



Behind the Web site: An inside look at the production of Web-based textual government information

Kristin R. Eschenfelder*

School of Library and Information Studies, College of Letters and Science, University of Wisconsin-Madison, Room 4259, White Hall, Helen C, 600 North Park Street, Madison, WI 53706-1403, USA

Abstract

This paper describes an exploratory, multisite case study of the production of textual content for state agency Web sites. The qualitative field study explored internal agency Web staff characterizations of textual Web content and staff perceptions of factors affecting the production of content. Study results suggest that staff characterize content in terms of its format, its style age, its rate of change, its degree of change, its owner, and the degree to which it is sensitive. Staff described nine factors affecting content production including information intensity, public education mission, public inquiry burden, top-down directives, existing maintenance burden, review and approval process, resources, management interest and goals, and support from other program staff. A better understanding of how internal agency staff perceive and treat content is important because staff play a large role in determining what content is produced and what characteristics the content contains. The inclusion or exclusion of certain characteristics in content has important implications for information usability, costs, citizen participation in agency policymaking, government transparency, and public trust in government.

© 2004 Published by Elsevier Inc.

Keywords: Government information; Government information management; Web site; Content management; Public administration; Controversial content; Content approval processes

* Fax: +1-608-263-4849.

E-mail address: eschenfelder@wisc.edu.

1. Introduction

Government information has long been seen as a public good. Dissemination of government information educates the citizenry, increases government transparency, and promotes economic development.¹ The current literature sees government Web sites as important tools allowing agencies to increase the amount of information actively disseminated to the public² while reducing paper printing and mailing costs.^{3,4}

Recently, government agency Web sites have come under increased scrutiny due to information censorship and cost concerns. In terms of censorship, critics charge that agencies change, remove, or avoid posting Web content that presents the agency, the current administration, or current policies in a critical light.^{5–7} Further, in the name of national security, agencies have removed potentially sensitive information from Web sites.^{8,9} From a cost perspective, e-government scholars and information managers point to problems with “information overload” as agencies place too much content on their Web sites¹⁰ and costs for maintaining Web content climbs to higher than expected levels.^{11,12}

These concerns about censorship and costs point to the need to understand more about how agencies decide what content to publish on the Web. But the literature currently contains few studies describing agency staff work practices related to Web content. To begin to fill this research gap, this paper addresses two questions:

1. How do agency staff characterize Web content?
2. What factors do Web staff perceive as affecting decisions to publish Web content?

This paper describes an exploratory case study of textual content production for Internet Web sites at four agencies in one Midwestern state.¹³ The study used an exploratory case approach^{14,15} and a grounded theory methodology¹⁶ to develop a rich understanding of how agency Web staff conceptualized text content and what factors Web staff perceived as affecting text content production.

This study’s focus on textual content sets it apart from other e-government research on applications-based content such as e-voting or service registration systems.^{17,18} While important, these studies tend to overlook textual content such as program descriptions, frequently asked questions (FAQs), reports, or instructions. But text content comprises a significant portion of agency Web content. Further, studies suggest that citizen demand for textual information is very high¹⁹, sometimes higher than demand for interactive applications.^{20,21} Finally, because little research has examined the production and management of textual content, it is not clear how management of textual content differs from management of applications content, or to what degree the lessons drawn from studies of e-government applications apply to textual content. The continued importance and ubiquity of textual content in government Web sites, the high public demand for textual content, and the lack of previous research focusing on production and management of textual content necessitate research focused on its unique attributes and management issues.

1.1. Text content in the literature

The e-government literature typically uses textual content as one of many independent variables to help explain outcomes such as costs,^{22,23} user satisfaction,²⁴ citizen trust or involvement in government,^{25,26} or government accountability.²⁷ In these cases, content is typically defined narrowly in order to facilitate assessment of relationships with other constructs. Fewer studies have explored the range and variation of textual content or investigated relationships between types of textual content and other variables.

Web site evaluation studies often examine textual content; but they tend to focus on users' perceptions, usability aspects, and externally perceivable aspects of content.²⁸ For example, studies measure the presence, absence, or quantity of certain types of textual content including official policies (e.g., security, privacy), ownership information, contact information, last update information, annotations, translations, and obviously outdated information.^{29–33} Other studies have examined how often Web pages are updated,^{34,35} the “density” of information on a page,³⁶ the range of subject matter of content,³⁷ the degree to which the language is “user-centered,” and the degree to which text comports with Web writing style best practices.^{38–41}

Less research has examined how internal agency staff perceive content. While understanding the user perspective is important for design purposes, understanding the staff perspective is important for management and policy reasons. For example this paper concludes that staff characterizations of content may affect publication decisions and through these decisions the amount and types of content available to the public.

