GILBANE

Content Technology Works

The Reality of Web 2.0: O'Reilly Media's SafariU Leads by Example

This Web-based custom publishing platform changes the dynamics of textbook publishing for higher education

Leonor Ciarlone
Senior Editor, The Gilbane Report
February 2006

O'Reilly Media spreads the knowledge of innovators through its books, online services, magazines, and conferences. Publisher of the iconic "animal books" for software developers, creator of the first commercial Web site (GNN), organizer of the summit meeting that gave the open source software movement its name, and prime instigator of the DIY revolution through its Make magazine, O'Reilly continues to create new ways to connect people with the information they need. Whether it's delivered in print, online, or in person, everything O'Reilly produces reflects the company's unshakeable belief in the power of information to spur innovation. A prime example is O'Reilly's SafariU, a Web-based custom publishing platform that enables higher education professors to create, publish, and share customized computer science and information technology textbooks and course materials.

O'REILLY®

Table of Contents

Introduction

Content Technology Works (CTW) is an industry initiative, administered by **The Gilbane Report**, to develop and share content technology best practices and success stories. The premise is that when given enough proven recipes for success, enterprise consumers will be able to adapt and replicate that success for themselves—increasing productivity and confidence.

Success stories are written by **The Gilbane Report,** with final editorial control resting entirely in the hands of the adopter. The result is that:

- Vendors do not control content.
- Success stories are as opinionated and as jargon free as the adopter prefers.
- Analysis is included from **The Gilbane Report** and invited contributors.
- The stories are not just about technology, but also focus on what matters to the adopter in terms of business requirements and other objectives.

CTW case studies provide organizations with best practices in content technologies and strategies for securing funding, measuring actual value, and driving adoption. For more information on the CTW program, see page 16 or visit http://gilbane.com/ctw_description.html.

Case Study Overview

O'Reilly Media spreads the knowledge of innovators through its books, online services, magazines, and conferences. Since 1978, O'Reilly has been a chronicler and catalyst of leading-edge development, honing in on technology trends and spurring their adoption by amplifying faint signals from the "alpha geeks" who are on the front lines of the future. An active participant in the technology community, the company has a long history of advocacy, meme-making, and evangelism.

This success story outlines the technology and business processes that enabled O'Reilly to expand into the higher education textbook publishing market by creating a Web-based custom publishing platform that produces more targeted, less expensive teaching materials than traditional textbooks.

Acknowledgments

We gratefully acknowledge the generous contribution that O'Reilly Media made to the development of this case study. The company allocated the time of talented and heavily committed management for the purpose of improving the understanding and adoption of enterprise content technology. We especially thank the individuals with whom we spoke when researching the case study. Their passion for getting things right for their customers is a significant part of the O'Reilly story.

Product, technology, and service names are trademarks or service names of their respective owners. For additional information on our editorial policy, see www.gilbane.com/editorial_policy.html.

Using This Case Study

This case study outlines the essential elements of applying a Web-based custom publishing platform to the needs of higher education professors and students. This is an individual story about one organization, O'Reilly Media. While O'Reilly's approach may not be universal, its success in solving critical problems is indisputable. It is not possible to generalize O'Reilly's approach into a universal formula, but there is much here that will be useful to other organizations with similar corporate goals.

In Their Own Words: The O'Reilly Perspective

Recognizing the Opportunity

What were the symptoms in the market that brought this need to your attention?

"Over the past few years, as we've created a database of technology information by producing our books using XML, we realized that the shift from paper to pixels opened up a host of new possibilities for delivering information. The textbook market is ripe for change. The current system of creating and delivering textbooks is expensive and offers little value for students and educators alike. We believe that the market is primed for innovation and we plan to lead the way." *CJ Rayhill, CIO and SafariU General Manager*

"Students and faculty want more choice and value in their instructional materials. SafariU continues blazing the path O'Reilly and Pearson started four years ago and adds even more flexibility to the university environment." *Gary June, Chief Marketing Officer of Pearson Education*

"The textbook industry is about to go through a major revolution. It's being challenged for the first time the way Napster challenged the recording industry and forced it to rethink the distribution of music." *Kent Sandoe, Professor of Management Information Systems, California State University, Chico*

How did you identify what specific content technologies were appropriate for SafariU?

"SafariU takes full advantage of three classic disruptive technologies that are now converging: digital information, Web services, and print-on-demand. XML is a core technology, because it enables a vast storehouse of content, which Mark Logic calls a "contentbase." With Web services, we built a front-end to that contentbase that allows users to reassemble or remix our information to suit their needs. Print-on-demand gives users the choice to publish their remixed information in print. As print-on-demand technology has matured, it's become more robust and affordable, and we believe it's now ready for prime time." *CJ Rayhill, CIO and SafariU General Manager*

Note: According to Mark Logic Corporation, a contentbase is a collection of content such as documents, emails, Web pages, etc., that is stored in a common repository in a consistent format to facilitate retrieval and processing.

