

# Handout for “Beyond WCAG for PDF: Understanding PDF/UA and the Matterhorn Protocol”

## Slide 1 – Title Slide

**Beyond WCAG for PDF:  
Understanding PDF/UA and the Matterhorn Protocol**

By Bouton Jones

IT Business Systems Analyst Senior  
[bouton.jones@gmail.com](mailto:bouton.jones@gmail.com)  
My pronouns are he and him



4:00 PM to 5:30 PM on May 13, 2026  
John Slatin AccessU 2026  
Trustee 118, St. Edward's University

**Presentation Title:** Beyond WCAG for PDF: Understanding PDF/UA and the Matterhorn Protocol

**Author:** Bouton Jones

Bouton Jones  
IT Business Systems Analyst Senior  
[bouton.jones@gmail.com](mailto:bouton.jones@gmail.com)  
My pronouns are he and him

4:00 PM to 5:30 PM on May 13, 2026  
John Slatin AccessU 2026  
Trustee 118, St. Edward's University

Many of us look at A11y from the perspective of web development, but ignoring office documents and PDFs would be a huge mistake. When we do apply A11y to PDFs, it's by adapting WCAG principles. That's a great place to start. But there are other relevant guidelines to follow. In this session, we'll compare the WCAG and PDF/UA guidelines and recognize which

individual guidelines apply to PDF Accessibility. We'll also look at the Matterhorn Protocol and see how it relates to PDF/UA. We'll also cover the necessity and techniques of embedding fonts in both MSO documents and PDFs.

This will build upon my earlier presentation, but it will not be necessary to have attended the earlier one to understand this next one.

#### HIGHLIGHTS:

- How do the WCAG and PDF/UA guidelines relate to each other?
- Which specific guidelines apply to PDFs?
- What are their respective limitations and potential pitfalls?
- What is the Matterhorn Protocol and how does it apply?
- What are the checkpoints unique to the PDF/UA and which checkpoints might not be relevant and in what situations?
- Why would you lose headings when exporting Word documents to PDF?
- Why and how do we embed fonts?
- What happens to the document title when it is exported from a Word document to PDF?
- How can OCR be deceptive?
- How should AT address the content in headers and footers and what does that mean to document authors?
- When are "article threads" an issue?
- What is XFA and when is it a problem?
- What are the Base 14 fonts?
- What is the PDF Association's recommendation on the best approach to generating universally-accessible PDF documents?

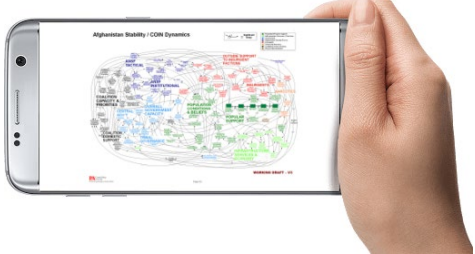
#### THREE TAKE-AWAYS:

- Making the most accessible PDFs possible
- Locating the best single checklist for PDF A11y
- Identifying which guidelines are N/A

## Slide 2 - Life is too short to be spent photographing PowerPoint slides.

**Life is too short to be spent photographing slides.**

A handout containing all the slides and speaker notes can be found at <http://bit.ly/2y0p5kJ>.



**NEXT SLIDE**



A handout containing all the slides and speaker notes can be found at <http://bit.ly/2y0p5kJ> . (or <http://www.boutonjones.com/pub/a11y.html> .)

## Slide 3 – QR Code for “Don’t Remediate Your PDFs” Handout

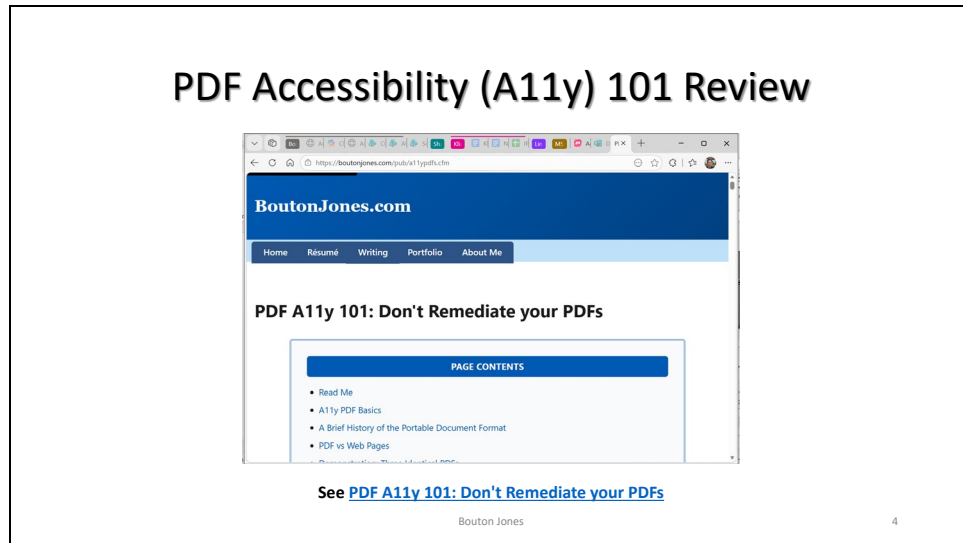


QR Code for the web page where the handout can be downloaded.

Short URL: <http://bit.ly/2y0p5kJ>

Long URL: <http://www.boutonjones.com/pub/a11y.html>

## Slide 4 - PDF Accessibility (A11y) 101 Review



At AccessU 2024 this session was part two of two sessions covering PDF accessibility.

For a better understanding of the material covered, you might want to read this web page.

See PDF A11y 101: Don't Remediate your PDFs at

<https://boutonjones.com/pub/a11ypdfs.cfm>

## Slide 5 - In Review (Slide 1 of 5): Do Not Scan to PDF

### In Review (Slide 1 of 5): Don't Scan to PDF



**PITFALL:** When you scan a printed document – e.g. a signed contract, police report, corporate policy, or anything requested for discovery – it is no more than an image of the hard copy. It will not be accessible to screen readers in its current form.



“A picture may be worth a thousand words --- but not if it's a picture of a thousand words that you're trying to read with accessible technology.”  
- David Ondich  
ADA Program Manager for the City of Austin

If you must scan to PDF, then perform OCR and then remediate.

Bouton Jones 5

**PITFALL:** When you scan a printed document – e.g. a signed contract, police report, corporate policy, or anything requested for discovery – it is no more than an image of the hard copy. It will not be accessible to screen readers in its current form.

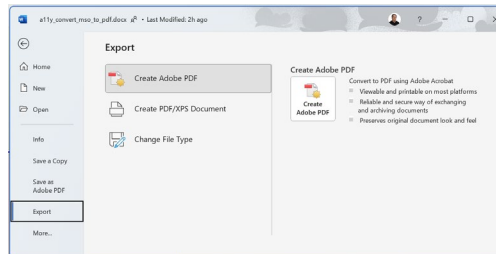
If you can't “copy and paste” text, it most likely is an image.

As David Ondich (The ADA Program Manager for the City of Austin) has said “A picture may be worth a thousand words --- but not if it's a picture of a thousand words that you're trying to read with accessible technology.”

## Slide 6 - In Review (Slide 2 of 5): Export to PDF

### In Review (Slide 2 of 5): Export to PDF

- **Best Option:** *Export > Create Adobe PDF (Word ribbon)* with Adobe Acrobat for full accessibility and compliance.
- **Next Best:** *Save As PDF* or *Export (Word built-in)* if Adobe is not available.
- **Worst:** *Print to Microsoft PDF* or *Scan to PDF* (manual)



See [Converting Microsoft Office Documents to PDF](http://bit.ly/2y0p5kl) for more information.

Sunday, May 10, 2026

Get the slides and notes at <http://bit.ly/2y0p5kl>

1

- **Best Option:** *Export > Create Adobe PDF (Word ribbon)* with Adobe Acrobat for full accessibility and compliance.
- **Next Best:** *Save As PDF* or *Export (Word built-in)* if Adobe is not available.
- **Worst:** *Print to Microsoft PDF* or *Scan to PDF* (manual)

See [Converting Microsoft Office Documents to PDF](http://bit.ly/2y0p5kl) for more information.

## Slide 7 - In Review (Slide 3 of 5)

### In Review (Slide 3 of 5)

- Only use PDFs when they are the best option.
- It's better to remediate and export an accessible source document to PDF than to remediate an inaccessible PDF.
- Testing
  - Perform both automated *and* manual testing.
  - Check the accessibility of your **source documents** before exporting to PDF.
  - Check the accessibility of your **PDFs** before publishing.

Bouton Jones

2

- Don't use PDFs unless they are the best option.
- It's better to remediate and export an accessible source document to PDF than to remediate an inaccessible PDF.
- Testing
  - Perform both automated and manual testing.
  - Check the accessibility of your source documents before exporting to PDF.
  - Check the accessibility of your PDFs before publishing.



## Slide 8 - In Review (Slide 4 of 5)

### In Review (Slide 4 of 5)

- Apply WCAG Guidelines to PDFs
  - Use Semantic markup
  - Use Alternative text or captions for non-decorative images.
  - Specify the document language
  - Color guidelines (e.g., contrast, color blindness, etc.,)
  - Avoid using tables for layout
- Electronic Forms
  - Make sure form fields are in the correct order
  - Make sure form fields are properly labeled
- Properly redact documents electronically (do not mask.)

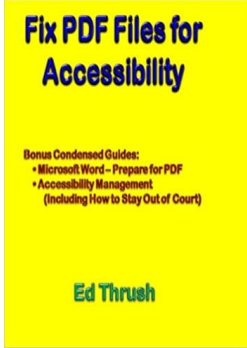
Bouton Jones

3

- Apply WCAG Guidelines to PDFs
  - Use Semantic markup.
  - Use Alternative text or captions for non-decorative images.
  - Specify the document language
  - Color guidelines (e.g. contrast, color blindness, etc.,)
- Electronic Forms
  - Make sure form fields are in the correct order
  - Make sure form fields are properly labeled
- Properly redact documents electronically (do not mask.)

## Slide 9 - In Review (Slide 5 of 5): A Book about PDF Remediation

**In Review (Slide 5 of 5)**



*Fix PDF Files for Accessibility: The Easy and Ultimate Guide to Applying Section 508 and WCAG 2.0 to PDF Files*

1st Edition  
by Ed Thrush  
Paperback : 148 pages  
ISBN-10 : 1514608936  
ISBN-13 : 978-1514608937  
<https://www.amazon.com/Fix-PDF-Files-Accessibility-Ultimate/dp/1514608936/>

Bouton Jones 1

*Fix PDF Files for Accessibility: The Easy and Ultimate Guide to Applying Section 508 and WCAG 2.0 to PDF Files*

1st Edition

by Ed Thrush

Paperback : 148 pages

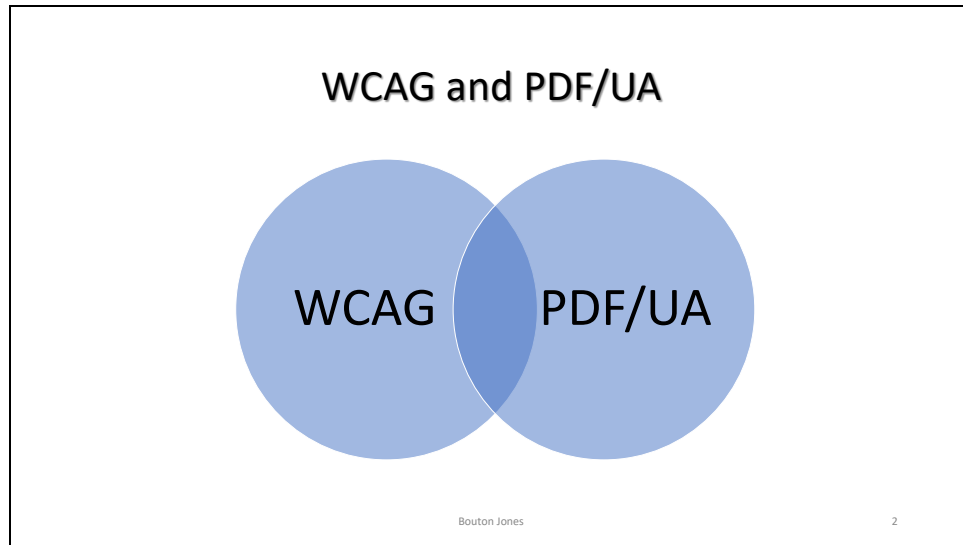
ISBN-10 : 1514608936

ISBN-13 : 978-1514608937

<https://www.amazon.com/Fix-PDF-Files-Accessibility-Ultimate/dp/1514608936/>

**AMAZON.com DESCRIPTION:** This book provides the easiest way to get to the required accessibility of a PDF file. It is rare for any PDF to come out perfect for accessibility and as a result almost all PDF files will need some type of work. In addition to the easy nuts and bolts methods to make PDF files work for the disabled, there are also condensed methods for preparation of documents commonly converted to PDF. In addition, there are instructions for managers and even contract language. You can get a legal history or technical lesson elsewhere, this book is for getting to work fixing PDF files for accessibility.

## Slide 10 - Title Slide: WCAG and PDF/UA



**IMAGE:** Very simple Venn diagram for WCAG and PDF/UA. Two overlapping circles. No data points.

## Slide 11 - Two Organizations and Their Guidelines

| Two Organizations and Their Guidelines                                                                                                                                                                         |                                                                                                                                                                                                                         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C)                                                                                                                                      | PDF Association                                                                                                                                                                                                         |
| <ol style="list-style-type: none"><li>1. Web Content Accessibility Guidelines (WCAG)</li><li>2. Authoring Tool Accessibility Guidelines (ATAG)</li><li>3. User Agent Accessibility Guidelines (UAAG)</li></ol> | <p>PDF/Universal Accessibility (PDF/UA) or ISO 14289-1</p> <ul style="list-style-type: none"><li>• How to produce accessible PDFs</li><li>• “Define compliant assistive technology and PDF reading software.”</li></ul> |

Bouton Jones 3

A table with two columns, a header row, and a data row.

First column

Header: The Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C)

Data:

- Web Content Accessibility Guidelines (WCAG)
- Authoring Tool Accessibility Guidelines (ATAG)
- User Agent Accessibility Guidelines (UAAG)

Second Column

Header: PDF Association

Data:

PDF/Universal Accessibility (PDF/UA) or ISO 14289-1

- How to produce accessible PDFs
- “Define compliant assistive technology and PDF reading software.”

## Slide 12 - Guidelines published by the International Organization for Standardization (ISO)



### Guidelines published by the International Organization for Standardization (ISO)



- Web Content Accessibility Guidelines (WCAG) 2.0
- World Wide Web Consortium (W3C)
- For websites, HTML-based material, and additional web technologies (*including PDFs*)
- Guidelines for A11y web designers, web developers, and CMS authors.
- 50 Checkpoints/Guidelines

- PDF/Universal Accessibility (PDF/UA) or ISO 14289-1 (i.e., PDF/UA-1)
- PDF Association
- For PDF documents
- Guidelines for producing accessible PDFs but also “define compliant assistive technology and PDF reading software.”
- 31 Checkpoints (many of which overlap with the WCAG)

Bouton Jones

4

### Two Sets of Guidelines:

- Web Content Accessibility Guidelines (WCAG) 2.0
- PDF/UA (Universal Access, or ISO 14289)

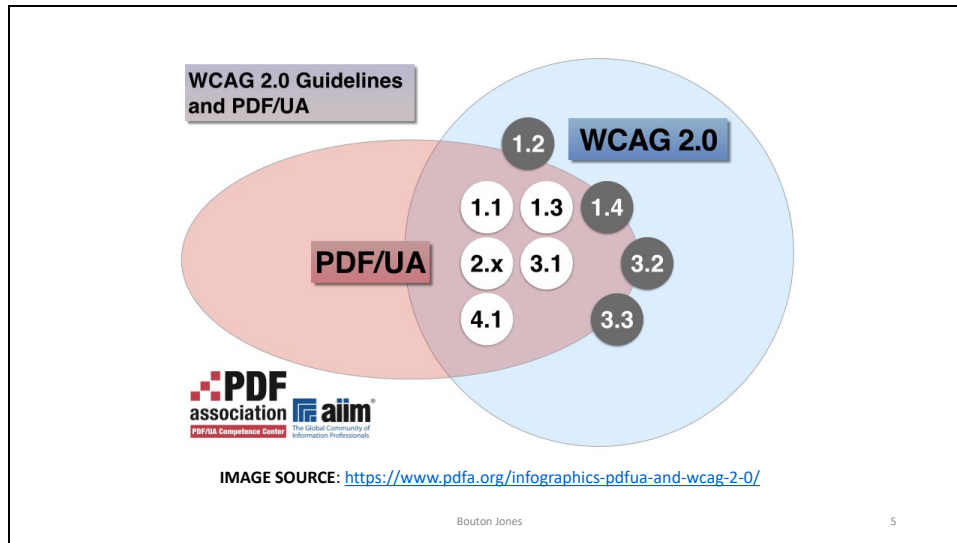
### WCAG

- Web Content Accessibility Guidelines (WCAG) 2.0
- World Wide Web Consortium (W3C)
- For websites, HTML-based material, and additional web technologies (*including PDFs*)
- Guidelines for A11y web designers, web developers, and CMS authors.
- 50 Checkpoints/Guidelines

### PDF/UA

- PDF/Universal Accessibility (PDF/UA) or ISO 14289-1
- PDF Association
- For PDF documents
- Guidelines for producing accessible PDFs but also “define compliant assistive technology and PDF reading software.”
- 31 Checkpoints (many of which overlap with the WCAG)

## Slide 13 – Venn Diagram from the PDF Association



**IMAGE:** an overly complex Venn Diagram for PDF/UA and WCAG.

**IMAGE SOURCE:** <https://www.pdfa.org/infographics-pdfua-and-wcag-2-0/>

**NOTE:** I didn't find this article helpful.

## Slide 14 - Update: PDF/UA-2 (ISO 14289-2)

### Update: PDF/UA-2 (ISO 14289-2)

- Officially published on March 15, 2024.
- Built on PDF 2.0;
- PDF/UA-2 is directly aligned with the PDF Association's [Well-Tagged PDF \(WTPDF\)](#) specification.
- Many tools (PAC 2026, Acrobat, Foxit, axesPDF, MadeToTag) support at least partial PDF/UA-2.
- PDF/UA 2 now available at no cost via the [PDF Association's sponsored bundle](#). The bundle includes the official ISO 14289-2:2024 standard, which defines accessibility for PDF 2.0.

Bouton Jones

6

- Officially published on March 15, 2024.
- Built on PDF 2.0.
- PDF/UA-2 is directly aligned with the PDF Association's Well-Tagged PDF (WTPDF) specification.
- Offers improved support for math content (MathML), annotations, and complex document structures.
- PDF/UA-2 introduces changes to tagging rules, math, SVG, annotations, and structure types.
- Many tools (PAC 2026, Acrobat, Foxit, axesPDF, MadeToTag) support at least partial PDF/UA-2.

