

**PDF-Access Meeting
Draft Meeting Minutes
March 8-9, 2005**

**National Federation of the Blind
Baltimore, Maryland**

Those in attendance included:

Bruce Bailey, U.S. Department of Education
Caesar Eghtesadi, Access World Solutions
Betsy Fanning, AIIM
Mark Gavin, APPLIGENT, Inc.
Virginia Gavin, APPLIGENT, Inc.
Brad Hodges, National Federation of the Blind
Michele Novak, Social Security Administration
Greg Psocky, Adobe
Lorretta Guarino Reid, Adobe
Dariusz Tybor, Social Security Administration

Via phone:

Renee Georges, AIIM
Richard Herring, Avaya
George Kercher, DAISY Consortium
Warner Krebs, Design Science
Bob Montengro, Alt Text Production
Sailesh Panchang, Deque
Skip Stahl, Cast
Neil Soiffer, Design Science

The March 8, 2005 meeting of the PDF/Accessible committee was called to order at 9:05 a.m.
The meeting participants introduced themselves.

The draft agenda (PDF-AC 2005-004) for the March 8-9, 2005 meeting was approved as presented. (Hodges/Guarino-Reid)

The draft meeting minutes (PDF-AC 2005-002) for the January 11-12, 2005 meeting was approved as presented. (Guarino-Reid/Hodges)

Each of the working groups presented status reports from the last meeting on their discussions.

Accessible PDF Characteristics Working Group

The 508 standards do not describe a standard. The US Department of Education is building a set of characteristics and identifying standards for documents. Their work is based on the 508 standards. The committee agreed that they want to leverage as much as possible the work already underway in many organizations to define the characteristics.

The committee reviewed the Department of Education's characteristics:

a) Images, photos, charts, graphical text and other non-text elements must include either textual equivalents in the document body or appropriate formatting such as alternative text.

In PDF there may be some things that will need to be ignored, however, the spirit of this is correct. This applies to PDF with some exception. Users need to ask if it is part of the content.

b) Movies and multimedia presentations containing audio dialog and cues must include synchronized captions.

Captions for deaf and scripted video for blind – committee discussed if they would be distinguishing between the two types and identified the need to address the unique characteristics of both. The committee discussed if PDF was the best way to present multimedia files. For the time being, the committee is discussing PDFs and their requirements. If the PDF includes a poorly constructed multimedia piece, then the file is not accessible. It was agreed that embedded objects need to satisfy the requirements. PDF supports the encapsulation of web pages including HTML. It was noted that we need to make sure there is an accessible playback mechanism. HTML supports external players and objects. Language can be used to say that a title exists. The specification needs to support long titles.

c) Audio description is required for music, sound, and non-verbal content when necessary to fully understand the information being conveyed.

Talking media to media – The intent is for a movie that is playing would have a secondary audio track describing the actions happening. This feature should be selectable or deselectable. The feature needs to be included but we need to revise the text to clearly describe what we want.

d) Blinking or flashing objects (including text and other elements) must cycle slower than three times per second or faster than 54 times per second.

Numbers identified because they are in 508 and WCAG.

e) The user of color is acceptable, however, documents must be designed so that all information required for meaning is not dependent on the ability to identify color.

agreed.

f) Foreground and background color and pattern combinations must provide high contrast.

agreed

g) The underlying structure of the document, which is relied upon by assistive technology, must provide a reading order that is equivalent to the visual layout.

The committee discussed if there was an underlying order that needs to be considered. The technique of chaining in PDF was also discussed. It is possible to have a file that contained an implicit order. The document needs linear structure or a chaining mechanism is required that allows it to be defined. A requirement that an unambiguous reading order is required was agreed to. The committee also agreed that structural elements need to be identified for navigation purposes.

h) Table creation tools integrated within the document authoring application must be used for presenting tabular data. Row and column headings must be identified when a method is available within the authoring tool, otherwise an alternative accessible version of the tabular data must be provided.

Charts would be graphical representation of tabular data.

i) Interactive documents, including form fields, controls, and dynamic content, must be completely useable without requiring the operation of a mouse or other pointing device and must be fully functional using only the keyboard. All elements must be identified with functional text that can be read by assistive technology.

The committee discussed how tax forms are addressed for accessibility. We need to specify the tools we are dealing with.

j) If the document cannot be made accessible in its original format, an accessible alternative version with equivalent information and functionality must be provided. The alternative version must be updated as often as the inaccessible original document, and the inaccessible document must make reference to the availability of an alternative version.

The committee agreed that this requirement is difficult and may not be wanted. Many forms are not covered by 508 but are intended to be completed online.

The National File Format technical committee produced a list of 5 principles that define user needs:

1. Should be sufficiently extensible to support transformation into alternative formats. E.g., should support the inclusion of layered content elements.
2. Support inclusion of all media that are contained in the print work. All non-text elements and all media included in the print work should be included and referenced in the NIMAS standard. Support the inclusion and contextual intent.
3. NIMAS should provide for the incorporation of structural information about the print work, e.g., chapters, sections, headers, etc. In other words, preserve the structural information that is represented in the layout.
4. Format should facilitate the efficient and automated conversion of print material into more accessible formats.
5. Should accommodate all fields of study including those that require special notation, e.g., math, music, etc.

Special Treatment Objects Working Group

The committee discussed adding a group to investigate user agents. Consensus of the committee is that we need to review standards for user agents.

