TECHNOLOGY BRIEF

The Adobe® XML Architecture

Introduction

As enterprises struggle to balance the need to respond to continually changing business priorities against ever-shrinking budgets, IT managers are looking to find ways to extend their investments in core business applications such as Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and Content Management (CM) systems. One of the most effective ways to do this is by replacing their slow, costly, and error-prone paper-based processes with faster, more accurate, and more efficient automated processes.

The Adobe XML architecture combines the powerful data and business logic capabilities of the eXtensible Markup Language (XML) with the rich presentation and enhanced security capabilities of Portable Document Format (PDF). This combination provides organizations with a step-by-step migration from manual, paper-based workflows to streamlined automated processes that fully integrate electronic documents and forms. Each step extends the value of your investment in core applications by increasing the quality of the data in the process and accelerating process cycles. These new processes still deliver the look and feel of familiar paper documents, thereby complying with regulatory requirements and increasing the rate of user adoption.

The Adobe XML Architecture

Adobe offers solutions for streamlining business processes with a robust set of document services, comprising both server and desktop applications, that create and integrate intelligent documents. These solutions extend the reach of core applications to users within and outside the organization, as documents can be accessed and interacted with across heterogeneous platforms and devices using the free and ubiquitous Adobe Reader® client.

Intelligent documents are the ideal way to connect people, processes, and applications in any workflow, because they offer a secure and reliable format for presenting content while including the business logic necessary for capturing and interacting with data stored in back-end systems. Additional benefits of using a document-centric approach for interacting with core applications instead of HTML-based interfaces include:

Meeting regulatory compliance requirements. Many government agencies require digital documents to look and feel exactly like the paper versions they are already using.

Overcoming users' resistance to change. People are accustomed to working with paper documents and forms, so creating electronic replicas of these familiar documents provides a comfort level that accelerates the adoption of new streamlined processes. This results in reduced training costs and greater compliance with organizational policies.

Supporting offline workflows. There are many people involved in business processes that cannot be directly connected to core applications, such as mobile workers or customers and partners that are outside the firewall. Because intelligent documents carry data and business logic within them, these users are able to fully participate in a process without being connected to the enterprise applications.

TABLE OF CONTENTS

1 Introduction
1 Adobe XML Architecture
3 XDP—Extending the flexibility of XML and PDF
4 Full Support for Arbitrary XML
5 Designing Enhanced Business Logic into Forms
6 Leveraging the New Architecture
7 Conclusion
8 Appendix I: Applying XML to a Business Process
10 About Adobe

Adobe®

The Adobe XML architecture is a key enabler of intelligent documents and forms.

XML is an open standard often used for data exchange and is a preferred format when data simply needs to move from one application to another. However, people need a way to interact with XML data that makes it easily consumable, while maintaining the ability for applications to work easily with the data when necessary.

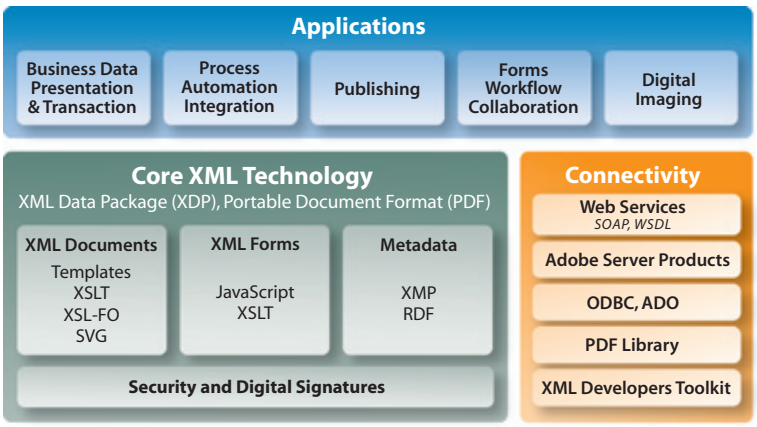
PDF has become a de facto standard for the secure and reliable exchange of documents. To date, more than 500 million copies of Adobe Reader have been distributed worldwide, which has contributed to users developing a high comfort level with PDF documents. As the presentation layer for intelligent documents, PDF provides a familiar and consistent user experience for people to interact with XML data.