1.2. Factors affecting the production of content

Managing government Web content is complicated due to Web content's frequently changing nature, the amount of discretionary material published, and the distribution of Web staff and Web content within agencies. Web content changes over time as facts, business goals, design practices, and customer expectations change; thus, maintaining existing content is an ongoing expense for organizations. Some content changes are simple, such as changing a phone number, but other changes are more complex, such as rewriting a program description because of a change in law.⁴²

The discretionary nature of agency Web content also complicates its management. While publication of certain materials may be required by statute, regulation, grants, or political pressure, a good deal of content is published at the discretion of program level agency staff, based on their perceptions of user needs and program goals. But resource limitations restrict the amount of content that agencies can publish and adequately maintain; and staff must prioritize what content should be published and maintained on the Web and what content should only be distributed on paper.⁴³

Further, the staff who create and maintain content files are often distributed within an agency, reporting to different managers in different programs. Managing content therefore requires coordination of disparate staff through standards, policies, and cross-program working groups or councils.^{44,45}

Research has found that lack of resources is a major impediment to management of agency Web sites.^{46,47} Agencies tend to underestimate resources, such as staff time and training, needed to maintain Web site content,^{48,49} and agencies typically develop Web content without consideration of who will maintain it with what resources.^{50,51}

Other findings emphasize that lack of management interest and lack of support from agency staff impede Web site efforts. Upper management and program staff may not be aware of the potential utility of Web content or of the ongoing management needs of Web content.^{52,53} Funding for projects depends on the interest of upper management, which may change with electoral cycles.⁵⁴ Further, lack of reward mechanisms may deter support of upper management and program staff.⁵⁵

In summary, the literature presents Web content characteristics in terms of users' perceptions and users' values. Less work has examined content from the perspective of agency staff. Further, Web site research has stressed the importance of resources and management and staff support, but little work has explored other factors potentially affecting textual content production. This study adds to the existing literature by providing an internal view of the creation of Web content and by identifying other factors that affect the production of content within agencies.

2. Study design and methods

The study used an exploratory case design, which is recommended when the phenomenon of interest is ill defined, is not clearly distinguishable from its context, and is likely composed of complex interacting constructs.^{56,57} The exploratory approach identifies an area for investigation and identifies constructs, generates theoretical tools, and identifies working hypotheses for use in future research.^{58,59} A qualitative approach was necessary to generate rich understandings of staff characterizations of content and staff perceptions of factors affecting content production.⁶⁰

The case sites included four state agencies from one Midwestern state including the Departments of Parks and Environmental Protection (PEP), Welfare, Labor, and Transportation.⁶¹ The study design purposefully selected large, public-facing agencies in order to control for variations introduced by agency size and orientation.

Because the study explored factors affecting Web content production, Web site size was also used as an agency sampling criteria. Three agencies had large Web sites (consisting of over 10,000 HTML files and 2–10,000 PDF files) and one agency had a small Web site (approximately 3000 HTML files and 2000 PDF files). This design allowed comparison of the factors affecting content production from the large agency sites with potentially contrasting results from the smaller agency.¹⁴ Table 1 summarizes demographics for each agency.

Within each agency, criteria sampling was used to select at least three programs based on information provided by the agency Web master.⁶³ Criteria sampling selected programs with high degrees of interaction with the public, at least one high resource program, at least one low resource program, and at least one program producing high profile or potentially controversial

Table 1
Agency demographics⁶²

Agency name	Budget 2001–2003 fiscal year	Number of FTE employees	Number of divisions	Web site size
PEP	9.5 million	2900	7	21,000 HTML 5000 PDF
Welfare	10.5 billion	6000	6	10,000 HTML 10,000 PDF
Labor	2 billion	2251	6	10,000 HTML 2500 PDF
Transportation	4 billion	3500	6	3,000 HTML 2000 PDF

content. Table 2 describes the number of participating programs per agency and the number of participants per agency.

2.1. Data collection and analysis

Data collection included both qualitative and quantitative data. Qualitative data stemmed primarily from semistructured interviews with thirty “Web staff.” Web staff included agency Web masters and other employees that planned, created, or approved Web content. Interviews took place from May through September of 2002 with scattered follow-up interviews occurring from Spring to Fall 2003. Interview transcripts were shared with the participants and mini case reports were shared with key informants at each agency in order to increase validity of the results.