The committee discussed explicitly labeling text that is embedded and should be accessible. It was determined that the writer should identify this information.

Focus (where to find the content you are currently reading) is a remaining problem. Issues exist around magnification tools. Readers read a paragraph and they get lost. The Social Security Administration is considering using a JAVA script in their documents to help with the focus.

The committee discussed if the existence of some sort of cursor location needs to be added to the specification. Conditional tabbing is an important issue.

The committee discussed the issue of radio buttons in Acrobat and that they do not act as they want them to be. Checked boxes are acting as radio button which requires the end user to label the field appropriately.

With regard to tables, DAISY has additional markup for tables, groups of columns, and repeating headers. Headers and footers are treated so that HTML can automatically repeat those on the repeated page. The committee discussed if the column grouping is different from column spans. The column group is to note visual groupings of columns in a table. The content and table models are from HTML. When DAISY borrows an element from another organization, DAISY never changes the semantics. The markup never drives the rendering. Here we are going from visual presentation to semantic. In DAISY, we are going from visual presentation to semantic and trying to capture the information.

PDF has a built in facility to export these formats. The committee discussed if Adobe had plans to make DAISY one of these standard export formats.

Mathematics is another object that needs to be considered. As XML areas are identified, we can leave a door open to leverage the work done in other organizations. DAISY will soon have an extension for mathematics using MathML. The committee may want to associate sub parts of XML and content. To do this, it is necessary to understand the relations on the page and may require a mechanism to do it. For example, MathML can be associated with the PDF formula tag and sub expressions in the MathML can be associated with tags inside the formula tag. PDF currently treats the formula tag similarly to a figure and is not accessible.

For modular extensions, if an existing standard exists, like MathML, we should use it but provide supporting materials for an endorsed modular extension for things like music or forms.

Everyone wants to read the documents natively using a PDF reader. It is believed that all the elements that are defined within DAISY's tagset are designed to make print accessible. It is a conversion DTD. There are individuals who want to export content from PDF and produce alternatives. An export mechanism is necessary. It will be difficult to test whether a document is accessible.

A comparison of the DAISY and Adobe PDF results in a small set of common tags. There are not always equivalents or approximate equivalents such as long descriptions. A table version of a pie chart may need to be included in a long description. There are issues with side notes and notice.

The committee needs to define what it means to create a document that is accessible. It also needs to look at annotations to determine what can be made accessible.

AIIM is to prepare an outline of the sections of PDF/A and PDF/E for the committee to review and begin to think of the structure of the PDF/Access standard.

The committee needs to review how W3C handles external references.

There are objects that are part of the file itself. There are functional requirements on them and their ability to interact with them. Perhaps objects in HTML might serve as the analogy that may be a place to look for a model on how to do this.

With regard to forms, we need to determine if we will allow extra pages be added to the form. Additional pages can be added to a PDF form with JavaScript. Basically, the committee needs to decide if they want to allow any scripting language. Dynamically modifying content is somewhat problematic for the current assistive technologies because of the complexity to edit the form.

If tagging is a pre-requisite to an accessible document, we will need to include the requirement. Additional discussion is needed on reading order which is required with tagging. We need to ask if accessibility requires tagging.

The committee discussed annotations. In its simplest form it is a post-it note. Other types of annotations highlight, underscores, strikethroughs, etc. Their existence can change the meaning of the content on the page. Another class is form fields which can also be buttons with slightly different behaviors. There are a number of different types of annotations that may alter the meaning of the page. There are allowances for annotations in the tag stream. Annotations are not literally part of the page content. They are in a separate location inside the PDF file. The committee needs additional discussion on static annotations and how they should be handled. In general, the committee needs to identify the accessibility requirements of annotations.

Forms can be accessible. A form field is a type that is called a widget. It has more information than a classic annotation such as a strikeout. A strikeout knows the position "quads" but not what is being struck out. The highlighter for example does not typically know what it is highlighting.

Annotations have bounding boxes. Page 842 of PDF Reference 1.6 discusses these annotations. It may not appear to refer to annotation elements. Annotations have locations.

Bookmarks are not contained in the structure. They are separate elements.

The committee is assuming that what is being done is based on the implicit tagged PDF definition and discussed if this is a safe assumption. Tagged PDF requires that the print stream be in reading order. A number of legacy documents will need to be reordered.

Some of the conformance information is defined in PDF/A. The committee discussed using PDF/A conformance and agree dot use PDF/A's words.

PDF/Accessible needs to specify the property of fonts. For example, type 3 fonts are problematic. Rules for fonts will be needed as users will most likely not know about the proper fonts.

The committee is to review PDF Reference 1.6 and identify those items that would not be allowed in PDF/Access.

PDF/Access will reference PDF Reference 1.6. The committee discussed how to handle the deficiencies in 1.6 and identify recommendations on how to handle them. Assistive technologies vary in their ability to do various things. PDF HTML, PDF and DAISY are all three expressive domains that appear to be linked in some way. It might be practical to address as many of the accessible constructs of HTML.

Actions:

1. Identify desired components in the current PDF Reference.
2. AIIM to prepare a strawman outline document.
3. Check with publishers of DTBook
4. AIIM to set up discussion group and amend web page.
5. Make recommendations to PDF Language Group – Greg and Lorretta
6. Mark Gavin to send his presentation on enhancements to PDF Reference
7. Committee members to bring examples of publications/textbooks that would not lend themselves to be accessible.

The meeting was adjourned at 11:20 a.m. (Bailey/Novak)