Adobe® InDesign CS3 XML Overview

Gary Cosimini

Business Development Director, Creative Solutions Business Unit
A Time of Change

Rapidly proliferating devices, competition for attention and industry evolution
How to Meet Content Delivery Challenges with So Many Targets?
How Does IT Meet these Challenges with these Workflows?
The Answer: Integrate Solutions Based on Media-Neutral Content
The Creative Solutions Toolkit

Create
- Creative Suite 3
- InCopy
- XMP
- ExtendScript

Manage
- Bridge
- Version Cue SDK
- InDesign Server
- Scene 7

Automate
- Connectivity
- Automate

Deliver
- HTML
- PDF
- Flash/Flex & AIR
- Digital Editions
Adobe Bridge and Version Cue CS3

A visual tool for organizing and editing files and XMP metadata. Bridge CS3 can be integrated into external data sources like DAMs and web services. Version Cue provides version management to all the CS3 applications.
Adobe InDesign Server CS3 Automates Document Layout

Server version of InDesign CS3 that powers publishing solutions

A robust, flexible, and scalable server-based layout and composition engine that automates professional design

- Brochure, collateral, and catalog creation
- Automated ad creation
- Editorial workflow automation
- Personalized marketing (VDP)
Today’s Agenda

- Why is XML important for publishers?
- Quick high-level overview on XML features in InDesign
- How does XML fit into my workflow?
- Cover key concepts for XML Import, Export, and Tagging
- Additional XML tools of the trade
- Tips, tricks, and “gotchas”

- NOT a replacement for existing documentation or tutorials
- Assumes you have already worked with the XML features in InDesign and know a little about XML and XML terminology
**What is XML? eXtensible Markup Language**

- Strictly speaking, XML is not a Mark-up Language despite its name.
- It is a general tagging scheme to create mark-up languages …many, many languages.
- XML is Rules & Tools: play by the Rules and you get to use the Tools.
- An international standard developed and maintained by the WWW Coordinating Committee (http://www.w3c.org).
- Based on another tagging language called SGML.

```xml
<tag> Text </tag>
```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<Root>
  <Card>
    <Name>Jane Doe</Name>
    <Address>New York City</Address>
  </Card>
  <Card>
    <Name>John Smith</Name>
    <Address>Seattle</Address>
  </Card>
</Root>
XML For…

- To speak about XML as a solution, it’s always necessary to qualify what the XML is for

- XML for Business Cards
- XML for Web (XHTML)
- XML for Print Job Definition Format (JDF)
- XML for Scalable Vector Graphics (SVG)
- XML for PDF (Mars)
- XML for InDesign page parts (INX)
- XML for Metadata (XMP)
- XML for eBooks (OEBPS)
- and so on…
Why Is XML Relevant To Publishing?

- Publishing is evolving from Proprietary formats to Open Formats
- Open formats give publishers more choices and preserve investments
- Flexible “meta-language” means publishers can leverage XML to create their own industry specific standards
  - (AdsML, NewsML, EXIF, DICOM, XMP, etc.)
- Used in database publishing solutions that use layout automation
- Used in cross media publishing solutions
Freely Available Technology Makes XML Attractive

- Open standard for data interchange encourages open source tools development
  - XSLT, XQuery, XSL-FO…
- XML technology supported by almost all software vendors
- Cross platform and portable
- All databases now support XML import and export
- Fundamentals of XML focuses on simplicity

- Recognized need for both a strict and a more “relaxed” approach (“valid” vs. “well-formed”)
  - Well-Formed XML conforms to a description of relationships of elements called a DTD (Document Type Definition) or to a Schema (same thing as DTD but in XML)
Numerous Long-term Benefits Implied

- Improving the process through automated layout on import
  - Database publishing (a.k.a. On Demand Publishing)
  - Personalized publishing (a.k.a. Variable Data Printing)

- Adding new processes & services to the digital design & production workflow
  - Crossmedia (Repurposing and Syndication of content)
  - Metadata
  - DRM (Digital Rights Management) Tracking & Auditing
  - Content or Digital Asset Management (CMS and DAM systems)

- Full roundtrip digital content workflows
  - Doing most of the above plus full bi-directional XML content creation and editing throughout the design and production workflow
But Adopting XML is Rarely Painless

- How you implement XML will greatly depend on what fits within your existing workflow, and to what degree you are willing to change (or can control) your workflow.

