



Multilingual

#24 Volume 13 Issue 4

COMPUTING & TECHNOLOGY

ON-LINE TRAINING
ACROSS LANGUAGES

Page 24

THE HUMANITARIAN
FACE OF TRANSLATION

Page 34

FONTS & SYSTEMS

DETERMINING LANGUAGE-RELATED
INFORMATION FOR FILE SYSTEMS

Page 53

A MERGER OF CONTENT MANAGEMENT
AND LOCALIZATION WORKFLOW

A MERGER OF CONTENT MANAGEMENT AND LOCALIZATION WORKFLOW

Algonquin Studios and E-Merge Strategies combine business models and software products into one multilingual solution

GREG NORTON, ADRIAN ROSELLI & DAVID THIEMECKE

In December 2001, Algonquin Studios and E-Merge Strategies completed a merger of two different companies with different products and targeted markets. Algonquin Studios had developed the QuantumCMS content management system, which was designed to support multiple Web sites on one platform. E-Merge Strategies had released a localization workflow software called Transmerge in early 2001. Once the paperwork was complete, the two companies had to bring together their different business models as well as their different software products. The participants discovered that the merger of the two systems paralleled what is going on in the overall Web development market: content management and localization are being intertwined as a total multilingual content management solution.

TRADITIONAL CONTENT NEEDS

The growing consumer demand for accurate, timely information via the Internet, plus the downturn in the world economy, has led to a need for a reduced-cost way of managing and maintaining corporate Web sites. Once it was the purview of HTML developers and Webmasters. Now, companies need a way to allow the people who directly produce the content to be able to support that content on the Web site while maintaining a consistent look and feel to promote a strong corporate brand. Their needs may include delegation of content authoring, ease of use and responsibility for approval. Content management systems have begun to fill this need. They range from high-end solutions that are largely enterprise-class document management systems to

the free or inexpensive WYSIWYG tools that manage HTML files.

WEB SITE CONTENT MANAGEMENT

How does a company keep its content fresh? This problem plagues organizations of all sizes. Most organizations understand the need to update their sites with regularity, but do not have the resources to do it effectively even when content is available. Often, one or two people on the staff know HTML, but maintaining the Web site is not their primary responsibility. Finding ways to reduce the skill set necessary to maintain the content can solve the problem, but it is important that the solution allow more sophisticated users to work with the system to provide advanced layout and integrate other applications into the Web site.

ENABLING AUTHORS AND DEVELOPERS TO DO THEIR JOBS

The Algonquin solution is to offer a limited-function WYSIWYG editor in the Web-based authoring tool. By focusing on the standard word-processing features such as bold, italics, bullet and ordered lists to which users are already accustomed, the developers can involve more people in an organization in the content management process. The key is to limit the ability of non-HTML-savvy users to impact the overall design applied to the site through carelessness. Users can select standard Web fonts and sizes, but these HTML styles



are designed to be overridden by a master stylesheet if a site-wide designer wants to maintain control. Similarly, an underline feature is not offered in the WYSIWYG because underlines are generally reserved solely for hyperlinks.

An HTML source editor allows more sophisticated users complete control and the ability to use advanced HTML, applets, client-side scripting and so on. Not only can anyone author content, but also, depending on the user's ability, he or she can style it and even insert his or her own HTML, while the system ensures it doesn't override the overall brand and usability of the site.

A distributed authoring and approval process allows large institutional and governmental clients to delegate authoring of their public, intranet and extranet sites on a site-by-site, section-by-section and page-by-page basis. Each section includes an approval process of author, editor and publisher. Some individuals may be assigned more than one of these roles. So, many users with specific business or

product knowledge can create and edit their own content, while site-wide publishers can address political concerns. Content can also be organized in multiple ways to meet the needs of various departments, thus allowing complex presentation and workflow processes.

BRANDING AND DESIGN

In addition to offering fresh content, sites also have to look up-to-date. Keeping the look current can mean anything from ensuring that a newly updated logo replaces an old one to making sure that an overhaul of the look and feel of the site applies to every page. Content management systems that separate the presentation from the content allow the greatest flexibility. The style can be changed completely without a need to modify layout in every page of the site.

The current buzzword in separating presentation from content as well as cross-platform data transaction is XML. Combining XML output with XSL stylesheets allows the delivery of rapid design changes, branded sites and even seasonal branding.