## Slide 15 - Should you follow PDF/UA-1 or PDF/UA-2?

### Should you follow PDF/UA-1 or PDF/UA-2?

- PDF/UA-2 (ISO 14289-2:2024) is not backward compatible with PDF/UA-1 (ISO 14289-1:2014).
- PDF/UA-1 (ISO 14289-1) remains fully applicable and is widely supported
- You are not obligated to use PDF/UA-2, but it is recommended for modern, complex documents needing better math or structure support.
- *Except where noted, I will reference PDF/UA-1.*

Bouton Jones

7

- PDF/UA-2 (ISO 14289-2:2024) is not backward compatible with PDF/UA-1 (ISO 14289-1:2014).
- PDF/UA-1 (ISO 14289-1) remains fully applicable and is widely supported
- You are not obligated to use PDF/UA-2, but it is recommended for modern, complex documents needing better math or structure support.
- *Except where noted, I will reference PDF/UA-1.*



## Slide 16 - Guideline Comparison

| Guideline Comparison      |                      |
|---------------------------|----------------------|
| Standard                  | Number of Guidelines |
| WCAG (not PDF applicable) | 27                   |
| Guidelines in Common      | 23                   |
| PDF/UA-1 Only             | 8                    |

**NOTE:** This is subjective. These numbers are rough estimates and subject to interpretation. These guidelines rarely correspond one-to-one. Your findings will be close but might vary from mine.

Bouton Jones 8

There is a large overlap in guidelines between WCAG and PDF/UA but each one also have guidelines that the other does not

**DISPLAY:** a two-column table for 1) the specific standard and 2) number of guidelines

- 27 WCAG guidelines that are not applicable to PDF
- 23 guidelines in common with both WCAG and PDF/UA
- 8 more guidelines for PDF/UA only

**NOTE:** This is subjective. These numbers are rough estimates and subject to interpretation. These guidelines rarely correspond one-to-one. Your findings will be close to mine but might vary a bit.

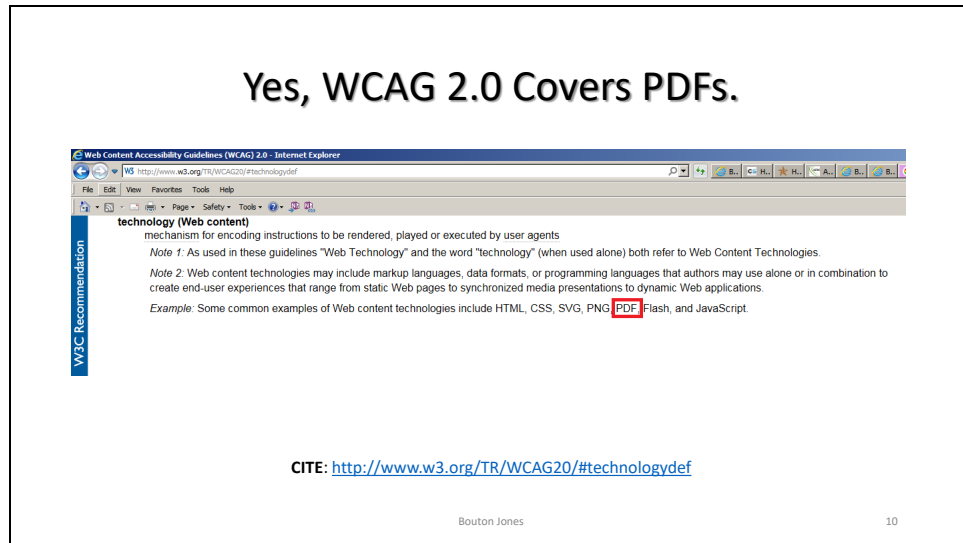
**CITE:**

- <https://www.w3.org/WAI/GL/WCAG20-TECHS/pdf.html>
- <https://www.pubcom.com/blog/2017-01-03/wcag-pdf-standards.html>

## Slide 17 - Title slide: WCAG for PDF

# WCAG for PDF

## Slide 18 - WCAG 2.x Covers PDFs.



**IMAGE:** Screen capture of a page on the W3C website.

The W3C includes "PDF" as a web technology in their Web Content Accessibility Guidelines (WCAG) 2.0. Here's a blockquote from their website:

### **technology (Web content)**

mechanism for encoding instructions to be rendered, played or executed by user agents

*Note 1:* As used in these guidelines "Web Technology" and the word "technology" (when used alone) both refer to Web Content Technologies.

*Note 2:* Web content technologies may include markup languages, data formats, or programming languages that authors may use alone or in combination to create end-user experiences that range from static Web pages to synchronized media presentations to dynamic Web applications.

*Example:* Some common examples of Web content technologies include HTML, CSS, SVG, PNG, PDF, Flash, and JavaScript.

**Source:** Web Content Accessibility Guidelines (WCAG) 2.0, Appendix A: Glossary, <http://www.w3.org/TR/WCAG20/#technologydef>

## Slide 19 - PDF Guidelines: WCAG Only

**PDF Guidelines: WCAG Only**

“WCAG 2.0 Requirements Beyond PDF/UA” (on the AIIM web site):  
[https://www.aiim.org/Global/AIIM\\_Widgets/Community\\_Widgets/Achieving\\_WCAG](https://www.aiim.org/Global/AIIM_Widgets/Community_Widgets/Achieving_WCAG)

Bouton Jones 1

“WCAG 2.0 Requirements Beyond PDF/UA” (on the AIIM web site):  
[https://www.aiim.org/Global/AIIM\\_Widgets/Community\\_Widgets/Achieving\\_WCAG](https://www.aiim.org/Global/AIIM_Widgets/Community_Widgets/Achieving_WCAG)

From “WCAG 2.0 Requirements Beyond PDF/UA” (on the AIIM web site)

--- Begin Blockquote ---

### WCAG 2.0 Requirements Beyond PDF/UA

WCAG 2.0 provides guidelines for general web content and is thus far broader in scope than PDF/UA. In some cases, therefore, PDF/UA conformance alone cannot assure conformance with WCAG 2.0 and direct reference to the WCAG 2.0 Guidelines is required.

The following list explains terms used in the Achieving WCAG 2.0 with PDF/UA table to help identify cases in which a WCAG 2.0 conformance claim cannot use PDF/UA alone to conform to WCAG 2.0 but requires validation against WCAG 2.0 directly in addition to applicable PDF/UA requirements.

- Time-based media alternatives, test, sensory and CAPTCHAs. PDF/UA does not address these items. Refer to WCAG 2.0.
- Audio and/or Video content. PDF/UA includes syntactical requirements but does not address the actual audio and video content. Refer to WCAG 2.0.
- Video captions. PDF/UA does not address video captions. Refer to WCAG 2.0.
- JavaScript design. While identifying some requirements for accessible scripting with PDF technology, PDF/UA is not a JavaScript accessibility standard. Refer to WCAG 2.0.

- JavaScript or Media content. As above, PDF/UA is not a JavaScript or multimedia accessibility standard. Refer to WCAG 2.0.
- Control device specific. PDF/UA is control device-independent. Refer to WCAG 2.0.
- Design consideration. In general, PDF/UA does not address visual appearance, content or design choices and relies on semantic equivalents. Refer to WCAG 2.0.
- Dynamic XFA-PDF. While prohibited in PDF/UA, dynamic XFA-PDF is not prohibited by WCAG 2.0.

--- End Blockquote ---

## Slide 20 - PDF Techniques in WCAG 2.0 (from the W3C)

### PDF Techniques in WCAG 2.x (from the W3C WAI)

|                                                                                                              |                                                                                                                  |
|--------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| 1. Applying text alternatives to images with the Alt entry in PDF documents                                  | 15. Providing submit buttons with the submit-form action in PDF forms                                            |
| 2. Creating bookmarks in PDF documents                                                                       | 16. Setting the default language using the /Lang entry in the document catalog of a PDF document                 |
| 3. Ensuring correct tab and reading order in PDF documents                                                   | 17. Specifying consistent page numbering for PDF documents                                                       |
| 4. Hiding decorative images with the Artifact tag in PDF documents                                           | 18. Specifying the document title using the Title entry in the document information dictionary of a PDF document |
| 5. Indicating required form controls in PDF forms                                                            | 19. Specifying the language for a passage or phrase with the Lang entry in PDF documents                         |
| 6. Using table elements for table markup in PDF Documents                                                    | 20. Using Adobe Acrobat Pro's Table Editor to repair mistagged tables                                            |
| 7. Performing OCR on a scanned PDF document to provide actual text                                           | 21. Using List tags for lists in PDF documents                                                                   |
| 8. Providing definitions for abbreviations via an E entry for a structure element                            | 22. Indicating when user input falls outside the required format or values in PDF forms                          |
| 9. Providing headings by marking content with heading tags in PDF documents                                  | 23. Providing interactive form controls in PDF documents                                                         |
| 10. Providing labels for interactive form controls in PDF documents                                          |                                                                                                                  |
| 11. Providing links and link text using the Link annotation and the /Link structure element in PDF documents |                                                                                                                  |
| 12. Providing name, role, value information for form fields in PDF documents                                 |                                                                                                                  |
| 13. Providing replacement text using the /Alt entry for links in PDF documents                               |                                                                                                                  |
| 14. Providing running headers and footers in PDF documents                                                   |                                                                                                                  |

CITE: <https://www.w3.org/WAI/WCAG22/Techniques/#pdf>

Bouton Jones 1

This is a huge web page! Here's an outline of the 23 techniques it covers.

1. Applying text alternatives to images with the Alt entry in PDF documents
2. Creating bookmarks in PDF documents
3. Ensuring correct tab and reading order in PDF documents
4. Hiding decorative images with the Artifact tag in PDF documents
5. Indicating required form controls in PDF forms
6. Using table elements for table markup in PDF Documents
7. Performing OCR on a scanned PDF document to provide actual text
8. Providing definitions for abbreviations via an E entry for a structure element
9. Providing headings by marking content with heading tags in PDF documents
10. Providing labels for interactive form controls in PDF documents
11. Providing links and link text using the Link annotation and the /Link structure element in PDF documents
12. Providing name, role, value information for form fields in PDF documents
13. Providing replacement text using the /Alt entry for links in PDF documents
14. Providing running headers and footers in PDF documents
15. Providing submit buttons with the submit-form action in PDF forms
16. Setting the default language using the /Lang entry in the document catalog of a PDF document
17. Specifying consistent page numbering for PDF documents
18. Specifying the document title using the Title entry in the document information dictionary of a PDF document
19. Specifying the language for a passage or phrase with the Lang entry in PDF documents
20. Using Adobe Acrobat Pro's Table Editor to repair mistagged tables
21. Using List tags for lists in PDF documents
22. Indicating when user input falls outside the required format or values in PDF forms

## 23. Providing interactive form controls in PDF documents

CITE: <https://www.w3.org/WAI/WCAG22/Techniques/#pdf>

## Slide 21 - PDF Guidelines: WCAG *and* PDF/UA

### PDF Guidelines: WCAG *and* PDF/UA

- “PDF Techniques for WCAG 2.0” (W3C WAI):  
<https://www.w3.org/WAI/GL/WCAG20-TECHS/pdf.html>
- “Accessibility Support Documentation for PDF” (W3C WAI) This web page on the W3C web site includes "Support Information for WCAG 2.0 Checkpoints": [https://www.w3.org/WAI/GL/WCAG20/implementation-report/PDF\\_accessibility\\_support\\_statement](https://www.w3.org/WAI/GL/WCAG20/implementation-report/PDF_accessibility_support_statement)
- “WCAG Mapping to PDF/UA” (from Adobe):  
<https://helpx.adobe.com/acrobat/using/create-verify-pdf-accessibility.html#WCAGmappingtoPDFUA>

Bouton Jones

2

- “PDF Techniques for WCAG 2.0” (W3C WAI): <https://www.w3.org/WAI/GL/WCAG20-TECHS/pdf.html>
- “Accessibility Support Documentation for PDF” (W3C WAI) This web page on the W3C web site includes "Support Information for WCAG 2.0 Checkpoints":  
[https://www.w3.org/WAI/GL/WCAG20/implementation-report/PDF\\_accessibility\\_support\\_statement](https://www.w3.org/WAI/GL/WCAG20/implementation-report/PDF_accessibility_support_statement)
- “WCAG Mapping to PDF/UA” (from Adobe):  
<https://helpx.adobe.com/acrobat/using/create-verify-pdf-accessibility.html#WCAGmappingtoPDFUA>



## Slide 22 - “WCAG Mapping to PDF/UA” (from Adobe)

| WCAG mapping to PDF/UA                             |                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                             |
|----------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WCAG 2.0                                           | ISO 14289-1 (file)                                                                                                                                                                                                                                                                                                                                                        | Techniques                                                                                                                                                                                                                  |
| 11.1 Non-text Content. (A)                         | <ul style="list-style-type: none"><li>• 7.3 addresses content requiring text alteration.</li><li>• 7.3.1 paragraph four addresses control descriptions.</li><li>• 7.3.6.2 addresses time-based media alternatives. Text, Sensory, and CAPTCHA use-cases are addressed via the technical means used.</li><li>• 7.1 paragraph 1, sentence 2 addresses decoration.</li></ul> | <ul style="list-style-type: none"><li>• PDF1</li><li>• PDF4</li></ul>                                                                                                                                                       |
| 1.2.1 Audio-only and Video-only (Prerecorded). (A) | <ul style="list-style-type: none"><li>• 7.3.6.2 addresses time-based media alternatives. Design-specific. It's necessary that authors and developers consider this provision and ensure conformance.</li></ul>                                                                                                                                                            | <ul style="list-style-type: none"><li>• General Techniques:<br/><a href="http://www.w3.org/WAI/WCAG20/quickref/#eq-media-equiv-av-only-alt">http://www.w3.org/WAI/WCAG20/quickref/#eq-media-equiv-av-only-alt</a></li></ul> |

URL: <https://helpx.adobe.com/acrobat/using/create-verify-pdf-accessibility.html#WCAGmappingtoPDFUA>

Bouton Jones

1

**IMAGE:** Screen Capture of a detail of the mapping on a web page on Adobe.com.

**URL:** <https://helpx.adobe.com/acrobat/using/create-verify-pdf-accessibility.html#WCAGmappingtoPDFUA>

## Slide 23 - Title Slide: PDF/UA and the Matterhorn Protocol

# PDF/UA and the Matterhorn Protocol

## Slide 24 – User Guide: “PDF/UA in a Nutshell”



**PDF/UA in a Nutshell**  
Accessible documents with PDF

Olaf Drümmer and Bettina Chang

- The ISO standard for universal accessibility
- Laws and regulations
- The history of PDF/UA
- File format requirements
- Compliant PDF programs
- Assistive technology
- PDF/UA creation tools
- Validating PDF/UA compliance
- What the experts say

PDF association

*PDF/UA in a Nutshell*  
By Olaf Drümmer and Bettina Chang  
Published by the PDF Association  
August 1, 2013  
21 pages  
Free  
<https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

“Rather than applying to the PDF file format alone, these clear specifications also define compliant assistive technology and PDF reading software.”

Bouton Jones 3

### Document:

#### *PDF/UA in a Nutshell*

By Olaf Drümmer and Bettina Chang

Published by the PDF Association

August 1, 2013

21 pages

Free

<https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

ISO standard 14289-1, known as PDF/UA, was first published in the summer of 2012. It is the first complete definition of a set of requirements for universally accessible PDF documents. Rather than applying to the PDF file format alone, these clear specifications also define compliant assistive technology and PDF reading software.

In signing the UN Convention on the Rights of Disabled Persons, many countries have chosen to work towards building a more inclusive society. They are now required to turn these inclusive principles into a reality on the national level and to remove existing barriers to social inclusion. One key area of activity here is in access to information and communications in the form of documents a subject which PDF/UA now stands to greatly simplify.

PDF/UA in a Nutshell will show you just what PDF/UA is. This brochure provides a clear summary of the key facts, the standard s requirements, the latest developments and the current legal situation regarding PDF accessibility, all in one place. In addition, it suggests

potential applications for the standard and lists the tools that can be used to create and validate PDF/UA-compliant documents.

CITE: <https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

## Slide 25 - Is “PDF/UA in a Nutshell” current?

### Is “PDF/UA in a Nutshell” current?

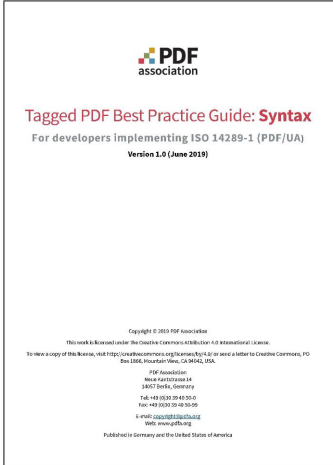
This document references PDF/UA-1.

An updated version of "PDF/UA in a Nutshell" is currently in development by the PDF Association to cover the new PDF/UA-2 (ISO 14289-2:2024) standard.

This document references PDF/UA-1.

An updated version of "PDF/UA in a Nutshell" is currently in development by the PDF Association to cover the new PDF/UA-2 (ISO 14289-2:2024) standard.

## Slide 26 – User Guide: “Tagged PDF Best Practice Guide: Syntax”



**Tagged PDF Best Practice Guide: Syntax**  
For developers implementing ISO 14289-1 (PDF/UA)  
Version 1.0 (June 2019)

Copyright © 2019 PDF Association  
This work is licensed under the Creative Commons Attribution 4.0 International License.  
To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1088, Mountain View, CA 94039, USA.

PDF Association  
Rheinstraße 14  
10627 Berlin, Germany  
Tel: +49 30 20 40 50 0  
Fax: +49 30 20 40 50 99  
E-Mail: [info@pdfa.org](mailto:info@pdfa.org)  
Web: [www.pdfa.org](http://www.pdfa.org)

Published in Germany and the United States of America

*Tagged PDF Best Practice Guide: Syntax*

"A publication of the PDF/UA Technical Working Group, this document is intended for developers implementing tagged PDF and PDF/UA. Others (including authors with some technical knowledge of PDF's accessibility mechanisms) may also benefit from this document.

"This guide is also intended to be useful for those performing detailed accessibility testing on PDF documents claiming conformance with PDF/UA, or on PDF documents claiming to be accessible according to some other specification."

URL: <https://www.pdfa.org/resource/tagged-pdf-best-practice-guide-syntax/>

Bouton Jones 2

### *Tagged PDF Best Practice Guide: Syntax*

A publication of the [PDF/UA Technical Working Group](https://www.pdfa.org/resource/tagged-pdf-best-practice-guide-syntax/), this document is intended for developers implementing tagged PDF and PDF/UA. Others (including authors with some technical knowledge of PDF's accessibility mechanisms) may also benefit from this document.

This guide is also intended to be useful for those performing detailed accessibility testing on PDF documents claiming conformance with PDF/UA, or on PDF documents claiming to be accessible according to some other specification.

Keep in mind that as of today most PDF documents do not follow these recommendations.

