The PPML Print Language in XML Workflows for Digital Print

Dave **deBronkart** <daved@podi.org>

Abstract

PPML is a new printer language gaining widespread support in variable data printing applications. PPML uses XML as its syntactical base, and can access page content files generated in many different formats. Applications will ultimately range from desktop to high-end digital printing presses.

PPML output is being integrated into a variety of complete workflows, several of which can use other XML methodologies. We envision a fully automated workflow of the future, in which data may flow directly from its source to automated production equipment, untouched by human hands, with every aspect of the process expressed in XML. In such a workflow, PPML will be the printer language by which content and page layout are delivered to the output system.

1. The context: digital printing

Digital printing is enabling an important convergence of applications and workflows that have previously been distinct. By combining the flexibility of an all-digital workflow with quality that had only been available on a printing press, today's digital presses make possible publishing applications that were previously impossible.

New XML-based standards are being introduced for control of print processes, because these new applications want to be automated, and conventional graphic arts workflows are largely manual. Nowhere is this more true than in personalized promotional printing, where variable content increases reader response and print quality is important.1 Several vendors built proprietary workflows for this application, because no standard method existed.

2. PPML and PODi

In 1999 PODi determined that a key factor holding back adoption of personalized print was the lack of standards.

PODi (http://www.podi.org) is a not-for profit industry initiative whose mission is to develop the digital print industry through market education and by encouraging standardization. Executive Board members are Adobe, CreoScitex, Electronics For Imaging, Hewlett-Packard, NexPress, and Xerox. General members include Barco, IBM, Indigo, Lexmark, Noosh, Océ, Pageflex, Scitex Digital Printing, and Xeikon. Banta, Canon, Epson, Hitachi-Koki, Mediaflex, Minolta, and Nimblefish are Observer members, who preview outputs of PODi activities.

The members developed a new non-proprietary langauge, Personalized Print Markup Language (PPML). Using XML as its syntactical base, the specification was first introduced in 2000, initial technology demonstrations were seen at the drupa 2000 exhibition in May 2000, and initial system shipments took place by year-end. A three-page overview of the language's purpose and applications is available at http://www.podi.org/standards/PPML. Work has now begun to integrate PPML into complete workflows.

At this point it is appropriate to establish the position of PPML in the overall workflow.

3. Using XML in different stages of document workflows

Automated document production requires transmitting several types of information:

- Content

- Page Layout

- Production instructions, common referred to as job ticketing.

For data-driven personalized printing, a separate issue is how (and where in the workflow) the database fields influence the generation of the finished documents.

3.1. PPML

PPML is used to convey content and layout.

3.1.1. Content

PPML uses the IANA type identifiers (e.g. "application/postscript", "image/jpeg") to identify the format of the source data for page content objects.

PPML itself does not *describe* page content; it merely conveys (or points to) content defined in some other language. (There is no construct in PPML to define a three inch red circle. Rather, PPML could point to a file "ThreeInchRedCircle.EPS", or it can enclose code to create such a circle, in PostScript or SVG or any print language supported on a given machine.)

Since the PPML language is intended to support a broad range of applications, it allows manufacturers to choose which data formats their machines will support. For common applications, such as graphic arts, office, and transactional printing, common feature sets are being defined, so that users will have predictability. The first such feature set is defined in PPML version 1.5, for graphic arts applications.

3.1.2. Page Layout

PPML is hierarchical. The root PPML element can contain JOB elements, which can contain DOCUMENTs, which contain PAGEs, which contain MARKs. Each MARK has a Position attribute.

In addition, optional PRINT_LAYOUT elements can define how PAGEs are distributed onto SHEETs, including IMPOSITION, SHEET_MARKs, etc.

3.2. JDF for job ticketing

The third component, beyond content and layout, is the job ticket. The de facto standard in development for job ticketing is the Job Definition Format (JDF), being developed by CIP4 (www.cip4.org) and well documented elsewhere in these proceedings.

One point is worth noting about JDF's interaction with PPML. As mentioned above, a PPML dataset can include instructions for positioning pages of content onto press sheets. Strictly speaking, this violates the concept of PPML as a vehicle for content and page layout. (PPML's PRINT_LAYOUT element specifies how something should be produced, which a content format should not do.) For that reason, the PPML specification states that PPML-consuming systems are entitled to ignore any imposition instructions in the PPML.

