

May 17-18, 2006

**PDF/UA (Universal Accessibility)
Working Group Draft Meeting Minutes**

Face to Face / Teleconference

Philadelphia Convention Center
1101 Arch Street, Philadelphia, Pa 19107
<http://www.paconvention.com/>

Wednesday, May 17 9:00 AM - 4:30 PM, Room 105A

Dial-In Instructions:

888-742-8686 (US)

303-928-2600 (International)

Conference ID: 7083945

Thursday, May 18 9:00 AM - 12:00 PM, Room 106B

Dial-In Instructions:

888-742-8686 (US)

303-928-2600 (International)

Conference ID: 5991077

NOTE: All attending members should load the [PDF/UA - Universal Accessibility](#) page.
We'll be working "live" throughout the meeting!

Agenda for Wednesday, May 17

Attendees

Duff Johnson - DJ, Chair Document Solutions

Bruce Bailey - BB - US Dept. of Education

Angy Brooks - AB - Independent Accessibility Consultant

Mark Gavin - MG - Appligent

Loretta Guarino-Reid - LGR - Adobe Systems

Richard Herring - RH - Independent Accessibility Consultant

Neil Soiffer - NS - Design Science

[Shumpei Tamaki - ST - Ricoh Japan \(observing\)](#)

Action Items indicated with an AI:

Duff Johnson opened the meeting at 9:13

Approval of PDF/UA 2006-115, May 17-18, 2006 Agenda [2006-115 Agenda](#)
Guarino-Reid / Soiffer

Introductions

Loretta Guarino-Reid, Mark Gavin, Richard Herring, Greg Pisocky, Bruce Bailey,
Representative from RICOH - Japan Japanese Strds, Ricoh Co. Ltd., TC-171, Neil Soiffer,
Angy Brooks, Duff Johnson

Approval of PDF/UA 2006-114, May 3, 2006 Minutes [[Meeting Minutes](#)]
Approval deferred until all had a chance to examine.

Report on / update and discuss [PDF/UA - Action Items](#)

The list of open/completed action items was updated.

Complete review and deletion of headings on the [PDF/UA - Working Draft](#)
Font rendering was deleted

Mark Gavin has to leave at 10:30 for an hour.

Skipping action items for now, addressing section deletion while Mark Gavin is present.

Low-level font rendering, rendering, and printing sections were deleted.

Work through "headings without pages", brainstorm and add notes to each to "start" the page

AI:

Orga and use of fonts - 50% complete

Scripting:

DJ: Tom Parker, Bryan Guignard, accessible scripting. Doc Solutions is working on scripting issues. Linear forms. Possibilities in script.

AI: GP: Contact Lori DeFurio and get the big sriptors. Scripting that changes content on the page . How to address the issue. Treat the section like others, make notes.

DJ: When we get to scripting notes regarding types of concerns 1) Javascript can cause page content to change.

DJ: Scope of the Wiki, reference maximally. Change maximal to optimal. Change to "as

accessible as possible.”

The Welcome paragraph was changed to read,

“Welcome to the PDF Universal Accessibility Committee. This committee is developing a specification for accessible PDF. The Committee’s goal is to set standard for PDF authoring such that conforming PDF files are accessible and usable to all including those who use assistive technologies.”

Moving on to the working draft and adding content to each section heading.

Table of Contents

- [Cover](#)
 - [Copyright Page](#)
 - [Foreword](#)
 - [Introduction](#)
1. [Scope](#)
 2. [Conformance](#)
 3. [Normative references](#)
 4. [Terms and definitions](#)
 5. [Symbols \(and abbreviated terms\)](#)
 6. [Technical requirements](#)

1. [Syntax](#)

Note: Do we want to say anything about the encrypting of metadata.

1. [Lexical Conventions](#) (likely to be deleted - check with Mark Gavin (potential deletion))
1. [Character Set](#) (the PDF file and what the bytes are and how they are interpreted).
2. [Comments](#)
2. [Filters](#)
1. [Crypt Filter](#) Note: discuss whether or not to allow encrypted xml metadata
3. [Encryption](#)

AI: RH to address this

LGR: Pointed out the distinction between encryption / security. Encryption vs. Security. “Any Encrypted PDF must be accessible.”