- There's pain in changing your workflow, and so the gain had better be worth it. Planned workflows are where XML projects tend to be more successful.

- Tech docs, some journals, books, directory, and catalog publishers who adopted SGML have also transitioned to XML. Unplanned workflow tend to have trouble finding the right ROI in XML workflows.
InDesign and XML Workflows

- **First Ask:**
  What Is The Nature of Your Business Workflow?

- **Planned: high ROI**
  - Majority of work is updating and republishing the same publications, which leverage templates and stylesheets extensively, on a regular, predictable timetable.

- **Unplanned: low ROI**
  - Majority of work cannot be planned and the requirements of the document will change client-by-client, job-by-job.

- **Ad-Hoc: medium ROI**
  - Somewhere in-between Planned and Un-planned. Have some major clients who provide steady work and are looking for ways to improve the process.
List of Top XML Features in InDesign

- **XML Import:**
  - How do I get XML into InDesign document?

- **Adding Structure to an ID Document:**
  - How do I work add XML structure and work with XML in an InDesign Document

- **XML Export:**
  - How do I get XML out of an InDesign document?

- **Some friendly advice:**
  - The devil is in the details!
  - There are a lot of choices to make.
  - Read the Manual!
  - Many of these options are cryptic and it helps to understand what InDesign is doing, and how to debug or troubleshoot it.
  - Read the technical white papers! The manual is a good place to start, but does not cover everything.
Major new InDesign CS3 XML features

- XML Rules (scripting) on Import
- XSLT on Import and Export
- Find/Change: search on the usual parameters but on Change, apply a Tag
Major new InDesign CS3 XML features (continued)

- Minor CS3 improvements:
  - CALS Table import
  - XML files linked in Links panel
  - AID namespace attributes support Table and Cell styles as well as Paragraph and Character styles
  - Automatic (hard-coded) character replacement on Export
  - GREP as another tool to search at a more sophisticated pattern matching level

- We provide an XML Technical Whitepaper to cover all of these topics

XML Import – Decisions, Decisions…

- XML is imported in the Structure pane, unlike other file formats

- Manual vs. Automated Layout

  - Manual (aka, just-get-the-job done):
    - Importing XML but formatting and laying out the XML content “by hand”
    - Automatic (aka, don’t make me repeat myself):
      - Importing XML content but automatically apply as much formatting on import as possible

- Merge Content vs. Append Content

  -Appending adds XML content to your document leaving the existing structure and content alone
  -Merging assumes you have created “placeholder” text or frames and wish to substitute it with new XML you are importing
XML Import: Tag to Style Mapping

- Mapping Tags-to-Styles can help automatically format text
- Quick demo: Biz Card
- Use AID attributes allow you to do the same thing as Map Tags-to-Styles but within the XML data
  - Apply paragraph and character styles with AID:PSTYLES and AID:CSTYLES
  - Better controls over importing XML tables with AID:TABLES (and TROWS, TCOLS, THEADER, CROWS, CCOLS, CCOLWIDTH, TFOOTER)
Practical Advice

- Use XSLT to transform the XML into a simpler, flatter structure
  - Also a good time to resolve “ordering” issues
  - Also a good time to resolve “typographic” issues with the data, esp. mapping text entities into typographic font equivalents

- XML Import cannot automatically add pages or create new frames, so
  - “Placeholder” content approach is best used for static designs where mapping is relatively straight-forward
Working With XML
Working With XML