**URL:** <https://www.pdfa.org/resource/tagged-pdf-best-practice-guide-syntax/>

## Slide 27 - The PDF/UA is based on WCAG

**Introduction to the Web Content Accessibility Guidelines 2.0**

The requirements for PDF/UA are based on the principles of Version 2.0 of the W3C consortium's Web Content Accessibility Guidelines (WCAG), also published in identical form as an ISO standard in October 2012 (ISO/IEC 40500). WCAG 2.0 guides the creation of accessible web content, focusing on the kinds of content and applications usually provided on websites.

Some WCAG 2.0 tips of a purely informative nature suggest how barriers to access can also be avoided in PDF documents, but a systematic explanation or guide is not provided. PDF/UA, on the other hand, exclusively and fully addresses the PDF format, bringing the ideas behind WCAG 2.0 into the world of PDF and setting PDF-specific rules.

CITE: <https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

Bouton Jones 1

Blockquote from the PDF Org website:

“The requirements for PDF/UA are based on the principles of Version 2.0 of the W3C consortium’s Web Content Accessibility Guidelines (WCAG), also published in identical form as an ISO standard in October 2012 (ISO/IEC 40500). WCAG 2.0 guides the creation of accessible web content, focusing on the kinds of content and applications usually provided on websites.”

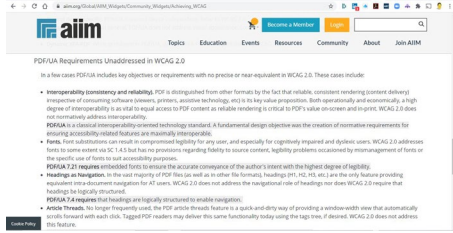
“Some WCAG 2.0 tips of a purely informative nature suggest how barriers to access can also be avoided in PDF documents, but a systematic explanation or guide is not provided. PDF/UA, on the other hand, exclusively and fully addresses the PDF format, bringing the ideas behind WCAG 2.0 into the world of PDF and setting PDF-specific rules.”

CITE: <https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

## Slide 28 - PDF Guidelines: WCAG Only

### PDF Guidelines: PDF/UA Only

“PDF/UA Requirements Unaddressed in WCAG 2.0” (AIIM):  
[https://www.aiim.org/Global/AIIM\\_Widgets/Community\\_Widgets/Achieving\\_WCAG](https://www.aiim.org/Global/AIIM_Widgets/Community_Widgets/Achieving_WCAG)



The screenshot shows a web browser displaying the AIIM website. The page title is 'PDF/UA Requirements Unaddressed in WCAG 2.0'. The content includes a list of key objectives or requirements that PDF/UA includes but WCAG 2.0 does not. The list includes: Interoperability (consistency and reliability), Fonts, and Headings as Navigation. The page is authored by Bouton Jones.

Bouton Jones

2

“PDF/UA Requirements Unaddressed in WCAG 2.0” (AIIM):

[https://www.aiim.org/Global/AIIM\\_Widgets/Community\\_Widgets/Achieving\\_WCAG](https://www.aiim.org/Global/AIIM_Widgets/Community_Widgets/Achieving_WCAG)

--- Begin Blockquote ---

In a few cases PDF/UA includes key objectives or requirements with no precise or near-equivalent in WCAG 2.0. These cases include:

- **Interoperability (consistency and reliability).** PDF is distinguished from other formats by the fact that reliable, consistent rendering (content delivery) irrespective of consuming software (viewers, printers, assistive technology, etc.) is its key value proposition. Both operationally and economically, a high degree of interoperability is as vital to equal access to PDF content as reliable rendering is critical to PDF's value on-screen and in-print. WCAG 2.0 does not normatively address interoperability. PDF/UA is a classical interoperability-oriented technology standard. A fundamental design objective was the creation of normative requirements for ensuring accessibility-related features are maximally interoperable.
- **Fonts.** Font substitutions can result in compromised legibility for any user, and especially for cognitively impaired and dyslexic users. WCAG 2.0 addresses fonts to some extent via SC 1.4.5 but has no provisions regarding fidelity to source content,



legibility problems occasioned by mismanagement of fonts or the specific use of fonts to suit accessibility purposes.

- PDF/UA 7.21 requires embedded fonts to ensure the accurate conveyance of the author's intent with the highest degree of legibility.
- Headings as Navigation. In the vast majority of PDF files (as well as in other file formats), headings (H1, H2, H3, etc.) are the only feature providing equivalent intra-document navigation for AT users. WCAG 2.0 does not address the navigational role of headings nor does WCAG 2.0 require that headings be logically structured.
- PDF/UA 7.4 requires that headings are logically structured to enable navigation.
- Article Threads. No longer frequently used, the PDF article threads feature is a quick-and-dirty way of providing a window-width view that automatically scrolls forward with each click. Tagged PDF readers may deliver this same functionality today using the tags tree, if desired. WCAG 2.0 does not address this feature.
- PDF/UA 7.12 requires that if present, article threads shall follow the logical structure (tags) of the document.

--- End Blockquote ---

## Slide 29 - PDF/UA Technical Requirements

### PDF/UA Technical Requirements

- All real (meaningful) content shall be tagged. Artifacts (non-meaningful content) shall not be tagged.
- The “tag tree” created by the individual tags (representing the document’s content) should be sorted to reflect the document’s logical reading order.
- All structure types in use shall be standard or mapped to standard
- A **manual check** of the document is required to ensure that there is no information conveyed using visual means alone (e.g., contrast, color etc.), and that there is no flickering, blinking, or flashing content.
- A document title must be given in the metadata. In the Document Properties, the Initial View settings must be set up so that the title (rather than the file name) appears in the window title (under Window Options).
- Any non-text elements (e.g., images, charts, graphs, infographics, logos, etc.) must have corresponding alternative text.

CITE: <https://accessible-docs.com/overview-pdf-ua-standard/>

Bouton Jones

3

#### “WCAG 2.0 compliance

“The W3C’s Web Content Accessibility Guidelines 2.0 (WCAG 2.0) was written with Web content in mind. The application of WCAG 2.0 to other electronic content requires interpretation to determine how to apply the WCAG 2.0 Principles, Guidelines and Success Criteria. There is not always a 1:1 mapping.

“In the case of PDF documents, the international standard (ISO14289-1:2014) defines requirements specifically for accessible PDF files. The requirements for the PDF/UA standard are complementary to the WCAG 2.0 success criteria, and are therefore compliant with its requirements. This means that a PDF file which is compliant to the PDF/UA standard will also be compliant to ADA, Section 508 and AODA requirements.

#### “PDF/UA technical requirements

“The following are just a few of the technical requirements as set out by the PDF/UA standard:”

- All real (meaningful) content shall be tagged. Artifacts (non-meaningful content) shall not be tagged.
- The “tag tree” created by the individual tags (representing the document’s content) should be sorted to reflect the document’s logical reading order.
- All structure types in use shall be standard or mapped to standard

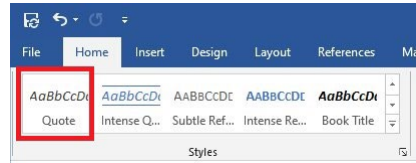
- A manual check of the document is required to ensure that there is no information conveyed using visual means alone (e.g. contrast, color etc.), and that there is no flickering, blinking or flashing content.
- A document title must be given in the metadata. In the Document Properties, the Initial View settings must be set up so that the title (rather than the file name) appears in the window title (under Window Options).
- Any non-text elements (images, charts, graphs, infographics, logos, etc.) must have corresponding alternative text.

**CITE:** <https://accessible-docs.com/overview-pdf-ua-standard/>

## Slide 30 - PDF/UA Beyond the WCAG (Slide 1 of 2)

### PDF/UA-1 *Beyond* the WCAG (Slide 1 of 2)

- Document Title
- Quotes and Blockquotes
- Embed Fonts
- Article Threads
- Captions
- Don't use background images or watermarks
- Don't use flickering, blinking or flashing content
- Don't use Dynamic XFA forms



Bouton Jones

4

- Document Title
- Quotes and Blockquotes
- Embedded Fonts
- Article Threads
- Captions
- Don't use background images or watermarks
- Don't use Dynamic XFA forms. (If you don't know what they are don't worry about it.)

## Slide 31 - PDF/UA Beyond the WCAG (Slide 2 of 2)

### PDF/UA-1 *Beyond* the WCAG (slide 2 of 2)

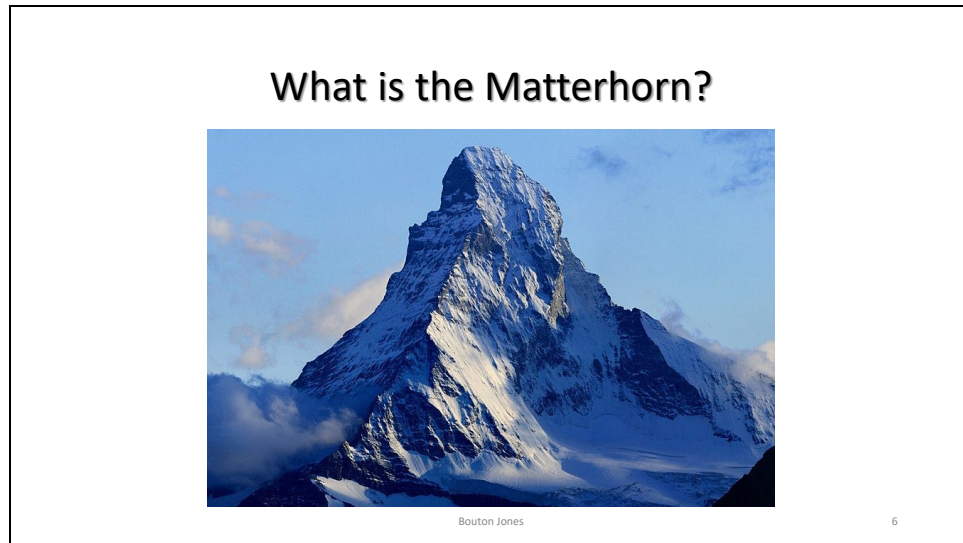
- Use *standard structure types* --- or map to standard structure types with a role map in the structure tree root.
- Tag any information --- that was conveyed by visual means alone (contrast, color, format, or layout) --- to reflect the intended meaning.
- Headings
  - Only use Heading levels 1 through 6
  - You *can* have more than one Heading level 1
  - Do not separate a Heading into two parts

Bouton Jones

5

- Use standard structure types or map to standard structure types, with a role map in the structure tree root.
- Any Information conveyed by contrast, color, format or layout must also be tagged to reflect the intended meaning.
- Headings
  - You can have more than 1 Heading level 1
  - Only use heading levels H1 – H6. Unlike HTML, the PDF format doesn't include additional heading levels (e.g. H7.) This is not an issue in Word
  - Do not separate a Heading into two parts

## Slide 32 - Chat Poll: What is the Matterhorn?



**IMAGE:** The Matterhorn mountain.

It is a mountain of the Alps with a summit 14,692 feet (or 4,478 metres) high, making it one of the highest summits in Europe.

**FROM Wikipedia:** “The first ascent of the Matterhorn was in 1865 from Zermatt by a party led by Edward Whymper; it ended when four of its seven members fell to their deaths on the descent. This disaster, later portrayed in several films, marked the end of the golden age of alpinism. .... It is estimated that over 500 alpinists have died on the Matterhorn, making it one of the deadliest peaks in the world.”

## Slide 33 - What the word “Matterhorn” means to me.

### What the word “Matterhorn” means to me:

- Massive
- Monolithic
- Insurmountable
- Dangerous
- Unavoidable
- One of the highest summits in Europe
- “One of the deadliest peaks in the world.” (From Wikipedia)

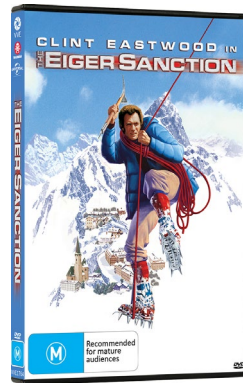
Bouton Jones

7

- Massive
- Monolith
- Insurmountable
- One of the highest summits in Europe
- “It is estimated that over 500 alpinists have died on the Matterhorn, making it one of the deadliest peaks in the world.” (From Wikipedia)

## Slide 34 - The Title “Matterhorn Protocol”

### The Title “Matterhorn Protocol”



Bouton Jones

8

It reminds me of the book and the movie *The Eiger Sanction*.

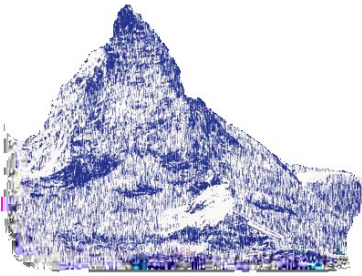
It stars Clint Eastwood as Dr. Jonathan Hemlock, a Classical art professor and collector who doubles as a professional assassin. He is coerced out of retirement to avenge the murder of an old friend. The film is full of action, danger, deception, and betrayal.



## Slide 35 - What is the Matterhorn Protocol?

### What is the Matterhorn Protocol?

Is it another set of **massive** and **insurmountable** A11y guidelines?



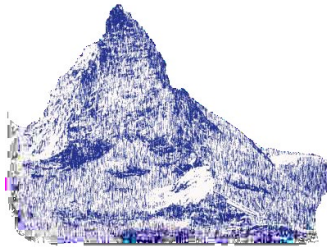
Bouton Jones 1

Is it another set of massive, insurmountable, deadly A11y guidelines to learn?

## Slide 36 - Matterhorn Protocol defined

### **Matterhorn Protocol:**

A list of all the possible ways to fail PDF/UA produced by the PDF Association's PDF/UA Competence Center. It consists of 31 Checkpoints comprised of 136 Failure Conditions.



*It's based on the PDF/UA.*

Bouton Jones

2

Matterhorn Protocol: a list of all the possible ways to fail PDF/UA produced by the PDF Association's PDF/UA Competence Center. It consists of 31 Checkpoints comprised of 136 Failure Conditions.

## Slide 37 - Download the Matterhorn Protocol PDF



**Matterhorn Protocol**  
PDF Association, PDF/A Competence Center | Document Version 1.02 | April 30, 2014

|                                                    |   |
|----------------------------------------------------|---|
| Document History                                   | 2 |
| Background                                         | 2 |
| About PDF/A                                        | 2 |
| Introduction to the Matterhorn Protocol            | 2 |
| Matterhorn Protocol Characteristics                | 2 |
| Key to the Readers in the Future Conditions Tables | 3 |
| Additional Reading                                 | 3 |
| PDF/A-1 Full-Form Conditions                       | 4 |
| CharAgreement 01: Best content support             | 4 |
| CharAgreement 02: Role Mapping                     | 5 |
| CharAgreement 03: Role Mapping                     | 5 |
| CharAgreement 04: Color and Content                | 5 |
| CharAgreement 05: Content                          | 5 |
| CharAgreement 06: Metadata                         | 5 |
| CharAgreement 07: Dictionary                       | 6 |
| CharAgreement 08: OCR Validation                   | 6 |
| CharAgreement 09: Appearance Table                 | 6 |
| CharAgreement 10: Character Mapping                | 6 |
| CharAgreement 11: Document Metadata                | 6 |
| CharAgreement 12: Document Characteristics         | 7 |
| CharAgreement 13: Document                         | 7 |
| CharAgreement 14: Document                         | 7 |
| CharAgreement 15: Content                          | 7 |
| CharAgreement 16: Content                          | 7 |
| CharAgreement 17: Document                         | 7 |
| CharAgreement 18: Document                         | 7 |
| CharAgreement 19: Document                         | 7 |
| CharAgreement 20: Document                         | 7 |
| CharAgreement 21: Document                         | 7 |
| CharAgreement 22: Document                         | 7 |
| CharAgreement 23: Document                         | 7 |
| CharAgreement 24: Document                         | 7 |
| CharAgreement 25: Document                         | 7 |
| CharAgreement 26: Document                         | 7 |
| CharAgreement 27: Document                         | 7 |
| CharAgreement 28: Document                         | 7 |
| CharAgreement 29: Document                         | 7 |
| CharAgreement 30: Document                         | 7 |
| CharAgreement 31: Document                         | 7 |
| CharAgreement 32: Document                         | 7 |
| CharAgreement 33: Document                         | 7 |
| CharAgreement 34: Document                         | 7 |
| CharAgreement 35: Document                         | 7 |
| CharAgreement 36: Document                         | 7 |
| CharAgreement 37: Document                         | 7 |
| CharAgreement 38: Document                         | 7 |
| CharAgreement 39: Document                         | 7 |
| CharAgreement 40: Document                         | 7 |
| CharAgreement 41: Document                         | 7 |
| CharAgreement 42: Document                         | 7 |
| CharAgreement 43: Document                         | 7 |
| CharAgreement 44: Document                         | 7 |
| CharAgreement 45: Document                         | 7 |
| CharAgreement 46: Document                         | 7 |
| CharAgreement 47: Document                         | 7 |
| CharAgreement 48: Document                         | 7 |
| CharAgreement 49: Document                         | 7 |
| CharAgreement 50: Document                         | 7 |
| CharAgreement 51: Document                         | 7 |
| CharAgreement 52: Document                         | 7 |
| CharAgreement 53: Document                         | 7 |
| CharAgreement 54: Document                         | 7 |
| CharAgreement 55: Document                         | 7 |
| CharAgreement 56: Document                         | 7 |
| CharAgreement 57: Document                         | 7 |
| CharAgreement 58: Document                         | 7 |
| CharAgreement 59: Document                         | 7 |
| CharAgreement 60: Document                         | 7 |
| CharAgreement 61: Document                         | 7 |
| CharAgreement 62: Document                         | 7 |
| CharAgreement 63: Document                         | 7 |
| CharAgreement 64: Document                         | 7 |
| CharAgreement 65: Document                         | 7 |
| CharAgreement 66: Document                         | 7 |
| CharAgreement 67: Document                         | 7 |
| CharAgreement 68: Document                         | 7 |
| CharAgreement 69: Document                         | 7 |
| CharAgreement 70: Document                         | 7 |
| CharAgreement 71: Document                         | 7 |
| CharAgreement 72: Document                         | 7 |
| CharAgreement 73: Document                         | 7 |
| CharAgreement 74: Document                         | 7 |
| CharAgreement 75: Document                         | 7 |
| CharAgreement 76: Document                         | 7 |
| CharAgreement 77: Document                         | 7 |
| CharAgreement 78: Document                         | 7 |
| CharAgreement 79: Document                         | 7 |
| CharAgreement 80: Document                         | 7 |
| CharAgreement 81: Document                         | 7 |
| CharAgreement 82: Document                         | 7 |
| CharAgreement 83: Document                         | 7 |
| CharAgreement 84: Document                         | 7 |
| CharAgreement 85: Document                         | 7 |
| CharAgreement 86: Document                         | 7 |
| CharAgreement 87: Document                         | 7 |
| CharAgreement 88: Document                         | 7 |
| CharAgreement 89: Document                         | 7 |
| CharAgreement 90: Document                         | 7 |
| CharAgreement 91: Document                         | 7 |
| CharAgreement 92: Document                         | 7 |
| CharAgreement 93: Document                         | 7 |
| CharAgreement 94: Document                         | 7 |
| CharAgreement 95: Document                         | 7 |
| CharAgreement 96: Document                         | 7 |
| CharAgreement 97: Document                         | 7 |
| CharAgreement 98: Document                         | 7 |
| CharAgreement 99: Document                         | 7 |
| CharAgreement 100: Document                        | 7 |

The PDF Association | PDF/A Competence Center | PDF/A-1 Matterhorn Protocol 1.02 | 1

Document Version 1.02 | April 30, 2014 | 13 pages

URL: [https://www.pdfa.org/wp-content/uploads/2014/06/MatterhornProtocol\\_1-02.pdf](https://www.pdfa.org/wp-content/uploads/2014/06/MatterhornProtocol_1-02.pdf)

Bouton Jones

3

Document Version 1.02 | April 30, 2014 | 13 pages

URL: [https://www.pdfa.org/wp-content/uploads/2014/06/MatterhornProtocol\\_1-02.pdf](https://www.pdfa.org/wp-content/uploads/2014/06/MatterhornProtocol_1-02.pdf)

## Slide 38 - Is the Matterhorn Protocol current?

### Is the Matterhorn Protocol current?

- Matterhorn Protocol 1.02 is still valid, but **it applies only to PDF/UA-1**, not PDF/UA-2.
  - The Matterhorn Protocol 1.02 lists 31 checkpoints and 136 failure conditions for PDF/UA-1.
- A direct, official "Matterhorn Protocol for PDF/UA-2" does *not* currently exist.
- The original Matterhorn Protocol remains an extremely useful checklist for PDF Accessibility.