Data collection also included content analysis of agency annual reports and agency information technology strategic plans. Analysis searched the documents for the terms “inform*” “educat*” and “Web” to identify references to public information or education goals and references to textual content. In addition, the researcher reviewed state and agency policies related to Web content and state almanacs describing agency demographics.

Quantitative data were also collected on the size of agency Web sites and agency publications required by the state statutes. Searches using the state search engine generated data about agency Web site size. Searches counted the number of HTML and PDF files using the case agencies’ URLs and file types as search delimiters.⁶⁴ To determine the number reports required by state statutes, the researcher searched the text

Table 2
Number of participants per agency

Agency	Number of participants	Number of different programs participating
PEP	8	5
Welfare	9	4
Labor	8	5
Transportation	5	3

of the statutes for references to an article in the statutes describing report submission procedures.⁶⁵

Analysis employed the grounded theory approach to identify constructs, generate theoretical tools, and identify working hypotheses for future research.^{66,67} Interview analysis employed open and axial coding. Open coding for this paper identified characteristics of content and factors affecting content production. Axial coding investigated links between the characteristics and factors. Content analysis of the documents used simple counts. When possible, analysis compared data from interviews with findings from the other sources.

3. Findings

3.1. Agency content environment

Most content was created by program staff that had received Web publication training. They typically were not information technology specialists, but many had paper publishing or journalism backgrounds. Because they had a high level of knowledge of program activities, these “Web staff” often created content on their own based on their assessment of user information needs. In other cases (e.g., highly specialized or scientific information), Web staff worked with other program experts to create Web content.

3.2. Web staff view of content

Across the four case sites, Web staff described content in terms of six characteristics: its format, its style age, its rate of change, its degree of change, its owner, and its political sensitivity.

3.2.1. Format

Web staff reported that they almost always created Web content from electronic files representing paper documents. Creation of Web content involved conversion of the electronic file to a Web friendly file format, layout, and writing style.

Different conversion strategies required different skills and resources. A basic PDF conversion was seen as requiring the least effort as it did not change the original format of the document. Conversion to HTML could require a variable amount of effort, depending on the degree to which the Web staff altered the document.

Web staff perceived that an important aspect of their work was advising document owners how to improve content usability, how to organize larger sets of content into navigable hyperlinked Web sites, how to comply with agency and state standards (e.g., accessibility), and how to best word sensitive material. Staff reported that their consultation with document owners typically took place over paper copies of the documents with the Web staffer and document owners marking up the paper copy with their ideas.

Staff perceptions of the value of the different output formats varied. Web staff with limited resources often favored PDF formats because PDF was easy to generate. As one participant described, “I used to HTML them, and now I just PDF them. It’s just too much work [to convert them to HTML.]” But in other cases, staff wanted to move away from PDF because of the format’s navigation and search limitations. As another participant described, “To just create it as a PDF file is not necessarily the best way to go.” But the PDF versus HTML decision was not based solely on resources; some content, such as legal documents, was better suited to PDF.

Some staff expressed a desire to create Web-specific content that did not draw from or reflect existing paper documents; however, they were typically not able to pursue these projects due to resource restrictions and the need to first make existing publications available on the Web first.

3.2.2. *Style age*

Participants also described content as having a “style age” based upon the era of Web writing and usability best practices under which it had been created. Some Web staff talked about the need to go back and rewrite content that had been created in an older style age. As one participant noted, “the hardest thing . . . is to get rid of some of the old stuff that’s been on [the Web site].” In many cases, the information contained in this content was still valid, but the content did not meet the Web staff’s perceptions of quality content.

3.2.3. *Rate/degree of change*

Staff also described content in terms of its rate of change and degree of change. Rate of change referred to how often Web staff needed to “touch” or make changes to any of the content to keep it up-to-date. The rate of change is similar to actual measures of change used in the literature,⁶⁸ but the measure reported here represents perceptions of how much work will be required to maintain the content, as opposed to how often the content actually changes.

The degree of change referred to how much text on the page required rewriting. Some changes were simple—for example, updating a telephone number. Other changes were more extensive—such as rewriting health recommendations to meet new federal guidelines. The degree to which content changed had implications for the work required to maintain it, and for what type of review the content would receive before it was published. Typically, content that changed significantly required review from content approvers before it could be published again, while content that only changed in minor ways did not require approval.