- InDesign can load a DTD and tag existing text or frames to elements
  - Once a DTD is loaded, you can periodically validate the structure, but note that ID is not continually validating or enforcing the structure in real time
- Structure pane view can also allow you to select, reorder, and place XML content
  - TIP: XML is SEPARATED from your design content, so you can import more XML content than is shown in your document layout view
- Remember, InDesign isn’t an XML Editor! It’s easy to break the structure if you don’t know what you’re doing
  - TIP: Use the Story Editor mode, instead of the Structure pane, to “safely” edit the text, esp. when adding or removing carriage returns
Style to Tag Mapping

- InDesign can map Character and Paragraph Styles to tags
  - Very useful if styles are applied logically and consistently
- TIP: Each time you do this, previously mapped tags are wiped out, including any custom attributes you may have added
XML Export

- Scenario 1:
  - You already have a structured document either by Importing XML or applying structure (tags) from the Tag Palette. Just Export XML.

- Scenario 2:
  - You have used paragraph and character styles judiciously in the InDesign document. First Map Styles to Tags, then Export XML.

- TIP: You can also export “child” nodes of your XML by selecting that node in the Structure pane first, and then Exporting
Export XML Tips

- Third-party plug-ins like PCIPage Scriba and Atomik EasyPress can deduce structure from layout and rulesets, and force XML exported from richly-designed but unstructured documents to conform to a DTD.
What is INX (InDesign Interchange format)?

- INX is a serialization of the contents of an InDesign document, story, or page part expressed in XML with InDesign-proprietary tags
- Invented for
  - Saving backwards to older versions of InDesign (easier to deprecate new features)
  - Exchanging stories with InCopy
  - Saving pieces of a page to Libraries and to InDesign “Snippets”
  - Filtering converted DTP files to clean out irregularities
- INX files include content, layout and XML structure, but not in a human-readable order
  - The ONLY way to reliably read, edit and write INX files is by using InDesign
- Could INX evolve to become a structured and logical document description? Yes.
What’s InDesign Tagged Text?

- Don’t confuse Tagged Text with XML
  - Tagged Text is a formatting language that happens to look like XML tags
    - Tagged Text is a kind of “old-school” approach to automating layout
      - Quark Xpress Tags, Microsoft Rich Text Format
    - It’s NOT the same thing and has little connection with InDesign XML specific features
- One could create Styles in Tagged Text or RTF, and then Map Styles to Tags, but XML Tagging can do the same thing more directly with fewer chances for error
A Few Programs that Export XML

- InDesign, of course
- FileMaker
- Mark Logic, Oracle, MS-SQL
  - Practically any database
- Microsoft Excel
  - As well as Office 2007 applications including Word and PowerPoint
- Adobe Acrobat
  - Import and export XML to PDF Forms
  - MARS plug-in on labs.adobe.com “rephrases” PDF in XML tags
- Soundbooth (Export Flash Cue Data)
- SoftCare K4 XML Export
- Many, many others
Some XML Editing Tools

- Altova XMLSpy
- Oxygen XML
- ElfData XML Editor
- Your favorite text editor
  - BBEdit
  - Text Wrangler
  - TextPad
- DreamWeaver and GoLive
- InDesign or InCopy*

* Note that InCopy does not have DTD validation; InDesign does
Other XML Resources

1. InDesign XML Content:  

2. Adobe User-to-User Forum:  
   InDesign Scripting

My Top Picks:

3. InDesign CS3 and XML: A Technical Reference

4. InDesign User Guide XML Chapter

5. Real-World Adobe InDesign CS2: XML Chapter 14
Parting Comments

- XML isn’t a solution, it’s just technology
- Lots and lots of choices to make, which means be prepared to do some homework, experiment, and expect to make mistakes
- Start small, gain some experience, and learn from your mistakes
- Begin by creating some tags in InDesign, tagging some text, exporting as XML, editing the XML, and then importing it back.
- Decide what kind of XML workflows best fit your business needs and workflows, and keep in mind the how you will achieve an ROI
- There aren’t yet any “out-of-the-box” solutions for XML
- If you’re not technically savvy, seek out assistance!
- Remember, most databases do not produce publishing quality data, and some clean-up work will be required
Thank you!
Demos

- Import XML, Map Styles
- Import XML/XML Rules
- Export XHTML/CSS
- Export XML/Digital Editions