The Vision

What did you want to be able to do by using content technologies for SafariU?

"The Web is opening up some amazing possibilities for us to reinvent content, reinvent collaboration. SafariU was built from the ground up to let people remix content. We set out to create a framework that lets you pick and choose the information you want, re-assemble it, mix it with your own material, and publish the result online, in print, or in both formats." *Tim O'Reilly, Founder and CEO*

"We're putting information online, not books. We consciously decided *not* to focus on ebooks, but rather to put the information in our print books into a XML database and develop a Web interface to that content. In the process of building SafariU, we developed tools and competencies that give us the capability to create other XML-driven information services. We see SafariU as just the beginning for new content delivery methods." *CJ Rayhill, CIO and SafariU General Manager*

Product Selection

Which vendor(s) did you select and what were the overriding considerations?

"After extensive testing, we chose MarkLogic Server as our XML content server and publishing platform. Scalability was a big factor for us and Mark Logic Corporation's product was the only one that held up under our tests at the time. MarkLogic Server uses standards such as XML and XQuery so we can customize it to fit our needs. And since our in-house developers already use an array of standards-driven and open source languages, they have the skills to do the customization we require." *CJ Rayhill, CIO and SafariU General Manager*

Measuring Success

What do your customers think of SafariU?

"We're still in the early stages of the product's use, so we don't have an extensive track record yet. What we *do* know is that a high percentage of customers who complete a project using SafariU are coming back for more. They get hooked!" *CJ Rayhill, CIO and SafariU General Manager*

"O'Reilly owns a tremendous quantity of well-written, valuable information. The fact that they're willing to disaggregate it, allowing a professor to go and re-aggregate the content in a way that is meaningful to students, to a particular audience, to a particular course – that's revolutionary." *Kent Sandoe, Professor of Management Information Systems, California State University, Chico*

"Being able to pull everything together and bundle it from one source is incredibly helpful. I didn't have to go through someone else's huge knowledgebase, or navigate the internet to find these things. All the SafariU content has one common look and feel. I can browse all the resources without having to acclimate to a bunch of different Web sites." *Jon Preston, Interim Chair of the Department of Information Technology, Clayton College and State University*

"The biggest attraction for me is that I can create the book my way. For whatever reason, I tend to view things differently from most textbook authors out there. I haven't found any books that present the material my students need – and in the way I think it should be presented – in a single textbook. For some topics, keyword searches enabled me to find things in places where I wouldn't have known to look on my own." *Charles Anderson, Assistant Professor, Division of Computer Science, Western Oregon University*

O'Reilly Media Company Background

 O'Reilly Media is inarguably one of the world's most renowned computer book and online media publishers. The company boasts a highly respected and recognizable brand characterized by the unique design of O'Reilly animal book covers (one of which features the company's tarsier mascot), a deep and consistent focus on leading-edge technologies, and the pioneering vision of founder and CEO Tim O'Reilly.

The largest independent publisher of technology books, O'Reilly holds a 15% share of the U.S. technical and computer publishing market. The company publishes its "animal books" in addition to series such as Essentials, Hacks, Cookbooks, and Pocket References. Subjects include open source technologies, operating systems, networking and system administration, programming languages, Web design, and more. Devotees of O'Reilly books are predominantly highly technical programmers and system administrators from the corporate sector. However, technology innovation and adoption over the past five years has brought more diverse audiences such as academics and mainstream consumers into the customer mix.

O'Reilly describes itself as a media company as opposed to a traditional publisher. The reason for this self-description is crystal clear: company assets include a 700+ library of O'Reilly-authored books (90% of which are electronically available through the Web-based SafariBooksOnline subscription service), a multi-faceted and highly-respected presence in the technical conference industry, and a portfolio of Web-based technical communities, blogs, and online articles that are indispensable to most software developers.

According to Chief Information Officer and SafariU General Manager CJ Rayhill, "O'Reilly Media takes the intellectual assets of innovators and brings them to the masses." In doing so, the company mixes traditional channels with emerging ones in ways that are unique, interactive, and usually pioneering. A snapshot of some of the company's accomplishments provides substantial evidence:

- 1992: Published the *Whole Internet User's Guide & Catalog*, a comprehensive introduction to the Internet for beginners and veterans.