Bouton Jones

4

- Matterhorn Protocol 1.02 is still valid, but **it applies only to PDF/UA-1**, not PDF/UA-2.
  - The Matterhorn Protocol 1.02 lists 31 checkpoints and 136 failure conditions for PDF/UA-1.
- A direct, official "Matterhorn Protocol for PDF/UA-2" does not currently exist.
- The original Matterhorn Protocol remains an extremely useful checklist for PDF Accessibility.

## Slide 39 - Matterhorn Protocol: PDF/UA-1 Checkpoints

### Matterhorn Protocol: PDF/UA-1 Checkpoints

- |                        |                              |                           |
|------------------------|------------------------------|---------------------------|
| 1. Real content tagged | Language                     | 21. Embedded Files        |
| 2. Role Mapping        | 12. Stretchable Characters   | 22. Article Threads       |
| 3. Flickering          | 13. Graphics                 | 23. Digital Signatures    |
| 4. Color and Contrast  | 14. Headings                 | 24. Non-Interactive Forms |
| 5. Sound               | 15. Tables                   | 25. XFA                   |
| 6. Metadata            | 16. Lists                    | 26. Security              |
| 7. Dictionary          | 17. Mathematical Expressions | 27. Navigation            |
| 8. OCR Validation      | 18. Page Headers and Footers | 28. Annotations           |
| 9. Appropriate Tags    | 19. Notes and References     | 29. Actions               |
| 10. Character Mappings | 20. Optional Content         | 30. XObjects              |
| 11. Declared Natural   |                              | 31. Fonts                 |

Bouton Jones

5

- |                               |                 |
|-------------------------------|-----------------|
| 1. Real content tagged        | 28. Annotations |
| 2. Role Mapping               | 29. Actions     |
| 3. Flickering                 | 30. XObjects    |
| 4. Color and Contrast         | 31. Font        |
| 5. Sound                      |                 |
| 6. Metadata                   |                 |
| 7. Dictionary                 |                 |
| 8. OCR Validation             |                 |
| 9. Appropriate Tags           |                 |
| 10. Character Mappings        |                 |
| 11. Declared Natural Language |                 |
| 12. Stretchable Characters    |                 |
| 13. Graphics                  |                 |
| 14. Headings                  |                 |
| 15. Tables                    |                 |
| 16. Lists                     |                 |
| 17. Mathematical Expressions  |                 |
| 18. Page Headers and Footers  |                 |
| 19. Notes and References      |                 |
| 20. Optional Content          |                 |
| 21. Embedded Files            |                 |
| 22. Article Threads           |                 |
| 23. Digital Signatures        |                 |
| 24. Non-Interactive Forms     |                 |
| 25. XFA                       |                 |
| 26. Security                  |                 |
| 27. Navigation                |                 |

## Slide 40 - Checkpoint 01: Real content tagged

### Checkpoint 01: Real content tagged

Use *standard structure types* --- or map to standard structure types with a role map in the structure tree root.

**Failure Condition 1-006** : The *structure type* and attributes of a structure element are not semantically appropriate for the structure element.

**48 PDF Structure Types:** Document, Part, Art, Sect, Div, BlockQuote, Caption, TOC, TOCI, Index, NonStruct, Private, P, H, H1, H2, H3, H4, H5, H6, L, LI, Lbl, LBody, Table, TR, TH, TD, THead, TBody, TFoot, Span, Quote, Note, Reference, BibEntry, Code, Link, Annot, Ruby, Warichu, RB, RT, RP, WT, WP, Figure, Formula, and Form.

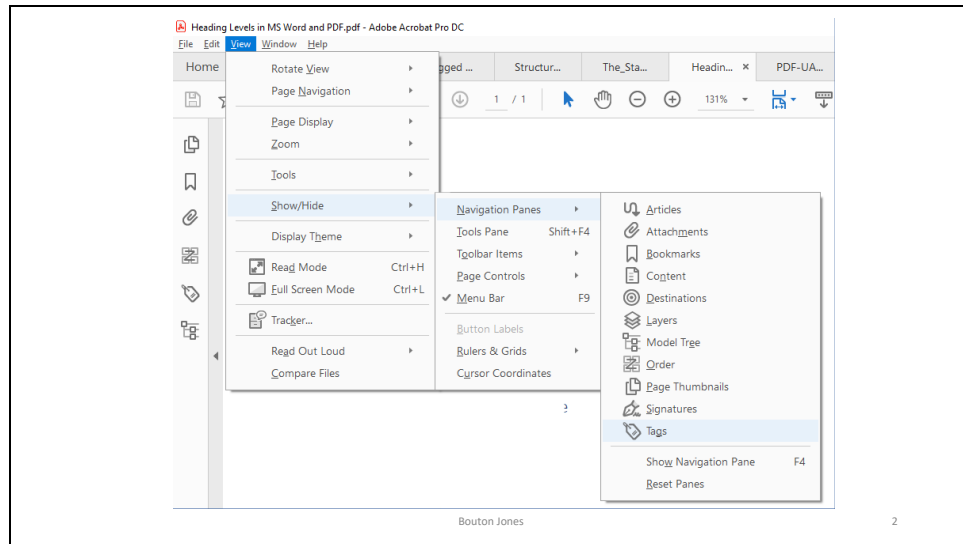
Bouton Jones

1

Failure Condition 1-006 : The structure type and attributes of a structure element are not semantically appropriate for the structure element.

**48 PDF Structure Types:** Document, Part, Art, Sect, Div, BlockQuote, Caption, TOC, TOCI, Index, NonStruct, Private, P, H, H1, H2, H3, H4, H5, H6, L, LI, Lbl, LBody, Table, TR, TH, TD, THead, TBody, TFoot, Span, Quote, Note, Reference, BibEntry, Code, Link, Annot, Ruby, Warichu, RB, RT, RP, WT, WP, Figure, Formula, and Form.

## Slide 41 - Navigation to the Tag Panel in Acrobat



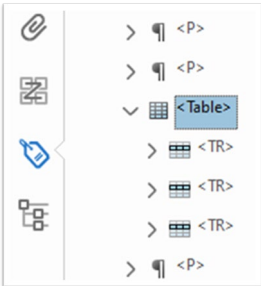
**IMAGE:** Screen capture showing the following navigation route.

- Select the View drop down menu
- Select the Show/Hide option
- Select the Navigation Panes option
- Select the Tags option

The Tags pane will appear on the right-hand side of the Acrobat window.

## Slide 42 – Table Structure Types

**Table Structure Types**



| HTML Tags | PDF Structure Types |
|-----------|---------------------|
| <table>   | table               |
| <tr>      | tr                  |
| <th>      | th                  |
| <td>      | td                  |
| <thead>   | thead               |
| <tbody>   | tbody               |
| <tfoot>   | tfoot               |

THEAD, TBODY, and TFOOT help with pagination.  
Adobe InDesign provides TFOOT support.

Bouton Jones3

**IMAGE:** Screen capture of a few table tags in Acrobat’s Tag pane.

Two Column Table with a column header for “HTML Tags” and “PDF Structure Types”. There are 7 rows and the contents for both cells in each row are identical. The repeated contents of the rows are: Table, TR, TH, TD, THead, TBody, Tfoot.

So, there is a one-to-one correspondence for table tags in HTML and PDF Structure Types for tables.

THEAD, TBODY, and TFOOT help with pagination.

Adobe InDesign provides TFOOT support.



## Slide 43 – Label and List Structure Types

### Label and List Structure Types

| PDF Structure Types                                                                                                                       | HTML Tags   |
|-------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>L</b> - the List tag, which contains one or more LI tags.                                                                              | <UL> & <OL> |
| <b>LI</b> - the List Item tag. List item tags can contain Lbl and LBody tags.                                                             | <LI>        |
| <b>Lbl</b> - the list item label. Contains distinguishing information such as a item number or bullet character.                          | *           |
| <b>LBody</b> - the list item body. Contains list item content, or in the case of a nested list, it may contain additional List tag trees. |             |

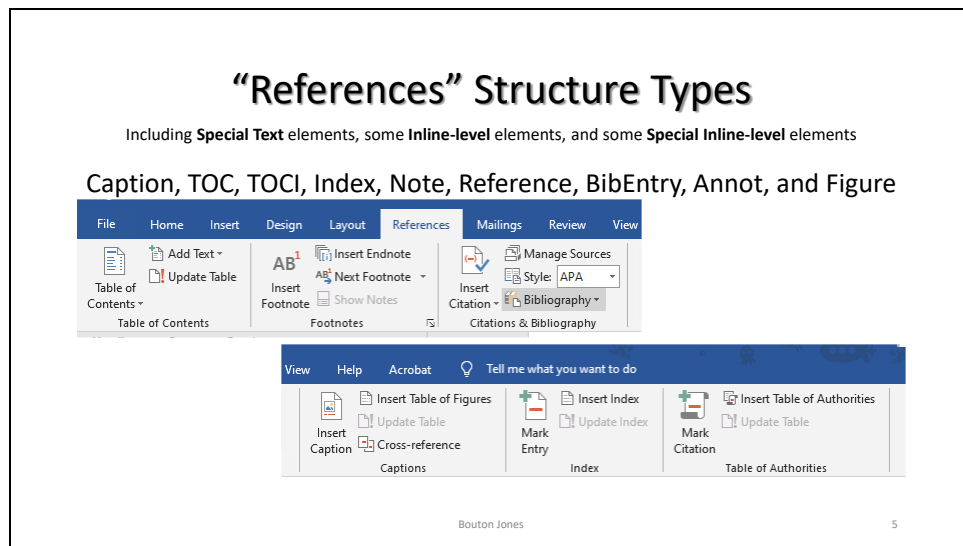
**\*NOTE:** HTML's <label> tag is only used in *forms*. It doesn't apply to lists.

Bouton Jones

4

- L - the List tag, which contains one or more LI tags.
- LI - the List Item tag. List item tags can contain Lbl and LBody tags.
- Lbl - the list item label. Contains distinguishing information such as a item number or bullet character.
- LBody - the list item body. Contains list item content, or in the case of a nested list, it may contain additional List tag trees.

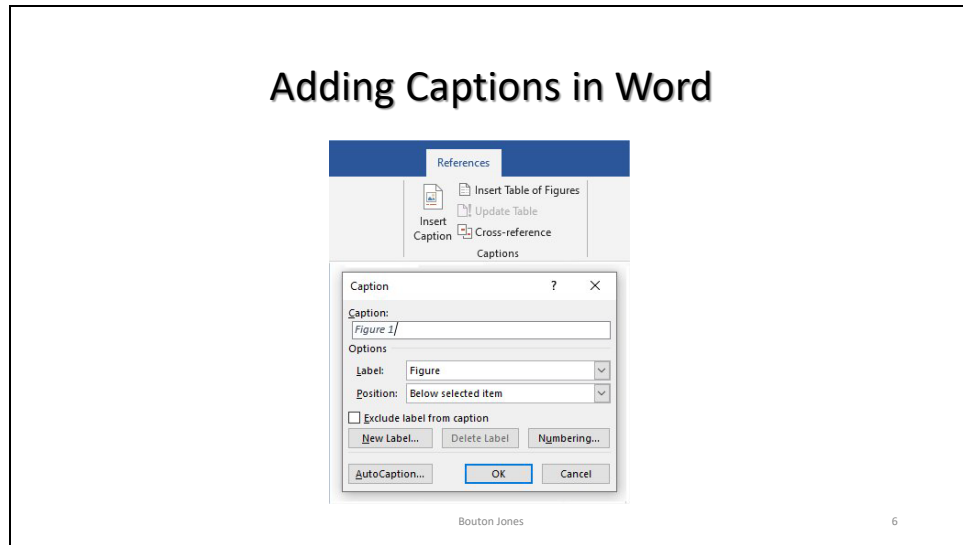
## Slide 44 – “References” Structure Types



- Caption
- TOC
- TOCI
- Index
- Note
- Reference
- BibEntry
- Annot
- Figure

**IMAGE:** Screen capture of Microsoft Word ribbon with the References tab selected. It shows the Word options that are equivalent to many of the PDF Structure Type (e.g. Table of Contents, Insert Footnote, Insert Citation, Bibliography, Insert Caption, Insert Index, etc.,)

## Slide 45 - Adding Captions in Word



**IMAGE:** Screen capture of a detail of the References tab in Word. A captions window below

## Slide 46 – Header and Paragraph Structure Types

### Header and Paragraph Structure Types

- H (heading, strongly structured)\*
- H1
- H2
- H3
- H4
- H5
- H6
- P

*\*"Due to a lack of suitable tools, the H structure element is impractical, and its use is not recommended." (CITE: Tagged PDF Best Practice Guide: Syntax from the PDF Association)*

Bouton Jones

7

- H (heading, strongly structured)\*
- H1
- H2
- H3
- H4
- H5
- H6
- P

*\*Due to a lack of suitable tools, the H structure element is impractical, and its use is not recommended." (CITE: Tagged PDF Best Practice Guide: Syntax from the PDF Association)*

## Slide 47 - Structure Types for Japanese Typography

### Japanese Typography Structure Types



- RB
- RP
- RT
- Ruby (*furigana*)
- Warichu
- WP
- WT



Image Source:  
<https://www.pinterest.com/pin/500884789788049202/>

Image Source:  
<https://www.w3.org/international/questions/qa-ruby-data/double-sided.png>

Bouton Jones

8

- RB
- RP
- RT
- Ruby
- Warichu
- WP
- WT

Warichu (literally “split annotation”): the Japanese word for the traditional East Asian typographic device of typesetting in small double lines editorial comments, notes, parenthetical comments, and other annotations that may or may not belong to the text proper.

## Slide 48 – NonStruct

### NonStruct

From the PDF Association's *Tagged PDF Best Practice Guide*:

Has no substantive role or meaning; interpretation is out of scope for consumers of tagged PDF.  
NonStruct's value is primarily as a utility for role mapping custom structure element types for which no corresponding standard structure element type is suitable.

A document containing custom structure elements for which there is no corresponding standard structure type, but where the content and child elements are intended to be real content. In such a case appropriate semantics require a writer to role map the custom element type to NonStruct.

NonStruct has no semantic significance, but its content and contained structure elements may well have significance. A consuming processor would be expected to pass these children to AT and other processors.

CITE: <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

Bouton Jones 9

From the PDF Association's Tagged PDF Best Practice Guide:

Has no substantive role or meaning; interpretation is out of scope for consumers of tagged PDF. NonStruct's value is primarily as a utility for role mapping custom structure element types for which no corresponding standard structure element type is suitable.

A document containing custom structure elements for which there is no corresponding standard structure type, but where the content and child elements are intended to be real content. In such a case appropriate semantics require a writer to role map the custom element type to NonStruct.

NonStruct has no semantic significance, but its content and contained structure elements may well have significance. A consuming processor would be expected to pass these children to AT and other processors.

**CITE:** <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

## Slide 49 - Private

### Private

From the PDF Association's *Tagged PDF Best Practice Guide*:

Akin to <NonStruct>, this element differs from <NonStruct> insofar as not only the structure element itself, but *the children of the structure element, including structure elements and content, are also ignored*.

CITE: <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

Bouton Jones

10

From the PDF Association's Tagged PDF Best Practice Guide:

Akin to <NonStruct>, this element differs from <NonStruct> insofar as not only the structure element itself, but the children of the structure element, including structure elements and content, are also ignored.

CITE: <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

## Slide 50 - Remaining Structure Types

### Remaining Structure Types

- BlockQuote
- Code
- Div
- P
- Quote
- Span

Bouton Jones

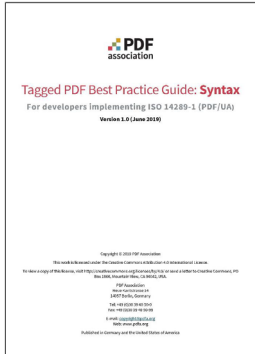
11

- BlockQuote
- Code
- Div
- P
- Quote
- Span



## Slide 51 - For more about Structure Types:

### For more about Structure Types:



See the PDF Association's Tagged PDF *Best Practice Guide* at <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

Bouton Jones 12

See the PDF Association's Tagged PDF Best Practice Guide at <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

**IMAGE:** First Page of the PDF Association's Tagged PDF Best Practice Guide .

## Slide 52 - Checkpoint 06: Metadata

### Checkpoint 06: Metadata

| Index  | Failure Condition                                                                     | Section   |
|--------|---------------------------------------------------------------------------------------|-----------|
| 06-001 | Document does not contain XMP metadata stream                                         | UA1:7.1-8 |
| 06-002 | The metadata stream in the Catalog dictionary does not include the PDF/UA identifier. | UA1:5     |
| 06-003 | Metadata stream does not contain dc:title                                             | UA1:7.1-8 |
| 06-004 | dc:title does not clearly identify the document                                       | UA1:7.1-8 |

We're going to cover this Checkpoint in two parts:

1. 06-001 and 06-002
2. 06-003 and 06-004

Bouton Jones

13

| Index | Failure Condition | Section |
|-------|-------------------|---------|
|-------|-------------------|---------|

|        |                                               |           |
|--------|-----------------------------------------------|-----------|
| 06-001 | Document does not contain XMP metadata stream | UA1:7.1-8 |
|--------|-----------------------------------------------|-----------|

|        |                                                                                       |       |
|--------|---------------------------------------------------------------------------------------|-------|
| 06-002 | The metadata stream in the Catalog dictionary does not include the PDF/UA identifier. | UA1:5 |
|--------|---------------------------------------------------------------------------------------|-------|

|        |                                           |           |
|--------|-------------------------------------------|-----------|
| 06-003 | Metadata stream does not contain dc:title | UA1:7.1-8 |
|--------|-------------------------------------------|-----------|

|        |                                                 |           |
|--------|-------------------------------------------------|-----------|
| 06-004 | dc:title does not clearly identify the document | UA1:7.1-8 |
|--------|-------------------------------------------------|-----------|

\*Dublin Core (DC): a set of fifteen "core" elements (properties) for describing resources. This fifteen-element Dublin Core has been formally standardized as ISO 15836, ANSI/NISO Z39.85,[2] and IETF RFC 5013. The core properties are part of a larger set of DCMI Metadata Terms. "Dublin Core" is also used as an adjective for Dublin Core metadata, a style of metadata that draws on multiple RDF vocabularies, packaged and constrained in Dublin Core application profiles. (CITE: [https://en.wikipedia.org/wiki/Dublin\\_Core](https://en.wikipedia.org/wiki/Dublin_Core) )

## Slide 53 - Checkpoint 06: Metadata

### Checkpoint 06: 001 & 002 XMP Metadata & PDF/UA Identifier

| Index  | Failure Condition                                                                             | Section   |
|--------|-----------------------------------------------------------------------------------------------|-----------|
| 06-001 | Document does not contain <i>XMP metadata</i> stream                                          | UA1:7.1-8 |
| 06-002 | The metadata stream in the Catalog dictionary does not include the <i>PDF/UA identifier</i> . | UA1:5     |
| 06-003 | Metadata stream does not contain dc:title                                                     | UA1:7.1-8 |
| 06-004 | dc:title does not clearly identify the document                                               | UA1:7.1-8 |

Bouton Jones

14

Index    Failure Condition    Section

06-001 Document does not contain XMP metadata stream    UA1:7.1-8

06-002 The metadata stream in the Catalog dictionary does not include the PDF/UA identifier.  
UA1:5

06-003 Metadata stream does not contain dc:title    UA1:7.1-8

06-004 dc:title does not clearly identify the document    UA1:7.1-8

## Slide 54 - Checkpoint 06: Failure Condition 06-001

### Checkpoint 06: 001 XMP



Adding Intelligence to Media

**Failure Condition 06-001:** Document does not contain *XMP metadata stream*.

“The Extensible Metadata Platform (XMP) provides Adobe applications with a common XML framework that standardizes the creation, processing, and interchange of document metadata across publishing workflows. You can save and import the document metadata XML source code in XMP format, making it easy to share metadata among different documents. You can also save document metadata to a metadata template that you can reuse in Acrobat.”

