

# **Guidelines for Producing Accessible eText**

**May 2018**



**Round Table on Information Access for  
People with Print Disabilities Inc.**

## **Guidelines for producing accessible eText**

Copyright © 2018 Round Table on Information Access for People with Print Disabilities

This work is licenced under the Creative Commons:

Attribution-Non-commercial-NoDerivs 3.0 Unported License.

To view a copy of this licence, visit [Creative Commons:](http://creativecommons.org/licenses/by-nc-nd/3.0/)

<http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Published by Round Table on Information Access for People with Print Disabilities Inc.  
13 Myee Crescent, Baulkham Hills, NSW 2153, Australia

Email: [admin@printdisability.org](mailto:admin@printdisability.org)

Web address: [Round Table \(Information Access for People with Print Disabilities\):](http://printdisability.org/)  
<http://printdisability.org/>

## **National Library of Australia Cataloguing-in-Publication entry**

Author: Round Table on Information Access for People with Print Disabilities Inc.

Title: Guidelines for producing accessible eText

ISBN: 978-1-925437-99-7

Subjects: People with disabilities--Services for.

Electronic publications--Handbooks, manuals, etc.

Dewey Number: 362.4

---

## **About these guidelines**

The Guidelines for producing Accessible eText are published by Round Table on Information Access for People with Print Disabilities Inc. (Round Table). Round Table is an umbrella organisation which brings together producers, distributors and consumers of information in alternative formats to print, blindness agencies, tertiary institutions and government departments in Australia and New Zealand.

These guidelines are available from Round Table in accessible formats.

## **Acknowledgements**

The original guidelines were compiled by the eText working group of the Round Table in 2009.

A comprehensive revision of the Guidelines were completed in May 2018 by a working group that comprised experts in the field of accessibility from Round Table member organisations within Australia and New Zealand. Working group comprised the following:

Christopher Goj, Blind Foundation, New Zealand

Dinesh Burah, VisAbility, Perth WA (Chairperson)

Jessica King, Vision Australia

Michelle Knight, Royal Institute for Deaf and Blind Children

Sonali Marathe, Royal Institute for Deaf and Blind Children

Vithya Vijayakumare, VisAbility, Perth, WA.

A special thank you to Sonali Marathe and Michelle Knight for their dedication and Brian Conway for proof reading the final version.

## **Disclaimer**

This document has been prepared from the collective knowledge and experience of the working group and sources believed to be accurate and complete at the time of compilation. While every effort has been made to ensure accuracy, the accessibility of the websites and the references, Round Table does not accept any responsibility for omissions or inaccuracies.

## Contents Page

|  |    |
|--|----|
| Guidelines for Producing Accessible eText.....   | 1  |
| About these guidelines.....                      | 2  |
| Acknowledgements .....                           | 2  |
| Disclaimer.....                                  | 2  |
| Contents Page.....                               | 3  |
| Introduction to these guidelines .....           | 5  |
| What is accessible eText?.....                   | 5  |
| Guidelines for accessible eText.....             | 6  |
| Background .....                                 | 6  |
| Who are these guidelines for? .....              | 7  |
| General principles.....                          | 8  |
| 1. Equivalent to print.....                      | 8  |
| 2. Easy to navigate.....                         | 8  |
| 3. Clear visual style .....                      | 8  |
| 4. Standards, guidelines and best practice.....  | 9  |
| General principles explained.....                | 10 |
| 1. Equivalent to print.....                      | 10 |
| 2. Easy to navigate.....                         | 13 |
| 3. Clear visual style .....                      | 14 |
| 4. Standards, Guidelines and Best Practice ..... | 16 |
| Common formats.....                              | 17 |
| Glossary.....                                    | 24 |

|  |    |
|--|----|
| Appendix 1: Accessibility Checklist .....                | 28 |
| 1. Equivalent to print .....                             | 28 |
| 2. Easy to Navigate .....                                | 28 |
| 3. Clear Visual Style .....                              | 29 |
| 4. Follow standards, guidelines and best practices ..... | 29 |
| Appendix 2: Guidelines for Formatting Plain Text.....    | 30 |
| Appendix 3: Tips for using Scanners and OCR.....         | 32 |
| Appendix 4: Standards, Guidelines and References.....    | 34 |

## Introduction to these guidelines

### What is accessible eText?

eText is structured electronic text which is accessible to people with a print disability, that is, to people who can't access information from regular print.

eText may be an alternative version of a print document that provides a person with a print disability with equivalent access to the information. The term eText can also refer to a document that is originally created in digital format and is accessible to all readers.

Typical formats for accessible eText include:

- **Plain text:** This may include textual "markup" to indicate structure
- **Word processor documents:** Common formats are Microsoft Word, RTF or OpenDocument
- **HTML and HTML5:** Hyper Text Markup Language, used widely by website developers
- **DAISY books:** Can contain the full structured electronic text of a book, and may contain audio
- **PDF:** It is possible to create accessible PDF files in Office, as long as accessibility features are built in
- **EPUB 3:** Can contain the full structured electronic text of a book, and may contain audio.

People with print disabilities use a range of assistive technologies to access eText.