4. [Document Structure](#)
1. [Document Catalog](#)
2. [Page Tree](#) (page tree content belongs in Document Catalog)
3. [Name Dictionary](#)
5. [Content Streams and Resources](#) (Delete (MG - it’s a catch all - delet content streams and resources))
1. [Content Streams](#)
2. [Resource Dictionaries](#)
2. [Graphics](#)

1. Path Construction and Painting (delete Path Construction and Painting) (Deleted Make clear that items are in the PDF spec are maintained - this spec documents addendums for the sake of accessibility.)
 1. Path Construction Operators
 2. Path-Painting Operators
 3. Clipping Path Operators
 2. Color Values (Not about use of color, it is about use of what color To Be Deleted)
 3. External Objects-may want to disallow from PDF/UA
 1. PostScript XObjects
 4. Images (Mark for Delete)
 1. Image Parameters
 2. Sample Representation
 3. Image Coordinate System
 4. Image Dictionaries
 5. Masked Images
 6. Inline Images
 5. Form XObjects (Mechanism for sharing items, AI: verify LGR)
 1. Form Dictionaries
 2. Group XObjects
 3. Reference XObjects
 6. Optional Content (layers: What happens for accessibility when layers are turned on/off - interaction with tagging, change of state issues, layer interactions, presence of layers)
 1. Optional Content Groups
 2. Making Graphical Content Optional
 3. Configuring Optional Content
 3. Text (GP)
 1. [Organization and Use of Fonts](#) (Pages at the lowest point of the tree: AI: GP)
 1. Basics of Showing Text
 2. Achieving Special Graphical Effects
 3. Glyph Positioning and Metrics
 2. Text State Parameters and Operators (Invisible text, outline text move to fonts GP)
 1. Text Rendering Mode (MG)
 2. Text Knockout (MG) (Cropping - delete)
 3. Simple Fonts
 1. [Character Encoding](#) (heading with an idea)
 4. Extraction of Text Content (Note refer to tagging chapter to see if adequately addressed - discussion of how to represent languages that do not have character encodings)
 1. Mapping Character Codes to Unicode Values (shall will be needed here. Character codes must have a match to a ToUnicode)
 2. ToUnicode CMaps
 5. [Non-linguistic Notation](#) (Content begun)
 6. [MathML](#) (Defer to this afternoon's discussion) Jacques Distler Incorporate PDF tags in the PDF generation setup. LaTeX, PDFTeX. At 3:05 Jacques Distler joined the discussion. Perfectionist /

ameliorist stance.

Notes: Annotate what the conclusions are.

A. Ameliorist vs. Perfectionist - LGR Perhaps establish a list of priorities. NS If it can be done in MathML in an accessible and appropriate fashion, then that should be the preferred method. How to embed both the tags and the namespace for interpretation. NS: 2 different approaches, namespace, tag collision. Tag it any old way - serves no purpose. How to interpret. Or extend the PDF tag set. If it is Math, then these are the tags you should use. Our discussion is how to support them. Audience consideration text books. For the Blind, if there is a Braille code that exists and you do not translate to those Braille codes, the blind do not consider these accessible. Regarding MathML, we should adopt at least a "should". For all types of equations. The requirement is At a minimum "All math shall be tagged as Math, and the tagging should be with MathML" The remaining point of interest, and explore spaces where we shall use MathML. The tag that shall be required is <Formula>.

Can attributes be used for this? Other interesting paths to pursue include contemplate how to improve upon the inefficiency of tags.

What's not on the table... Conformance levels. Simple / Complex issue.

The issue of namespace comes into play here. It won't take long before the tag names are repeated.

AI: GP to reinitiate correspondence with Peter Murray Rust to replicate this effort with Chem ML

4. Transparency (LGR: Effects of transparency for low vision, color contrast issues,)
5. Interactive Features
 1. Viewer Preferences (These are document preferences which control the way a document is presented on screen or in print. These are document settings not the application. Subjects under this heading include AI: DJ Bookmarks, mag levels, etc.
 2. [Document-level Navigation](#) (Place regarding bookmarks, (Duff to move to lower branches)
 1. Destinations (Determined for us)
 2. Document Outline
 3. Thumbnail Images
 3. Page-Level Navigation
 1. Page Labels (Could be a DAISY issue. Refer to George K. Page level, page index.
 2. Articles (If used, use in a way that is consistent with the structured reading order)
 3. Presentations (slide show, page duration, control time to read a page need a way to override a duration parameter. This brings up issues of time out, timed response).

Agreed, work through 1:00 then break.

