Electronic Reference Library, Phase 2
Analysis of Web-Based Content Management Systems for Managing
Iowa DOT Engineering Specifications: Assessing Best Practices

Submitted to: Iowa Department of Transportation
                                  800 Lincoln Way
                                  Ames, IA 50010

Prepared by: Center for Transportation Research and Education
                                  Iowa State University
                                  2901 S. Loop Drive, Ste 3100
                                  Ames, Iowa 50010

Principal Investigator: Geoffrey Sauer
                                  English Department
                                  Iowa State University
                                  294-3085 email: gsauer @iastate.edu

Co-Principal Investigator: Lee Honeycutt
                                  English Department
                                  Iowa State University
                                  294-4615 email: honeyl @iastate.edu

Co-Principal Investigator: Rebecca Burnett
                                  English Department
                                  Iowa State University
                                  294-5654 email: rburnett@iastate.edu

Co-Principal Investigator: Russ Walters
                                  Civil, Construction, Environmental Engineering
                                  Iowa State University
                                  294-2171 email: rwalters@iastate.edu

Project Starting Date: May 1, 2005

Project Completion Date: June 30, 2005

Project Cost: $44,723

TABLE OF CONTENTS

Analysis of Web-Based Content Management Systems for Managing DOT Engineering Specifications: Assessing Best Practices

Background

This proposal addresses construction of an online content management system that will support distributed collaborative authorship and editing of the Iowa DOT's Specifications Section documents. The project builds on previous work between Iowa DOT and Iowa State's Civil, Construction and Environmental Engineering Department on the Electronic Reference Library (ERL). Just as the ERL vastly improved distribution of these documents, the proposed content management system will streamline their production and editing by consolidating the text into an XML-based database system that supports refined version control, detailed editorial workflows, and multiple output formats.

Problem Statement

The ERL has made sustained improvements in how Iowa DOT distributes its specification documents, including the Standard Specifications, Supplemental Specifications, Materials Instructional Memorandums (IMs), and Standard Road Plans. The CD-ROM version of the library provides clients with a hyperlinked version of the specifications that facilitates quick navigation to referenced specification sections, word search capability, and quick printing of individual sections. However, the ERL has done less to address limitations in the current document production workflow resulting from current word processing systems. Additionally, editorial workflows for reviewing and approving changes to the specifications are now handled manually, through email or telephone notification. A content management system—be it an

open-source solution such as Plone or a commercial application like Microsoft's SharePoint—would automate this process by giving authors access to only those portions of the specifications for which they are responsible and by setting editing and publishing notifications according to a defined hierarchy of writers and editors. And because all text is stored in an XML-based database, distributing the specifications to a variety of document formats, such as XHTML, PDF, MS Word, can be achieved on the fly through XML stylesheet transformations.

Under the proposed contract, Iowa State University will analyze the current editorial workflow of Iowa DOT Specifications Section and recommend an enterprise content management systems that will best meet the present and future needs of the Electronic Reference Library.

In recent years, many organizations have felt the need to reassess the design of their information systems to incorporate single-sourcing using online content management systems, and many coherent strategies have been developed to guide and assist integration of varied data and information for diverse audiences. The most successful of these strategies are user based. As information resources grow in scope and acquire new database elements, real-time updates, video displays and other disparate components, the potential and assumed audience continually expands to the point where it is difficult at any given moment to define neatly a 'typical' user or their information retrieval task. Often, this pattern of growth of an information system results in audience segmentation and differentiation. Users identify themselves differently when they visit a web portal in search of different task information, and it is important for designers to identify the most common 'roles' audiences play when entering into a web portal in search of information.

Research Approach

This contract will begin a multi-year project to develop a next-generation content management system for the Electronic Reference Library (ERL). It will begin during this contract by collecting data from three primary sources: workflow analysis, log analysis and surveys/questionnaires. It will then set up a prototype server (compatible with the Blade production servers used at Iowa DOT) to permit the research team to install and configure prospective content management systems for the future ERL. The prototype server built under this contract will operate beyond this contract for the duration of the entire project, as the actual CMS is created and current content is migrated to it. When the final system has been implemented and the current standards documents migrated, it will be transferred from the prototype server to the ERL's virtual server at DOT, replacing the current ERL web-and-CD-ROM production system.