3.2.4. *Owner*

Web staff also identified content in terms of its “owner” and how much material the owner had previously asked a Web staffer to publish. In each of the four agencies, the owner of the content was the person who had written the original document or who provided the intellectual content. This could be the Web staffer or a subject expert within the program.

Web staffers were very aware of how often a staff expert provided content for Web publication. Some program experts asked for too much content to be placed on the Web, while others did not produce content unless badgered to do so. As one Web staffer explained, “depending on what area we’re speaking of, some individuals . . . provide way too much, and think that there can never be enough [up on the Web]. Other individuals . . . it is sometimes difficult to get enough updated materials from them. . . .”

3.2.5. *Sensitivity*

Web staff also described content in terms of its sensitivity—the likelihood that the content would generate public controversy, or the degree to which the content might increase scrutiny of the program. In an era of budget cuts, Web content often carried bad news—describing cuts to programs or services, or restrictions of eligibility for benefits. Staff were wary of generating content that might end up in a negative newspaper story. As one participant exclaimed, “you don’t want to be in the press! You know, you don’t want someone to write stories about you.”

Developing sensitive content took great care. Web staff had to ensure that their content reflected the shifting state and federal laws. “We have to be very careful how we say it . . . on our Web site. Just because something might pass we cannot say, ‘This might pass.’” They also had to be extremely careful with language, making certain that text would be interpreted as objective. One Web staffer described how she educated staff during conversions about the potential sensitivity of their content. “They don’t necessarily understand or see the politics in it. And then I think once you explain it to them, they do.”

Sensitive content demanded more resources because it required careful wordsmithing and higher levels of review prior to publication. While lower level managers reviewed most content, sensitive content was typically forwarded to upper level managers. One Web staffer described the intense review his content received. “Yeah, I had four managers that my [content] would go through . . . because it was a very high profile program. . . . And, it took forever to get [the content up]. . . .”

Because sensitive content required special handling and could have negative consequences, some programs avoided it. One Web staffer explained how her program avoided publishing information that might attract attention. “We don’t have problems because our [programs] try to avoid putting out anything that’s controversial.”

While participants expressed frustration with the long review processes for sensitive content, they believed the managers should have the opportunity to review the content—or at least be alerted to its publication.⁶⁹ As one Web staffer described, “I will make sure my divisional administrator, who has to ultimately respond to that [content] being there, [knows about it and I make sure] that he endorses what we are doing.” As another participant described, “We’re pretty politically savvy at the program level, so if it’s controversial, we’ll know to either hold onto it or to touch base if it’s needed.”

One participant described a case in which higher level review led to changes in the content. But the participant did not feel the changes had a negative impact and described the situation as “not a big deal.” “If it’s something that’s very controversial and the

Secretary's Office wants it out, we just take something out—and we explain to the [program] that it's not a huge deal.”

3.3. Factors affecting content production

Data analysis produced nine factors affecting content production including: the information intensity of the agency or its programs, the public education mission of the agency, the desire to decrease time spent on public inquiries, top-down directives, existing content management burdens, content review and approval processes, staffing resources, management interest, and support from other program staff. The last three factors had been previously described in the literature.^{70–72} The findings from this paper add depth to that literature by helping to explain how these factors affect textual content production.

3.3.1. Information intensity

“Information intensity” represented the degree to which agencies produced information products for the public. This factor emerged from interviews where Web staff explained that their programs' relatively extensive Web content stemmed from their production of paper information. One Web staffer from a research program described how her program produced more content for the Web than other programs because its mission was “to produce information.”

It was difficult to develop a good measure of the amount of information that each agency produced because each agency organized and managed publications in a different way. Publication counts at each agency might include different types of information, making cross-agency comparisons difficult. The state depository library reported the number of publications each agency submitted, but review of available agency online publication repositories suggests that these numbers do not reflect the full number of publications produced by each agency. And while it was possible to obtain the number of reports required by the state legislature each year, it was not possible to produce summary information about the number of federal reports each agency produced.⁷³ Table 3 summarizes these data.