**CITE:** <https://helpx.adobe.com/acrobat/using/pdf-properties-metadata.html>

Bouton Jones15

Failure Condition 06-001: Document does not contain XMP metadata stream

“The Extensible Metadata Platform (XMP) provides Adobe applications with a common XML framework that standardizes the creation, processing, and interchange of document metadata across publishing workflows. You can save and import the document metadata XML source code in XMP format, making it easy to share metadata among different documents. You can also save document metadata to a metadata template that you can reuse in Acrobat.”

**CITE:** <https://helpx.adobe.com/acrobat/using/pdf-properties-metadata.html>

## Slide 55 - Fixes for Missing XMP Data

### Fixes for Missing XMP Data

Instructions to Add Metadata with an XMP File:

- Taggedpdf.com: <https://taggedpdf.com/508-pdf-help-center/pdfua-identifier-missing/>
- Accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>

Bouton Jones

16

Instructions to Add Metadata with an XMP File:

- Taggedpdf.com: <https://taggedpdf.com/508-pdf-help-center/pdfua-identifier-missing/>
- Accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>

## Slide 56 - Checkpoint 06: 002 PDF/UA Identifier

### Checkpoint 06: 002 PDF/UA Identifier

**Failure Condition 06-002:** The metadata stream in the Catalog dictionary does not include the *PDF/UA identifier*.

This is an issue for PDFs created in:

- Microsoft Word 2013 for Windows
- Any version of Adobe InDesign (formerly Adobe PageMaker)

Test with:

- Adobe Acrobat
- PDF Accessibility Checker (PAC) 2026 (free testing tool from Access for all)
- axesPDF tool from axes4

Bouton Jones

17

Failure Condition 06-002: The metadata stream in the Catalog dictionary does not include the PDF/UA identifier.

A missing PDF/UA Identifier can be an issue for PDFs created in:

- Microsoft Word 2013 for Windows
- Any version of Adobe InDesign (formerly Adobe PageMaker)

Test with:

- Adobe Acrobat
- PDF Checker (PAC) 2026 (free testing tool from Access for all)
- axesPDF tool from axes4

## Slide 57 - An example of a complete PDF/UA flag

An example of a complete PDF/UA flag:

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="Adobe XMP Core 5.6-c015
91.163280, 2018/06/22-11:31:03 "
  >
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    <rdf:Description rdf:about=""
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:pdfuaid="http://www.aiim.org/pdfua/ns/id/"
    <dc:title>
      <rdf:Alt>
        <rdf:li xml:lang="x-default">PDF/UA Document</rdf:li>
        <rdf:li xml:lang="en">PDF/UA Document</rdf:li>
      </rdf:Alt>
    </dc:title>
    <pdfuaid:part>1</pdfuaid:part>
  </rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>
```

CITE: <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>

18

IMAGE: XML code for a complete PDF/UA flag

---- BEGIN CODE ---

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="Adobe XMP Core 5.6-c015
91.163280, 2018/06/22-11:31:03 ">
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  <rdf:Description rdf:about=""
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:pdfuaid="http://www.aiim.org/pdfua/ns/id/"
  <dc:title>
  <rdf:Alt>
    <rdf:li xml:lang="x-default">PDF/UA Document</rdf:li>
    <rdf:li xml:lang="en">PDF/UA Document</rdf:li>
  </rdf:Alt>
  </dc:title>
  <pdfuaid:part>1</pdfuaid:part>
  </rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>
```

---- END CODE ---

IMAGE and CODE SOURCE: <https://www.pdfa.org/wp-content/uploads/2019/06/TaggedPDFBestPracticeGuideSyntax.pdf>



## Slide 58 - Why must we include the PDF/UA Identifier?

### Why include the PDF/UA Identifier?

- “This enables reader applications and assistive technologies to fully process PDF/UA conforming files to maximize accessibility.” (CITE: <https://accessiblewebsiteservices.com/accessible-pdfs-how-to-add-a-pdf-ua-identifier/> )
- “In order for a PDF document to be technically accessible, the [PDF/UA](#) identifier (metadata) must be set.” (CITE: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing> )

So, why is that?

Bouton Jones

1

- “This enables reader applications and assistive technologies to fully process PDF/UA conforming files to maximize accessibility.” (CITE: <https://accessiblewebsiteservices.com/accessible-pdfs-how-to-add-a-pdf-ua-identifier/> )
- “In order for a PDF document to be technically accessible, the PDF/UA identifier (metadata) must be set.” (CITE: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing> )

So, why is that?

## Slide 59 - The PDF Association on the PDF/UA Identifier

### The PDF Association on the PDF/UA Identifier

From an email from [info@pdfa.org](mailto:info@pdfa.org):

The PDF/UA flag is a quality indicator. It prompts AT to expect high-quality tagging and other structures required for accessibility, as well as alt. text, validated disambiguation of real content and artifacts, validated reading-order and so on.

But how does including the PDF/UA identifier improve the experience of a PWD?

Bouton Jones

1

From an email from [info@pdfa.org](mailto:info@pdfa.org):

The PDF/UA flag is a quality indicator. It prompts AT to expect high-quality tagging and other structures required for accessibility, as well as alt. text, validated disambiguation of real content and artifacts, validated reading-order and so on.

But how does including the PDF/UA identifier improve the experience of a PWD?

## Slide 60 - IMHO

### IMHO

1. Should *not* fix the PDF/UA identifier unless your document meets **all** other PDF/UA compliance criteria
2. You do *not* need to fix the PDF/UA identifier unless your Accessibility policy **explicitly** requires full PDF/UA compliance.

Bouton Jones 2

### IMHO

- You should *not* fix the PDF/UA identifier unless your document meets **all** other PDF/UA compliance criteria.
- You do *not* need to fix the PDF/UA identifier unless your Accessibility policy **explicitly** requires full PDF/UA compliance.

## Slide 61 - Solutions for PDF/UA identifier missing (1 of 3)

### Solutions for PDF/UA identifier missing (1 of 3)

#### 1. Adobe Acrobat DC

- A. Open the PDF document in Adobe Acrobat DC.
- B. Select More Tools > Print Production > Preflight
- C. In the Preflight dialog, be sure PDF Standards is selected from the dropdown.
- D. Activate the wrench icon above the list of Preflight checks. This should open the Document info and Metadata settings.
- E. Select the Set PDF/UA-1 entry option, then click the Fix button.
- F. In the next dialog box, you will need to re-save your file with a different name.

CITE: <https://accessiblewebsiteservices.com/accessible-pdfs-how-to-add-a-pdf-ua-identifier/>

Bouton Jones

3

#### 1. Adobe Acrobat DC

- Open the PDF document in Adobe Acrobat DC.
- Select More Tools > Print Production > Preflight
- In the Preflight dialog, be sure PDF Standards is selected from the dropdown.
- Activate the wrench icon above the list of Preflight checks. This should open the Document info and Metadata settings.
- Select the Set PDF/UA-1 entry option, then click the Fix button.
- In the next dialog box, you will need to re-save your file with a different name.

CITE: <https://accessiblewebsiteservices.com/accessible-pdfs-how-to-add-a-pdf-ua-identifier/>

## Slide 62 - Solutions for PDF/UA identifier missing (2 & 3)

### Solutions for PDF/UA identifier missing (2 & 3)

2. Add Metadata with an **XMP** File

- Instructions at taggedpdf.com: <https://taggedpdf.com/508-pdf-help-center/pdfua-identifier-missing/>
- Instructions at accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>

3. In **InDesign** (using the **MadeToTag** plugin from axaio)

- Instructions at accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>
- About MadeToTag: <https://www.axaio.com/doku.php/en:products:madetotag>

Bouton Jones 1

### 2. Add Metadata with an XMP File

- Instructions at taggedpdf.com: <https://taggedpdf.com/508-pdf-help-center/pdfua-identifier-missing/>
- Instructions at accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>

### 3. Automatic approach in InDesign (MadeToTag)

Instructions at accessible-pdf.info: <https://accessible-pdf.info/tutorials/general/pdf-ua-identifier-is-missing>

## Slide 63 - Solutions for PDF/UA identifier missing (4)

### Solutions for PDF/UA identifier missing (4)

4. Use the **axesPDF** tool from axes4 (<https://www.axes4.com/>) to add the identifier to compliant documents.

4. Use the axesPDF tool from axes4 (<https://www.axes4.com/>) to add the identifier to compliant documents.

## Slide 64 - It's worth repeating ...

### It's worth repeating ...

A missing PDF/UA Identifier can be an issue for PDFs created in:

- Microsoft Word 2013 for Windows
- Any version of Adobe InDesign (formerly Adobe PageMaker)

So, it might not be relevant to you.

Bouton Jones

3

A missing PDF/UA Identifier can be an issue for PDFs created in:

- Microsoft Word 2013 for Windows
- Any version of Adobe InDesign (formerly Adobe PageMaker)

So, what does that mean for you?

If you don't use these programs or don't maintain legacy documents created with these programs, this is probably not an issue for you.

## Slide 65 - Checkpoint 06: 003 & 004 (Dublin Core) Title

### Checkpoint 06: 003 & 004 DC:Title

| Index  | Failure Condition                                                                     | Section   |
|--------|---------------------------------------------------------------------------------------|-----------|
| 06-001 | Document does not contain XMP metadata stream                                         | UA1:7.1-8 |
| 06-002 | The metadata stream in the Catalog dictionary does not include the PDF/UA identifier. | UA1:5     |
| 06-003 | Metadata stream does not contain <i>dc:title</i> *                                    | UA1:7.1-8 |
| 06-004 | <i>dc:title</i> does not clearly identify the document                                | UA1:7.1-8 |



**\*Dublin Core (DC):** a set of fifteen "core" elements (properties) for describing resources. This fifteen-element Dublin Core has been formally standardized as ISO 15836, ANSI/NISO Z39.85,[2] and IETF RFC 5013. The core properties are part of a larger set of DCMI Metadata Terms. (CITE: [https://en.wikipedia.org/wiki/Dublin\\_Core](https://en.wikipedia.org/wiki/Dublin_Core))

Bouton Jones

4

06-003: Metadata stream does not contain dc:title

06-004: dc:title does not clearly identify the document

\*Dublin Core (DC): a set of fifteen "core" elements (properties) for describing resources. This fifteen-element Dublin Core has been formally standardized as ISO 15836, ANSI/NISO Z39.85,[2] and IETF RFC 5013. The core properties are part of a larger set of DCMI Metadata Terms. "Dublin Core" is also used as an adjective for Dublin Core metadata, a style of metadata that draws on multiple RDF vocabularies, packaged and constrained in Dublin Core application profiles. (CITE: [https://en.wikipedia.org/wiki/Dublin\\_Core](https://en.wikipedia.org/wiki/Dublin_Core) )



## Slide 66 - In HTML, the Title Tag is in the HEAD Section

### In HTML, the Title Tag is in the *HEAD* Section

```
1 <html>
2 <head>
3   <title>Heading Levels in HTML</title>
4   <link rel="stylesheet" type="text/css"
5     href="styles.css" />
6 </head>
7 <body lang="EN-US">
8
9 <p class=MsoTitle>MS Word Title</p>
10 <p class=MsoSubtitle>MS Subtitle</p>
11 <h1>MS Word Heading Level One</h1>
12 <h2>MS Word Heading Level Two</h2>
13 <h3>MS Word Heading Level Three</h3>
14 <h4>MS Word Heading Level Four</h4>
15 <h5>MS Word Heading Level Five</h5>
16 <h6>MS Word Heading Level Six</h6>
17 <p class=MsoHeading7>MS Word Heading Level Seven</p>
18 <p class=MsoHeading8>MS Word Heading Level Eight</p>
19 <p class=MsoHeading9>MS Word Heading Level Nine</p>
20
21 </body>
22 </html>
```

5

IMAGE 1 of 2: HTML Markup

---- Begin Markup ----

```
<html>
<head>
  <title>Heading Levels in HTML</title>
  <link rel="stylesheet" type="text/css" href="styles.css" />
</head>
<body lang="EN-US">

<p class=MsoTitle>MS Word Title</p>
<p class=MsoSubtitle>MS Subtitle</p>
<h1>MS Word Heading Level One</h1>
<h2>MS Word Heading Level Two</h2>
<h3>MS Word Heading Level Three</h3>
<h4>MS Word Heading Level Four</h4>
<h5>MS Word Heading Level Five</h5>
<h6>MS Word Heading Level Six</h6>
<p class=MsoHeading7>MS Word Heading Level Seven</p>
<p class=MsoHeading8>MS Word Heading Level Eight</p>
<p class=MsoHeading9>MS Word Heading Level Nine</p>

</body>
```

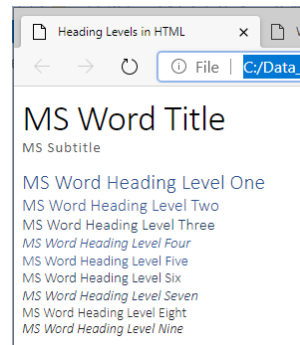
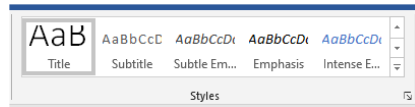
</html>

---- End Markup ----

IMAGE 2 of 2: Browser Tab displays the page title as being “Heading Levels in HTML”

Slide 67 - In Word, users typically add the Title Tag into the body of the document.

In Word, users add the Title Tag into the *body* of the document.



## Slide 68 - Question: How Does Word's **Title** Tag Export to PDF?

### Question: How Does Word's **Title** Tag Export to PDF?

1. As the PDF's Title *Property* (in the META data?)
2. Or as the top-most heading level?

Bouton Jones

7

2. Does it export as the PDF's Title Property (not to be confused with the filename?)
3. Does it export as the top-most heading level?

## Slide 69 – What the Heck!?

**What the Heck!?**

It does *not* export to the PDF title in the Meta Data

Bouton Jones 8

It does *not* export to the PDF title in the Meta Data

## Slide 70 – Another Surprise

### Another Surprise

Nor does Word's Title export to the PDF's top-most heading level. It exports to a *paragraph* tag.

This is the Title

This is Heading Level Two

This is Also Heading Level Two

<P>

This is the Title

<H2>

This is Heading Level Two

Navigation

Headings Pages Results

This is Heading Level Two

This is Also Heading Level Two

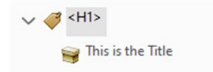
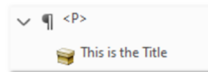
Bouton Jones 9

Nor does Word's Title export to the PDF's top-most heading level. It exports to a *paragraph* tag.

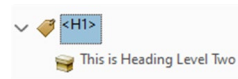
## Slide 71 - Two Options for the Title Paragraph

### Two Options for the Title Paragraph

1. Change the Paragraph tag into an H1 tag.



2. Remove the Paragraph tag and the title text inside it.

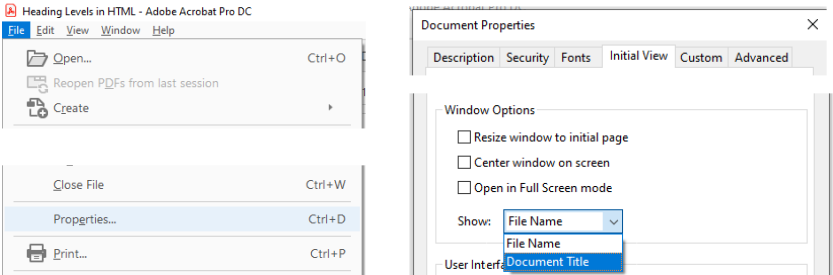


1. Change the Paragraph tag into an H1 tag.
2. Remove the Paragraph tag and the title text inside it.

## Slide 72 - How to Display the Document Title

### How to Display the Document Title

Select **File** drop-down menu > **Properties** option > select the **Initial View** tab in the **Document Properties** window > select **Document Title** in the **Show** field.



The screenshot shows the Adobe Acrobat Pro DC interface. On the left, the **File** menu is open, displaying options: **Open...** (Ctrl+O), **Reopen PDFs from last session**, **Create**, **Close File** (Ctrl+W), **Properties...** (Ctrl+D), and **Print...** (Ctrl+P). The **Properties...** option is highlighted. On the right, the **Document Properties** window is open, with the **Initial View** tab selected. Under the **Window Options** section, there are three checkboxes: **Resize window to initial page**, **Center window on screen**, and **Open in Full Screen mode**. Below these, the **Show:** dropdown menu is open, showing **File Name** and **Document Title**, with **Document Title** selected. The **User Interface** section is partially visible at the bottom.

Bouton Jones 11

Select File drop-down menu > Properties option > select the Initial View tab in the Document Properties window > select Document Title in the Show field.



