Very difficult to read
- Not applicable

Q22. Any other comment on the text size, spacing, colours, in this document?

## **Section B**

*Now please consider the document as a whole, its organisation and layout.*

Q23. The structure of this document is ...

- Very easy to understand
- Easy to understand
- Difficult to understand
- Very difficult to understand

- Not applicable

### **Section C**

*Now please consider the content of the document and how easy it is to understand.*

Q24. The text is easy to understand?

- Yes, the text is very easy to understand
- Yes, the text is easy to understand
- No, the text is difficult to understand
- No, the text is very difficult to understand
- Not applicable

### **Section D**

Q25. Are you able to read the document easily with your device?

- Yes, there is no problem to open and read the document on my device
- There are some problems to open and read the document on my device
- I cannot open and read the document on my device
- Not applicable

Q26. Is this document compatible with your regular assistive technology tools?

- Yes, there is no problem to open and read the document with my usual tools
- There are some problems to open and read with my usual tools
- I cannot open and read this document with my accommodation tools
- Not applicable

### 7.3. FACIL'iti survey

This FACIL'iti survey is only used on FACIL'iti remediated documents. This survey evaluates the tool and the accessibility level of remediated documents using FACIL'iti and includes the following questions:

Q1. What is your personal ID?

Q2. The document being assessed is:

Q27. Please rate how easy you found FACIL'iti to setup.

- Very easy
- Easy
- Neither easy nor difficult
- Difficult
- Very difficult

Q28. Please rate the range of accessibility options on FACIL'iti

*where 1 is a poor range of options and 5 is an excellent range of options.*

- 1.
- 2.
- 3.
- 4.
- 5.

Q29. Did you find it easy to modify your profile and preferences on FACIL'iti?

- Very easy
- Easy
- Neither easy nor difficult
- Difficult
- Very difficult

Q30. Did using FACIL'iti make the document easier to read?

- Yes
- No
- Unsure

Q31. Could you use FACIL'iti easily on both computer and mobile device?

- Yes
- No
- Unsure

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms, and their related entities (collectively, the "Deloitte organization"). DTTL (also referred to as "Deloitte Global") and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see [www.deloitte.com/about](http://www.deloitte.com/about) to learn more.

Deloitte is a leading global provider of audit and assurance, consulting, financial advisory, risk advisory, tax and related services. Our global network of member firms and related entities in more than 150 countries and territories (collectively, the "Deloitte organization") serves four out of five Fortune Global 500® companies. Learn how Deloitte's approximately 330,000 people make an impact that matters at [www.deloitte.com](http://www.deloitte.com).

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms or their related entities (collectively, the "Deloitte organization") is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.

© 2021 Deloitte Tax & Consulting