- 1993: Launched GNN (Global Network Navigator), an Internet "information center" that provided a portal to the Whole Internet catalog, news aggregation, and an online marketplace.

- 1999: Created the O'Reilly Network, fostering the development of a variety of technical portals for developers focusing on technologies such as XML, Linux, Perl, Java, and Web Services.

- 2001: Launched SafariBooksOnline as a joint venture with The Pearson Technology Group, a division of Pearson plc (NYSE:PSO). SafariBooksOnline provides a collection of e-reference libraries to programming and IT professionals, business managers, and general knowledge workers.

- 2005: Launched the O'Reilly Radar, a blog featuring conversation from CEO Tim O'Reilly, CTO Rael Dornfest, and other O'Reilly executives.

O'Reilly's intense focus on the importance and impact of various technologies and computing trends, in particular the Internet, has transformed what was once a small technical writing consultancy into a major global brand in the technical publishing industry. According to Nielson Bookscan, "O'Reilly was the only major publisher to show growth quarter by quarter throughout 2004 and to show 2005 Q1 numbers above a year ago."[1] Nielson Bookscan aggregates point-of-sale data from about 70% of U.S. bookstores.

Given the need for O'Reilly to effectively react to market consolidation in both the publishing and retail industries while managing a disruptive, but opportunistic Internet influence, the company's success reflects a corporate savvy in balancing focused strategy with product diversity.

The O'Reilly Vision: Defining the Publishing and Internet Intersection

O'Reilly Media epitomizes the vision of "agile publishing," a term pundits have used since the mid-1990's to describe the process of repurposing structured content to enable and drive dynamic, automated publishing processes. In the 90's, these discussions were largely predictions. Today, in the words of a favorite William Gibson quote cited frequently by Tim O'Reilly, "The future is here, it's just not evenly distributed."

O'Reilly understands and exploits the fact that software developers, and increasingly mainstream consumers view the Web as *the* source for "the most current information – on demand." In fact, the company has embraced the Internet as the digital infrastructure for

---

[1] O'Reilly Radar, *Book Sales as a Technology Trend Indicator*. April 23, 2005.
http://radar.oreilly.com/archives/2005/04/book_sales_as_a.html

information to respond to two critical trends in the publishing industry that counter traditional *hardcopy-centric* publishing cycles and revenue models:

1. Retail vendor consolidation has significantly restructured hardcopy delivery channels. In fact, the four largest booksellers controlled approximately half of retail book sales in year 2000.[2] Consequently, hardcopy books have a severely limited "shelf-life" and in turn, reduced revenue per title. According to CJ Rayhill, "What used to be an 18 month shelf-life for a book is now 4-6 months."

2. Publisher consolidation has significantly restructured the competitive landscape, spurring publishers to implement strategies for coopetition from a horizontal as well as vertical perspective.

Survival in the publishing arena calls for innovation as well as agility. O'Reilly clearly possesses both. Research firm Outsell, Inc., recently labeled the company an "agile publishing vendor," noting its ability to provide "agile content," a characteristic of publishers who have implemented nimble processes and technology platforms that enable modular and flexible products.[3]

Since 2001, SafariBooksOnline has been a premier example of the intersection between publishing and the Internet, demonstrating that electronic content delivery infinitely extends the shelf-life and value of books while simultaneously providing additional revenue streams. Today, the Safari brand offers a series of e-reference libraries and other e-reference tools for enterprises, academic institutions, and individuals.

Advocates of Chris Anderson's Long Tail concept,[4] O'Reilly can point to the tangible benefits of its application. According to SafariBooksOnline Managing Director Sean Devine, "Many books are alive years after they have been pulled from the retail channel. We find that inside Safari, books get used longer and the most-read books are older ones not available anywhere else."

The success of SafariBooksOnline also demonstrates the power of coopetition in a consolidated publishing industry. A joint venture with the Pearson Technology Group, the application includes titles from multiple publishers, validating Tim O'Reilly's premises that "sharing increases value" and "openness will become the norm."

O'Reilly's definition of "agile content" is intrinsically tied to its vision for the publishing and Internet intersection. Defined by O'Reilly's Web 2.0 principles,[5] the company foresees a variety of Web-based information products that:

- Treat content as a corporate asset, using the "data as the next Intel Inside" perspective

- Are designed for remixing and publishing on-demand

- Engage "collective intelligence" by encouraging user contributions

[2] United States Small Business Administration's Office of Advocacy and Jack Faucett Associates. *Small Business Survival in Competition with Large Multi-Unit Retail Firms.* January 2000.

[3] Strohlein, Marc. Outsell, Inc. *HotTopics: Agile Publishing and the New Breed of Publisher.* March 2005.