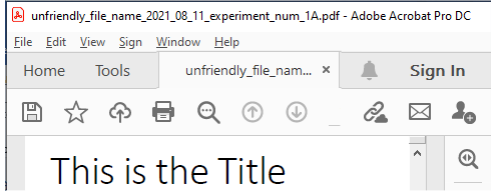
## Slide 73 - Example: Displaying the File Name

### Example: Displaying the File Name

By default, the Title is empty and the PDF is set to display the File Name instead of the Title

Title:

Show: File Name

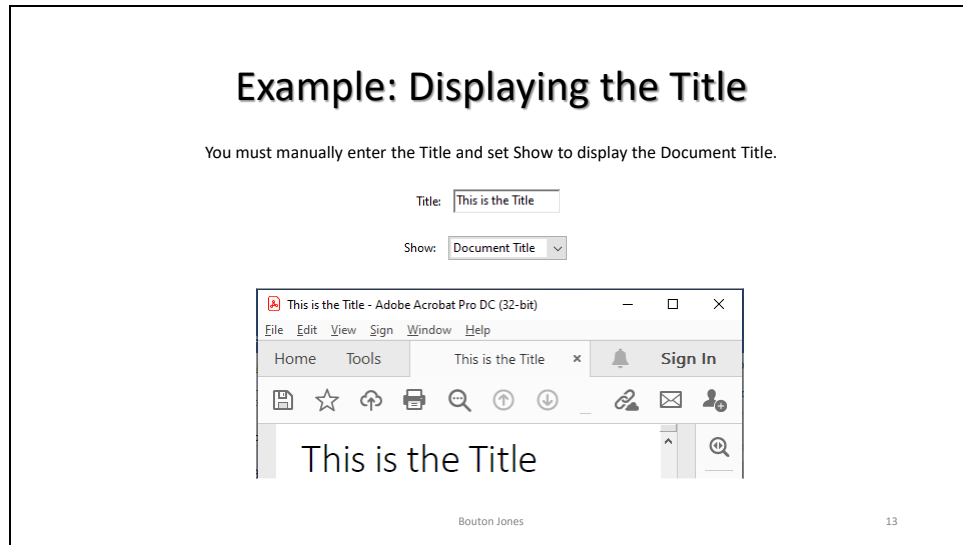


The screenshot shows the Adobe Acrobat Pro DC interface. The title bar indicates the file name 'unfriendly\_file\_name\_2021\_08\_11\_experiment\_num\_1A.pdf'. The menu bar includes File, Edit, View, Sign, Window, and Help. The toolbar shows icons for Home, Tools, Save, Star, Copy, Print, Comment, Up, Down, Share, Email, and Sign In. The main content area displays the text 'This is the Title'.

Bouton Jones 12

By default, the Title is empty, and the PDF is set to display the File Name instead of the Title

## Slide 74 - Example: Displaying the Title



You must manually enter the Title and set Show to display the Document Title.

## Slide 75 - Checkpoint 08: OCR Validation

**Checkpoint 08: OCR**

| Index  | Failure Condition                               | Section    |
|--------|-------------------------------------------------|------------|
| 08-001 | OCR-generated text contains significant errors. | UA1:7.1-10 |
| 08-002 | OCR-generated text is not tagged                | UA1:7.1-10 |



Bouton Jones

14

- OCR-generated text contains significant errors.
- OCR-generated text is not tagged

**IMAGE:** The acronym O. C. R. being viewed under a magnifying glass.

Optical character recognition or optical character reader is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo or from subtitle text superimposed on an image.

**CITE:** Wikipedia

## Slide 76 - How Reliable is OCR?

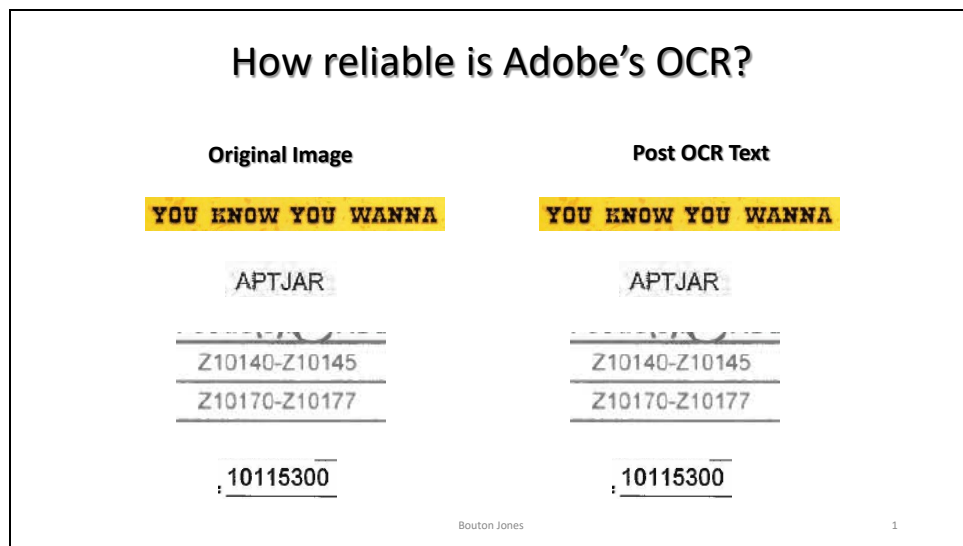


Table with two columns and four rows

Column One: Original Image

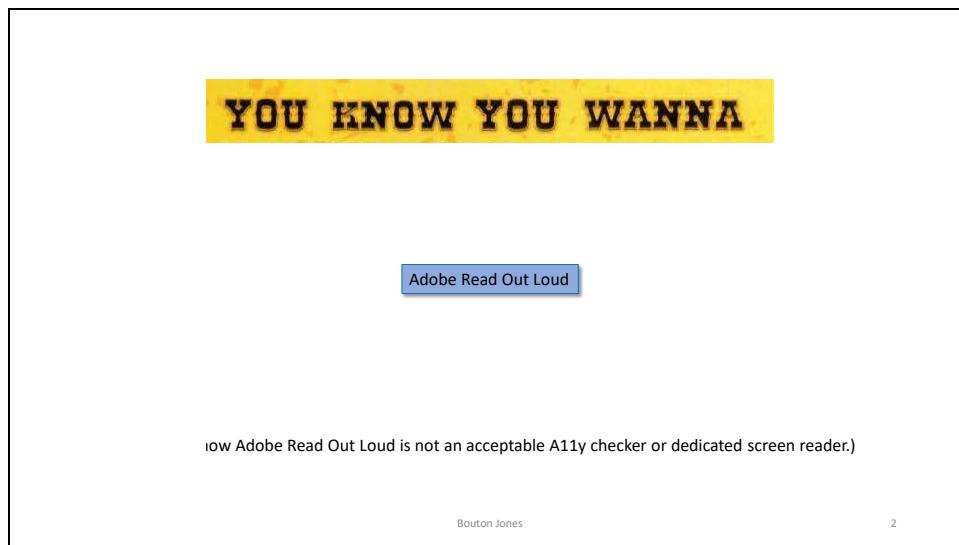
Column Two: Post OCR Text

The screen captures in column one and column two appear identical.

After performing OCR, usually the resulting renderable text looks exactly like the scanned image of text made with a decorative font or Word Art (e.g. "You Know You Wanna.") It's visually indistinguishable to the original to me. But now I can select the new text to copy or edit.

Because electronic signatures were not an option for the last three examples, the documents were filled out online, printed, signed, and scanned.

## Slide 77 - How would a screen reader identify the following text?



**IMAGE:** Stylized text in a decorative font – “You Know You Wanna”

How would Adobe’s Read Out Loud identify the Post OCR text above?

As “You Know You Wanna?”

Or something completely incomprehensible?

**AUDIO FILE:** Adobe Read Out Loud function saying “YDV H!fDW YOU WAH!fA?”

Sorry. I know Adobe Read Out Loud is not acceptable an A11y checker or dedicated screen reader.

## Slide 77 - How Reliable is OCR (really?)

| How reliable is Adobe's OCR (really?) |                                 |
|---------------------------------------|---------------------------------|
| Post OCR Text                         | Post OCR Text Copied and Pasted |
| YOU KNOW YOU WANNA                    | YDV H!fDW YOU WAHIfA            |
| APTJAR                                | Af>T JAR                        |
| Z10140-Z10145                         | <b>2</b> 10140- <b>2</b> 10145  |
| Z10170-Z10177                         | Z10170-Z10177                   |
| :10115300                             | _1_0_1_1_5_3_0_0                |
| Bouton Jones 3                        |                                 |

Table with two columns and four rows

Column One: Post OCR Text

Column Two: Post OCR Text

Often the Post OCR text can appear exactly the same as the original text in graphic but in reality, the characters in the new text will not correspond with the original text.

For example: The text in graphic might have been meant to be “You know you wanna” but when converted in OCR to actual text the characters are “YDV H!DW YOU WAHIfA.” And to a sighted person they may appear to be “You know you wanna” but a screen reader will identify it as “YDV H!DW YOU WAHIfA.”

## Slide 78 - Clarification on Adobe's OCR

### Clarification on Adobe's OCR

Adobe improved its OCR in 2024 and again in 2025, but decorative fonts still fail often, especially where  *Kerning*  is extreme.

**Kerning:**

the typography process of adjusting the horizontal space between two specific characters to improve visual balance and readability. It corrects uneven spacing, such as bringing letters like 'A' and 'V' closer together, to create a harmonious look.

Bouton Jones

4

Adobe improved its OCR in 2024 and again in 2025, but decorative fonts still fail often, especially where kerning is extreme.

**Kerning:**

The typography process of adjusting the horizontal space between two specific characters to improve visual balance and readability. It corrects uneven spacing, such as bringing letters like 'A' and 'V' closer together, to create a harmonious look.

## Slide 79 - What does the “Recognize Text” (OCR) function do?

**POLL QUESTION: What does the “Recognize Text” (OCR) function do?**

- A. Does it convert the image of text in the image-only PDF into true text that can be edited like a Word document?
- B. Does Acrobat place an invisible layer of "renderable text" --- i.e. vector (computer generated) text --- on top of the image layer?

Bouton Jones

5

Does it convert the image of text in the image only PDF into true text that can be edited like a Word document?

Or does Acrobat place an invisible layer of "renderable text" --- i.e. vector (computer generated) text --- on top of the image layer?



## Slide 80 – Poll Question: What does the “Recognize Text” (OCR) function do?

**POLL QUESTION: What does the “Recognize Text” (OCR) function do?**

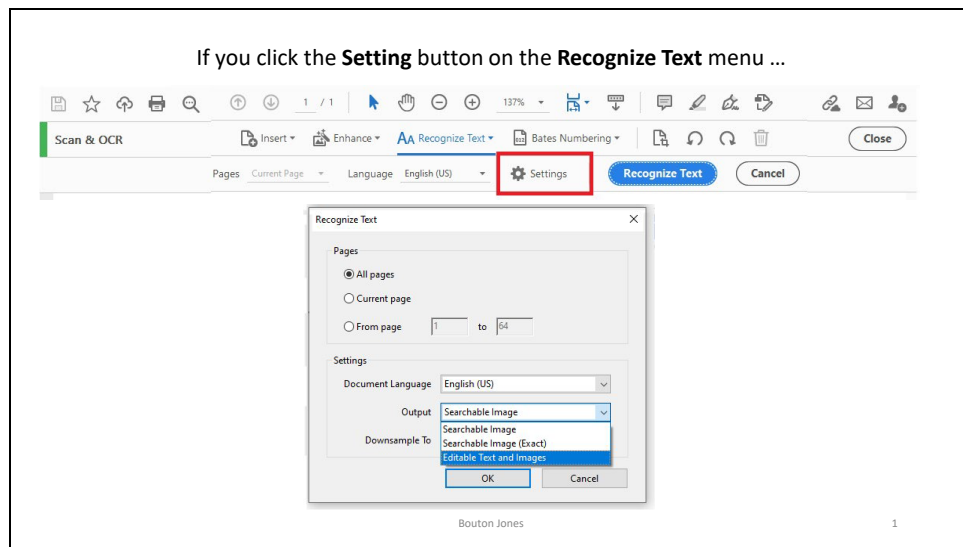
- A. Does it convert the image of text in the image-only PDF into true text that can be edited like a Word document?
- B. Does Acrobat place an invisible layer of "renderable text" --- i.e. vector (computer generated) text --- on top of the image layer?

Bouton Jones

1

- A. Does it convert the image of text in the image only PDF into true text that can be edited like a Word document?
- B. Or does Acrobat place an invisible layer of "renderable text" --- i.e. vector (computer generated) text --- on top of the image layer?

## Slide 81 - Output Options for the Recognize Text (OCR) Function in Adobe Acrobat 2017



If you click the Setting Button on the Recognize Text menu, you can choose among three options in the Recognize Text dialog box.

- Searchable Image
- Searchable Image (exact)
- Editable Text and Image

## Slide 82 - Two Output Options:

| Editable Text and Image                                                                         | Searchable Image and<br>Searchable Image (Exact)                                                        |
|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| Converts the graphical text into vector text. The vector text will replace the “text in image.” | Places an invisible “layer” (not <i>really</i> a layer) on top of the original image.                   |
| This setting creates and embeds a custom font that looks like the original font in the image.   | This setting does not alter the text in the image or embed a custom font.                               |
| You can now edit the text.                                                                      | You can’t easily edit the document, but it can be read by assistive technology (such as JAWS and NVDA.) |
| <small>Bouton Jones</small>                                                                     | <small>2</small>                                                                                        |

### 1. Editable Text and Image:

Converts the graphical text into vector text. The vector text will replace the “text in image.”

It creates and embeds a custom font that looks like the original font in the image.

You can now edit the text.

The file size is larger

### 2. Searchable Image

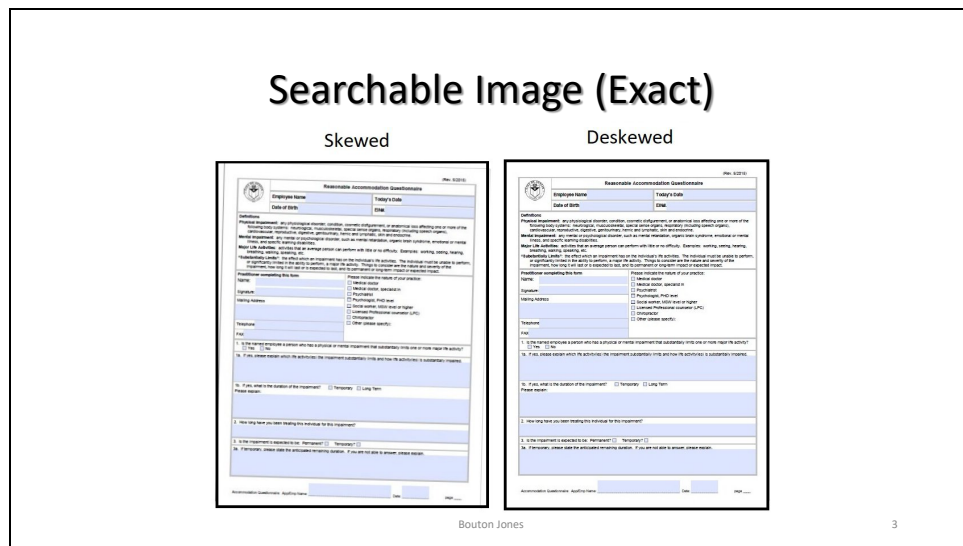
Places an invisible “layer” (not really a layer) on top of the original image.

You can’t easily edit the document but it can be read by assistive technology (such as JAWS and NVDA.)

### Additional Resources on the Output Options

- Understanding OCR options in Adobe Acrobat: “Searchable Image”, “Searchable Image (Exact)”, and “Editable Text and Images” (Superuser) - <https://superuser.com/questions/1276594/understanding-ocr-options-in-adobe-acrobat-searchable-image-searchable-imag>
- Searchable Image vs. Searchable Image (Exact) - Quality of OCR (Adobe Community) - <https://community.adobe.com/t5/acrobat/searchable-image-vs-searchable-image-exact-quality-of-ocr/td-p/4734730?page=1>

## Slide 83 – Searchable Image (Exact)



**IMAGES:** An original skewed page and a deskewed version side by side

### Deskew:

The process of straightening an image that has been scanned or photographed crookedly — that is an image that is slanting too far in one direction, or one that is misaligned. (Vangie Beal on webopedia)

### CITE:

[https://www.webopedia.com/TERM/I/image\\_de\\_skew.html#:~:text=By%20Vangie%20Beal%20The%20process,Also%20called%20deskew.%5C](https://www.webopedia.com/TERM/I/image_de_skew.html#:~:text=By%20Vangie%20Beal%20The%20process,Also%20called%20deskew.%5C)

The Searchable Image setting will deskew the image before OCR.

The Searchable Image (Exact) setting will NOT deskew the image before OCR.

Often deskewing the image will lead to a smaller file size.

But there may be a reason that you don't want to alter the document. (Ask a lawyer)

## Slide 84 - Checkpoint 09: Appropriate Tags

### Checkpoint 09: Appropriate Tags

| Index  | Failure Condition                                                                                                         | Section   |
|--------|---------------------------------------------------------------------------------------------------------------------------|-----------|
| 09-001 | Tags are not in logical reading order.                                                                                    | UA1:7.2-1 |
| 09-002 | Structure elements are nested in a semantically inappropriate manner. (e.g., a table inside a heading).                   | UA1:7.2-1 |
| 09-003 | The structure type (after applying any role-mapping as necessary) of a structure element is not semantically appropriate. | UA1:7.2-1 |
| 09-004 | A table-related structure element is used in a way that does not conform to the syntax defined in ISO 32000-1, Table 337. | UA1:7.2-1 |
| 09-005 | A list-related structure element is used in a way that does not conform to Table 336 in ISO 32000-1.                      | UA1:7.2-1 |
| 09-006 | A TOC-related structure element is used in a way that does not conform to Table 333 in ISO 32000-1.                       | UA1:7.2-1 |
| 09-007 | A Ruby-related structure element is used in a way that does not conform to Table 338 in ISO 32000-1.                      | UA1:7.2-1 |
| 09-008 | A Warichu-related structure element is used in a way that does not conform to Table 338 in ISO 32000-1.                   | UA1:7.2-1 |

Bouton Jones

4

09-001: Tags are not in logical reading order.

09-002: Structure elements are nested in a semantically inappropriate manner. (e.g., a table inside a heading).

09-003: The structure type (after applying any role-mapping as necessary) of a structure element is not semantically appropriate.

09-004: A table-related structure element is used in a way that does not conform to the syntax defined in ISO 32000-1, Table 337.

