Preparing PDF Files for ALSTAR

White Paper
June 16, 2008

Preparing PDF Files for ALSTAR

Summary

Two issues face registrants participating in ALSTAR.

1. What constitutes the "label" for ALSTAR?

ALSTAR defines the label as the container label components which hold all text and graphics required to accompany the product when sold. For example the label might include 1) a front-panel, a back-panel, and a supplemental booklet. The label would not include the carton unless the product was required to be sold in the carton.

2. How should PDF files be created for ALSTAR?

PDF files submitted to ALSTAR should be created directly from the authoring software and contain no PDF document security. Additionally, PDF files should conform to the PDF/A-1b specification but this is not an absolute requirement. ALSTAR disallows scanned labels but exceptions may be made on a case by case basis.

Technical Discussion: What constitutes the "label" for ALSTAR?

Even though states often have their own definitions of what they require to be submitted for state licensing, ALSTAR defines the label as the container label components which hold all text and graphics required to accompany the product when sold. For example the label might include 1) a front-panel, a back-panel, and a supplemental booklet 2) front and back of a bag 3) a front-panel, a back-panel fix-a-form booklet, and base. The label would not include the carton unless the product was required to be sold in the carton.

Normally a printer's proof of each container label component is prepared as a separate PDF file by a typesetter using desktop publishing software such as QuarkXpress or InDesign. The registrant must assemble the individual components into a single PDF for submission to ALSTAR.

Exception: Today some registrants submit 8.5 x 11 representations of their container labels to states rather than printer's proofs of the actual container label. If this kind of label has been sufficient in the past, it may be submitted to ALSTAR.

Technical Discussion: How should PDF files be created for ALSTAR?

This section suggests some best practices for PDF creation which can be tested and perhaps formally adopted at a later date.

PDF files can be created quickly and easily from most label authoring software, for example QuarkXpress, InDesign, and Microsoft Word. Sometimes the ability to create PDF files is built into the authoring software itself. But in most cases a separate PDF creation application works together with the authoring application to create a PDF file. Behind the normal default settings, there are many PDF creation options which can affect how the resulting PDF performs. For example a PDF file with many colors and images and intended to be printed at high resolution would use different settings than a black and white business document.

The most familiar PDF creation tool is Adobe Acrobat . This discussion offers explanations in terms of Acrobat 8 Professional's functionality. There are many other PDF creation tools – some free – which do most or all of the things discussed here.

Three Methods to Create PDF Files of Product Labels

1. The Best Option. Create PDF files directly from the authoring software and comply with the PDF/A-1b specification. PDF/A-1b is a recently approved ISO standard for long-term electronic document archival. This is a small change in the way most registrants already create PDF labels.

What is PDF/A and why is it the best option? PDF/A is an electronic file format created for the long-term archiving of electronic documents. It was published as an ISO standard in 2005. In concept, it is similar to standards requiring the use of certain kinds of paper for long-term archiving of paper records. PDF/A describes a subset of features found in the PDF file format. It disallows certain features not suited to long-term archiving. PDF/A can be displayed in all standard PDF viewers.

Key elements of PDF/A
- Exactly reproducible for many years in future
- 100% self contained, all information necessary for display is embedded in file
- No audio or video content
- No javascript or executable files embedded in the PDF
- All fonts embedded
- Color spaces must be device-independent so the PDF will display the same on all computers.
- No encryption (i.e., there can be no security set to disallow, for example, changing the document, copying text, etc.)
- Metadata must be standards based. Metadata is indexing information (like author name, and file creation date) placed in the document properties.
- Based on the PDF 1.4 (Acrobat 5) file format

Conformance levels:
At the moment there are two PDF/A conformance levels: PDF/A-1a and PDF/A-1b.
While PDF/A-1a is a more robust specification there seems to be a consensus that
compliance to this level is not normally practical today. The major difference is the
inclusion of document structure (i.e., tags). Document structure can be useful, for
example, to aid in the display of documents on devices with small screens -- like mobile
phones. As mentioned earlier, ALSTAR recommends conformance to PDF/A-1b which
ensures reliable reproduction of the visual appearance of the label (see details below).
A new conformance level, PDF/A-2, is in development and reportedly will be based on
the PDF 1.6 (Acrobat 7) file format and will bring additional features including support for
JPEG2000 compression.

Comments specific to creating a PDF/A-1b file:
The simplest method to create a PDF/A-1b file in Adobe Acrobat is to choose the pre-set
PDF/A-1b distiller settings when creating the PDF file. This process requires choosing
between the RGB and CMYK color formats. Generally, RGB creates a smaller file and
is normally used in electronically distributed materials. But since the PDF files that
registrants will submit to ALSTAR are originally created to print containers, it may be that
CMYK is the color format already in use. For ALSTAR either color format is acceptable.

If starting from an existing PDF file, created for example at a typesetting vendor, Acrobat
has a PDF/A compliance check under its Preflight menu. The easiest way to make an
existing non-compliant PDF file compliant is to redistill the PDF with the correct settings.
Unfortunately this won't work in all situations. Sometimes it will be necessary to go back
to the original authoring software file and redistill using the PDF/A-1b settings.