4. Annotations (a lot to be said, for any given annotation what has to be in the PDF. AB to extract annotation info. LGR to do further brain storm)
 1. Annotation Dictionaries
 2. Annotation Flags
 3. Border Styles

4. Appearance Streams
5. Annotation Types
5. Actions (Already begun, move forward)
1. Action Dictionaries
2. Trigger Events
3. Action Types
6. Interactive Forms (Held over from parts of WCAG)(AB to provide links) -

TU text string (*Optional; PDF 1.3*) An alternate field name to be used in place of the actual field name wherever the field must be identified in the user interface (such as in error or status messages referring to the field). This text is also useful when extracting the document's contents in support of accessibility to users with disabilities or for other purposes (see Section 10.8.2, "Alternate Descriptions").

1. Interactive Form Dictionary
2. Field Dictionaries
3. Field Types
4. Form Actions
5. Named Pages (delete)
6. Forms Data Format (delete)
7. XFA Forms (Huge task, PDF spec points to XFA spec) (GP: to do the XFA forms sections)

Lunch Break

7. Digital Signatures (the guidelines for Digital Signatures are similar to those being used by formfields - treat the Dig Sig as yet another type of form field).
 1. Transform Methods
 2. Signature Interoperability
 3. Permissions
 4. Legal Content Attestations
8. Measurement Properties (Measurement Page level - nothing to say which differs from the PDF Specification - to be removed)
9. JavaScript (The javascript has content in it, and discussion can continue)
6. Multimedia Features
 1. Multimedia (values that are important to accessibility) (AI: LGR to review Multi Media features)
 1. Viability
 2. Renditions
 3. Media Clip Objects
 4. Media Play Parameters
 5. Media Screen Parameters
 6. Other Multimedia Objects
 1. Sounds (suggest requiring a Text equivalent description at the level of the sound object)
 2. Movies (refer to multi-media section, mark for deprecation, The PDF Spec deprecates this)
 3. Alternate Presentations (deleted)

4. 3D Artwork (Rotation. Limited Mobility, Requires further discussion)
1. 3D Annotations
2. 3D Streams
3. 3D Views
4. Coordinate Systems for 3D Annotations

(Adjourned at 4:40 - Discussion to Resume Here Tomorrow)

Day Two: May 18, 2006

The previous meeting minutes were approved

Approval of PDF/UA 2006-114, May 3, 2006 Minutes
Johnson / Guarino-Reid

Discussion resumed with document interchange

7. Document Interchange
 1. [\[pdfua-metadata\]](#)
 1. Document Information Dictionary
 2. Metadata Streams
 2. Marked Content (deleted)
 1. Property Lists
 2. Marked Content and Clipping
 3. Logical Structure

From the PDF Reference:

PDF's *logical structure* facilities (PDF 1.3) provide a mechanism for incorporating structural information about a document's content into a PDF file. Such information might include, for example, the organization of the document into chapters and sections or the identification of special elements such as figures, tables, and footnotes. The logical structure facilities are extensible, allowing applications that produce PDF files to choose what structural information to include and how to represent it, while enabling PDF consumers to navigate a file without knowing the producer's structural conventions.

PDF logical structure shares basic features with standard document markup languages such as HTML, SGML, and XML. A document's logical structure is expressed as a hierarchy of *structure elements*, each represented by a

dictionary object. Like their counterparts in other markup languages, PDF structure elements can have content and attributes. In PDF, rendered document content takes over the role occupied by text in HTML, SGML, and XML.

A PDF document's logical structure is stored separately from its visible content, with pointers from each to the other. This separation allows the ordering and nesting of logical elements to be entirely independent of the order and location of graphics objects on the document's pages.

1. Structure Hierarchy (Delete the heading)

Reference to the tree like structure that is displayed in Acrobat's tags panel. Logical structure is at a lower level than the Tags. Tags are the values that get put into logical structure.

LGR: Relationship between content, logical structure and reading order. Content data structure used for rendering the page. Order in which items get marked as the page is rendered. Logical structure is a separate tree. It is a tree, one per document. A substructure can be built of logical structure nodes. Leaves function as pointers to content and other objects. Tags are items related to logical structure. A walk through the structure tree defines reading order. (Source of two reading orders when examining a PDF file, walking the tree differs from how the page appears).

Tagged PDF defines a particular use of the logical structure mechanism.

Logical structure is similar to XML, no meaning until a grammar is placed on top of it. Tagged PDF can be thought of the grammar that sits atop the logical structure.

Logical structure with any random set of tags implies a reading order and grouping (heirarchy).

See PDF Reference Section 10.6

Important to note: The use of the term tags refers to the use of tags to represent logical structure as opposed to an emerging use of the term where the term tagging is used to apply meta data information to a PDF file.

Important to note, for the purposes of accessibility some items indicated as Optional for PDF fidelity will be required in order to achieve accessibility.

The sub level heading elements to Logical structure can be deleted for the purposes of our discussion, but the heading Logical Structure is to be maintained.