Workflow Analysis

Interviews of authors and editors (including Pat Magoon, Beth Richards, Laurel Raasch, Michele Regenold, Dan Harness and Tom Reis, among others) will map current workflows, with an eye toward mapping these into a content management system. This will permit the recommendation to incorporate participatory design principles to ensure maximum usability for the resulting CMS.

Log Analysis

In addition to workflow analysis, new log analysis tools developed by e-commerce interests over the past few years make possible detailed quantitative analyses of readers' paths through a website. Some of these tools can provide amalgam information of great use to information architects: the number of pages viewed by each

reader, the amount of time spent online, how often readers return to a site, the most-followed paths taken by readers navigating through the site, the most common referrers, and the search terms people enter into search engines. All of this helps information designers assess how to organize information to serve end users. For example, Figures 1.1 and 1.2 below show data gathered using such tools, allowing for analyses that track the 'stickiness' of a site—how long a site remains compelling to visitors. Site stickiness is one key to discerning whether or not site visitors remain long enough to gather the information that designers perceive as most important.
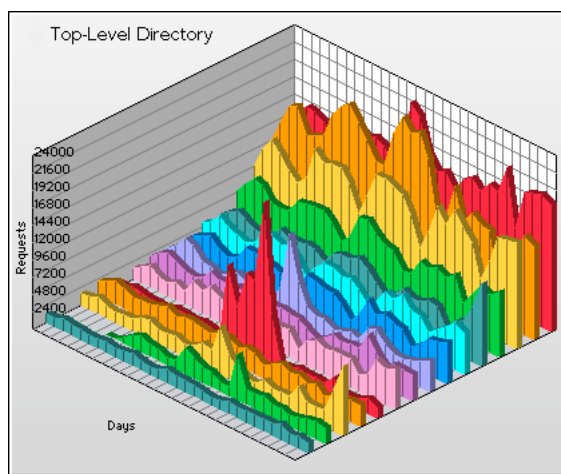
Figure 1.1: graphical representation of one month's readership of particular subcollections on a popular web portal
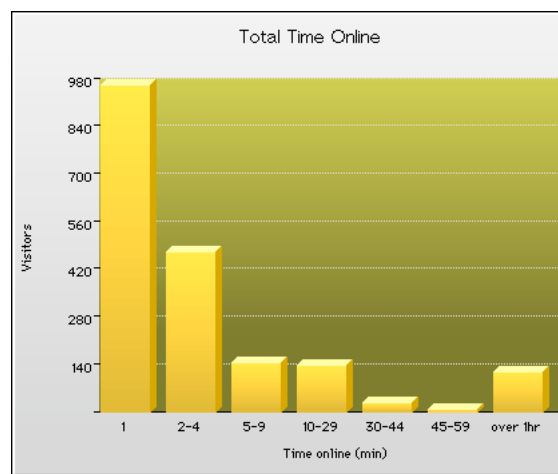
Figure 1.2: a table of time spent (in minutes) by one day's visitors browsing pages on a popular web portal

By studying mean entry points, exit points, the average and standard deviations for the amount of time users spend on particular pages in the site, by reviewing the mean paths (the most-frequently followed courses taken by readers through the ERL website) we can determine a great deal about usage. By looking at logs of search engine queries (as in Figure 1.3 below) and loyalty indexes (as in Figure 1.4) we can determine a great deal about users' expectations as they navigate the site, and we can extract information about how many users actually find their goal information.

