

# DELIVERABLE

**Project Acronym:** EODOPEN

**Grant Agreement Number:** 607666-CREA-1-2019-1-AT-CULT-COOP2

**Project Title:** EODOPEN | eBooks-On-Demand-  
Network Opening Publications for  
European Netizens

**Project Website:** <https://eodopen.eu/>

---

## Activity A13: Compiling training material for delivery formats and training library staff

*D13a Workshops and trainings for librarians and interested  
communities on nationwide levels*

*D13b Educational material used in workshops and trainings*

---

**Author(s): Andreja Hari and Alenka Kavčič Čolić**

Project Co-funded by the Creative Europe Programme of the European Union		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the project partners and the Commission Services	

---

### DOCUMENT INFORMATION

---

Activity number: A13

Activity title: Compiling training material for delivery formats and training library staff

Contractual date of activity: M37-M48

Actual date of activity: M37-M60

Author(s): Andreja Hari and Alenka Kavčič Čolić

Contributor(s):

Working group: WG4

Working group title: Delivery formats of digitised material for special needs

Working group leader: National and University Library

Dissemination Level: P

---

### HISTORY OF VERSIONS

---

Version	Date	Status	Author (organisation)	Description/Approval Level
1	Dec. 2023-Aug. 2024	Draft	Andreja Hari (NUK) Alenka Kavčič-Čolić (NUK)	Draft of all educational modules
2	Sep. 2024	Final	Andreja Hari (NUK) Alenka Kavčič-Čolić (NUK)	Final version

---

## EODOPEN PROJECT SUMMARY

---

Libraries all over Europe face the difficult challenge of managing tremendous amounts of 20<sup>th</sup> and 21<sup>st</sup> century textual material that has not yet been digitised due to the complex copyright situation. These works cannot be accessed by the general public and are hidden deep in library stacks, as they are often out of print or have never even been printed at all, while reprints or facsimiles are out of sight.

The **EODOPEN** project focuses on making 20<sup>th</sup> and 21<sup>st</sup> century library collections digitally visible by **directly engaging with communities** in the selection, digitisation and dissemination processes. As a leading partner, the University Library of Innsbruck, joined by 14 European libraries from 11 nations, has set itself the goal of making 15,000 pieces of textual material digitally available, and of reaching more than one million people in Europe by 2024.

Among other goals, such as building a common portal to display the project outcomes, EODOPEN aims to stimulate interest in and improve access to 20<sup>th</sup> and 21<sup>st</sup> century textual material, including grey and scientific literature. EODOPEN continuously carries out social media campaigns in order to attract new audiences. Furthermore, the participating libraries establish contact with commemorative institutions all over Europe, as well as with researchers and doctoral study boards, history associations and local publishing houses, in order to obtain suggestions from a broad audience.

In collaboration with local institutions, all of the project partners select hidden library treasures, **deal with rights clearance questions** and put new content online, while dissemination activities display the digital content via international channels.

In addition, EODOPEN aims to provide alternative delivery formats suitable **for blind or visually impaired users**. An international survey gathers data from a broad European public about the use of e-books. By evaluating this data, the project broadens its scope to alternative delivery formats in order to fulfil the needs of **blind or visually impaired users**.

In order to promote best practice in rights clearance among the library community, EODOPEN provides handouts and tools to make 20<sup>th</sup> and 21<sup>st</sup> century books available beyond the project's lifetime. In this regard, the project partners cooperate closely to develop an online tool for the documentation of rights clearance, especially suited for out-of-print and orphan works. Interactive workshops investigate needs related to **dealing with rights clearance** questions in order to implement the requirements of the international community in establishing the online tool.

---

## ABSTRACT

---

This handbook compiles the experiences gathered during the implementation of the EODOPEN project from 2019 to 2024. It includes two key deliverables: D13a Workshops and Training for Librarians and Interested Communities on a Nationwide Level, and D13b Educational Material Used in Workshops and Trainings. The aim of this handbook is to equip librarians and other cultural workers with the skills, competence, and knowledge to provide services for print-disabled users, create alternative delivery formats, and understand copyright issues and the Marrakesh Treaty at a national level.

The target communities for the educational material are users of mobile devices and individuals who are blind or partially sighted. The document is organised to provide librarians from EODOPEN partner countries with guidance on topics relevant to digitisation processes: accessibility, copyright, and digitisation and delivery formats. Additionally, it offers detailed instructions with illustrations on how to create accessible document formats, including EPUB and HTML.

The handbook is divided into five modules. Each module explains the content, goals, and training scenarios for delivering the content, and the training material needed. They are supported with practical examples, specially produced video recordings and/or slides, and an evaluation test.

**Statement of originality:**

This educational material contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

# **TABLE OF CONTENTS**

<b>DOCUMENT INFORMATION .....</b>	<b>2</b>
<b>HISTORY OF VERSIONS .....</b>	<b>2</b>
<b>EODOPEN PROJECT SUMMARY .....</b>	<b>3</b>
<b>ABSTRACT .....</b>	<b>4</b>
<b>TABLE OF CONTENTS .....</b>	<b>5</b>
<b>INTRODUCTION AND PURPOSE .....</b>	<b>8</b>
<i>Module's Overview .....</i>	<i>9</i>
<b>I. MODULE: ACCESSIBILITY AND READING DIGITAL .....</b>	<b>10</b>
<i>About the Module .....</i>	<i>10</i>
Training Scenarios/How to Deliver the Content .....	10
Duration of the Module.....	10
Training Material Needed .....	11
1 <i>Accessibility and User Needs in Regard to Digital Objects.....</i>	<i>11</i>
2 <i>Print Disability.....</i>	<i>13</i>
2.1 <i>Why Care About Information Accessibility? .....</i>	<i>13</i>
2.2 <i>Ways that People with Print Disabilities Read .....</i>	<i>14</i>
3 <i>What is Assistive Technology? .....</i>	<i>15</i>
4 <i>Mobile Devices .....</i>	<i>15</i>
5 <i>Accessibility Standards.....</i>	<i>16</i>
5.1 <i>Accessibility Standards and Various Formats .....</i>	<i>16</i>
5.2 <i>Summary of Key Accessibility Principles.....</i>	<i>17</i>
6 <i>Further resources .....</i>	<i>21</i>
<b>II. MODULE: COPYRIGHT .....</b>	<b>23</b>
<i>About the Module .....</i>	<i>23</i>
Training Scenarios/ How to Deliver the Content .....	23
Duration of the Module.....	23
Training Material Needed .....	24
1 <i>European Legislation Relevant for Increasing Digital Accessibility in Libraries .....</i>	<i>24</i>
2 <i>Marrakesh Treaty .....</i>	<i>28</i>
2.1 <i>History of the Marrakesh Treaty .....</i>	<i>30</i>
2.2 <i>The Marrakesh Treaty and European Union legislation .....</i>	<i>32</i>
3 <i>Further resources .....</i>	<i>34</i>
<b>III. MODULE: DIGITISATION AND FORMATS .....</b>	<b>37</b>
<i>About the Module .....</i>	<i>37</i>
Training scenarios/ how to deliver the content .....	37
Duration of the module.....	37
Training material needed .....	37
1 <i>Digitisation in Libraries .....</i>	<i>38</i>
1.1 <i>Full Text Generation: Optical Character Recognition (OCR).....</i>	<i>41</i>
2 <i>Formats and Accessibility.....</i>	<i>50</i>
3 <i>Conversion .....</i>	<i>51</i>
4 <i>Further resources .....</i>	<i>57</i>
<b>IV. MODULE: CREATING ACCESSIBLE DOCUMENT FORMATS .....</b>	<b>58</b>
<i>About the Module .....</i>	<i>58</i>

Training Scenarios/How to Deliver the Content.....	58
Duration of the Course .....	58
Training Material Needed .....	59
Additional Notes.....	59
<b>1 Basic Word Accessibility.....</b>	<b>59</b>
1.1 OCR Clean-up and Reading Order .....	59
1.2 Style and Contrast .....	63
1.3 Headings and Formatting .....	64
1.4 Specifying the Language .....	68
1.5 Bulleted and Numbered Lists .....	69
1.6 Images and Image Descriptions.....	70
1.7 File Name, Properties and Exporting.....	73
<b>2 Advanced Word accessibility .....</b>	<b>76</b>
2.1 Page Numbering.....	76
2.2 Tables .....	78
2.3 Hyperlinks.....	82
2.4 Mathematical Expressions and Special Characters .....	83
2.5 Headers and Footers .....	86
2.6 Footnotes and Endnotes .....	86
2.7 Text Boxes, Charts and Objects .....	87
<b>3 Templates .....</b>	<b>89</b>
3.1 What is a Word Template?.....	89
3.2 Why Make the Templates Accessible? .....	89
3.3 What to Do with Templates .....	90
<b>4. The Microsoft Accessibility Checker.....</b>	<b>90</b>
4.1 Start Accessibility Checker.....	90
4.2 Understand the Inspection Results .....	92
<b>5 Further resources.....</b>	<b>92</b>
<b>V. MODULE: CREATING ACCESSIBLE EPUB AND HTML FORMATS .....</b>	<b>94</b>
<i>About the module .....</i>	<i>94</i>
Training Scenarios/How to Deliver the Content.....	95
Duration of the Course .....	95
Training Material Needed .....	95
Additional Notes.....	95
<b>1 What is needed before using the WordToEpub tool? .....</b>	<b>95</b>
<b>2 WordToEpub Preferences Before Converting the First Book.....</b>	<b>96</b>
2.1 Default Conversion Options .....	97
2.2 Special Markers .....	98
2.3 Metadata .....	98
2.4 Default Pagination .....	99
2.5 Default Cover Option.....	99
2.6 Styling .....	100
2.7 User Interface Options .....	100
2.8 Updates .....	101
2.9 Word Add-in and Writer Extension .....	101
<b>3 Converting the Book to EPUB.....</b>	<b>102</b>
3.1 Opening the Word Document in WordToEpub .....	102

3.2 Settings Before Conversion .....	102
3.3 During the Conversion and Conversion Report .....	109
4 <i>Editing EPUB in Sigil</i> .....	111
4.1 Editing Content .....	112
4.2 Editing Visual Appearance .....	116
4.3 Editing Metadata .....	117
4.4 Saving the File and Continuing the Workflow in WordToEpub .....	118
5 <i>Accessibility Check</i> .....	119
5.1 EpubChecker .....	119
5.2 Ace by DAISY .....	121
5.3 Thorium Reader and Read Aloud Test .....	124
6 <i>Converting the Book to HTML</i> .....	126
7 <i>Further resources</i> .....	128
<b>VI. SUMMARY .....</b>	<b>129</b>
<b>VII. ANNEX 1: ASSESSMENT QUESTIONNAIRES BY MODULES .....</b>	<b>131</b>
<i>Module 1: Accessibility and Reading Digital</i> .....	131
<i>Module 2: Copyright</i> .....	131
<i>Module 3: Digitisation and Formats</i> .....	131
<i>Module 4: Creating Accessible Documents</i> .....	132
<i>Module 5: Creating Accessible EPUB and HTML formats</i> .....	132

## **INTRODUCTION AND PURPOSE**

In Working Group 4 on delivery formats of digitised material for special needs, we analysed alternative delivery formats for two different communities: users of mobile devices and blind and partially sighted users. Based on our observations and previous research under Activities 11 and 12, we found that it is much easier to adapt different delivery formats for sighted users, even when accessing e-books through mobile devices, than for blind and partially sighted users.

During the EODOPEN project, we realised that librarians face significant challenges in providing access to blind and partially sighted users. Consequently, under Activity 13 of the project and in consultation with project partners, we decided to focus on the production processes to provide access to digitised content specifically for blind and partially sighted communities. We noticed that this area is crucial for many libraries, and with the increasing number of users in these communities, a compilation of training materials would greatly benefit our partners.

Additionally, we recorded a few training videos to facilitate easier content dissemination. We are grateful to the Daisy Consortium for allowing us to use and adapt some of their meticulously prepared educational materials. We believe that our educational material will equip librarians and other cultural workers with the skills, competence, and knowledge to provide services for print-disabled users, create alternative delivery formats, and address copyright issues and the Marrakesh Treaty at a national level.

Project Activity No. 13 includes two deliverables: D13a Workshops and training for librarians and interested communities on a nationwide level, and D13b Educational material used in workshops and trainings. D13a includes recorded videos, slides, and training sessions prepared and implemented for project partners. D13b, the educational material used in workshops and trainings, is a handbook containing information on how to deliver the relevant courses, showcasing the topics developed during the project. For each of the topics in D13b, links to the videos, and presentations' slides related to D13a contents are provided.

D13b will be available for all partners to translate into their national languages and to use or reuse for their own purposes through training courses or workshops. During the project's lifetime, we also organised several training courses covering topics related to the production of alternative delivery formats.

## Module's Overview

In our handbook, *Educational material used in workshops and trainings* (D13b) we developed the following five modules:

- I. Accessibility and Reading Digital
- II. Copyright
- III. Digitisation and Delivery Formats
- IV. Creating Accessible Document Format
- V. Creating Accessible EPUB and HTML File Formats

Modules I to III contain slides and video presentations while all the topics in Modules IV and V provide the possibility of self-learning due to their complexity.

Each module has the following structure:

1. Introduction on the module: goals to achieve, brief presentation of the content, and instructions on how to deliver the educational content and the material needed for the course
2. Presentation of sub-topics with examples
3. Assessment questionnaire for all modules - All assessment questionnaires are available in Annex 1 and could be used by the trainers and/or trainees as an assessment tool for the content learned in each module.
4. Further literature and sources.

# **I. MODULE: ACCESSIBILITY AND READING DIGITAL**

## **About the Module**

In this module, we explain the meaning of accessibility in the digital world, its importance, and how to improve it. Special emphasis is given to users of mobile devices and persons with print disabilities. There are different types of print disabilities, which require different approaches, including the use of assistive technologies. Furthermore, relevant accessibility standards are described.

The goals of this module are:

- To become acquainted with accessibility and user needs in regarding digital objects.
- To learn about print disability, why libraries should care about the accessibility to information.
- To find out how people with print disabilities read in digital space.
- To learn about assistive technologies that aid people with print disabilities in reading digitally.
- To understand the importance of mobile devices in accessing digital reading.
- To learn how accessibility standards empower digital reading and understand the key accessibility principles.

### **Training Scenarios/How to Deliver the Content**

- Training organised for 10 to 30 persons.
- Introduction to the course, integration of the group (10 minutes)
- Ask participants how much they know about print disabilities, assistive technologies and accessibility standards. (15 minutes)
- Describe accessibility and user needs regarding digital objects. (20 minutes)
- Describe print disability and the ways people with print disabilities read. (30 minutes)
- Describe assistive technology. If possible, have some available so participants can have first-hand experience. (30 minutes)
- Describe the topic of mobile devices. (10 minutes)
- Describe accessibility standards and key accessibility principles. (40 minutes)
- Ask participants to fill in the assessment questionnaire in Annex 1.

### **Duration of the Module**

The duration of the module is approximately 3-4 hours with a break in the middle.

### Training Material Needed

- Projector
- Speakers
- Slides
- Video recordings
- Assessment questionnaire in Annex 1

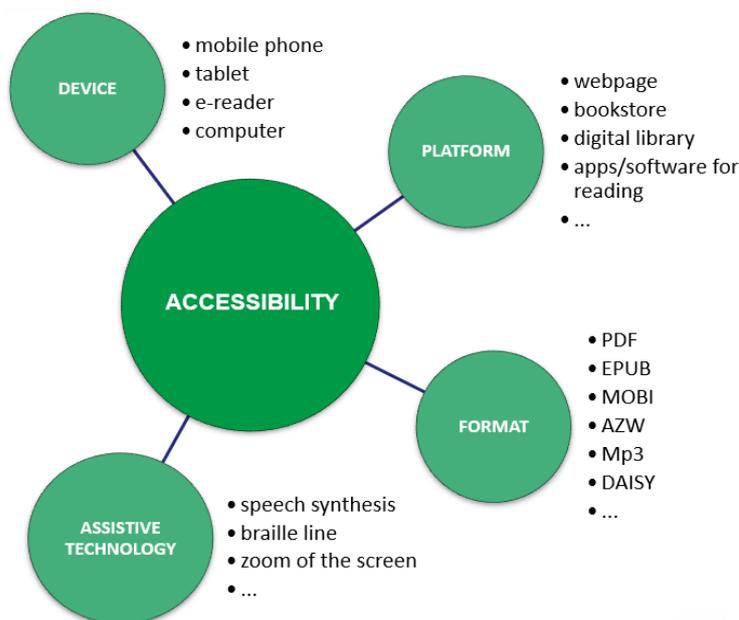
## 1 Accessibility and User Needs in Regard to Digital Objects

“The accessibility of publication depends on how suitable and easily adaptable it is to the different ways in which people read” (Ganner. J. et al., 2023). This applies to all users, not only those with special needs or with print disabilities, but also people with situational barriers (e.g., a broken arm, strong sunlight) and users of mobile devices. Enabling all users to adapt digital objects to their reading needs is becoming increasingly necessary. In the publishing field there is support for creating born-accessible publications.

In Europe, this is further encouraged by the European Accessibility Act (EAA), a directive that aims to improve accessibility of products and services, including e-books, e-readers, and e-Bookstores.

Accessibility is comprised of the words **access** and **ability**, meaning it is important to have access to an object, open or enter it, interact with or leave it, and possess the skills to engage with it. By providing people accessible publications and information, such as those held in a local library, all library users can have broader access to knowledge, regardless of their disabilities. This inclusivity allows everyone to integrate into society, whether in employment, leisure, political or cultural contexts. Currently, that ideal situation is far from reality; only about 5-7% of published information is fully accessible to people with disabilities, and only about 20% of webpages are accessible.

In the field of digital objects, we must consider the accessibility of not only the digital objects themselves (e.g., e-books) but also every element that contributes to reading experience. This includes the accessibility of the devices people use to read, the platform they use to obtain digital objects, the format and accessibility of the digital objects, and, in the case of people with disabilities, the assistive technology they use for reading (**Picture I-1**). For example, the experience of a person using a mobile phone to access a digitised book in PDF format from a digital library is different from that of a person using a computer to access the same book and reading it with speech synthesis as assistive technology.



**Picture I-1:** Four elements that contribute to the accessibility are presented - devices, platforms, formats, and assistive technology.

Libraries also create digital objects, either born-digital or as a result of digitisation. The accessibility of born-digital content needs to be ensured, while digitisation presents more challenges in providing the most accessible objects possible. This topic, on how to make digitised objects accessible in DOCX, EPUB and HTML formats, is covered in modules IV and V. The outcome of these two modules is adapted digitised objects that can be read on any device and support assistive technologies (see **Picture I-2**).



**Picture I-2:** Original physical publication and adapted digitised version which can be read on various devices, with various adjustments and with assistive technology.

If libraries wish to offer more accessible digitisation outputs to their users, at least two main groups of user needs should be considered: users with print disabilities and users of mobile devices. These groups are explained in the following chapters, but they should not be treated separately, as the users can fit into both categories. For example, a person with low vision reading on a mobile device or a person without print disabilities using a screen reader to listen to a book. The chapters will focus only on specific differences.

## 2 Print Disability

People with print disabilities are those who find it difficult to access printed content for various reasons. These include people who:

- Are blind.
- Have low vision or a perceptual or reading disability, such as dyslexia, and are unable to read printed works to the same degree as a person without an impairment or disability.
- Are unable, due to physical disability, to hold or manipulate a book or to focus or move their eyes for reading.
- Have a learning disability, which may create barriers in accessing printed materials.
- Are elderly and may have a combination of the above-mentioned barriers.

When referencing people with disabilities, it is important to use person-first language, whereby the person is not defined by their disability. So, refer to the person first, and the impairment or disability second. For example, say **persons who are blind** or **people with print disabilities** rather than **the blind** or **the print disabled**.

### 2.1 Why Care About Information Accessibility?

Persons with disabilities may face difficulties in reading documents created by you. Many documents contain hidden obstacles that can sometimes deny or restrict access to users with disabilities, particularly persons with blindness, low vision, colour blindness, reading disabilities, and certain mobility impairments.

People with disabilities use digital documents in different ways. Some may want to have it read aloud using Text-To-Speech software or read it on Refreshable Braille Devices, while others may want to magnify the text and change the foreground and background colours to suit their visual disability. Some readers may use the keyboard to navigate through your documents, while others may use touch, voice commands, a modified mouse, Head Stylus, or even Eye Tracking technology.

Keep in mind that the content you create will be consumed by people in different ways. If you do not consider this variety while creating content, many will find it hard or impossible to use your creations.

The World Health Organization (2011) estimates that about 15% of the world's population lives with some form of disability. Therefore, a significant number of people are likely to be denied the right to information if the content is not in a format that they can adapt to their own needs for reading. In addition to being the right thing to do, digital accessibility is also a regulatory requirement according to many international and domestic conventions and laws.

## 2.2 Ways that People with Print Disabilities Read

People with print disabilities use many creative approaches to access print that is otherwise difficult or impossible for them to read. These various solutions are typically achieved using digital publications and are sometimes described as reading with eyes, ears, or fingers!

For example:

- **Eyes:** Readers with low vision may use a form of magnification to see content in large print. A person with a physical disability may read the book visually, using clever eye-tracking technology to turn the pages. For example, [watch the video about one of the magnification software with some examples](#).
- **Ears:** Others may use their ears and listen to a narrated version of the publication (audiobook) or listen to a computer voice converting the text to audio (text-to-speech). For instance, [watch the video about reading a document using a screen reader](#).
- **Fingers:** Persons with blindness may read content in braille using their fingers. For instance, [watch the video about braille, reading physical braille books and digital reading with braille display](#).

Not all people with print disabilities read in the same way or face the same barriers when accessing digital documents. Their barriers can vary daily, influenced by fatigue or the specific reading problems they are facing. For people with vision impairments, these problems can include difficulty focusing on text, reduced contrast sensitivity, reduced field of vision, sensitivity to movement, perceptual differences or contextual factors. Due to these changing factors, it is important that these readers have the ability to adapt the visual presentation of the text according to their current best reading options. For example, they might read with a dark theme due to glare or light sensitivity, or change margins and spacing between lines, words, or characters due to a reduced field of vision or difficulty focusing on text. Additionally, assistive technology often enables them to read independently.

### 3 What is Assistive Technology?

Assistive or adaptive technology commonly refers to a device or piece of equipment that can improve the functional capabilities of individuals with disabilities. There are many types of assistive technology, but a few examples related to reading include:

- **Screen reading software**—used by readers who are blind or have low vision to read out text that is on the screen.
- **Read aloud**—used by someone with a reading disability (such as dyslexia) where the computer reads out the publication text while highlighting details on-screen to support reading and comprehension.
- **Screen magnifiers**—used by someone with low vision to enlarge everything on the display
- **Refreshable braille display**—the text of a publication is represented in braille on an electronic device.

### 4 Mobile Devices

Mobile devices have become an integral part of everyday life for Europeans, and we anticipate that they will increasingly dominate as the primary technology for accessing information, including digitised content, in the future.

According to Eurostat<sup>1</sup>, in 2023, 90% of internet users accessed the internet via mobile devices, with 85% using mobile or smartphones and 63% utilising tablets or laptops. Mobile devices account for approximately half of web traffic globally. In the last quarter of 2023, excluding tablets, mobile devices generated 58.67% of global website traffic, consistently maintaining around 50% since early 2017 and surpassing it in 2020<sup>2</sup>.

According to GSMA<sup>3</sup>, in 2022 there were approximately 496 million unique mobile subscribers in Europe, with a penetration rate of 90% of population. It is projected that in 2030, this number will rise to 507 million, with a penetration rate expected to reach 92%.

Due to rising usage of mobile devices, creating digital objects suitable for any device is important to consider. People who use various mobile devices often find it difficult to read digital documents if the content isn't responsive to the different screen sizes (reflowable).

---

<sup>1</sup> [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital economy and society statistics - households and individuals#Devices used to connect to the internet](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#Devices_used_to_connect_to_the_internet)

<sup>2</sup> <https://www.statista.com/statistics/277125/share-of-website-traffic-coming-from-mobile-devices/>

<sup>3</sup> <https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2023/11/GSMA-Mobile-Economy-Europe-2023.pdf>

Additionally, due to mobility and reading outdoors, they often face problems with reading in bright sunlight or other conditions. Having the ability to change the visual presentation of the text, similar to what was described for people with print disabilities, is essential.

## 5 Accessibility Standards

Catering to the needs of individuals with diverse requirements may seem daunting to the content creators. Thankfully, you don't have to consider the needs of every disability, nor do you have to test your products with all the assistive technologies.

There are globally accepted standards and best practices for creating accessible digital content. Some of the most adopted standards are [WCAG](#), [Section 508](#), and [PDF/UA](#). Adherence to any of these ensures that documents will be usable by everyone, including persons with disabilities, without any significant barriers.

The accessibility guidelines and best practices for creating digital documents are intended to achieve the following objectives:

1. **Creating a structured and navigable document** – It should be possible for all readers to easily identify and move to any position in the document, such as a chapter or subsection. Tables, lists, notes, etc. should be created using best practices instead of customised methods.
2. **Provision of text descriptions for graphical content** – Pictures, flow charts, and maps should include text descriptions so that visually impaired readers do not miss important aspects of the document.
3. **Providing an adaptable format that is marked up semantically** – Readers should be able to adapt the visual presentation of the document to suit their needs. The meaning of different text elements should be conveyed not only through visual presentation, e.g., colour, alignment, but also through the appropriate use of built-in styles.

### 5.1 Accessibility Standards and Various Formats

- **EPUB, HTML:** If you are creating e-books using the EPUB file format, follow the EPUB Accessibility Guidelines. If you are creating content that will be delivered via web browsers, conform to the Web Content Accessibility Guidelines (WCAG). Your target should be to comply with WCAG 2.1 Level AA. If delivering content via a custom desktop or mobile platform app, ensure the user interface is tested for accessibility. Manual testing using assistive technology will be essential.
- **PDF, PDF/UA:** If you are creating PDFs, ensure the files are properly tagged and conform to the PDF Accessibility Guidelines. PDF files can be further tested against the

WCAG, Section 508, or PDF/UA specifications. Testing and fixing accessibility errors requires additional knowledge and tools.

- **DOCX, RTF:** Use the Accessibility Checker withing Microsoft Office Word. More about this is explain in [Module IV](#), which will focus on making Word documents accessible.

## 5.2 Summary of Key Accessibility Principles

The following is a summary of the various accessibility principles found in the guidance for EPUB, PDF, and WCAG listed above. These principles remain the same whether you are creating web pages, e-books or apps.

### *Structure (TOC, page numbers)*

People want to be able to navigate to any section or page of their choice while reading. It is essential to provide a mechanism that allows the reader to quickly and easily navigate to any section, sub-section, or page number by properly formatting headings and the table of contents (TOC). Clear, well-formatted headings are crucial for ensuring your publications meet global accessibility standards. Many people use screen readers, which can navigate through properly created heading levels, although this type of navigation only works when the author/publisher uses appropriate heading styles.

The headings, sections, and sub-sections in the content should be formatted using the **heading styles** built into the authoring applications. If the content is being created using HTML or XML, the section and sub-section names should be coded with **heading tags** such as <h1>, <h2>, and so forth.

Always use heading styles in a logical order and do not skip levels. For example, heading 1 should always be followed by heading 2, and heading 2 should be followed by heading 3 (or another Heading 2). If this is not done, the document will not pass accessibility tests. Screen reader users may think they have missed a heading or may get confused if headings are not presented in a logical order.

Often, authors and designers make the section and sub-section names visually prominent using a bigger font, bold/underline style, or a different colour. This kind of styling does not work for assistive technology users.

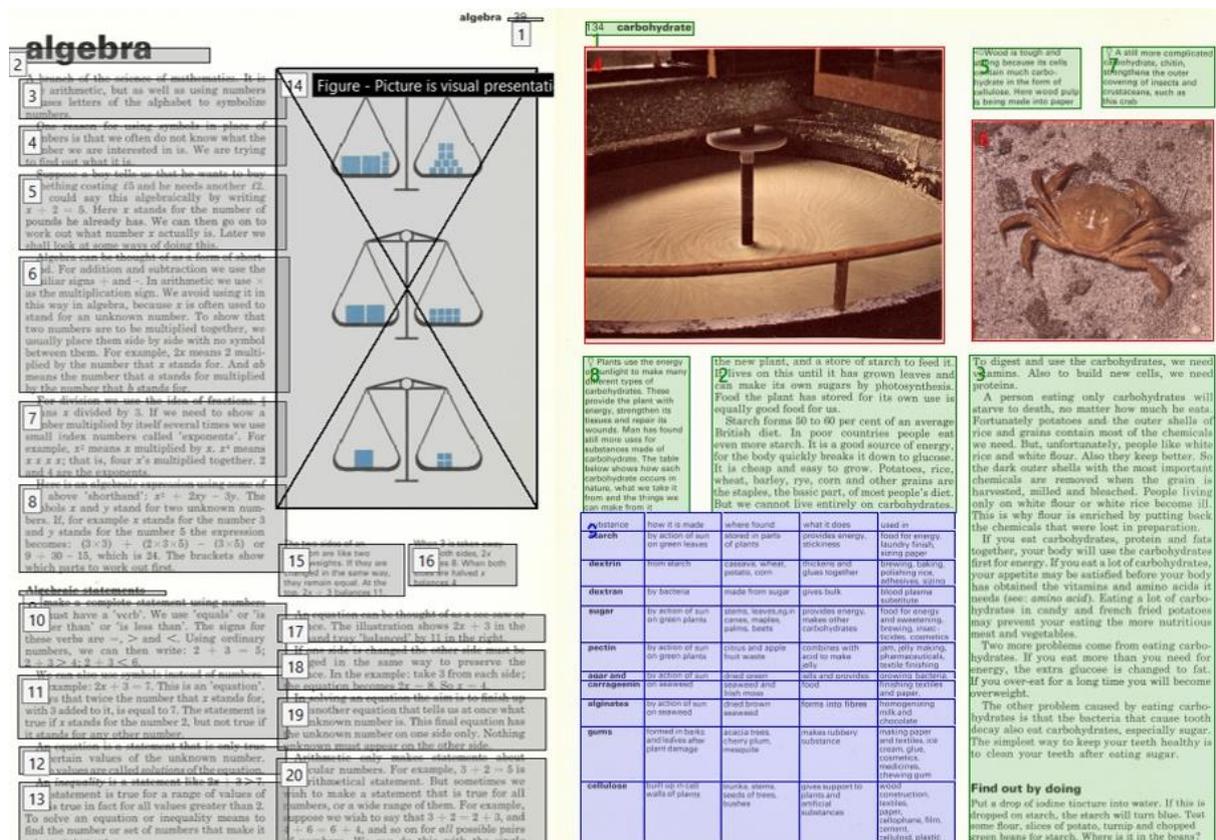
When sections and sub-sections are styled using appropriate heading styles/tags, reading apps can generate a table of contents or index that is highly useful to all users. A table of contents enables navigation to any part of the publication, keeping track of where they are and understanding the overall structure.

When print material is published in digital format, it is important to include the **print page number** in the e-book. If you are using the EPUB format, you will find that print page numbering is supported, making it possible to navigate to specific locations in the print book. This is important for students with print disabilities in a mainstream classroom, as it allows them to go to the same section in the digital publication while their peers use the print version of the same book.

A good, well-structured table of contents in a digital publication is appreciated by all users, irrespective of ability, as it allows everyone to save time by quickly jumping to the page or section they want, without needing to scroll through the book.

### Reading Order

For complex layouts, it's important to define the reading order for your material and to check that the user experience functions as intended. Some authoring applications mistakenly apply reading order based on the order in which content was edited.



**Picture I-3:** Two examples of working with reading order after or during digitisation: the first in Adobe Acrobat Pro and the second in Abby FineReader 15, with different type of elements.

The correct reading order of your material may not be exactly in line with the visual presentation. Consider how people should read the panels if they were placed in a single

column. Either of the examples in the image above could be valid reading orders depending on the content. You should always bear in mind that assistive technology needs to understand and guide the reader through the content, and that a logical reading order is imperative for these accessibility requirements.

### *Images*

Blind and low vision users cannot see and understand any non-textual content such as pictures, charts, or maps that you may have included in the publication. However, screen readers can read out a text description called Alt-Text (alternative text) provided by you in place of the graphical content or rich and complex material. In the absence of Alt-Text, any information presented within an image will not be available to readers who are blind or visually impaired.

#### Hints and Tips for Writing Alt-Text:

- Use Alt-Text to convey the important content or function of the object.
- The image's purpose and information need to be conveyed clearly, but as succinctly as possible. Rambling Alt-Text is not helpful.
- Be concise. Typically, a few words are all you need, although sometimes a short sentence or two may be required.
- Often, the object is described in the surrounding text. In such cases, your Alt-Text can be very brief. It should not repeat the information already provided in the document.
- Screen readers generally announce the type of content the Alt-Text is applied to, like an image or a table, so you do not need phrases like "image of" or "table of" in your description.
- Decorative images need not be described. Ensure that the Alt-Text is empty.
- Consider the context of the image. The Alt-Text should contain the information that the image conveys. Simply repeating information in the existing caption or associated text is not helpful.

A comprehensive set of [Image Description Guidelines are available from the DIAGRAM Center.](#)

If you think your audience needs more information, you can write an extended description. For complex objects like charts, people often write the longer description in the body of the text below the object or provide a link to the extended description at the end of the publication. It is also common to prefix the description in the body of the work with the words **Image description:** to provide context outside of the Alt-Text.

## *Tables and Charts*

It is important that you do not insert tables and charts into your content as images. Understanding the purpose of tables and charts (and when they should be avoided) will enable you to prepare them with accessibility in mind and help you to appreciate how they can be inaccessible.

Hints and Tips for Creating Accessible Tables:

- Keep the table as simple as possible.
- Avoid using merged cells, split cells, and nested tables as assistive technologies do not cope well with this level of complexity.
- If appropriate, designate a header row and/or header column.
- Do not insert an image of a table, as this will not be accessible. Use the correct table feature for the format being used.
- Define the table and column width in percentages so that they can adjust to different screen and page sizes.

## *Use of Colour*

Text colour alone should not be used to convey information in a document. People with visual disabilities, such as low vision and colour blindness, are likely to miss out on this information.

Wherever possible, use the heading styles. Alternatively, the coloured text can be underlined. If using colour in charts, supplement the colour coding with texture, different line styles, text in graphs, or shades of one colour to improve accessibility. Printing a colour document in black and white is the best test to see if you have lost any meaning.

Also, take care of colour contrast and avoid putting very similar colours on top of one another. Good contrast between the text and background colour makes the document easier to read for everyone, especially those with visual impairments.

## *Science, Technology, Engineering, Arts and Mathematics Publications*

Publications in the STEM (or STEAM) domains are often textbooks, reference titles or instruction manuals. In addition to more sophisticated structure (including features such as parts, sub-headings, references, and glossaries), they will often include more complex images, diagrams, and scientific expressions.

Techniques and technologies are now available to make even these complex titles accessible to the widest range of readers. For example, some formats support extended descriptions, used when Alt-Text is not sufficient to provide an equitable experience (an infographic, for example). Mathematical and chemical expressions can be included using industry-standard

techniques that render well on a conventional display and can also be explored by a reader using braille.

Best practices for including these types of contents continue to evolve and improve, making production simpler and more cost-effective, and enhancing the user experience.

### *Additional Points to Consider*

- Ensure that the font size is sufficiently large across the document. It's generally good practice to ensure a minimum size of approximately 11 point, 13-14px, or 0.90em.
- With formats like EPUB, it's important not to specify a font size at the base <p> text in the CSS<sup>4</sup>, as this can prevent users from changing the text size.
- Define the language of the content to ensure assistive technology uses the correct voice and pronunciation for the content, and the correct braille code is applied on a braille display.

## **6 Further resources**

Benetech. (2017). *Poet Image Description*. <https://poet.diagramcenter.org/>

Benetech. (2019). *Diagram Center*. <http://diagramcenter.org/>

DAISY Consortium. (30. 3. 2020). *Examining the Accessible Mobile Reading Revolution* [Video]. YouTube. <https://www.youtube.com/watch?v=v8d2PwfVTWg>

DAISY Consortium. (12. 5. 2020). *Publishing, accessibility, W3C standards – where are we and how did we get here?* [Video]. YouTube. <https://www.youtube.com/watch?v=PLSM-YnIYig>

DAISY Consortium. (15. 6. 2020). *The European Accessibility Act - Considerations for the publishing industry* [Video]. YouTube. <https://www.youtube.com/watch?v=H3PSCZtm3s>

DAISY Consortium. (27. 7. 2020). *The Art and Science of Describing Images* [Video]. YouTube. <https://www.youtube.com/watch?v=k9TkZEyeww>

DAISY Consortium. (8. 12. 2020). *The Art and Science of Describing Images Part 2* [Video]. YouTube. <https://www.youtube.com/watch?v=inXK7YfpN04>

DAISY Consortium. (16. 2. 2021). *The Art and Science of Describing Images - Part 3* [Video]. YouTube. <https://www.youtube.com/watch?v=A36MZfg5fIQ>

DAISY Consortium. (7. 12. 2021). *European Accessibility Act - Update* [Video]. YouTube. <https://youtu.be/70aOZTpEd58>

Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services. (2019). *EUR-Lex*,

---

<sup>4</sup> Cascading Style Sheets (CSS) is a stylesheet language used to style a document written in HTML or XML.

PE/81/2018/REV/1x. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882>

Ganner, J. et al. (2023). *Books without barriers: a practical guide to inclusive publishing*. Professional Editors and Australian Publishers Association.

<https://www.iped-editors.org/resources-for-editors/books-without-barriers/>

World Health Organization. (n. d.). World report on disability 2011.

<https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/world-report-on-disability>

World Wide Web Consortium (W3C). (7. 3. 2024). *Web content accessibility guidelines (WCAG) 2 overview*. <https://www.w3.org/WAI/standards-guidelines/wcag/>

## **II. MODULE: COPYRIGHT**

### **About the Module**

In this module, we present the legal context for providing access to copyrighted materials to people with visual impairments. For libraries, it is very important to be aware of the existing copyright regulations, as well as the limitations and exceptions concerning blind and partially sighted users.

The goals of this module are:

- To become acquainted with existing copyright regulations that could be applied to the use of digitised works.
- To understand the benefits of the Marrakesh Treaty for blind and visually impaired users in the library.
- To find out the limitations and exceptions concerning blind and partially sighted users.
- To learn how the Marrakesh Treaty has been implemented in the legislation of the European Union.
- To learn about other European legislation that concerns people with visual disabilities.

### **Training Scenarios/ How to Deliver the Content**

- Training organised for 10 to 30 persons.
- Introduction to the course, integration of the group (10 minutes)
- Ask participants how much they know about the legislation related to access to copyrighted works by disabled people. (10 minutes)
- Describe the existing copyright regulations in Europe. (20 minutes)
- Describe the existing obstacles access copyrighted works by blind and partially sighted users. (15 minutes)
- Explain how the Marrakesh Treaty, with its limitations and exceptions, helps people with visual impairments access copyrighted works. (20 minutes)
- Describe other European regulations that increase accessibility for people with visual disabilities. (15 minutes)
- Ask participants to fill in the assessment questionnaire in Annex 1.

### **Duration of the Module**

The duration of the module is 3 hours with a break in the middle.

### Training Material Needed

- Projector
- Prepared slides
- Assessment questionnaire in Annex 1

## **1 European Legislation Relevant for Increasing Digital Accessibility in Libraries**

The development of the information society in Europe in the late 90s had a significant impact on increasing accessibility to digital content in cultural organisations, including libraries. The World Wide Web became one of the most important channels for the publication of digital content, most of which represented reproductions of analogue publications. To foster access to information for all European citizens, digitisation was promoted and financially supported by the European Commission in the first decade of the second millennium. However, access to copyrighted works posed many problems, even for non-commercial purposes such as research and education.

The European Union sought to find solutions to copyright regulations that were not entirely justified for non-commercial use in research and education. The first important milestone in this field was the so-called **Copyright Directive**—Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society. This Directive introduced exceptions or limitations to reproduction rights in the following cases (Limitations and Exceptions..., 2020, p. 3-4):

- In respect of private use and non-commercial purposes upon fair compensation;
- In respect of specific acts of reproduction made by publicly accessible libraries, educational establishments or museums, or by archives, which are not for direct or indirect economic or commercial advantage;
- Use for the sole purpose of illustration for teaching or scientific research, as long as the source is indicated;
- Uses for the benefit of people with a disability, which are directly related to the disability and of a non-commercial nature, to the extent required by the specific disability;
- Quotations for purposes such as criticism or review;
- Use by communication or making available, for the purpose of research or private study.

The goal of the Copyright Directive was to harmonise exceptions and limitations to copyright across Europe. These exceptions were very broad and general, and were not mandatory for

European states. Each European country decided on the implementation of exceptions and limitations they considered relevant.

In 2005, the EU Parliament adopted the **Digital Libraries Initiative i2010** (Communication from the Commission of 30 September 2005 to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions – i2010: Digital Libraries [COM(2005) 465 final – Official Journal C 49 of 28.2.2008]). Its aim was to make Europe's cultural, audiovisual, and scientific heritage accessible to all. The initiative combines cultural diversity, multilingualism, and technological progress. It addressed three priority areas: online accessibility, digitisation of analogue collections, and the preservation and storage of digital content. Additionally, it focused on scientific information.

