Center for
Technology in Government

Using XML for
Web Site Management:

An Executive Briefing on
streamlining workflow,
reducing costs, and enhancing
organizational value

May 2007

Center for Technology in Government
University at Albany, SUNY
187 Wolf Road, Suite 301
Albany, NY 12205
Phone: (518) 442-3892
Fax: (518) 442-3886
E-mail: info@ctg.albany.edu
www.ctg.albany.edu

Overview

A typical government agency Web site contains thousands of pages and links, online transactions, and critical reports. It needs to be accurate, up-to-date, and available 24/7 to a wide audience from many locations using different devices. Unfortunately, the technologies and processes used to establish Web sites have not kept pace with the efficiencies needed to manage them today and into the future.

Much of the problem is due to the fast pace of change. The computer age is a phenomenon of the past half-century, the desktop PC of the past quarter-century, and the Internet just the past decade. Within that time, many ways of creating, storing, and managing information have come and gone (see Figure 1).

Most Web sites today exhibit similar longevity issues because they are structured with individual HTML files, proprietary databases, or content management systems tied to specific software and hardware.

An XML-based Web site offers a solution because XML is not owned by a specific vendor or dependent on specific hardware or software. It is simply a standard for structuring content (e.g., data, text, images, etc.), and presenting that content in multiple ways (such as Web pages, documents, mobile device displays, etc.). The open-standard, non-proprietary format of XML does not become unusable as technology advances because it is designed for change. It makes it possible to react quickly and reuse content as needed. Organizations using XML realize benefits in **information consistency, data longevity, workflow management** and **productivity.**

"Information outlives technology ... and yet, as of today, too much of our intellectual heritage is tied up in fragile, proprietary, binary word processor files ... XML is the solution."
Tim Bray, co-inventor of XML and director of Web technologies at Sun Microsystems

Overview (cont.)

Figure 1. Lifespans of Computer Hardware and Word Processors
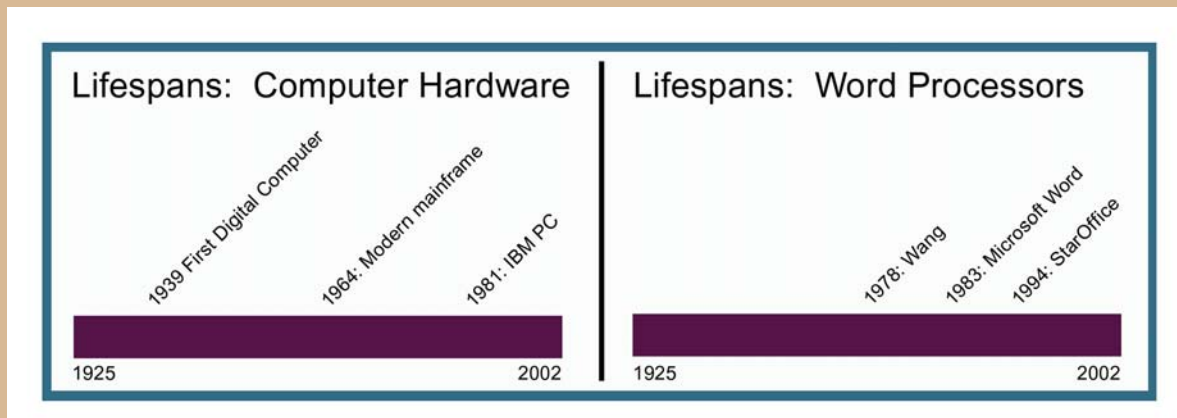
"XML, Information Technology, and Intellectual Capital," XML Testbed Project Day on January 25, 2006 by Tim Bray, Director of Web Technologies, Sun Microsystems.

The practical impact can also be seen in some key issues addressed by an XML-based Web site:

■ **Accessibility**

Section 508 of the Rehabilitation Act of 1973 and NYS Policy P04-002 require Web sites to be accessible to persons with disabilities, but can be very labor-intensive bringing thousands of non-compliant HTML pages into compliance. XML can ease that burden since Web pages are generated automatically and uniformly. A change in one file can bring dozens or hundreds or even thousands of Web pages into compliance.

■ **Enterprise Perspective**

The single-source orientation of XML lends itself to sharing data across an enterprise. Rather than creating another information silo in a Web site, an underlying XML structure makes that Web site available throughout the enterprise.

