

It has come to our attention that there is some confusion on the technical side about what the definition of a PDF with XML Form data and a PDF without XML Form data is. (In previous conversations we have used the terms "Structured" and "Unstructured", but I'm hoping we can move away from this terminology since 'Structure' is a different concept in PDF, unrelated to XML Forms. I haven't come up with a suitable alternative yet, maybe "XML Form based" and "Non XML form based"? TBD.)

I've put together a few pointers for those interested in the technical details and for those who may want to actually generate these PDFs at the language level, rather than through the use of Adobe Tools.

The first document to look at is the PDF reference:

http://partners.adobe.com/public/developer/pdf/index_reference.html
<http://partners.adobe.com/public/developer/en/pdf/PDFReference16.pdf>

Search for the word "XFA".

These provide an top level overview, and describe how XFA should be embedded in a PDF.

8.6.1 - page 635

8.6.7 - XFA Forms

- The PDF Reference specifies that the XFA resource can be one stream or a collection of streams in an array. For performance reasons XFA should always be stored as an array of streams, not as a single stream, even though both are supported.

Note that there are several varieties of XFA Form with PDF. You can create all these from Designer. There are:

Non interactive, which have no interactive fields. These are essentially print forms

Interactive, which do have interactive fields. These consist of:

- Static PDF forms - these can be thought of as "rendered once" forms in that after the PDF is generated once it does not get regenerated. So if you end up importing XML data that has many more elements than are currently rendered, the PDF will not adjust to this new data.

- Dynamic PDF Forms - these PDFs are regenerated on the fly to keep up with the dynamic nature of XFA. So if you import new data with a different number of elements than are currently loaded in the form, the PDF will adjust. These will likely be the forms we base PDF/H XML Forms around.

There is also a history with Acroforms worth mentioning here. Originally, before Acrobat 6, forms in PDF were exclusively implemented using Acroforms. This is a PDF only based technology where form elements are entirely described using only PDF constructs. In Acrobat 6 and 7 XFA was integrated into PDF. The PDF specification provides some details on how this integration was done. In particular note the section on Page 683 about PDF interactive Form fields being present for every field in the XFA resource. Basically the Acroform and XFA structures are mirrored.

The next document to look at is the XFA specification here:

http://partners.adobe.com/public/developer/xml/index_arch.html
http://partners.adobe.com/public/developer/en/xml/xf_spec.2.2.pdf

This document is the full XFA Specification. I'm not sure anyone needs to read the entire document, I certainly haven't. But it might be a good idea to browse this and look at the sections below:

1 Introduction to XML Forms Architecture

3 Object Models in XFA

- the subsection "The DOMs used in XFA" is probably interesting

4 Exchanging Data Between an External Application and Basic XFA Form

6 Template features for Designing Forms with repeating sections

- this is the interesting information on form design for the Actors, Payers and any other section that repeats

8 Dynamic forms

- this provides interesting details on how the algorithm for data binding works.

10 Scripting

11 Forms that Initiate Interactions with Servers

14 Security, Control, and Digital Signatures - good reading for the security subcommittee in particular look at the subsection "Signed Forms". The other subsections aren't particularly relevant.

These two chapters describe how XFA sets up the connections between an XML Schema, XML data, or a WSDL file, and the form itself:

18 Connection Set Specification

19 Data Description Specification

This is another document that talks about schemas in Designer

http://www.adobe.com/devnet/livecycle/articles/lc_designer_XML_schemas.pdf

Hopefully this is a good start for anyone interested in these details. Please contact me if you have any questions or need any clarification. Please let me know if you think it might be useful for me to give a quick overview and Demo on XFA, Designer, and PDF. If enough people think it is, I'll see if I can put something together. Once we get to the 3rd section of the CCR we'll want other participants to do some of the next few forms/subsections so it might be useful for those participants to get familiar with the language and tools.

Thanks

Anatole