In 2006, following the i2010 and Copyright Directive, the European Commission adopted the **Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation** (2006/585/EC) (Recommendation, 2006). The Recommendation put major emphasis on digitisation in libraries, online accessibility, and preservation. It also adopted licencing mechanisms in areas such as orphan works and out-of-print or out-of-distribution works. The Recommendation promotes the development of the European Digital Library.

The need for joining efforts to develop a European Digital Library was very vivid in the first decade of this millennium. By that time, Europe was searching for joint values and how to reflect the different nations and cultures that were essential to European States. Kolding Nielsen (2007) reports on the findings of CENL investigation, which showed that by 2007 only 1% of the holdings of European national libraries have been digitised (approximately 4.7 million of items or 17 million of pages), mainly newspapers, special collections, and rare, fragile, or heavily used materials like manuscripts or rare books. The priority of digitisation was access, not preservation.

In November 2008, funded by the European Commission, the European Digital Library Network (EDLnet) prototype was launched. It was succeeded by Europeana in February 2009.

EDLnet provided access to 4.5 million digital objects, which increased in Europeana to 10 million of digital objects in 2010. These were in the public domain.

By this time, the Google Books project was spreading through European States, raising problems of copyright of digitising libraries. Consequently, the European Commission conducted a consultation on how best to operate a digital library of Europe's scanned-in books (Masons, 2009).

In 2014, in the case of the Technische Universität Darmstadt vs. Eugen Ulmer, the Court of Justice of the European Union ruled that "... libraries in EU member states may digitise books in their collections without the consent of copyright owners, allowing the work to be available

at on-site electronic reading points ...” (EU Libraries ..., 2014; Essers, 2014), although without allowing users to print out the works or store them on USB stick.

Special attention was called to **orphan works** and **out-of-commerce works**. “Orphan works’ are works in which copyright still exists, but where the rightholder, whether it be the creator of the work or his/her successor in title, cannot be identified or located. This therefore limits the potential reach and impact of the works to a variety of audiences who may benefit from these works, such as engagement with works in a new, digital format.” (Mass digitisation ..., 2019). Out-of-commerce works are also in copyright works, but they are not any more available for purchase.

**The Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works**<sup>5</sup> permitted the making of orphan works available to the public, its reproduction, digitisation, indexing, cataloguing, preservation and restoration. Major emphasis was given on diligent search.

All adopted European regulations until 2013 did not include exceptions and limitations that would help access digital and digitised contents for users with visual impairment. In 2013 in Marrakesh, the World Intellectual Property Organization (WIPO) organized a meeting at which members adopted **The Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired, or Otherwise Print Disabled (Marrakesh Treaty)**. According to this Treaty, for accessing digital contents by blind and partially sighted exceptions and limitations to copyright were permitted: making a copy in an accessible format, distribution of the accessible copy, and cross-border exchange of the accessible copy. (More on Marrakesh Treaty in next chapter).

Copyright legislation needed further harmonisation among European Member States. This task was meant to be solved by the **Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market** and amending Directives 96/9/EC and 2001/29/, known as the **Digital Single Market Directive**. This Directive included new exceptions to copyright for the following research, innovation and education purposes:

- Text and data mining of all legally accessed materials in libraries.
- Online and cross-border teaching.
- Preservation of cultural heritage.
- Digitisation and making available of out-of-commerce works - After the rather unsuccessful efforts to increase access to orphan works, out-of-commerce works are approached differently in this Directive: an exception to copyright will apply unless collective management organisations representative of creators in a particular

---

<sup>5</sup> <https://eur-lex.europa.eu/eli/dir/2012/28/oj>

category of work can give licenses. In this case, they will be able to give licenses to cultural heritage institutions for the digitisation and making available of works not only by their members, but also by non-members.

- Protection of works in the public domain.
- Free access to clippings and extracts from press publications for non-profit or public educational institutions.
- Use of protected content by online content-sharing service providers.

However, the Digital Single Market Directive did not address or include the following issues that remain open to resolution (Limitations and exceptions, 2020):

- Access to in-copyright content by people with disabilities. Libraries should have the ability to reproduce publications, converting them into accessible formats. Only Member States that are signatories of the Marrakesh Treaty have adopted the exceptions and limitations for access to in-copyright materials by blind and partially sighted people. However, this is usually limited to specialised libraries for the blind and partially sighted or specialised institutions.
- Reproduction of in-copyright contents for data analytics, content/text data mining. This requires amendments to the Copyright Directive and Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases to ensure lawful access to information to citizens, the possibility to analyse that information by a computer, to extract facts, data and findings, reutilise and share them anywhere, and the ability to circumvent the technical protection measure (TPM) which is blocking legal access to the in-copyright work for the mentioned purposes.
- eLending: to ensure that European citizens may purchase or lend any e-book, and that authors are fairly remunerated through the public lending right where it is not included in the licence fee paid to publishers by libraries. In addition, this should include digitisation and lending of any book if it is not available electronically.
- The duration of copyright laws in European states should be harmonised. Currently, it varies from state to state for different copyrighted materials.
- Cross-border exchange of copies of copyrighted materials for research purposes.
- Making digital copies available on the library premises for access on different devices.
- Mass digitisation of out-of-commerce works, including cross-border cooperation. The DSM Directive proposes access to out-of-commerce works through licensing schemes on a Member State level between users and Collective Management Organisations.
- Revision of Orphan Works Directive concerning artistic works embedded in another copyright work is needed.
- Exceptions for research copying should apply to both commercial and non-commercial research.

- Replacement collection copies: When an item is destroyed, damaged, or digitally corrupted, a library should be able to source replacement copies anywhere within the EU.
- Harmonisation across Member States of copyright related to access to in-copyright contents for educational purposes.
- For educational, research, and cultural heritage purposes, people should have the automatic right to circumvent TPM where TPMs prevent the enjoyment of activities lawful under copyright exception. (LIBER, 2020, p. 14)

## 2 Marrakesh Treaty

“Copyright presents three main obstacles. First, the making of a copy in an accessible format, such as braille, could infringe the reproduction right. Second, the distribution of the accessible copy could infringe the distribution or making available to the public right. Third, the cross-border exchange of the accessible copy could infringe the importation or exportation right. Overcoming the copyright obstacle requires exceptions.” (Dreyling and Hackett, 2022, p.248)

**The Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired, or Otherwise Print Disabled** (Marrakesh Treaty) is an international treaty adopted in 2013 by the World Intellectual Property Organization (WIPO). “It is the first WIPO treaty to focus on user rights, and the first copyright treaty to include a clear human rights perspective.” (Dreyling and Hackett, 2022, p.246). Its primary goal is to improve access to printed materials for people with print disabilities by facilitating the production and exchange of accessible formats across borders. It also recognises the need to maintain a balance between the protection of the rights of authors and the larger public interest, particularly education, research and access to information, and highlights that such a balance must facilitate effective and timely access to works for the benefit of persons with print disabilities.” (Dreyling and Hackett, 2022, p. 255)

According to the Marrakesh Treaty “...people with print disabilities include those who are blind, those who have a visual impairment or a perceptual or reading disability and those who are unable to focus the eyes or to hold or manipulate a book.” (Coates et al., 2018, p.10) “An accessible format is any format that allows a print-disabled person to read the work as comfortably as a person without a disability.” (Coates et al., 2018, p.13) Examples of accessible formats are Braille, large print, and books in audio form, audio-only format as DAISY books, full-text DAISY books, EPUB3, etc.

More than 90% of all published materials are not accessible to blind and partially sighted individuals (Marrakesh Treaty, 2024). Only about 7% of published works are made available in

accessible formats for print-disabled people. In developing countries, the percentage is lower (less than 1%) (Coates et al., 2018, p.10).

The Marrakesh Treaty seeks to remove copyright barriers that prevent access to print works for print-disabled people. It contains an “... express reference to two widely adopted human rights treaties, the UN Declaration of Human Rights and the UN Convention on the Rights of Persons with Disabilities, placing human rights objectives, non-discrimination and equal opportunities at the centre of the treaty.” (Dreilyng and Hackett, 2022, p. 254-255). It consists of a preamble, 22 articles and 13 agreed statements.

The Treaty defines authorised entities and beneficiaries. The authorised entities are those authorised or recognised by the government to provide education, training, or information access to beneficiary persons on a not-for-profit basis. The beneficiary persons are blind and visually impaired individuals and persons with physical disability that prevents them from holding or turning the pages of a book.

The Treaty applies to published literary and artistic works in the form of text, notation, or illustrations, including in audio form, such as audiobooks. In addition, it also refers to educational multimedia DVDs. Films are not included.

Key provisions of the Marrakesh Treaty are stated in Articles 4 through 7, and they include:

1. **Limitations and Exceptions:** The treaty encourages countries to adopt limitations and exceptions to copyright laws to allow the creation and distribution of accessible format copies without seeking permission from copyright holders.

Authorised entities shall be permitted, without the authorisation of the copyright rightsholder, to make an accessible format copy of a work, obtain an accessible format copy from another authorised entity, and supply those copies to beneficiary persons by any means, including by non-commercial lending or by electronic communication by wire or wireless means, and undertake any intermediate steps to achieve those objectives, when all of the following conditions are met (Marrakesh Treaty, 2015, p.9):

- The authorized entity wishing to undertake said activity has lawful access to that work or a copy of that work.
- The work is converted to an accessible format copy, which may include any means needed to navigate information in the accessible format, but does not introduce changes other than those needed to make the work accessible to the beneficiary person.
- Such accessible format copies are supplied exclusively to be used by beneficiary persons.

- The activity is undertaken on a non-profit basis.
2. **Cross-Border Exchange:** It allows for the import and export of accessible format copies between countries without the need for permission from copyright holders, thereby addressing the issue of duplication of efforts and ensuring wider access to accessible materials.
  3. **Accessibility Standards:** The treaty promotes the use of internationally recognised standards for accessible formats, making it easier for organisations and individuals to produce materials that are compatible with a variety of assistive technologies.
  4. **Safeguards for Copyright Holders:** The Marrakesh Treaty includes provisions to safeguard the interests of copyright holders, ensuring that the exceptions and limitations provided do not unduly interfere with their rights.

### 2.1 History of the Marrakesh Treaty

(Dreyling and Hackett, 2022)

- 1977 - The World Council for the Welfare of the Blind, with the support of Brazil, first raised the issue at WIPO and UNESCO.
- Creation of a joint WIPO and UNESCO Working Group to consider the use of exceptions in international copyright conventions; IFLA was represented in this Working Group.
- The Working Group adopted a model of limits and exceptions for the benefit of persons who are blind or visually disabled (WIPO, 1982).
- 1994 - Adoption of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), establishing the World Trade Organization.
- TRIPS made the adoption of strict levels of IP protection at national level, a pre-condition for participation in the global trading system.
- A global civil society movement emerged on access to knowledge that identified WIPO, in particular, as an important forum in need of change (Franz, 2010).
- 2000 - IFLA and World Blind Union (WBU) met the WIPO Secretariat in Geneva to discuss the potential for widening access to accessible format material.
- 2001 - At IFLA WLIC, IFLA and WBU called for an international treaty for persons who are blind and visually impaired.
- 2004 - Joint policy statement: Policy Position Agreed by the WBU, the DAISY Consortium, and IFLA Libraries for the Blind Section 2004).
- 2006 - The United Nations Convention on the Rights of Persons with Disabilities was adopted and hailed as the first human rights treaty of the 21st century. It represents a major step in ensuring that the rights of disabled people are protected. The purpose of the Convention is to promote, protect, and ensure the full and equal enjoyment of

all human rights and fundamental freedoms by all persons with disabilities, and to promote respect for their inherent dignity.<sup>6</sup>

- IFLA and EIFL produced a draft Treaty on Access to Knowledge and Technology (CPTech, 2005b) with provisions in favour of persons with disabilities, as well as libraries, education, and other public interest purposes.
- The Access to Knowledge Coalition (A2K) issued the Geneva Declaration on the Future of the World Intellectual Property Organization (CPTech 2005a) – advocating for more flexibility regarding IP protection.
- 2007 - The Geneva Declaration kickstarted a discussion among WIPO Member States, which ended in the adoption of a Development Agenda by WIPO – promoting balanced IP system.
- 2008 - Limits and exceptions were formally established on the agenda of the IFLA Standing Committee on Copyright and Related Rights (SCCR).
- May 2009 - Brazil, Ecuador, and Paraguay formally introduced a proposal at SCCR for a treaty relating to limits and exceptions for persons who are blind and visually impaired, based on a text prepared by WBU – to improve access to protected works.
- Five-year negotiations among IFLA, EIFL, the Library Copyright Alliance, The Canadian Library Association, and the Deutscher Bibliotheksverband/German Library Association faced strong opposition from right holders.
- European civil society groups focused their advocacy on the European Parliament to put pressure on the European Commission, which had the negotiating mandate at WIPO. The EBU worked closely with members of the European Parliament.
- The European Parliament adopted a resolution in support of an international treaty, which influenced the Commission to relax its position (European Parliament 2012; New 2011).
- December 2012 - WIPO General Assembly agreed to convene a diplomatic conference in Marrakesh, Morocco.
- 17-28 June 2013 - A diplomatic conference, a specially convened event, was held to negotiate the final stages of a treaty (WIPO General Assembly 2013).
- The Treaty met the key demands set by the WBU. It also gave libraries a key role in the successful implementation of a landmark treaty, providing a new opportunity to help end the book famine.

---

<sup>6</sup> <https://learning.daisy.org/?lesson=the-united-nations-convention-on-the-rights-of-persons-with-disabilities-crpd>

## 2.2 The Marrakesh Treaty and European Union legislation

The European Union has provided a strong legal framework for ensuring legal certainty and the protection of rightsholders. The Directives 96/9/EC, 2001/29/EC, 2006/115/EC, and 2009/24/EC of the European Parliament and Council harmonise the rights of rightsholders in the area of copyright and related rights. Together with the Directive 2012/28/EU, they provide exceptions and limitations to those rights that are in the public interest, offering a fair balance of rights and interests between rightsholders and users.

However, the mentioned Directives do not remove accessibility barriers for blind, visually impaired, or otherwise print-disabled persons. The European Union (EU) has implemented the provisions of the Marrakesh Treaty through its own regulations. In September 2017, the EU passed the Directive (EU) 2017/1564 on certain permitted uses of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired, or otherwise print disabled.

After the adoption of the Marrakesh Treaty in 2013, the European Parliament has passed two important legislative acts:

1. **Directive (EU) 2017/1564** of the European Parliament and of the Council of 13 September 2017 on certain permitted uses of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print-disabled, and amending **Directive 2001/29/EC** on the harmonisation of certain aspects of copyright and related rights in the information society.

The Directive aims to further harmonise Union law applicable to copyright and related rights within the framework of the internal market, by establishing rules on the use of certain works and other subject matter without the authorisation of the rights holder, for the benefit of persons who are blind, visually impaired, or otherwise print-disabled.

This directive mandates EU member states to adopt legislation that enables the creation and distribution of accessible format copies of copyrighted works for persons with print disabilities without the need to obtain prior authorisation from copyright holders. It aligns the EU's legal framework with the provisions of the Marrakesh Treaty and ensures consistent implementation across EU member states.

2. **Regulation (EU) 2017/1563** of the European Parliament and of the Council of 13 September 2017 on the cross-border exchange between EU and third countries of accessible format copies of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired, or otherwise print-disabled.

The Regulation lays down uniform rules on the cross-border exchange of accessible format copies of certain works and other subject matter between EU and third countries that are parties to the Marrakesh Treaty without the authorisation of the rights holder, for the benefit of persons who are blind, visually impaired, or otherwise print-disabled, within the field harmonised by Directives 2001/29/EC and (EU) 2017/1564, to prevent jeopardising the harmonisation of exclusive rights and exceptions in the internal market.

The Directive and Regulation for the implementation of the Marrakesh Treaty in EU law were published in the Official Journal on 20 September 2017 and entered into application on 12 October 2018.

Additional regulations aimed at increasing the accessibility of works for print-disabled users are:

1. **Directive 2019/790 (Copyright Directive):** This directive, adopted in 2019, updates EU copyright rules to better reflect the digital age. It includes provisions related to the accessibility of works for persons with disabilities. Article 7 of the directive requires member states to ensure that cultural heritage institutions make their collections accessible to persons with disabilities.
2. **EU Web Accessibility Directive:** While not specifically related to copyright, this directive requires public sector websites and mobile applications to be accessible to all users, including those with disabilities. It aims to ensure that people with disabilities can fully participate in the digital society and access online information and services.
3. **European Accessibility Act (EAA):** This act, adopted in 2019, aims to improve the accessibility of products and services across the European Union, including digital content and technologies. While it primarily focuses on physical products and services, it also addresses the accessibility of digital content, which can benefit blind and partially sighted users.

These regulations and directives, along with **Directive 2017/1564**, form part of the broader legal framework within the European Union aimed at improving access to printed materials for people with print disabilities and providing them equal access to cultural, educational, and digital resources.

Libraries can make their accessible format works discoverable for other libraries through metadata in the catalogue record and by sharing information with other libraries. The accessible book service should be based on a not-for-profit basis.

### 3 Further resources

3 reasons why libraries should care about the EU-Digital Single Market Directive. (2019).

<https://www.ifla.org/news/3-reasons-why-libraries-should-care-about-the-eu-digital-single-market-directive/>

Coates, J., Felsmann, C., Hackett, T., Keninger, K., Martinez Calvo, F., Owen, V., Taylor, A., Pereyaslavskaya, K., Berg, F.; Owen, V. (Ed.). (2018). *Getting started. Implementing the Marrakesh Treaty for persons with print disabilities. A practical guide for librarians.*

EIFL. Available at: <https://www.eifl.net/resources/getting-started-implementing-marrakesh-treaty-persons-print-disabilities-practical-guide>

Commission Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation (2006/585/EC) [Official Journal L236 of 31. 8. 2006]

[https://www.eumonitor.eu/9353000/1/i4nvhdhfc8bljza\\_j9vvik7m1c3gyxp/vik0kdvwcbz1?layout=print&printmo=1](https://www.eumonitor.eu/9353000/1/i4nvhdhfc8bljza_j9vvik7m1c3gyxp/vik0kdvwcbz1?layout=print&printmo=1)

Communication from the Commission of 30 September 2005 to the European Parliament, the council, the European Economic and Social Committee and the Committee of the Regions – i2010: digital libraries [COM(2005) 465 final – Official Journal C 49 of 28.2.2008]. Source: i2010: Digital libraries <https://eur-lex.europa.eu/EN/legal-content/summary/i2010-digital-libraries.html>

Cultural heritage: digitisation, online accessibility and digital preservation. Consolidated Progress Report on the implementation of Commission Recommendation (2011/711/EU) 2015-2017. (2018). European Commission.

<https://www.digitalmeetsculture.net/wp-content/uploads/2019/06/ReportonCulturalHeritageDigitisationOnlineAccessibilityandDigitalPreservation.pdf>

Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works Text with EEA relevance <https://eur-lex.europa.eu/eli/dir/2012/28/oj>

Directive 2019/790/EU of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019L0790>

Dreyling, J. and Hackett, T. (2022). Success for People with Print Disabilities: The Marrakesh Treaty. In: Coates, J., Owen, V. and Reilly, S. (2022). *Navigating Copyright for Libraries: Purpose and Scope*, Berlin, Boston: De Gruyter Saur, pp. 245-271.

<https://doi.org/10.1515/9783110732009>

- Essers, L. (2014). Libraries may digitize books without permission, EU top court rules. PCWorld. <https://www.pcworld.com/article/435213/libraries-may-digitize-books-without-permission-eu-top-court-rules.html>
- European Directive on Copyright in the Digital Single Market: what it is about and why libraries should care. (2019). <https://www.ifla.org/news/european-directive-on-copyright-in-the-digital-single-market-what-it-is-about-and-why-libraries-should-care/>
- EU consults on copyright problems of digitising libraries (2009). Pinsent Masons: <https://www.pinsentmasons.com/out-law/news/eu-consults-on-copyright-problems-of-digitising-libraries>
- EU Libraries May Digitize Books Without Permission. (2014). Liberties. Accessed on 18 June 2024 at: <https://www.liberties.eu/en/stories/cjeu-libraries-may-digitize-books/2444>
- Getting Marrakesh Ratification Right – a Toolkit. [2019]. IFLA. [https://cdn.ifla.org/wp-content/uploads/2019/05/assets/clm/publications/marrakesh\\_toolkit.pdf](https://cdn.ifla.org/wp-content/uploads/2019/05/assets/clm/publications/marrakesh_toolkit.pdf)
- Kolding Nielsen, E. (2007). Digitization of library material in Europe. Problems, obstacles and perspectives anno 2007. DF Revy, 8 december 2007, pp. 10-12.
- Limitations and Exceptions in EU Copyright Law for Libraries, Educational and Research Establishments: A Basic Guide. (2020). Liber. <https://libereurope.eu/wp-content/uploads/2020/11/A-Basic-Guide-to-Limitations-and-Exceptions-in-EU-Copyright-Law-for-Libraries-Educational-and-Research-FINAL-ONLINE.pdf>
- The Marrakesh Treaty : An EIFL guide for libraries. (2015). Vilnius: EIFL. Available at: <https://www.eifl.net/resources/marrakesh-treaty-eifl-guide-libraries-english>
- Marrakesh Treaty: Marrakesh Treaty Ratification and Implementation Campaign. (2024). Toronto: World Blind Union. Available at: <https://worldblindunion.org/programs/marrakesh-treaty/>
- Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired, or Otherwise Print Disabled. (2013). WIPO.
- Mass digitisation of cultural heritage in Europe. (2019). <https://www.bournemouth.ac.uk/research/projects/mass-digitisation-cultural-heritage-europe>
- Owen, V. (2018)
- Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation (2006/585/EC) [Official Journal L236 of 31. 8. 2006] [https://www.eumonitor.eu/9353000/1/i4nvhd fcs8bljza\\_i9vvik7m1c3gyxp/vik0kdvwc\\_bz1?layout=print&printmo=1](https://www.eumonitor.eu/9353000/1/i4nvhd fcs8bljza_i9vvik7m1c3gyxp/vik0kdvwc_bz1?layout=print&printmo=1)
- Scollo Lavizzari and C., Tran, Q. (2016). IPA Guide to the Marrakesh Treaty. A WIPO Treaty to Facilitate Access to Published Works by Visually Impaired Persons and Persons with Print Disabilities, of 27 June 2013. A guide for publishers and their trade associations

in membership with the IPA. <https://internationalpublishers.org/wp-content/uploads/2023/09/ipa-guide-to-the-marrakesh-treaty.pdf>

World Blind Union – Marrakesh Treaty Advocacy Course. *Daisy Learning*. (s. a.). Daisy Consortium. Available at: <https://learning.daisy.org/?course=world-blind-union-marrakesh-treaty-advocacy-course>

## **III. MODULE: DIGITISATION AND FORMATS**

### **About the Module**

In this module, we present the topic of digitisation, output delivery file formats, and their conversion. The module is additionally based on the results from the EODOPEN project research work done by the National and University Library (Slovenia) and contains practical examples.

The goals of this module are:

- To become aware of the difficulties that could arise especially in mass digitisation.
- To help planning digitisation projects.
- To understand the significance of output file formats.
- To be acquainted with conversion tools.

### **Training scenarios/ how to deliver the content**

- Training organised for 10 to 30 persons.
- Introduction to the course, integration of the group (10 minutes)
- Ask participants how much they know about the digitization, different delivery file formats and their conversion. (10 minutes)
- Describe the digitization in libraries. (20 minutes)
- Describe the delivery file formats also in regard to accessibility. (20 minutes)
- Describe about file format conversion and some of the most often used tools. (15 minutes)
- Ask participants to fill in the assessment questionnaire in Annex 1.

### **Duration of the module**

The duration of the module is approximately 2 hours with a break in the middle.

### **Training material needed**

- Projector
- Prepared slides (attached)
- Video recordings
- Assessment questionnaire in Annex 1

## 1 Digitisation in Libraries

The first digitisation projects in libraries started in the 1990s. This was when we got the World Wide Web, which provided new forms of access to library digital holdings. Initially, digitisation was done manually, page by page, with slow digital cameras or flat-bed scanners. It was a time of learning, and most of the staff involved in digitisation dedicated a lot of time to each scan.

After 2000, digitisation became part of many national and European strategies. The digital libraries initiative i2010, adopted on 30 September 2005 by the European Commission, seeks to optimise the benefits of the new information technologies for economic growth, job creation, and the quality of life of European citizens. The Commission made digital libraries a key aspect of i2010. This initiative addressed three priority areas: online accessibility, digitisation of analogue collections, and the preservation and storage of digital content. Additionally, it combined cultural diversity, multilingualism, and technological progress.

In the early years, digitisation was strongly supported by the European Union, both legally and financially. As a consequence, quantity was prioritised over quality, and there were no guidelines or standards for this task.

In 2008 the European Digital Library Network was developed, later becoming Europeana in and constituting in an institution known as Europeana Foundation. It has had a significant impact in digitisation through the adoption of quality and interoperability standards.

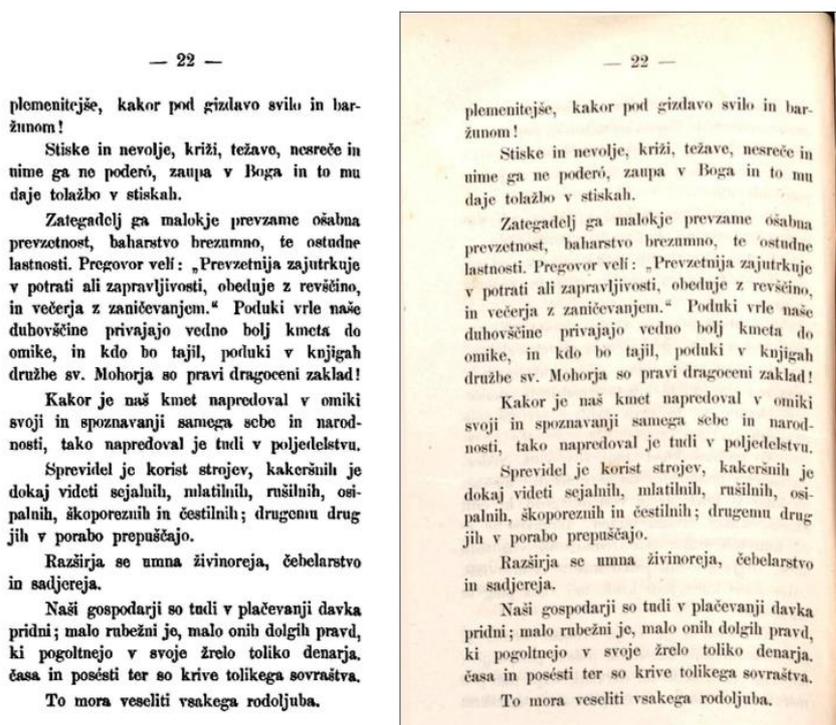
This document is not intended to cover the basic aspects of digitisation. There are digitisation handbooks, available on the web (Youngs, 2016; 5 Easy Steps, s.a.; Digitisation Toolkit, s. a.; Scientific Library Services, 2009; etc.). All of them recommend clarifying the aims, objectives, and target audiences of the digitisation project, the type of material to be digitised, and the minimal resources needed to produce the digital collection. All these should be addressed as part of the overall organisational strategy.

Usually, the digitisation workflow includes the following phases:

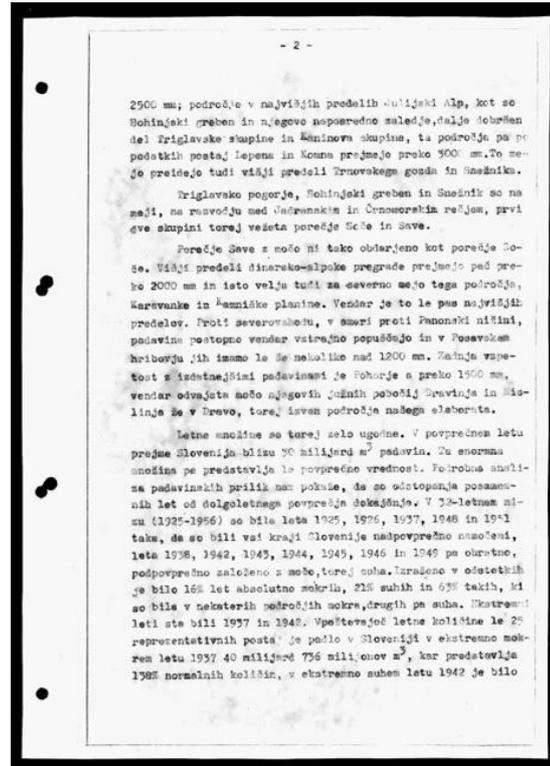
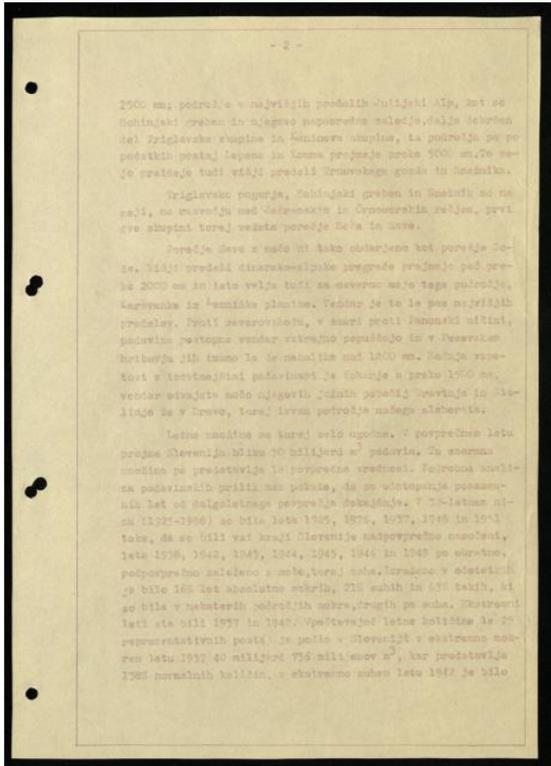
- **Materials Selection and Acquisition:** Setting clear selection criteria according to the aims and goals of the digitisation project; checking if the same content has already been digitised; checking copyright issues.
- **Metadata Definition:** This includes descriptive, structural, administrative and technical metadata, preservation, file naming, persistent identifiers, watermarks, digital signatures, etc.
- **Image Capturing / Scanning / Conversion:** Defining the resolution and bit depth.
- **Image Processing:** Cropping, touch-up, colour-adjustment, contrast adjustment, de-skewing (i.e. straightening out a crooked scan), etc.
- **Full-text Generation:** Optical Character Recognition (OCR).
- **Additional Metadata Generation.**

- Quality Evaluation.
- Digital Collection Management.

Today, digitisation is one of the basic services provided by libraries. Many national libraries have prioritised providing digital access to their old collections. Robotic scanners, which automate the digitisation processes and introduced mass or large-scale digitisation, have proven very useful for scanning large quantities of similar book or periodicals. However, the quality can vary considerably. To highlight these differences, three examples from National and University Library (Slovenia) are presented, showing problems with warping of the text, colour, and contrast. After some manual intervention, better results can be achieved, not only improving images for reading but also affecting the OCR. (Picture III-1, III-2 and III-3).



Picture III-1: Changes in the warped text (left) and original (right). Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-GWA1NIVS>



Picture III-2: Changes in the colour and contrast (right) and original (left) to achieve better OCR. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-4CX01LRB>



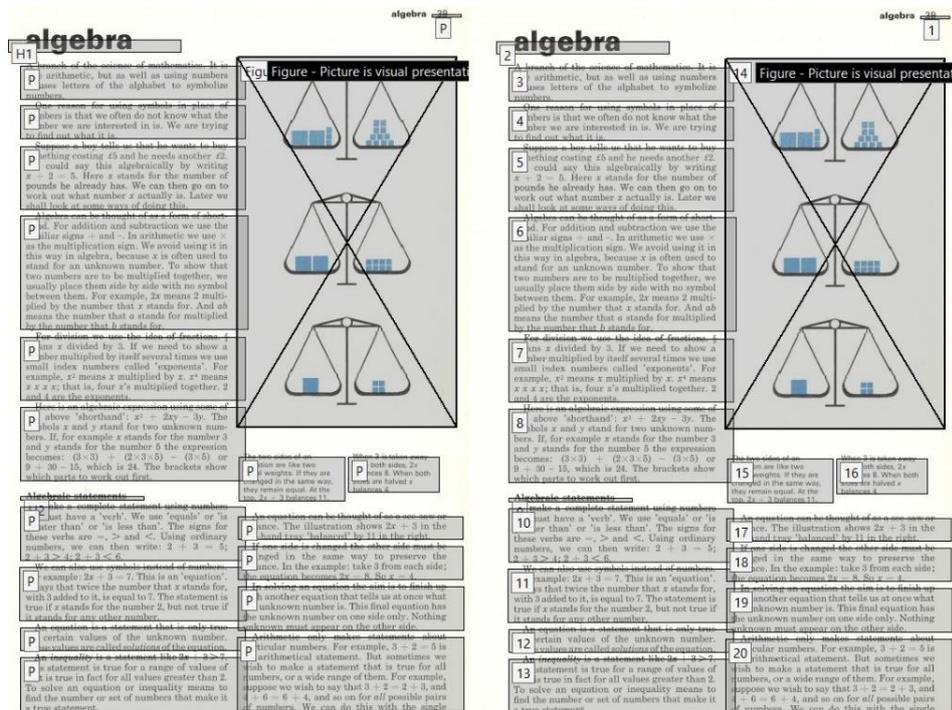
Picture III-3: Changes in the colour and brightness (left) and original (right). Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-5Q4ZPJZW>

The problem is that with such a large quantity of pages, imperfections can easily be missed as each publication has its own characteristics. This becomes evident when users complain that

they cannot access old digitised materials by searching through their full-text only. Upon investigation, the main reason for poor OCR results was imperfections caused by low-quality scanning. Additional imperfections are produced by specific language characteristics and print properties. Some of these are discussed in the following chapters.

### 1.1 Full Text Generation: Optical Character Recognition (OCR)

“OCR or Optical Character Recognition, is a process by which characters typed or printed on a page are electronically scanned, analysed, and, if found recognizable on the basis of appearance, converted into a digital character code capable of being processed by a computer” (Reitz, 2004, p. 505). OCR technology has advanced significantly, making it easy to implement on modern texts. However, this technology cannot be applied to all types of old typographies, but only to those for which additional development has been undertaken (for instance, for Blackletter types or certain languages). Digitised texts are usually processed with various software, including binarization, text aligning, geometric correction, brightness and contrast adjustment, and dust cleaning. For full text generation, page segmentation is also crucial, especially on pages with complex layout pages (Pictures III-4 and III-5). With these processes, we get a clean scan that can be processed with OCR technology.



Picture III-4: Two examples of text segmentation with Adobe Acrobat Reader Pro<sup>7</sup>.

<sup>7</sup> Source: Hari (2022), Digitalne in digitalizirane vsebine za uporabnike s posebnimi potrebami. *Knjižničarske novice*. <https://knjiznicarske-novice.si/digitalne-in-digitalizirane-vsebine-za-uporabnike-s-posebnimi-potrebami/>

134 carbohydrate

Plants use the energy of sunlight to make many different types of carbohydrates. These provide the plant with energy, strengthen its tissues and repair its wounds. Man has found still more uses for substances made of carbohydrates. The table below shows how each carbohydrate occurs in nature, what we take it from and the things we can make from it.

Wood is tough and strong because its cells contain much carbohydrate in the form of cellulose. Here wood pulp is being made into paper.

If a still more complicated carbohydrate, chitin, strengthens the outer covering of insects and crustaceans, such as this crab.

Plants use the energy of sunlight to make many different types of carbohydrates. These provide the plant with energy, strengthen its tissues and repair its wounds. Man has found still more uses for substances made of carbohydrates. The table below shows how each carbohydrate occurs in nature, what we take it from and the things we can make from it.

The new plant, and a store of starch to feed it. It lives on this until it has grown leaves and is to make its own sugars by photosynthesis. Food the plant has stored for its own use is equally good food for us.

Search forns 50 to 60 per cent of an average British diet. In poor countries people eat even more starch. It is a good source of energy, for the body quickly breaks it down to glucose. It is cheap and easy to grow. Potatoes, rice, wheat, barley, rye, corn and other grains are the staples, the basic part, of most people's diet. But we cannot live entirely on carbohydrates.

To digest and use the carbohydrates, we need vitamins. Also to build new cells, we need proteins.

A person eating only carbohydrates will starve to death, no matter how much he eats. Fortunately potatoes and the outer shells of rice and grains contain most of the chemicals we need. But, unfortunately, people like white rice and white flour. Also they keep better. So the dark outer shells with the most important chemicals are removed when the grain is harvested, milled and bleached. People living only on white flour or white rice become ill. This is why flour is enriched by putting back the chemicals that were lost in preparation.

If you eat carbohydrates, protein and fats together, your body will use the carbohydrates first for energy. If you eat a lot of carbohydrates, your appetite may be satisfied before your body has obtained the vitamins and amino acids it needs (see: amino acids). Eating a lot of carbohydrates in candy and french fried potatoes may prevent your eating the more nutritious meat and vegetables.

Two more problems come from eating carbohydrates. If you eat more than you need for energy, the extra glucose is changed to fat. If you over-eat for a long time you will become overweight.

The other problem caused by eating carbohydrates is that the bacteria that cause tooth decay also eat carbohydrates, especially sugar. The simplest way to keep your teeth healthy is to clean your teeth after eating sugar.

carbohydrate	how it is made	where found	what it does	used in
starch	by action of sun on green leaves	stored in parts of plants	provides energy (calories)	food for energy, laundry starch, tracing paper
dextrin	from starch	cassava, wheat, potatoes, corn	thickens and glues together	brewing, baking, softening soap, synthetic dyes
glycogen	by bacteria	made from sugar	gives bulk	food, plasma substitute
sugar	by action of sun on green plants	sugarcane, leaves of corn, maple, palm, beets	provides energy and sweetening, makes other carbohydrates	food for energy and sweetening, brewing, soap, scuba, cosmetics
pectin	by action of sun on green plants	straw and apple fruit waste	combines with acid to make jelly	jam, jelly making, pharmaceuticals, textile finishing
agar and carrageenan	by action of sun on seaweed and kelp	dried seaweed and kelp	sets and stabilizes food	ice cream, jellies, softening toothpaste, slapping plaster and paper
alginate	by action of sun on seaweed	dried brown seaweed	forms into fibres	homogenizing milk and chocolate
gums	formed in bark and leaves after stem damage	acacia tree, cherry plum, mesquite	makes sticky substance	printing paper and textiles, ice cream, glue, cosmetics, medicines, chewing gum
cellulose	built up in cell walls of plants	trunk, stems, seeds of trees, bushes	gives support to plants and animal tissues	book construction, cellulose, paper, cellophane, film, rayon, celluloid, plastic

**Find out by doing**  
Put a drop of iodine tincture into water. If this is dropped on starch, the starch will turn blue. Test some flour, slices of potato, turnip and chopped green beans for starch. Where is it in the bean?

Picture III-5: Example of page segmentation with element differentiation by Abbyy FineReader 15<sup>8</sup>.

Library materials from the 19<sup>th</sup> century and earlier require more attention due to poor quality of paper and other imperfections. The real paradox is that precisely this material is the object of mass or large-scale digitisation. Thousands of old newspaper pages are being automatically digitised. It would be a huge effort to provide a bibliographic description for every article or notice contained in them, and it is not practical to do so. In this context, the optical character recognition cannot replace detailed bibliographic descriptions but can significantly aid in retrieving information from scanned documents. In fact, we cannot imagine any large-scale digitisation project without OCR.

Since such imperfections can also be found in library materials from the 20<sup>th</sup> and 21<sup>st</sup> centuries, we have decided to give attention to this issue as well. Examples are provided below.

The texts processed by OCR can be indexed and searched through digital library search interfaces. This enables easier document finding and reduces documentation procedures. Digital Libraries offer scanned library materials in several file formats, such as PDF, HTML, EPUB or XML, along with the corresponding text resulting from OCR.

<sup>8</sup> Source: Hari (2022), Digitalne in digitalizirane vsebine za uporabnike s posebnimi potrebami. *Knjižničarske novice*. <https://knjiznicarske-novice.si/digitalne-in-digitalizirane-vsebine-za-uporabnike-s-posebnimi-potrebami/>



SPRETNJE ROKE 171

palico jambora (8) gibljivo pritrđimo na jambor z majhnim ušesastim in kljukastim vijakom. Trikotno jadro

(9) pritrđimo na jambor in na vodravno palico jambora z močnim, toda ne predebelim sukancem. Za jambor uporabimo svilo ali podobno blago, ki ne prepušča zraka.

Mere sestavnih delov našega modela niso omenjene, ker si bo vsakdo zgradil jadrnico take velikosti, kot si jo pač želi. Važno je samo, da je krizno ogrodje čim lažje in da znaša dolžina prečnice (2) najmanj tri petine dolžine središčne letvice (1). Pri enostavnem trikotnem jadrju naj bo višina jambora enaka do 1,2-kratni dolžini središčne letvice.

Manjše modele lahko preizkusimo v večji sobi in sicer s pomočjo zračnega curka iz sesalca za prah. Na ta način lahko najlažje določimo mesto, kamor na središčni letvici namestimo jambor.

Naprave za gledanje negativov in diapozitivov

Za natančen pregled negativa ali diapozitiva je zelo praktična steklena opalna plošča, ki propušta svetlobo enakomerno. Naša naprava je prirejena za gledanje vseh negativov in diapozitivov do velikosti 6x9 cm.