■ **Collaboration and Integration**

Because XML structures information in an open, standard, and sharable format, it allows for easier collaboration and integration within and across organizations. As the use of XML expands, these opportunities will increase and prove more beneficial.

■ **Positioning for the Future**

The Web is already moving away from the desktop to mobile and wireless devices. New devices require new formats and standards to effectively transmit and display information. As an open specification, XML accommodates any new format or device.

Understanding the Problem

The root problem in Web site management is fundamentally one of content and workflow. The content that appears on the Web typically comes from many different sources and appears in many different formats (HTML pages, PDF documents, etc.). Workflows are established to keep track of these source files and their Web displays while keeping the content consistent, accurate, and timely. However, these workflows usually comprise many redundant tasks as shown in Figure 2, and it becomes difficult, if not impossible, to ensure that all information is reliable and up-to-date.

Using XML for Web site management, on the other hand, attacks the root problem directly. Content is not stored in multiple formats and locations to be transformed by the Webmaster into formats suitable for the Web. Instead, XML maintains content in a single, standardized, non-proprietary source that is accessible to users in whatever format is required—for word processing, as an HTML page, PDF document, etc. Workflows are simplified and resources optimized since the redundant tasks associated with tracking and converting multiple source files are no longer needed as seen in Figure 3.

Figure 2. Workflow in Non-XML Based Web Site

Multiple Workflow Dependencies and Overlaps
Multiple "Source" Documents (RTF, PDF, HTML)
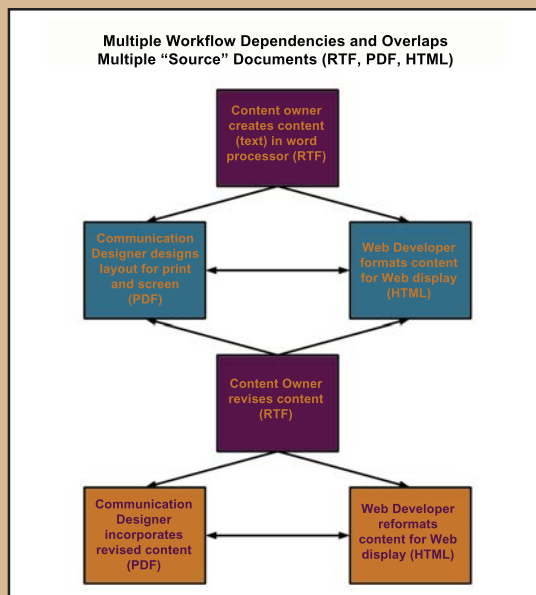
The diagram above shows how much activity is spent passing documents back and forth while trying to keep them all consistent and up-to-date. Many of the tasks consist of manual reformatting of the content for the Web, while checking that it's still accurate.

Figure 3. Workflow in XML Based Web Site

Fewer Workflow Dependencies and Overlaps
Single "Source" Document (XML)
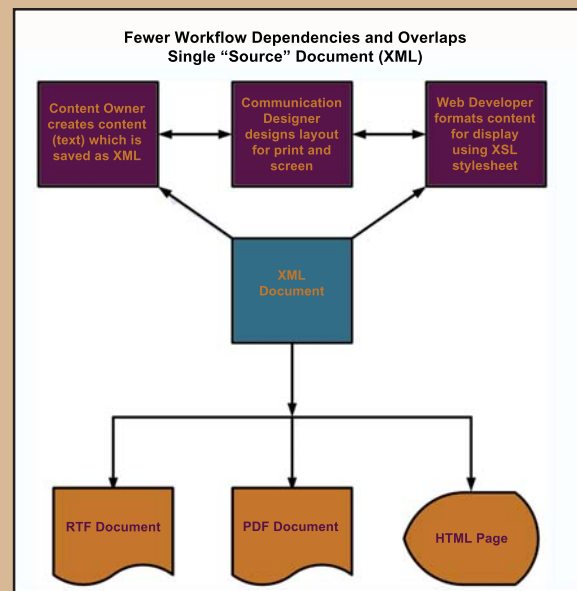
The diagram above shows how the single XML source document at the center of the process eliminates much of the redundancy and checking activity associated with the non-XML based workflow. Many of the manual tasks are automated.