3.3.2. Public education goals

Content analysis of agency strategic plans showed substantial variation in emphasis on public education or public information. PEP and Labor texts referred to public education

Table 3
Paper publications produced by agencies

Agency	Number of required legislative reports	Number of publications submitted to depository library year 2001
PEP	127	244
Welfare	202	40
Labor	250	101
Transportation	32	51

or public information goals more than welfare or Transportation. PEP's materials made five references to informing or educating, calling on the agency to "Provide and promote information and education..." and "Use education and enforcement to protect public health, safety, and the environment..." Welfare's strategic plan included three references to informing or educating the public including: "communicate timely, useful, accessible information to individuals, communities and the general public" and "provide health information to assist consumers in decisions..." Labor's strategic plan included five references to educating or informing the public. It would "provide information and assistance" and support the "education of employers and public." Transportation's strategic plan included only one reference, noting the need for "public outreach" and "education" for highway safety issues.

3.3.3. *Public inquiry burden*

Interview data suggest that the desire to reduce the resources spent on answering public e-mail or phone inquiries encouraged content production. Participants described generating new content in response to inquiries. For example, one Web staffer explained, "I am one of the contact people for email that comes in and if we start getting a number of questions about the same thing, I think, 'Do we need something out there?'" Participants believed that the Web content reduced the amount of time they had to spend on public phone and e-mail inquiries. As one participant described, "[the Web] has reduced my level of phone calls by probably 95% over the last couple of years." A content approver explained that programs rarely had too little content because phone inquiries provided an incentive to publish. "If there is [information] missing, [the programs] will put it up there ... to get the phone calls out of their office."

3.3.4. *Top-down directives*

Agencies were relatively autonomous in deciding when and how to publish because each agency maintained an independent Web server, and the state had not implemented a content management system that might centralize content publication.

While participants typically described the development of content as a bottom-up activity, they reported that actions taken by other governmental institutions or upper level managers within the agency affected content generation in several ways.

The state information technology standards affected the usability characteristics of content by creating minimum usability and accessibility requirements and encouraging use of Web design best practices (e.g., short paragraphs, creation of agency template). The standards also required or prohibited certain types of content, for example, they prohibited advertising and required inclusion of contact information.

As noted earlier, the state legislature, via state statutes, required agencies to produce certain annual or biennial reports (as shown in Table 3). Further, all the agencies received federal funds, which often came with reporting requirements.

The agencies operated in turbulent statutory and regulatory environments where changes to federal or state laws often required changes to program content. For example, PEP reported sixty-nine state statutory language changes in its 2003 biennial report.⁷⁴ One participant

described how new legislation spurred content, “We’re . . . legislatively driven, as far as what changes are made or what content might appear. If a new law goes into effect we try to . . . get something on our Web site as quickly as we can, as a way to inform the public.”

Agency strategic plans contained general references to the need to provide information to the public about programs and services, but had few direct references to specific areas of text content. Most of the documents described plans for, or descriptions of, applications development or infrastructure upgrades. PEP documents mentioned development of textual content for land use plans, regulations, and recreational information. Welfare documents called for “strengthening the use of the Web as a communications . . . tool,” but they did not identify specific areas of textual content. Labor documents referred to text content for workplace safety and laws, workers’ compensation, and accessibility laws. Transportation documents mentioned text for one well-known construction project and for highway safety education.

Most content planning was done on an informal basis by program level Web staff, their colleagues, and their managers. PEP and Transportation requested that each program in the agency create a plan defining content production priorities for the year; however, submission of the document was not enforced. Welfare and Labor did not require content plans. Interviews indicated that informal planning occurred within programs by program staff based on staff perception of user information needs and program goals. But because users or constituents could vary widely from program to program, perceived user needs and program goals varied. Some Web staff were highly focused on service to the public, while others’ priorities were providing information to business partners or regulated entities.

Occasionally, preparation of sensitive content was taken over by high-level agency managers. In these instances, higher level managers told programs what to create or what to remove from content. A Web staffer described how content for a high profile program was taken over by high-level agency managers. “. . . We ha[d] a site up for the [preliminary] study, for the past two years . . . and when it was passed to the Governor to make his announcement on it . . . [the managers] kind of took control of our content and started telling us, ‘You have to put this out and put this out and put this out.’”

3.3.5. *Existing maintenance burden*

Interview data suggest that another factor affecting content production was the difficulty of maintaining existing content. Past research has noted the costs associated with the ongoing management and updating of content.⁷⁵ Participants in this study described how perceived maintenance burdens discouraged production of new content. For example, Web staffers described how they specifically discouraged creation of content whose maintenance costs would be too great, “We’ve restrained people on some of their ideas because we know they would be too difficult to maintain.” Another Web staffer explained how concerns about the amount of work needed to maintain content discouraged people from creating content. “A lot of people, I think, are reluctant to get involved [with content] because they’re afraid that—well, [if] you put a page up, you’ve got to maintain it.”