Ogradje naprave izdelamo iz 10 mm debelega mehkega lesa. Vse potrebne mere in način spajanja posameznih sestavnih delov vidimo na sliki 1. V desni strani izžagamo krožno odprtino s premerom 33 mm.

Okvir A je iz 3 mm debele vezane plošče in ga pritrđimo na ogrodje naprave s štirimi vijaki za les s poglobljenimi glavami. Kvadratna odprtina okvirja meri 105 x 105 mm, lahko pa je tudi večja, kar velja tudi za ogrodje naprave. S spodnje strani okvirja namestimo stekleno opalno ploščo D. V srednjo polovico dna in v zgornjo polovico zadnje stene izvrtamo več 8 mm lukenj J za odvod toplenga in dovod hladnega zraka v notranjost naprave. Dve letvici na dnu omogočata dostop

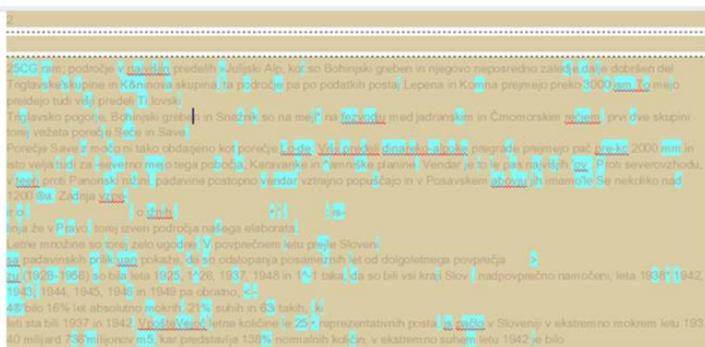
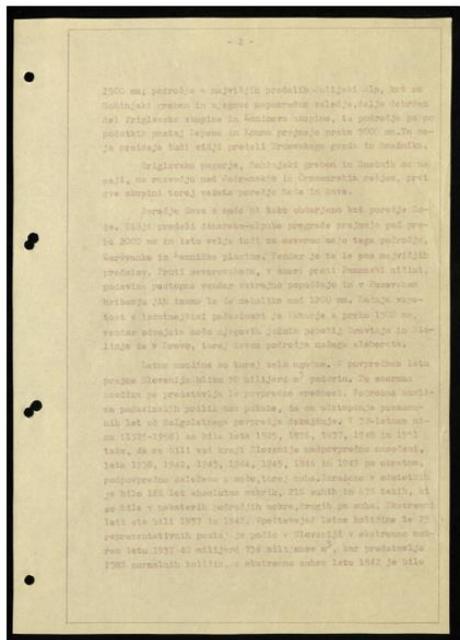
hladnega zraka skozi odprtine na dnu, medtem ko vroči zrak izteka skozi odprtine ha zadnji steni naprave.

Z vijaki (s polkrožnimi glavami) pritrđimo na dno in na zadnjo steno po eno zrcalo E velikosti 100 X 45 mm, ki usmerjata svetlobo žarnice F na stekleno opalno ploščo.

Uporabili bomo 15 do 23 W cevasto žarnico, po možnosti s kar se da dolgo žarilno nitko. Potrebujemo še okovje za žarnico G, vtičač in 2 m dvopolnega PVC kabla. Iz 5 mm debele vezane plošče izžagamo obroč H s premerom 45 mm, na katerega namestimo okovje za žarnico G. Obroč H prevrtamo s 6 mm svodrom, da skozi to odprtino speljemo PVC kabel. Žarnico privijemo v okovje, leseni obroč H pa z zunanje strani s tremi vijaki pritrđimo na : ogrodje naprave.

Picture III-6: Example of text order in a more complex structure with multiple columns. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-JAULUIGA>

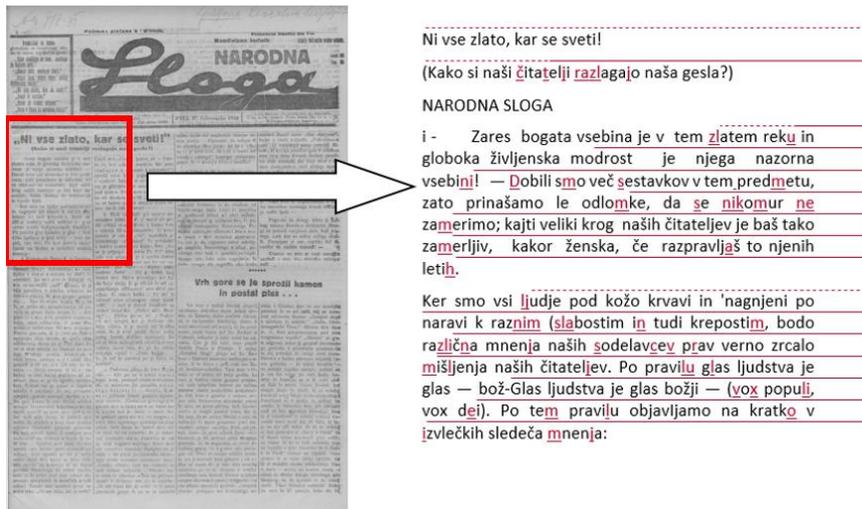
The quality of OCR depends on the software used and the quality of scanning. If the scan is not clear or has imperfections (an example is given in Picture III-7), the OCR will be very poor, necessitating more manual corrections to improve full-text search conditions. At this point, we must decide which digitisation technique is more suitable for our collection.



Picture III-7: Example where text is poorly visible and OCR is also poorly recognised. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-4CX01LRB>

Some more examples that affect OCR quality are presented.

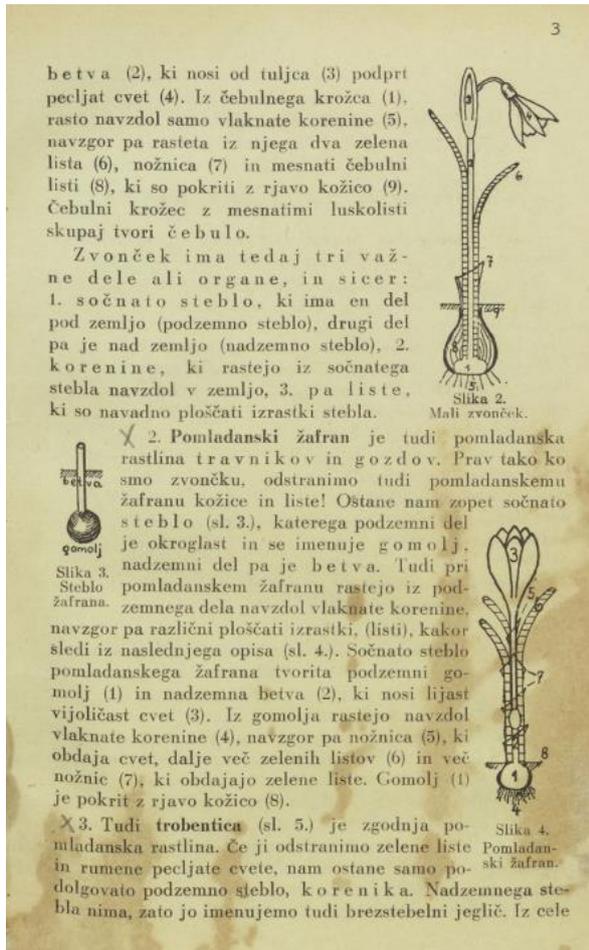
### 1. Imperfections without additional cleaning, which can be time-consuming (See Picture III-8)



Picture III-8: Example of poor OCR quality. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-4BIISUVV>

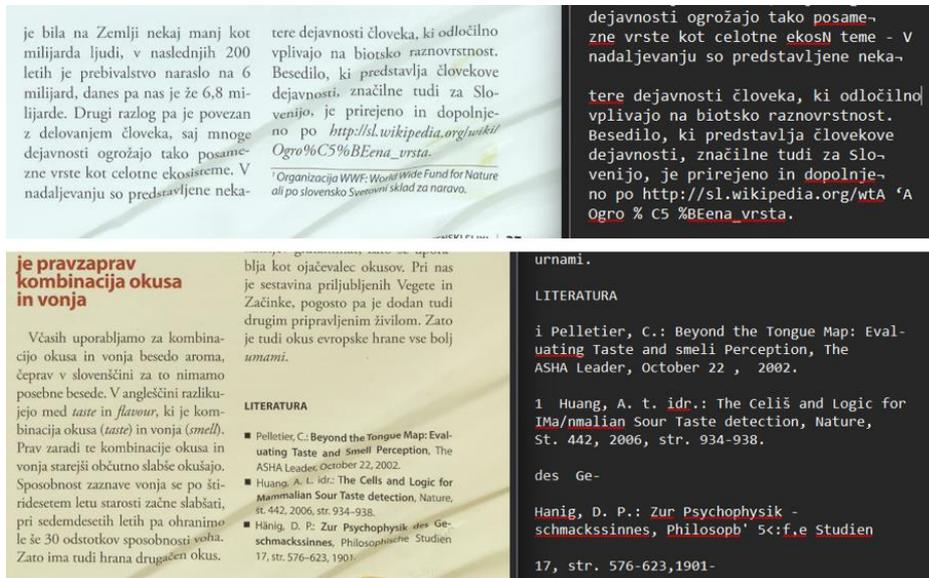
### 2. Imperfections due to damages or foxing stains on the paper, or caused by humidity

Additional spots and damages require special treatment, depending on the colour and the visibility of the letters. Humidity can influence on old periodical pages paper by warping it, and this cannot be corrected even by covering the page with a glass plate while scanning. The result is misaligned characters and unclear words (Picture III-9).



**Picture III-9:** Example of a page with stains visible on the paper. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-OECROINJ>

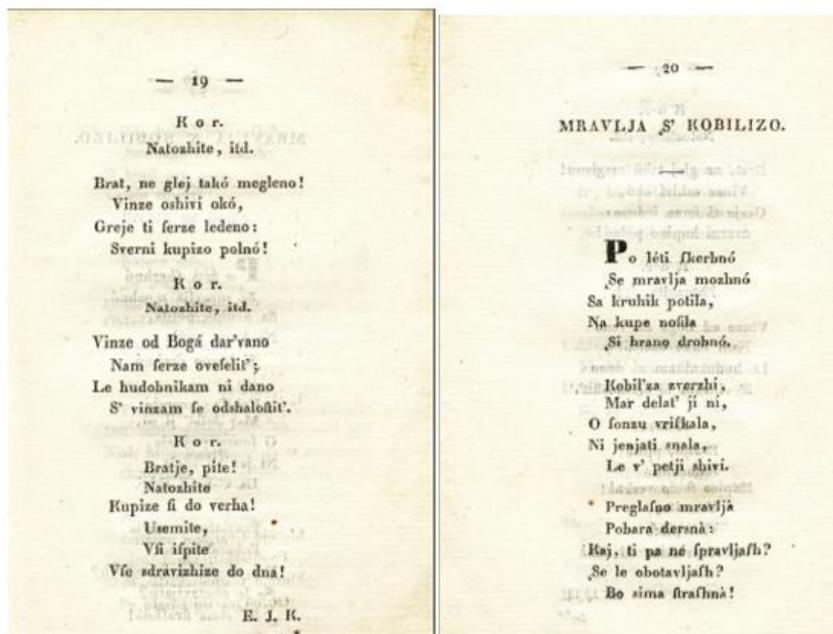
### 3. Damaged, crumpled etc. paper (See Picture III-10)



**Picture III-10:** Two examples of warped text due to crumpled paper and the OCR errors it causes.  
Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-ITXPT3BS>

### 4. Imperfections due to the reflection of text from the opposite side

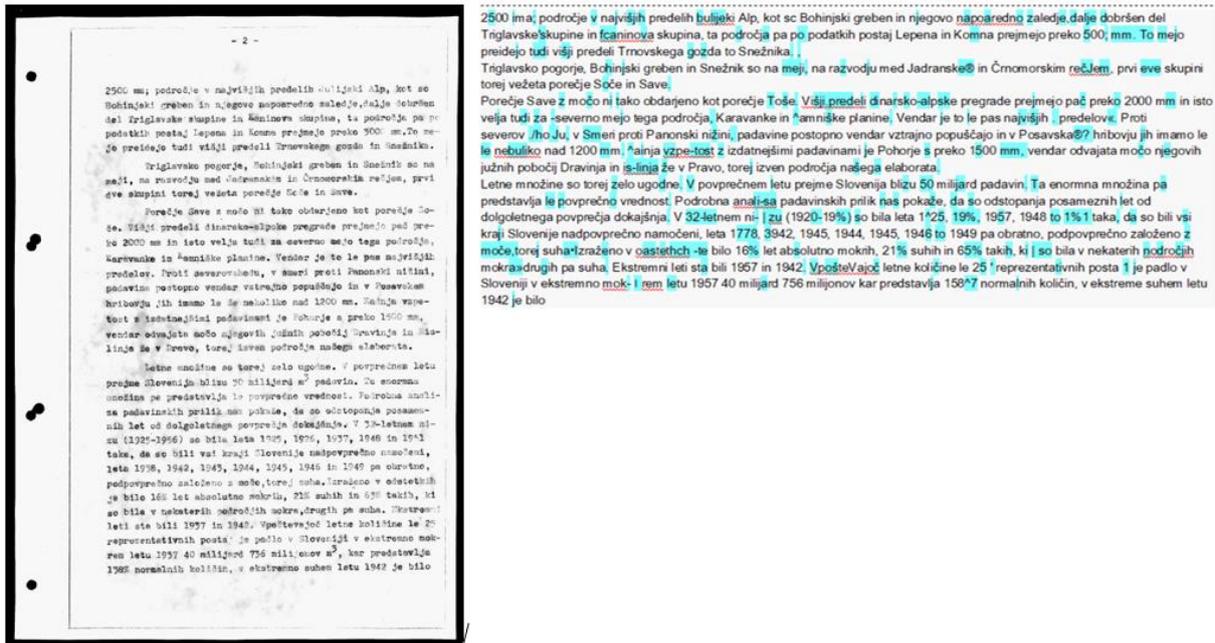
These imperfections can occur when the printing ink was not dry and the letters are reflected from the opposite page. A similar issue might occur when the paper is so thin that the ink shows through from the other side of the leaf. The character shapes are disturbed and the OCR results are of low quality. (See **Picture III-11**)



**Picture III-11:** Two pages of Kranjska čbelica nr. 2, from 1831, showing the reflection of the letters from the neighbouring page.

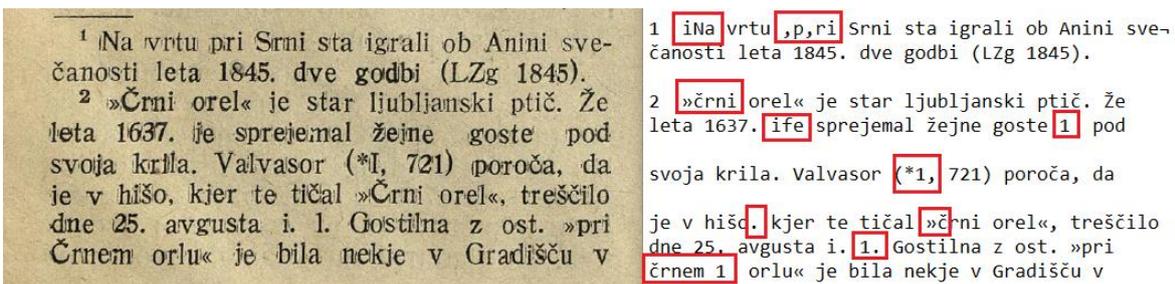
## 5. Poor printing: blurred, broken, faded characters

In **Picture III-12** is a clear example of a printed page using a typewriter, which was used as educational material and OCR-processed with existing OCR technologies. It is evident that the characters the tool cannot recognize influence in the quality of OCR. Due to the rudimentary printing technology, old printed books often contain blurred, broken, or dotted characters, or even characters that are bound together (**Picture III-13**).



**Picture III-12:** Manuscript from 20<sup>th</sup> century with the text resulting from OCR, but only after some manual intervention. Many mistakes still remain. Source:

<https://www.dlib.si/details/URN:NBN:SI:DOC-4CX01LRB>

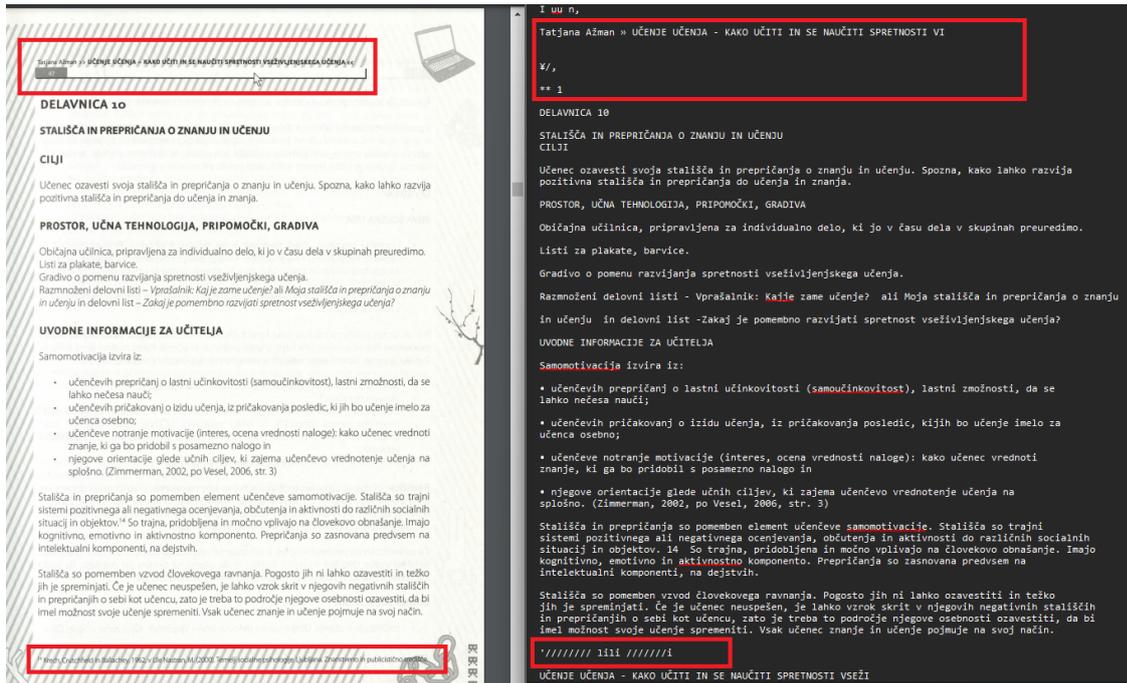


**Picture III-13:** Examples of OCR mistakes due to poor printing. Source:

<https://www.dlib.si/details/URN:NBN:SI:DOC-9H36VQ3X>

## 6. Poor printing: decorations on the paper causing problems in text recognition

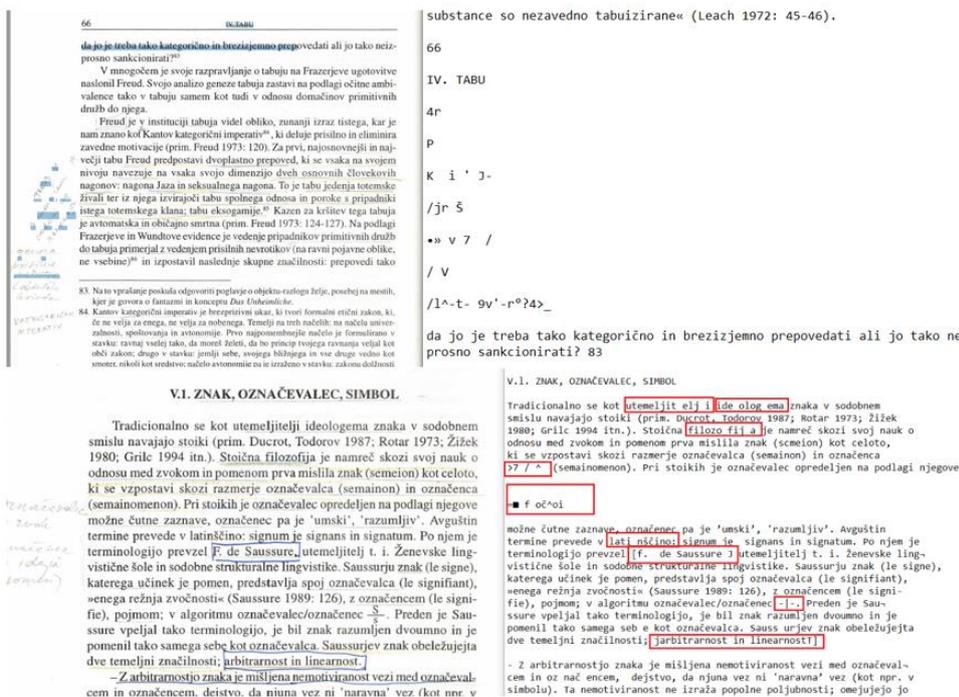
In **Picture III-14**, an example is given of a page where the text over the background graphic causes problems in text recognition or isn't even recognized.



Picture III-14: Example of decorated paper causing problems in text recognition. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-JZ02PZNF>.

## 7. Annotations in the text

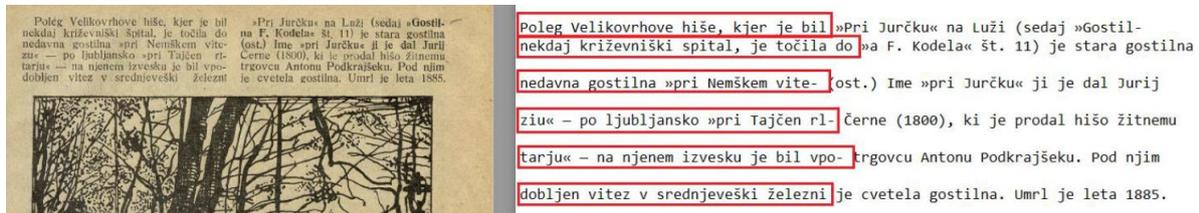
The annotations include texts and drawings added manually as well as stamps used within libraries. They can disturb the segmentation of the text and OCR (Picture III-15).



Picture III-15: Two examples of annotations in the text and the resulting OCR errors. Source: <https://www.dlib.si/details/URN:NBN:SI:DOC-HWODIV2H>

## 8. Page layout complexity

Due to complex layouts, pages can be segmented incorrectly. The effects are significant since the text is not ordered correctly (**Picture III-16** and **Picture III-17**).



**Picture III-16:** Incorrectly recognised part of a page with text in two columns. Source:

<https://www.dlib.si/details/URN:NBN:SI:DOC-9H36VQ3X>

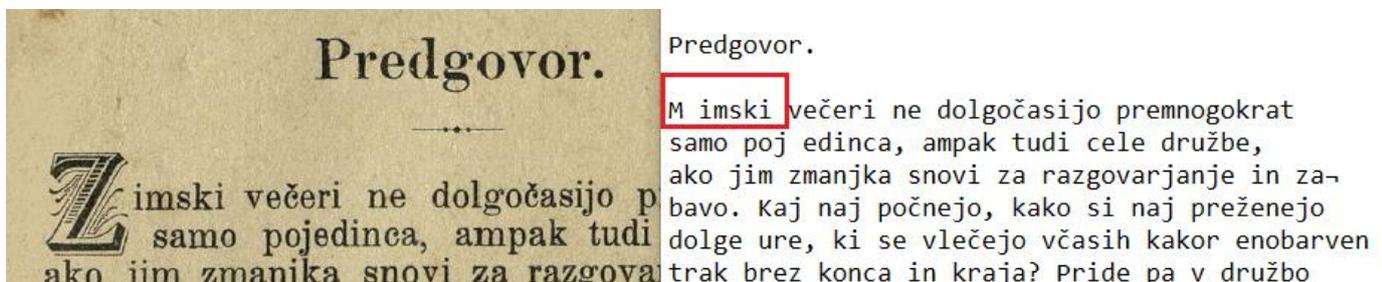


**Picture III-17:** Another example of incorrectly recognised text in two columns. Source:

<https://www.dlib.si/details/URN:NBN:SI:DOC-ITXPT3BS>

## 9. Unrecognisable initials

In some texts, even in modern printed publications, we find decorated initials that are magnified and separated. The OCR tools may detect these as images rather than letters. An example is given in **Picture III-18**.



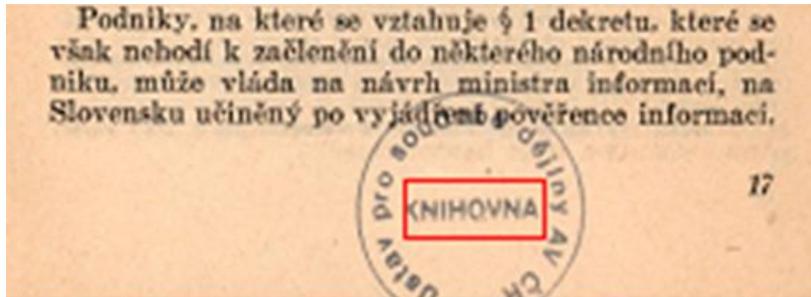
**Picture III-18:** Initials as images at the beginning of the text. Source:

<https://www.dlib.si/details/URN:NBN:SI:DOC-JDE8NLA6>

## 10. Existing stamps in the publication

Stamps in books from library or private collections are very common. Owners mark their library holdings with their stamps. The OCR tool could detect them as part of the text if they

contain readable words. On the other hand, the stamp could impede the clarity of the text behind it (**Picture III-19**).



**Picture III-19:** Example of stamp in publication, affecting OCR. Source: Československá filmová společnost. *Znárodněný film a pokusy o jeho "odnárodnění": k návrhu zákona, kterým se sahá na samu podstatu dekretu presidenta republiky o znárodnění filmu. (1947), p. 17.* Available at: <https://kramerius.lib.cas.cz/>.

## 2 Formats and Accessibility

Digital publications can be produced, distributed, and consumed in different formats. From an accessibility perspective, there are some differences between them.

PDF publications retain the same layout as a print publication, are very familiar to most readers, and have wide support in assistive technology. They are less ideal for people who want to personalise the visual presentation (for example, fonts, spacing) or who read on mobile devices. Many companies provide services to make existing PDF files accessible. According to our EODOPEN research and other similar studies, problems occur with digitised PDF files as they lack, for example, tags to segment the content, disturbed reading order, unclean text, etc. After digitisation, most of the PDFs would need remediation to fit the accessibility standards.

Publications that are read as web pages (HTML) are accessed using a regular web browser and are likely to be a familiar experience for assistive technology users comfortable with using websites. Special reading apps are not required, and the publications can be compatible with assistive technology tools such as read-aloud. As with any website, it depends on the design and the features provided. Publications on websites are often a more basic reading experience and may be missing some of the features required by people with print disabilities, such as bookmarking, changing fonts, adjusting spacing, and page navigation.

E-books are most commonly distributed in the publishing industry using the EPUB format, which is designed for desktop and mobile reading. The standard leverages modern web technologies and supports a wide range of accessibility features, including reflowing content, print page navigation, extended descriptions, and accessible math expressions. The EPUB format also supports accessibility metadata to assist readers with print disabilities. EPUB

publications are read in an app, and for commercial titles, this is typically provided by the distribution platform. The accessibility of these reading apps can vary.

The Kindle Package Format (KPF) is a proprietary format used by Amazon for their e-book retail and lending services. The reflowable book type is the most common of the KPF choices available and offers the best accessibility features. In many cases, the best results are achieved by submitting well-formatted accessible EPUB3 versions to Amazon who then ingest this file and use their own systems to convert to KPF.

Other formats that are very accessible to readers with print disabilities are DOCX and RTF formats, created for use with Microsoft Office Word or similar text processing software. These formats are compatible with mobile devices as well as larger screens and support assistive technologies.

During digitisation, libraries often produce TXT files, which are raw text without any formatting or design, and all graphical elements (pictures, charts, etc.) are not present. These files are mostly useful to readers who are blind, as the files are not cluttered and are compatible with assistive technologies.

It is important to note that while some formats offer more accessibility features, none are automatically accessible. Attention to the relevant accessibility standards for the formats being used is essential.

No single format meets every use case. Publishers or librarians striving for best practices will offer their publications in more than one format.

In the Module IV and V, you can learn more on how to create accessible DOCX, EPUB, and HTML formats from a digitised book.

### **3 Conversion**

E-book users have various needs due to different types of mobile devices or their specific requirements, such as those of blind and partially sighted users. Users themselves can convert the offered file format into a different one according to their needs. Alternatively, file conversion can be carried out by libraries or other institutions, depending on the access format required by the user. Thus, users do not need conversion apps and can access the format of their choice.

It should be noted that the document to be converted must be suitably prepared, with particular attention to structure and navigation. If the document to be converted is not properly structured or made according to accessibility standards and best practices, then the resulting document will inherit the same issues. For example, if a conversion from PDF to EPUB is done on a PDF without marked heading, then the EPUB will lack them as well, meaning no

navigation options, no table of contents, etc. Therefore, if we decide to convert one format to another, we need to be aware of the importance of the original file used for conversion.

Listed below is some of the most commonly used open-source software, tested during the EODOPEN project in the deliverable *D12a Technical report on the implementation of conversion services*.

### Calibre version 6.1.0

**Available at:** <https://calibre-ebook.com>

**Cost:** Free

**About:** The open-source software is developed by various contributors from the user community.

Calibre is a “cross-platform open-source suite of e-book software. Calibre supports organising existing e-books into virtual libraries, displaying, editing, creating, and converting e-books, as well as syncing e-books with a variety of e-readers. Editing books is supported for EPUB and AZW3 formats.”<sup>9, 10</sup> Calibre primarily focuses on text-based documents and does not support conversion into audio formats.

It supports the following operating systems: Windows, macOS, and Linux.

Calibre also features a viewer application that allows users to read e-book, and adjust basic features such as font, size, and margins. An interesting feature is the ability to adjust the text according to the size of the viewer window, even on a computer.

**Input format options:** AZW3, DOCX, EPUB, PDF, RTF.

**Output format options:** EPUB, AZW3, MOBI, DOCX, FB2, HTMLZ, LIT, LRF, PDB, PDF, PMLZ, RB, RTF, SNB, TCR, TXT, TXTZ, and ZIP.

**First impressions:** The conversion process is simple, enabling bulk conversion of multiple books simultaneously, and offering a wide range of output formats. However, it is recommended to modify settings and metadata as needed. It should be noted that Calibre does not support the conversion of a single book into multiple formats simultaneously.

### RoboBraille

**Available at:** <http://robobraille.org>

**Cost:** Free online tool for personal use, subscription fee for institutional use.

**About:** The web service is developed by RoboBraille, Denmark. “RoboBraille is an e-mail and web-based service, capable of automatically transforming documents into a variety of

---

<sup>9</sup> Wikipedia. (2021). Calibre (software): [https://en.wikipedia.org/wiki/Calibre\\_\(software\)](https://en.wikipedia.org/wiki/Calibre_(software))

<sup>10</sup> The software works similarly as Sigil as it enables entrance to each XHTML file, navigation file, styles used in the book, media included in the book etc. Preview is also enabled so that the changes are immediately visible.

alternate formats for the visually and reading impaired.”<sup>11</sup> It offers braille, audio, e-book, and accessibility services. The braille services provide conversion to contracted or uncontracted Braille for 18 Braille code languages. Audio services offer conversion into MP3 files and DAISY talking books, with audio available for 28 languages. E-book services enable document conversion primarily to EPUB and MOBI formats. Accessibility services allow the conversion of inaccessible documents (GIF, TIFF, JPG, PNG, BMP, PCX, DCX, J2K, JP2, JPX, DJV, image-only PDF) into more accessible formats (tagged PDF, DOC, DOCX, Word XML, XLS, XLSX, CSV, TXT, RTF, and HTML).

**Input format options:** DOC, DOCX, PDF, PPT, PPTX, TXT, XML, HTML, HTM, RTF, EPUB, MOBI, TIFF, TIF, GIF, JPG, JPEG, BMP, PNG, PCX, DCX, J2K, JP2, JPX, DJV, TEX, ZIP, and ASC. The maximum size of the file is limited to 64 MB, which may cause difficulties when working with larger documents.

**Output format options:** EPUB, MOBI, tagged PDF, DOC, DOCX, Word XML, XLS, XLSX, CSV, TXT, RTF and HTML, MP3, DAISY full text and audio, DAISY Math full text and audio.

**First impressions:** Conversion can be cumbersome when a user needs to convert a single input format to multiple output formats. This process requires the user to upload the file each time and receive a separate mail for each output format. Furthermore, the output format is stored on the RoboBraille server for a few days before being deleted, necessitating the user to convert the file again if they require it in the same format. It is important to note that different input formats offer varying options for output formats; not all input formats provide the same conversion choices.

### **Balabolka version 2.15.0.821**

**Available at:** <http://www.cross-plus-a.com/balabolka.htm>

**Cost:** Free

**About:** Balabolka is a Text-To-Speech (TTS) software designed for converting text into audio files. It supports synthetic voices installed on a computer system and can convert from various text file formats. The software is compatible with Microsoft Windows XP/Vista/7/8/10/11 operating systems. Additionally, the software allows users to adjust a voice's parameters, such as rate and pitch. A special substitution list can also be applied to enhance the quality of the voice's articulation<sup>12</sup>. The output audio can be generated as a single file or split into smaller audio files, facilitating the swift creation of audiobooks.

It is important to note that Balabolka supports non-commercial voices. However, if users choose to utilise commercial voices, they are required to contact the voice creator and

---

<sup>11</sup> RoboBraille (s.a.) Introduction to RoboBraille. Available on 20. April 2023 at <https://www.robobraille.org/introduction-robobraille/>

<sup>12</sup> Cross+A (s.a.) Balabolka. Available on 20. April 2023 at <http://www.cross-plus-a.com/balabolka.htm>

purchase the license for commercial use or audio broadcast.<sup>13</sup> The guidelines for using commercial voices for non-commercial purposes are not entirely clear, posing potential challenges for libraries, as there is no institutional solution similar to RoboBraille’s agreement between parties. Balabolka is better suited for individual purposes, and libraries could assist users by providing accessible input format options.

**Input format options:** ODT, DOC, FB2, EPUB, RTF, PDF, HTML, and clipboard content.

**Output format options:** MP3, WAV, WMA, OGG, MP4.

**First impressions:** The software for creating text-to-speech-based audiobooks is user-friendly and requires basic knowledge. It enables text editing before conversion, allowing for additional changes such as adding elements (e.g., alt-text which is typically a hidden element not visible in the text and needs to be manually inserted) or removing elements (e.g., certain parts that are unnecessary in the audiobook, such as a list of all original pages in the table of contents). It also provides additional settings to optimize the audio output quality.

### Mathpix’s Snip

**Available at:** <https://mathpix.com>

**Cost:** Free, free educational subscription, or pro subscription for \$4,99 per month.<sup>14</sup>

**About:** The software, developed by Mathpix, includes Snip, a tool designed for extracting text, particularly mathematical equations, from images and documents. It enables conversion for PDFs using AI-powered document conversion technology.

**Input format options:** PDF

**Output format options:** LaTeX, DOCX, HTML, Overleaf, Markdown and more

**First impressions:** In terms of output formats, the recognition of Math as MathML is excellent, even for complex equations. However, there are some issues encountered. The tool utilises its own OCR recognition, resulting in differences from the original content, and it may struggle with complex structures. Additionally, some pages may be missing in the output formats.

### WordToEpub 1.0.10

**Available at:** <https://daisy.org/activities/software/wordtoepub/>

**Cost:** Free

**About:** The WordToEpub software, developed by The DAISY Consortium, is used to convert documents from various word processors (Microsoft Word, Google Docs, LibreOffice Writer, or Apple Pages) to EPUB and HTML formats. The supported operating systems include

---

<sup>13</sup> Balabolka 2.15.0.844. Available on 20. April 2023 at <https://www.techspot.com/downloads/5814-balabolka.html>

<sup>14</sup> Mathpix Products Pricing. Available on 20. April 2023 at <https://mathpix.com/pricing>

Windows 7, 8, 10, and 11. An additional feature called EpubToWord has been available since version 1.0.10 (February 2023) of the software, allowing for the conversion of EPUB format to Word. This option is useful for users who are accustomed to editing documents in word processors rather than utilising more complex software like Sigil or Calibre, where access to the underlying code requires at least basic or advanced knowledge of HTML and CSS. It is worth noting that, according to the developers, the conversion from EPUB to Word is only applicable to DRM-free documents<sup>15</sup>.

**Input format options:** DOC, DOCX<sup>16</sup>

**Output format options:** EPUB, HTML. Additionally, it is also connected to Calibre for further conversion to other formats.

**First impressions:** If the input format is well-structured and properly prepared, the resulting EPUB and HTML formats appear excellent and offer good accessibility. The tool functions as a wizard and guides users through a few steps in simple or advanced mode. The software's development focuses on ensuring accessibility for end users and is continually improving.

### EasyConverter Express 3.0.1.55

**Available at:** <https://yourdolphin.com/EasyConverter-Express> or Microsoft store

**Cost:** Free 30-day trial or one-time payment of 71.99 €, also available as an add-in or as EasyConverter Express Online for 150 €.

**About:** The software, developed by Dolphin Computer Access, is used for simple 3-step conversion.

**Input format options:** DOC, DOCX, or Office 365 documents

**Output format options:** Braille, large print, MP3, and text only EPUB.

**First impressions:** While there were no issues encountered with the EPUB output format, the software encountered difficulties with the MP3 format. Sometimes, it was necessary to attempt the conversion two or three times to obtain the desired result, or in one case, the software would stop at 10%.

According to the results from the EODOPEN research, the most favourable conversion results were obtained when manual interventions were applied to the input file. Therefore, it is advisable to ensure that the input format aligns with the WCAG criteria for accessibility before utilising conversion services to achieve optimal output results. Inaccessible input will yield inaccessible output.

The conversion outputs excelled when dealing with DOCX, RTF, and EPUB as input formats. The DOCX format demonstrates excellent compatibility with most conversion services,

<sup>15</sup> Documents which are not protected by Digital Rights Management services (DRM).

<sup>16</sup> For EpubToWord the input format is ePub and output DOCX format.

offering a wide range of options and consistently delivering good results. RTF gave similar outcomes, but content was often lost. The EPUB format demonstrated good results when the input format was well-prepared. It is advisable to use DOCX, RTF, or EPUB as input formats. All three formats excelled in usability for mobile devices and blind and partially sighted users. Mathematical expressions were best handled when MathML was used in the input formats. During conversion, MathML was managed well in most cases, though occasionally mathematical expressions were lost.

When evaluating audio formats generated through conversion services, it is crucial to be mindful of potential limitations in speech synthesis, particularly how it manages the conversion from text to audio. Unfortunately, the results from the Balabolka conversion service do not include an institutional license for speech voices, requiring organisations to reach out to the rights holder(s) for licensing, especially for commercial voices. In this scenario, the conversion can be utilised for personal use, which may not be ideal for libraries.

On the other hand, RoboBraille, the second tool tested for creating audiobooks, offers an institutional license for a submission fee. When utilising conversion to audio, emphasis should be placed on text order, language, alt-text, and, before conversion, manual adaptation—especially concerning mathematical expressions or special characters. Additionally, regarding mathematical expressions and language segments, DAISY audiobooks created with RoboBraille handled them more effectively than the MP3 format.

To provide users with audiobooks, we suggest offering them accessible DOCX format files and informing them about the use of conversion services like RoboBraille and Balabolka for generating MP3 and DAISY formats.

Regarding the choice of conversion services best suited for formats adaptable to different screen sizes, we recommend considering Calibre, RoboBraille, or WordToEpub. However, as mentioned previously, careful consideration of the input format is essential.

For users with special needs, we recommend WordToEpub, as it addressed most accessibility problems effectively. Once again, the accessibility of the input format should be considered.

As a result of the evaluation, the following recommendations were proposed:

- Consider input format before the conversion: It is advisable to use DOCX, RTF, or EPUB as input formats.
- Ensure the accessibility of the input format before conversion: Ensure that the input format aligns with the WCAG criteria for accessibility to achieve optimal output results. Inaccessible input will yield inaccessible output.
- Select conversion service based on desired output format: Conversion services offer a wide range of input and output formats and the quality of output format may vary

when using the same input format with different conversion services. Additionally, not all conversion services provide options of the desired output format.

- Consider using open, non-proprietary file formats.
- To provide users with audiobooks, offer them accessible DOCX format files and inform them about using the RoboBraille and Balabolka conversion tools/services for generating MP3 and/or DAISY formats. Also, when preparing text for users to create audiobooks, avoid special characters and mathematical expressions by converting them into words (e.g., “one plus three equals four”; “ $\alpha$  – alpha,  $\beta$  – beta,” etc.). Alt-text cannot be directly extracted from images by the conversion services; it must be written in line with other text.

## 4 Further resources

5 Easy Steps to Digitization in Your Library. A Beginner’s Guide. Softlink. Available at: <https://ic.softlinkint.com/blog/5-easy-steps-to-digitization-in-your-library-a-beginners-guide/>

D11 Guidelines and recommendations for the provision of alternative and special formats based on the survey on special needs of users and technical requirements <https://www.dlib.si/details/URN:NBN:SI:DOC-8HF207XU>

D12a Technical report on the implementation of conversion services <https://www.dlib.si/details/URN:NBN:SI:DOC-GU9T82TJ>

D12bc Report on trial implementations for mobile devices and print disabled users <https://www.dlib.si/details/URN:NBN:SI:DOC-W8RSOJGW>

*Digitisation Toolkit.* (s. a.). State Library of Queensland. <https://www.slq.qld.gov.au/how-do-i/contribute-state-library-collections/how-to-contribute/digitisation-toolkit>,

Reitz, J. M. (2004). Optical character recognition. *Dictionary for Library and Information Science*. Westport, CT: Libraries Unlimited.