[4] Anderson, Chris. Wired Magazine. *The Long Tail.* http://www.wired.com/wired/archive/12.10/tail.html

[5] O'Reilly, Tim. *What is Web 2.0?* September 30, 2005.
http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

The Problem: Change the Dynamics of Textbook Publishing

It is difficult to find an industry in need of a better and, in this case, *different* publishing solution than higher education. The issues are multi-faceted and exacerbated by cost. According a United States Government Accountability Office (GAO) report, "In the last two decades, college textbook prices have increased at twice the rate of inflation. Increasing at an average of 6 percent per year, textbook prices nearly tripled from December 1986 to December 2004."[6]

According to the same report, "While there are hundreds of college textbook publishers, there has been substantial industry consolidation in recent years, with sales at five of the largest publishers representing over 80 percent of the market in 2004."

Could there be a more appropriate time to change the dynamics of textbook publishing? Is there an opportunity to intertwine Web 2.0 principles with the Long Tail concept and increased publisher coopetition? O'Reilly certainly thinks so, and embarked on a detailed research initiative in 2004 to understand the details of the problem. The company conducted hundreds of interviews with professors and students and came away with an understanding of key industry issues:

- Teachers are looking for one source for trusted, copyright-cleared course materials available both in print and online.

- Students and teachers are publicly campaigning against high textbook costs, superfluous ancillary materials, and unnecessary new editions.

- Teachers are complaining that textbooks don't fit their classes. Most report that students end up paying for a whole book when only a few chapters are relevant to their class.

- Teachers report that traditional textbooks are an anachronism; teaching in most U.S. universities is chapter-based, rather than book-based. In addition, students are used to finding information and doing schoolwork online.

- Teachers want to harvest the knowledge of their peers and share teaching materials with peers in other colleges and universities.

O'Reilly's desire to inspire innovation in textbook publishing for higher education led the company to step up to the challenge in its typically bold and pioneering fashion. The development of SafariU, O'Reilly's XML-driven, Web-based custom publishing platform, is bringing textbook publishing into the 21st century.

The Need: Design a Custom Textbook Information Service

Computer science and information technology educators have incorporated O'Reilly materials into university courses for many years. Students have used O'Reilly books, tutorials, and articles as input to research papers and thesis' even longer. Today, numerous professor-authored blogs and Web syllabus pages directly reference O'Reilly books and SafariBooksOnline as recommended resources.

Although not designed as textbooks, it is clear that O'Reilly *tech*books hold a coveted spot in the "top ten information sources" for those teaching and studying emerging technology topics. In response to requests from academics, the company expanded its focus on content for computer science education, driven by the results of internal research on higher education pain points. The next step was to make sure that content delivery would not be hindered by the traditional textbook publishing model.

---

[6] United States Government Accountability Office. *GAO Report to Congressional Requesters on College Textbooks.* July 2005.

A team of O'Reilly executives, business analysts, and technical professionals set out to design the service. Led by CJ Rayhill, Chief Information Officer and SafariU General Manager, individuals such as Program Director Allen Noren and Marketing Director Nathalie Mainland worked closely with a professor-driven focus group during the brainstorming process. The design objectives quickly became clear:

- Deliver a Web-based information service like no other, enabling professors to create, publish, and share custom textbooks.

- Allow access to *all* the information sources that are so difficult for professors to locate and aggregate with current research, publishing, and consumption techniques.

- Provide an integrated "one-stop shopping" platform for computer science and information technology content with unparalleled search, query, retrieval, remix, and publishing capabilities within a single interface.

The potential beneficiaries are vast. There are more than one million academic faculty members employed at U.S. colleges and universities.[7] In the student arena, college enrollment is projected to rise to 17.5 million by the year 2010, an increase of 20 percent from 1998.[8] The O'Reilly team was up for the challenge. While the initial focus of SafariU is Computer Science and Information Technology, the SafariU platform can be extended for use in other academic disciplines.

Criteria for Success: Time and Cost Savings for Professors and Students

"Teach Exactly What You Want – With Just One Book"
"Spend Less. Learn More"

These mantras define the O'Reilly team's cost and time-driven success factors for SafariU. Each is based on the "wish lists" of hundreds of university and college professors:

1. Consolidate source materials to deliver a repository of copyright-cleared content with a breadth and depth that eliminates the need to search for information in multiple places.