The advantages of combining PDF and XML include:

- XML data can be embedded within digital documents and forms that are high-fidelity replicas of their paper counterparts, while still enabling the direct integration of XML data with enterprise applications.
- PDF can include business logic, such as calculations and data validations, resulting in improved quality of captured data, thus reducing process delays.
- People can participate in processes while they are offline, because XML data travels within the PDF document.
- Users can extend and customize metadata within PDF.
- PDF files include a variety of security options, from access restrictions to electronic signatures, that help protect sensitive company information.
- XML data can be locked down by the PDF container to create documents of record and to comply with the PDF/A industry standard for electronic document archiving as recently mandated by the US National Archives & Records Administration. The association of XML data to its display format is critical to maintaining an audit trail, combatting fraud, and protecting the abuse of raw data stored in databases, which is especially critical in highly regulated industries.
- Users can participate in XML workflows any time, anywhere, on any device through universal clients such as Adobe Reader or Web browsers.

The Adobe XML architecture leverages the capabilities of XML and PDF to support a variety of business applications, while offering connectivity to systems through a variety of industry-standard and Adobe technologies (see Figure 1).

Figure 1



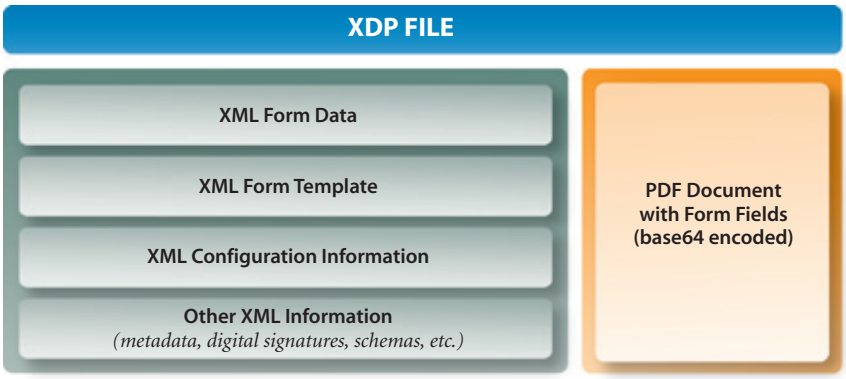
XDP—Extending the Flexibility of XML and PDF

With the Adobe XML architecture, depending on the requirements of the process, documents can be created in PDF and then processed as PDF is today, or expressed in an XML Data Package (XDP) and processed as XML. An XDP file is simply an XML file that packages a PDF file in XML, along with XML form and template data.

Because an XDP file is an XML file, standard XML tools, system interfaces, and Web services can work with it, and the XML data is directly accessible. Additionally, a person can view and interact with the PDF document using the free Adobe Reader or the Adobe Acrobat® family of products.

An XDP file contains several distinct blocks of information, as shown in Figure 2.

Figure 2



XML Form Data. This component is the user data encoded according to an arbitrary XML schema chosen by the form developer during the design phase. The schema can be an industry standard, the enterprise's standard, or completely customized. Some examples of industry-standard schemas are ACORD (insurance), XBRL (finance), HL7 (healthcare), and SF424 (eGovernment).

XML Form Template. This component contains all the form intelligence, including the mapping of XML form data to PDF form fields as well as all the business logic that controls the interactive behavior of the document, such as calculations and data validations.

XML Configuration Information. The XML form template uses this component as a global reference for database and Web services SOAP connections.

Other XML Information. XDP files can include custom XML information such as a schema file to facilitate validation, XML digital signatures, content metadata to facilitate archiving, or data used by a custom digital document application.

PDF Document. XDP files provide all the traditional PDF benefits of precision document layout and high fidelity printing by embedding the PDF in an XML element.

An XDP file is an XML file, so all XML tools, XML system interfaces, and Web services can work directly with it. The XML data is directly accessible.

