

allyant

PDF/UA-2:

The latest on the new standard

AHG 2023

Paul Rayius



A few notes & comments

Before we get into the details...

allyant.com



Housekeeping

1. This presentation is not “all-encompassing”
2. This presentation is not “How to do this in Acrobat”
 - A. Support for some things has started (in some processors) but is not fully implemented yet
 - B. Once published, processor support will likely be more widely implemented
3. UA-2 does not replace UA-1
 - A. UA-1 is for PDF 1.7 files
 - B. UA-2 is for PDF 2.0 files

Release Date?

1. Final Draft International Standard (FDIS) has been sent for publication
2. ISO editors have to do their “thing”
3. It **must** be published by March 2024

Why update PDF/UA?

It was needed

1. Technology advances over time/ new functionalities (PDF/UA-1 was published in 2014)
2. For clarification/ to fill “gaps” in some areas
3. To make corrections in PDF/UA-1

PDF/UA is not all-encompassing

There's stuff PDF/UA doesn't do

1. Does say how to make stuff accessible/ doesn't regulate the content. For example:
 - A. Table summaries are not required
 - B. Tables of Contents are not required – if present, they are not required to be linked!
2. Some examples of what PDF/UA does not mandate:
 - A. Color use
 - B. Contrast
 - C. ECMAScript (JavaScript) may introduce accessibility issues
 - D. Text content, which may affect those with certain cognitive impairments
 - E. Deciding whether something is “real” content or “artifact” content (for tagging purposes)
3. Consult WCAG or other standards as needed

Let's get down to business...

allyant.com



Compliance

For a PDF to be PDF/UA-2 compliant it:

1. Must meet all requirements of ISO 32000-2 (the PDF specification)
2. Cannot contain anything deprecated in ISO 32000-2
3. Shouldn't use anything that's not in ISO 32000-2 and/ or in UA-2
4. Must meet all requirements of ISO TS 32005 (What?! We'll come back to this)
5. Must meet all requirements of PDF/UA-2 (Clause 8 - the real "meat" of the document)
6. Embedded files (if necessary for understanding the "main" document) must be accessible

What is ISO TS 32005?

ISO 32005 - inclusion rules specification

What is it?

“Inclusion rules” for tags - what can be inside
what

Why?

1. 32000-2 only has rules for tags in PDF 2.0
2. 32005 has rules for PDF 1.7 tags and PDF 2.0 tags
3. Meeting 32005 allows for the use of both types of tags

So, what's different between
PDF/UA-1 and **PDF/UA-2**?

allyant.com



Overview

1. Some tags (and other things) went away (not mentioned in 32000-2 or were deprecated)
2. Some new tags were added (in 32000-2)
3. Some tagging rules have changed (or been clarified, at least)
4. There's other stuff, too...

“Other” things

Before we get into tagging...

1. Deprecated: Sound and Movie annotations (probably not used all that often in PDF)
2. New things:
 - A. Projection and RichMedia annotations
 - B. Pronunciation help: “PhoneticAlphabet,” “Phoneme,” and “PronunciationLexicons”
 - C. “Short” (in TH) - a “short form” of long TH text

Tags not mentioned in 32000-2

1. Allowed/ required to use:

- A. Art (Article)
- B. Index
- C. Quote/ BlockQuote
- D. TOC/ TOCI
- E. Reference
- F. BibEntry
- G. Code

2. Should not use: Private

3. Prohibited: Note

Shiny new tags!

1. DocumentFragment
2. H# (where # is greater than 6) - no longer role mapped to P (Hn, in the spec.)
3. Title (tag) - NOT necessarily the same as in the metadata
4. FENote - replaces Note, as Note is no longer allowed
5. Sub (Subdivision) - identify semantic subdivisions within a Block level element
Example: legal docs with line numbering (Sub can contain nested tags)
6. Em (emphasis) - Example: “**cats** are cute” vs. “cats are **cute**”
7. Strong - to convey importance - for example, more important items in a list of tasks
8. Artifact
9. Aside

Some changes to tagging “rules”

1. Use the most semantically appropriate tag, regardless of content “encoding” (Figure or no?)
2. Headings
 - A. Allowed to skip heading levels to ensure proper semantics (like, in a DocumentFragment)
Use with Caution!
 - B. H tag (not numbered) - not allowed in PDF/UA-2 docs
3. Lbl - required for anything that “labels” anything else - including things like form labels and in lists (unless the label is an artifact - like, line numbers?)
4. Figure - can contain sub-structures
5. Caption
 - A. Uses - extended beyond tables and images
 - B. Placement (as a child)

More stuff...

1. Namespaces - a more “generic” role mapping
 - A. Some applications support this now
2. Support for MathML
3. Aria Roles on tags - helpful for 1.7 tags and/ or to give “richer” semantics
4. Declarations - conformity to WCAG, for example

Tables and lists

1. Tables

- A.** Shall be regularly shaped
- B.** There's an algorithm (from PDF 2.0) to help with headers

2. Lists

- A. Continued/ Continued From
- B. ListNumbering
 - 1) New attributes: Unordered, Ordered, Description
 - 2) Required for ANY list that has Lbls
 - 3) "None" is not allowed

Annotations and navigation

1. Reference tags can contain link Annotations directly (TOC tagging, footnotes)
2. Ref entry and structure destinations
 - A. Ref entry required in TOCs and Reference/ footnote “combinations”
 - B. Internal destinations/ link annotations must use Structure Destinations
3. Annotations
 - A. Contents entry - not necessarily “mandated” but a “should” in a lot of scenarios and a “shall” in others (depending on annotation type and context, etc.)
 - B. Change to Tab Order requirement - S (Structure), A (Annotations), or W (Widgets - two “passes,” widgets first then other annotations)

Forms: tagging

1. Lbl - within Form tags, contains the text label for the widget (Form tag contains Lbl and widget)
2. One widget (form OBJR) per Form tag

Forms: tooltips, contents, and a functionality note

1. TU (tooltip):

- A. UA-1 - required
- B. UA-2 - can provide more context - considered more “supplementary” in UA-2 files
- C. Often insufficient, especially in a field with multiple widgets (set of radio buttons)
In those cases, label, Contents, or both are required

2. Contents entry:

- A. Required if the label is missing or insufficient
- B. Should not be the same text as the label

3. Export values (on radio buttons and checkboxes)

- A. Are not intended to be descriptive
- B. Should not be used by AT (this is a major change to current real-world functionality)

Other notables

1. A PDF cannot always be both PDF/UA-1 and PDF/UA-2 compliant (due to some changes to 32000-2 and PDF/UA-2)
2. Some examples/ reasons
 - A. Differences between the PDF 1.7 and PDF 2.0 standards
 - B. Document tag, in the 2.0 namespace, is required as the only child of the Tags root
 - C. Namespaces (not in v.1, required in v.2)
 - D. Structure Destinations (not in v.1, required in v.2)
 - E. Use of Figure tags for graphics (required in v.1, not in v.2)
 - F. Use of H tag (allowed in v.1, not allowed in v.2)
 - G. Reference tags containing (local) Annotations (in UA-2)
 - H. ListNumbering (handled differently between the two versions)

If you want to chat after...

Visit the Allyant booth in the vendor hall!

allyant

simple // seamless // accessibility

allyant.com