2. Provide multiple content delivery methods, including bound printed books and electronic versions.

3. Provide the ability to create and share supplemental materials from multiple sources and formats.

The Need: Enable Web-Based Remixing and Multi-Channel Publishing

A Web-based custom textbook service was only the beginning of the SafariU vision. In fact, the definition of *service* for the platform is much deeper than the term implies. Referring again to the SafariU mantras, in particular "Teach Exactly What You Want," the design team was determined to seamlessly incorporate O'Reilly's Web 2.0 principles into SafariU core architecture.

This would take the project's time and cost savings success criteria to an entirely new level, delivering the value of a content engine that would deliver remixing capabilities to the professor desktop. Successful execution of the design vision would address a broader spectrum of higher education pain points, including:

- Teaching is more chapter-based than book-based

---

[7] American Council on Education. *An Overview of Higher Education in the United States: Diversity, Access and the Role of the Marketplace.* 2004. http://www.acenet.edu/bookstore/pubInfo.cfm?pubID=327

[8] National Center for Education Statistics. *"Projections of Education Statistics to 2010,"* 2000. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000071

- Textbooks just don't "fit" classes

- Out of print books are unavailable

Removing these kinds of barriers would provide higher education with what Lawrence Lessig describes as a "remix culture,"[9] or a society which allows and encourages derivative works. A key theme of the O'Reilly 2005 Emerging Technology Conference, remixing is a familiar concept to disc jockeys, musicians, and teenagers. It is not commonly applied to textbooks, or for that matter, *most* books. The O'Reilly team set out to change this.

"To remix" however, is simply one part of the puzzle; "to publish" is another. SafariU's publishing capabilities needed to be anything *but* traditional, as defined by O'Reilly internal research and professor-driven focus groups. Publishing in print and/or online with support for a variety of input and output formats would be the norm. Subject matter and chapter-specific database querying and search would be expected. Only the power of an XML content server would meet these goals. Having chosen MarkLogic Server as the SafariU platform, the O'Reilly team was in an excellent position to deliver these requirements.

Criteria for Success: Professor-Driven Custom Textbooks

"One of the key ideas from the Creative Commons that I really embrace is the idea that all creativity is rooted in reuse. The smartest thing that any publisher can do is to make sure that we allow our customers to surprise us with ways that they have remixed our ideas and our material with their own." Tim O'Reilly, Founder and CEO

An innovative company from the beginning, O'Reilly is well-known for its ability to research, implement, and advocate standards-based technologies. The company's deep investment in XML as the authoritative "gold standard" source for intelligent, electronic content would enable the full spectrum of what the SafariU team wanted to deliver.

Together with full-text search capabilities, XQuery would provide the access, manipulation, and dynamic transformation layer that would unlock the power of the repository. As CJ Rayhill noted, "The better the key, the easier it is to open the door." Based on MarkLogic Server's transparent delivery of XML and XQuery benefits, the "key" was in place, making the SafariU design a potent reality.

The Need: Enable Collaborative Learning Communities

The final piece of the SafariU vision is perhaps the holy grail of designing and delivering Web 2.0 applications: engaging users and enabling collective intelligence. Sharing and peer review, two key tenets of Web 2.0, facilitate conversation and information exchange. Implemented as core functionality, the result is a library of collective intelligence that encourages contribution and participation in a user community.

In a university setting there are two types of content that are invaluable and potentially reusable course components:

- The syllabus, which establishes a critical connection between a professor, the course, the required and optional course materials, and students.

- Supplementary course materials, which include Web articles, lecture notes and reviews, slides, code samples, and exams.

Unfortunately, these content types are often isolated, available only to the professor who develops and assembles them. Without an environment to create, store, and share this content, sharing and peer review is impossible.

---

[9] Wikipedia. *Definition of Remix culture*. http://en.wikipedia.org/wiki/Remix_culture

The O'Reilly team's desire to "go beyond the book" would change this situation. Applying the music industry's "playlist" concept to textbooks, the team designed a Web-based peer network called the "Learning Object Exchange." This environment would allow professors to share syllabi and supplementary materials in the context of a particular course; in effect, creating a playlist of accessible and reusable learning objects.

The Learning Object Exchange would create the environment for a collaborative learning community and put professors in charge of its breadth and depth. Some universities have envisioned the concept of a "dynamic syllabi," a means to stage, manage and enhance a course through an online platform."[10] Others have simply longed for an environment to access and share supplementary materials relevant to multiple courses. SafariU would deliver on both requirements.

Criteria for Success: A Library of Reusable Learning Objects

Viral adoption of SafariU's Learning Object Exchange will result in a library of reusable learning objects that are created, shared, and used by professors to increase the quality and quantity of supplementary materials for a particular subject matter or course. Dynamic electronic syllabi that link to all course materials including book chapters and supplementary materials will provide students with a flexible electronic bookshelf from which they can access custom textbooks.