Consider the following factors when deciding whether PDF or XDP is the appropriate representation for a given process:

Use PDF when:

- File size and transmission time are important
- Documents are large
- Document is packaged with supplementary data or images

Use XDP when:

- Documents and form data must travel in XML workflows
- Form data needs to be manipulated with XML tools
- Documents will be stored in an XML repository

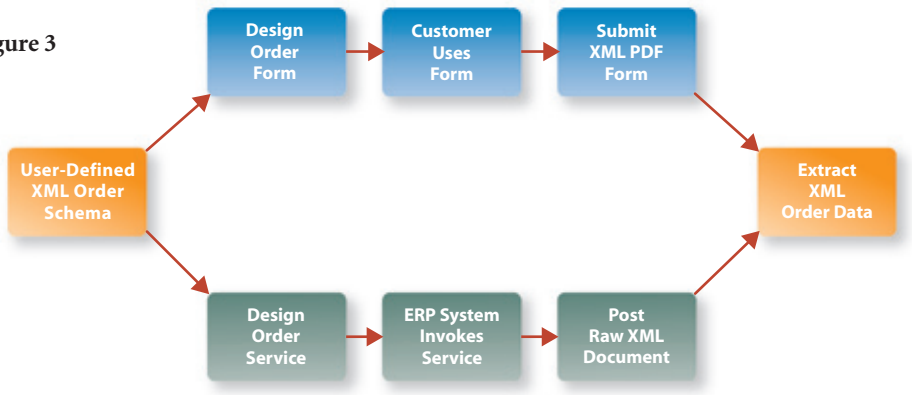
Both formats are treated identically by Acrobat and Adobe Reader and can be used online and offline.

Full Support for Arbitrary XML

The Adobe XML architecture also supports user data in an arbitrary XML format. So if your organization uses a standard schema for purchase orders, account information, or project schedules, your PDF forms can directly use XML data structured according to this schema. Figure 3 shows the advantage of this approach. Once you have a schema for something like an order, interface designers can create form layouts for human interactions while application developers can use the same schema to build Web services interfaces for the back-end system integration.

For example, customers at companies who haven't invested in enterprise applications can download the form, fill it out, and submit it via e-mail, or alternatively, a user can simply click a Submit button on the form to send the data to a Web services interface via a SOAP message. For an organization that has implemented ERP, the application can directly connect to the service interface, generate the appropriate XML data, and invoke the service. But the end result is the same. The order processing system gets a completed XML document that conforms to the desired schema. When you want to accommodate human interaction in an automated process, this approach reduces the cost of requirements analysis, interface development, and back-end integration.

Figure 3



One Form, Two Processes
The XML within the form is used to directly interact with enterprise systems, but can also be rendered to PDF when people need to participate in the process.

The Adobe XML architecture uses an object-oriented approach to map user XML form data and data from dynamic sources to particular fields in the form layout. You can also apply an XSLT transform to the user form data before performing this mapping. This option is useful if your organization needs to support two different schemas for the same information—for instance, an internal standard and an industry standard. You can reuse all the form logic, and then simply apply the transform to convert data from one format to the other.

Designing Enhanced Business Logic into Forms

While supporting arbitrary XML form data incrementally improves the ease of integration, the addition of XML-based business logic significantly increases the efficiencies of business processes and the quality of the data used in those processes.

A new Adobe form designer that is in development will enable enterprises to more easily take advantage of a wide range of capabilities that are included in the Acrobat 6.0 product family and Adobe Reader 6.0. In addition to providing a graphical interface for designing highly structured form templates that include the ability to align each form element precisely, the new tool will also enable form authors to import XML schemas and graphically map them to form fields. Advanced palette tools, widget inspectors, and interactive wizards will greatly simplify adding logic, such as setting valid data ranges, restricting access to certain fields, or setting up Web services connections.

These capabilities deliver a powerful client platform with a rich user interface, the ability to execute intelligent actions, and direct connectivity to back-end systems, without the need for server-side processing.

The capabilities provided in the new design tool will include the following:

User Interaction. You can simplify filling out forms for users by adding enhanced functions, such as defining a formula for calculating fields from data in other fields. Formatting can also be applied dynamically, such as when tabbing out of a field or making additional entries.

Validation Constraints. To reduce data entry errors, form designers can include data validation within a form. For example, you can specify constraints on the data type or range of a single field, or define particular relationships that must hold among data in a group of fields. Another example is including restrictions based on a user's role, such as disabling a customer's ability to access certain fields that are reserved for internal administrators.

Data Connections. Users often need to dynamically retrieve the most up-to-date version of certain information, such as customer contacts, production volume, stock prices, and approval status. Forms allow you to retrieve such information using ADO, ODBC, or SOAP.