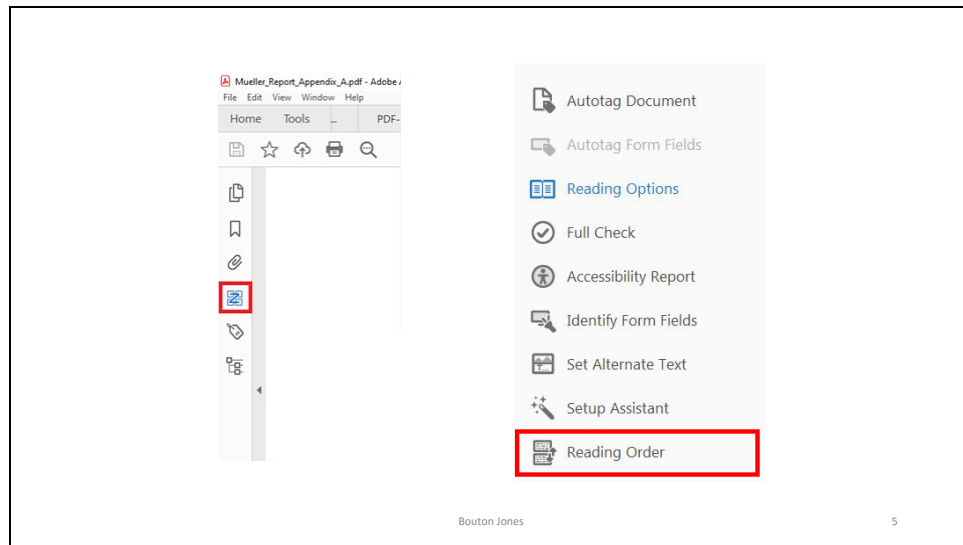
09-005: A list-related structure element is used in a way that does not conform to Table 336 in ISO 32000-1.

09-006: A TOC-related structure element is used in a way that does not conform to Table 333 in ISO 32000-1.

09-007: A Ruby-related structure element is used in a way that does not conform to Table 338 in ISO 32000-1.

09-008: A Warichu-related structure element is used in a way that does not conform to Table 338 in ISO 32000-1.

## Slide 85 – Navigating to the Reading Order panel



## Slide 86 - Checkpoint 14: Headings

### Checkpoint 14: Headings

| Index                                                                                                                                                      | Failure Condition                                                                                      | Section     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|-------------|
| 14-001                                                                                                                                                     | Headings are not tagged.                                                                               | UA1:7.4-1   |
| 14-002                                                                                                                                                     | Does use numbered headings, but the first heading tag is not H1.                                       | UA1:7.4.2-1 |
| 14-003                                                                                                                                                     | Numbered heading levels in descending sequence are skipped (Example: H3 follows directly after H1).    | UA1:7.4-1   |
| 14-004                                                                                                                                                     | Numbered heading tags do not use Arabic numerals.                                                      | UA1:7.4.3-1 |
| 14-005                                                                                                                                                     | Content representing a 7 <sup>th</sup> level (or higher) heading does not use an “H7” (or higher) tag. | UA1:7.4.3-1 |
| 14-006                                                                                                                                                     | A node contains more than one H tag.                                                                   | UA1:7.4.4-1 |
| 14-007                                                                                                                                                     | Document uses both H and H# tags.                                                                      | UA1:7.4.4-3 |
| <b>NOTE:</b> In weakly-structured documents, headings always take the form “Hn” (e.g., H1, H2, Hn) without intervening whitespace or numerical separators. |                                                                                                        |             |

Bouton Jones

6

| Index         | Failure Condition                                                                                                                                                                            | Section     |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>14-001</b> | Headings are not tagged.                                                                                                                                                                     | UA1:7.4-1   |
| <b>14-002</b> | Does use numbered headings, but the first heading tag is not H1.                                                                                                                             | UA1:7.4.2-1 |
| <b>14-003</b> | Numbered heading levels in descending sequence are skipped (Example: H3 follows directly after H1).                                                                                          | UA1:7.4-1   |
| <b>14-004</b> | Numbered heading tags do not use Arabic numerals.                                                                                                                                            | UA1:7.4.3-1 |
| <b>14-005</b> | Content representing a 7th level (or higher) heading does not use an “H7” (or higher) tag.                                                                                                   | UA1:7.4.3-1 |
| <b>14-006</b> | A node contains more than one H tag.                                                                                                                                                         | UA1:7.4.4-1 |
| <b>14-007</b> | Document uses both H and H# tags.<br><br>NOTE: In weakly-structured documents, headings always take the form “Hn” (e.g., H1, H2, Hn) without intervening whitespace or numerical separators. | UA1:7.4.4-3 |

## Slide 87 - Update: Skipping Heading Levels in PDF/UA-2

### **Update:** Skipping Heading Levels in PDF/UA-2

Unlike PDF/UA-1, PDF/UA-2 permits skipping heading levels (H1 to H3, for example) under specific circumstances, providing flexibility in document structure.

Unlike PDF/UA-1, PDF/UA-2 permits skipping heading levels (H1 to H3, for example) under specific circumstances, providing flexibility in document structure.



## Slide 88 - Why would you lose headings when exporting from Word to PDF?

Why would you lose headings when exporting from Word to PDF?



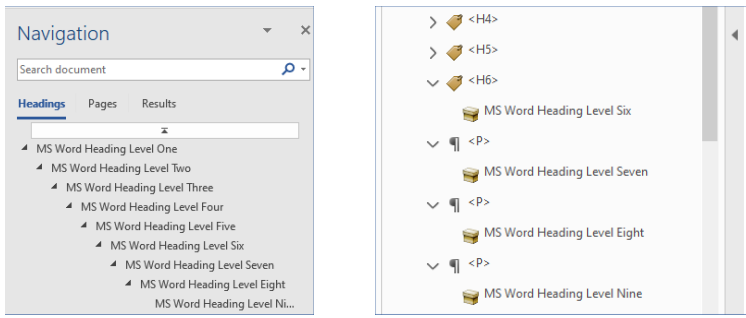
Bouton Jones

8

**IMAGE:** guillotine. an apparatus designed for efficiently carrying out executions by beheading. The device consists of a tall, upright frame with a weighted and angled blade suspended at the top. The condemned person is secured with stocks at the bottom of the frame, positioning the neck directly below the blade.

## Slide 89 - Exporting Heading Levels from Word to PDF

**Exporting Heading Levels from Word to PDF**



Try to stop at 6.

Bouton Jones 1

First screen capture: MS Word document showing 9 levels of headings

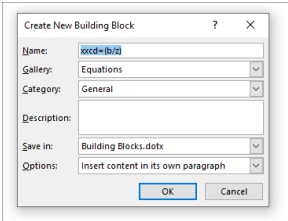
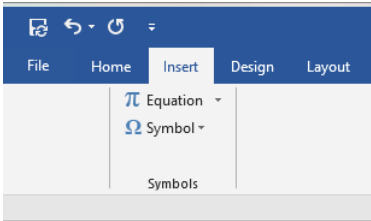
Second screen capture: PDF tag panel showing the tag panel for a PDF that was exported from the previously mentioned MS Word document

The headings at level 7, 8, and 9 have exported as paragraphs in a PDF.

## Slide 90 - Checkpoint 17: Mathematical Expressions

### Checkpoint 17: Mathematical Expressions

| Index  | Failure Condition                                                          | Section   |
|--------|----------------------------------------------------------------------------|-----------|
| 17-001 | Content is a mathematical expression but is not tagged with a Formula tag. | UA1:7.7-1 |
| 17-002 | Formula tag is missing an Alt attribute.                                   | UA1:7.7-1 |
| 17-003 | Unicode mapping requirements are not met.                                  | UA1:7.7-2 |



Bouton Jones 2

17-001: Content is a mathematical expression but is not tagged with a Formula tag.

17-002: Formula tag is missing an Alt attribute.

17-003: Unicode mapping requirements are not met.

**IMAGES:** Two Screen captures showing

- The Insert tab selected in the Word ribbon. The Symbols section shows the Equation and Symbol functions
- The Create New Building Block window

## Slide 91- **Update:** Mathematical expressions in PDF/UA-2

### **Update:** Mathematical expressions in PDF/UA-2

PDF/UA-2 (ISO 14289-2) revolutionizes mathematical expression accessibility by mandating the embedding of MathML directly into tagged PDF 2.0 documents, allowing screen readers to accurately read and navigate complex formulas.

#### **MathML:**

an XML-based industry standard used to display complex mathematical formulas and notations on the web, ensuring they are rendered accurately and remain accessible to screen readers. It acts as a markup language that allows browsers to understand and format equations structurally rather than treating them as static images.

Bouton Jones

1

PDF/UA-2 (ISO 14289-2) revolutionizes mathematical expression accessibility by mandating the embedding of MathML directly into tagged PDF 2.0 documents, allowing screen readers to accurately read and navigate complex formulas.

#### **MathML:**

an XML-based industry standard used to display complex mathematical formulas and notations on the web, ensuring they are rendered accurately and remain accessible to screen readers. It acts as a markup language that allows browsers to understand and format equations structurally rather than treating them as static images.

## Slide 92 - Embed MathML

### Embed MathML

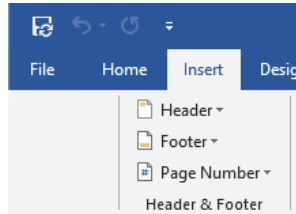
1. As an Associated File (AF) linked to a <Formula> tag
2. via direct tagging with structure elements from the MathML namespace.

1. As an Associated File (AF) linked to a <Formula> tag
2. via direct tagging with structure elements from the MathML namespace.

## Slide 93 - Checkpoint 18: Page Headers and Footers

### Checkpoint 18: Page Headers and Footers

| Index  | Failure Condition                                                           | Section   |
|--------|-----------------------------------------------------------------------------|-----------|
| 18-001 | Headers and footers are not marked as pagination artifacts.                 | UA1:7.8-1 |
| 18-002 | Header or footer artifacts are not classified as Header or Footer subtypes. | UA1:7.8-1 |



*These three semantic tags are listed under the **Insert** tab in Word.*

Bouton Jones

1

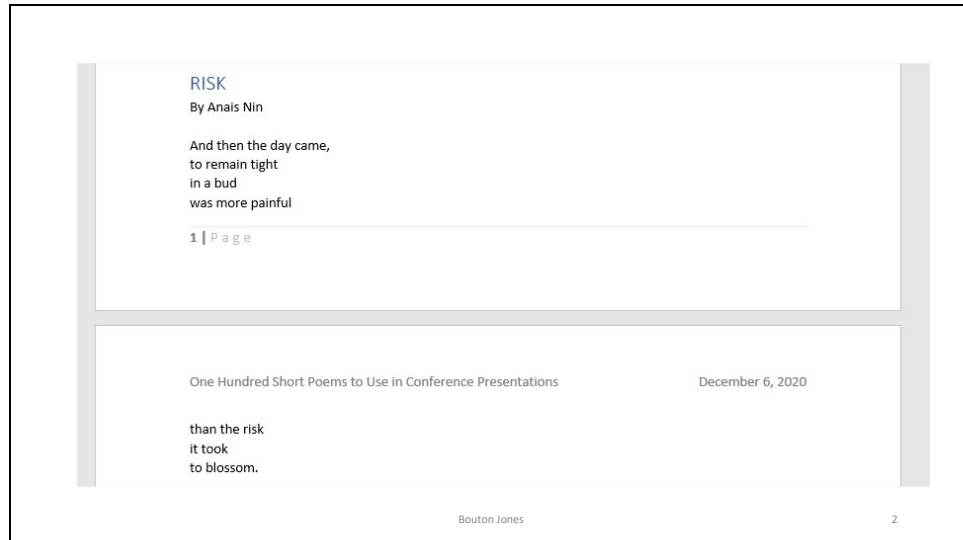
18-001: Headers and footers are not marked as pagination artifacts.

18-002: Header or footer artifacts are not classified as Header or Footer subtypes.

**IMAGE:** These three semantic tags (Header, Footer, and Page Number) are listed under the Insert tab in Word.

This is specific to Word. It doesn't apply to HTML.

## Slide 94 - Will assistive technology read the headers and footers?



**IMAGE:** Screen capture of PDF showing the header and footer of a document

If the PDF was exported from a properly tagged Word document, screen readers should not read the headers and footers. But you should confirm this before and after properly exporting the source document to PDF.

## Slide 95 - So, what if the information in the header and footer is important?



If the information in the header and footer is important to the understanding of the content, you must repeat that information in the body of the document to ensure an equitable experience for the screen reader users.

**SCREEN CAPTURE:** First page of a document repeating information from the header --- i.e. the document title and the published date --- in the body of the document (specifically on the first page.)



## Slide 96 - Checkpoint 22: Article Threads

### Checkpoint 22: Article Threads

22-001: Article threads do not reflect logical reading order.

In PDFs, articles are optional electronic threads that the PDF author defines within that PDF. Articles lead readers through the PDF content, jumping over pages or areas of the page that are not included in the article. It is the same way that you skim through a traditional newspaper or magazine, following one specific story and ignoring the rest. When you read an article, the page view zooms in or out so that the current part of the article fills the screen.

CITE: [https://helpx.adobe.com/acrobat/using/navigating-pdf-pages.html#article\\_threads](https://helpx.adobe.com/acrobat/using/navigating-pdf-pages.html#article_threads)

Bouton Jones

4

22-001: Article threads do not reflect logical reading order.

"In PDFs, articles are optional electronic threads that the PDF author defines within that PDF. Articles lead readers through the PDF content, jumping over pages or areas of the page that are not included in the article. It is the same way that you skim through a traditional newspaper or magazine, following one specific story and ignoring the rest. When you read an article, the page view zooms in or out so that the current part of the article fills the screen." (CITE: [https://helpx.adobe.com/acrobat/using/navigating-pdf-pages.html#article\\_threads](https://helpx.adobe.com/acrobat/using/navigating-pdf-pages.html#article_threads))

## Slide 97 - Defining Article Threads

### Defining Article Threads

"You create an article by defining a series of boxes around the content in the order in which you want the content read. The navigational path you define for an article is known as the article thread. You create a thread connecting the various boxes, unifying them into a continuous text flow.

....

"Most desktop publishing programs allow you to generate article threads automatically as you convert the files to Adobe PDF. If the file you're viewing has articles, you can show the names of the articles on a tab and navigate easily through them."

CITE: <https://helpx.adobe.com/acrobat/using/pdf-articles.html>

Bouton Jones

5

"You create an article by defining a series of boxes around the content in the order in which you want the content read. The navigational path you define for an article is known as the article thread. You create a thread connecting the various boxes, unifying them into a continuous text flow.

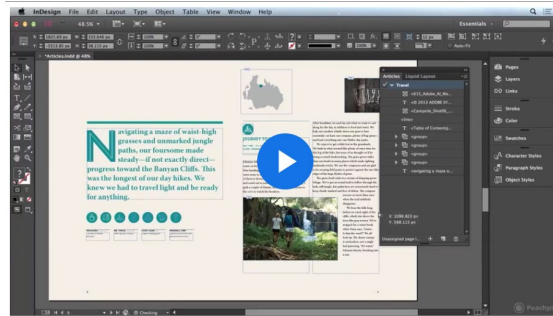
"Note:

"Most desktop publishing programs allow you to generate article threads automatically as you convert the files to Adobe PDF. If the file you're viewing has articles, you can show the names of the articles on a tab and navigate easily through them."

CITE: <https://helpx.adobe.com/acrobat/using/pdf-articles.html>

## Slide 98 - Adobe InDesign Video: "The Articles panel"

### Adobe InDesign Video: "The Articles panel"



URL: <https://helpx.adobe.com/indesign/how-to/indesign-articles-panel.html>

Bouton Jones

1

URL: <https://helpx.adobe.com/indesign/how-to/indesign-articles-panel.html>

## Slide 99 - Checkpoint 23: Digital Signatures

### Checkpoint 23: Digital Signatures

| Electronic Signature                                                                                                                                   | Digital Signature                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| An electronic symbol attached to a contract or other record, used by a person with an intent to sign.                                                  | A cryptographically secure electronic signature that guarantees an electronic document is authentic. They use a certificate-based digital ID issued by an accredited Certificate Authority (CA) or Trust Service Provider (TSP.) |
| It can be a scanned image of a signature. Or a text field or text fields.                                                                              |                                                                                                                                                                                                                                  |
|                                                                       |                                                                                                                                                |
| Digital Signatures are a type of Electronic Signature. All Digital Signatures are Electronic Signatures but not all Electronic Signatures are digital. |                                                                                                                                                                                                                                  |

Bouton Jones 2

### Types of eSignatures

**Electronic Signatures:** An electronic symbol attached to a contract or other record, used by a person with an intent to sign.

**Digital Signature:** A cryptographically secure electronic signature that guarantees an electronic document is authentic. They use a certificate-based digital ID issued by an accredited Certificate Authority (CA) or Trust Service Provider (TSP.)

Digital Signatures are a type of Electronic Signature. All Digital Signatures are Electronic Signatures but not all Electronic Signatures are digital.

## Slide 100 - Testing your Digital Signature Fields

### Testing your Digital Signature Fields

"No test specific to digital signatures is required, however other provisions apply (form fields)."

Digital signatures are form fields with a lot of functionality.

I suggest testing digital signatures "manually" with screen readers and keyboards (not mice.)

Bouton Jones

3

"No test specific to digital signatures is required, however other provisions apply (form fields)."

Digital signatures are form fields with a lot of functionality.

I suggest testing digital signatures "manually" with screen readers and keyboards (but not mice.) Can you read the document (including the digital signature field?) Can you apply the digital signature using your screen reader and keyboard (but not your mouse?) Can you confirm that you successfully sign the document using a screen reader and a mouse?

## Slide 101 - Checkpoint 25: XFA

### Checkpoint 25: XFA

"XFA (also known as XFA forms) stands for XML Forms Architecture, a family of proprietary XML specifications ... to enhance the processing of web forms. .... The XML Forms Architecture ... has been deprecated in PDF 2.0.

"XFA forms are saved internally in PDF files or as XDP (XML Data Package) files which can be opened in Adobe's LiveCycle Designer software. An XDP can package a PDF file, along with XML form and template data. XDP provides a mechanism for packaging form components within a surrounding XML container."

CITE: <https://en.wikipedia.org/wiki/XFA>

Bouton Jones

4

What are XFA Forms?

"XFA (also known as XFA forms) stands for XML Forms Architecture, a family of proprietary XML specifications that was suggested and developed by JetForm to enhance the processing of web forms. It can be also used in PDF files starting with the PDF 1.5 specification. The XFA specification is referenced as an external specification necessary for full application of the ISO 32000-1 specification (PDF 1.7). The XML Forms Architecture was not standardized as an ISO standard,[3] and has been deprecated in PDF 2.0."

"XFA forms are saved internally in PDF files or as XDP (XML Data Package) files which can be opened in Adobe's LiveCycle Designer software. An XDP can package a PDF file, along with XML form and template data. XDP provides a mechanism for packaging form components within a surrounding XML container."

<https://en.wikipedia.org/wiki/XFA>

## Slide 102 - XFA Update

### XFA Update

For 2026:

- Adobe Acrobat removed support for *dynamic* XFA in 2023.
- Static XFA rendering still exists in some enterprise contexts but is rare.
- Many federal agencies now reject XFA entirely.

Dynamic XFA is no longer supported by current Acrobat releases.

For 2026:

- Adobe Acrobat removed support for *dynamic* XFA in 2023.
- Static XFA rendering still exists in some enterprise contexts but is rare.
- Many federal agencies now reject XFA entirely.

Dynamic XFA is no longer supported by current Acrobat releases.