These capabilities will be a significant driver for the time and cost saving benefits SafariU is poised to deliver. As one professor noted, "There are a thousand other people in the world teaching the same class I am. Why are we all re-inventing the wheel?"

Preventing the reinvention of the wheel is the definitive SafariU criteria for success – ultimately fusing the benefits of consolidated source materials, remixing, collaboration, and dynamic, multi-channel publishing into a single platform.

Solution Components

Thinking Outside the Book: Introducing SafariU

After months of coding, designing, prototyping, and "testing, testing, testing," the O'Reilly team launched a limited beta of SafariU in 2004 and a full beta in 2005. Today, SafariU accesses over 2,200 books from O'Reilly and numerous publishing partners as well as over 5,000 O'Reilly articles. O'Reilly estimates that the XML content repository is 1.6 million pages, which could double over the next 5-10 years according to CJ Rayhill.

One of the most critical design and prototyping criteria for SafariU was and remains usability. Industry pressures as well as some amount of technology trepidation make delivering ease of use paramount for viral adoption. As one professor noted during a research conference for higher education, "If colleagues have to run in order to keep up with technology-based educational tools, the resulting breathlessness may lead to the following conclusion: stay with the ancient remedies so as to avoid the treadmills of change."[11]

One of the most important SafariU brand messages is "Custom Publishing Made Easy." The following capabilities prove this reality:

---

[10] Georgetown University CrossRoads Project. *"Technology and Learning Web page,"* 1998. http://www.georgetown.edu/crossroads/webcourses.html

[11] Baumeister, Hans-Peter. *Networked Learning and Knowledge Management - A Systemic Challenge for Universities.* April 2004. http://www.shef.ac.uk/nlc2004/Proceedings/Individual_Papers/Baumeister.htm

Create Your Custom Textbook

SafariU lets professors utilize powerful full text search and query technologies to pick and choose exactly the materials they want for a custom textbook. Once assembled, these materials can be remixed and aggregated with personal and public supplementary materials.

Publishing a custom textbook to multi-channel outputs is straightforward. Professors can have the book printed, bound with their name on it, and delivered to their doorstop or preferred bookstore. Dynamic conversion capabilities from file formats such as DOC, RTF, PPT, PS, JPG, TIF, and HTML produce a print-ready, book-formatted PDF file for review.

Create Your Online Syllabus

SafariU lets professors transform a custom textbook into an online syllabus and vice versa. Once the online syllabus is available, students can electronically order a course-length subscription, gaining full access to up to 10 complete books even if the professor has only referenced or printed one or more chapters in the custom textbook. Access to an unlimited amount of syllabus-driven supplementary materials is also included.

Professors can copy, re-order, and update syllabi for the life of the course, adding links for additional content books sources (up to 10) as well as new supplementary materials. Even when the course is competed, professors have full access to their syllabi library, enabling future use for new or similar classes.

Exchange Materials: The Learning Object Exchange

"Share your teaching materials with peers." This is a powerful message for an industry coming to grips with new technologies, market pressures, and an overriding mandate to collaborate. When professors make their custom textbook projects public, they enable peers to see the Tables of Contents for their projects, encouraging conversation, collaboration, and content reuse.

Unlocking the Content: The Role of MarkLogic Server

O'Reilly chose MarkLogic Server as the core architecture for the SafariU custom publishing platform. Certainly, an important driver of the initial evaluation was project-based. However, the purchasing decision was very much strategic: MarkLogic Server will be the fundamental architecture upon which O'Reilly builds future Web 2.0-centric content applications. As CJ Rayhill noted, "It's nice when you can justify a product for a project, but it's even better when you get an entire infrastructure."

Some of the more significant reasons why MarkLogic Server was *the* choice for O'Reilly included the product's ability to:

- Support large numbers of users and significant content volume while still providing high performance regardless of increases.

- Deliver a standards-driven architecture that enables modular application development, content exchange, and powerful content-based querying.

- Provide an XML-driven "gold-source" standard repository that can digest, convert, and normalize structured and unstructured content while also preserving the original formats of professor-driven supplementary materials.

- Provide leading-edge, automated, and format-independent capabilities for indexing, search, and query functionality.

- Support a deep metadata architecture for components such as syllabi and learning objects.

O'Reilly built SafariU on a modular XML, XQuery, and Java/JSP architecture. The approach gives them control, flexibility, and "sandbox" environments to experiment with the design of future information services. With it's ability to build applications and support business processes based on a reusable contentbase, MarkLogic Server is O'Reilly's platform of choice for content-driven applications.