Automated Submission. In many cases, it may be more convenient for users to submit forms directly to an automated system instead of e-mailing them. Forms can be directly submitted via HTTP POST or SOAP.

The Adobe XML architecture is consistently implemented across the Adobe Acrobat family, Adobe Reader, and Adobe server products to deliver a complete solution that connects documents, people, and processes across the enterprise.

Leveraging the New Architecture

Adobe's new generation of PDF is a complete solution platform for processes based on intelligent documents. As in the past, the PDF and XDP specifications will be published and made publicly available. In addition, the software development kit will include XML features for the new generation of forms and interfaces for common languages such as Java, JavaScript, and Visual Basic.

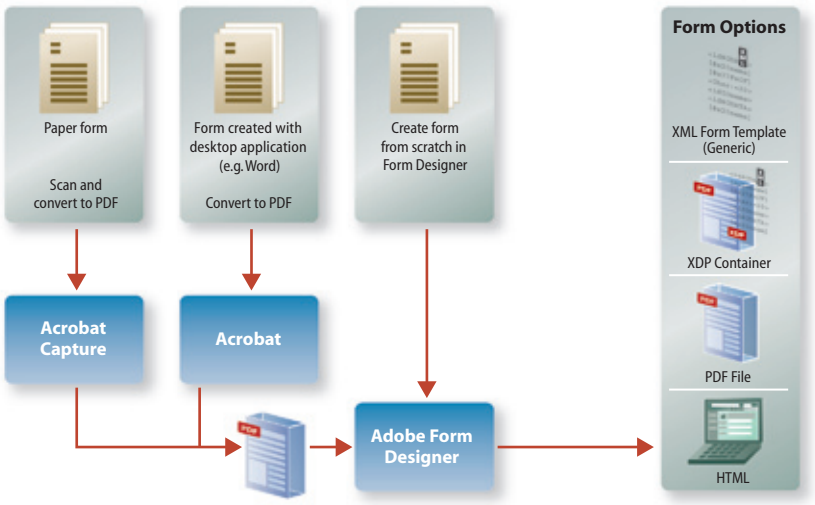
You can start taking advantage of PDF and XML electronic forms capabilities and proceed incrementally as time and budget allow, while achieving ROI at each step with increasing levels of forms sophistication.

- As a first step, simply converting paper-based forms to PDF reduces the distribution costs by enabling forms to be e-mailed or downloaded from a Web site. Users can then print a form, fill it in, and send it back. Paper forms can be scanned and converted to Adobe PDF using Adobe Acrobat Capture®. Forms can also be converted into PDF from native file formats such as Microsoft Word.
- As a second step, users can fill in forms online and then print them. This step improves the legibility of the captured information and, when calculations and validations are included in the form, the accuracy of the data is improved.
- The third step is to enable electronic submission of the form, including applying digital signatures. The XML data can be directly integrated into back-end systems, thus reducing errors associated with rekeying.
- The final step enables forms to initiate and continue through complex business processes, with data being added to the form as it moves through the workflow and integrated into multiple systems along the way. Advanced network services, such as Web services or data connectivity, can also be addressed for more sophisticated business interactions.

With current Adobe Form Designer, you can import Adobe PDF forms that have been scanned or converted from other file formats or design them from scratch using a graphical interface that provides precision layout capabilities. Form authors can then add intelligence, digital signature capabilities, data binding, and more to the template. The form can be deployed in multiple formats, as shown in Figure 4.

When the new form designer is available, you can easily migrate existing forms to enable the support of arbitrary XML and the ability to save forms as either PDF or XDP.

Figure 4



Conclusion

The Adobe XML architecture provides a robust and scalable technology framework required for enterprises to exploit the unique capabilities of XML and PDF.

- As a standard language for exchanging data with enterprise systems, XML offers powerful and flexible data integration, business logic, and routing capabilities.

- PDF enables users to participate in any XML workflow by presenting data in a secure and high-fidelity format for human interaction, across multiple platforms and devices.

- By embedding XML data within a PDF document, users can work online or offline, initiate processes based on the data within the form, and archive data and documents for process integrity.

- By packaging XML and PDF in an XDP container, files can be manipulated with any type of XML tool to access document content in the most appropriate format for each process.