IP Address	Search Terms	Date/Time
128.95.116.105	internship	9/20/00 13:45
128.95.116.105	kate long	9/20/00 13:50
128.95.116.105	carolyn plumb	9/20/00 13:51
128.95.202.127	math	9/20/00 14:08
128.95.196.196	Faculty listing	9/20/00 15:56
132.203.14.159	Japanese e-mail	9/21/00 7:42
132.178.181.180	target audience	9/21/00 7:57
132.178.181.180	target audience	9/21/00 7:58
132.178.181.180	Purpose	9/21/00 7:58
63.224.249.125	f716	9/21/00 12:38
63.224.249.125	A412 F716	9/21/00 12:50
63.224.249.125	Adobe FrameMaker	9/21/00 12:51
207.251.193.32		9/21/00 17:24
208.130.242.25	Language Courses	9/21/00 20:08
208.130.242.25	Foreign Languages	9/21/00 20:09
63.194.84.57	gif animator download	9/21/00 20:30
24.4.254.155	technical communications artical	9/22/00 7:17
24.4.254.155	technical communication	9/22/00 7:17
64.12.105.174	paparrazi	9/22/00 7:57
128.208.93.156	EWC	9/22/00 17:42
128.95.202.89	smtp	9/23/00 16:44
205.188.192.36	technical communication children	9/24/00 15:33
205.188.192.36	children	9/24/00 15:33
205.188.192.36	child	9/24/00 15:34
205.188.192.36	child@	9/24/00 15:34

Figure 1.3: a chart of the search terms typed by visitors to a university website, analyzed quarterly in order to plan future upgrades to the site.

Figure 1.4: a chart of visitor loyalty, displaying how many times visitors come back to a website (to complement statistics about total web 'hits').

Surveys/Questionnaires

Surveys and questionnaires, which can have both qualitative and quantitative components, will also be used as components of the research. For example, audience segmentation occurs two ways: an information provider can prescript division based on previous assumptions, or an information source can provide the opportunity for an audience to identify and segment itself (a large-scale university website provides an example of this latter approach). The use of open-ended surveys and question-naires will allow the investigators to both prescript some data collection as well as set the stage for the audience to self-segment, providing an important addition to the quantitative data.

This study will make use of an existing Specifications list of 2,200 e-mail ad-dresses of contractors, consultants, DOT employees and city and county employees who currently make use of the ERL on a regular basis. The study will create an online survey, which will be sent to these 2,200 users, in order to collect information to up-date and complement the 2002/03 survey run about a previous version of the ERL.

It will then analyze the results from this survey for their implications for a next-generation ERL.

Objectives

Through workflow analysis and complementary data gathering efforts (e.g. usage logs, questionnaires and surveys), we will gather detailed information on the efficacy of information structures currently delivered via the ERL. We will determine tasks performed by authors and readers of the site as well as the expectations brought to these tasks. We will analyze authors' experience in performing these tasks and the effectiveness of the site's information structure in supporting the efficient completion of these tasks. We will use this analysis to provide recommendations to help guide future design and organization of a CMS-driven future incarnation of the ERL.

The current contract has five objectives:

Analyze the editorial workflow of all contributors to the current specification documents.

(1) Analyze data from past surveys of ERL users, as well as conduct a more in-depth online survey of current users.

(2) Review several of the leading enterprise-level content management systems to explore their relevance to the Iowa DOT workflow and to analyze their fit with Iowa DOT's technological infrastructure.

(3) Perform proof-of-concept testing of these systems regarding version control and archiving, as well as XML-based transformations to other document formats.

(4) Submit a preliminary set of recommendations on how Iowa DOT should plan to proceed, including a brief review of the literature on single-sourcing and content management systems.

Benefits

The immediate benefits will be information about how to plan the future CMS re-design of the site, based upon actual users' experiences of the existing system.

The data on how people accomplish specific tasks will give us information on where things break down and from this information we will be able to systematize and produce advice on site organization. This will in turn allow the state to realize new tangible benefits from contemporary developments in content management. Additionally, because the project will produce a report of the current situation, Iowa DOT will be able to use its own resources to continue to meet the challenges posed by shifting technologies.

This proposal is a first stage of a multi-year project dedicated to the transition of the current ERL into a modern CMS.