Best Practices and Organizational Changes

The O'Reilly team offers the following best practices advice to others facing content challenges:

- Adopt a Web2.0 perspective on application development.

- Conduct a thorough vendor analysis according to your specific needs, providing sample data if possible.

- Acknowledge and plan for developer learning curves and do not discount the importance of professional services. Budget wisely for resource development and outside assistance.

- Keep your customers in the loop during design, prototyping, testing, and beta phases using surveys, analytics, and "high-touch" technical support. Follow the "Release early, release often" philosophy pioneered by the open source movement.

Results

Cost Savings for Educators and Students

SafariU is available to educators at no cost, providing the ability to create, remix, and deliver custom textbooks to students at a lower cost than a traditional textbook. Educators must meet a minimum annual amount of services driven by student subscriptions, currently set at $700 per year. This criterion is easily met with a single class of 20 students.

The savings are profound and address student anxiety about the high cost of textbooks. A typical SafariU printed custom textbook is 250 to 300 pages, costs students between $46 and $52, and can be professionally printed, bound, and delivered in as little as two weeks. During academic year 2003-2004, first-time, full-time students attending 4-year private, nonprofit colleges were estimated to spend $850 for books and supplies in their first year.[12] SafariU will have a tangible impact on reducing that cost.

For students who prefer electronic access to custom textbooks, a subscription to a SafariU syllabus costs $10 per month. The subscription provides access to the custom textbook, the complete content of up to 10 of the books referenced by the course syllabus, *plus* an unlimited amount of supplementary information that the instructor chooses to include. The expiration-driven subscription model works well for cost-conscious professors and students, while offering

---

[12] United States Government Accountability Office. *GAO Report to Congressional Requesters on College Textbooks.* July 2005.

a significant amount of value for the investment. An already devoted professor audience is ecstatic with the results:

- "SafariU brings two exponential improvements over the past. One is time savings by avoiding the wild goose chase on the Web. The second is the quality of the material. I've produced a course packet for the students in a very short period of time. And the quality of that packet is superb." *Kent Sandoe, Professor of Management Information Systems at California State University Chico*

- "As one of the first SafariU users, I was able to create a custom textbook in a single weekend. I started *Introduction to Perl Programming* on Friday and had it ready for class Monday morning." *Scott Gever, Adjunct Instructor, Perl Programming, Foothill College at Los Altos Hills*

A New Collaborative Paradigm for Higher Education

Widespread praise suggests that SafariU is onto something big – something that could change the dynamics of textbook publishing. Accolades for SafariU transform the rather dismal research published over the past five years from sources such as the United States Government Accountability Office (GAO), the American Council on Education, and The Futures Project into a more optimistic opportunity for change.

O'Reilly's Web 2.0-powered vision for SafariU predicts viral adoption based on the platform's ability to engage the collective intelligence of users, enable content remixing, and facilitate conversation and information exchange. It's a powerful, insightful, and forward-thinking recipe for change – and already proven in various Web initiatives such as Wikipedia, iTunes, and Flickr.

O'Reilly's execution of their vision is based on a systemic approach to increasing SafariU awareness, conducting educational seminars in state campus systems, and enabling more and more collaborative learning communities. Success, according to CEO and Founder Tim O'Reilly, "will turn textbook publishing on its head by transferring the publisher's authority to the instructor, allowing content to be mixed and matched from multiple sources, facilitating sharing in the academic community, and changing the pricing model."

Conclusions: The Gilbane Report Perspective

When the Gilbane team evaluates a potential case study for our Content Technology Works initiative, we specifically look for elements of the deployment that are useful lessons for other adopters of content technologies. We think that the following universal factors are key to the success that O'Reilly Media has achieved thus far.

- **"Data is the next Intel Inside"** – we could not agree more with this Web 2.0 principle. Having spent over 20 years researching and writing about content technologies, The Gilbane Report has consistently focused on how content technology can be used for enterprise business applications and how content and computing will evolve. Today, the power of "content as data" is clearly the success factor for a myriad of business applications, commercial products, and community and government services.

- **Content applications rely on XML** – as XML heads towards its eighth birthday (February 2006), its application to structured content is evident throughout most industries. Although regularly applied to data exchange during its first five years, it is the more recent years that demonstrate the value of content intelligence, flexibility, and reuse as enabled by XML and related standards. That value is reaping significant ROI for those making the investment.

- **Collaboration and content technologies are intrinsically intertwined** – creating and exchanging content is critical to enterprise applications focused on streamlined business processes, productivity, and knowledge management. In commercial, next-generation content applications such as SafariU, the importance of the connection between content and collaboration cannot be overstated.