Typical reading methods include:

- **Synthetic speech:** using screen reader software (e.g. JAWS, VoiceOver, NVDA, Window Eyes, ChromeVox, Talkback), built-in screen readers in mainstream software (e.g. Read Aloud in Adobe Acrobat) or within portable devices (e.g. Victor Reader Stream).
- **Refreshable Braille:** using a portable braille display connected to a desktop computer, laptop, tablet device or within a stand-alone braille notetaker.

- **Viewing on screen:** using specialised screen magnification software, tools that customise colour combinations, tool that aid in tracking when reading and built-in zoom functions within software and operating systems. These can be on mainstream technology or specialised devices such as desktop/handheld video magnifiers.

## **Guidelines for accessible eText**

The Guidelines for Accessible eText, referred to as the Guidelines, have been produced to provide document creators with an understanding of accessibility principles, and some best practice accessibility methods across a variety of commonly used electronic formats.

People are significantly disadvantaged if they cannot access information. Electronic formats provide the opportunity for equity of information access, but the reality is that many electronic documents are not designed to be fully accessible by readers with a print disability. Electronic documents need to be created in accordance with a standards-based approach to ensure accessibility.

Round Table's goal has been to make the Guidelines clear and easy to read in the hope they will be widely adopted.

## **Background**

Guidelines for Accessible eText were created to reflect the many formats and changing processes in creating these documents.

Version 1.0 of **Guidelines for preparation of text materials on computer disk for people with print disabilities** was produced by a sub-committee of Round Table in July 1995.

In 2006 Round Table Executives set up a working group to review the guidelines and make further improvements.

A comprehensive review of The Guidelines were undertaken in June 2017 to reflect best practices that have evolved, make references to new technologies and provide checklists for transcribers. In compiling the Guidelines, the working group consulted widely with organisations, professionals producing eText, and also undertook literature reviews of related guidelines and publications.

The related standards, guidelines and production manuals reviewed are referred to in **Appendix 4: Standards Guidelines and References.**

## **Who are these guidelines for?**

Anyone preparing electronic documents will find the guidelines useful. They have been written for professional transcribers as well as a general audience. Possible users include:

- People with a print disability who may find the guidelines useful in explaining their accessibility requirements.
- People wanting to create accessible documents.
- Specialist accessible format production agencies and educators when creating printed text in accessible electronic versions.
- Government agencies when preparing public documents.
- Organisations that may adopt the guidelines as part of their internal procedures for publication of reports and documents.
- Conference organisers who may specify that abstracts and papers should comply with the Guidelines.



## **General principles**

Provided below is a summary of four general principles that should be followed when creating accessible documents. Refer to Appendix 1 for an Accessibility Checklist.

### **1. Equivalent to print**

An accessible eText version of a print document should provide the same information as the original. This includes the following:

- Include all meaningful elements of the print document.
- Ensure accuracy.
- Indicate changes to the print with producers' notes.
- Include publishing information.
- Describe images and visual elements.
- Express special characters and languages in simple English.

### **2. Easy to navigate**

An accessible eText document should be navigable in order for it to be accessible by a person with a print disability. Also consider:

- Arranging text in a linear reading order.
- Including structural markup (e.g. heading levels).
- Ensuring consistency of style and formatting within the document.

### **3. Clear visual style**

An accessible eText document should have a clear, legible visual style that considers the needs of people with a range of print disabilities. It is recommended that users control visual appearance wherever possible.

#### **4. Standards, guidelines and best practice**

An accessible eText document should follow relevant standards and best practices where they exist, only deviating from these where necessary for particular reasons such as limitations in assistive technology. Consideration should also be given to applying applicable standards and adjusting for individual needs and preferences.

## **General principles explained**

### **1. Equivalent to print**

An accessible eText version of a print document should provide the same information as the original print document. Changes to the original print document beyond differences from the print beyond minor corrections of spelling should be indicated with a producer's note.

#### **1.1 Include all meaningful elements of the print document**

For an eText document to be accessible, all meaningful content in the print document must be included.

When producing a full book, information from the cover and preliminary pages such as the synopsis, review quotes, and notes about the author should be included as part of the eText document. Common practise is to include these at the beginning of the document.

If informative content is included in headers or footers, ensure this is repeated elsewhere in the document.

If documents are converted from print, it is preferable to use a manual contents page. A document converted from print contains two sets of page numbers (original print pages and on-screen pages in the word processor document), so automatic tables of contents can be confusing.

Print page numbers are important as they allow an eText reader to navigate to a particular page number, use the index and contents list, or reference the page where a particular quote or concept is located.

#### **1.2 Describe images and visual elements**

Where an image or visual element adds meaning to the text, that meaning should be expressed in words, or described, within the accessible eText file.

Text explanations can be included as a producer's note in the text, as "alt" attributes attached to an image, or as captions below an image.

For some readers, it may be more appropriate to provide images in an alternative format such as a tactile or enlarged graphic. If an image has been produced as a physical enlarged or tactile image, this should be explained with a producer's note.

The key to conveying visual information in eText format is to interpret what information the illustration is adding to the text, then describing this information in a simple, structured and straightforward manner.

**Note:** A good guideline is to imagine reading the document to another person, and to think about what would be said in place of the image.

