LaTeX to PDF Tagging Issues Hit List

10/27/14 Margo Miller

The following list is a prioritized list of tagging issues to be resolved in LaTeX-to-PDF conversion:

# PDF Tags Critical Elements

## Tagged Content

1. All content selection groups for the October 27th version of the PDF need to be corrected back to the selection method in the July PDF. Image examples below of what the PDF should look like:

|  |  |
| --- | --- |
| GOOD Content Grouping: | BAD Content Grouping: |
|  |  |

1. ALL content needs to be tagged with either a valid tag or as an artifact.
   1. Watch for equation/formula numbers, page numbers, decorative document elements, etc.
2. Paragraphs, headings, and figures must all be tagged correctly (e.g. paragraphs cannot be tables, figures and codes mapped to the <Formula> tag cannot be form fields, etc.)
3. Some type of character is causing empty tags to be generated – need to identify.
4. Paths should be tagged as artifact unless part of a figure – in which case they should be inside the <Figure> tag.
5. Hyperlinks should go to the proper destinations (Web links, cross-document links, bookmarks, footnotes, etc.)
   1. *Hyperlinks to external web pages, or to cross-document content are mostly generating the OBJR link which is good. The link is not tagged and will report a tagged annotation error, but that particular error is addressed in the “Nice to Correct” section below.*

## Headings

The title needs to have an **H1** tag

All subsequent headers should be numbered according to proper nesting (e.g. 1.0 Introduction is **H2**, 1.1 Overview is **H3**, 1.1.1 Project History is **H4**, etc.)

## Bookmarks

Bookmarks should be automatically generated, with proper nesting, and should be the same as the heading tags.

## Fonts

Check to make sure the templates are using the NREL standard fonts as defined in the Style Guide. Other “generated” font types will show up in the document, and that is fine as long as all fonts display “embedded subset”.

## Tables

ONLY tables should use the table tags (<Table>, <TR>, <TH>, <TD>). Table tagging should not be used to illustrate document layout. The correctness of the table tagging is addressed in the “Nice to Correct” section below.

# Nice to Correct

## Tables

Table tags all need to be nested inside the original <Table> tag. Table caption may be right before the table, or as the first value of the tag – either way is fine.

Table data cells need to accurately represented as unique values, not lumped into a single string.

## Figures

Items tagged as Figures should have the alt-text value entered. Is there a way to have this alt-text carried through from LaTeX? If not, Author can use the Alt-Text tool to quickly navigate from figure to figure to enter these values.

## Lists

List Item <LI> tags need to be nested under the <L> tag, not under another <LI>.

## Tagged Annotations

Linked content should generate a tagged annotation for hyperlinks.

# Final Steps in Acrobat

These final checks will need to be done in Acrobat, if not already set from conversion. Provided the PDF has all the critical elements corrected, these steps could be performed by Publication Services so the author/communications rep would not need a license for Acrobat Pro.

## Document Properties

Description tab - **Document Title** must be present. It appears the title is being transferred to the converted PDF, so this is good.

Fonts tab - All fonts must show **embedded** subset

Initial View tab:

Navigation tab:

**Pages Panel and Page** for documents under 16 pages, with Page Thumbnails embedded.

**Bookmarks Panel and Page** for documents 16 pages or larger.

Page layout: **Single Page**

Magnification: **Fit Width**

Window Options – Show: **Document Title**

Advanced tab: Reading Options Language: **English**