One strength of the use of XML and XSL for shuttling data and parsing into HTML is that the software can support a number of other output methods besides HTML such as WML, HDML and XHTML. Content can be output to RDF/RSS feeds for use as syndicated news feeds for other sites. Even print can be an output for content from QuantumCMS now that applications such as Adobe InDesign support XML data feeds into InDesign templates.

CONTENT TYPES

Many content systems focus on “page-level” content, meaning that users edit an HTML page and all the content on that page. Algonquin took a different approach by organizing its content system around concepts. This is the idea that users would want to create content based on types — raw documents, news stories, articles, biographies, products and so on — and would want to be able to manage and compare individual elements of these concepts. Combining this with the ability to reuse content in different styles allows users to construct multiple views appropriate to different areas of a site.

One of Algonquin’s primary marketing tools is the case study. Using structured content in the system to group graphics and captions together into a consistent content framework, Algonquin can render a view on its own home page of a highlighted client, offer secondary captioning as sidebars throughout the site and present the full case study. Since all the content is tied together, it can be updated once, and the change will be reflected throughout the site. The same content for case

studies is also used on the QuantumCMS marketing Web site, although it has a completely different look and feel. This approach to structured content also allows clients to create views which compare and contrast content without having to construct complicated searches.

ALTERNATE AND LEGACY DATA SOURCES

The final bogeyman in the content management world is how to integrate data from systems outside of new content solutions. Most organizations have legacy data stored in third-party or proprietary systems and have no intention of or interest in moving their core business data into a system that manages their Web content. At the same time, these companies may want to make this data available to clients, vendors and employees through existing Internet or intranet channels. Algonquin provided a number of options for integration, from a tight systems level that maps objects in the content system to objects in the third-party applications to a style-only integration method that mimics the look and feel of existing sites. One large government client’s parking ticket and fee data, for example, stored in an NCR Teradata database management system was integrated by having the content system construct the navigation, style and other elements of a page and then render a traditional Web application in that style.

LOCALIZATION WORKFLOW

E-Merge developed its localization workflow system with the middle-market translation buyer in mind: companies that run their own translation projects using a mix of internal and external resources. The mix often changes from language to language as well as from project to project. These middle-market companies usually cannot afford or do not see enough return on investment to integrate an enterprise-wide translation system into their translation processes. The goal was to provide a tool that leveraged the internal assets that the company may have for some or all of the languages and that integrated the external resources into the workflow in a manner that did not tax the management capabilities of the company.

CENTRALIZATION

When asked what was the most “painful” part of the translation process, our clients continually answered: managing the communication. The decentralized nature of translation production requires communicating with multiple workers, often in multiple locations and time zones.

To simplify the approach, the basic idea is that all final work must be centralized. If the work is centralized, user information must also reside in a central location to enable consistent tracking and communication. The user’s e-mail address must be current and in the system, and the system must be the final delivery point for all work. From these two basic rules an orderly process of localization can begin, and from this point consistent project tracking, auditing and scheduling can be completed.

CATALOGING OF CONTENT

The next most common frustration that clients voiced was that their translations could not be used from one project to another. Their content was being captured in one output media or another, and the correlation between source and target text was not being imposed during production. In order to allow reuse, another rule was applied to the process: all text was to be presented with its corresponding target translation. Context would be provided as a tertiary option to the viewer. Taking content down to the basic level of source-target pairs provides the building blocks for flexible multilingual content creation.

TOOL DEVELOPMENT

When this Web-based system’s architecture was being implemented, network-based translation systems were already well developed and entrenched on translators’ desktops. So, the focus was on making sure that the Web-based system worked with the existing translation systems. Users in the system could use their own tools as they wished; and storing all content in unformatted source-target pairs ensured that work downloaded from the system would be readable by any translation tool. The focus of development was on the tracking of the work, not the work itself, and on administration of the system. The system that was developed was open enough for integration with other existing systems, but had enough regimented workflow to ensure an efficient production process from assignment through output and final delivery.

BRINGING THE TWO TOGETHER

Part of the integration of products is the integration of skills, teams and business processes. The new Algonquin organization began by addressing how to bring the products together while integrating the experience of both teams. The company polled existing clients who used one system to gauge their interest in the other. Government and health-care content-management clients saw immediate benefit in adding the localization workflow, while regional manufacturing clients, for example, did not. Some of the clients

who used the Transmerge system to manage their localization processes already had content management platforms, but those content systems either did not allow regional offices to manage their own content or did not create a simple process, if any, to manage localization. Most clients with content systems experienced both issues.

While integrating these two, the company also could not forget the translation vendors. Many organizations hire professional freelance translators to use the localization workflow. Causing a translator to take twice as long, no matter how robust the system, can immediately impact the other cost benefits of a merged system.