There is a situation that will probably affect a few ALSTAR participants. Some fonts will
not allow themselves to be embedded in a PDF file. The only known workaround is to
re-create the document using a different font. But since this could reflow the text on the
label and cause various layout problems, it may be easier to leave this font un-
embedded. Perhaps a solution to this situation will emerge in the future.

2. The Second Best Option. Create PDF files directly from the authoring software but
not compliant with the PDF/A-1b standard. This is the way most registrants create PDF
labels today.

If PDF/A is not an option for some reason, it is possible to create a PDF file which meets
some or all the PDF/A constraints. The full PDF/A specification seems to be unavailable
for free on the web but can be purchased from various sources. Distiller settings which
might be especially important in the context of ALSTAR include:
- Fonts: Embed all fonts. Subset fonts below 100%. Make sure that no fonts are set
  to "Never Embed." (Subsetting the embedded fonts is desirable because it is rare
  that a label (or any document) would use all the characters in a font set.
  Embedding the unused characters increases the file size.)
- Compatibility: Choose compatibility with Acrobat 5.0 (PDF 1.4)
- File security: Do not apply PDF security of any type to the document.
- Attached files: Do not attach files to the PDF.

3. The Last Resort Option. Scan a printed label into PDF. In general, ALSTAR disallows scanned labels. Exceptions may be made on a case by case basis. Contact the ALSTAR Project leader for further information.

Actually, a scanned PDF can also be PDF/A-1b compliant. Still, scanned labels are less desirable for ALSTAR because:
- File sizes of scanned documents are normally much larger.
- Files are often more difficult to read.
- The text is not searchable or copy/paste-able (unless OCRed, which still may have accuracy problems).
- Acrobat's document compare function is not a practical option on scanned files.

Other Issues

1. File Size. Regardless of which PDF creation option is used, downsampling the images in the file (i.e., reducing the resolution) to reduce the file size may be useful.

Container labels often include colorful, high resolution graphics which tend to inflate file sizes. Such settings are necessary to produce the physical container label but have no advantages for ALSTAR. Large file sizes, in fact, are a significant negative for electronic distribution.

The level of downsampling is set when the PDF is created. However, if the typesetter creates the PDF file at a high resolution, it is possible to downsample after the file is created. In Acrobat, shrinking the file is straightforward. Select Reduce File Size (which uses default settings to shrink the file size). This process normally results in screen and print displays which are adequate for the purposes of ALSTAR. If less downsampling is necessary to avoid fuzziness, the PDF Optimizer is another Acrobat function which allows custom downsampling settings. PDF Optimizer is also a good tool if you want to shrink many PDF files in a single batch process.

2. Marked Up Labels. Registrants also sometimes submit annotated labels to show changes from previous labels. Some PDF annotation methods comply with PDF/A-1b, some do not. Feedback from ALSTAR participants can be used to write guidance for annotations.

Adobe Acrobat provides many different annotation possibilities. As this document is written, it is unclear to ALSTAR which kinds of annotations are allowed in PDF/A-1b.

At this point it looks like there are three possibilities.
- Registrants use any annotation type which is acceptable to the states. The marked up label is considered a temporary document and discarded after use.
- Registrants use only annotations which are allowed in PDF/A.
- Registrants use any annotation type and then use Acrobat's annotation summary tool to re-render the label into a new PDF file which shows the original document and the annotations. This file could comply with PDF/A-1b.

3. Page Dimensions. It is a goal to produce the PDF label with the actual dimensions of the container label. Control of the on-screen or print magnification level is chosen by the person viewing the label.

The magnification level for viewing and printing is often thought of as a problem with the way the the PDF file was created. And IT IS possible to control the text size at the time of PDF creation by using shrink-to or expand-to-fit page.

However, the magnification level of on-screen viewing or printing is more properly the responsibility of the person viewing the label. This way, the original dimensions of the label text are not lost – which can be important to state label reviewers to determine label readability, etc.

4. PDF File Security. The agreement between the ALSTAR registrants and states is that the PDF files will be uploaded to ALSTAR with no PDF document security. This will allow states the maximum flexibility to use the label, e.g., to allow electronic stamps and electronic document comparison. However before a label is made available on any public site, ALSTAR will apply PDF document security: allow printing only.

5. US EPA Guidance. The US EPA has already issued guidance on preferred PDF creation settings and will be updated in the future. The current guidance on the EPA's web site is from May, 2003. Recommendations here are for ALSTAR only and not exactly the same as the EPA's.

In EPA's 2003 guidance, they discuss layout and formating parameters for Text PDF Product Labels, the labels submitted to EPA for registration. While these topics are interesting to ALSTAR, the guidance has limited cross-over to the Final Printed Labels submitted by registrants to states. For example the requirements for file naming, margins, and colors are different in the two types of labels.

Additionally the guidance makes recommendations for PDF creation methods which pre-date the PDF/A specification that is recommended in this document.