

INDUSTRY DEVELOPMENTS AND MODELS

Active Documents: Changing How the Enterprise Works

Joshua Duhl

Susan Feldman

IDC OPINION

The role of information itself has played second fiddle to the technology that manages and finds it. "Active documents" change the role of information and how it interacts with both software applications and people. These new documents combine information with the processes that surround it. Using business rules, they embed processes such as retrieval, data acquisition, transactions, workflow, or archiving in the document itself. Rather than interacting with the application, users will interact with the document. This means that users will no longer need to learn the myriad of applications that fragment access to and use of information: The document becomes the de facto interface. As the potential inherent in active documents is explored, the changes in these interactions will be profound. This study examines what IDC calls active documents — a new kind of adaptive document that evolves as systems and information change. It discusses the impact they will have on companies, content infrastructure, the intercompany information chain, and the broader content marketplace. IDC believes that active documents will be a disruptive technology that will:

- Affect all business applications, business processes, and workflows — and in the process, make companies more responsive and adaptive
- Change the way in which companies are able to share, exchange, and integrate information with constituents of the intercompany information chain
- Blur the lines between dynamic content and programming
- Create a new category of applications for: integrating systems; gathering, presenting, and analyzing information; and personalizing content
- Create new opportunities for companies to provide software tools that create, track, manage, and archive active documents

IN THIS STUDY

What Are Active Documents?

Imagine an aircraft manual that is both reference book and maintenance record. Hook it up to supply catalogs to automatically restock depleted inventories every time a part is taken off the shelf and used to repair the aircraft. Track the repairs made and have the document offer related assemblies that need to be checked if one is found to be worn. Include interactive diagrams that can answer questions or offer alternative views for the technician and the aircraft designer.

Electronic documents are evolving into unprecedented roles. New technologies that utilize XML are emerging that enable documents to become more than the static, passive containers of words, graphics, and images they have been historically. They are business process plus information — archive and interactive knowledge base rolled into one.

These new documents are living and dynamic. They can interact with other content, data, applications, and documents. They are highly structured, yet they are capable of changing their structure and presentation on the fly for different users, contexts, or uses. Some actively launch corporate or interenterprise workflows. With varying degrees of embedded business logic programming, they can invoke functions at various points in a workflow process to automatically gather, retrieve, extract, integrate, and analyze information from a variety of sources, and send it on if needed. Some are defined to relate to other documents and to incorporate pieces from those documents on the fly to create new documents, making them adaptable and able to automatically, dynamically evolve throughout their lifetimes or the lifetimes of the things they describe.

Active documents are so new that there is no standard for what they look like or what they do. They can be found today in various forms and sizes, and they serve a variety of purposes: A simple active document might be a purchase order that also validates inventory levels and perhaps triggers a reordering of the item, or suggests alternatives for out-of-stock items. However, an active document can be as complex as an aircraft maintenance manual or other technical documentation that self-updates with approved changes from multiple component vendors as updates are made to components and repair processes.

Active documents can:

- Participate in transactions between companies — providing both the interface to the user and to the content — automatically routing themselves and participating in an electronic supply chain business process
- Be launched from other active documents (e.g., a maintenance manual that allows users to order parts directly)

- ☒ Include both text and interactive graphical elements, such as diagrams and graphs, allowing users to zoom or fly through a wire-frame model, or drilldown into a sales data visualization

Fundamentally, active documents are a disruptive technology, bringing with them a wave of sweeping change that will affect not only the software systems that create and interact with them, but also the human systems that define workflow and business processes. IDC believes that this new advance will have a dramatic effect on companies, content infrastructures, the intercompany information chain, and the broader content marketplace.

SITUATION OVERVIEW

Active Documents Defined

Documents are discrete containers of information expressed in words, sounds, or pictures. Active documents are XML documents that combine highly structured XML with executable behavior. The code, which may be XML itself or another scripting language, is contained in or attached to the document, but executed by an external processor. To the user, however, it appears that the document contains both information and the actions that are necessary to act on the information. Thus, a document can launch a workflow process or collect data from a database and populate a table so that the most recent stock quotes, for instance, that show stocks under \$20 per share whose prices have increased or decreased by x% are displayed. These documents can create audit trails of actions taken based on their information. They can automatically aggregate data from multiple active documents, electronic forms, or data sources to provide a composite, roll-up, or summary view. They can update themselves as technical manuals change. In other words, the document becomes dynamic and active in its own representation, interactions, and presentation. The point is that the document now becomes, for the user, the gateway into the multiple applications that are related to the information. Users eventually will not have to know how to use the database, search, or ERP applications: They will interact with the document. This will streamline knowledge work enormously.