Realizing the Benefits

Once the root problem in Web site management is addressed through the use of XML, key benefits occur:

■ **Consistency**

Information is a primary asset of an organization. But managing information assets can be complicated. The same content that appears on the Web site may also appear in a printed publication or in a PowerPoint™ slide show; it may have originated in an email or a press release; or it may be available in HTML, PDF, and Word™ formats. In all these cases, the information must remain consistent, accurate, and reflect a positive public image. XML guarantees that consistency since all content comes from a single source regardless of where it appears or in what format.

■ **Integrity and Longevity**

Perhaps the greatest benefit of XML derives not from how information is distributed today, but how it is preserved for tomorrow. Every organization has experienced problems with data trapped in proprietary, obsolete, and unusable formats— software that is no longer supported and technologies that have been left behind. XML, however, is open standard and independent of any programming language or device, so its lifespan is not tied to specific hardware or software. With XML, information is maintained separately from its distribution format, which ensures information integrity and longevity.

■ **Workflow Management**

An XML-based Web site offers a better approach to information management concerns. By standardizing the format of Web content, streamlining the workflow involved in its production, and automating its distribution and maintenance, XML eliminates much of the administrative overhead and uncertainty associated with Web site management. A more efficient Web site management system, based in XML, translates to improved responsiveness and better services delivered from the site. Information can be updated more easily and consistency is guaranteed.

■ **Productivity**

An XML-based Web site allows for smarter use of human resources and IT skills. Some characteristics of XML, such as re-usability of content and automatic creation of outputs, lead to streamlined business processes that in turn produce significant time and cost savings. The automation of tasks and elimination of workflow redundancies increase the productivity of an organization's IT workforce. Their time can then be allotted to higher value projects that deliver a greater organizational return-on-investment. In addition, low maintenance costs make a Web site more sustainable over time.

Figure 4 illustrates an example of the benefits in converting to an XML-based framework. The example is taken from one of the five New York State agencies that participated in CTG's XML Testbed in which they developed prototypes and business cases for converting their Web sites to XML.

It demonstrates how the Office of Cultural Education, New York State Education Department designed its prototype to:

- **Standardize its content in a single-source document**, which ensured the integrity and consistency of that content since it resided in only one place.

- **Generate multiple delivery formats** for different appearances and uses of that content on the Web site.

- **Streamline the workflow** associated with content creation and Web page delivery since most of the work is done automatically via XML.

**Figure 4. Benefits of Converting to an XML-based Framework**
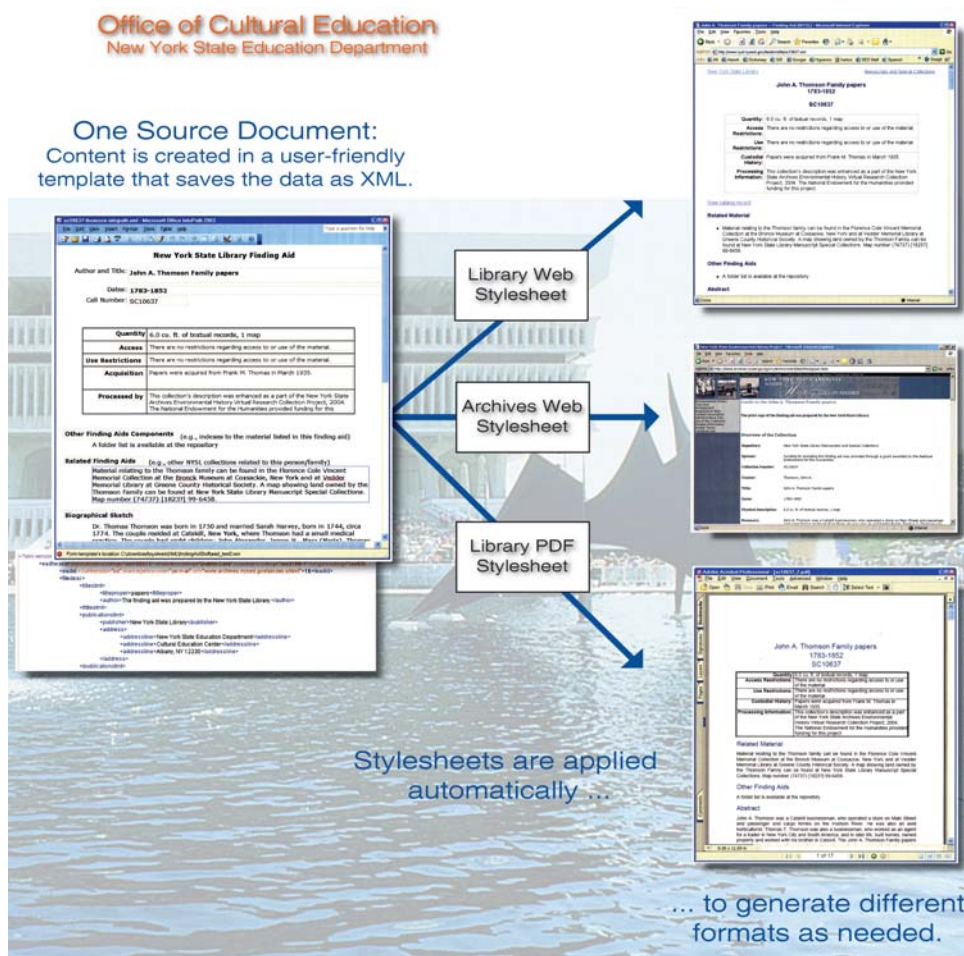
Assessing the Value