Some general guidelines for describing images are:

- Identify the purpose of the photo, illustration, diagram, etc. in its context.
- Identify the key concepts conveyed in the image and explain the concepts.
- Describe the overall image first, then provide further detail as necessary.
- Organise the details of descriptions in a linear order, for example from left to right or following a process.
- Separate information into easily readable pieces by using bullet points or line breaks.
- Include all titles, captions and labels provided in the original text.
- Avoid making judgements, assumptions or interpretations of the images as much as possible. If a judgement is required, state within a producer's note that it is the producer's interpretation.

Refer to the Round Table's **Guidelines on Conveying Visual Information** for more information and examples.

### **1.3 Ensure accuracy**

The eText version must provide equivalent information to the print. On occasion this may require the correction of errors such as spelling and typographic errors in the hard copy version.

**Appendix 3: Tips for using Scanners and OCR** contains a list of errors that are commonly introduced in the Optical Character Recognition process.

## **1.4 Mark changes to the print with producer's or transcriber's notes**

A producer's note should be used to comment about any changes made to the original version when converting it to eText, such as formatting changes or omissions. For example, if a line of text is unreadable or omitted from the print, insert a producer's note such as "The following line was not legible in the print and has been omitted from this eText version."

Producer's notes can appear at the beginning or end of the file, or within the text where the note relates to a particular point.

Make it clear where text is a producer's note rather than part of the original text, for example by adding the words "Producer's Note" and "End of Producer's Note".

## **1.5 Express special characters and languages simply**

### **Special characters**

Ideally, use a Unicode character set and represent all special characters in the document as their correct Unicode equivalent. Representing a character correctly in Unicode is less ambiguous than replacing the character with words. However, some assistive technologies or screen reader software do not reliably interpret some of the Unicode characters, for example the numeral sign № is not reliably read by current screen readers. If a document is being prepared for a particular person whose technology cannot interpret Unicode, these characters need to be written out in words (for example, replace ° with "degrees", £ with "pounds").

Avoid using proprietary Windows characters like "smart quotes" (change to " and ") and ellipses (change to "three dots", for example "...").

### **Languages**

In word processing formats, HTML, DAISY and EPUB 3, it is possible to indicate which language a document is written in. In these formats, ensure that the language is set correctly and that changes in language (for example, French words in an English document) are marked consistently. This allows screen readers to read content more accurately.

## **1.6 Include publishing information**

Publishing information related to the original document should be included either within the eText file itself or in a separate "readme" file. This information should include the title, author, publisher, copyright owner, special notes about the eText version (e.g.,

omissions), transcription date, and any identifier of the text such as ISBN or ISSN and the identity and copyright permissions of the producing agency.

## **2. Easy to navigate**

An accessible eText document must be fully usable by a person with a print disability.

### **2.1 Arrange text in a linear reading order**

Text should be arranged so that it can be read in a linear way; that is, pieces of information following each other in logical sequential order.

All "floating" elements like textboxes, images, icons, and tables should be formatted as "inline", within the flow of the text.

#### **Footnotes**

- Refer to the position of the footnote in the text.
- Avoid adding footnote text where it appears visually on the print page (i.e., at the end of the page) as this can interfere with reading and comprehension.
- Common practice is to move all footnotes to the end of the file. In Word and EPUB 3, links can be created to the footnotes at the end of the document.
- Alternatively, footnotes can be presented in line with the original text; that is, the footnote text can appear in the original position of the footnote reference, or at the end of the relevant sentence or paragraph.

#### **Text in boxes and margin notes**

If a print document contains text in boxes or margin notes, these should be inserted in the most logical place in the text, within the linear flow of the document. Avoid inserting features like text boxes in Microsoft Word which separate content from the main reading order.

The beginning and end of each example of text in a box should be marked in some way so that it can be perceived by someone using a screen reader by inserting, for example, "Box" and "End of Box".

#### **Headers and footers**

Print documents often include title or pagination information in the header and footer sections of the page. An eText version should only include this information where it is informative and not repeated elsewhere in the text.

## 2.2 Include structural markup

Structural elements of the text, such as headings, paragraphs, lists, tables and emphasised words should be "marked up" so that assistive technology can interpret their significance, meaning and context. Structural markup methods include using "Styles" in word processing software, codes in plain text documents, or semantically-correct HTML elements.

**Note:** mark the main headings in a Word document with the Heading 1 style, then modify the visual appearance of that Style, rather than selecting each heading individually and applying appropriate font properties (bold, bigger text size). In HTML, use structural elements like `<h1>` and control the visual appearance with stylesheets, rather than `<font size=+2>`.

The resulting document might **look** exactly the same but the significance of the visual structuring can be interpreted for different reading methods. For example, a screen reader can extract a list of headings in a document and use them for navigation.

Common structures should be treated consistently, for example, all level 2 headings should be structurally marked in the same way, even if they don't look the same visually in print.

### Mathematics

Mathematics can be conveyed unambiguously using MathML or LaTeX. The usefulness of these formats depends on the user's reading technology and familiarity.

## 3. Clear visual style

An accessible eText document should have a clear, legible visual style that considers the needs of people with a range of print disabilities.

### 3.1 Use a clear visual design

Clear visual design of an eText document for viewing on screen requires similar design decisions to creating clear print documents. Please refer to [Round Table Clear Print Guidelines: http://printdisability.org/guidelines/guidelines-for-producing-clear-print-2011/](http://printdisability.org/guidelines/guidelines-for-producing-clear-print-2011/)

General principles are:

- Use a simple and clear typeface.
- Use a large default text size, a minimum of 12 point.