RoboBraille conversion courses <https://access.sensus.dk/elearn/rb/en-gb/>

Scientific Library Services and Information Systems (LIS): DFG Practical Guidelines on Digitisation (2009). Deutsche Forschungsgemeinschaft. Available at: <https://www.dfg.de/resource/blob/168794/95cc7a6260e13f32ab711a1cefa02d97/praxisregeln-digitalisierung-en-data.pdf>,

Youngs, K. [2016]. *Managing the Digitisation of Library, Archive and Museum Materials*. National Preservation Office, The British Library. Available at: <https://collectionstrust.org.uk/wp-content/uploads/2016/11/Managing-the-Digitisation-of-Library-Archive-and-Museum-Materials.pdf>

## **IV. MODULE: CREATING ACCESSIBLE DOCUMENT FORMATS**

### **About the Module**

This course will focus on how librarians can prepare accessible Microsoft Word documents from digitised publications. It covers 14 elements based on accessibility guidelines and good practices used in the publishing industry. The elements are divided in two sections: basic accessibility and advanced accessibility. Both sections contain processes needed to enable easier access for users of mobile devices and better navigation, structure, and understanding of the content for blind or partially sighted users. The level of accessibility provided varies based on staff, knowledge, and financial resources available. The course is structured this way to cater to different levels of knowledge in accessibility, ensuring that at least basic accessibility can give users better access to digitised content.

The goals of this module are:

- To learn about criteria for basic Word accessibility.
- To learn about additional criteria for advanced Word accessibility.
- To learn how to use templates for quicker work.
- To learn using and understanding the Microsoft Accessibility Checker

Sources used for this course were mostly:

- Research results from EODOPEN project and practical experiences in adapting materials in different formats.
- Adapted learning materials, with permission, which are publicly available on the DAISY Learning website, created by the DAISY Consortium.
- Gathered sources and literature during the duration of the EODOPEN project.
- Practical project experience from Slovenia by creating accessible EPUB publications from Microsoft Word.

### **Training Scenarios/How to Deliver the Content**

This course is provided as a self-learning educational material, comprised of theory, graphical elements, video tutorials, and exercises. By adapting it, it could also be used as a training for interested target groups (maximum 10 people per course). However, as the course is lengthy, detailed, and includes practice work, we decided to keep it as a self-learning course.

### **Duration of the Course**

The duration of the course is approximately 5-6 hours.

## Training Material Needed

Needed for this course:

- Microsoft Office Word 2021 (or similar versions).
- [Mathpix snipping tool](#) for exercises on mathematical expressions.
- Practice document 1, which is plain text from a digitised book and is intended to transform theory into practice as most chapters contain exercises. To work with this document, an original digitised document is also needed for comparison. The PDF can be downloaded from the Digital library of Slovenia. Each exercise in this module assumes the previous ones were already applied, and if all exercises are completed to this document, you can also use it for Module V: Creating accessible EPUB and HTML format.
- Practice document 2, which contains additional exercises that are not included in the first document.
- Assessment questionnaire in Annex 1

## Additional Notes

This course and all the material are based on Microsoft Office Word version 2021, enabling the export of various types of documents (DOC, DOCX, TXT, RTF, PDF). If the participants are using a different version of Word or different document processing software (e.g., free LibreOffice), the path to the functions may be different, but most software enables the functions explained in this course.

For those who wish to participate in *Module V: Creating Accessible EPUB and HTML Formats* or if the result of the production is a new edition of the book, we recommend complying with basic and advanced accessibility to ensure born accessible publications.

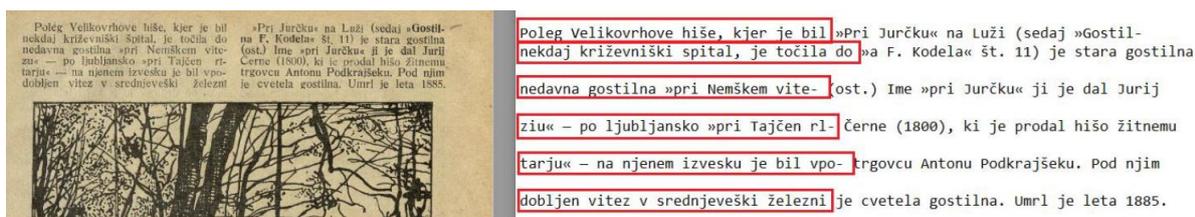
## **1 Basic Word Accessibility**

### 1.1 OCR Clean-up and Reading Order

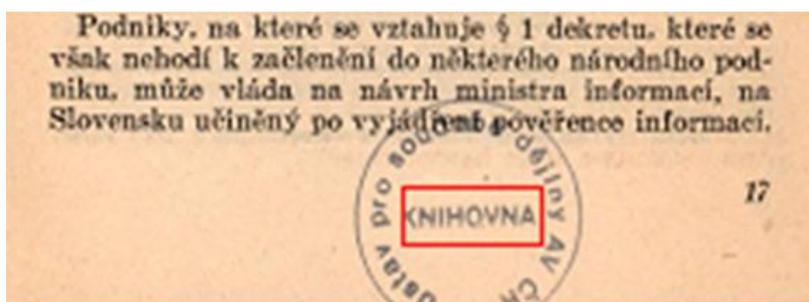
When digitising publications, the OCR process can contain mistakes related to individual characters, sets of characters, or the order of the text (especially with multi-column text). These mistakes can include incorrect or missing characters, incorrect sets of characters, extra empty spaces or tabs, divided words in lines of text, special characters (e.g., ◆, €, ≥, β, ©) and similar issues.

The most common problems with digitised works are:

- **Text is in multiple columns or spanning two pages:** This problem often occurs in complex structured works and requires manual work to determine the correct reading order. For example, see **Picture IV-1**.
- **Special characters, initials:** These may be incorrectly recognised or not recognised at all by OCR and thus may be missing.
- **Library stamps:** These can cause incorrectly recognised parts of the text where the stamp is located over the text, usually requiring manual work to correct. For example, see **Picture IV-2**.
- **User annotations in the original publications:** These writings can result in additional recognised characters that are not part of the original publication and need to be removed during remediation.
- **Correct form of mathematical expressions:** OCR recognition usually handles mathematical, chemical, etc. expressions poorly. This is also a task for OCR clean-up, but this topic is covered separately under Advanced Word Accessibility.



**Picture IV-1:** Example of a text in two columns, incorrectly recognised by OCR.



**Picture IV-2:** Library stamp covering part of the text, potentially disrupting OCR recognition.

When working in Microsoft Word, it is best to follow the following procedure:

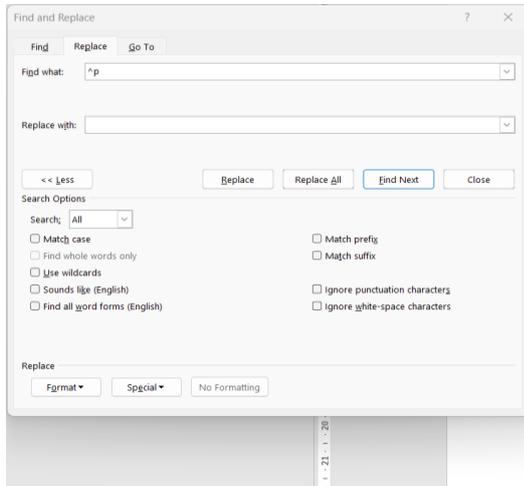
- Copy and paste the OCR text to a Notepad to remove all the formatting and design elements that shouldn't be included.
- Open a blank Word document (or template) and paste the text copied from the Notepad.
- Use "Find and Replace" option in Word to clean the text.
- Manually go through the text and correct any additional detected errors.

### Use “Find and Replace”

The “Find and Replace” feature of Word is an extremely powerful tool. It can be used to clean-up the whole document in minutes. To open “Find and Replace,” you can use the keystroke **Ctrl + H** or click on **Home > Editing > Replace**. For advanced settings in the “Find and Replace” dialog click on “More” to reveal all options. Some of the recommended and common use cases are listed below.

**Table IV-1:** Recommended use cases of “Find and Replace” after digitisation

Description	Find what	Replace with
Remove split words from the book original	^p	One space (press Spacebar once) (Manually replace the paragraphs which are not needed or use “replace all” but this option will also delete the paragraphs which are in the original).
Remove what is left from split words	- (there is one space after the character)	Leave this field empty. (Do not use “replace all” as it will also delete the characters mid sentence).
Remove blank lines (empty paragraphs)	^p^p	^p
Remove doubled spaces	Two spaces (press Spacebar twice)	One space (press Spacebar once). Use “replace all” until the result reaches 0 replacements.
Remove unwanted characters (shown is one example)	~	Leave this field empty.
Replace wrongly recognised characters (shown is one example)	VV	W



M Alpine Province of Carniola. M

The duchy of Carniola the country of wonders, as

Sir Humphry Davy has called it, is situated in the high Alps of Austria near the Adriatic Sea. It is the only Slavish country in the Alps. Mountains towering toward the sky and covered with eternal snow, lakes of crystal purity, green alpine pastures, fragrant forests, the healthy and bracing air that blows from the Alps, all these entice every year thousands and thousands of strangers to this country. It is well provided for the tourists by summer re-treats and winter sports. Direct express trains with dining and sleeping cars, modern hotels, alpine huts and societies for the information of strangers are aiding the traveller.

The duchy is Slovene (South Slavish), her inhabitants have a pretty high degree of culture and instruction and take much interest in the welfare of strangers.

The Capital of the province, Laybach (Sl. Ljubljana) has 45.000 inhabitants and is a charming modern town.

Her principal mountains are: the Julian Alps, the Karawanken and the Alps of Kamnik (Stein) or of the Carnic Alps.

**Picture IV-3:** Example of using “Find and Replace” to clean split words in Practice document 1.

**Table IV-2:** Other common use cases of “Find and Replace”

Description	Find what	Replace with
Remove Tab stops (use of Tab key)	^t	One space (press Spacebar once)
Remove Section break	^b	^p
Remove Manual line break	^l	One space (press Spacebar once)
Manual page break	^m	^p

### Exercise and Practical Example

Go to **Practice document 1** and follow the “Find and Replace” recommended options in Table 1 to create clean document.

1. Clean the split words and paragraph breaks.
2. Remove extra spaces, tabs, and double empty lines.
3. Remove any characters that are disturbing or not needed in the document.
4. Check dubious characters (e.g., vv → w).

Check video: [1.1 OCR Clean-up and Reading Order](#)

## 1.2 Style and Contrast

Preferred fonts for the text are Arial, Verdana, or Helvetica. Avoid decorative fonts which are hard to read by some readers. Text should be left-aligned, in one column, and consistently oriented throughout the whole document. Avoid using too much bold and italic text, or text in caps.

Wherever possible, use the heading styles (see more at chapter [1.3 Headings and Formatting](#)). Alternatively, coloured text can be underlined. If using colour in charts, supplement colour coding with texture, different line styles, text in graphs, or shades of one colour to improve accessibility. Printing a colour document in black and white is the best test to see if any meaning is lost.

Take care with colour contrast, avoiding very similar colours on top of one another. Good contrast between the text and background colour makes the document easier to read for everyone, especially those with visual impairments.

Text colour alone should not be used to convey information in a document. People with visual disabilities, such as low vision and colour blindness, are likely to miss out on this information.

If possible, avoid using text over an image. When adapting digitised content, it is better to use alternative text and describe the text on the image (for more about Alt-Text, see chapter [1.6 Images and Image Descriptions](#)).



**Picture IV-4:** Examples of good and bad contrast and font usage.

### *Exercise and Practical Example*

Go to Practice document 2 and edit the prepared graph according to the exercises in the chapter *Style and Contrast*.

Check video: [1.2 Style and Contrast](#).

### 1.3 Headings and Formatting

Clear, well-formatted headings can significantly enhance the accessibility of your Word documents. Many people use screen readers to create a list of headings, allowing them to skim the document to find the content they want. However, this type of navigation only works when the document’s author uses appropriate heading styles. Screen readers and text-to-speech tools are programmed to recognize them.

To add a heading style to text in Word, **select the text, choose the Home tab and in the Styles box, pick the heading style you want.**



**Picture IV-5:** Image of Home Tab Menu and Styles Box on the right side.

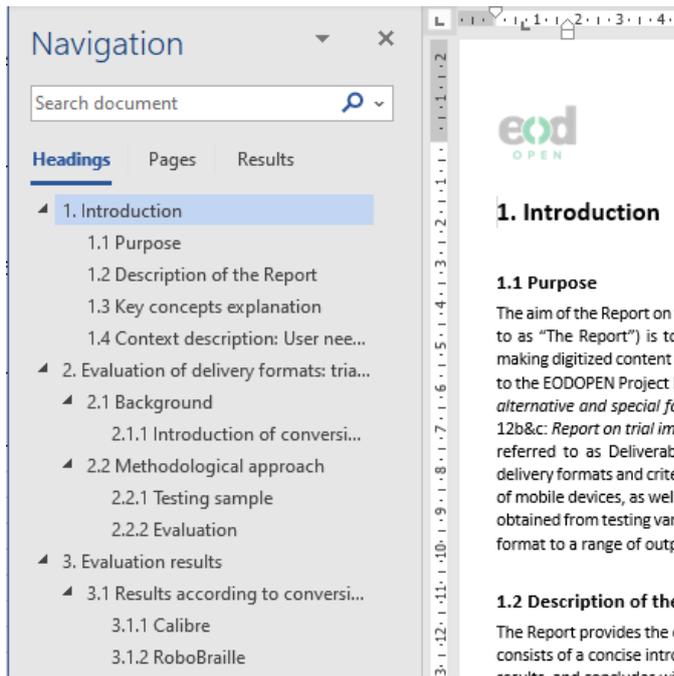
When you save your document in another format for download, such as HTML or PDF, Word retains the heading styles so everyone can still benefit from your headings.

Always use heading styles in a logical order and do not skip levels. For example, heading 1 should always be followed by heading 2, and heading 2 should be followed by heading 3 or another heading 2. If this is not done, the document will not pass the Accessibility Checker tests. Additionally, screen reader users may think they have missed a heading or become confused by the order.

**Main text should have the “Normal” style applied throughout the whole document.**

To determine heading styles when adapting printed publications, follow the table of contents or try to determine logical order from the chapter formatting (bigger or smaller letters, bold or italic letters, etc.).

Keeping the **Navigation Pane** open is very helpful while applying the heading styles. Click on the **View > tick the button Navigation Pane** to open the window which displays the list of all headings in the document (**Picture IV-6**).



**Picture IV-6:** *Opened Navigation Pane for easier control of the heading styles and levels.*

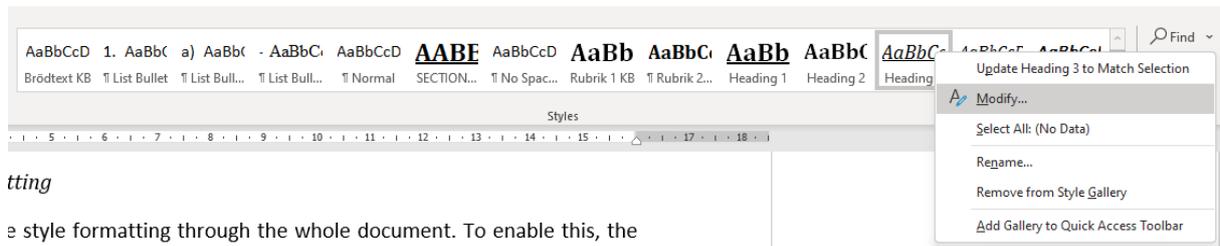
Once the whole document has the correct headings marked, the Table of Contents can be created at the beginning of the document. Chose the line where you want the Table of Contents, go to **References > Table of contents**, and choose an automatic table. This way the navigation is easier from the beginning of the document as the Table of Contents contains hyperlinks to the marked headings.

### *Spacing, Alignment and Margins*

People sometimes press the Enter key repeatedly to create white space between paragraphs. The Tab key is also commonly used to position text or create an indentation effect. These blank lines and white space are annoying to screen readers while reading the document. Such formatting also creates issues in converting the document to other formats. Use the Word built-in features such as indentations, line spacing, and Styles to achieve the desired visual presentation.

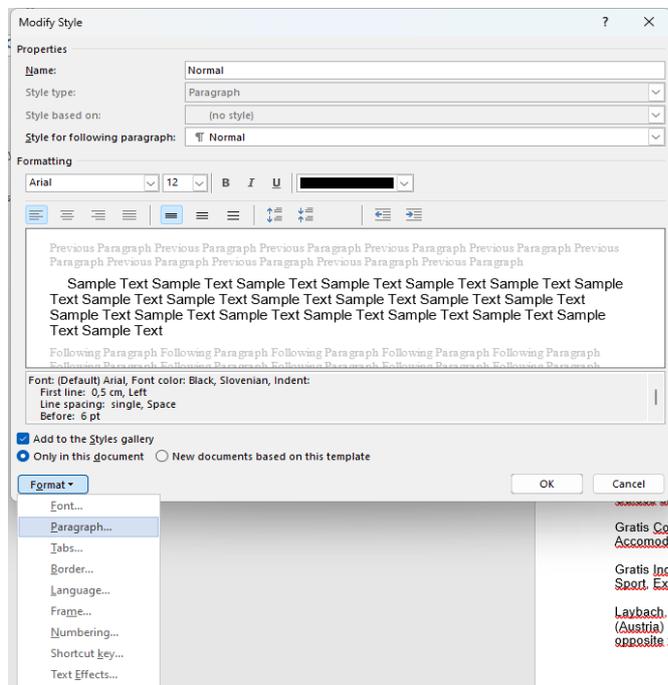
### *Setting Universal Formatting*

Readers prefer consistent style formatting throughout the whole document. To enable this, the best way is to prepare heading styles in advance, before working on the document. To do so, go to **Home > in the Styles Box, right-click** on each style separately (Normal, heading 1, heading 2, etc.) and **choose “Modify”** to change all settings for the chosen style.



**Picture IV-7:** The path to modifying the Styles.

A new window will appear with basic settings and a preview. Additional settings can be accessed with the “Format” button. Under the “Paragraph” option, for example, you can determine the previously mentioned settings like spacing, alignment, indentations, and margins.



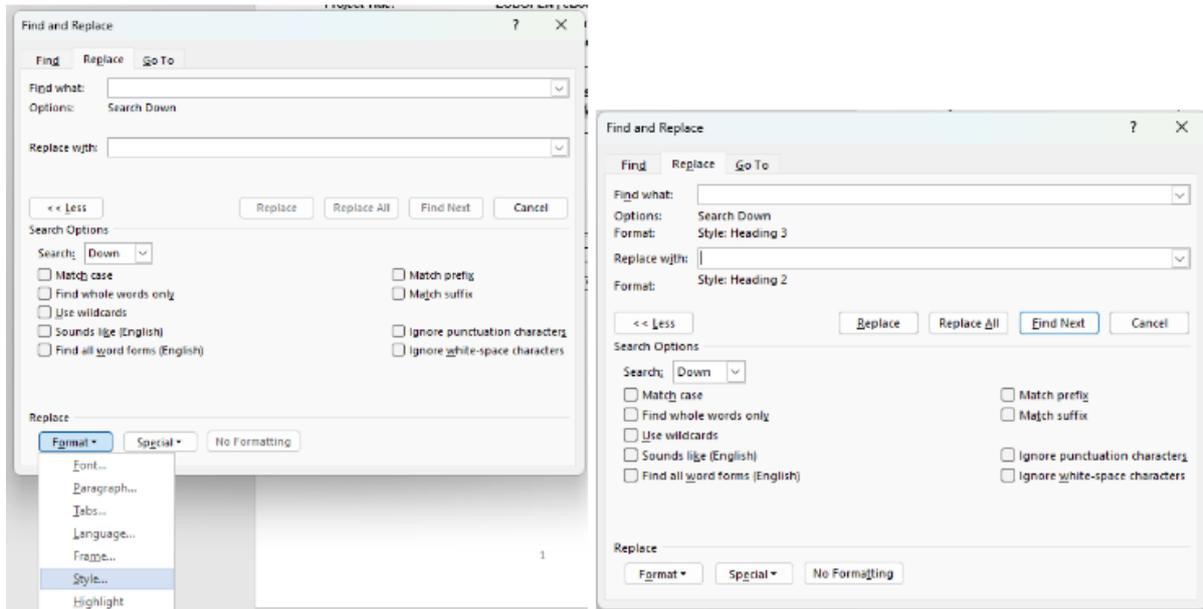
**Picture IV-8:** View of the window to modify styles.

### Find and Replace Styles

The “Find and Replace” feature described in previous chapter can also be used to replace styles. For example, if you have applied the Heading 3 style on certain subsections and now you decide to style them as Heading 2, you can do it all at once using the steps listed below.

1. Open Find and Replace (Ctrl + H)
2. Click in the Find what field and then click on More.
3. Click on Format and then on Style
4. Now select Heading 3. “Style: Heading 3” will appear under the Find what field.
5. Click in the Replace with field and then click Format > Style.

6. Now choose Heading 2.
7. Click on Replace All.



**Picture IV-9:** Using “Find and Replace” to change heading style levels.

You can use these steps to replace any style used in the document. Users who are not aware of this power tip spend a lot of time manually changing styles throughout the document.

### *Exercise and Practical Example*

Go to the original publication and Practice document 1. Add the following:

1. Modify the “Normal” style so that the text is left-aligned, indented to the first line in the paragraph, set font size 12, font family Arial and line spacing of 1,15 or 1,5. The language of the style is English.
2. Modify the “Heading 1” style so that the text is left-aligned, with space before and after set to 12pt, font size 16, bold, font family Arial. The language of the style is English. “Heading 2” should have same settings but without bold.
3. Mark the chapters with heading styles while keeping the Navigation Pane opened.
4. Insert a Table of Contents to the beginning of the document.
5. Change all Heading 1 styles into Heading 2 styles. Remember, the outcome is not accessible. The exercise is for testing purpose only.

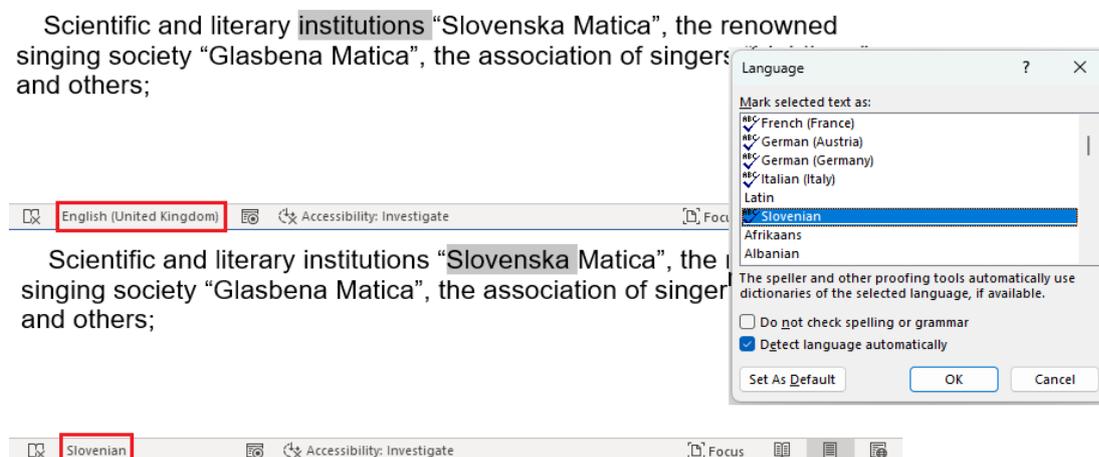
Check video: [1.3 Headings and Formatting](#).

## 1.4 Specifying the Language

It is essential to specify the language of the text in the document. This helps assistive technology like screen readers and read-aloud tools choose the correct TTS voice for reading.

When language is not specified, the system might misidentify the text’s language. In such cases, a screen reader may use a different voice, for example, a French voice to read English text, rendering the document inaccessible for such assistive technology users.

In **Picture IV-10** you can see two examples. Note that the language of the selected text is shown in the Microsoft Word Status Bar. If you click on the same Status Bar, a new window to choose language for the selected text will open.



**Picture IV-10:** Example of selected text with English and Slovenian languages.

To define the language, do the following. Note that defining the language is essential even when the document has text in only one language.

1. Select all text in the document (Ctrl+A).
2. Click on the **Review** tab and then click on the Language drop-down. Select **“Set Proofing Language”** or click the aforementioned Status Bar at the bottom.
3. From the list, select the main language of the document text.
4. If the document contains words or phrases in other languages, select them one by one and perform steps 2 and 3 listed above.

The language of the main text can also be set when modifying the “Normal” style. See more about this in chapter [1.3 Headings and Formatting](#).

**Example:** The following short audio recording presents text read out aloud when a change in language was not specified, compared to a second example where the main text was set to

English and additional text was marked in German. In the second example, the text is read with two different voices.

Check video: [1.4 Example of audio file](#). In the video, there are two examples: one with a single language defined and one with two languages defined. The option with two languages defined is more accessible to blind readers.

### *Exercise and Practical Example*

Go to the Practice document 1. Add the following:

1. Set the language of the main text by selecting all and choosing the language.
2. Set the language in the “Normal” style.
3. Set the language for short instances in other languages.

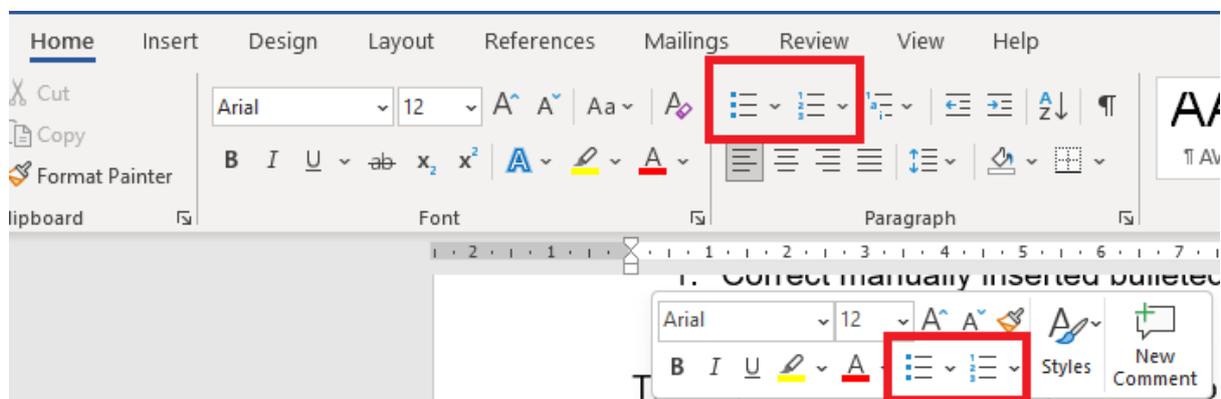
Check video: [1.4 Specifying the Language](#).

## 1.5 Bulleted and Numbered Lists

Use the bullets and numbering feature to create lists in the Word document. The list numbers or bullets should not be manually typed. When the lists are created using the automatic styles, assistive technology users are informed of the beginning and end of lists, and also of the number of list items. This helps them better understand of the content.

To convert a list where the numbers (e.g., a), b), etc.) have been typed manually, select the entire content of the list then choose one of these options:

- Go to the HOME tab, section Paragraph, and choose the appropriate style.
- After selecting the content, quick settings will automatically appear where you can choose an appropriate style.



**Picture IV-11:** An example showing where to find options to apply lists.

### *Exercise and Practical Example*

Go to Practice document 2 and complete the exercise in the chapter *Bulleled and Numbered Lists*.

Check video: [1.5 Bulleted and Numbered Lists](#).

## 1.6 Images and Image Descriptions

Blind and low vision users cannot see and understand the non-textual content (e.g., pictures, charts, maps) in the document. However, screen readers can read out the text description called Alt-Text provided by you in place of the graphical content. You should add Alt-Text to pictures, clip art, charts, tables, shapes, SmartArt, graphics, and embedded objects to help blind and low vision users fully understand your content.

### *Inserting the Images*

When you insert images to your document they shouldn't be copied and pasted from a different location. Instead, they should be inserted through **Insert > Pictures** and then from a location on the device.



**Picture IV-12:** An example showing where to find options to insert images.

This means that when you adapt a digitised book to an accessible format, you should go through the entire book and save the images on your device. You can either save individual images or take screenshots of each one and save them to the device.

The most useful tool that Windows offers for taking screenshots is the Snipping Tool. Press **ALT+SHIFT+S**, then draw the rectangle around the image, and the image is automatically stored on your computer, usually at the location: `C:\Users\name of the user\Pictures\Screenshots`.

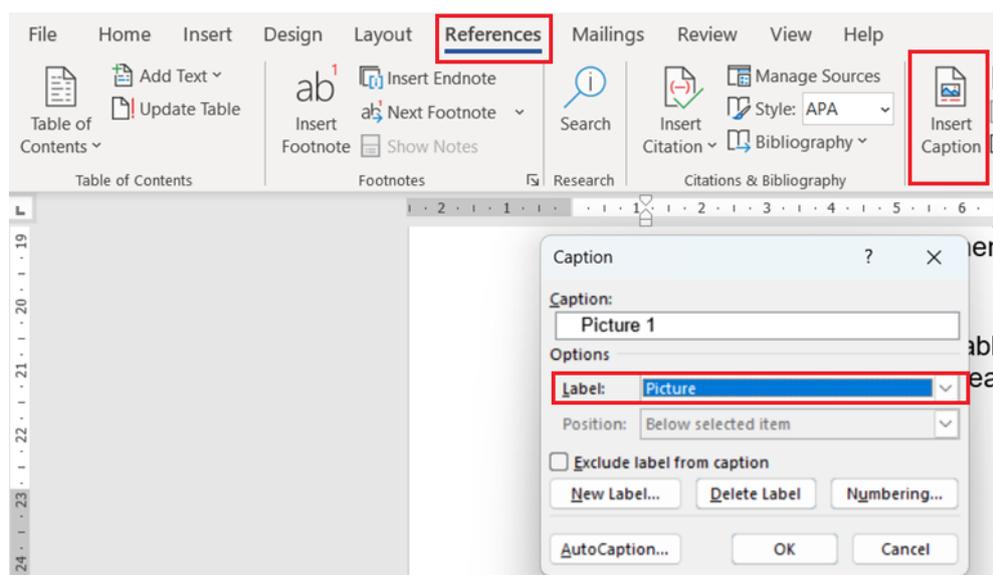
Before inserting the images, make sure you do all the editing on the original image and not within Word (e.g., cropping, rotating, changing the quality of the image, etc.). When using

conversion services later on, they take the original image and ignore any editing you have done in the Word file. Once you have completed that, you can add images to your document as described above.

Note where in the document the picture should be placed. If you have a book with complex structure, it is important to place the image close to the text that mentions it or topic of the paragraph. The image should not disrupt the text flow (e.g., in the middle of the sentence).

### *Adding Captions to the Images*

Depending on the original book you are trying to adapt into accessible material, some images also have captions. To add a caption to the bottom of the image, place cursor at the location where you want to insert it, then go to **References > Insert Caption**, choose image or Figure caption, and press OK. Then add the text caption as it is in the original book.



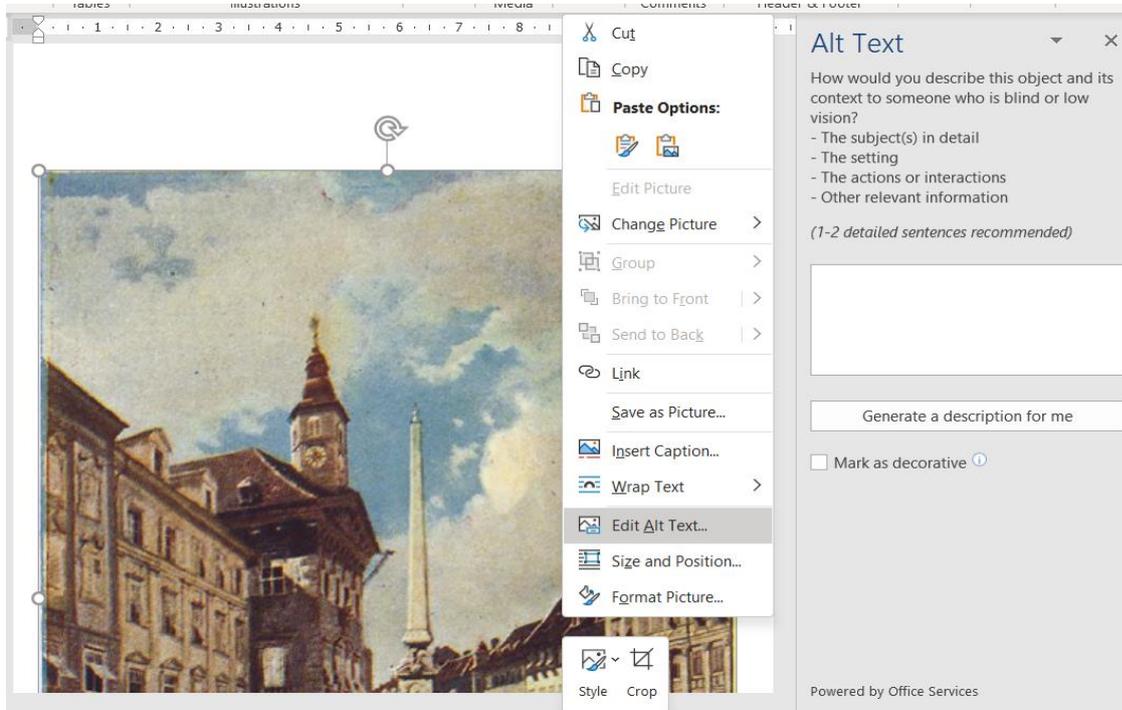
**Picture IV-13:** An example showing where to find options to insert captions to the images.

### *Adding Image Descriptions*

The topic of Image description was already covered in *Module 1: Accessibility and Reading Digital*, under chapter [3.2 Summary of Key Accessibility Principles](#). At this point, we will look at how to add these in a Word document.

To add Alt-Text to an object, **right-click** on the object and choose **edit Alt-Text**. Note that in different versions of Microsoft Word, the option may be placed differently. A new window opens on the right side. Click inside the empty field and type or paste the image description text, then click Close. If the image is decorative, you can also tick the option **Mark as Decorative**.

The option **Generate a Description for Me** could be a starting point, but it still needs manual checking or sometimes additional editing. Also, the generator cannot determine the connection between the image and the surrounding text, meaning it cannot provide image description in the right context. Its success rate could also vary among different languages.



**Picture IV-14:** An example showing where to find options to insert Alt-Text to the images.

If you think your audience needs more information, you can write the description just below the object. For complex objects like charts, people often write the description below the object. It is also common to prefix the description text with the words “Image description:”

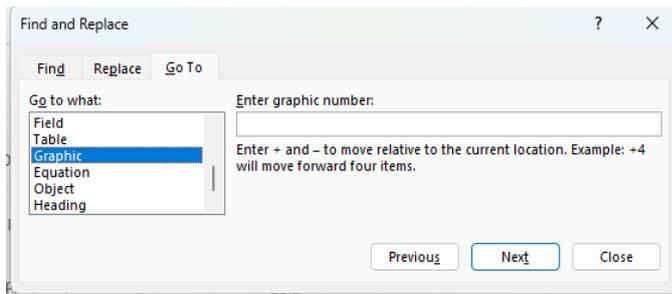
*Tip to Quickly Search for Images and Add Alt-Text*

Earlier in this course, you learned to provide Alt-Text to images. This requires you to select the image, **right-click** on it and select **Edit Alt Text**.

If your document has many images, you can use this tip to keep the Alt-Text pane open, so you will not need to open the context menu for each image and click on Edit Alt-Text.

1. Right-click on any image and then click on Edit Alt-Text. The Alt-Text pane will open to the right of the document. Provide suitable Alt-Text to this image and do not close the Alt-Text pane.
2. Open the Go dialog using the keystroke **Ctrl + G**.
3. In “Go to what” select “Graphic” and click on Next, then on Close.

4. Now press **Ctrl + Page Down**. The next image in the document will be highlighted. The Alt-Text pane will be open and you will be able to check or modify its Alt-Text.
5. Keep pressing Ctrl + Page Down to find the next image in the document and modify its Alt-Text. You can press **Ctrl + Page Up** to find the previous image and check its Alt-Text.



**Picture IV-15:** Go to window to start the navigation between images.

### *Exercise and Practical Example*

Go to the original book and Practice document 1.

1. Save all images from the original book and fix their orientation.
2. Insert the saved images in the Word document.
3. Add captions to the images.
4. Add image descriptions to each image.
5. Try the tip to quickly navigate between images.

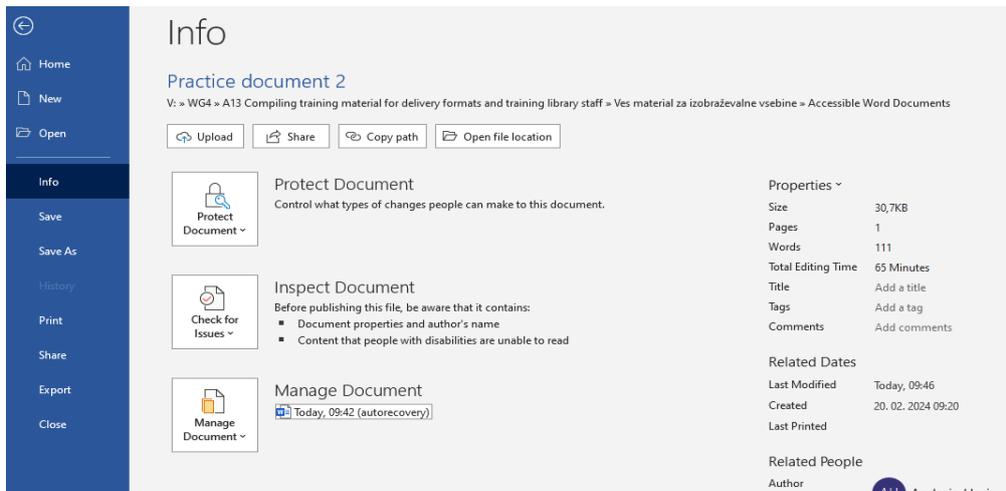
Check video: [1.6 Images and Image Descriptions](#).

## 1.7 File Name, Properties and Exporting

Giving your documents meaningful filenames and document properties makes them easier to find for everyone. These steps are especially important for meeting new accessibility guidelines, like the US Section 508 Refresh, the new EU directive on Accessibility, and many others around the world. A good file name provides clues to a document's content and age.

To rename a document in File Explorer, **right-click** the file and choose **Rename**. Type the new name and hit **Enter**.

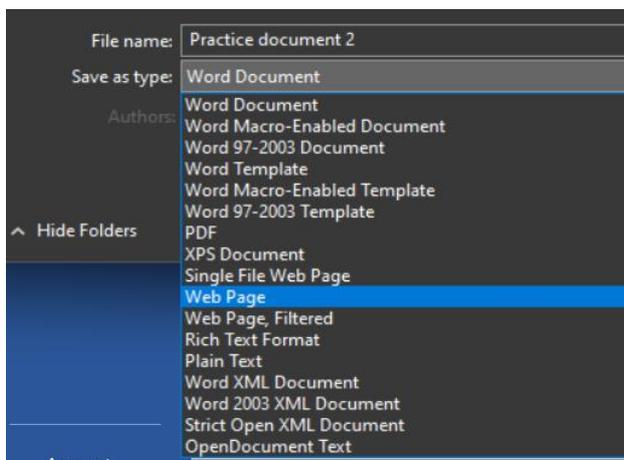
When the document is open in Word, you can add a **TITLE** and **Author name** to the document properties, which makes the file easier for others to find. Adding these properties is also part of the US 508 accessibility guidelines. To add these and other information go to **Home > Info** and the fields will be displayed on the right side of the window.



**Picture IV-16:** An example showing where the info settings are located.

### Exporting and Use of Different Formats

Microsoft Office Word enables saving the document into various formats. The most common are: Word Document (DOCX format) or any older variations of it, PDF format, Web Page (HTML format), Rich Text Format (RTF), Plain Text (TXT format).