The development of new publishing models within O'Reilly Media is forward-thinking and admirable. They are engaging their customers in new ways while simultaneously delivering strategic improvements in content delivery and usage. It is no wonder that O'Reilly is considered as one of the most innovative and successful technical publishers in the industry. **The Gilbane Report** thanks O'Reilly and its employees for sharing their story.

A Supplier's Voice: Mark Logic

The Gilbane Report appreciates Mark Logic Corporation's contribution of the content for this section.

O'Reilly faced many daunting challenges to bring SafariU from vision to reality. Some of these challenges are unique and some are shared with other organizations that seek innovative ways to access and effectively use their vast volumes of content. For example:

- A major U.S. airline that is automating the intricate process of assembling flight operations manuals and keeping them up to date while giving pilots an online, searchable "electronic flight bag" that replaces cumbersome, frequently out-of-date, and hard-to-use paper manuals.

- A branch of the U.S. Army that is creating an online knowledge base that enables front-line troops to access and share vital information directly with each other in the field.

- A large medical information publisher that is speeding the creation and delivery of new online physician reference portals while simplifying the process of managing disparate information products.

What these organizations have in common with O'Reilly is content, of course, and lots of it. And, like O'Reilly, they all share a vision to leverage the XML standard to unlock the full value of their content, making it both agile and reusable. These are typical of the coming generation of content applications that require a robust, scalable, and high-performance platform built specifically for content.

What sets MarkLogic Server apart from other solutions is that it can load content 'as-is' regardless of format or schema and without time consuming schema analysis and 'shredding'; it has a complete implementation of the powerful XQuery language; and it is a scalable, high-performing, and robust 'enterprise server'. All of these capabilities were vital to the SafariU development team. With MarkLogic Server, each component of the application is developed and integrated using open standards, so applications can be assembled rapidly. This was a major consideration for the O'Reilly development team.

SafariU is an excellent example of the reality of content applications that embody the principle of 'collective intelligence' inherent in the Web 2.0 paradigm:

- SafariU uses an XML content server as an infrastructure component, enabling a Web-based custom publishing platform based on an intelligent contentbase; with the MarkLogic Server as the foundation, SafariU will deliver unparalleled performance regardless of volume.

- SafariU goes way beyond basic search, utilizing XQuery as a means to write applications in a high-level, declarative language and execute database-style queries against the contentbase.

MarkLogic Server has been selected by leading enterprise organizations, information publishers, and government agencies as the platform for new content applications that help them create innovative products and services, enhance revenue, reduce costs, extract greater value from existing content, and streamline business processes. MarkLogic Server enables these organizations to meet today's content challenges head-on: to integrate content from different sources; to repurpose content into multiple products; to build custom documents and views; to deliver content through multiple channels; and, to search-and-discover previously unknown information.

About Mark Logic Corporation

Mark Logic Corporation is the provider of the industry's leading XML content server. Mark Logic works with providers of information products to accelerate new product creation, deliver products through multiple channels, integrate content from different sources, repurpose content into multiple products, build custom publishing systems, and mine content to find previously undiscovered information. MarkLogic Server does this by enabling companies to query, manipulate, and render XML content using the W3C-standard XQuery language. Designed for high performance and scalability, MarkLogic Server can deliver millisecond response times against multi-terabyte contentbases. Mark Logic is privately held and backed by Sequoia Capital and Lehman Brothers. For more information, please visit www.marklogic.com.

About Content Technology Works

When we first conceived of an initiative that would develop and distribute success stories that placed recipe over ingredients and favored no supplier, technology or computing standard, we also recognized that our most significant hurdle would be to recruit vendors to subsidize such an independent and open process.

Since the CTW program was first conceived in late 2003, we have sought out suppliers who were passionate about and committed to content technology as a game-changing force in the markets that they serve. Our CTW partners know that public, open and unfettered access to successful enterprise deployments, regardless of the technology mix, only benefit the commercial aspirations of organizations that offer material, dependable and predictable value.

The Gilbane Report team wishes to thank these diverse and often competing organizations for their generous support and sponsorship of the development, promotion and distribution of CTW material. They are: Software AG (TECdax:SOW), Sun Microsystems (NASDAQ:SUNW), Artesia Digital Media, a Division of Open Text, Astoria Software, ClearStory Systems (OTCBB:INSS), Context Media, Convera (NASDAQ:CNVR), IBM (NYSE:IBM), Idiom, Mark Logic, Open Text Corporation (NASDAQ:OTEX), SDL International (London Stock Exchange:SDL), Vasont Systems, Vignette (NASDAQ:VGN), WebSideStory (NASDAQ:WSSI).