The combined sales force had to be aware of what developments were in the pipeline and needed to feel comfortable when reporting suggestions and feedback from clients, prospects and leads to the developers and executives. They also needed to keep the developers apprised of the marketplace and to help ensure that features were not developed that were not in demand while other, more high-demand features were still in progress. The marketing team had to help keep existing clients informed of plans for the product and to understand the value the client assigned to those features. Developers needed to be confident that they were building a capable system and were not introducing too much bloat.

Bringing two teams with disparate skills together required much communication, but offered both teams new perspectives on how to accomplish their tasks. Now they can share knowledge of their respective processes and compare experiences and notes, potentially creating a more powerful final product.

FIRST STEPS

Satisfying all the needs identified through developers, clients and the sales force, as well as general market research, required a good deal of planning to prioritize and implement.

A top priority was to reduce costs for existing clients in order to broaden product appeal. This was accomplished conservatively by building bridges for content to flow bidirectionally between the products. A site built in the content management system would easily make the round trip through the localization workflow system. But the localization workflow still had to import and export static sites for customers unwilling to make the move to the full content-management platform; and content-management clients had to be able to choose to localize when ready.

A second, parallel phase was the merger of the two products. Obviously, customers want to be able to keep their multilingual content fresh. The merged product can submit smaller chunks of content to localization when the content gets changed, as opposed to parsing entire pages, sections or sites for every change. Customers can spread out the cost of translation across the lifetime of the Web site, instead of at publishing milestones.

The freelance translators like having a steadier, more manageable flow of work without losing any features from the original localization workflow system and without forcing them to learn a brand-new tool. Having one code base also simplifies ongoing maintenance for developers, thus keeping costs lower for clients as well.

Obviously, adding new features is anything but helpful to clients if they lose the benefit of using a content system. A round-trip path was required to localize English content that fits into the existing content-managed sites as well as any new ones that are developed. Content management users, as defined by their roles, needed the ability to decide when and how to localize content, using the interface to which they were already accustomed.

The first step was to identify all the places where content flowed into and out of both systems. We found a good match in submitting flat-file sites from the content management system to the localization workflow. The localization workflow could already shepherd static Web sites in flat files through translation. On the other side, the content management system already let users import and export all or part of a site to XML in flat files. Algonquin added another parser to the localization system to read and write the XML files coming out of the content management system. After localization, the content management system reads the translated site file and places the localized content where the administrator chooses in an existing site or into a new site.


Clients still choose the pieces of content within a site to localize, but they submit it to translation much more easily now. Clients can do this as often as they choose. While this method works best with one source language, clients can extend it to more.

The goal was that authors would be able to author content and write updates in any language and to have the content system shepherd the changes to each locale. Coordinators would set up a scope of content to localize and rely on the system to submit content to translation continuously.

To begin, both product teams examined the internal concepts in each product. The teams identified how the content management system represented content and directed the localization process to operate on that representation. Then they upgraded existing concepts in the content management system to support locales, such as dates, times, calendars, measuring units and systems, currency, addresses, numbering systems, conversions and so forth. The teams preserved the proven localization process already in place. As a result, the localization system became a satellite application integrated into the content management system, with a workflow and translation repository. Features were added to the content system to import and export static Web sites so that existing localization-only customers could continue to leverage the workflow system with their existing solutions.

Authors can write content in any language and submit their final drafts to be localized to each supported language provided they fall in the designated scope of content to translate. For example, authors writing separate stories in French, English and Spanish would submit them to be translated into each of the other languages at the same time.

Translators can work on a flatter, more continuous load of content across more languages. For customers of the localization application in the multilingual content management system, localization bills will become part of the everyday cost of operating their Web sites. Users benefit from more current, complete Web sites, regardless of their preferred language.

During this process, non-English user interfaces were developed for the content management system. The result, at the end of the integration, was a multilingual content management system. 

Greg Norton, former president of E-Merge Strategies, is vice president of sales at Algonquin Studios. He can be reached at norton@algonquinstudios.com



Adrian Roselli is the vice president of interactive media for Algonquin Studios. He can be reached at roselli@algonquinstudios.com



David Thiemecke is vice president of technology at Algonquin Studios. He can be reached at thiemecke@algonquinstudios.com



**400 Brisbane Building
403 Main Street
Buffalo, NY 14203**



**Phone: (716) 842-1439
Fax: (716) 842-1762
www.algonquinstudios.com**