**Picture IV-17:** An example of exporting formats.

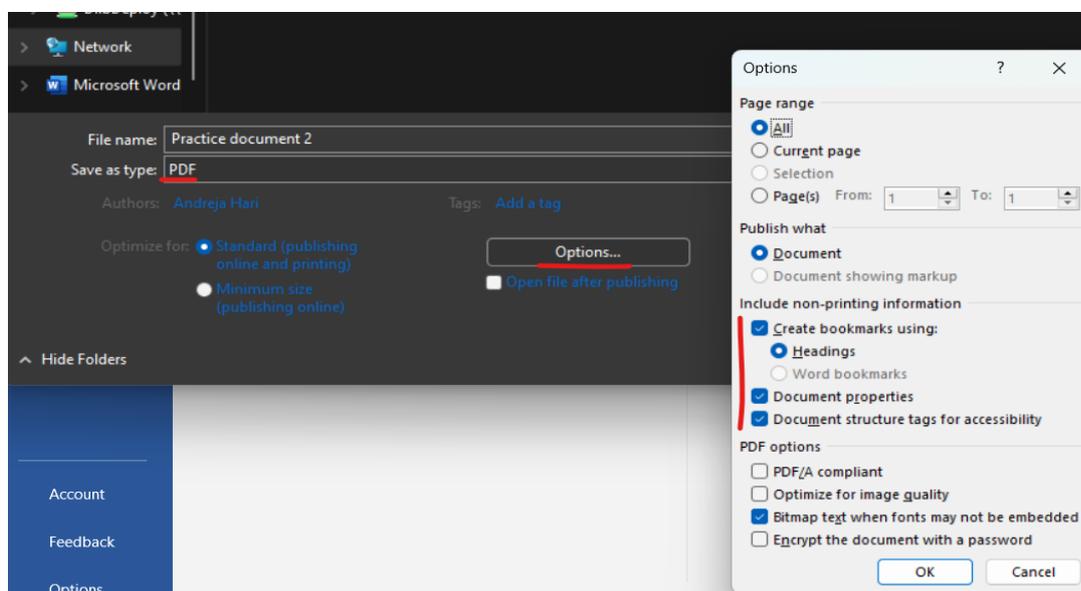
Before saving the document in any format, check the accessibility of the document using the build-in Microsoft Accessibility Checker. That way, you ensure the document doesn't have any accessibility issues, and if it does, you can solve them before saving the document. More about this is described in this module under chapter [4. The Microsoft Accessibility Checker](#).

If a library plans to use the Microsoft Office Word file for further conversion of the format, either in textual or audio formats, it is suggested to save it as a DOCX format. This format is mostly supported by various conversion software and tools (for example, Robobrace, WordToEpub, etc). For more information about the conversion of file formats see *Module III: Digitisation and Formats*, chapter [2 Formats and Accessibility](#).

Creating HTML format is covered in Module V: *Creating Accessible EPUB and HTML Format*, chapter 2.5 *Converting the Book to HTML*, but exporting the Word file at this point into HTML will give similar results. Librarians can try both options, decide which one works best for them, and choose the most accessible outcome.

*Additional tip for saving Word file as PDF*

When saving as PDF, make sure to open additional **Options**. In a new window tick the options to create bookmarks according to marked headings, to add document properties which you added in this chapter and to add document structure tags for accessibility. This way you can provide stronger accessibility of a PDF format.



**Picture IV-18:** An example of settings when exporting PDF format.

*Exercise and practical example*

Go to original Practice document 1.

1. Add properties to the document (at least the book's author and title) and save as DOCX format.
2. Save the document as PDF in regard to accessibility.
3. Try saving the document in other formats.

Check video: [1.7 File Name, Properties and Exporting](#).

## 2 Advanced Word accessibility

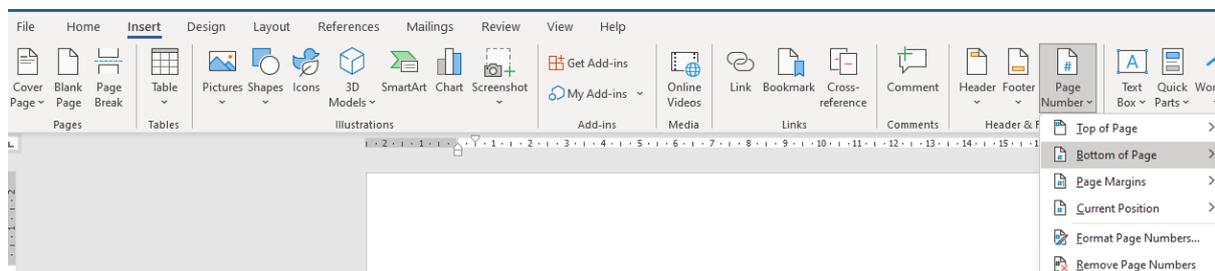
### 2.1 Page Numbering

It is advisable to insert page numbers in Microsoft Word documents as it helps users in various scenarios, such as keeping track of reading progress, making references to the document, and collating the pages after printing.

In Microsoft Word, the page number is generally inserted either in the header or footer. You can choose from the available styles for inserting page numbers or customise the visual presentation of the page numbers. You can also choose to insert both the current page number and the total page number in the “Page X of Y” format.

It is also possible to divide the document into different sections and insert different page numbering in different sections. For example, you may have page numbers in Roman numerals (i.e., i, ii, iii) at the beginning and Arabic numerals (i.e., 1, 2, 3) in the rest of the content.

To insert page numbers in Microsoft Word, click on Insert and then on Page Number. Now you can choose the location (i.e., top or bottom of the page) and then the numbering style.



**Picture IV-19:** An example where the settings for page numbering are located.

### *Page numbering strategy for Print to Digital Conversion*

When a print book is converted to a digital format, the goal should be to create a true copy that retains all text, pictures, and even page numbering.

The first strategy is to adjust the text in the Microsoft Word document so that every page has the same content as the print copy. This requires manual work but can be done by inserting page breaks (Ctrl+Enter), choosing a larger page size, and adjusting the margins and text size. However, in some cases, it may be extremely difficult to accommodate the desired text on each page while maintaining an acceptable visual presentation.

The alternative strategy, which we recommend, is to manually include “print page” at the top/beginning of the page. This method is also useful for conversion to different formats like EPUB by using the WordToEpub tool.

## Tips for Writing Page Numbers for Conversion when Using WordToEpub in Module V: Creating Accessible EPUB and HTML Format

In this case, you need to decide how to number the pages in the intermediary Microsoft Word document so that the final output has the same number of pages as the print book.

For example, WordToEpub allows you to write the print book page number in the Microsoft Word document exactly after the word or sentence as it appears in the print book (e.g., page 4). When such a document is converted using WordToEpub, these numbers with the chosen style are treated as page numbers, and the Microsoft Word pagination is ignored.

Caution is advised with the use of the word “page” as it may appear in the main text and be treated as a page number during the EPUB conversion. Therefore, the typing of the original page number must be very specific. A suggested format is: “Page: 4” or something similar.

This strategy can be illustrated with an example: You digitised a book, performed OCR, and now have the content in Microsoft Word. You notice that the Word document is shorter than the original book. You should now check the pages in the print book and make insertions as follows:

1. Locate the first sentence on the first page of the printed book. Now locate the same sentence in the Word document and type “Page: 5” before it. Most books start with page 5, but verify if this is different for your book. Also, ensure that sentences are completed before adding a new page.
2. Repeat these steps for each page in the print book.
3. The same method can be used in the middle of a sentence if a new page splits the sentence.
4. Caution is needed when a chapter starts on a new page. In this case, the “Page: #” should be inserted in the same line as the chapter title and styled with the heading style (see **Picture IV-20**).

Page: 1 Alpine Province of Carniola.

The duchy of Carniola the country of wonders, as Sir Humphry Davy has called it, is situated in the high Alps of Austria near the Adriatic Sea. It is the only Slavish country in the Alps. Mountains towering toward the sky and covered with eternal snow, lakes of crystal purity, green alpine pastures, fragrant forests, the ‘healthy and bracing air that blows from the Alps, all these entice every year thousands and thousands of strangers to this country.

It is well provided for the tourists by summer retreats and winter sports. Direct express trains with Page: 2 dining and sleeping cars, modern hotels, alpine huts and societies for the Information of strangers are aiding the traveller.

**Picture IV-20:** Example of page numbering at the beginning of a chapter that starts on a new page and in the middle of a sentence.

When this document is converted with WordToEpub, the resulting EPUB will have pages for navigation, with each page containing the same text as the print book. WordToEpub provides several other page numbering options, but based on the experience of the National and University Library (Slovenia) for converting print books to accessible formats, this option worked the best.

Be aware that these custom page numbers inserted within the text may not be treated as desired or may be ignored by Braille and other conversion tools. Your organisation should finalise a strategy for including page numbers in Word documents, considering all factors. This strategy should be made a part of the accessible book creation workflow.

### *Exercise and Practical Example*

Go to original publication (PDF) and Practice Document 1. Add the following in the practice document:

1. Write the original pages from the book into the document using the format “Page: #” where # is replaced with the page number. Add it to the beginning of the page, even if the book has the page number at the bottom of the page.
2. Pay attention to where chapters start on a new page, as the page number should be inserted differently in these cases.
3. Pay attention to where sentences are interrupted by images or not completed at the end of the page. Sentences should be concluded before the page break is inserted.

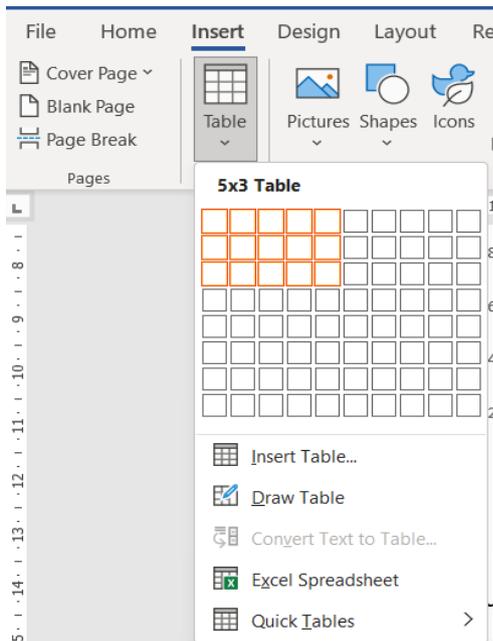
Check video: [2.1 Page Numbering](#).

## 2.2 Tables

Keep the following points in consideration while creating tables in Word documents:

- Use the default Microsoft Word functionality to create tables.
- Keep the table as simple as possible.
- If appropriate, designate a header row.

Do not draw a table using lines, or insert an image of a table. The correct approaches are to use **Insert > Table**, or **Convert Text to Table**.

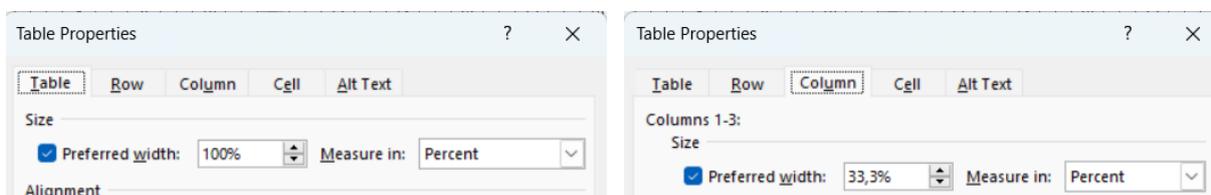


**Picture IV-21:** An example showing how to insert a table.

Screen readers and other assistive technologies do not cope well with complex tables. Avoid using merged cells, split cells, and nested tables, as these will result in warnings from the accessibility checker.

When adapting digitised books, instead of using merged cells, it is better to repeat the same text in two separate cells.

**Define the table and column width in percentages** so that they adjust to different screen and page sizes. To do this **right click** on the table and choose **Table Properties**. Set the percentage in the **Table** tab and the **Column** tab.



**Picture IV-22:** An example how to adapt the table and column width to percentages.

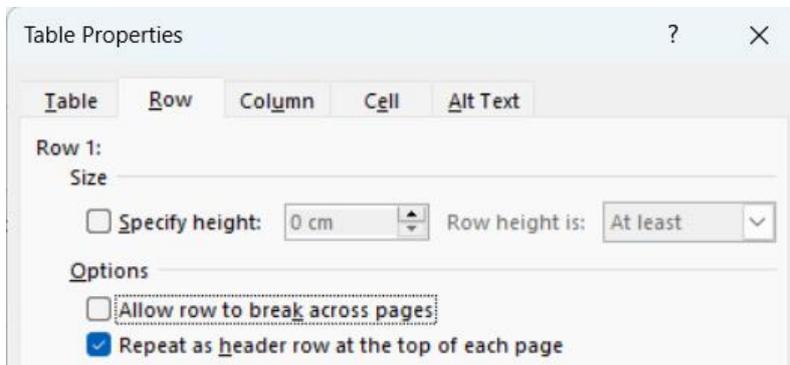
Here are a few ways to check the accessibility of your tables. First, try navigating through your table using only the tab key. If you can tab smoothly through the table, cell by cell and row by row, a screen reader should have no trouble with it. In English-language tables, the tab key should move from left to right, starting in the top-left cell and ending in the bottom-right cell.

### *Designating Header Rows and Providing Alt-Text to the Table*

Next, consider using a **designated header row** for your table. Designated header rows make it easier for a screen reader to navigate your table, allowing the screen reader to announce the column name before reading the data.

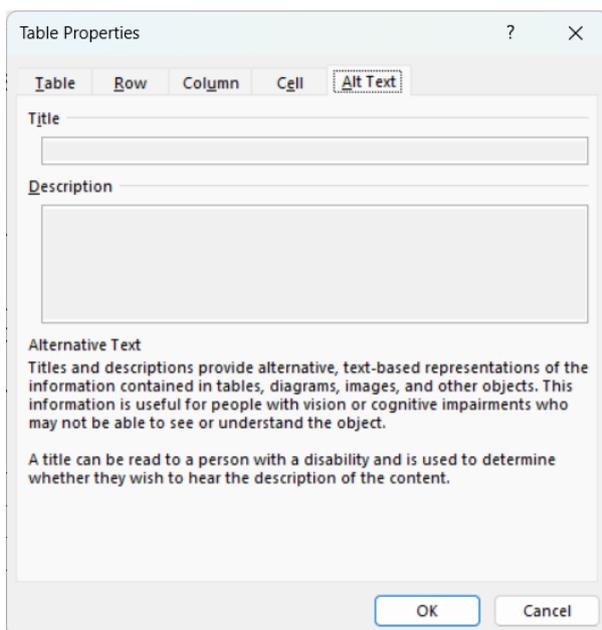
To designate a row as a header, first select it, then:

- Right-click and select Table Properties.
- Select the Row tab.
- Select **repeat as header row at the top of each page**. Ensure **allow row to break across pages** is unchecked.



**Picture IV-23:** An Example of How to Designate a Header row.

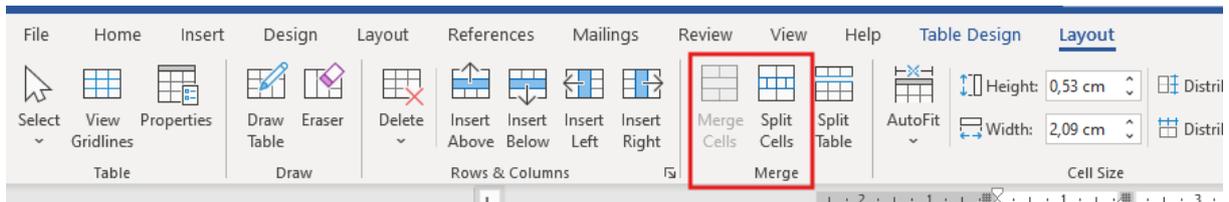
To provide Alt-Text for a table, in the same dialog, switch to the Alt Text tab, provide the text description, and click OK.



**Picture IV-24:** An example showing how to provide Alt-Text for the table.

### Fixing Split and Merged Cells

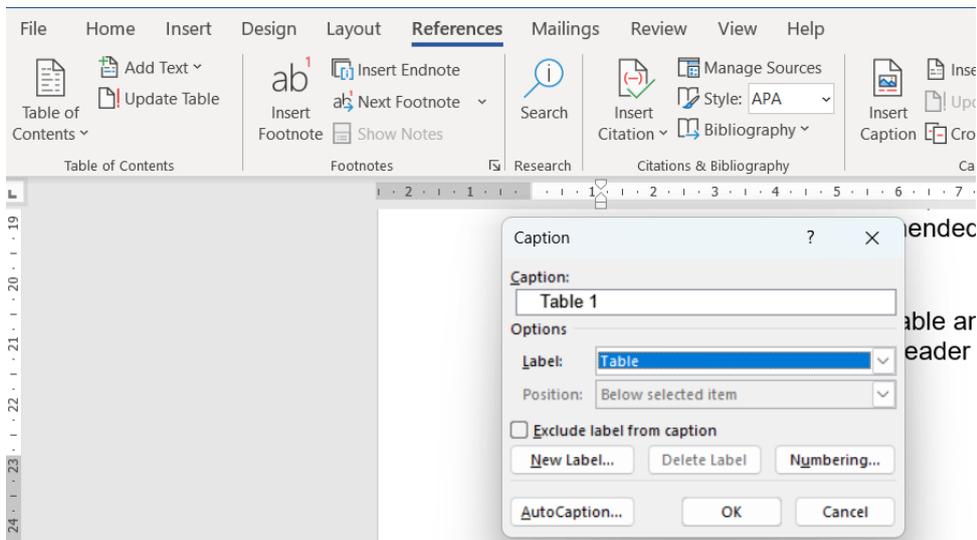
To fix split cells, where two cells occupy the space of a single cell, select the cells, right-click, and choose **Merge Cells**. Alternatively, when the table is selected, go to **Layout** and under the **Merge** section, use the options to merge or split cells.



Picture IV-25: An example showing how to split and merge cells.

### Adding Captions to the Tables

Depending on the original book you are trying to adapt into accessible material, some tables may also have captions. To add a caption to the top of the table, place the cursor at the location where you want to insert it, then go to **References > Insert Caption**, choose Table Caption, and press OK. Then add the text caption as it is in the original book.



Picture IV-26: An example showing how to provide Alt-Text for the table.

### Exercise and Practical Example

Go to Practice Document 2 and solve the exercise in the chapter on *Tables*.

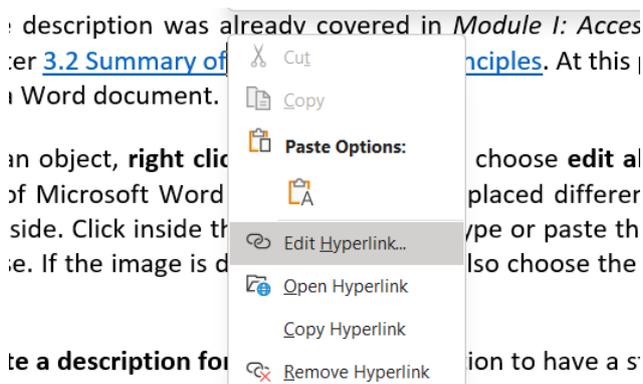
Check video: [2.2 Tables](#).

## 2.3 Hyperlinks

If you have hyperlinks in the document, they can be added by **selecting the text** that should appear as the hyperlink. Then press **CTRL+K**, add the link, and press **OK**.

The displayed text should be in ordinary language to make it more understandable for users who rely on screen reading programs. To change the display text **right-click** the hyperlink on the page, choose **Edit Hyperlink**, type a description in the **Text to Display** box, and click OK.

When adding display text, avoid phrases like “Click here” or “Learn more”. Screen reader software users often rely on a list of hyperlinks to browse the document. If the display text for links uses the same generic phrase, such users will not be able to differentiate between the purposes of the hyperlinks.



**Picture IV-27:** Options available for an existing hyperlink.

### Hyperlinks to Specific Locations Inside the Document

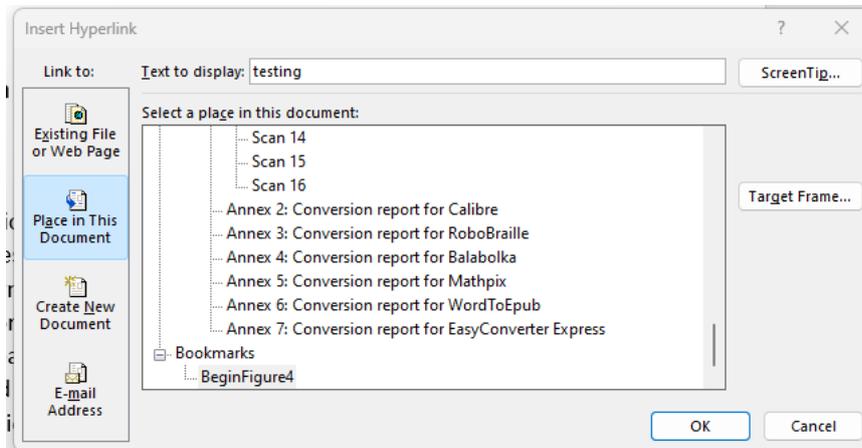
The document could have references to specific parts of the document (chapter, topic, etc.) that should be converted into hyperlinks for easier access (see Picture IV-28).

sample consisting of 16 scanned pages (refer to [Annex 1](#)). Among the five delivery formats, two were automatically generated digitized results (PDF and PDF/UA), while the remaining three had manual corrections applied by project partners (RTF, DOCX and EPUB). Among the last three, EPUB format has been most accessible. These chosen delivery formats served as the input formats for the conversion services described in chapter [2.1.1 Introduction of conversion services used for testing](#). Not all results from Deliverable 12b&c were selected, as

**Picture IV-28:** Example of hyperlinks used in the document.

When using heading styles in your document, these are automatically set as places in the document to which you can create direct links. Sometimes inserting bookmarks is also an option, then using hyperlinks to lead the user to a specific location (e.g., beginning of the content, index etc.). To achieve this, place the cursor before the location, insert a bookmark by going to **References > Bookmark**, and in the dialogue box, insert the name of the bookmark without spaces. Similarly to inserting hyperlinks, you can link the selected text to the

bookmark. In the same dialogue for hyperlinks, choose the location **Place in This Document**, find the correct bookmark, and choose OK.



**Picture IV-29:** A window showing how to add a hyperlink to a specific place in this document.

### *Exercise and Practical Example*

Go to Practice Document 2 and complete the exercise in the chapter on *Hyperlinks*.

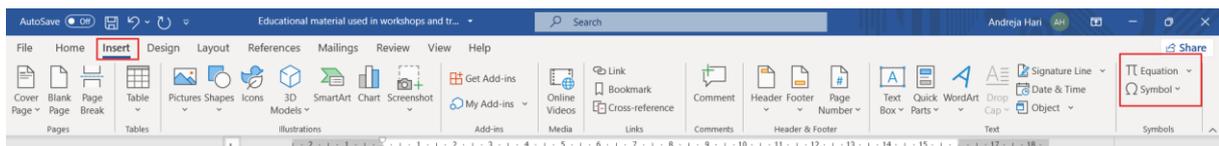
Check video: [2.3 Hyperlinks](#).

## 2.4 Mathematical Expressions and Special Characters

Assistive technology users face difficulties when accessing mathematical expressions or special characters, such as Greek letters, super- and subscript, fractions, etc. Most of the time, such elements appear incorrectly recognised during digitisation, resulting in messy text or incorrect representations of the equations. Additionally, images of equations are not advised for accessibility purposes.

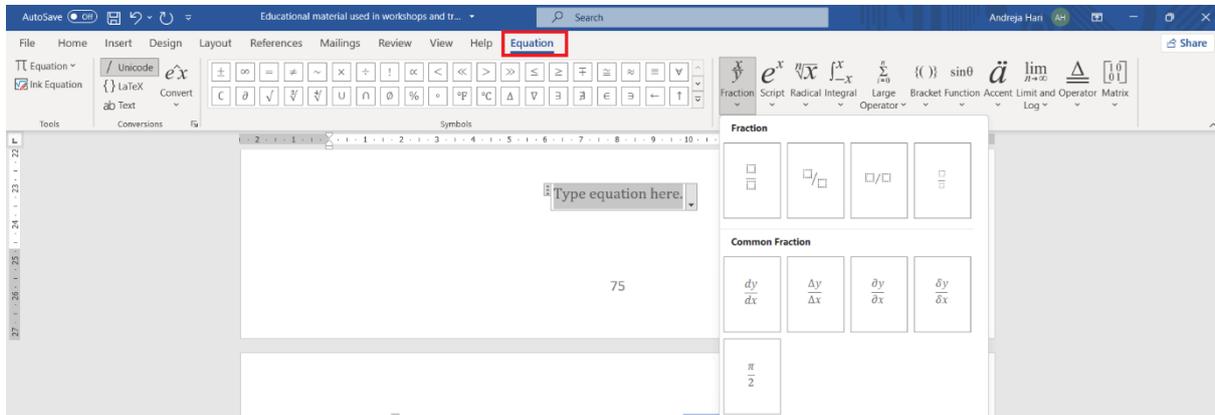
For equations, the most commonly used formats are LaTeX or MathML, or a combination of both. LaTeX presents the equation in a single line, while MathML is based on HTML.

Microsoft Office Word enables the insertion of such equations within the text. This is available under the tab **Insert > Equation** or **Symbol**.



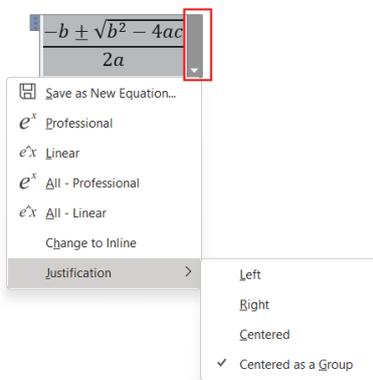
**Picture IV-30:** The path to where you can access the insertion of equations and symbols.

When choosing the option **Equation**, a box appears in the row where the cursor is currently positioned, and you can start building your equation step by step using the additional options in the equation menu bar on top (see **Picture IV-31**).



**Picture IV-31:** The equation menu bar at the top, the equation box in the middle of the document, and the opened options for adding fractions.

Once the equation is added, clicking the arrow on the right side of the equation box gives you some additional options on how the equations should look, whether to change to inline position, or settings regarding justification.



**Picture IV-32:** Options for the equation box.

$$L = \frac{(v_1 + D_1 + D_2) \cdot v_1}{v_1 - v_2}$$

$$L = ((v_1 + D_1 + D_2) \cdot v_1) / (v_1 - v_2)$$

$$\$L = \frac{\left(v_1 + D_1 + D_2\right) \cdot v_1}{v_1 - v_2}\$$$

**Picture IV-33:** An example of an equation, presented in three different writing options: MathML, linear Unicode and LaTeX.

The example above in MathML has the following code behind it.

```

<math xmlns="http://www.w3.org/1998/Math/MathML" display="block">
  <mi>L</mi>
  <mo>=</mo>
  <mfrac>
    <mrow>
      <mfenced open="(" close=")" separators="|">
        <mrow>
          <msub>
            <mi>v</mi>
            <mn>1</mn>
          </msub>
          <mo>+</mo>
          <msub>
            <mi>D</mi>
            <mn>1</mn>
          </msub>
          <mo>+</mo>
          <msub>
            <mi>D</mi>
            <mn>2</mn>
          </msub>
        </mrow>
      </mfenced>
    </mrow>
    <mo>.</mo>
    <msub>
      <mi>v</mi>
      <mn>1</mn>
    </msub>
  </mrow>
  <mrow>
    <msub>
      <mi>v</mi>
      <mn>1</mn>
    </msub>
    <mo>-</mo>
    <msub>
      <mi>v</mi>
      <mn>2</mn>
    </msub>
  </mrow>
</mfrac>
</math>

```

In cases of complex mathematical expressions, instead of adding them manually in Word, it is faster and easier to insert the math by using Mathpix's Snipping Tool. Snipping Tool is a desktop app that allows you to OCR content from your screen and copy math and chemistry to your clipboard. The tool quickly and accurately creates the equation in MathML and other formats. With the option to copy the equation to Word, this can expedite the equation-

building process. The video below shows an example of how fast the creation is compared to manual creation of the equation.

Check video: [2.4 Example of using Mathpix Snipping Tool](#).

### *Exercise and Practical Example*

Go to Practice Document 2 and complete the exercise in chapter on *Mathematical Expressions and Special Characters*.

Check video: [2.4 Mathematical Expressions and Special Characters](#).

## 2.5 Headers and Footers

Screen readers and other assistive technology do not automatically read the information in the document header and footer. Therefore, for accessibility, it is necessary that important information is not contained in the header or footer alone; it should be available in the document content as well.

Watermarks should be avoided as they can make text reading difficult for persons with weak eyesight. If a watermark is necessary, the same information should be included in the document text, as a watermark cannot be detected by screen readers.

Ensure that vital information such as "Author name," "Document number," "Respond by X Date," "Confidential," and "Do Not Distribute" is not present only in the header, footer, or as a watermark. Such information should be provided near the beginning of the document.

Note that accessibility guidelines do not require the deletion of the header or footer. You only need to ensure that the information in the header and footer is also present in the document text.

### *Exercise and Practical Example*

Go to Practice Document 2 and complete the exercise on chapter *Headers and Footers*.

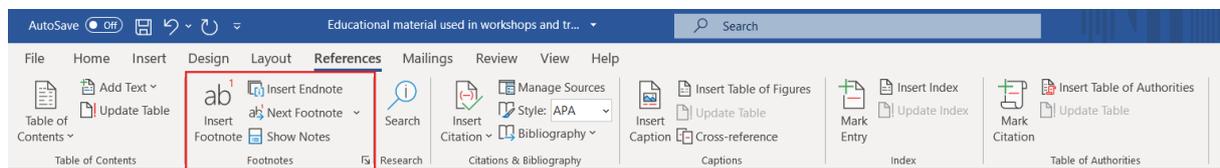
Check video: [2.5 Headers and Footers](#).

## 2.6 Footnotes and Endnotes

For ensuring accessibility, Footnotes and Endnotes should be created using Microsoft Word features. Do not manually type the note reference number or create space at the bottom of pages or the end of the document to write the note text.

Follow the steps listed below:

1. In a Microsoft Word document, **place the cursor where you want to insert a Footnote/Endnote**. Typically, this will be after a word or the end of a sentence. You can click at the location with a mouse or use the keyboard to move the cursor to the desired location.
2. In **References** tab click on **Insert Footnote** or **Insert Endnote**.
3. The cursor will move to the Footnote/Endnote area. **Now, type or paste the Footnote/Endnote text** here. Microsoft Word creates space for Footnote text at the bottom of pages and for Endnote at the end of the document.



**Picture IV-34:** The path to where you can access the insertion of footnotes and endnotes.

Microsoft Word automatically numbers the notes. For example, if this is the first Footnote, “1” will be inserted after the word/sentence and also before the Footnote text.

When you follow this process, a link is established from the Footnote reference number to the Footnote text. One can hover the mouse pointer over the Footnote reference number to see the Footnote text. Screen readers provide keystrokes to read the Footnote. When the Word document is converted to other formats, the Footnote reference number works as an internal hyperlink.

### *Exercise and Practical Example*

Go to Practice Document 2 and complete the exercise on chapter *Footnotes and Endnotes*.

Check video: [2.6 Footnotes and Endnotes](#).

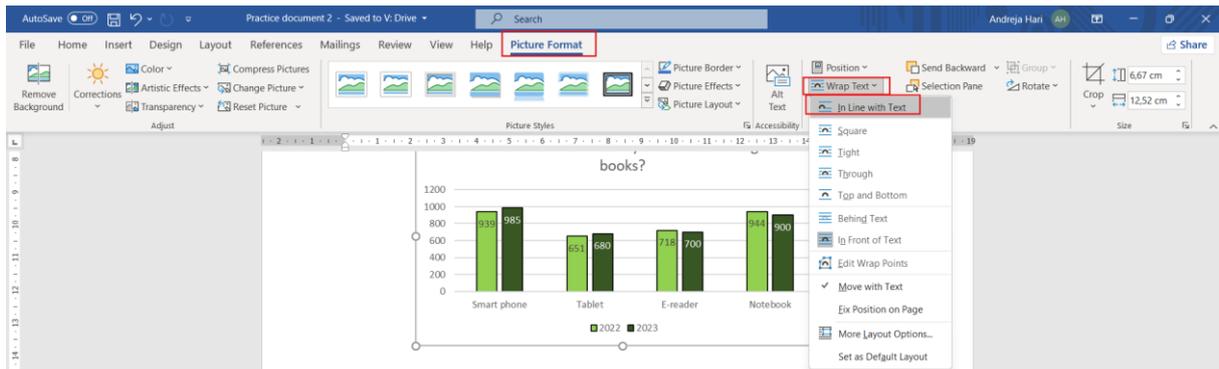
## 2.7 Text Boxes, Charts and Objects

### *Object Wrapping*

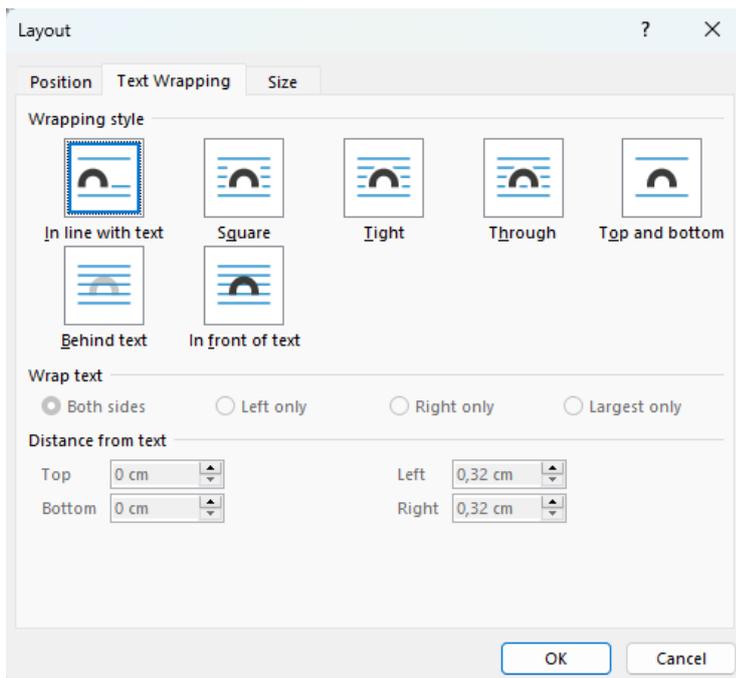
Screen reading software may ignore floating images, charts, and other objects or read their Alt-Text in the wrong order. It is recommended to use the “in line with text” or the “top and bottom” option in text wrapping.

To change the text wrapping of images:

- Click on the object, then click on the **Picture Format** menu bar, click **Wrap Text**, and select **In Line with Text**.
- **Right-click** on the object, then click on **Size and Position**. Switch to the **Text Wrapping** tab and select **In Line with Text**.



**Picture IV-35:** The path to change the wrapping of the image according to the first option.



**Picture IV-36:** Options available for wrapping according to the second option.

Note that this option may be located in different menus depending on the type of object selected.

### *Avoid Using Text Boxes to Create Sidebars or Highlight Text*

Text boxes should be avoided as screen readers often skip them or cannot read their contents. You can use the Borders and Shading feature of Word to highlight paragraphs and make them look like text boxes.

To address this issue, you can insert the content of the text box or sidebar at a logical location in the document. Use left and/or right indenting to make this content stand out from the surrounding text. You can also give it a border or background colour of your choice. This way, the visual presentation can mostly be retained while ensuring accessibility for assistive technology users.

### *Exercise and Practical Example*

Go to Practice Document 2 and solve the exercise on chapter *Text Boxes, Charts, and Objects*.  
Check video: [2.7 Text Boxes, Charts, and Objects](#).

## **3 Templates**

### *3.1 What is a Word Template?*

Word templates are sample documents upon which other documents are based. Word template files have the extension “dotx”, “dot”, or “dotm”. Templates reduce the time taken to create a new document and ensure that the documents based on them have the same structure, font, styles, and page layout.

Organisations often create templates for documents that need to be created frequently, such as monthly or quarterly activity reports. By creating templates, organisations can regulate the content and control the appearance of the documents created by their staff.

If a library provides a service of adapted books to its users, it is recommended to create a template based on all the previous chapters and use it for each adapted book.

### *3.2 Why Make the Templates Accessible?*

If the Word templates are not accessible, then documents based on them will also not be accessible. For example, if sections and sub-sections in the Word template do not have heading styles, the documents based on this template will not be easy to navigate or understand for persons with disabilities. Similarly, templates may not have been tested for

colour contrast issues, making documents based on the template difficult to read for persons with visual impairments.

When templates are not accessible, other formats created from them, such as PDFs, are also not accessible and require more time and effort to remediate. Additionally, persons with disabilities may not be able to use certain templates for creating new documents if the templates use inaccessible fields and text boxes.

### 3.3 What to Do with Templates

The process of creating an accessible template is the same as creating an accessible document.

1. While creating the document that will later be converted to a template, ensure all best practices and guidelines mentioned in this course for structure, styles, images, colour, tables, hyperlinks, and language are followed.
2. Use Microsoft Accessibility Checker and fix any issues flagged by it (see more in chapter [4. The Microsoft Accessibility Checker](#)).
3. Test the document with a screen reader.
4. Save the document as a template. Check if the template is easy to use for everyone.
5. If possible, get the template tested by persons with disabilities and fix any accessibility issues detected.

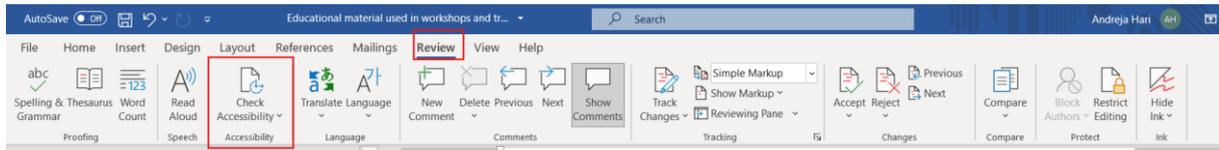
## 4. The Microsoft Accessibility Checker

The Accessibility Checker tool finds accessibility issues in your Word documents. It generates a report of issues that could make your content difficult for people with disabilities to understand. The Accessibility Checker also explains why you should fix these issues and how to fix them.

Ensuring that your documents pass the tests of this Accessibility Checker is generally sufficient to ensure that people with different disabilities will not have major difficulties using the content. Note that accessibility for all users cannot be guaranteed; however, documents that pass the Accessibility Checker and manual testing will certainly be more accessible and easier to use.

### 4.1 Start Accessibility Checker

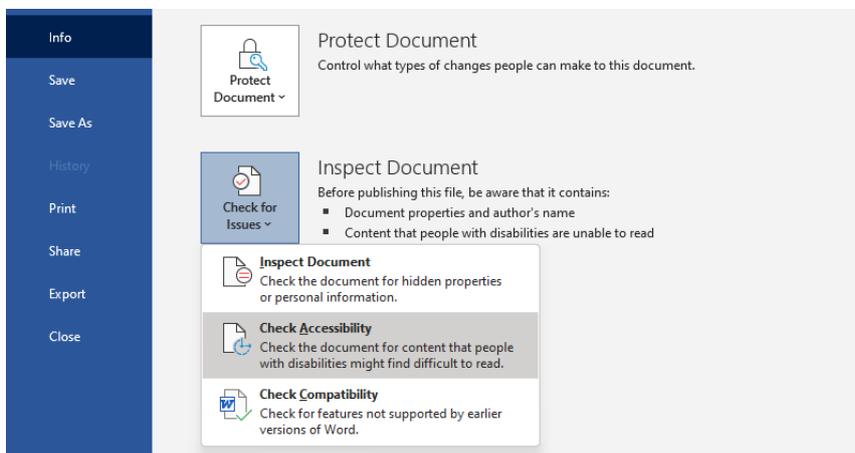
It very easy to open and use the Word Accessibility Checker. Go to the **Review** tab and look for the **Check Accessibility** button. Click it to open the Accessibility Checker.



**Picture IV-37:** How to access the Accessibility Checker.

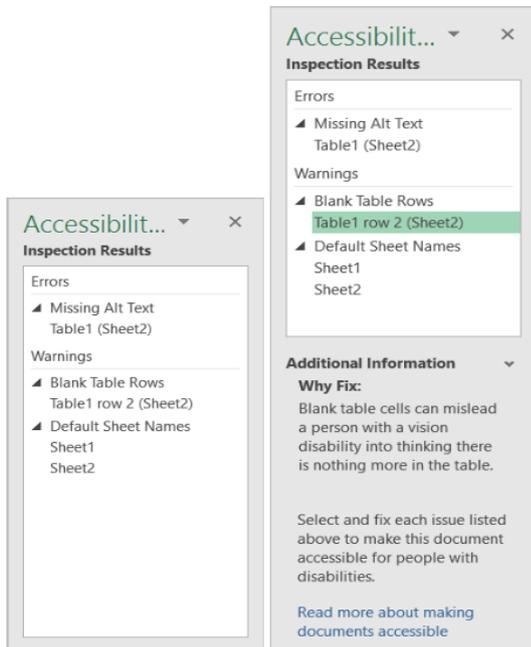
If you have an older version of Word, such as Word 2010, and do not see the Check Accessibility button on the Review tab on the Ribbon, follow these steps to open the Accessibility Checker:

1. Click **File > Info**.
2. Select the **Check for Issues** button.
3. In the Check for Issues drop-down menu, select **Check Accessibility**.



**Picture IV-38:** How to access Accessibility Checker in an older version of Word.

The Accessibility Checker Task Pane appears next to your content and shows the inspection results. To see details about an issue, including why and how to fix it, select an issue under Inspection Results. Results appear under Additional Information, and you are directed to the inaccessible content in your file.



**Picture IV-39:** An example of the Accessibility Checker report.

## 4.2 Understand the Inspection Results

After Accessibility Checker inspects your content, it reports the inspection results based on the severity of the issues found, categorised as follows:

- **Errors:** Issues reported as Errors include content that is very difficult or impossible for people with disabilities to understand.
- **Warnings:** In many cases, Warnings mean that the content is challenging for people with disabilities to understand.
- **Tips:** Tips let you know that, even though people with disabilities can understand the content, it could be better organized or presented to improve their experience.

## 5 Further resources

Benetech. (2017). *Poet Image Description*. <https://poet.diagramcenter.org/>

Benetech. (2019). *Diagram Center*. <http://diagramcenter.org/>

*Books for all: a starter kit for accessible publishing in developing and least developed countries*. (b. d.). Accessible Books Consortium.