3.3. Xpath, XSL, XSLT, FOP

Today's production processes for personalized documents (with reusable content) are largely mapped onto conventional workflows. In these workflows, the variable data records trigger the selection of reusable objects, using logic rules embedded in a template.

But we eagerly anticipate the development of a new breed of workflows, in which the data itself is translated or transmuted into content objects or even into fully formatted documents. For these new-breed applications, there will be great use for Xpath, XSL, and XSL-Translate. The PDF-generating tool FOP may also be used to directly generate PDF content objects, which can be referenced within the PPML stream.

4. The evolution of standards

PODi's mission includes encouraging evolution by promoting interoperability, through standards. Our driving principles are fast time to market, and ensuring commercial implementation.

PODi's method of prioritizing activities is explained in our article "Trends that will change the business of print" in the *Seybold Report on Publishing Systems*, Sept. 8, 2000. The guiding principle is to "choose a problem that's worth solving, and can be solved, while the problem's still there to be solved." In other words, the scope of the project must be kept to a manageable size, and vendors must see clearly defined payback for their investments.

In this context, we can ask, which XML standards for digital print will actually achieve widespread adoption?

We believe that the answer lies in understanding the forces that drive standardization, and that the answer points first to PPML.

In his landmark book *The Innovator's Dilemma*, Harvard Business School professor Clayton Christensen wrote about how new "disruptive technologies" undermine an established technology. The book is a classic, read almost as widely as Geoffrey Moore's *Crossing the Chasm*. His principles are backed by massive amounts of data regarding product life cycles and technology adoption.

In a new working paper, Christensen has extended his research to document how and why industries switch from vertical integration (complete systems) to "disintegration," i.e. modular components distributed horizontally across an industry. The driving force, he finds, is this: when technology has fully satisfied market needs, then a vendor cannot assure market position by being better. Instead, deals are won by being faster or more flexible and delivering exactly the configuration an individual customer wants.

At the moment when that happens, standards become valuable - because they enable a shift to modularity, which in turn enables delivering product faster or more flexibly.

That is the state, today, of the digital printing industry. Quality is now more than good enough for many applications. So the deciding factor in choosing equipment will not just be print quality, it will be suitability for the application. We are now at the point where modularity will be valuable - and a standard interface to the RIP will be a compelling factor.

There are two important bonuses. First, since many software companies are adopting PPML as an output option, so every buyer of a PPML machine has a range of different applications available. Second, PPML is designed to be suited for cross-market applications, so a print shop that adopts PPML workflows will be positioned to capitalize on new opportunities as print applications converge.

And since PPML page streams can be generated directly from XML data (as described above), without going through intermediate (manual) applications, it will become possible to generate PPML output as automatically as Web pages are generated today, directly from XML.

5. Conclusion

The world of digital printing is evolving rapidly. As vendors and users strive to eliminate sources of friction from today's workflows and system configurations, the value of modular configurations and hands-off workflow increases. The first point supports the adoption of PPML as a RIP interface; the second point shows the value of having a print language that's entirely expressed in XML.

When PPML (for content and layout) is combined with JDF (for job ticketing) the industry will have achieved an all-XML workflow for digital print.

1 Personalized print typically requires that the print system vary the selection of page content elements based on the values of database fields.

Glossary

PPML   Personalized Print Markup Language

Biography

Dave **deBronkart**
    Senior Consultant
    Print On Demand Initiative (PODi)
    West Henrietta
    USA
    Email: daved@podi.org

Dave deBronkart - Dave deBronkart is Senior Consultant with PODi, the Digital Printing initiative, a not-for-profit industry initiative focused on developing the market for digital printing through market education and standards activities. Throughout his career in the graphic arts he has focused on productivity enhancements, both as consultant and employee at industry leading firms including Compugraphic (now Agfa), Indigo, Scitex, Itek, and Du Pont. For the past year he has served as Program Director for PODi's initiative to develop

PPML, an XML-based standard for personalized printing. A graduate of MIT's Sloan School of Management, Mr. deBronkart is based in the Boston area.