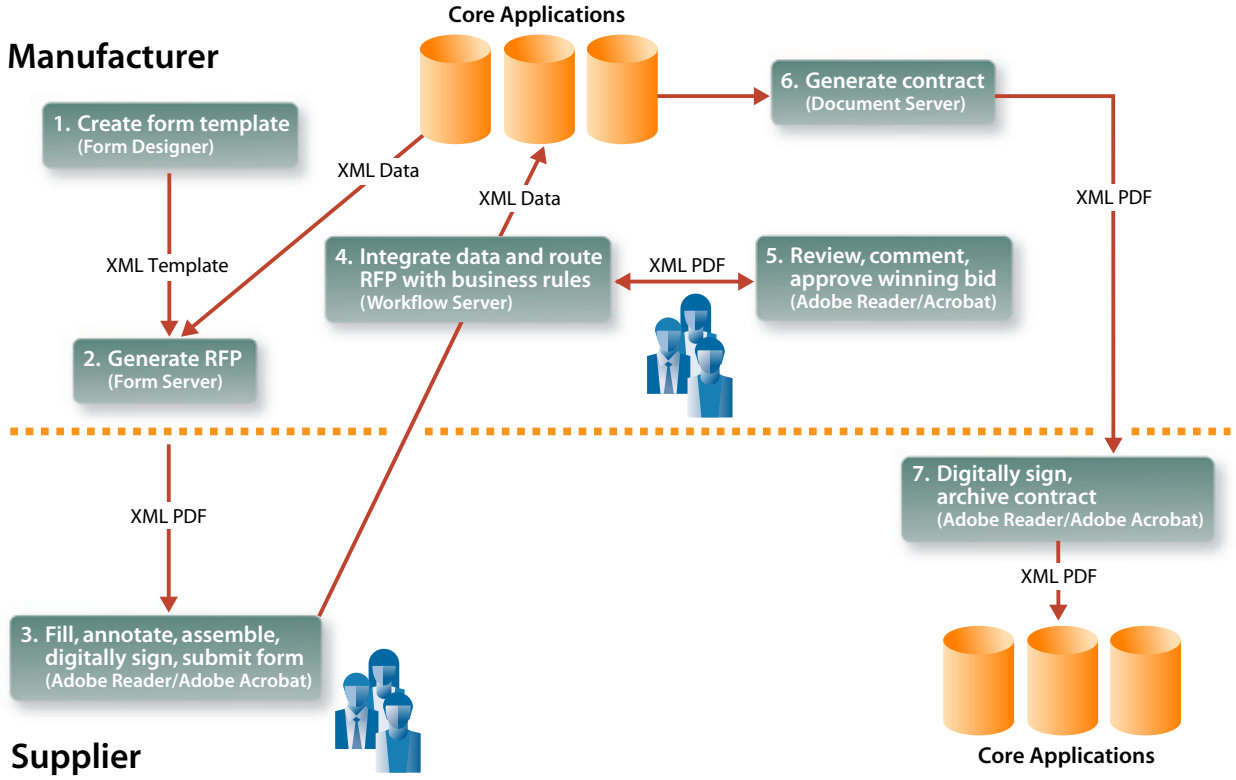
The Adobe XML architecture provides the fastest and easiest way for enterprises to move to fully automated and integrated document-based processes to improve customer service, reduce operation costs, and gain competitive advantage.

Appendix I: Applying XML to a Business Process

The following is a simplified example of a Request for Proposal (RFP) process used by a manufacturing enterprise that solicits bids from multiple parts suppliers. The company wants to automate the process to simplify and accelerate the evaluation process.

1. The manufacturer designs a form template once that can be used in future RFP processes.
2. The RFP is generated as a PDF file, prepopulated with each supplier's existing data, and sent to the prospective suppliers to fill in and submit.
3. Each supplier fills in the form, adds comments, assembles it with supporting documentation, and digitally signs the form so that the manufacturer can validate its authenticity and integrity.
4. When the supplier clicks the Submit button on the form, the data is automatically integrated with the manufacturer's enterprise applications.
5. At the same time, the completed document and attachments are routed to the appropriate people for evaluation. As the document moves through the workflow, it can be managed with automated deadlines and escalations to ensure timely processing. Data can be transferred between the document and the core applications at any point throughout the process, so the systems contain up-to-the-minute information.
6. When a winning proposal has been selected, a contract that combines data captured from the RFP with the manufacturer's standard terms and conditions is dynamically generated.
7. The supplier receives the contract in PDF to add a digital signature, and the locked-down document can be archived in the supplier's system as a document of record.