[https://www.researchgate.net/publication/328280949\\_A\\_Starter\\_Kit\\_for\\_Accessible\\_Publishing\\_In\\_Developing\\_And\\_Least\\_Developed\\_Countries](https://www.researchgate.net/publication/328280949_A_Starter_Kit_for_Accessible_Publishing_In_Developing_And_Least_Developed_Countries)

Ganner, J. et al. (2023). *Books without barriers: a practical guide to inclusive publishing*. Professional Editors and Australian Publishers Association. <https://www.ipeditors.org/resources-for-editors/books-without-barriers/>

- DAISY Consortium. (n. d.). *Creating accessible Word documents*. <https://daisy.org/info-help/guidance-training/daisy-tools/creating-accessible-word-documents/>
- DAISY Consortium. (15. 3. 2021). *Word Document Accessibility 101* [Video]. YouTube. <https://www.youtube.com/watch?v=oxjS7zpm6uw>
- DAISY Consortium. (13. 4. 2021). *Word Document Accessibility Part 2* [Video]. YouTube. <https://www.youtube.com/watch?v=x065kj23wcg>
- DAISY Consortium. (2024). *DAISY Learning*. <https://learning.daisy.org/>
- EBU clear print guidelines. (2016). European Blind Union. <http://www.euroblind.org/sites/default/files/media/ebu-media/Guidelines-for-producing-clear-print.pdf>
- European Blind Union. (n. d.). *Making information accessible for all*. <https://www.euroblind.org/publications-and-resources/making-information-accessible-all>
- Mcnaught, A. (2013). *Libraries and alternative format research*. [https://www.researchgate.net/publication/267266966\\_Libraries\\_and\\_alternative\\_format\\_research](https://www.researchgate.net/publication/267266966_Libraries_and_alternative_format_research)
- Microsoft. (n. d.). *Make your Word documents accessible to people with disabilities*. Microsoft Support. <https://support.microsoft.com/en-us/office/make-your-word-documents-accessible-to-people-with-disabilities-d9bf3683-87ac-47ea-b91a-78dcacb3c66d>
- Rayini, J. (2017). Library and information services to the visually impaired persons. *Library Philosophy and Practice*, 2017(1). [https://www.researchgate.net/publication/317590601\\_Library\\_and\\_information\\_services\\_to\\_the\\_visually\\_impaired\\_persons](https://www.researchgate.net/publication/317590601_Library_and_information_services_to_the_visually_impaired_persons)
- Sutton, J. (2002). *A guide to making documents accessible to people who are blind or visually impaired*. American Council of the Blind. <https://www.sabeusa.org/wp-content/uploads/2014/02/A-Guide-to-Making-Documents-Accessible-to-People-Who-are-Blind-or-Visually-Impaired.pdf>
- University of Kent. (2024). *Accessible content*. <https://www.kent.ac.uk/guides/accessible-content>

## **V. MODULE: CREATING ACCESSIBLE EPUB AND HTML FORMATS**

### **About the module**

This course will focus on how librarians can prepare accessible EPUB and HTML publications from already created accessible Word documents from Module IV. This module is divided into seven sections, primarily focusing on preparation before conversion, key points during conversion, and necessary steps after conversion. Everything is based on accessibility guidelines and best practices used in the publishing industry.

All processes in this module are designed to ensure that users of mobile devices have easier access and blind or partially sighted users have better navigation, structure, and understanding of the content, whether in EPUB or HTML format. The course is structured so that you can choose the workflow for one of the presented formats or both.

The formats were chosen because they are both based on markup, supported by various reading software or browsers, allow users to adapt the visual appearance of the book, and can be used on various devices due to reflowing text, independent of screen size.

Summary of the Course:

1. What is needed before using the WordToEpub tool?
2. WordToEpub preferences before converting the first book
3. Converting the book to EPUB
4. Editing EPUB in Sigil
5. Accessibility check
6. Converting the book to HTML
7. Further resources

Sources used for this course were mostly:

- Research results from the EODOPEN project and practical experiences in adapting materials in different formats.
- Adapted learning materials, with permission, publicly available on the [DAISY Learning webpage](#), created by the DAISY Consortium.
- Gathered sources and literature during the duration of the EODOPEN project.
- Practical project experience from Slovenia in creating accessible EPUB publications from Microsoft Word. Practical example was presented in one of the EODOPEN meetings: [5th EODOPEN project meeting 29+30 November 2021 - EBook accessibility for all.](#)

## Training Scenarios/How to Deliver the Content

This course is provided as self-learning educational material, comprised of theory and graphical elements. By adapting it, it could also be used as training for interested target groups (maximum 10 people per course). However, as the course is lengthy, detailed, and includes practical work, we decided to keep it as a self-learning course.

## Duration of the Course

The duration of the course is approximately 5-6 hours.

## Training Material Needed

For this course, you will need:

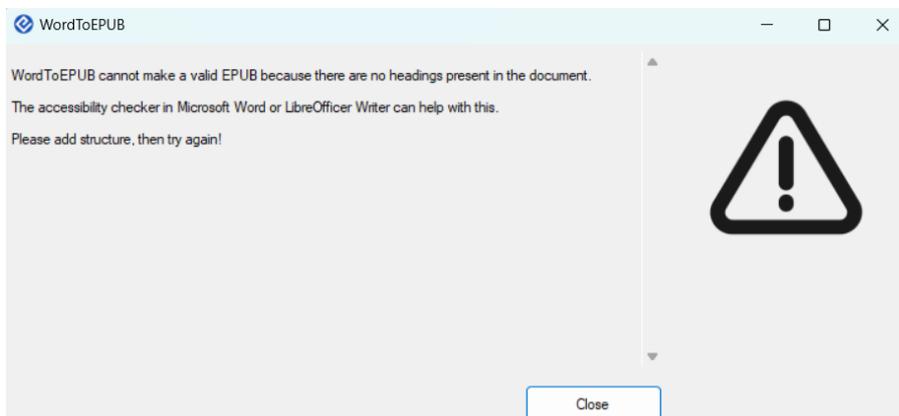
- Basic knowledge of HTML and CSS.
- A finished accessible Word document from [Module IV: Creating Accessible Document Format](#)
- Software for conversion to EPUB and accessibility check: [WordToEpub](#), [Sigil](#), [EpubChecker](#), [Ace by DAISY](#), Thorium Reader (available in Microsoft Store) or other eBook reading software.
- Software for conversion to HTML and accessibility check: WordToEpub, any browser.
- Assessment questionnaire in Annex 1

## Additional Notes

This course and all the material are based on the WordToEpub tool, which is created and supported by The DAISY Consortium. For this course, we used version 1.0.11. If the participant is using a newer version, some functions could be different.

## **1 What is needed before using the WordToEpub tool?**

The source Word document should be prepared according to Module IV to achieve the best results regarding accessibility. Ensure you check the Word document's accessibility before starting the conversion process. The WordToEpub tool will notify you during the conversion if there are any problems with the source document. For example, if there is no heading structure inside the source document, it will automatically reject the conversion (see **Picture V-1**).



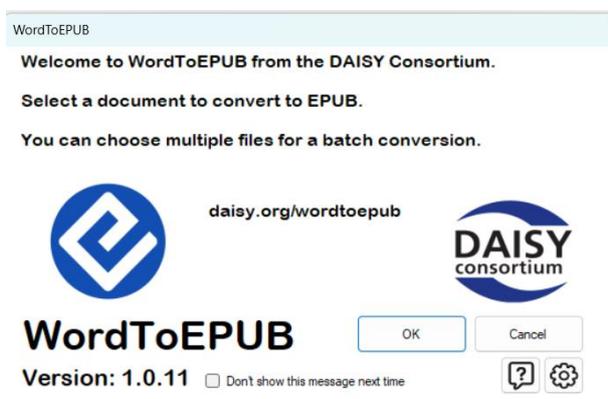
**Picture V-1:** WordToEpub rejected a document without a heading structure.

If a library wishes to use a specific cover for all adapted books, it is recommended to use a cover template that includes, for example, library or project logos. The title and author can be added as well or later during the conversion process. If you need the cover template, prepare it beforehand. The templates included in WordToEpub are sized 1000x1600px or around 1200x1600px.

## 2 WordToEpub Preferences Before Converting the First Book

Before starting to use the WordToEpub tool, some initial preferences/settings need to be applied to simplify further conversions. This step only needs to be done once and will remain set for all future conversions. Follow the steps described below.

To access the preferences, first open the tool. When the greeting window opens, choose the **settings** button in the bottom right corner. You can also tick the option **Don't show this message next time** if you wish to avoid it and go directly into conversion (from where you can still access the preferences).

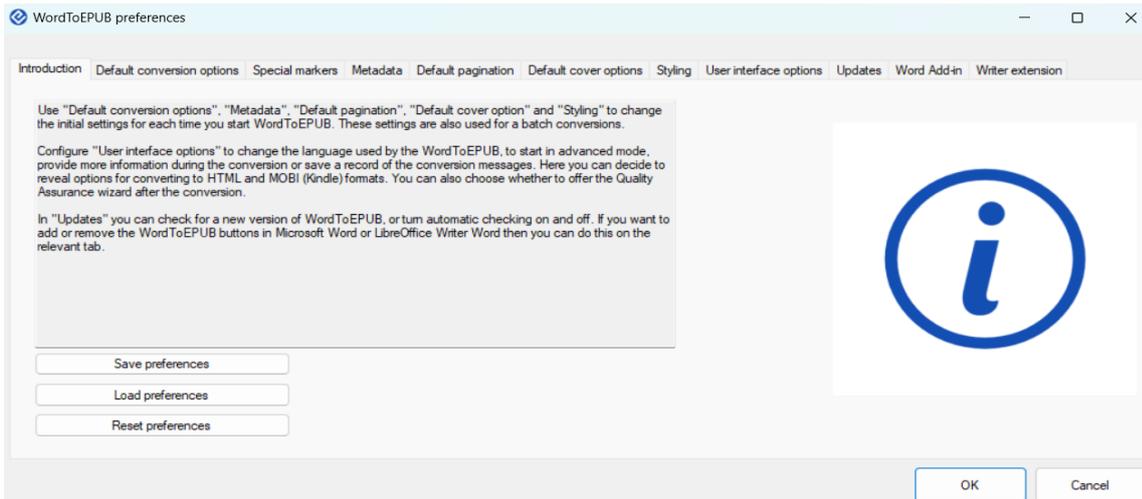


**Picture V-2:** Greeting window of WordToEpub.

Preferences open in a new window where the following tabs are available: Introduction, Default Conversion Options, Special Markers, Metadata, Default Pagination, Default Cover Options, Styling, User Interface Options, Updates, Word Add-in, and Writer Extension. As the

Introduction tab explains, the first tabs are intended to change the initial settings for each time you start WordToEpub, and these settings are also used for batch conversions. The last four tabs are intended for the user to set settings for WordToEpub, and add/remove buttons in Microsoft Word or LibreOffice Writer.

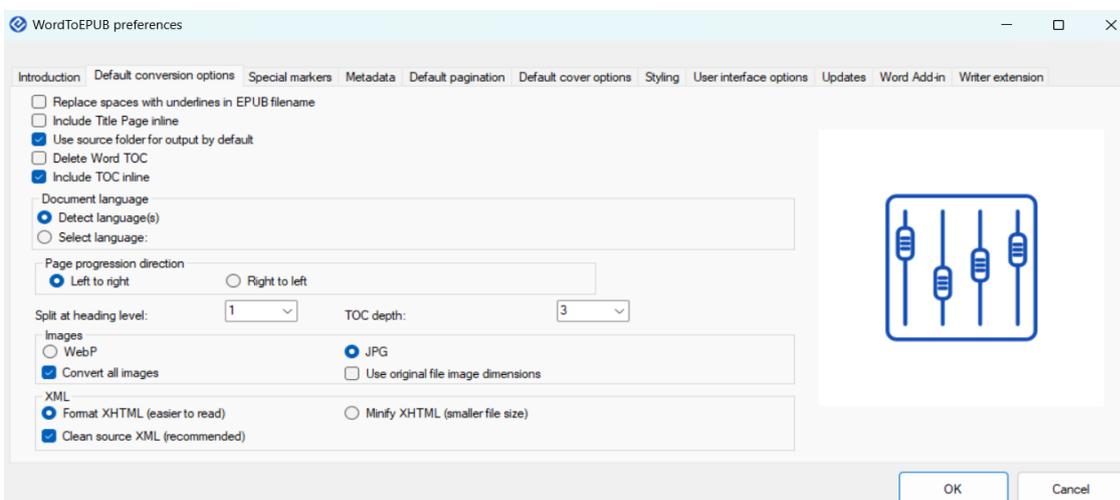
Each of the settings and preferred choices are further described in the following chapters.



**Picture V-3:** WordToEpub preferences window.

## 2.1 Default Conversion Options

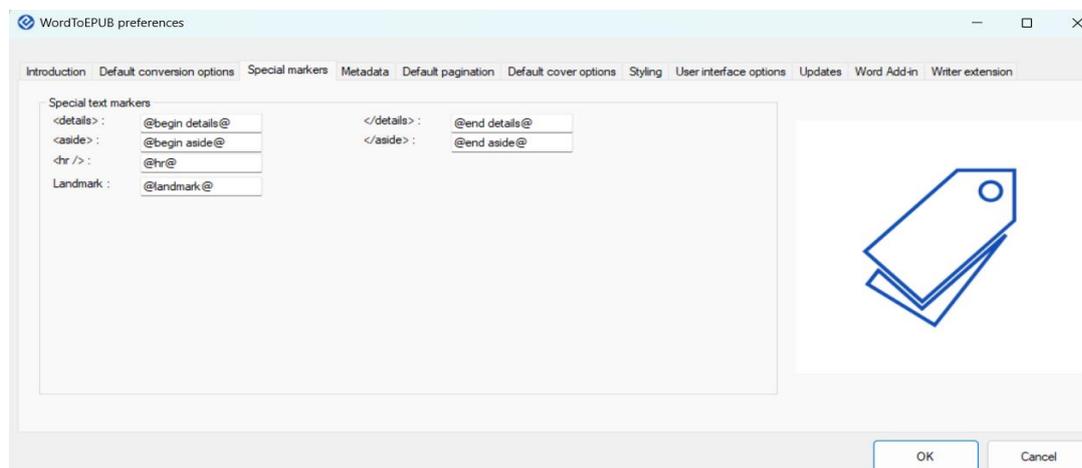
In this tab (**Picture V-4**), the most important preferences are: options to enable saving the outcome in the same folder as the source file, to include the table of contents according to marked headings in Microsoft Word, to detect language(s), to keep page progression direction from left to right, to split the EPUB at level 1 and Table of Contents to depth level 3, to convert images in the document to JPG, to use format XHTML (which is easier to read), and to recommend XML clean source.



**Picture V-4:** Window of preferences for default conversion options.

## 2.2 Special Markers

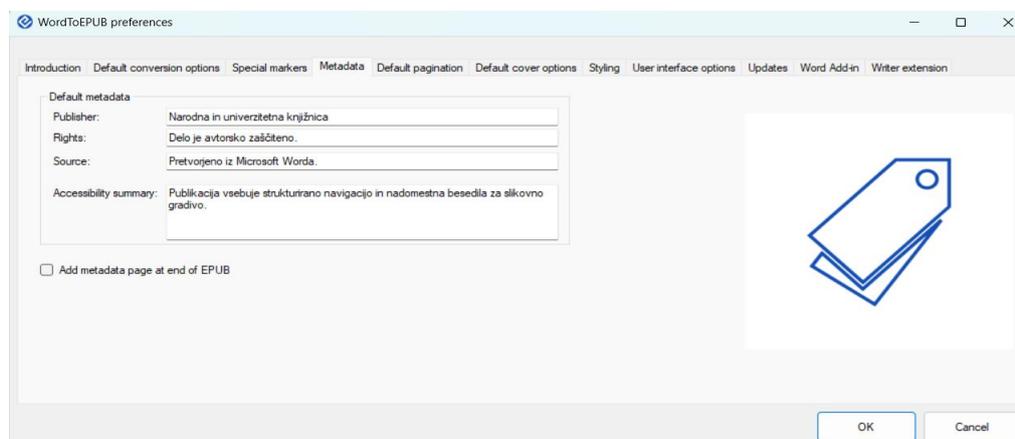
In this tab (**Picture V-5**) no changes are needed. It is a list of special markers in case they are marked in Microsoft Office Word before conversion.



**Picture V-5:** Window of preferences for special markers.

## 2.3 Metadata

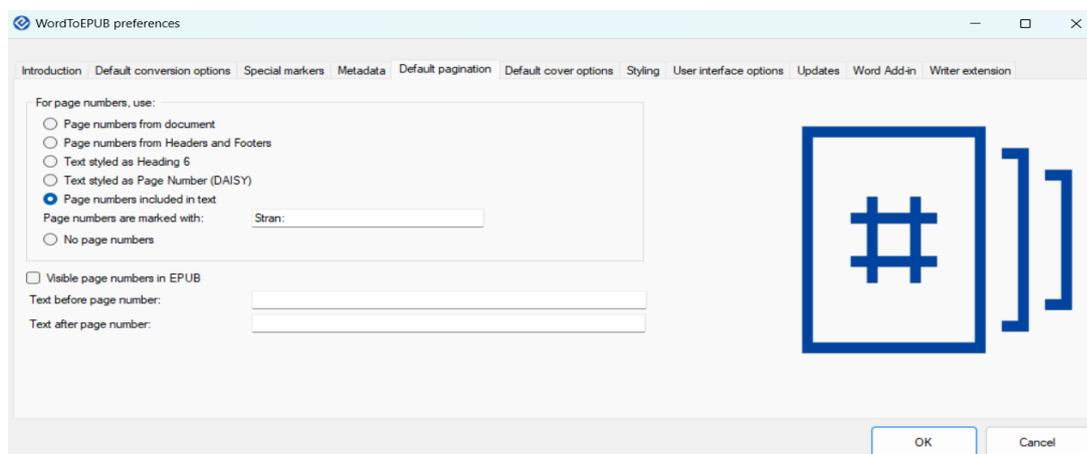
In this tab (**Picture V-6**), some basic metadata can be inserted, which will then be used for all converted publications. In the example below, the name of the National and University Library is inserted into the metadata for the publisher. Additionally, the rights can be adapted based on the copyright status: either "The work is copyright protected" or "The work is in the public domain." It is also possible to use CC licences. Under "Source," the example has written: "Converted from Microsoft Word." Under "Accessibility Summary," the example has written: "The publication contains structured navigation and alternative text for graphic materials." Other descriptions can also be included. There is also an option to add a metadata page at the end of the EPUB, but at the National and University Library of Slovenia, we decided not to use this option.



**Picture V-6:** Window of preferences for metadata.

## 2.4 Default Pagination

In this tab (Picture V-7), you can choose how to show the pagination of the book. Since this module is about converting a digitised book into an adapted version, the best choice is to use the option “page numbers included in text” and written as we marked the original pages inside the Word document (example: Page: ). For more about adding original pages to the Microsoft Word document, see *Module IV: Creating Accessible Document Format*, chapter Advanced Word Accessibility, sub-chapter [Page Numbering Strategy for Print to Digital Conversion](#). Other options are more useful for documents that are originally digital.



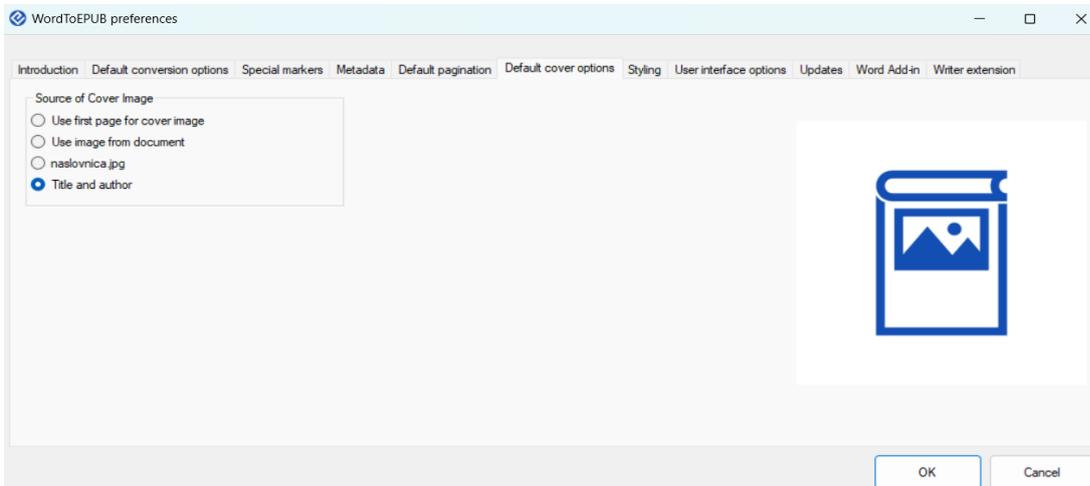
**Picture V-7:** Window of preferences for default pagination.

## 2.5 Default Cover Option

In this tab (Picture V-8) you can choose how the cover of the EPUB publication will look. Either you can choose the **first page** for the cover image, **an image from the document**, or your **personalised cover** (e.g., naslovnica.jpg) which we mentioned in the previous chapter and should also contain the title and author written on the cover.

Last option is **Title and Author**, which allows you to design the cover directly inside the tool. What is needed is only the background cover image/template with library and/or project logos. This option will be described further during the conversion process in chapter [Cover Image](#).

If you decide that the publication will not have a cover, that option is also available further during the conversion process.

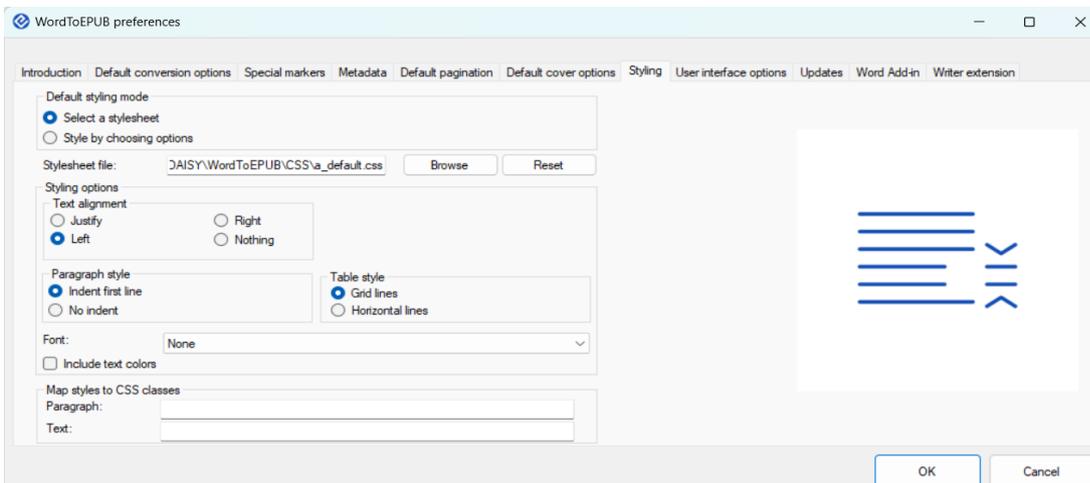


**Picture V-8:** Window of preferences for default cover options.

## 2.6 Styling

In this tab (**Picture V-9**) you can choose one of the stylesheets offered by the tool (accessed by choosing the **Browse** button) or add a stylesheet you created yourself. From the ones offered, suggested is the stylesheet **a\_default.css**. If this one isn't available anymore, choose other style or keep the one that is chosen by default.

Furthermore, you can define the text to be left-aligned, have the first line in a paragraph indented, and add grid lines to tables. The other options are best left unchosen.

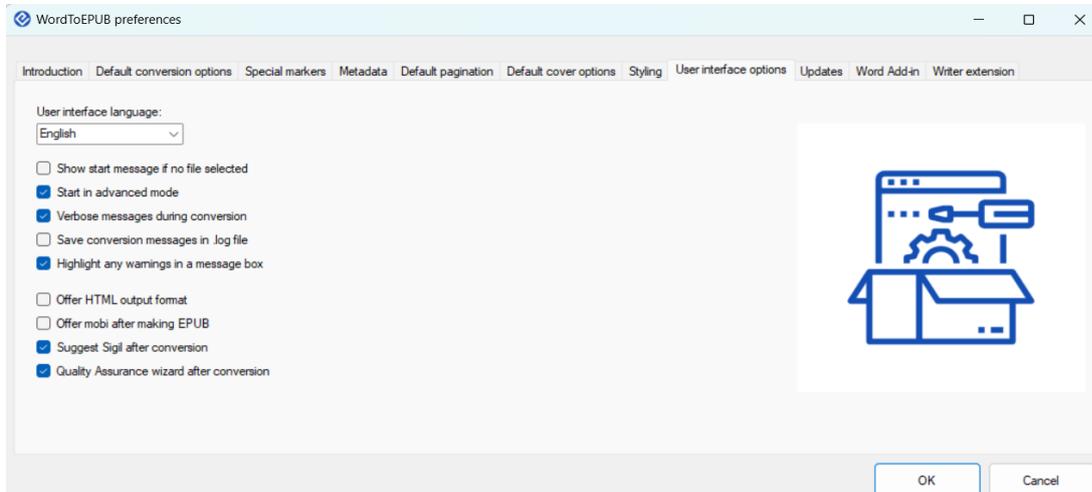


**Picture V-9:** Window of preferences for styling.

## 2.7 User Interface Options

In this tab (**Picture V-10**), the language of the user interface is offered in 16 different languages. The recommended ticks to choose are: Start in Advanced Mode, Verbose Messages During Conversion, Highlight Any Warnings in a Message Box, Suggest Sigil After Conversion,

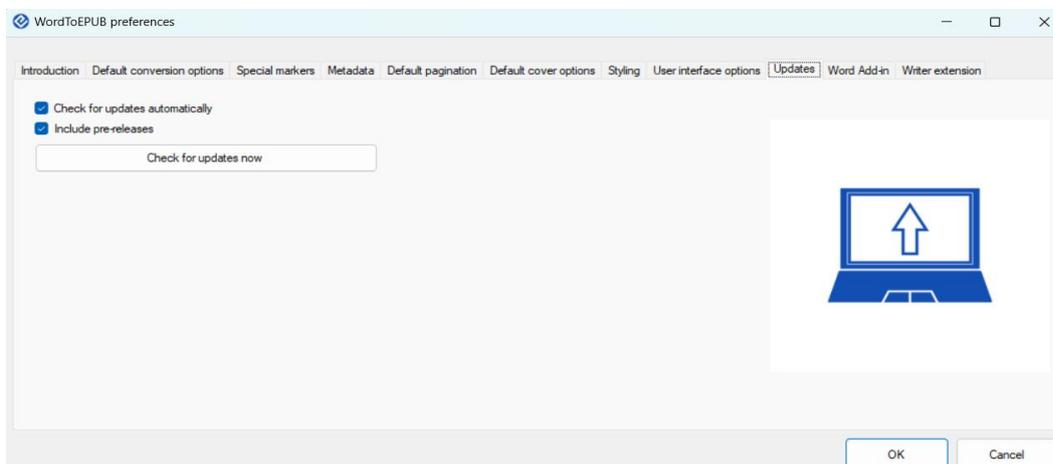
and Quality Assurance Wizard After Conversion. Options for HTML and MOBI formats are optional, depending on whether you choose to convert to those formats as well.



**Picture V-10:** Window of preferences for user interface options.

## 2.8 Updates

In this tab (**Picture V-11**), choose to check for updates automatically as the software is continually developing. If you want to test new releases before they become official, you can also choose to include pre-releases.



**Picture V-11:** Window of preferences for updates.

## 2.9 Word Add-in and Writer Extension

In these last two tabs, you can install the extensions for Microsoft Word and/or LibreOffice Writer. With the extensions installed, you can run WordToEpub directly from Microsoft Word or LibreOffice.

## 3 Converting the Book to EPUB

### 3.1 Opening the Word Document in WordToEpub

To open the document in WordToEpub and start the conversion process, several options are available:

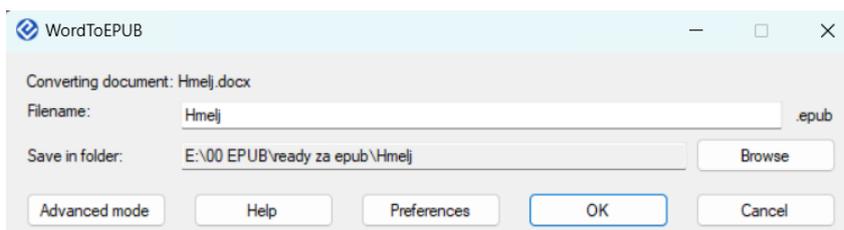
- **Drag the Word document** and drop it onto the icon of the WordToEpub tool.
- **Right-click** on the Word file > **Show More Options** > **Convert with WordToEpub**.
- **Open the tool** and **navigate to the location** where the document is saved on the computer.
- Use installed extensions **directly from the Word editor**.

No matter which path you chose, the simple or advanced settings for conversion will open which are further described in the following chapters.

### 3.2 Settings Before Conversion

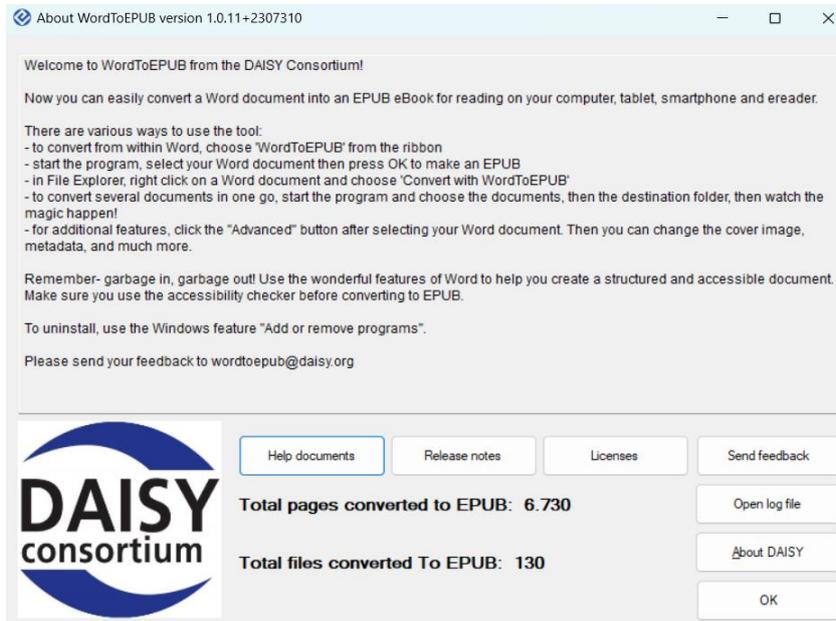
Once the document is opened with the tool, the process starts immediately. A window with a simple mode of settings appears, as shown in **Picture V-12**. These simple settings use the preferences already set for the conversion. The only changes you can make are to the filename and saving location of the output format.

The available buttons are: Advanced Mode, Help, Preferences, OK to start the conversion process, or Cancel.



**Picture V-12:** Window of the simple mode in WordToEpub.

The Help button is useful as it provides further guidance on using the tool, help documents, release notes, licenses, statistics about conversion, and other information.

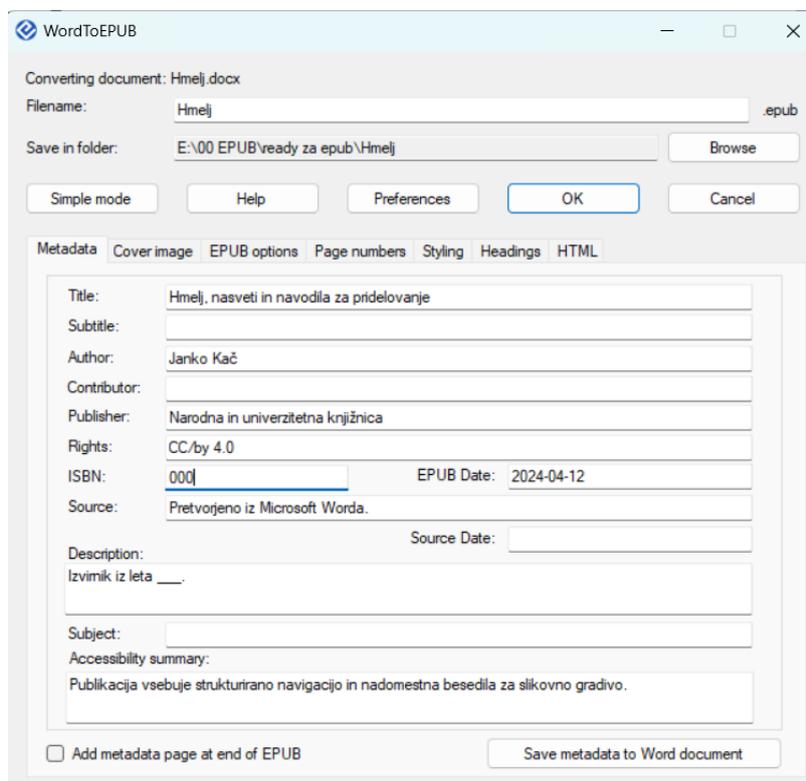


**Picture V-13:** Window of the help options in WordToEpub.

It is suggested to open the tool in advanced settings mode. If you have already set advanced mode in preferences, the tool will automatically open in advanced mode. The window looks the same at the top but expands below with additional tabs for options: Metadata, Cover Image, EPUB Options, Page Numbers, Styling, and Headings.

After checking all the tabs, which are further described below, you can click OK to start the conversion process.

If you choose the option to convert to HTML in preferences, an additional tab appears at the end, as shown in **Picture V-14**. For further conversion to HTML, see chapter: [IV. Converting the Book to HTML](#).



**Picture V-14:** Window of advanced settings in WordToEpub.

### Metadata

This tab (**Picture V-14**) presents the metadata of the individual publication. If at least some basic metadata were inserted in the process of making the Word document (see [7. File Name, Properties and Exporting](#)), they are automatically inserted. The same happens with the information you provided in the preferences (Publisher, Rights, Source, Accessibility Summary).

At this point it is important to **check if the data is correct, edit, or add any missing data** (e.g., description at the image states the year of the original publication).

If you will get the ISBN number later in the process for the publication, it is recommended to enter a placeholder here (e.g., 000), as it will be easier to discover and replace later (see Chapter 4: Editing EPUB with Sigil > [Editing Metadata](#)).

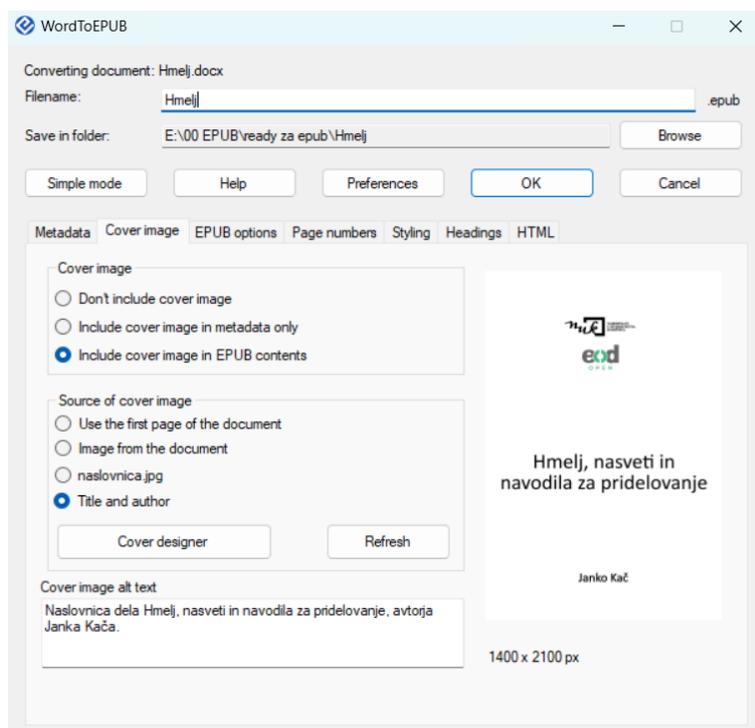
When all the metadata is correct, click **Save Metadata to Word Document**. That way, if the conversion needs to be repeated, the metadata will already be in the Word document.

Then go to the next tab, which is the Cover Image.

### Cover Image

This tab (**Picture V-15**) is intended for further work and settings regarding the cover image.

1. If you do not want to use the cover, choose the option **Don't include cover image**.
2. If you want the publication to have a cover, it is recommended to use the option **Cover Image in EPUB content**, which will then appear as a first element in the EPUB book. Additionally, to design the cover inside the tool, use the source of the cover image **title and author**.



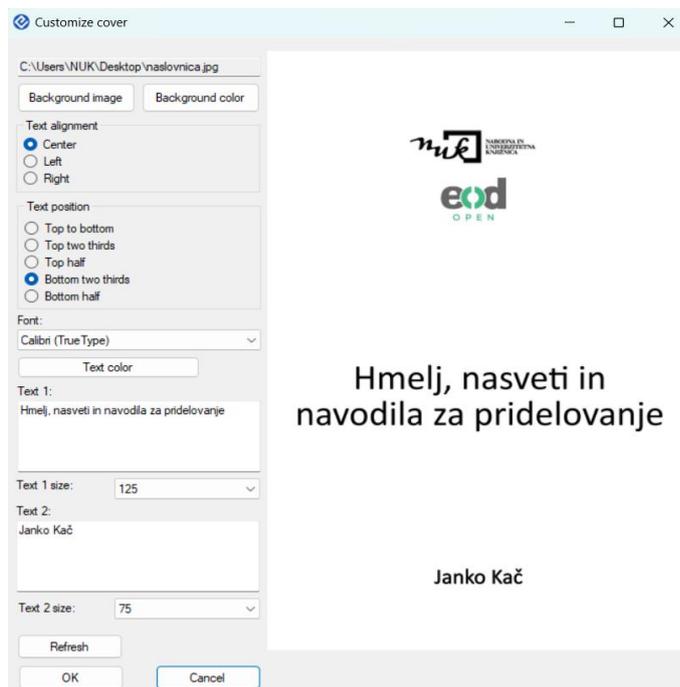
**Picture V-15:** Window of cover image settings in WordToEpub.

Further design of the cover is inside the cover designer (Picture V-17). There you can add your own background image (the example contains a white background and logos of the National and University Library of Slovenia and EODOPEN). The text for the author and title is automatically filled out from the metadata, but you can edit the text if desired.

You can also choose text alignment and position options, fonts, text colours, and text sizes.

A refresh button is useful to see the changes immediately. The same refresh button is available outside the cover designer.

Lastly, alternative text for the cover is also needed. The example from Picture V-16 only contains text explaining that the image is a cover of the book *Hmelj, nasveti in navodila za pridelovanje* and the author's name.

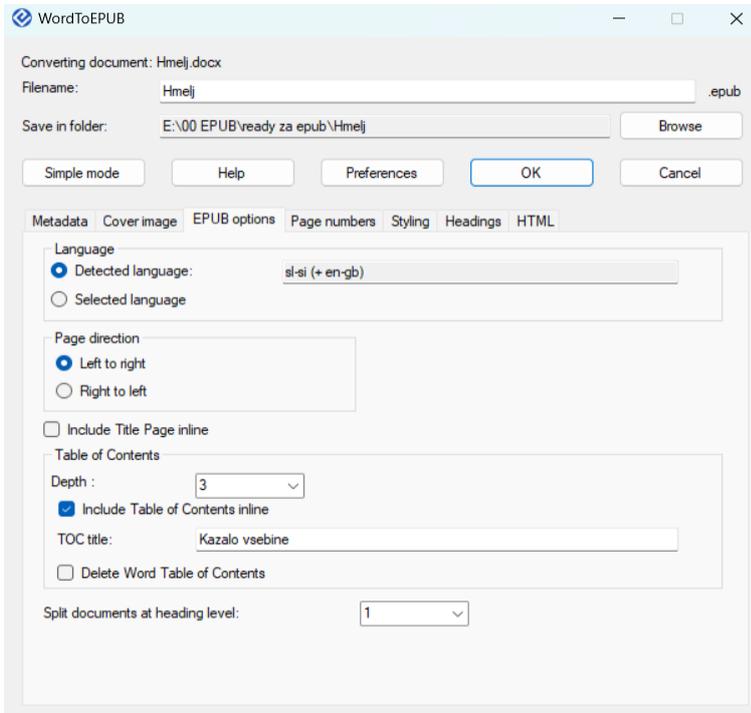


**Picture V-16:** Window of cover designer settings in WordToEpub.

### EPUB Options

This tab (**Picture V-17**) is intended to check predetermined preferences and to add other elements. Recommended settings are as following:

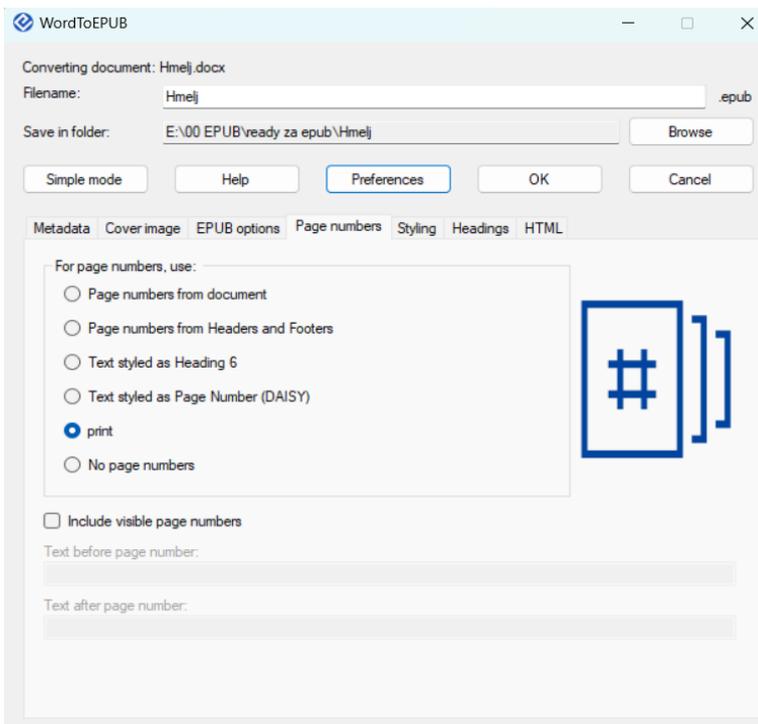
- **Check detected language** - the language should be the same as the main text of the book or the language that was marked in the Microsoft Word document.
- **Page direction should be left.**
- Ensure the option **include Table of contents inline** is checked, which will make the Table of Contents (TOC) appear at the beginning of the book. You can determine the title of the TOC in the same language as the document (e.g., Slovenian: Kazalo vsebine). The **depth of the TOC should be set to 3** (the number represents the marked headings until level 3). Additionally, if there is already a TOC inside the Word document, tick the option **Delete Word Table of Contents**.
- **Split the document at level 1** - this means the EPUB will be split into separate files according to heading level 1.