- Use 1.2 line spacing, 6 point or 10 point for paragraphs and columns.
- Ensure there is adequate contrast between text and background.
- Avoid placing text over background images.
- Ensure all text is left-aligned and set horizontally.
- Use a consistent and logical layout. Unusual spacing or layout can make it difficult to locate information on the page.
- Use **bold** to emphasise text. Avoid using italics and underline.
- Do not convey information solely through images or colours, and ensure images have a description or ""alt text"" that is meaningful.
- Give tables a bold, visible border and ensure adequate space between the border and contents of a cell.
- Ensure that structural elements like headings and lists stand out and have a consistent appearance.

### **3.2 Allow users to control visual appearance**

Use stylesheets to control text and background colour so these can be adjusted by the user. In formats where it is possible, use relative rather than absolute text sizing so that the reader can easily resize text.



## 4. Standards, Guidelines and Best Practice

An accessible eText document should follow relevant standards and best practices where they exist, only deviating from these where they are needed for particular reasons such as limitations in assistive technology like screen reader software and hardware, hearing aids and braille devices.

A document that is produced according to standards is more likely to work with a wider range of assistive technology, and is easier to modify—whether to other eText formats including webpages or to other formats like braille, large print or synthetic audio.

### 4.1 Follow applicable standards

Standards that may be applicable and provide additional information to these Guidelines include:

[Accessible Microsoft Word: http://webaim.org/techniques/word/](http://webaim.org/techniques/word/)

[HTML5: http://webaim.org/blog/future-web-accessibility-html5-semantic-tags/](http://webaim.org/blog/future-web-accessibility-html5-semantic-tags/)

[WCAS \(Web Content Accessibility Checklist\): https://www.w3.org/TR/WCAG20/](https://www.w3.org/TR/WCAG20/)

[PDF accessibility best practice:](https://www.adobe.com/content/dam/acom/en/accessibility/products/acrobat/pdfs/acrobat-xi-pro-accessibility-best-practice-guide.pdf)

<https://www.adobe.com/content/dam/acom/en/accessibility/products/acrobat/pdfs/acrobat-xi-pro-accessibility-best-practice-guide.pdf>

[DAISY: http://daisy.org/](http://daisy.org/)

[Inclusive Publishing Hub: https://inclusivepublishing.org/](https://inclusivepublishing.org/)

### 4.2 Adjust for individual need and preference

When eText is being prepared with a particular person in mind, individual need should take precedence over standards. If eText might be used by multiple people, this can be achieved by producing the file in a standards-based way and customising it for individual need. An example is using Unicode to represent special characters, then replacing these characters for people whose technology can't interpret them. Some examples of where there would be an individual need or preference include:

- Assistive technology limitations, e.g., a screen reader that is unable to read Unicode.
- Software availability, e.g., a person might not necessarily have Microsoft Word

- Varying levels of computer literacy
- Reading preferences.

### **4.3 Ensure consistency**

These guidelines are not a prescriptive formatting guide. Sometimes a producer or document author will need to make choices about document-specific formatting, such as heading styles. An important guiding principle is consistency of style and formatting within a document or across a set of documents. For organisations creating digital documents regularly, it may be helpful to develop an organisational style guide and/or create document templates to help maintain consistency.

### **Common formats**

The Common formats section summarises a range of common formats used in the production of accessible eText. These formats are:

- Microsoft Word (including Rich Text Format (RTF)).
- Plain text.
- DAISY text.
- EPUB 3.
- HTML5.
- PDF.

Where several alternative formats of a document are provided, the best practice approach is:

- Place the most accessible document first.
- Include the download file format and file size as part of the link label, e.g., Document Title (rtf 750 kb).

A brief description of the format and additional considerations are outlined below.

## 1. Microsoft Word (including RTF)

The general accessibility principles described earlier in this document should be adhered to across formats. Also refer to **Appendix 1: Accessibility Checklist**.

Word processing formats, such as Microsoft Word, RTF and OpenDocument are widely supported by current assistive technology. It is preferable to use RTF as this is supported by a wider range of software, however Word or OpenDocument can be used if requested by individuals.

RTF can be created in a variety of word processing programs, and files in Word and OpenDocument can be saved as RTF. This format can be useful for providing downloadable content on the web, so that a whole document is available as one file.

Text boxes can be represented in Rich Text Formats by marking the beginning and end of the box with distinctive styles, marking the contents of the box with a distinctive style.

### Additional considerations

- Do not use floating "textboxes" in Word.
- Use list styles.
- Aim for a Flesch Reading Score of 60 or below.

## 2. Plain text

Plain text is the most universally accessible file format. It is useful when producers are unsure of which assistive technology will be used to access the file. Different conventions exist for formatting plain text documents. Plain text has a unique set of formatting guidelines. Refer to **Appendix 2 Guidelines for formatting plain text**.