Active documents blur the concept of what has traditionally been a "document" by adding code and actions to the document, and by extending the document to encompass links to related information as well as to collections of data or to processes. For example, an active document may discuss the history and financial outlook of a company, and automatically update the stock price of the financials from a database.

Because they mix code and content, active documents are a good example of a trend we note at IDC: As we move toward a new "information worker" (to use Microsoft's terminology) environment, the next generation of applications have begun to mix not only code and content, but types of technologies as well. The lines between database and unstructured information applications, portal, and collaborative applications are no longer clear.

XML and XML-based structures are the key to active documents. XML can represent almost any data or content. Since it is text, an XML document can function as a container for content and metadata (information about the content), or it can be an executable program (SVG) where elements function as computer instructions and associated data. As a result, the distinction between dynamic content and programming is beginning to blur.

Active documents expand the concept of what a document is: The document may modify itself and affect other things around it through the execution of the attached behaviors (code). Active documents are able to interact more directly with software and devices in their environment, and to adapt themselves over time to changes in the software environment. As the nature of XML forces the separation of presentation from content, these documents are able to adapt the visual presentation of their content through embedded code such as scripts, templates, or other XML-defined transformation documents written in XSLT. Active documents can also participate in structured external workflows and be routed or guided by internal rules, actions, or workflows embedded within the document (as code or textually defined directives).

Today, active documents are confined largely to active forms, making validation and function calls to databases or to Web services for retrieval. These are early days for the concept. The features vary widely from one product to another, and the applications may be difficult for novice users to master to develop anything more than a simple active document. To create a complex document that has a complicated workflow, is a compound document, or that integrates with many back-end systems, today's active document author must know where the desired content resides, how to make calls on that repository or ERP system, how to write scripts, and how to create a logical flow of actions or workflow. To create effective active documents, the author must understand the workflow within or across the organization(s) thoroughly. This complexity cries out for specialized services engagements. We can also expect new roles — perhaps that of a "business process analyst" — to arise within the enterprise.

In the future, these documents will contain both the relevant information and the code to launch such processes as review, approve, initiate action, secure, retrieve, trigger, notify, track, monitor, archive, send email, route, procure and initiate a transaction, update, call an external service, create data summaries, roll up, print, or visualize information.

The key elements of active documents, then, are that they are:

- Interactive — and interactive with both the user and with a variety of underlying applications and/or data and content sources
- Containers for both information and code, and process
- Dynamic, changing as either the information or the process changes

Today, uses of active documents applications fall loosely into three categories: active forms, active narratives, and visual active documents.

Active forms are intended for use in business processes and workflows. They gather data and enable business transactions. Representative applications include Microsoft

InfoPath, Adobe Form Designer, Adobe Form Server, Adobe Document Server for Reader Extensions and Adobe Reader 6.0, FileNet Forms Manager, Cardiff Software's LiquidOffice, PureEdge's XML E-forms, and other electronic forms and data gathering tools. They participate in and affect business processes and workflows between front-end tools, back-end systems, and humans.

Active narratives are intended as long-lived, ever-changing, always-current documents. They describe the most current state of a physical thing or process. These documents tend to be large and are created through more traditional publishing processes, using XML authoring tools such as Corel XMetaL, Arbortext Epic, or Adobe Framemaker, and XML document repositories such as Astoria Software's Astoria CMS, CambridgeDocs' Content Transformation Server, IXIASOFT's TEXTML Server, and XyEnterprise's Content@ and XML Professional Publisher software. Through purely XML technologies, or XML combined with behavior, they can change, update, and modify their structure and content over time. This keeps the most current version of a long document current. Because they are tagged, content search, extraction, and replacement become a fundamental part of their existence, with the ability to extract and replace, in whole or in part, virtually any part of their content.

Visual active documents are predominantly visual rather than text based, like the first two categories. They use XML-based presentation languages such as Scalable Vector Graphics (SVG) to define dynamic two-dimensional interfaces and scenes. When combined with dynamic data, they present highly interactive graphics — diagrams, charts, or images — that enable the user to navigate large and complex sets of information. All of these categories can be intertwined. For example, narrative documents can launch active forms: A parts order form could be launched, populated, and sent directly from the manual that describes how to install or troubleshoot a malfunctioning refrigerator. Narratives can be affected by an active form (or a series of them), so that aggregated engineering change orders will prompt automated or human changes to a narrative document. The data in a form can be presented by means of a visual document, and part of the content of a narrative document can be visual (e.g., an interactive diagram of an engine assembly). It is even possible to make text somewhat visual; for example, mousing over or clicking on a description of an engine part would highlight or animate the corresponding part in the diagram. If a concept-based language analyzer is included in the active document infrastructure, the ideas in a paragraph can be extracted and used to find and pop up similar paragraphs or tagged illustrations that are related.