**Picture V-17:** Window of EPUB options settings in WordToEpub.

### Page Numbers

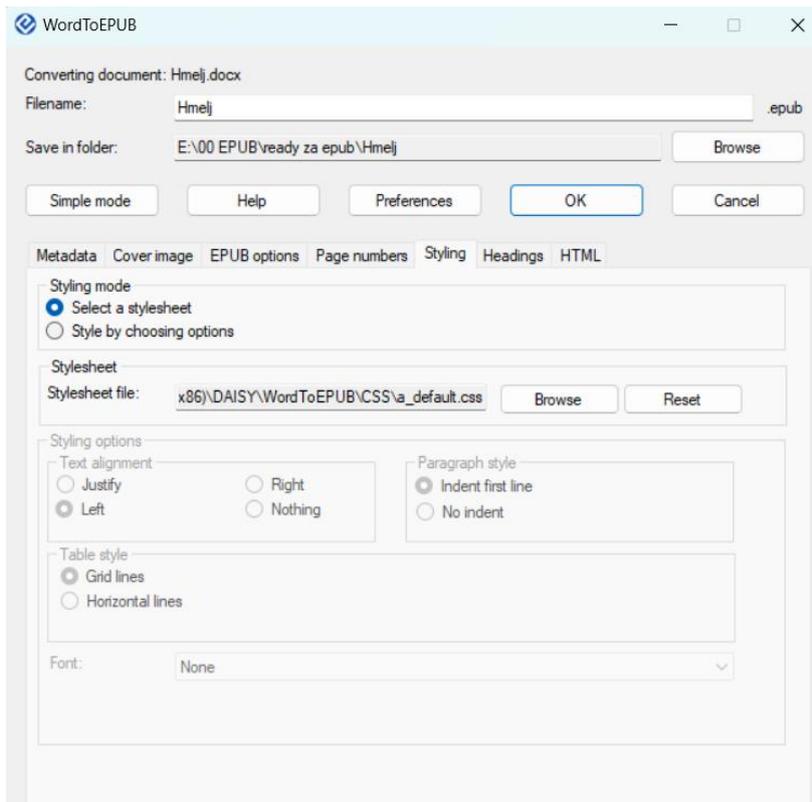
This tab (**Picture V-18**) is intended to check how the original page numbers will be dealt with. We determined this already in the preferences, but here the option **print** should be selected.



**Picture V-18:** Window of page numbers settings in WordToEpub.

## Styling

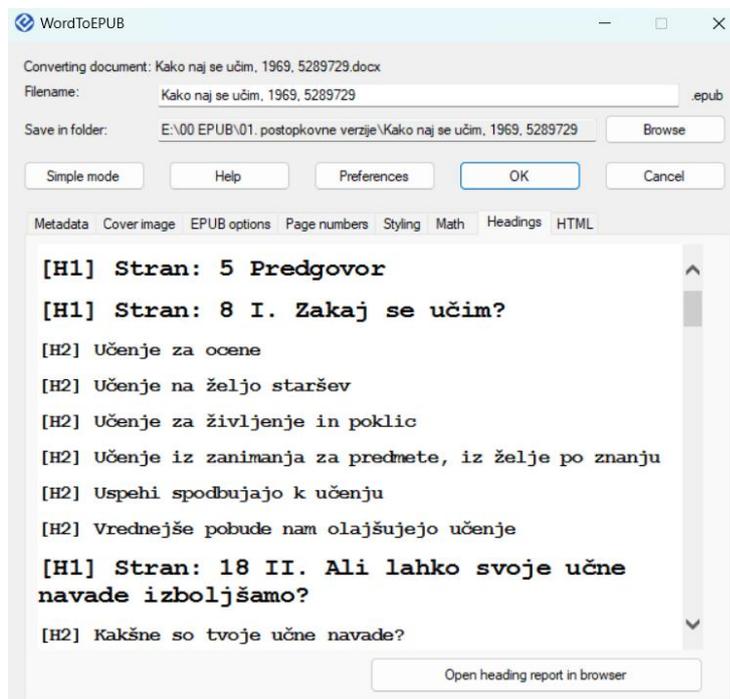
This tab (**Picture V-19**) is intended to check or determine styling - choose a stylesheet. We determined this already in the preferences, but here the option for styling mode should be **Select a Stylesheet**. Below it shows the stylesheet a\_default.css. At this point, you can still change it to a different one, regardless of the set preferences.



**Picture V-19:** Window of styling settings in WordToEpub.

## Heading

This tab (**Picture V-20**) is intended to check if the heading order is correct. In case of any issues, the errors are highlighted in red. The example shows that there are no issues with the heading order.



**Picture V-20:** Window of headings tab in WordToEpub.

In case of complex heading structure, the option to open a heading report in the browser is useful (**Picture V-21**).

Heading report from WordToEPUB for Kako naj se učim, 1969, 5289729.docx

```

h1 Kolofon
h1 Stran: 5 Predgovor
h1 Stran: 8 I. Zakaj se učim?
h2 Učenje za ocene
h2 Učenje na željo staršev
h2 Učenje za življenje in poklic
h2 Učenje iz zanimanja za predmete, iz želje po znanju
h2 Uspehi spodbujajo k učenju
h2 Vrednejše pobude nam olajšujejo učenje
h1 Stran: 18 II. Ali lahko svoje učne navade izboljšamo?
h2 Kakšne so tvoje učne navade?

```

**Picture V-21:** Window of headings report in a browser.

### 3.3 During the Conversion and Conversion Report

Once you check all the tabs and apply the settings, you can start the conversion process by clicking **OK**. The tool then analyses and converts the content of the book, creating the EPUB according to your preferences and settings.

Once the conversion is completed, it offers you the conversion report so you can immediately check if there are any issues. Here is an example of a conversion report, with the parts that should always be checked marked.

```
Now converting: Hmelj.docx
Starting...
Making a safe copy of your document
File copy success
Getting the names of heading styles
Reading metadata
Analyzing headings
Analyzing languages...
Simplify Word markup
Numbered headings
Marking the pages
Number of pages marked: 40
Captions: 8...
Language set to: sl-si
Added language markup into content: en-gb
Described images: 6, decorative images: 3, missing alt text: 0
Generating metadata for the EPUB
Building your EPUB with pandoc .
Upgrading the EPUB with your choices
Generating metadata for the EPUB
Finished (3 s)
The Word document of 40 pages has been converted to:
Hmelj.epub
in the folder:
E:\00 EPUB\ready za epub\Hmelj
Your publication in numbers:
Number of words: 11938
Number of paragraphs: 159
Number of images: 9
Number of described images: 6
Number of decorative images: 3
Number of languages: 2
Filesize (bytes): 377.554
No warnings :)
```

At the same time, the tool already offers the next step required in the conversion process >  
**Edit EPUB with Sigil.**

Once you click that, the EPUB will open in the software Sigil. More about this follows in the next chapter.

**At this point, leave the tool opened at all times until all the steps are completed!**

In case you mistakenly close the tool, simply follow the next chapters by opening the generated EPUB in the corresponding software as described.

## 4 Editing EPUB in Sigil

If you follow the steps in the WordToEpub tool and click Edit EPUB in Sigil after the first step in the conversion, or if you open the EPUB file in Sigil, the book will open inside the software.

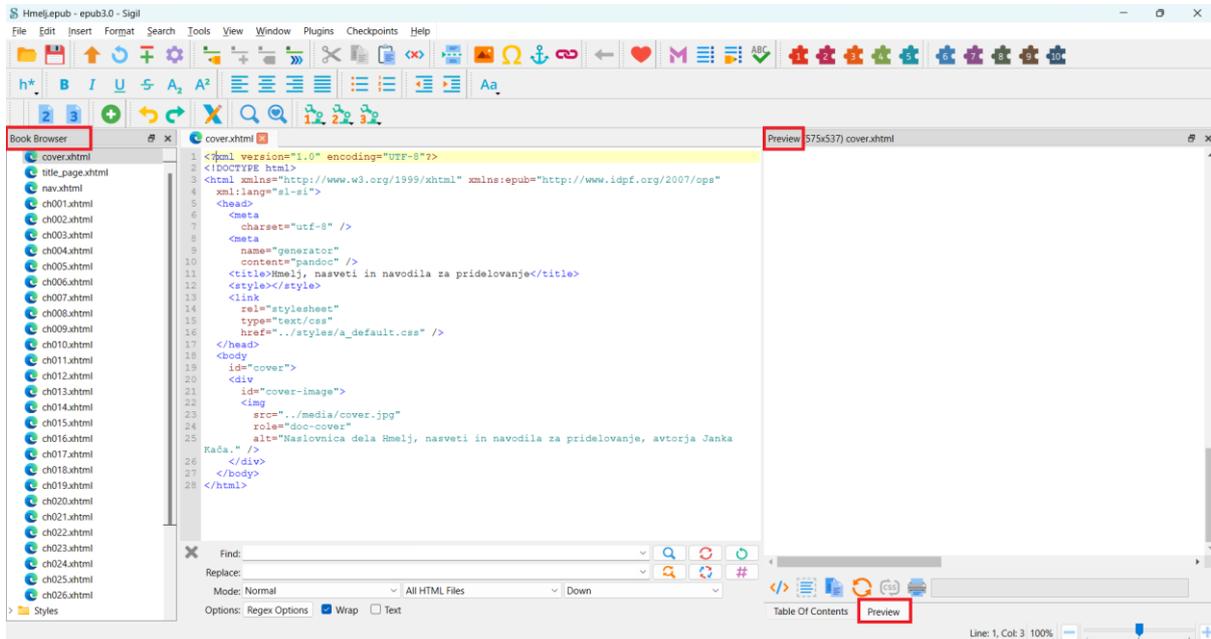
To work with this software, some basic knowledge of HyperText Markup Language (HTML) is required. It is important to know that each HTML element consists of content, opening, and closing tags. For example, the title of this chapter would be: `<h2>4 Editing EPUB in Sigil</h2>`. The content of this element is the title of the chapter, and the opening and closing tags for heading level 2 are `<h2>` and `</h2>`.

If you need more help or assistance, check: [Learn more about HTML elements.](#)

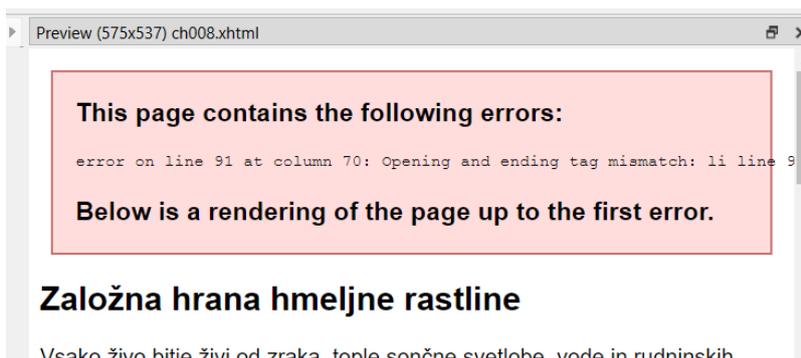
When working with Sigil, here is some starter information for using the software.

The user interface is divided into three sections (**Picture V-22**):

- **Book Browser** on the left: This shows the entire package of the eBook. It contains XHTML files for each chapter divided by heading level 1, which is the actual content of the book. Other elements include styles, images, and other multimedia files, toc.ncx, and an important file we will use later, content.opf. To open each file, double-click on it, and the content will appear in the editing space.
- **Space for editing** in the middle: This is where you can edit all the files you opened from the book browser by double-clicking on them. Each element opens in a separate tab. Search and replace options can also be useful in some cases. More about working in this space will be further explained in the following chapters.
- **Preview window** on the right: All changes made in the editing space are immediately visible on the right. It also displays errors if the code is not correctly edited (**Picture V-23**). Ensure the Preview tab at the bottom is active and not the Table of Contents tab.



**Picture V-22:** User interface of Sigil.



**Picture V-23:** Example of an error notification in the preview window, which explains exactly where the error is.

After WordToEpub creates the EPUB file, it is your responsibility to check if the conversion was done properly. Sigil is a space where you can still fix anything regarding the content, visual appearance, metadata, etc. These are further presented below.

#### 4.1 Editing Content

WordToEpub converts the content from Microsoft Word as you have marked it in the source file. Sometimes mistakes happen when preparing the source file or during the conversion. Below are some examples of what should be checked or fixed regarding content:

- **Check Split Files According to Heading 1**

In the settings of WordToEpub, it was specified that the files should be split by heading 1. By clicking on the XHTML files in the book browser and viewing them in the space for editing or preview, it is important to see if the split of the files is correct.

- **Check Landmarks**

WordToEpub automatically creates landmarks for the cover, table of contents, and beginning of the content. The language of these landmarks should correspond to the main language of the document. Check inside the nav.xhtml file to ensure all the landmarks are in your language; if not, correct it (the black text). To find the region faster, type "landmarks" in the search bar.

```

162 <nav
163   epub:type="landmarks"
164   id="landmarks"
165   aria-label="Mejniki"
166   hidden="hidden">
167   <ol>
168     <li>
169       <a
170         href="text/cover.xhtml"
171         epub:type="cover">Naslovnica</a>
172     </li>
173     <li>
174       <a
175         href="nav.xhtml#toc"
176         epub:type="frontmatter">Kazalo vsebine</a>
177     </li>
178     <li>
179       <a
180         href="text/ch001.xhtml"
181         epub:type="bodymatter">Začetek vsebine</a>
182     </li>
183   </ol>
184 </nav>

```

**Picture V-24:** Example of landmarks in the Slovene language.

- **Check original page numbers**

Check if all original page numbers are present by opening the file nav.xhtml. A bit lower in the code, there is a page list that looks like **Picture V-25**. Check if the numbers in black text are in the correct order and all present.

```

185 <nav
186   role="doc-pagelist"
187   id="page-list"
188   hidden="hidden"
189   aria-label="Seznam strani"
190   epub:type="page-list">
191   <ol>
192     <li>
193       <a
194         href="text/ch002.xhtml#page3">3</a>
195     </li>
196     <li>
197       <a
198         href="text/ch002.xhtml#page4">4</a>
199     </li>
200     <li>
201       <a
202         href="text/ch003.xhtml#page5">5</a>
203     </li>
204     <li>
205       <a
206         href="text/ch005.xhtml#page6">6</a>

```

**Picture V-25:** Example where list of original pages is visible.

- **Check Where the Page Breaks are**

You can check this in each XHTML file of the content. As you wrote the original pages in the source files, some additional changes may need to be made if a page appears in the middle of a sentence or if a mistake was discovered.

This is how a page break should look:

```
<span
epub:type="pagebreak"
role="doc-pagebreak"
id="page6"
aria-label="Stran 6. "></span>
```

If you need to move the page somewhere else in the text, the whole span element needs to be moved.

- **Check Figures and Alternative Text**

The figure element in the current version of the tool WordToEpub does not always properly convert captions from the Word document and may need manual adjustment. The figure usually consists of the image (src, style, and alt attributes) and figcaption. Figcaption may not be present if the image had no caption in the source file. The alt attribute is empty only when the image is decorative and does not add additional value to the surrounding text. Example: alt=""

This should be checked to ensure the form is correct. The basic form for the figure element should be as follows:

```
<figure>

<figcaption>Figure 1: Text under the image.</figcaption>
</figure>
```

```
33
34 <figure>
35 
39 <figcaption>Slika 1: 1. Trta se ovija na desno. 2. Kos hmeljevine s
plezalnimi dlačicami. 3. Plezalna dlačica. 4. Lat moških cvetov. 5. Prašnični
cvet. 6. Prašnik. 7. Cvetni prah. 8. Lat pestičnih cvetov. 9. Pestično
socvetje ali storžek. 10. Krovni list z dvema pestičnima cvetoma. 11. Vejica
storžkov. 12. Krovni list storžka s plodom. 13. Žleza. 14. Žleza v prerezu.
15. Plod. 16. Prerezano seme. 17. Štor se obreže na čepe pri s – s
40 </figcaption>
41 </figure>
42
43
```



Slika 1: 1. Trta se ovija na desno. 2. Kos hmeljevine s plezalnimi dlačicami. 3. Plezalna dlačica. 4. Lat moških cvetov. 5. Prašnični cvet. 6. Prašnik. 7. Cvetni prah. 8. Lat pestičnih cvetov. 9. Pestično socvetje ali storžek. 10. Krovni list z dvema pestičnima cvetoma. 11. Vejica storžkov. 12. Krovni list storžka s plodom. 13. Žleza. 14. Žleza v prerezu. 15. Plod. 16. Prerezano seme. 17. Štor se obreže na čepe pri s – s

Hmelj je večletna rastlina, katere življenjska doba niha med 10 in 30 let, kakršna je pač sorta in kakršni so pogoji okolja in obdelave. Naš golding se drži do 12 let, se pravi, da do te dobe še povrne vanj vloženi trud.

**Picture V-26:** Example of a corrected figure element with Alt-Text and caption present.

- **Check Ordered and Unordered Lists**

When marking lists back in Microsoft Word, the layout might seem correct, but sometimes the code needs to be fixed, especially when there are lists spread over two pages and a page break could disrupt the list order. Let's see two examples.

The first example is an ordered list which is marked with small characters (a, b, c), and each list item has an additional attribute for paragraph, which is actually not needed, so we cleared

all `<p>` and `</p>` tags. There is only a small difference in the layout. With paragraph tags, there was more space between each list item (**Picture V-27**), and without them, the list items are close together (**Picture V-28**).

```

27 <ol
28   type="a"
29   <li>
30     <p>rez na balin, ki ga izvajamo pri starih in krepkih rastlinah in pri
31     takšnih, ki so prišle že preveč navrh. Pri tem načinu rezi porežemo vse trte tesno ob
32     glavi (štoru), tik pod najnižjim kolobarjem očes:</p>
33   </li>
34   <li>
35     <p>običajna rez, priporočljiva pri šibkejših rastlinah, ki so še globoko v
36     zemlji. Pri tem načinu rezi odrežemo dve trti tik nad prvim venčkom očes, ostale pa tik
37     ob glavi:</p>
38   </li>
39   <li>
40     <p>visoka rez, ki se izvaja na rastlinah, ki so že pregloboko v zemlji in pri
41     prvič obrezanih. Dve trti odrežemo nad drugim ali celo tretjim kolobarjem očes, ostale
42     tik ob glavi. Tako tudi mladimo stare nasade.</p>
43   </li>
44 </ol>

```

a. rez na balin, ki ga izvajamo pri starih in krepkih rastlinah in pri takšnih, ki so prišle že preveč navrh. Pri tem načinu rezi porežemo vse trte tesno ob glavi (štoru), tik pod najnižjim kolobarjem očes;

b. običajna rez, priporočljiva pri šibkejših rastlinah, ki so še globoko v zemlji. Pri tem načinu rezi odrežemo dve trti tik nad prvim venčkom očes, ostale pa tik ob glavi;

c. visoka rez, ki se izvaja na rastlinah, ki so že pregloboko v zemlji in pri prvič obrezanih. Dve trti odrežemo nad drugim ali celo tretjim kolobarjem očes, ostale tik ob glavi. Tako tudi mladimo stare nasade.

Nož naj bo vedno prav oster. Odrezane trte čim prej pobereмо in odstranimo s hmeljnika in jih kompostiramo. Glavno pravilo pri hmeljski rezi

**Picture V-27:** Example of a numbered list where additional paragraph creates more space before and after each list element.

```

<ol
  type="a"
  <li>
    rez na balin, ki ga izvajamo pri starih in krepkih rastlinah in pri takšnih,
    ki so prišle že preveč navrh. Pri tem načinu rezi porežemo vse trte tesno ob glavi
    (štoru), tik pod najnižjim kolobarjem očes;
  </li>
  <li>
    običajna rez, priporočljiva pri šibkejših rastlinah, ki so še globoko v
    zemlji. Pri tem načinu rezi odrežemo dve trti tik nad prvim venčkom očes, ostale pa tik
    ob glavi;
  </li>
  <li>
    visoka rez, ki se izvaja na rastlinah, ki so že pregloboko v zemlji in pri
    prvič obrezanih. Dve trti odrežemo nad drugim ali celo tretjim kolobarjem očes, ostale
    tik ob glavi. Tako tudi mladimo stare nasade.
  </li>
</ol>

```

a. rez na balin, ki ga izvajamo pri starih in krepkih rastlinah in pri takšnih, ki so prišle že preveč navrh. Pri tem načinu rezi porežemo vse trte tesno ob glavi (štoru), tik pod najnižjim kolobarjem očes;

b. običajna rez, priporočljiva pri šibkejših rastlinah, ki so še globoko v zemlji. Pri tem načinu rezi odrežemo dve trti tik nad prvim venčkom očes, ostale pa tik ob glavi;

c. visoka rez, ki se izvaja na rastlinah, ki so že pregloboko v zemlji in pri prvič obrezanih. Dve trti odrežemo nad drugim ali celo tretjim kolobarjem očes, ostale tik ob glavi. Tako tudi mladimo stare nasade.

Nož naj bo vedno prav oster. Odrezane trte čim prej pobereмо in odstranimo s hmeljnika in jih kompostiramo. Glavno pravilo pri hmeljski rezi je: Rezi krepke rastline na kratko, slabe na dolgo!

**Picture V-28:** Example of a numbered list where additional paragraph was removed.

The second example shows a page break in the middle of the list item, which causes the unordered list `<ul>` to close before the page break and open again on the new page (**Picture V-29**). This should be fixed so that the page break is after the list or fixed so that there is no closing and opening of the unordered list before and after the page break.

```

26 <ul>
27   <li>
28     dušika 110 kg = 550 kg apnena dušika ali 700 kg čilskega solitra;
29   </li>
30   <li>
31     fosforne kisline 60 kg = 350 kg superfosfata;
32   </li>
33 </ul>
34 <span
35   epub:type="pagebreak"
36   role="doc-pagebreak"
37   id="page16"
38   aria-label="Stran 16." "></span>
39 <ul>
40   <li>
41     kalija 100 kg = 250 kg kalijeve soli in apna 190 kg.
42   </li>
43 </ul>

```

### U gnojenju na spiso

Po rezi nasade pognojimo. Hmelj je za gnoj prav zahtevna rastlina, ki odvzema zemlji velike množine hrane, če vzamemo, da se ona tretjina hrane v jeseni po obiranju vrača v podzemsko dele rastline, moramo dodati, če hočemo povsem nadomestiti hrano, ki jo je prejšnje leto pobrala rastlina iz zemlje, na vsak ha hmeljišča še:

- dušika 110 kg = 550 kg apnena dušika ali 700 kg čilskega solitra;
- fosforne kisline 60 kg = 350 kg superfosfata;
- kalija 100 kg = 250 kg kalijeve soli in apna 190 kg.

V primeri s pšenico porabi hmelj dvakrat več dušika in fosforne kisline, kalija trikrat več, apna pa celo petnajstkrat več. Če hočemo obdržati hmeljev

**Picture V-29:** Example of an unordered list with page break in the middle of the list. Arrows show the closing and opening tags which shouldn't be present as they disrupt the order of the list.

```

26 <ul>
27   <li>
28     dušika 110 kg = 550 kg apnena dušika ali 700 kg čilskega solitra;
29   </li>
30   <li>
31     fosforne kisline 60 kg = 350 kg superfosfata;
32   </li>
33   <li>
34     kalija 100 kg = 250 kg kalijeve soli in apna 190 kg.
35   </li>
36 </ul>
37 <span
38   epub:type="pagebreak"
39   role="doc-pagebreak"
40   id="page16"
41   aria-label="Stran 16." "></span>

```

ro rezi nasade pognojimo. Hmelj je za gnoj prav zahtevna rastlina, ki odvzema zemlji velike množine hrane, če vzamemo, da se ona tretjina hrane v jeseni po obiranju vrača v podzemsko dele rastline, moramo dodati, če hočemo povsem nadomestiti hrano, ki jo je prejšnje leto pobrala rastlina iz zemlje, na vsak ha hmeljišča še:

- dušika 110 kg = 550 kg apnena dušika ali 700 kg čilskega solitra;
- fosforne kisline 60 kg = 350 kg superfosfata;
- kalija 100 kg = 250 kg kalijeve soli in apna 190 kg.

V primeri s pšenico porabi hmelj dvakrat več dušika in fosforne kisline, kalija trikrat več, apna pa celo petnajstkrat več. Če hočemo obdržati hmeljev nasad na primerni višini, moramo gledati, da v čim zadostnejši meri povrnemo zemlji odvzete snovi. Kar je za ljudi kruh, to je za hmelj hlevski

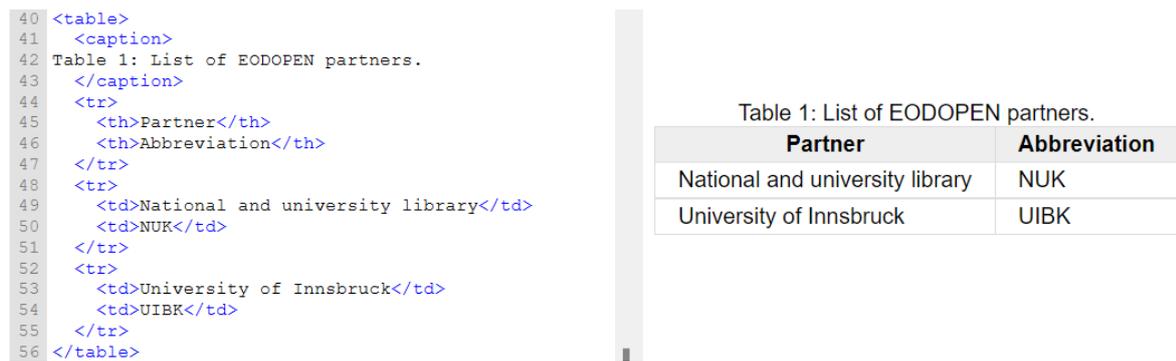
**Picture V-30:** Example of a fixed unordered list by moving the page break to the end of the list.

- **Check Table Structure**

It can happen that if you did not structure the table correctly in the source file, the table isn't correctly converted to EPUB. Check table data `<td>`, table rows `<tr>`, and table header `<th>` to ensure the elements are correct. The most common problem occurs when the table header is not defined. **Picture V-31** shows an example of a correctly structured table.

If the table has a caption, that element should be present as well inside the table element. The basic structure of a table should look like:

```
<table>
<caption>
text
</caption>
<tr>
<th>text</th>
<th>text</th>
</tr>
<tr>
<td>text</td>
<td>text</td>
</tr>
<tr>
<td>text</td>
<td>text</td>
</tr>
</table>
```



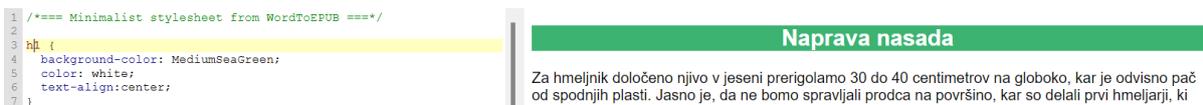
```
40 <table>
41 <caption>
42 Table 1: List of EODOPEN partners.
43 </caption>
44 <tr>
45 <th>Partner</th>
46 <th>Abbreviation</th>
47 </tr>
48 <tr>
49 <td>National and university library</td>
50 <td>NUK</td>
51 </tr>
52 <tr>
53 <td>University of Innsbruck</td>
54 <td>UIBK</td>
55 </tr>
56 </table>
```

Partner	Abbreviation
National and university library	NUK
University of Innsbruck	UIBK

**Picture V-31:** Example of correct table structure, containing caption, table header, table rows and table data.

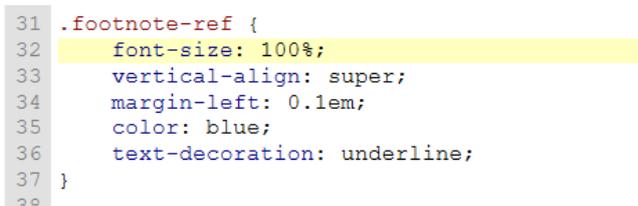
## 4.2 Editing Visual Appearance

If the CSS offered through WordToEpub is not enough and you wish to change the visual appearance of any elements in the whole book at once, go to the styles folder in the book browser and double-click on the stylesheet `a_default.css` or any other that you have chosen during the conversion. From here, you can edit the appearance to your liking (**Picture V-32**).



**Picture V-32:** Editing stylesheet. By adding background colour, text colour, and alignment, the heading titles appear differently.

We believe the current CSS only has one part that should be fixed for further uses. This is how big the footnote number appears. The `a_default.css` has a `font-size` value currently at 75%, and you can easily change it to 100% so it is more visible (**Picture V-33**). `Footnote-ref` appears twice in the stylesheet, so change it both times to 100%.



**Picture V-33:** Example of how to change the font size of footnote reference number.

To avoid this work for each separate book, create your own CSS stylesheet or edit the already existing one. If you do so, this part can be skipped the next time since you will use your own stylesheet during the conversion.

### 4.3 Editing Metadata

You can additionally edit metadata inside Sigil, especially if you get some additional information like the ISBN number (see [Metadata](#) chapter where we inserted 000). To access metadata, open the `content.opf` file and edit any black text to your preferences (**Picture V-34**).

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <package version="3.0" unique-identifier="epub-id-1" prefix="ibooks: http://vocabulary.itunes.apple.com/
3 rdf/ibooks/vocabulary-extensions-1.0/" xml:lang="sl-si" xmlns="http://www.idpf.org/2007/opf">
4 <metadata xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:opf="http://www.idpf.org/2007/opf">
5 <dc:identifier id="epub-id-1">urn:uuid:32a4b04f-06ee-47cc-90d8-c90cd33c9830</dc:identifier>
6 <dc:title id="epub-title-1">Hmelj, nasveti in navodila za pridelovanje</dc:title>
7 <meta property="title-type" refines="#epub-title-1">main</meta>
8 <dc:date id="epub-date">2024-04-24</dc:date>
9 <dc:language>sl-si</dc:language>
10 <dc:creator id="epub-creator-1">Janko Kač</dc:creator>
11 <meta name="cover" content="cover_jpg" />
12 <meta property="dcterms:modified">2024-04-24T06:35:23Z</meta>
13 <dc:identifier id="pub-id">urn:isbn:000</dc:identifier>
14 <dc:source id="src-id">Pretvorjeno iz Microsoft Worda.</dc:source>
15 <meta property="source-of" refines="#src-id">pagination</meta>
16 <dc:publisher>Narodna in univerzitetna knjižnica</dc:publisher>
17 <dc:description>Izvirnik iz leta 1951.</dc:description>
18 <dc:rights>CC/by 4.0</dc:rights>
19 <meta property="schema:accessibilitySummary">Publikacija vsebuje strukturirano navigacijo in nadomestna
20 besedila za slikovno gradivo.</meta>
21 <meta property="schema:accessMode">textual</meta>
22 <meta property="schema:accessMode">visual</meta>
23 <meta property="schema:accessModeSufficient">textual</meta>
24 <meta property="schema:accessibilityFeature">structuralNavigation</meta>
25 <meta property="schema:accessibilityFeature">displayTransformability</meta>
26 <meta property="schema:accessibilityFeature">readingOrder</meta>
27 <meta property="schema:accessibilityFeature">tableOfContents</meta>
28 <meta property="schema:accessibilityFeature">unlocked</meta>
29 <meta property="schema:accessibilityFeature">printPageNumbers</meta>
30 <meta property="schema:accessibilityFeature">alternativeText</meta>
31 <meta property="schema:accessibilityHazard">none</meta>
32 </metadata>
33 </package>