## 3. DAISY text

DAISY is a standard for structured digital books that has been developed by the DAISY Consortium, which is made up of print disability organisations around the world. A DAISY book can include audio, text and images. The treatment of text in a DAISY book is based on XML, and allows a closer representation of the print than any other electronic text format. For example:

- A reader can navigate directly to a particular page number. DAISY has a page number structure which differentiates between normal numbered pages, preliminary pages, and pages in a special sequence, like a series of images inserted between

two numbered pages. DAISY software and DAISY readers provide the ability to navigate directly to a page. In other formats, include the print page number on a separate line, marked by words (Page, Print Page) or letters (p, pp).

- Footnotes can be marked as a separate structure that the reader can choose to navigate to directly, skip over or access in line with the text.
- DAISY has its own footnote structure which provides an ideal way of representing footnotes. Readers can choose to read a document with footnotes switched on or off, and can navigate directly to a footnote. In HTML, a common practice is to make the footnote reference a link to the footnote text, which appears at the bottom of the document. The footnote text then includes a link back to the related reference.
- Text boxes can be represented by using the "sidebar" structure.

Full text DAISY books consist of marked-up text that can be accessed and navigated using [Daisy Playback Tools: http://www.daisy.org/tools/splayback](http://www.daisy.org/tools/splayback)

The DAISY website contains guidelines for producing DAISY books: [Daisy: http://www.daisy.org/](http://www.daisy.org/)

#### **4. EPUB 3**

EPUB 3 is the latest publishing format that is built upon the Open Web Standards. EPUB 3 is an offline version of a website which can be downloaded and read on a computer and on any mobile device such as smartphones and tablets.

EPUB 3 requires features and standards that should be followed when creating content that will enable reading systems to discover and render them to the user. These include features that are found on a website such as listening to audio, watching video, interactivity, animation, etc.).

EPUB 3 can be opened and read on programs like Radium, Adobe Digital Editions, VitalSource Bookshelf for Android or iBooks.

A few of the key features of EPUB 3 are listed below:

- Multimedia support: Native (video and audio) with the ability to include captions, subtitles and other timed tracks.
- Accessible SVG (Scalable Vector Graphics): Supports ALT Text; images do not pixelate when scaling.

- Media overlay: Audio and text synchronisation using Synchronised Multimedia Integration Language (SMIL).
- Global language support: Support for a diverse range of languages and writing modes which includes, writing from left-to-right/right-to-left and vertical writing
- Mathematical Markup Language (MathML): Supports the integration of mathematical content for textbooks and publications.
- Music XML: Supports the integration of musical content from musical scores and publications.
- Provides enhanced metadata information.

The International Digital Publishing Forum (IDPF) website contains guidelines to produce accessible EPUB 3. [IDPF: https://idpf.github.io/a11y-guidelines/](https://idpf.github.io/a11y-guidelines/)

### **How to create new EPUB 3 or convert existing documents to EPUB 3**

EPUB 3 can be created by converting an existing eText document or creating from scratch. We have outlined below guidelines when converting existing documents or creating new EPUB 3 documents.

#### **Converting a Word document to EPUB 3**

Shown below is a high level overview of converting a Word document to EPUB 3 using open source programs like Daisy Pipeline 2 and TOBI:

##### **Daisy Pipeline 2:**

- Create your accessible document in word, following the accessible checklist as a guideline Appendix 1: Accessibility Checklist.
- Create a DAISY book or a DTBook XML file using the ""Save as Daisy"" plugin.
- Convert DAISY book to EPUB 3 in Daisy Pipeline 2 using the ""DAISY 2.02 to EPUB 3"" or ""DTBookXML to EPUB 3"" options.

##### **TOBI Authoring Tool:**

- Create your accessible document in Word and follow the accessible checklist as a guideline Appendix 1: Accessibility Checklist.
- Create a DAISY book or a DTBook XML File using the ""Save as Daisy"" plugin

- Use Tobi to create EPUB 3 using the "Daisy to EPUB 3" input and follow the steps involved.
- Tips: Users can create media overlay (human narration or synthetic audio) using TOBI to enhance accessibility.

Last but not least, validate your EPUB 3 content by using the [ACE Accessibility Checker Tool: https://inclusivepublishing.org/toolbox/accessibility-checker/](https://inclusivepublishing.org/toolbox/accessibility-checker/) from Daisy.

Please note: Ensure Daisy Pipeline 2 is installed along with TOBI. Tobi uses Daisy Pipeline 2 to convert Daisy content to EPUB 3.

Please refer to the links below for further information:

- [DAISYPedia TOBI Convert Daisy to EPUB 3: http://www.daisy.org/daisypedia/tobi-convert-daisy-books-epub-3](http://www.daisy.org/daisypedia/tobi-convert-daisy-books-epub-3)
- [DAISYPedia Word to Accessible EPUB 3 conversion: http://www.daisy.org/daisypedia/how-convert-microsoft-word-document-accessible-EPUB-3](http://www.daisy.org/daisypedia/how-convert-microsoft-word-document-accessible-EPUB-3)

### **Other advanced editing tools**

Adobe InDesign CC is widely used by content creators and publishers to create accessible reflowable and fixed layout EPUB 3 publications. Publishers must make use of the accessibility guidelines when they are exporting their content from InDesign in EPUB 3 format in order to provide optimum usability and accessibility to all readers.

Other more advanced editing tools that are used to create EPUB 3 are oXygen XML editor, Sigil, Calibre and DITA for Publishers EPUB Plugin. However, knowledge of HTML5 and CCS is required to use the above software.