2. Structure Types

1. Grouping Elements
2. Block-Level Structure Elements

1. Headers
2. Lists
3. **Tables**
4. Inline
5. Link
6. Annotation
7. Ruby and Warichu

3. Illustration Elements

3. Structure Content
4. Structure Attributes
5. Example of Logical Structure

4. **Tagged PDF**

Reconsider wording of the statement based on proposed changes to tags and attributes.

1. **Tagged PDF and Page Content - Tables page moved here but is more appropriate under Standard Structure types. One item of note MG has pointed out, where the PDF reference stipulates optional behavior that should be examined and deemed required for accessibility purposes. For example, on Page 814 the following note concerning artifacts appears:**

PDF's *logical structure* facilities (PDF 1.3) provide a mechanism for incorporating structural information about a document's content into a PDF file. Such information might include, for example, the organization of the document into chapters and sections or the identification of special elements such as figures, tables, and footnotes. The logical structure facilities are extensible, allowing applications that produce PDF files to choose what structural information to include and how to represent it, while enabling PDF consumers to navigate a file without knowing the producer's structural conventions.

The term *should* should be changed to *shall*.

Note: To aid in text reflow, it is recommended that artifacts be defined with property lists whenever possible. Artifacts lacking a specified bounding box are likely to be discarded during reflow.

PDF logical structure shares basic features with standard document markup languages such as HTML, SGML, and XML. A document's logical structure is expressed as a hierarchy of structure elements, each represented by a dictionary object. Like their counterparts in other markup languages, PDF structure elements can have content and attributes. In PDF, rendered document content takes over the role occupied by text in HTML, SGML, and XML. A PDF document's logical structure is stored separately from its visible content, with pointers from each to the other. This separation allows the ordering and nesting of logical elements to be entirely independent of the order and location of graphics objects on the document's pages.

2. Basic Layout Model - Section (at this point in time the PDF Reference is sufficient)

3. Standard Structure Types (someone needs to add questions). Concrete document structural education. Focus on required tags. RH: go through today's spec and determine what is sufficient and simply being ignored. Requests for addition to the tag set, may need to clarify, how to add MathML.

AI: RH to examine the current language, determine what is sufficient. What is minimum: Tables, lists, headings, "Tables must be tagged as Tables"

Headings...

Problem list: Relationship of currently tagged as artifacts under current page model and how clients can be structured.

Logical Read order

BB: At a minimum, PDF/UA shall conform to the tagged PDF specification.

A list of can and cannot actions also needs to be developed... i.e., alt text on nodes

Alternate Text should not incur on a parent node in case the node would include interactive elements.

The current minimum list appears to be adequate.

4. *Standard Structure Attributes Page created. Think about preliminary copy for this entry. Leave Standard Attribute owners section as is.*

5. *Accessibility Support*

1. *Natural Language Specification - existing page (Linguistic vs. Non Linguistic considerations.*
2. *Text Equivalent - existing page*
3. *Replacement Text - Page created, optional and required elements to be specified.*
4. *Expansion Of Abbreviations And Acronyms*
5. *Tab Order*
6. *[pdfua- Focus]*
2. *Keyboard*
3. *Others*
1. *Content/Structure Separation*
2. *Visual And Audio Contrast*
3. *Time Limits*
4. *Seizure*
5. *Minimize Errors*
6. *Meaning*
7. *Media Equivalent*
8. *Other Possible Headings (need to be organized)*
9. *Color*
10. *Contrast*
11. *[pdfua- Text Formatting]*
1. *Spacing (discuss negative)*
2. *Read Order*
 - *Annex A (normative) Annex Title*
 - *Annex B (informative) Annex Title*
 - *Bibliography*

Review of Action Items

Duff Johnson reviewed the list of outstanding action items and updated the action item page accordingly.

Accept the Annual Report

Guarino-Reid/Gavin

Upcoming Meeting Schedule

Next Meeting May 31, 2006 3:00 pm EST

June 21, 2006 3:00 pm EST

July 26, 2006 3:00 pm EST

August 8, 2006 3:00 pm EST

August 23, 2006 3:00 pm EST

Sept. 06, 2006 3:00 pm EST

Sept. 20, 2006 3:00 pm EST

Oct. 04, 2006 3:00 pm EST

Oct. 17-18, 2006 3:00 pm EST - tentative face to face. Toronto perhaps.

Goal for the next Face to Face: A reviewed statement under each item in the Table of Contents.

Adjourn

Meeting Adjourned 12:10 pm EST.

Johnson / Herring