Timeline

The usability research will adhere to the following schedule:

Project Plan Items: By 6/30/05	Scheduled
Analyze 2002/03 Survey Data	5/5/05
Meet with DOT IT Staff to Model Server	5/15/05
Create 2005 Online Survey	5/25/05
Write Brief Review of Relevant Lit About CMSes	5/30/05
Evaluate Current Workflow	6/15/05
Review Possible CMS Packages	6/1/05
Deploy 2005 Survey	6/5/05
Test Prototype 2005 Survey	6/10/05
Configure "Build" Server	6/15/05
Collect ERL Web Log Data	6/15/05
Install Prospective CMS: Drupal	6/15/05
Install Prospective CMS: Plone	6/20/05
Install Prospective CMS: SharePoint	6/25/05
Analyze 2005 Survey Data	6/30/05
Write Analysis of ERL Web Log Data	6/30/05
Write Analysis of 2002/03 Survey Data	6/30/05
Write Analysis of 2005 Online Survey Data	6/30/05

Proposed Technology Transfer Activities

The research team will assist the Specifications Section in presenting progress reports on ERL CMS development as requested within the effort available in the contract. The following forums are possible targets for such presentations.

- Iowa DOT Conferences
- CTRE Conferences
- Iowa County Engineers Association
- Iowa Public Works Association
- Construction Trade Associations

During the scope of this project, early training about CMSes will be provided so that the DOT staff can begin to consider how to integrate various documents into the next-generation ERL, including such documents as the Standard Specifications, General Supplemental Specifications, Supplemental Specifications, Developmental Specifications, Material I.M.s, Construction Manual, and Statewide Urban Standard Specifications.

Relationship to Other Research Projects

There are two types of research projects that are of particular relevance to this effort:

(1) previous studies on the design and delivery of online documentation, and

(2) general projects on content management and the usability of electronic information.

Members of the research team for this project are or have been involved in both types of research in the past.

Budget

The project budget asks for $44,722. The budget includes the following items:

Analysis of Web-based Content Management Systems for Managing Iowa DOT Engineering Specifications
Budget Estimate
May 1, 2005 to June 30, 2005

Staff Detail

Faculty			Rate*	# of hrs	Salary	Fringe
Geoffrey Sauer			$ 30.58	272	$ 8,316.92	$2,179
Lee Honeycutt			$ 30.66	272	$ 8,340.46	$2,185
Rebecca Burnett			$ 39.88	60	$ 2,392.96	$627
Russ Walters			$ 42.39	40	$ 1,695.69	$444
Hourly students	1	@	$ 9.00	40	$ 360.00	
Total Salary/Fringe benefits:					$21,106	$5,435

Budget Summary by Category

Salaries/Hourly (from detail above)			$21,106
Payroll Benefits (from detail above)			$5,435
Equipment >$5,000			$7,626
Supplies/Materials			$900
Project supplies		$450	
Software		$450	
Other Direct Costs			$2,000
Honoraria/Services (IT Support/Setup)		$2,000	
TOTAL DIRECT COSTS			$37,068
Indirect Costs	@	26.0%	$7,655
TOTAL ALL COSTS			$44,722

1. ISU employees are salaried. Estimate is based on the FY2005 base rates. Annual increases (July 1) and/or midyear promotions or rate changes may affect the level of effort possible under this budget. Hourly rates can be estimated by dividing annual salary by 2080.

2. Fringe rates are estimated as follows: Faculty - 26.2%; P&S - 31%; Merit - 40%; Research Asnts - 11.3%; non student hourly - 12%. Actual fringe will be charged.

3. This project is administered by CTRE (Center for Transportation Research and Education), an ISU non-academic research center that has no ISU department funding for administering this research. Charges must be made to the project to cover basic staff support services, phone calls, copies, etc for the project. Internal accounting systems ensure that such charges are made directly to the appropriate project.

4. ISU charges indirect on ALL direct costs except equipment items over $5000 and each subcontract's cost over $25,000. Indirect rate is determined by a negotiated agreement between Iowa State University and the Department of Health and Human Services. A copy of it can be found at the website http://ospa.iastate.edu/Documents/F&A_Rate_Agreement_2004.pdf . Basic institutional information can be found at http://ospa.iastate.edu/instinfo.htm