Please refer to the links below for further information on EPUB 3 editing tools:

- [DAISYPedia EPUB 3 Conversion with InDesign: http://www.daisy.org/daisypedia/using-adobe-indesign-create-accessible-epub-3-files](http://www.daisy.org/daisypedia/using-adobe-indesign-create-accessible-epub-3-files)
- [DAISYPedia tools on creating EPUB 3: http://www.daisy.org/daisypedia/tools-creating-epub-3-files](http://www.daisy.org/daisypedia/tools-creating-epub-3-files)

For further information regarding EPUB 3 e-reading system support/accessibility, refer to the Book Industry Study Group (BISG), IDPF and Daisy Consortium's [EPUB 3 Support Grid: http://epubtest.org/](http://epubtest.org/).

## 5. HTML5

HTML5 is the latest hypertext markup language for creating accessible websites. It incorporates standard features that were also available in previous version of HTML.

HTML5 has included new support for multimedia features (<video> <audio>) tags and new sets of html tags (<header>, <footer>, <nav>, <section>, <aside>, <article>, etc.,) to enhance accessibility. It is simple to create, easy to navigate and supported by most assistive technology.

Assistive technology can interpret information from structural elements such as headings, lists, and anchor tags. As an example, most screen readers can extract a list of heading tags from a HTML page and use them like a table of contents.

The general accessibility Checklist described in **Appendix 1: Accessibility Checklist** should be followed as a general guide when creating your base document.

## 6. Portable Document Format (PDF)

PDF is a file format developed by Adobe that makes it possible to reliably create, combine, and control text and graphical documents with the original formatting intact.

The general accessibility checklist described in **Appendix 1: Accessibility Checklist** should be followed as a general guide when creating your base document. If this base document was created in MS Word or RTF, it can be saved as a PDF, which will be mostly accessible.

### Additional Considerations

- Document title and language must be defined.
- Bookmarks should be used in long documents in addition to headings to navigate content.
- All text must be searchable (ensure the PDF is not a scanned image).
- When converting from Word, make sure the ""Document structure tags"" for accessibility check box is selected, and then click OK.





## Glossary

**Accessible:** Usable by a person with a print disability, in the sense of perceivable, operable, understandable and robust.

**Alternative Format:** Any text that has been reproduced in either large print, electronic text (eText), braille, or audio format.

**Assistive Technology:** Technology (hardware and software) that enables individuals with a print disability to access information. Also known as "adaptive technology" and "access technology".

**Caption:** Written description that normally accompanies a picture, video, chart or diagram.

**Character Set:** A defined list of characters recognised by computer software. Each character is represented by a number. The ASCII character set, for example, uses the numbers 0 through 127 to represent all English characters as well as special control characters. European ISO character sets are similar to ASCII, but they contain additional characters for European languages. Unicode is a widely used character set for representing characters in any language that is backwards compatible with ASCII.

**Colour Contrast:** The contrast between the brightness and hue of text and background colours.

**Cascading Style Sheets (CSS):** A style sheet language used to control the presentation of a document written in a markup language. Its most common application is to style documents written in HTML and XHTML. CSS is used to define colours, fonts, layout, and other aspects of document presentation. It is designed primarily to enable the separation of document content (structured by HTML or a similar markup language) from document presentation (written in CSS).

**DAISY:** An open standard for accessible digital books. Refer to [Daisy:](http://www.daisy.org) <http://www.daisy.org>.

**Endnote:** A note placed at the end of an article, chapter, or book that comments on or cites a reference for a designated part of the text.

**EPUB:** An open platform e-book format that can be downloaded and read on a computer and any device like a smartphone and tablets.

**EPUB 3:** Latest ePub format which is built upon the Open Web Standards—HTML 5, CSS3, WAI-ARIA, JavaScript, SVG, SMIL.

**Flesch Reading Score:** The Flesch reading score of the text indicates how easy the content is to comprehend, i.e., whether it is readable. The higher the score, the easier it is to understand the text.

**Footnote:** A note placed at the bottom of a page of a book that comments on or cites a reference for a designated part of the text.

**Heading Styles:** Heading styles are usually hierarchical. Most texts can be divided into main sections such as chapters, then into smaller sections, then into sub-sections, and so forth. Headings are used to label each piece of the text. The major headings through the document should be assigned Heading 1. Heading 2 applies to subheadings, and so on. HTML allows a maximum of 6 levels of heading.

**HTML:** Latest Markup Language for website creation.

**IDPF:** International Digital Publishing Forum.

**Large Print:** Print that is enlarged and reformatted for clarity, designed to be optimally legible for people with low vision.

**LaTeX:** A document markup language and document preparation system for the TeX typesetting program, which can be used to lay out complex mathematical equations.

**Markdown:** A human-readable markup system based on plain text email conventions:  
[Daring Fireball website: http://daringfireball.net/projects/markdown/](http://daringfireball.net/projects/markdown/)

**Markup:** Characters or text that are placed in a document to indicate how a document looks or how the document is structured to be accessible and usable.

**MathML:** Mathematical Markup Language (MathML) is an application of XML developed by the W3C for describing mathematical notations and capturing both its structure and content. It aims at integrating mathematical formulae into text documents.