## Slide 103 - XFA Conclusion

### XFA Conclusion

- Dynamic XFA is *unsupported in modern Acrobat*; editing XFA isn't supported — you must **flatten/convert**.
- Built in browser PDF viewers still don't support XFA; use Acrobat/Reader or render to HTML5.

Bouton Jones

2

- Dynamic XFA is *unsupported in modern Acrobat*; editing XFA isn't supported — you must **flatten/convert**.
- Built in browser PDF viewers still don't support XFA; use Acrobat/Reader or render to HTML5.



## Slide 104 - So, if you don't use Adobe LiveCycle ...

So, if you do *not* use Adobe LiveCycle ...



Bouton Jones

1

XFA forms shouldn't be a problem.

IMAGE: Man bent backward to dodge bullets fired at him.

## Slide 105 - Checkpoint 30: XObjects

### Checkpoint 30: XObjects

| Index  | Failure Condition                                             | Section    |
|--------|---------------------------------------------------------------|------------|
| 30-001 | A reference XObject is present.                               | UA1:7.20-1 |
| 30-002 | Form XObject contains MCIDs and is referenced more than once. | UA1:7.20-2 |

**Form XObjects:** a way of describing objects (text, images, vector elements,...) within a PDF file. It is meant for repetitive objects that only get stored once in a PDF but referenced multiple times. In a way, Form XObjects are similar to a mini-PDF embedded inside the main PDF document ... (CITE: <https://www.prepressure.com/pdf/basics/form-xobjects> )

**Reference XObject:** a form XObject which contains an optional Ref sub-dictionary. The sub-dictionary specifies a page in some target PDF file. (CITE: <https://www.adobe.com/devnet-docs/acrobatetk/tools/PrefRef/Windows/AVDisplay.html> )

**SEE:** <https://www.prepressure.com/pdf/basics/form-xobjects>

Bouton Jones

2

**30-001** A reference XObject is present. UA1:7.20-1

**30-002** Form XObject contains MCIDs and is referenced more than once.  
UA1:7.20-2

Form XObjects: a way of describing objects (text, images, vector elements,...) within a PDF file. It is meant for repetitive objects that only get stored once in a PDF but referenced multiple times. In a way, Form XObjects are similar to a mini-PDF embedded inside the main PDF document ... (CITE: <https://www.prepressure.com/pdf/basics/form-xobjects>)

Reference XObject: a form XObject which contains an optional Ref sub-dictionary. The sub-dictionary specifies a page in some target PDF file. (CITE: <https://www.adobe.com/devnet-docs/acrobatetk/tools/PrefRef/Windows/AVDisplay.html>)

SEE: <https://www.prepressure.com/pdf/basics/form-xobjects>

## Slide 106 - Checkpoint 31: Fonts

### Checkpoint 31: Fonts

```
38 HTML,
39 BODY {
40     font-family: Calibri, "Trebuchet MS", Arial, Georgia, "Times New Roman", Helvetica, Serif;
41     font-size: 100%;
42     line-height: 1.125em;
43     margin-top: 0px;
44     margin-bottom: 0px;
45 }
```

**Businesses continue to be at great risk of a legal action regarding the accessibility of digital technologies, including websites, mobile applications, and kiosks.** The number of demand letters and lawsuits filed on this basis has continued unabated for quite a few

Bouton Jones 3

In HTML and CSS, you can choose alternative fonts

--- Begin CSS ---

```
HTML,
BODY {
    font-family: Calibri, "Trebuchet MS", Arial, Georgia, "Times New Roman", Helvetica, Serif;
    font-size: 100%;
    line-height: 1.125em;
    margin-top: 0px;
    margin-bottom: 0px;
}
```

--- End CSS ---

If the Calibri is not installed on the user's computer, the browser will use Trebuchet MS. If Trebuchet MS is not installed, the browser will use Arial. And so on and so forth.

But in Word you can only select one font to use at a time.

## Slide 107 – If You Are Missing an Embedded Font

If You are Missing an Embedded Font:



Two Screen Shorts showing garbage characters.

## Slide 108 - Adobe Acrobat's Base 14 Fonts

### Adobe Acrobat's Base 14 Fonts

|                                                                                                                                                                                                                                       |                                                                                                                                                                                          |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Monospaced Fonts</b> <ul style="list-style-type: none"><li>• Courier New</li><li>• <b>Courier New Bold</b></li><li>• <i>Courier New Oblique</i></li><li>• <b><i>Courier New Bold Oblique</i></b></li></ul>                         | <b>Proportional Fonts (Sans Serif)</b> <ul style="list-style-type: none"><li>• Helvetica</li><li>• Helvetica Bold</li><li>• Helvetica Oblique</li><li>• Helvetica Bold Oblique</li></ul> |
| <b>Proportional Fonts (Serif)</b> <ul style="list-style-type: none"><li>• Times New Roman</li><li>• <b>Times New Roman Bold</b></li><li>• <i>Times New Roman Italic</i></li><li>• <b><i>Times New Roman Bold Italic</i></b></li></ul> | <b>Remaining Fonts</b> <ul style="list-style-type: none"><li>• Symbol (e.g. Σμβολ.)</li><li>• <u>ZapfDingbats</u></li></ul>                                                              |

Bouton Jones 2

The Base 14 Fonts are defined in ISO 32000-1:2008(E) in Section 9.6.2.2.

They are also referred to as the “Standard 14 Fonts,” the “Standard Type 1 Fonts”, and the “Standard Fonts.”

**Monospaced Fonts:** Courier, Courier Bold, Courier Oblique, Courier Bold-Oblique

**Proportional Fonts (Sans Serif):** Helvetica, Helvetica Bold, Helvetica Oblique, Helvetica Bold-Oblique

**Proportional Fonts (Serif):** Times Roman, Times Bold, Times Italic, Times Bold-Italic

**Remaining Fonts:** Symbol, Zapf Dingbats

## Slide 109 - Base 14 Fonts Update

### Base 14 Fonts Update

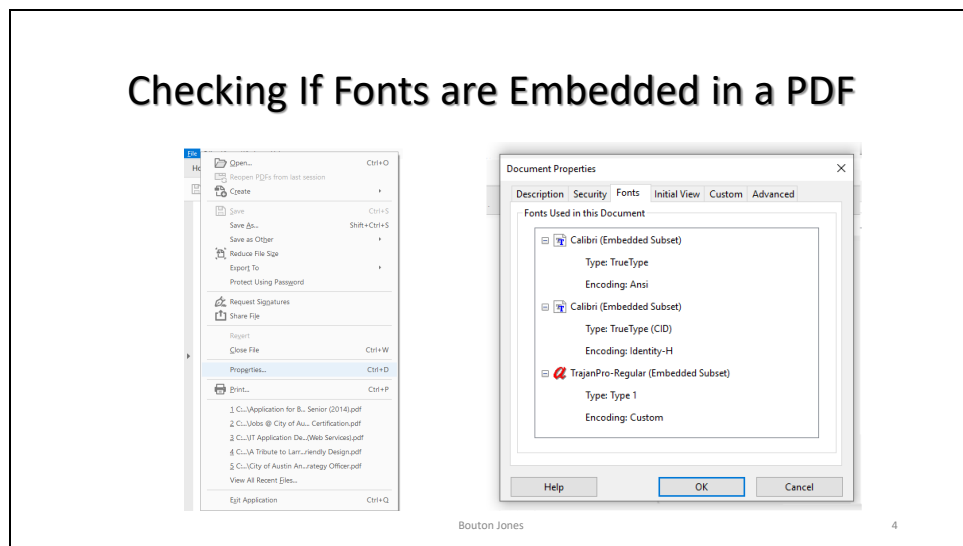
- ISO 32000-2 (PDF 2.0) revised some recommendations; Base 14 are still guaranteed, but embedding is still best practice for accessibility.
- Acrobat and Word use updated font substitution engines as of 2024 to address changing font technologies and improve performance.

Bouton Jones

3

- ISO 32000-2 (PDF 2.0) revised some recommendations; Base 14 are still guaranteed, but embedding is still best practice for accessibility.
- Acrobat and Word use updated font substitution engines as of 2024 to address changing font technologies and improve performance.

## Slide 110 - Checking If Fonts are Embedded in a PDF.



To check whether the fonts are all embedded in your PDF file or not:

- Open your PDF file
- Click File> Document Properties
- Click on the Fonts Tab to display the list of all fonts
- All fonts are either Type 1 or TrueType fonts
- All fonts should show as “Embedded Subset”

**CITE:** <https://kbpdfstudio.qoppa.com/how-to-embed-fonts-in-a-pdf-document/>

## Slide 111 - Embedding Fonts in Word Documents

### Embedding Fonts in *Word Docs*

1. In the **Files** tab select the **More** option.
2. Choose the **Options** option.
3. In the **Words** pane, select the **Save** option.
4. Select the “Embed fonts in the file” checkbox
5. Leave the “Embed only the characters used in the document (best for reducing file size)” checkbox.
6. Select the “Do not embed common systems fonts” checkbox.
7. Click the OK button to close this window.
8. You will need to use the “Save As PDF” or “Export to PDF”

The UI varies among releases of Microsoft Word. Microsoft is likely to change it for future versions of Word

Bouton Jones

5

1. In the **Files** tab select the **More** option.
2. Choose the **Options** option.
3. In the **Words** pane, select the **Save** option.
4. Select the “Embed fonts in the file” checkbox
5. Leave the “Embed only the characters used in the document (best for reducing file size)” checkbox.
6. Select the “Do not embed common systems fonts” checkbox.
7. Click the OK button to close this window.
8. You will need to use the “Save As PDF” or “Export to PDF”

The UI varies among releases of Microsoft Word. Microsoft is likely to change it for future versions of Word



## Slide 112 - Embedding Fonts in PDFs

### Embedding Fonts in *PDFs*

1. Click File > Print
2. Click Adobe PDF
3. Click the Properties button to the right of the Printer Name text box
4. Select the tab Adobe PDF Settings
5. Edit the Default Settings
6. Click Fonts
7. For Subset embedded fonts when percent of characters used is less than: Set the percentage to 100%
8. Select the Embed all Fonts option
  - A. For Embedding, select the folder with the fonts you want to embed from the drop-down list
  - B. Make sure the fonts you need to embed are in the Always Embed box and not in the Never Embed box

Bouton Jones

6

Pro tip: a font can only be embedded if it contains a setting by the font vendor that permits it to be embedded.

1. Click File > Print
2. Click Adobe PDF
3. Click the Properties button to the right of the Printer Name text box
4. Select the tab Adobe PDF Settings
5. Edit the Default Settings
6. Click Fonts
7. For Subset embedded fonts when percent of characters used is less than: Set the percentage to 100%
8. Select the Embed all Fonts option
  - a. For Embedding, select the folder with the fonts you want to embed from the drop-down list
  - b. Make sure the fonts you need to embed are in the Always Embed box and not in the Never Embed box

CITE: <https://www.printivity.com/insights/2020/09/13/how-to-embed-fonts-in-pdfs/>

See also: VIDEO "How to Embed Fonts in a PDF using Acrobat pro-2017" (1:21)  
<https://www.youtube.com/watch?v=Su50hSBGcW8>

## Slide 113 - How much does embedding fonts increase the file size?

### How much does embedding fonts increase the file size?

**NOTE:** This was a quick and dirty check --- not a rigorous scientific study. Your results might vary.

| File                          | Size   |
|-------------------------------|--------|
| PDF with no embedded fonts    | 239 KB |
| PDF with one embedded font    | 248 KB |
| PDF with two embedded fonts   | 250 KB |
| PDF with three embedded fonts | 250 KB |
| PDF with four embedded fonts  | 260 KB |

**NOTE:** A common typography best practice is to limit the number of fonts you use on web pages and documents to 2 – 3.

Bouton Jones

7

| File                          | Size   |
|-------------------------------|--------|
| PDF with no embedded fonts    | 239 KB |
| PDF with one embedded font    | 248 KB |
| PDF with two embedded fonts   | 250 KB |
| PDF with three embedded fonts | 250 KB |
| PDF with four embedded fonts  | 260 KB |

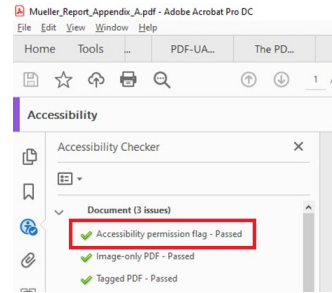
NOTE: This was a quick and dirty check --- not a rigorous scientific study. Your results might vary.

NOTE: A best practice in web design is to limit the number of fonts you use on web sites. It's probably a good idea to limit the number of fonts you use in a document whether embedded or not.

## Slide 114 - Accessibility Permission Flag

### Accessibility Permission Flag

Setting the flag ensures that screen readers and other assistive technologies can access, copy, and interpret text for users with disabilities, even when document security settings (like password protection) are enabled. It specifically prevents security restrictions from blocking essential access for text-to-speech tools.



Bouton Jones

8

Setting the flag ensures that screen readers and other assistive technologies can access, copy, and interpret text for users with disabilities, even when document security settings (like password protection) are enabled. It specifically prevents security restrictions from blocking essential access for text-to-speech tools.

## Slide 115 - Setting the Accessibility permission flag in Adobe Acrobat

### Setting the Accessibility permission flag in Adobe Acrobat

1. Open the PDF and select **All tools > Prepare for accessibility**.
2. Click **Check for accessibility** (or "Full Check").
3. In the results panel on the left, expand **Document** and locate **Accessibility permission flag**.
4. Right-click on "Accessibility permission flag" and select **Fix**.

Bouton Jones

9

1. Open the PDF and select **All tools > Prepare for accessibility**.
2. Click **Check for accessibility** (or "Full Check").
3. In the results panel on the left, expand **Document** and locate **Accessibility permission flag**.
4. Right-click on "Accessibility permission flag" and select **Fix**.

## Slide 116 - What is Else is new in PDF/UA-2?

### What is Else is new in PDF/UA-2?

- Offers improved support for annotations and complex document structures.
- PDF/UA-2 introduces changes to tagging rules, SVG, annotations, and structure types.
- Adds structure element attributes, annotation rules, MathML, structure destinations, and Associated Files.

Bouton Jones

1

- Offers improved support for annotations and complex document structures.
- PDF/UA-2 introduces changes to tagging rules, SVG, annotations, and structure types.
- Adds structure element attributes, annotation rules, MathML, structure destinations, and Associated Files.

## Slide 117 - New Tagging Elements & Structure

### New Tagging Elements & Structure

**New Tags:** Introduction of specialized tags such as <DocumentFragment> (for partial documents), <FENote> (Footnote/Endnote), <Title>, and <Aside>.

**Semantic Emphasis:** Enhanced support for tagging emphasis within text using <Em> and <Strong>.

**Improved <Figure> Tags:** Figure tags can now contain internal sub-structures, allowing for more detailed descriptions of complex diagrams and images.

**Formula Tagging:** Detailed requirements for mathematical expressions, including improved support for MathML.

Bouton Jones

2

**New Tags:** Introduction of specialized tags such as <DocumentFragment> (for partial documents), <FENote> (Footnote/Endnote), <Title>, and <Aside>.

**Semantic Emphasis:** Enhanced support for tagging emphasis within text using <Em> and <Strong>.

**Improved <Figure> Tags:** Figure tags can now contain internal sub-structures, allowing for more detailed descriptions of complex diagrams and images.

**Formula Tagging:** Detailed requirements for mathematical expressions, including improved support for MathML.

## Slide 118 – One Last Point: The PDF Association’s Recommendation

### One Last Point: the PDF Association’s Recommendation

#### PDF/UA: Creation tools

In principle, universally-accessible PDF documents can either be generated directly during document creation, using a suitable program, or indirectly generated by adapting an existing PDF document.

The indirect approach tends to require a great deal of work, as all tags and numerous other settings will need to be provided manually.

This work can also become void as soon as a new version of the PDF document replaces the old one, if any changes to the document’s content need to be made. The adaptation process must then be started again from scratch.

The direct approach is almost always the better option; post-creation PDF edits should be avoided or at least kept to a minimum. Either way, it is essential that the document creation program can perform the functions required and that the document creator can make use of them. Although only a few programs currently support PDF/UA in its entirety, a wide range of options still exist in the form of the programs listed alphabetical below:

CITE: <https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

Bouton Jones 1

Blockquote from “PDF/UA in a Nutshell”:

“In principle, universally-accessible PDF documents can either be generated directly during document creation, using a suitable program, or indirectly generated by adapting an existing PDF document. The indirect approach tends to require a great deal of work, as all tags and numerous other settings will need to be provided manually.”

“This work can also become void as soon as a new version of the PDF document replaces the old one, if any changes to the document’s content need to be made. The adaptation process must then be started again from scratch.”

“The direct approach is almost always the better option; post-creation PDF edits should be avoided or at least kept to a minimum. .”

CITE: <https://www.pdfa.org/resource/pdfua-in-a-nutshell/>

## Slide 119 - Additional Resources

### Additional Resources

- The PDF Association: <https://www.pdfa.org/>
- The PDF Accessibility group on Facebook:  
<https://www.facebook.com/groups/PDFAccessibility/>
- Tagged PDF: <https://taggedpdf.com/>

Bouton Jones

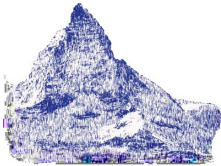
2

- The PDF Association: <https://www.pdfa.org/>
- The PDF Accessibility group on Facebook:  
<https://www.facebook.com/groups/PDFAccessibility/>
- Tagged PDF: <https://taggedpdf.com/>



## Slide 120 – Closing Summary

Closing Summary



Bouton Jones 1

Follow both sets of guidelines --- WCAG and PDF/UA --- and use the Matterhorn Protocol as a checklist for the PDF/UA.

## Slide 121 – Any Questions?



Slide 122 - Thank You. This concludes this session.



scott-jones

3

## Slide 123 - Appendices



Additional content for the Q & A.

## Slide 124 – A Video Showing How to Hack Improperly Redacted PDFs.



Watch the [View on YouTube](#)

Transcript:

I will now demonstrate how to hack a PDF that has not been properly redacted. First, I select the text that needs to be unredacted. Then, I press control C on my keyboard. This will copy the text. I open up Notepad. (That's a text editor that is part of Windows.) You can use any text editor to do this, but for simplicity's sake I'm going to use Notepad. I hold down the Control key on my keyboard and press V. That pastes the text inside Notepad. Because Notepad is a text editor, it strips it [the text] of all visual formatting. This is purely *text*. And you'll notice that the "redacted" name appears here. Well, this document is not properly redacted.

Instead of redacting the document, this user *masked* the content. One way to mask --- and I'm talking about masking here not redacting --- is just select the text --- that's a black text on a white background --- and change the background to black. Another method is to insert a shape over the object.

Again, this is *not* redacting the document. It's just covering the text with the shape, another layer. The text layer is unaffected. It still can be retrieved. It can be copied and pasted in either of these situations. Whether you have a black background or if you're masking the text, the text that you meant to redact is still there.

You can copy and paste it. You can also run a screen reader, and it would read the text unimpeded with no indication that it's meant to be redacted.

## Slide 125 - Redacting PDF Techniques



See <https://boutonjones.com/pub/hack.cfm>