Here are some key factors to consider when determining whether Web site management practices provide a good return on time, money, and resources invested:

- **Accuracy** of information.

  √ Are mechanisms in place to guarantee that all information on the site is correct?

  √ Can inaccurate information be identified and removed or corrected efficiently?

- **Consistency** of message.

  √ Does the Web site information conform to similar information in other formats, such as printed materials and presentations?

  √ Do the Web pages present a consistent image and can the look and feel of the site be easily maintained?

- **Timeliness** of updates and new postings to the Web.

  √ Does information appear when, where, and how it should appear?

  √ Is some information not making it to the Web simply because it takes too long?

- **Integrity and ownership** of content.

  √ Do the words and images—the information—on the Web site reside in multiple organizational locations and multiple formats and are they uniform throughout locations?

  √ Is the information in an open, accessible, non-proprietary format that will be usable in the future regardless of technological changes?

- **Business processes** involved in getting information to the Web.

  √ Are they formalized and fully understood by everyone involved?

  √ Do they contain redundant, low-value tasks?

  √ Are they structured solely or mostly to work around shortcomings of the technology?

- **Strategic positioning** for now and the future.

  √ Does the Web site meet the current Section 508 accessibility guidelines for persons with disabilities and can it be easily kept in conformance with new policies and mandates?

  √ Does the technical infrastructure accommodate the mobile devices, security, enterprise integration, and cross-organizational collaboration that characterize the Web of today and tomorrow.

A Web site based on XML addresses these factors at a lower cost with higher long-term value as compared to traditional HTML-based sites or proprietary content management systems. Figure 5 illustrates CTG's own ROI (return on investment) in converting its Web site from an HTML to an XML basis.

Figure 5. Return on Investment for CTG in Converting to an XML-based Web site

Cost Trends with Existing Web Technology
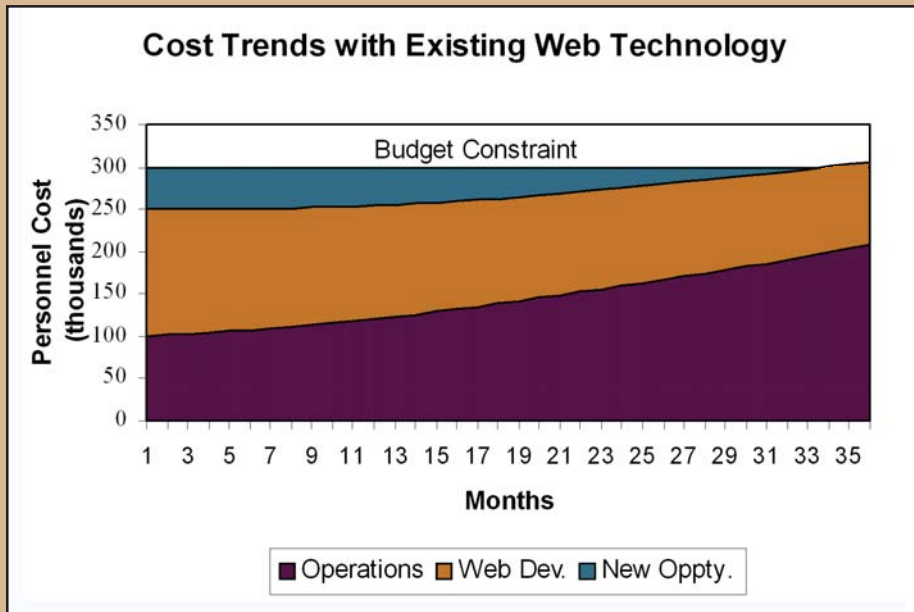
Non-XML Based Web Site