[W3C Math Home: http://www.w3.org/Math/](http://www.w3.org/Math/)

**Metadata:** information about a document that is used to facilitate the understanding, use and management of the document.

**Optical Character Recognition (OCR):** The electronic identification and digital encoding of printed or handwritten characters by means of an optical scanner and specialised software.

**Print Disability:** An individual with a print disability is unable to access the information disseminated in a regular print format because they:

1. are blind or vision impaired.
2. have physical disabilities which limit their ability to hold or manipulate information in a printed form.
3. have perceptual or other disabilities which limit their ability to follow a line of print or which affect their concentration or,
4. cannot comprehend information in a print format due to insufficient literacy or language skills.

**Producer:** Someone responsible for the production of alternative format material.

**Producer's Note:** A note included by the producer to indicate differences from the print or the producer's interpretation of visual material.

**Project Gutenberg:** Project Gutenberg is the first and largest single online collection of free electronic books, or e-books. [Project Gutenberg: http://www.gutenberg.org](http://www.gutenberg.org)

**Semantic Markup:** Using the most meaningful tag to describe the type of content.

**SMIL:** Synchronized Multimedia Integration Language.

**Styles:** A style is a set of formatting characteristics that can be applied to text in a document to quickly change its appearance.

**Stylesheet:** A specification, describing how elements of another document should be displayed.

**Subscripts and Superscripts:** A number, figure, symbol, or indicator that appears smaller than the normal line of type and is set slightly below or above it; subscripts appear below the baseline, while superscripts are above. Subscripts and superscripts

are typically used in formulas, mathematical expressions, and descriptions of chemical compounds or isotopes, but have many other uses as well.

**SVG:** Scalable Vector Graphic.

**Tactile Graphics:** Images that use raised surfaces so that a vision impaired person can feel them. They are used to convey non-textual information such as maps, paintings, graphs and diagrams.

**Unicode:** Provides a unique number for every character.

**W3C:** World Wide Web Consortium. The main international standards organisation for the World Wide Web: [W3C: http://www.w3.org/](http://www.w3.org/)

**WAI-ARIA:** Web Accessibility Initiative – Accessible Rich Internet Applications

**WCAG:** Web Content Accessibility Guidelines (WCAG) are published by the W3C's Web Accessibility Initiative. They give guidance on making content accessible, primarily for users with a disability, but also for all user agents, including highly limited devices, such as mobile phones: [Web Accessibility Initiative: http://www.w3.org/WAI/](http://www.w3.org/WAI/)

**XML:** Stands for **Extensible Markup Language**. A general-purpose markup language used to encode structured information. A document written in XML contains content (text) and structure (elements or "tags" which describe the content's purpose), but does not describe how the content should be displayed. This separation of content and presentation allows an XML document to be easily converted between different presentations such as a webpage, print document, braille or synthetic speech.

## Appendix 1: Accessibility Checklist

The accessible checklist below outlines key elements that should be considered when creating accessible documents. The checklist follows the four general principles discussed in The Guidelines.

### 1. Equivalent to print

#### Images

- Images have a description or "alt text". Alt text descriptions are included in the **Description field**, not in the **Title field** in the MS Word dialog, for example
- Images are inline with the text

### 2. Easy to Navigate

#### Headings

- Headings are formatted using pre-set styles
- Headings are logically nested in a hierarchical fashion

#### Hyperlinks

- Hyperlinks and email links are active/live
- Link text is descriptive, not just URLs

#### Page Numbers

- Page numbers are included in documents with more than one page

#### Tables

- Tables are solely used to represent tabular data
- Tables have a bold, visible border and there is adequate space between the border and contents of a cell
- Header rows are specified in tables
- Header rows are repeated across pages

## **Table of Contents**

- Large documents have a table of contents

## **3. Clear Visual Style**

### **Font**

- Default text size is a minimum of 12pt
- Bold is used to emphasise text rather than italics or underline

### **Layout**

- Layout is consistent, logical and in linear order
- Minimum 1.2 line spacing, 6pt or 10pt space between paragraphs

### **Text**

- Text has a clear typeface
- Text is legible and not placed over background images
- Text is left-aligned and set horizontally
- High contrast between text and background

### **Other considerations**

- Information is not solely conveyed through colour
- Use simple language: short sentences and words

### **Accessibility Checker**

- Document is checked for accessibility using the native accessibility checker

## **4. Follow standards, guidelines and best practices**

- Please refer to appendix 4

## Appendix 2: Guidelines for Formatting Plain Text

### Document setup

- The first line of a document should be the title (omit any blank lines before the title).
- Provide a list of the codes at the beginning of the file, or in a separate readme file.

### Line breaks

- Insert a blank line between paragraphs
- Leave two blank lines before all headings and one blank line between consecutive headings. Leave a blank line after all top level headings.
- Indicate the beginning of a print page on a new line. Leave a blank line before a page indicator.
- Blank lines can be added between list items or rows of a table if needed for clarity.

### Emphasis

Different types of emphasis include block capitals, bold, italic and underline. These can often be omitted unless they are needed to understand the meaning of the document. A common internet convention uses asterisks and underscores for **\*bold\*** and *\_italics\_*.

