

Beyond 'Post': Adobe Forms vs. InfoPath?

Make sense out of two forms-based technologies, Adobe's PDF-based forms software, and Microsoft's newly released InfoPath application.

By Andrew Binstock

During the last few years the concepts of workflow management and collaboration have been enjoying steady adoption in commercial enterprises. Their emergence is not just the product of new standards and new technologies, but rather, it's driven by the inherent cost and inefficiency of handling paper forms and the desire for traceability and accountability in document handling.

As these electronic versions of what was once hard copy takes root in the enterprise, software giants have begun releasing enterprise-scale packages. The loudest of these has been Microsoft with its InfoPath technology. Adobe, whose Acrobat technology and PDF specification have made it the market leader in forms technology, is the other major player in this market. Let's compare their offerings and examine their differences.

Microsoft InfoPath

Microsoft's forms technology derives from XForms, a specification that reached recommendation status in the Worldwide Web Consortium (W3C). XForms specifies how forms should appear and how data should be handled using XHTML (that is, HTML that is XML-compliant, or well-formed HTML).

Microsoft rarely mentions this standards connection any longer; instead, referring only to the XML derived from its own schemas, which it generates in InfoPath. The product was first released in 2003, which was followed by its first service pack. InfoPath requires installation of a separately purchased software client, which can be obtained either as a standalone product or as part of Microsoft Office Professional Enterprise Edition 2003, the high-end version of the Office product suite.

Users employ InfoPath when filling out documents, such as expense reports, that are sent to them via a pre-established template. The template, the form, and the filled out data are all XML that uses the schemas published by Microsoft. The InfoPath document can make use of Web services for purposes of formatting information or validating user input. Employees who want to only read filled-out InfoPath documents (with the .mht file extension) can do so using Microsoft Internet Explorer. (The Firefox browser does not read .mht files.) However, should they want to modify the contents, they then need the full InfoPath client software stack.

One important limitation of the InfoPath product is that the documents are pure XML text, sent in the clear. There is no concept of encryption, password protection, or security of any kind. Anyone with a copy of the document can examine its contents.

In addition, InfoPath is narrowly focused on the form itself. Workflow is provided by separate Microsoft technology, namely the Biz Talk server. This comparatively expensive software also uses XML, but its model is rather different than InfoPath. BizTalk is an enterprise communication hub with some workflow capabilities. It is not a document-centric system. As a result, it requires an entirely different set of skills to administer than does InfoPath.

Implementations of BizTalk server require Microsoft Windows Server 2003 or Windows 2000 Server, Microsoft SQL Server, the .NET Framework, and Microsoft Visual Studio .NET. So, while forms might be easy to use, but insecure, the workflow solution from Microsoft is large, complicated, and not principally designed for forms and document workflow.

Adobe Forms

The Adobe Intelligent Document Platform and LiveCycle enterprise workflow system are based on the company's widely used products, document services, and open standards. The documents employ the open PDF standard as well as XML, while the server side runs on all leading J2EE servers.

As millions of users worldwide know, the Adobe Reader is available at no cost for dozens of different platforms. Adobe Reader can open and display forms and execute embedded actions without any additional software. And because PDF documents have numerous provisions for security (such as password protection; prevention of copying, extracting, or printing contents; and digital signatures), the Acrobat Reader has built-in functionality to enforce these security restrictions at the destination. PDF files are also intelligent documents. They can embed logic that validates user input, cross-foots totals, generate printed barcodes from entered data, and performs a multitude of other nifty functions.

An important distinguishing trait of the Adobe solution, however, is the back-end systems for collaboration and workflow management. While Microsoft uses a heavyweight solution that performs enterprise collaboration as a sideline, the Adobe LiveCycle product is purpose-built for workflow management. And because it's built on Java technology, it does not require one vendor's software ecosystem to run. Any J2EE server, even an open-source implementation, can run Adobe's LiveCycle server.

The document services provided by LiveCycle are flexible and extensive. The LiveCycle Workflow, for example, coordinates the movement of documents throughout the company and can integrate the document management process with a company's core systems, business rules engine, and Web services. The LiveCycle Policy Server enforces corporate rules regarding confidentiality, privacy, and accountability of electronic documents, even for forms that are offline and outside the firewall.

Finally, the Adobe LiveCycle Form Manager serves as a repository for forms and templates. But as more than a container for forms, the Manager can prepopulate forms for specific users or groups of users. This function, which is configured via an administrator's portal, reduces errors and decreases the time users spend filling out forms and repeatedly providing the same information.

Comparing the Two

Microsoft provides basic document creation with InfoPath and limited workflow capabilities with BizTalk Server. The documents this set-up generates are pure XML that is sent over the wire as plain text. It has no encryption or security mechanism.

In contrast, the Adobe Intelligent Document Platform endows every step in the process with document-specific benefits. The universal client, Acrobat Reader, is available at no cost on all platforms. It renders documents identically on all systems and it enforces the security, encryption, and management policies embedded in the document. In addition, it provides forms-filling capabilities that are highly interactive in nature and can be scripted by developers.

The documents themselves have intelligence--not only at the level of security but in their ability to interact with Adobe servers to ascertain routing.

The back-end systems, consisting of multiple Live Cycle tools that can be mixed and matched as needed by the enterprise, provide extensive collaboration, workflow management, policy enforcement, intelligent document generation, and efficient forms handling.

Another salient difference between the two companies is the ecosystem around the products. Adobe's PDF file format is now generated by hundreds of products that leverage the standard's capabilities. There are many open-source implementations, as well as commercial libraries, that make it possible for IT departments to create their own forms and documents with special features. In comparison, Microsoft's InfoPath has a supporting ecosystem consisting of a few handfuls of ISVs. (A search on SourceForge.net shows two projects for InfoPath and hundreds for PDF.)

Adobe's extensive developer programs enable programmers to design software that plugs into the LiveCycle product chain at numerous points, including form design, form interaction, and the various aspects of document management. Links of interest to developers appear in the next section.

Resources

The home page for Adobe's Intelligent Document Platform is:
<http://www.adobe.com/enterprise/idp.html> Developer resources for this platform are explained at: <http://www.adobe.com/enterprise/developer/main.html>

The specification for the PDF file format from Adobe is available here:
http://partners.adobe.com/public/developer/pdf/index_reference.html, while is XML Forms document is available at
http://partners.adobe.com/public/developer/en/xml/xfapecification_2.1_draft3.pdf

The specification for XForms can be found at: <http://www.w3.org/TR/xforms/>

The home page for Microsoft InfoPath 2003 is
<http://www.microsoft.com/technet/prodtechnol/office/office2003/plan/inpthfaq.msp#EGAAA>

Home page for Microsoft BizTalk Server 2004 is:
<http://www.microsoft.com/biztalk/default.msp>

Andrew Binstock is the principal analyst at Pacific Data Works LLC. Previously he was the director of PricewaterhouseCoopers' Global Technology Forecasts. He writes the business integration column for SD Times. His latest book, 'Programming with Hyper-Threading Technology: How to Write Multithreaded Software for Intel IA-32 Processors,' is now available from Intel Press.