CTG found that continuing to maintain its Web site with existing HTML-based technology would squeeze out opportunities for new technical projects as Web staff would be forced to devote increasing amounts of time to operational maintenance. This was not a viable strategy for future success.

Cost Trends with New Web Technology
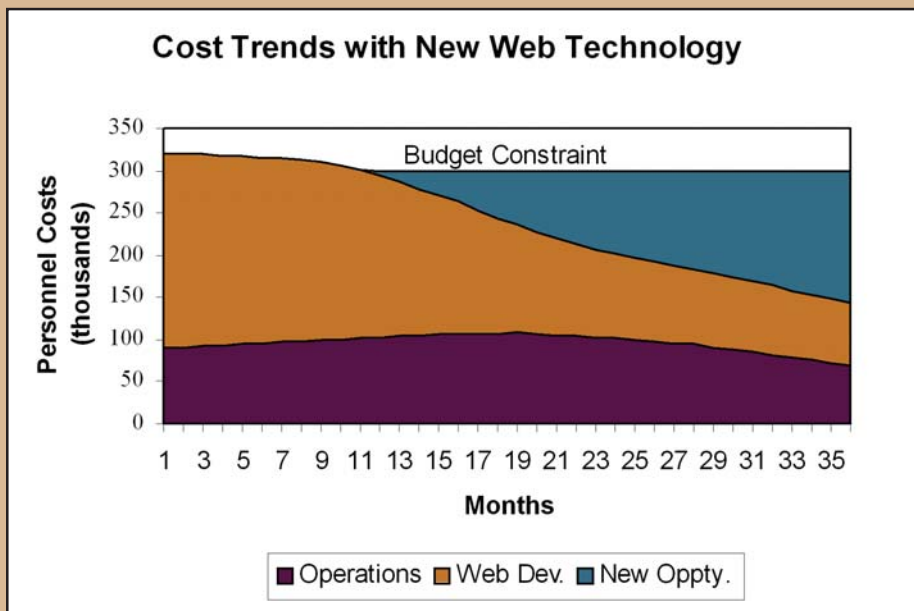
XML Based Web Site

By converting to an XML-based technology for its Web site, CTG saw the prospects for new development opportunities enhanced dramatically. As routine maintenance tasks were streamlined and automated, operational activity leveled off as a small percentage of overall time. Productivity increased while budgets remained steady.

Making it Happen

Many of the new directions and developments on the Web have a basis in XML, which is becoming a critical technology for all types of information services. The features of XML that have been emphasized throughout this Executive Briefing—open standard, reusability, technologically neutral—make it an ideal strategy for preparing for the future, while achieving efficiencies today.

Executives have an organizationwide perspective on their Web sites. By developing an understanding of XML, they can make better decisions that lead to the benefits outlined in this briefing. The key is to create an environment in which adoption and implementation of XML for Web site management is seriously examined and integrated into an organization's strategic plan.

Additional Resources

This Executive Briefing is one component in a suite of publications and resources that CTG has produced as guides for the use of XML in Web site management. The other components, referenced below, provide a more in-depth look at the organizational, technical, and business implications of using XML and are directed specifically to the management and staff involved in these areas. Refer to these additional resources for more information:

■ For further details on how to begin adoption and implementation of XML for Web site management, see *Using XML for Web Site Management: Getting Started Guide* (www.ctg.albany.edu/ publications/guides/ xml_getting_started).

■ For more technical tips and resources, see *The XML Toolkit* (www.thexmltoolkit.org/), a Web site containing a library of XML resources, including detailed examples for implementing XML in any Web environment.

■ For additional information on how adopting an XML-based Web site provided a good return on investment to CTG, see *Return on Investment in Information Technology: A Guide for Managers* (www.ctg.albany.edu/ publications/guides/roi).

UNIVERSITY
AT ALBANY
State University of New York