Over time, with advances in linguistic-processing capabilities, active documents may gain the ability to change their structure, self representation, or presentation without human intervention.

The active document unites applications and acts as a de facto single interface to them. It brings together the major enterprise applications and their technologies, including database technologies, visualization, search and retrieval, language understanding, collaborative technologies, and document and content management systems.

Fundamentally, active documents are a disruptive technology, bringing with them a wave of sweeping change that will affect not only the software systems that create and interact with them, but also the human systems that define workflow and business processes within companies as well as between them.

To qualify as an active document application, IDC specifies that the application must:

- Embed code and/or business rules within the content of a document
- Enable a user's document interactions to launch actions, make calls on other applications, or initiate processes without express commands being given at the time of use
- Change content based on actions that were taken
- Optionally, create an audit trail and provide rollback and versioning to create a history of the changes and the actions that were taken at any point
- Change appearance based on actions that were taken; the user's rights, roles, or privileges; or as determined by the document's content itself.

Products on the market today that fit the description of an "active documents application" include XML authoring tools, forms design and layout tools, XML content repositories, schema design tools, and some initial document transformation tools and technologies.

FUTURE OUTLOOK

There is no doubt that the current information or knowledge work environment is ripe for change. Enterprises realize that they do not have a handle on their information, and that their processes are a series of unrelated actions confined by the separate silos of each enterprise application. They realize that substantial time and effort is wasted on cumbersome or overwhelming paper-based processes, manual information gathering and data entry, erroneous data entry, redundant copying and pasting of textual information, changing the form or output appearance of content, and slow form processing infrastructures and workflows. Moreover, a strong business case can be made for uniting and streamlining these processes with the information that is related to them across all applications. In the current status quo, information is lost, and knowledge workers are simply wasting too much time moving from one application to another. This time cuts deeply into worker productivity, and it puts the organization at risk if it acts on incomplete information. For this reason, we see recent developments like Adobe's Acrobat 6.0 and Microsoft's Office 2003 acting as the leading edge into a new work environment. Whether enhancements to these applications or a new, entirely different approach to organizing information and process interactions will emerge is an open question.

Will the arrival of a natural user interface (NUI) with a different, supporting information infrastructure disrupt active documents? For a NUI and related infrastructure to develop, we believe that more robust, practical agent-based systems, better speech recognition, haptic interfaces, and above all a better, standardized human-computer

interface must be developed. Today, the practical application of technologies that would support a NUI appear to be too far in the future for us to expect the NUI to arrive anytime soon. In the interim, the active document appears to be a solid step in the right direction. Potentially, it could unite or dislodge the established enterprise conglomeration of applications that we see today.

Active documents are a new technology. They will form the basis for task-specific horizontal, vertical, or industry-specific workflows that can be customized more easily. Because of the nature of this foundation technology, IDC expects it to drive the adoption of XML and Web services standards faster. In fact, active documents will become the user interface to Web services.

ESSENTIAL GUIDANCE

To survive the shift to a single enterprise interface, software vendors must be prepared to:

- First, XML-enable their applications, then XML-orient their applications
- Continue to expand the ability to embed code and define rules within these documents, and make this easier for nontechnical developers to do
- Provide the Web services that will link their applications to active documents
- Establish partnerships with services and software vendors driving active document adoption

Service vendors must:

- Understand the impact on the enterprise that this development will have
- Advocate adoption of active documents, and develop compelling demonstrations of and uses for them
- Develop centers of competence and expertise in active document-based workflows
- Incorporate active document-based solutions into existing business process reengineering, workflow, and content management practices
- Establish partnerships with software vendors that are active document-enabled

LEARN MORE

Related Research

- Microsoft Office System 2003: A System Is Born?* (IDC #30319, October 2003)
- Worldwide Interactive Data Visualization Tools Forecast, 2002–2007* (IDC #30215, September 2003)

- ☒ *Worldwide XML Authoring Software Forecast, 2003–2007* (IDC #29399, May 2003)

- ☒ *Worldwide Authoring Tools Market Landscape* (IDC #28775, February 2003)

Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit www.idc.com to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/offices. Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or sales@idc.com for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2003 IDC. Reproduction is forbidden unless authorized. All rights reserved.

Published Under Services: Content Management and Retrieval Software;
Streaming and Rich Media Software