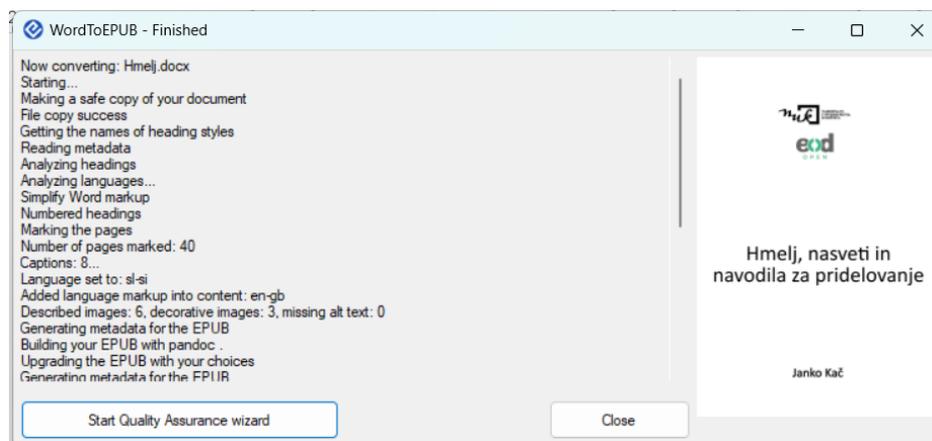
```

**Picture V-34:** A part of the code which contains metadata information. It shows the part where ISBN information can be additionally changed or added.

#### 4.4 Saving the File and Continuing the Workflow in WordToEpub

Once you apply any changes inside Sigil, whether it is regarding content, visual appearance, or metadata, go to **File > Save**. With this, any changes are saved.

Now go back to the WordToEpub tool, click **Next** to continue to the next steps, which are about checking accessibility. Click **Start Quality Assurance wizard** and continue learning about this step in the next chapter.



**Picture V-35:** WordToEpub interface to start the Quality Assurance wizard.

## 5 Accessibility Check

This section covers the automatic checking of the accessibility of your EPUB document and additional manual checking for any specifics.

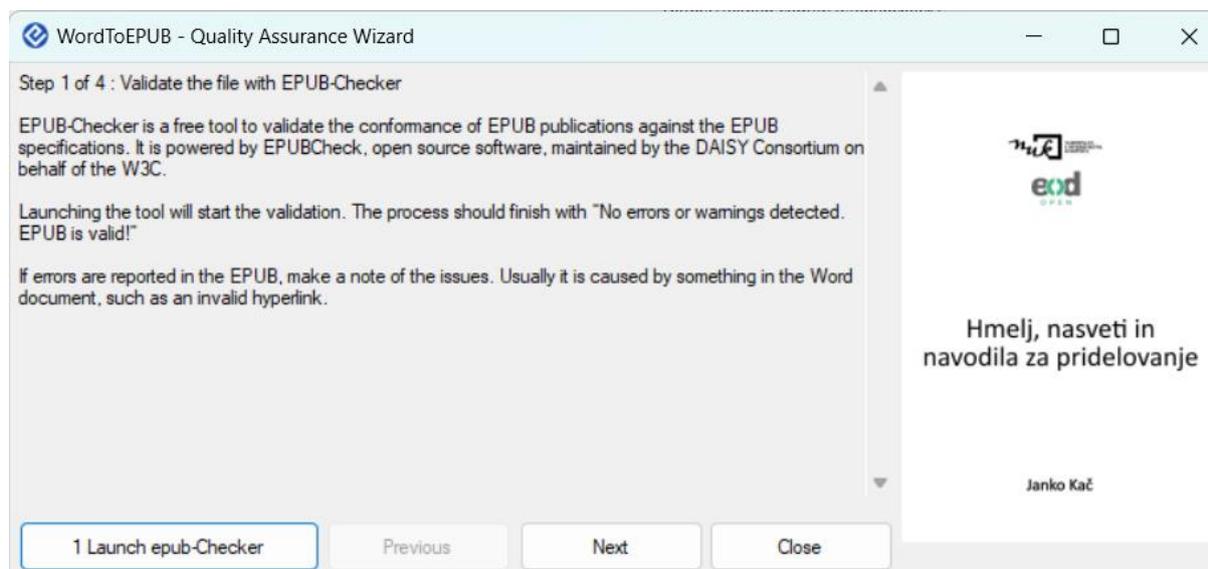
Automatic checks can detect most problems in the code, but some elements still need manual checks or human input. For example, it checks if Alt-Text is present but not the quality or correctness of it.

The Quality Assurance Wizard offers the following three steps:

- EpubChecker
- Ace by DAISY
- Thorium reader

After clicking the first option in WordToEpub **Launch EpubChecker**, additional software opens and immediately launches the checking.

If you accidentally closed the WordToEpub tool, open the EPUB file in the software corresponding to the following chapters.

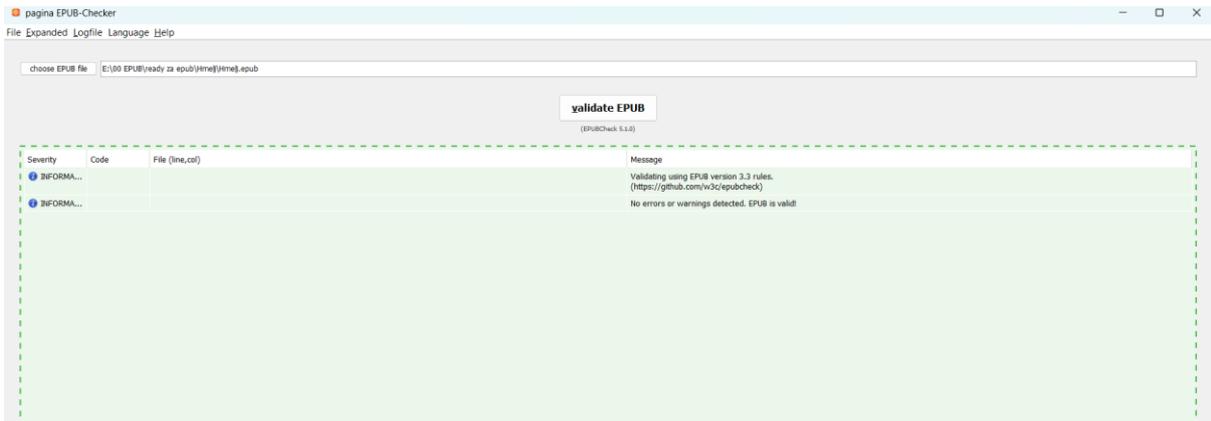


**Picture V-36:** Interface of WordToEpub before starting EpubChecker.

### 5.1 EpubChecker

EpubChecker is an open-source conformance checker tool to evaluate EPUB files. It is owned by W3C and can detect problems with the structure, package description file, associated metadata, content markup, and the consistency of internal references. It reports errors in criteria: fatal error, error, warning, info, and usage.

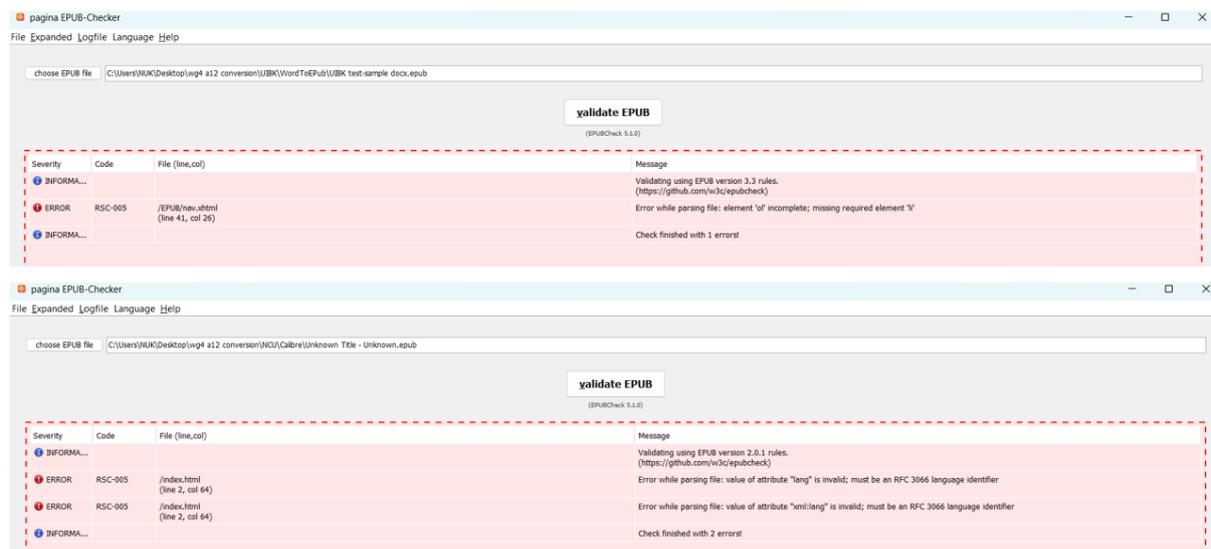
If there are no errors detected, the screen becomes green as in **Picture V-37**.



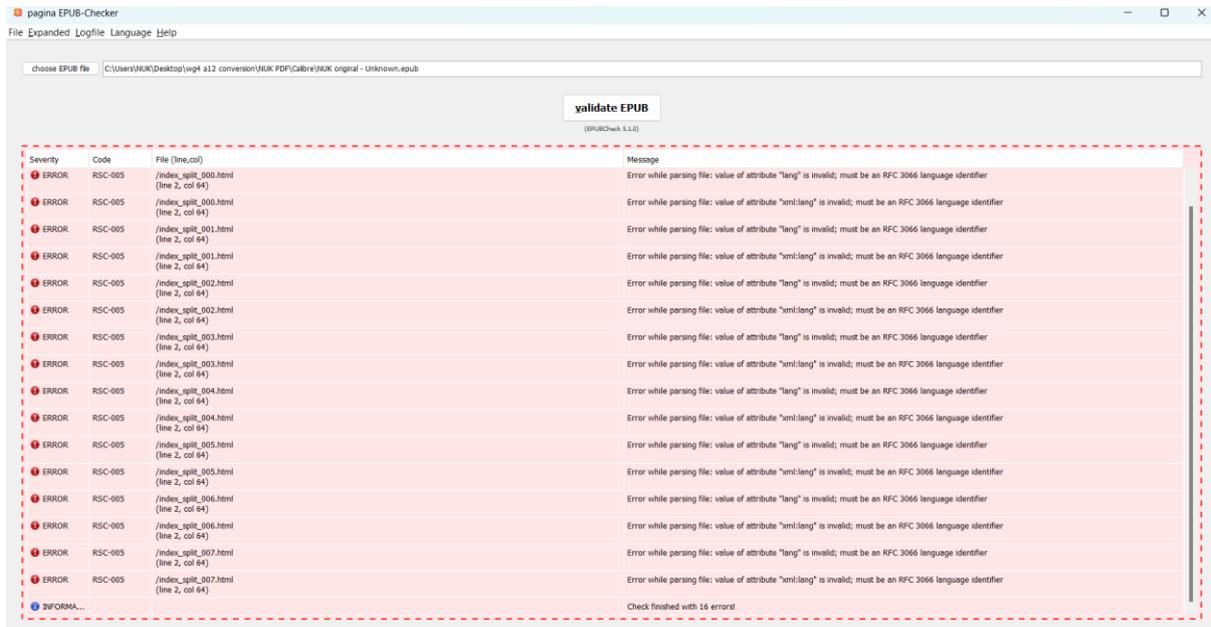
**Picture V-37:** Interface of EpubChecker with no errors detected.

In case errors appear, the software provides information about what is wrong and where to fix it. At this point, you should return to Sigil, correct what is wrong, and start the evaluation with this software again to check if the problem has been resolved.

Following are some examples that resulted in errors. All examples are results that came out from EODOPEN WG4 testing of conversion services. Some resulted in more errors than others.

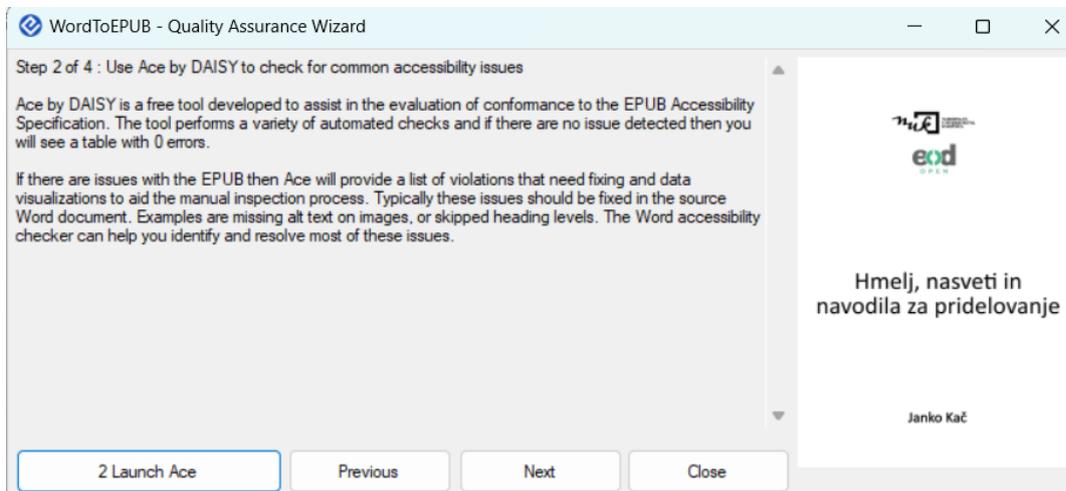


**Picture V-38:** Two examples with only a few errors detected by EpubChecker.



**Picture V-39:** Example with many errors detected by EpubChecker.

After the tool results in no errors, return to WordToEpub and click **Next** to go to the second phase of the Quality Assurance wizard. When clicking the option **Launch Ace**, additional software opens.



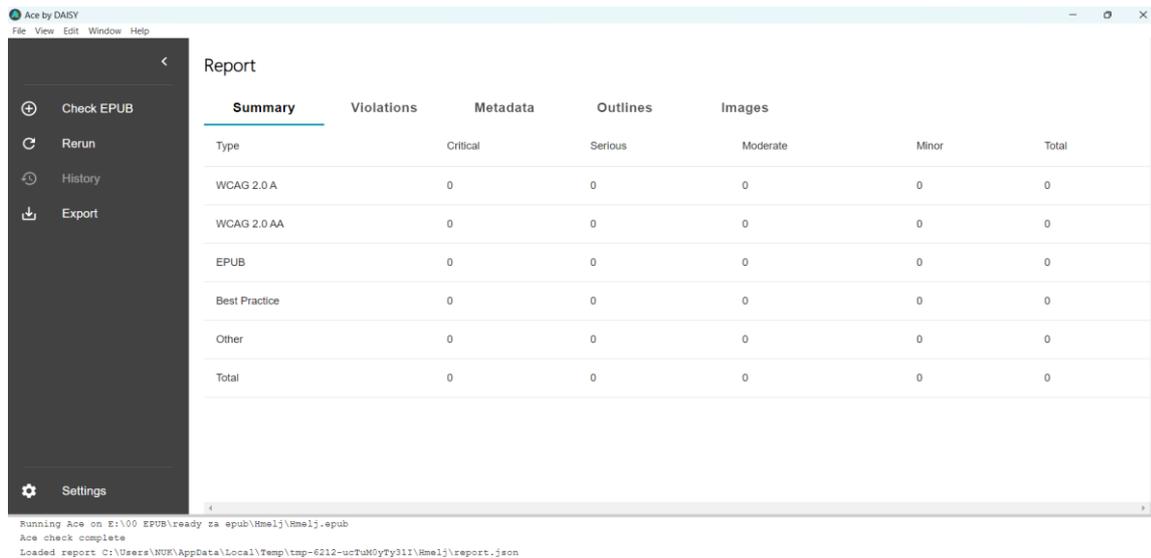
**Picture V-40:** Interface of WordToEpub before starting Ace by DAISY.

## 5.2 Ace by DAISY

Ace by DAISY is an open-source tool designed to check the accessibility of EPUB files and checks conformance to the EPUB Accessibility Specification. It is developed by The DAISY Consortium. It provides reports of errors in criteria: critical, serious, moderate, minor, and also provides a total number of errors. It checks conformance according to WCAG 2.0 on level A and level AA, EPUB specifications, and best practice.

This software currently does not automatically start, so you should open the converted EPUB file to launch the checking procedure. Click **Check EPUB** to browse for the file or just drag and drop the file in the software.

If there are no errors detected, the total number of reported errors should be zero and should look like **Picture V-41**.

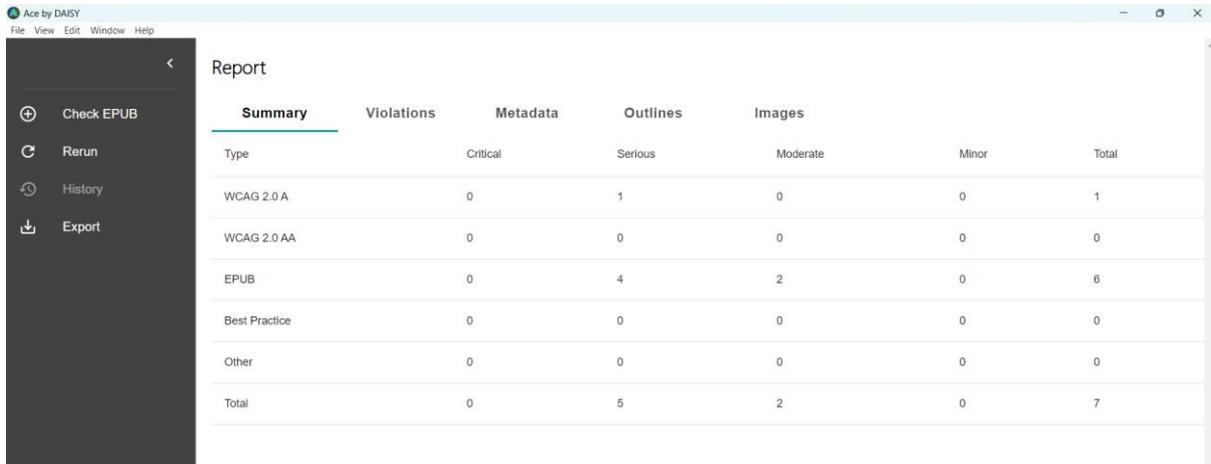


**Picture V-41:** Interface of Ace by DAISY when no errors are detected.

In case errors appear, the software provides information about what is wrong under the tab **violations** and where to fix it. At this point, you should also return to Sigil, correct what is wrong, and start the evaluation with both software again to check if the problem has been resolved.

The software also enables you to check the tab **metadata**, check **outlines** like heading structure, and **images** which appear in the document along with their Alt-Text provided.

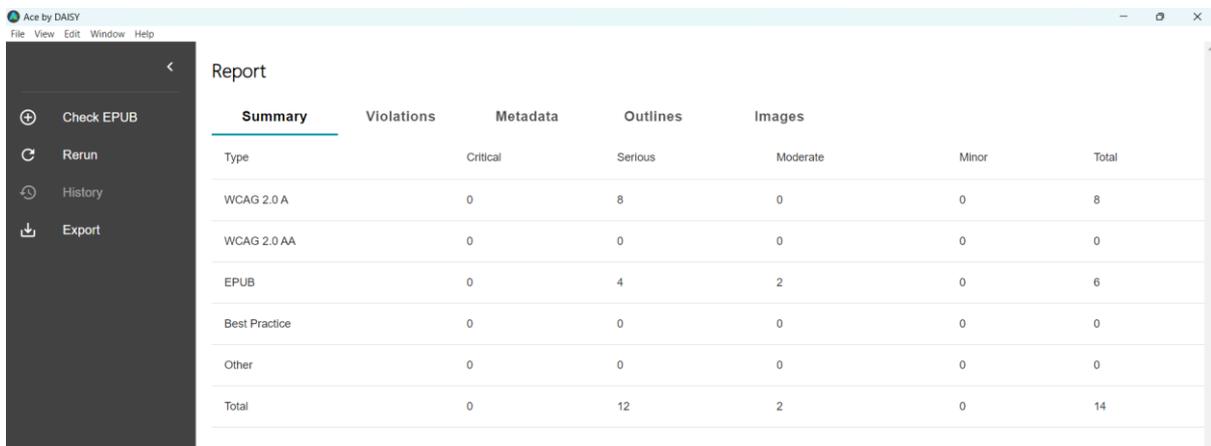
Following are some examples from the previous chapter which resulted in errors with EpubChecker and had issues pointed out also in Ace by DAISY. All examples are results that came out from EODOPEN WG4 testing of conversion services. Some resulted in more errors than others.



The screenshot shows the 'Report' window in Ace by DAISY. The 'Summary' tab is selected, displaying a table of violations. The total number of errors is 7.

Type	Critical	Serious	Moderate	Minor	Total
WCAG 2.0 A	0	1	0	0	1
WCAG 2.0 AA	0	0	0	0	0
EPUB	0	4	2	0	6
Best Practice	0	0	0	0	0
Other	0	0	0	0	0
<b>Total</b>	<b>0</b>	<b>5</b>	<b>2</b>	<b>0</b>	<b>7</b>

**Picture V-42:** Example of evaluated EPUB book by Ace by DAISY and resulting in 7 total errors.



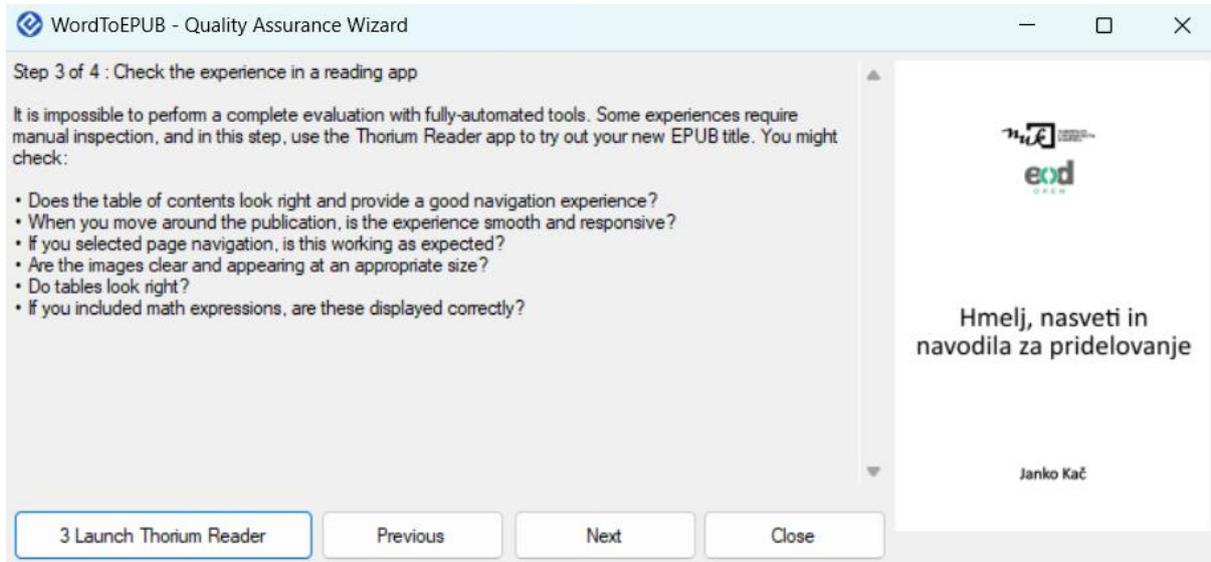
The screenshot shows the 'Report' window in Ace by DAISY. The 'Summary' tab is selected, displaying a table of violations. The total number of errors is 14.

Type	Critical	Serious	Moderate	Minor	Total
WCAG 2.0 A	0	8	0	0	8
WCAG 2.0 AA	0	0	0	0	0
EPUB	0	4	2	0	6
Best Practice	0	0	0	0	0
Other	0	0	0	0	0
<b>Total</b>	<b>0</b>	<b>12</b>	<b>2</b>	<b>0</b>	<b>14</b>

**Picture V-43:** Example of evaluated EPUB book by Ace by DAISY and resulting in 14 total errors.

After the tool results in no errors, return to WordToEpub and click **Next** to go to the third phase of the Quality Assurance wizard. The tool also provides you with information what to check in the next, manual step.

After clicking the option **Launch Thorium Reader** in WordToEpub, additional software opens.



**Picture V-44:** Interface of WordToEpub before starting Thorium Reader.

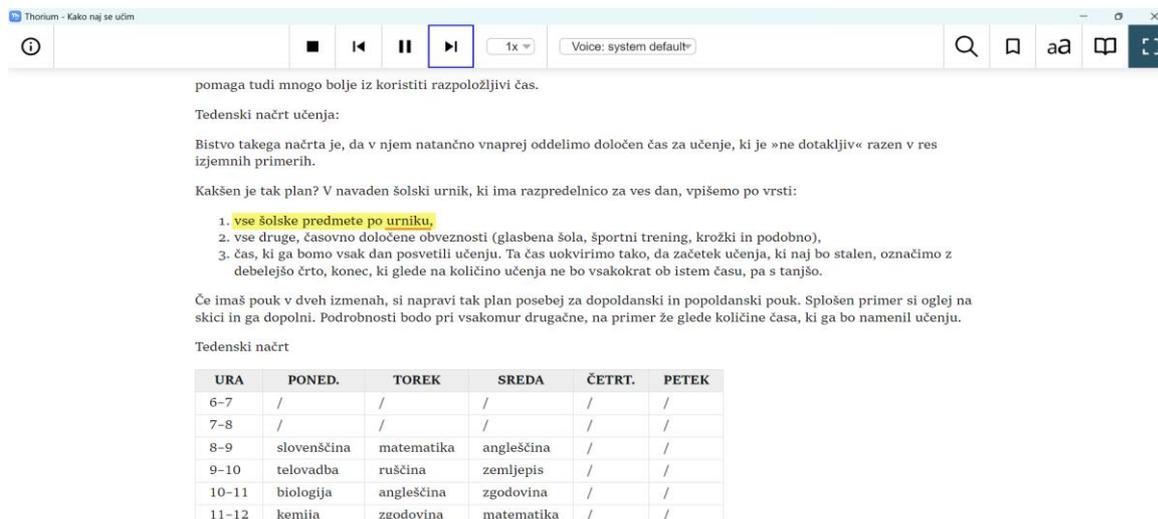
### 5.3 Thorium Reader and Read Aloud Test

Thorium Reader is a free EPUB reader, developed by EDRLab. It supports not only EPUB files but also PDF, audiobooks, DAISY 2.02, and DAISY 3 books. The software is developed with the goal of making it accessible primarily to people with visual impairments and dyslexia. The software includes a read aloud option, which is useful for manual testing of the book.

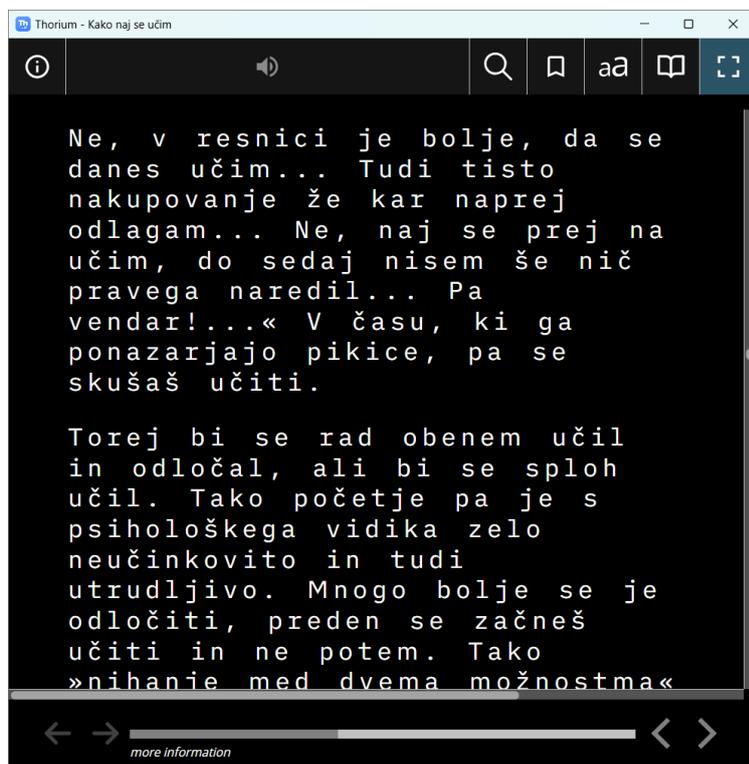
Additionally, you can test the visual look of the book, the structure of the book such as lists and tables, check the landmarks, table of contents, and navigation to the original pages. When changing the visual appearance of the book, you can ensure it is presented to the users correctly and that no content is missing.

The software is also constantly being developed, so some features may look different over time.

If you know or use similar applications or software, try to perform the test there as well.



**Picture V-45:** Example of using read aloud option in Thorium Reader. The active lines are highlighted.

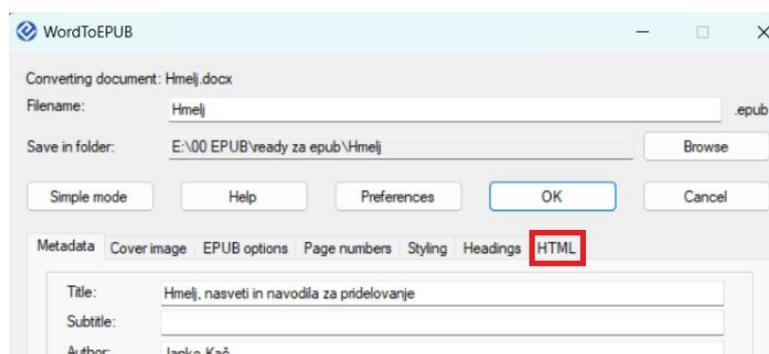


**Picture V-46:** Example of visual adjustments (changed fonts, size, spacing, dark theme etc.) on a narrow window in Thorium Reader.

You can also check the final publication in other software that provides eBook reading. It is especially important to check on a mobile device to see how the content is displayed on smaller screens.

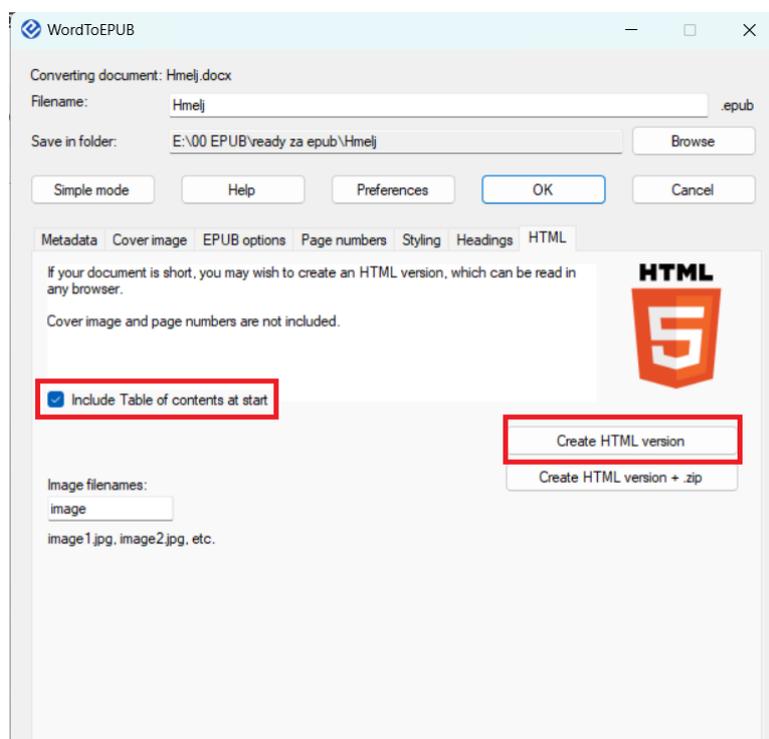
## 6 Converting the Book to HTML

If the option for conversion to HTML is chosen in the advanced preferences from chapter: [2 WordToEpub Preferences Before Converting the First Book](#), under the tab **User Interface Options**, the option to convert to HTML appears when the conversion of the book is started.



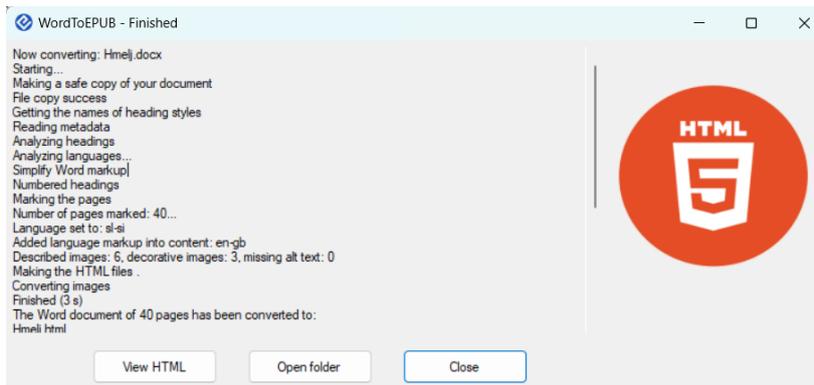
**Picture V-47:** Starting interface of WordToEpub where the tab HTML is available.

When clicking the tab **HTML**, the setting for conversion will open, and the process is much shorter than for conversion to EPUB. Enable the checkbox **include Table of contents at start** and click **Create HTML version**.



**Picture V-48:** Interface of WordToEpub with settings for conversion to HTML.

As with converting to EPUB, the process starts. In the window appearing after the conversion, we are offered options to **view HTML** in any browser, **open the folder** where the HTML was saved, or **close** the tool. On the left side, the full conversion report is available.



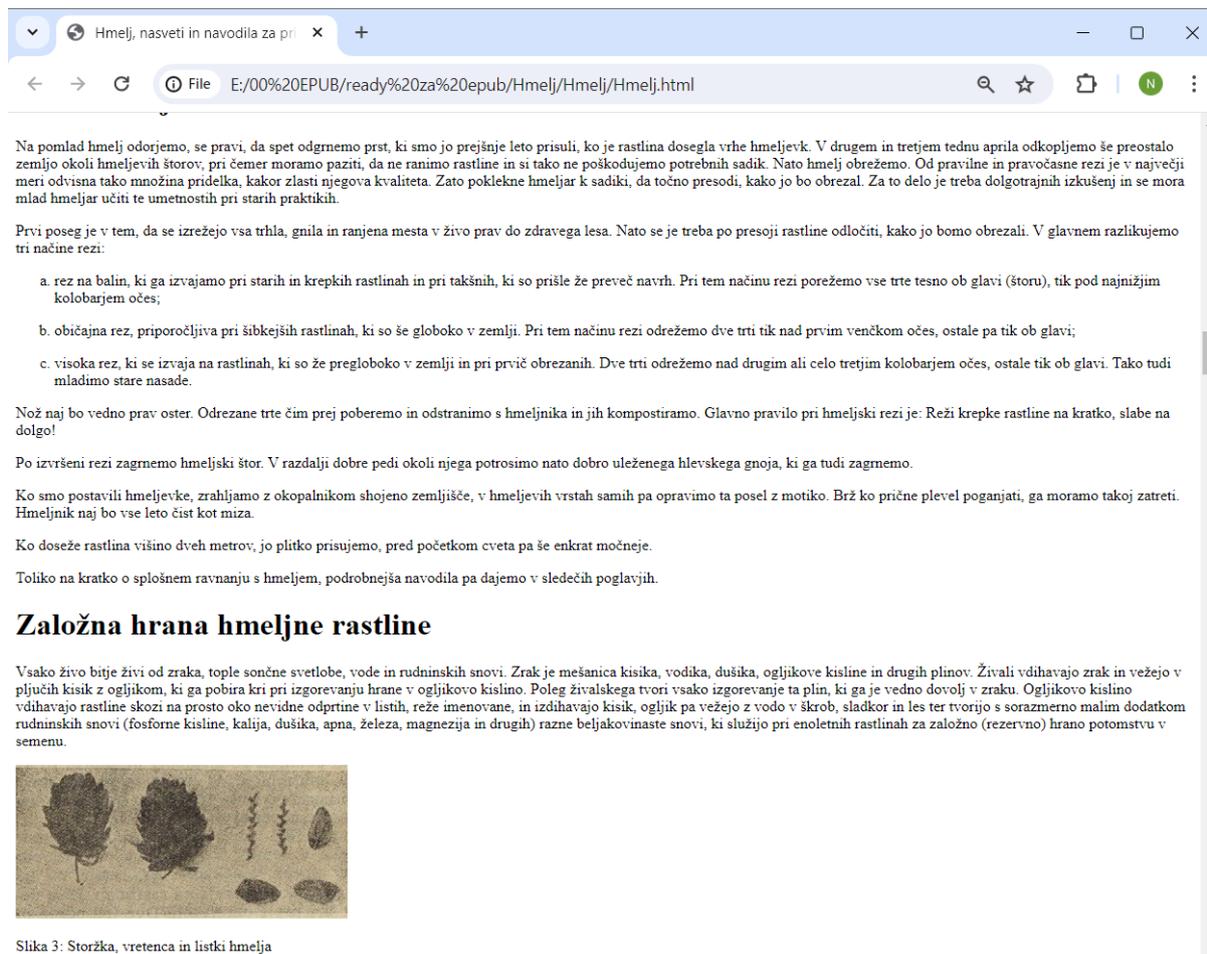
**Picture V-49:** Interface of WordToEpub after finishing the conversion to HTML.

An example of the full report is added below, with parts that should always be checked marked.

```

Now converting: Hmelj.docx
Starting...
Making a safe copy of your document
File copy success
Getting the names of heading styles
Reading metadata
Analyzing headings
Analyzing languages...
Simplify Word markup
Numbered headings
Marking the pages
Number of pages marked: 40...
Language set to: sl-si
Added language markup into content: en-gb
Described images: 6, decorative images: 3, missing alt text: 0
Making the HTML files .
Converting images
Finished (3 s)
The Word document of 40 pages has been converted to:
Hmelj.html
in the folder:
E:\00 EPUB\ready za epub\Hmelj\Hmelj
Your publication in numbers:
Number of words: 12852
Number of paragraphs: 259
Number of images: 9
Number of described images: 6
Number of decorative images: 3
Number of languages: 2
Filesize (bytes): 96.074
No warnings :)
    
```

If you click the option to view HTML, the book will open in your default browser.



Hmelj, nasveti in navodila za pri x +

E:/00%20EPUB/ready%20za%20epub/Hmelj/Hmelj/Hmelj.html

Na pomlad hmelj odrtorjemo, se pravi, da spet odgrnemo prst, ki smo jo prejšnje leto prisuli, ko je rastlina dosegla vrhe hmeljev. V drugem in tretjem tednu aprila odkoplremo še preostalo zemljo okoli hmeljevih štorov, pri čemer moramo paziti, da ne ranimo rastline in si tako ne poškodujemo potrebnih sadik. Nato hmelj obrežemo. Od pravilne in pravočasne rezi je v največji meri odvisna tako množina pridelka, kakor zlasti njegova kvaliteta. Zato pokleknite hmeljar k sadiki, da točno presodi, kako jo bo obrezal. Za to delo je treba dolgotrajnih izkušenj in se mora mladi hmeljar učiti te umetnosti pri starih praktikih.

Prvi poseg je v tem, da se izrežejo vsa trhla, gnila in ranjena mesta v živo prav do zdravega lesa. Nato se je treba po presoji rastline odločiti, kako jo bomo obrezali. V glavnem razlikujemo tri načine rezi:

- rez na balin, ki ga izvajamo pri starih in krepkih rastlinah in pri takšnih, ki so prišle že preveč navrh. Pri tem načinu rezi porežemo vse trte tesno ob glavi (štoru), tik pod najnižjim kolobarjem očes;
- običajna rez, priporočljiva pri šibkejših rastlinah, ki so še globoko v zemlji. Pri tem načinu rezi odrežemo dve trti tik nad prvim venčkom očes, ostale pa tik ob glavi;
- visoka rez, ki se izvaja na rastlinah, ki so že pregloboko v zemlji in pri prvič obrezanih. Dve trti odrežemo nad drugim ali celo tretjim kolobarjem očes, ostale tik ob glavi. Tako tudi mladimo stare nasade.

Nož naj bo vedno prav oster. Odrezane trte čim prej pobereemo in odstranimo s hmeljnika in jih kompostiramo. Glavno pravilo pri hmeljski rezi je: Reži krepke rastline na kratko, slabe na dolgo!

Po izvršeni rezi zagnemo hmeljski štor. V razdalji dobre pedi okoli njega potrosimo nato dobro uležena hlevskega gnoja, ki ga tudi zagnemo.

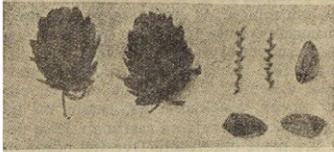
Ko smo postavili hmeljeve, zrahljamo z okopalnikom shojeno zemljišče, v hmeljevih vrstah samih pa opravimo ta posel z motiko. Brž ko prične plevel poganjati, ga moramo takoj zatreti. Hmeljnik naj bo vse leto čist kot miza.

Ko doseže rastlina višino dveh metrov, jo plitko prisujemo, pred početkom cveta pa še enkrat močneje.

Toliko na kratko o splošnem ravnanju s hmeljem, podrobnejša navodila pa dajemo v sledečih poglavjih.

### Založna hrana hmeljne rastline

Vsako živo bitje živi od zraka, tople sončne svetlobe, vode in rudninskih snovi. Zrak je mešanica kisika, vodika, dušika, ogljikove kisline in drugih plinov. Živali vdihavajo zrak in vežejo v pljučih kisik z ogljikom, ki ga pobira kri pri izogrevanju hrane v ogljikovo kislino. Poleg živalskega tvori vsako izogrevanje ta plin, ki ga je vedno dovolj v zraku. Ogljikovo kislino vdihavajo rastline skozi na prosto oko nevidne odprtine v listih, reže imenovane, in izdihavajo kisik, ogljik pa vežejo z vodo v škrob, sladkor in les ter tvorijo s sorazmerno malim dodatkom rudninskih snovi (fosforne kisline, kalija, dušika, apna, železa, magnezija in drugih) razne beljakovinske snovi, ki služijo pri enoletnih rastlinah za založno (rezervno) hrano potomstvu v semenu.



Slika 3: Storzka, vretenca in listki hmelja

**Picture V-50:** Example of a book, viewed in Chrome browser.

## 7 Further resources

- DAISY Consortium. (21. 4. 2020). *Create EPUB publications from Word with a simple tool anyone can use* [Video]. YouTube. <https://www.youtube.com/watch?v=4M3liZDbKjo>
- DAISY Consortium. (2. 8. 2020). *WordToEPUB Extended Tutorial - Accessible EPUB in Seconds* [Video]. YouTube. <https://www.youtube.com/watch?v=7skf3NkHoJU>
- DAISY Consortium. (15. 12. 2020). *Do more with WordToEPUB* [Video]. YouTube. <https://www.youtube.com/watch?v=B231FE0DQ2k>
- DAISY Consortium. (11. 10. 2021). *EPUB Accessibility 101* [Video]. YouTube. [https://www.youtube.com/watch?v=ibfD\\_vArrpl](https://www.youtube.com/watch?v=ibfD_vArrpl)
- DAISY Consortium. (9. 11. 2021). *Validating and Conformance Checking EPUB* [Video]. YouTube. <https://www.youtube.com/watch?v=RoNbftMtEQ>
- DAISY Consortium. (22. 11. 2021). *Creating and Editing Accessible EPUB* [Video]. YouTube. <https://www.youtube.com/watch?v=mTd6YIP52oY>
- DAISY Consortium. (2024). *DAISY Learning*. <https://learning.daisy.org/>
- W3Schools. (2024). *HTML elements*. [https://www.w3schools.com/html/html\\_elements.asp](https://www.w3schools.com/html/html_elements.asp)

## VI. SUMMARY

This handbook compiles the experiences gathered during the implementation of the EODOPEN project from 2019 to 2024. It includes two key deliverables: D13a *Workshops and Training for Librarians and Interested Communities on a Nationwide Level*, and D13b *Educational Material Used in Workshops and Trainings*. The aim of this handbook is to equip librarians and other cultural workers with the skills, competence, and knowledge to provide services for print-disabled users, create alternative delivery formats, and understand copyright issues and the Marrakesh Treaty at a national level. The target communities for the educational material are users of mobile devices and individuals who are blind or partially sighted.

The handbook is divided into five modules. In **Module I: Accessibility and Reading Digital**, we explain the meaning of accessibility in the digital world, its importance, and how to improve it. Special emphasis is given to users of mobile devices and persons with print disabilities. Different types of print disabilities require different approaches, including the use of assistive technologies. Furthermore, relevant accessibility standards are described.

In **Module II: Copyright**, we present the legal context for providing access to copyrighted materials to people with visual impairments. For libraries, it is very important to be aware of the existing copyright regulations, as well as the limitations and exceptions concerning blind and partially sighted users.

In **Module III: Digitisation and Formats**, the topics of digitisation, output delivery file formats, and their conversion are presented. The module is additionally based on the results from the EODOPEN project research conducted by the National and University Library (Slovenia) and contains practical examples.

**Module IV: Creating Accessible Document Formats** focuses on how librarians can prepare accessible Microsoft Word documents from digitised publications. It covers 14 elements based on accessibility guidelines and good practices used in the publishing industry. The elements are divided into two sections: basic accessibility and advanced accessibility. Both sections contain processes needed to enable easier access for users of mobile devices and better navigation, structure, and understanding of the content for blind or partially sighted users. The level of accessibility provided varies based on staff, knowledge, and financial resources available. The course is structured to cater to different levels of knowledge in accessibility, ensuring that at least basic accessibility can give users better access to digitised content.

**Module V: Creating Accessible EPUB and HTML Formats** provides librarians with instructions on how to prepare accessible EPUB and HTML publications from the accessible Word documents created in Module IV. This module is divided into seven sections, focusing on preparation before conversion, what to pay attention to during conversion, and what is

needed after conversion. Everything is based on accessibility guidelines and good practices used in the publishing industry. All processes in this module are designed to ensure that users of mobile devices have easier access and that blind or partially sighted users have better navigation, structure, and understanding of the content in either EPUB or HTML format. The course is structured so that you can choose the workflow for one of the presented formats or both. EPUB and HTML formats were chosen because they are both based on markup, are supported by various reading software or browsers, allow users to adapt the visual appearance of the book, and can be used on various devices due to reflowing text, independent of screen size.

Each module explains the content, goals, and training scenarios for delivering the content, and the training material needed. They are supported with practical examples, specially produced video recordings and/or slides, and an evaluation test.

## **VII. ANNEX 1: ASSESSMENT QUESTIONNAIRES BY MODULES**

The trainers and/or trainees could use the questionnaires as an assessment tool of the content learned in each module.

### **Module 1: Accessibility and Reading Digital**

1. What does 'accessibility' mean in the context of publications in the digital environment?
2. What is a print disability?
3. How much of all published information is accessible to people with print disabilities?
4. Why is accessibility important for people with print disabilities?
5. How do mobile devices influence accessibility?
6. What types of assistive technology do you know?
7. Are you aware of any accessibility standards?
8. What does the acronym WCAG stand for?
9. What are the key accessibility principles?

### **Module 2: Copyright**

1. What were the main shortcomings of the Copyright Directive (2001/29/EC)?
2. How has the Digital Libraries Initiative i2010 contributed to digitisation in libraries?
3. Why is Europeana significant for European Union member states?
4. What is the meaning of 'orphan work'?
5. How do the exceptions and limitations to copyright legislation influence accessibility?
6. Did the Digital Single Market Directive address the problem of digitisation of out-of-commerce works properly?
7. Could you list at least three open issues relevant to libraries that should be addressed by European copyright regulation?
8. What was the impact of the Marrakesh Treaty on people with disabilities?
9. Could you list the three exceptions addressed by the Marrakesh Treaty?

### **Module 3: Digitisation and Formats**

1. Could you list the essential phases of a digitisation project?
2. What does image processing comprise?
3. Could you list at least three imperfections in the original book that could affect the quality of digitisation?

4. What is the meaning of OCR?
5. What elements in the book significantly impact OCR results?
6. Which access file formats do you know?
7. Can conversion services improve access to digitised contents? How?
8. What conversion services do you know?

## **Module 4: Creating Accessible Documents**

1. When doing OCR clean-up after digitisation, which parts are most important to remove from the document? Which tool in Word is useful for that?
2. Which fonts are most suitable for the text and what should be avoided?
3. When applying colours inside the document, what do you need to pay attention to?
4. What do we use to design and format the document (chapters, main text, etc.)?
5. When applying heading styles, what do we need to pay attention to in order to enable smooth navigation for the user?
6. Why is it important to specify the language of the document and its sections?
7. What kinds of lists are available in the document and how do we apply them?
8. How are images inserted in the document?
9. What two things do we add to the images to provide better accessibility to the reader?
10. What information about the document should we provide before saving it?
11. When adapting a digitised publication, how and where can we add the original pages?
12. When adding tables to the document, what should we avoid?
13. When adding tables to the document, what needs to be added or applied?
14. When adding hyperlinks to the document, what should we pay attention to?
15. Besides adding hyperlinks to websites, what other types of hyperlinks can we add to the document?
16. During digitisation, maths can become inaccessible. How can we fix that and provide more accessible maths?
17. What can we use instead of text boxes?
18. What kind of wrapping is most suitable for accessibility?
19. What is the purpose of templates?
20. How can you check the accessibility of your documents?

## **Module 5: Creating Accessible EPUB and HTML formats**

1. What is needed before using the WordToEpub tool to convert DOCX format into an EPUB?
2. Where can you add or edit the metadata of the final book before and after the conversion?
3. If you want to use the same cover for all your publications, what should you prepare?

4. How are the settings in the WordToEpub tool connected to the marked pages inside the document when defining page numbers?
5. Which seven tasks should we go through after the conversion to edit the content?
6. What do we edit if we want to change the visual appearance of the EPUB?
7. Which three phases of accessibility check are needed after the conversion?
8. After converting to HTML format, where can you check or view the book?