Figure 5



This diagram illustrates the steps above and shows how Adobe server and desktop products will work with the XML architecture throughout this example process.

Intelligent data capture

The new Adobe form design tool will provide an easy-to-use interface for designing XML forms that enable precise alignment of form elements, incorporation of rich visual elements, and the addition of business logic, such as calculations, validations, and form field binding.

The Adobe Form Server run-time engine merges XML data from enterprise applications, legacy systems, and other databases into the templates to be delivered across any channel: Web, e-mail, mobile devices, and PDAs. Because the Adobe Form Designer stores forms as an abstract description of the layout and behavior, Form Server can deliver the form in the format appropriate for any device (and browser version), without needing to create new templates for each channel.

Additionally, the XML forms can be deployed either as PDF documents or as a series of HTML screens to help users step through the form filling process. The XML data can then be integrated into applications through a variety of mechanisms, including directly to a database via ADO or ODBC; via Web services using SOAP; or to any server, including Form Server and Web servers using HTTP POST.

Business process management

When more complex processes require human intervention, Adobe Workflow Server uses a rules- and roles-based process model for managing workflows to ensure process integrity and consistency. Workflow Server includes a graphical designer that specifies the process in Figure 5, adding the necessary decision points and possible exception handling to this primary flow.

The design tool then generates an XML process description and installs it in Workflow Server. In addition, third-party tools can easily extract the stored data from the XML for purposes such as efficiency analysis and process optimization. As with Form Server, Workflow Server also uses XML for integrating data back into the core applications.

Adobe products can also integrate with third-party workflow engines, including those that are included with enterprise applications.

Dynamic document generation

Adobe Document Server generates, assembles, and manipulates documents that contain customized data, content, and graphics. In the process represented in Figure 5, Document Server generates the final contract. It could also be used to create the original RFP document, and automatically include related documentation, such as bills of materials, technical drawings, and demand forecasts.

Using XML for both the integration and business logic associated with document generation makes it easy for any application to use Document Server to create complex, personalized documents that contain dynamic content from enterprise applications and static template elements, such as corporate logos and boilerplate text. Document Server also leverages XML for presentation capabilities by supporting automated customization of content produced using World Wide Web Consortium (W3C) Extensible Stylesheet Language Formatting Objects (XSL-FO) and graphics produced using W3C Scalable Vector Graphics (SVG).

Client interaction

Whenever the RFP process requires people to view or interact with the data, they can use the Adobe Acrobat family or Adobe Reader to view the rendered PDF document.

For example, the supplier would use these clients to fill in the RFP form, apply a digital signature, and submit the form. The manufacturer could then route the completed PDF form to employees to review and comment on, while protecting the source content of the original documents. The information stored in the comments is stored as XML data as well, so it can be customized or used to interact with core applications.

Adobe Reader lets users view and print any PDF document, and fill in and submit online forms that have been created with Acrobat Professional 6.0. As well, files can also have additional rights enabled using Adobe Document Server for Reader Extensions, which lets users of Adobe Reader save forms locally, work offline, add comments, and apply digital signatures. This allows the manufacturer in the example to have any supplier participate in the electronic RFP process, with no additional costs for the supplier.

About Adobe

Founded in 1982, Adobe Systems Incorporated helped launch the desktop publishing revolution. Twenty years later, Adobe software products and solutions lie at the heart of Network Publishing, an equally innovative communications model. With Adobe tools for Network Publishing, enterprises, creative professionals, and home and office users can make reliable, visually rich information available to anyone, anywhere, on any device. One of the world's largest PC software companies, Adobe generates annual revenues exceeding US$1 billion. Today, over 3,000 employees across the world share Adobe's commitment to helping people communicate better. Headquartered in San Jose, CA, Adobe is traded on the Nasdaq National Market under the symbol ADBE.

FOR MORE INFORMATION

Visit *www.adobe.com/xml*

Adobe Systems Incorporated • 345 Park Avenue, San Jose, CA 95110-2704 USA • *www.adobe.com*

95001839 6/03

Adobe

Tools for the New Work™