### Lists

- Retain numbering and letters to indicate listed items. Insert a full stop character after the number or letter to clearly differentiate it from the following text.
- Indicate bulleted lists with asterisks. For multi-level lists, use – (hyphens) and – (dashes) to indicate different levels. For example:

- Countries
  - Australia
  - New Zealand
  - India

## Tables

Indicate the beginning and end of a table with the words "Table" and "End of Table." If the table has a title, use "Table: Title"

Use a producer's note to explain how the table has been presented.

Separate cells with a distinctive character (| is commonly used). Indicate blank cells with a hyphen ( - ) or a word like "nil" or "blank".

### A basic way of presenting tables

Table

Producer's note: The following table has three columns with the headings: "A", "B", "C". The columns are separated by the | character. End of Producer's note.

A | B | C

Mazda | 1985 | \$15,000

Toyota | 2003 | \$30,000

Honda | 2013 | \$50,000

End of Table



## **Appendix 3: Tips for using Scanners and OCR**

The advent of scanners and Optical Character Recognition (OCR) software has enabled the easy transfer of printed text into an electronic format. However, production of a well-formatted and correct document will usually require additional proofreading and formatting.

There are a number of OCR software packages and apps available and the following suggested tips are generic in nature and not specific to any particular software.

### **General tips**

- Bound books cannot always be laid perfectly flat for scanning. The curve where the spine occurs can create distortion. If it is feasible to unbind or remove the binding of a book, do so or else you may need to use manual pressure to ensure that each page is as flat against the scanner glass as possible.

### **Formatting**

- **Literal layout vs linear formatting:** When using OCR to create a document for an electronic format, it is important to first have a clear understanding of these Guidelines for Accessible eText. Always keep in mind that the eText format is a representation of the document rather than a literal recreation of the actual layout. For example, text in multiple columns will be reduced to a single column, and text boxes will be removed.
- **Reading Order:** Not all printed material is linear in its presentation. Ensure that a logical reading order is maintained. Some software will allow you to select blocked areas of a page in the order that OCR is to occur.
- **Columns:** Where there are columns on the printed page, it is important that the OCR software recognises the separate columns and does not "read across the page."
- **Tables:** ensure that rows and columns are correct in complex tables, e.g. where a single cell of a table in print goes over more than one line. Some OCR software will allow the marking of rows and columns of tables to ensure that each cell is correctly identified.

## **Proofreading**

There are a number of common errors which can occur during OCR, not all of which will be picked up by a spell checker. Careful proofreading is still required for an accurate document.

If your OCR has a spell and/or grammar checker, utilise these as well as those in your word processor. Different software will pick up different errors.

Listed below are some common errors which are worth looking for.

- Misrepresented letters: These are not always picked up in spell checkers. For some documents, search and replace can be a useful tool.
  - 1 for l (numeral one and capital i)
  - O and 0 (capital o and numeral zero)
  - b and h (e.g., he instead of be and vice versa)—grammar checkers will sometimes identify these
  - m for rn (e.g., modem instead of modern)
- Punctuation:
  - Spacing: e.g. additional spaces before periods, question marks, etc.
  - Quotation marks: Ensure that open and close quotes are balanced i.e., all open quotes have corresponding close quotes. Open quotes can often be missed at the beginning of paragraphs by the OCR software.

## **Foreign languages**

In most OCR programs you can choose the language or languages for recognising text in the source document. This ensures that accents on characters are recognised correctly.

---

## Appendix 4: Standards, Guidelines and References

*Microsoft Word - Creating Accessible Documents*, WEB AIM last updated July 22, 2016, accessed 4 April 4, 2017 <[WebAim: http://webaim.org/techniques/word/](http://webaim.org/techniques/word/)>

*Future Web Accessibility: HTML5 Semantic Elements*, May 4, 2010 accessed April 4, 2017 <[WebAim: http://webaim.org/blog/future-web-accessibility-html5-semantic-tags/](http://webaim.org/blog/future-web-accessibility-html5-semantic-tags/)>

*Web content Accessibility Guidelines (WCAG) 2.0*, 11 December 2008 accessed April 4, 2017 <[W3C: https://www.w3.org/TR/WCAG20/](https://www.w3.org/TR/WCAG20/)>

Adobe® Acrobat® XI Pro Accessibility Guide: *Best Practices for PDF Accessibility*, 2012 accessed April 4, 2017 <[Adobe Acrobat: http://www.adobe.com/content/dam/Adobe/en/accessibility/products/acrobat/pdfs/acrobat-xi-pro-accessibility-best-practice-guide.pdf](http://www.adobe.com/content/dam/Adobe/en/accessibility/products/acrobat/pdfs/acrobat-xi-pro-accessibility-best-practice-guide.pdf)>

*Constructing a POUR website* last updated August 28, 2013, accessed April 4, 2017 <[WebAim: http://webaim.org/articles/pour/](http://webaim.org/articles/pour/)>EPUB 3 Accessibility Guidelines IDPF last updated August 5, 2016 accessed April 5, 2017: <[IDPF: https://idpf.github.io/a11y-guidelines/](https://idpf.github.io/a11y-guidelines/)>

What is EPUB 3 2011, O' Reilly Media, accessed May 15, 2017 <[What is EPUB 3: http://shop.oreilly.com/product/0636920024897/ViewLarger.do?sortBy=publicationDate](http://shop.oreilly.com/product/0636920024897/ViewLarger.do?sortBy=publicationDate)>