3.3.6. *Review and approval process*

Interview data and analysis of agency Web management processes suggest that the nature of the review process affected content production. Typically, Web staff created content and then passed it on to managers for review. The number of levels of managerial review required varied by content type and by agency, with different agencies having different requirements.⁷⁶ In each agency, significant changes to existing content, or the creation of new content, required more review—typically, review by both local managers, divisional managers, and high-level agency managers. PEP and Transportation required the more levels of review than Labor or Welfare. PEP required the approval of an agency review board that met every few weeks. Transportation required approval by the agency secretary's office.

The burdensome nature of the PEP and Transportation review processes may have discouraged generation of content that qualified as a “significant change” or “new content.” But this does not mean that staff developed less content overall. Rather, staff developed work-arounds to avoid having content classified as a “significant change” or “new content.” For example, instead of creating a new Web page for some information, they would place the information on an existing page and put a graphical box around it to draw attention to it. In this manner, they were able to publish while sometimes avoiding the delays inherent in the approval process.

Content approvers saw the approval process as a tool to ensure that only important, high-quality content was produced. Other Web staff agreed that content should be reviewed but complained of the delay imposed by review. For them, the approval process was red tape to be managed in order to publish content swiftly.

3.3.7. *Resources*

This study confirms earlier findings that resources affect e-government efforts. Most participants complained that budget cuts had negatively affected content generation. As one participant explained, “Nobody has the time around here [to create content] because we're . . . in a terrible budget crisis.” A review of agency statistics in the state almanac showed that all agencies but Labor lost full-time positions from 2002 to 2004. In interviews, participants also reported losing a large number of part-time Web positions, which were in addition to the full-time position cuts reported in the almanac.

This study measured resources in terms of “span of service” and the percent of time each Web staffer could devote to Web work. Span of service was the ratio of total staff to Web staff within a program. If Web staff were only creating content for a small number of program staff, then arguably they could produce more content. Similarly, if they could spend a large percentage of their time on Web content, then presumably they could produce more content. One author who could devote less than 10 percent of her work time to the Web and served a large number of program experts expressed frustration about the amount of content she was able to produce. “I don't want to say we're lagging behind [in content], but our bureau has a lot of program areas. I mean it's not like a bureau that has maybe five program areas . . . you can [easily] get information out on five.”

Table 4 describes the Web staff span of service at each agency. Labor Web staff had the smallest span of service while Transportation has the most.

The percent of time that Web staff could spend on content work varied both by agency and by program within agencies. Most PEP Web staff reported spending 25 percent or more of their time on Web content. In contrast, Welfare Web staff spent 5 percent or less of their time on Web content.⁷⁷ Most Labor Web staff reported spending less than 25 percent of their time on Web work. Web staff at Transportation reported devoting at least 40 percent of their time to Web work. Therefore, although Transportation has the largest span of service, its few Web staffers spent more of their time on content than other agency Web staffers.

3.3.8. Program management interest and goals

Study results also confirm previous findings that management interest is a major factor in e-government projects. Participant depictions of program level management interest varied by program, with some program managers encouraging content generation and others discouraging the production of anything more than basic program descriptions.

3.3.9. Support from other program staff

The study results confirm the importance of support from other program staff previously reported in the literature.⁷⁸ Support from staff colleagues was important because the colleagues often provided new content and updates for existing content. In some cases, colleagues would not provide content, forcing the Web staff to knowingly leave out of date content up on the Web site. One Web staffer explained, “I’ve been fighting to get them to replace the information that’s up there. [I ask] ‘When are we going to get the new stuff?’ And they always plead, ‘Other things are more pressing.’”

4. Discussion

4.1. A richer conceptualization of textual Web content

This study sought to increase understanding of how government agency Web staff characterize content and what factors Web staff perceive as affecting the decision to publish content.

Table 4
Span of service for each case agency

Agency	Span of Service: Number of FTE program staff per Web staffer	Typical percent of work time program Web staff spend on Web content
PEP	28	25% or more
Welfare	33.5	5% or less
Labor	9.5	25% or less
Transportation	129	40% or more

Table 5 summarizes the content characteristics described by the participants. Several of the characteristics are unique to the internal staff perspective including format, degree of change, degree of sensitivity, and owner.

A richer understanding of staff characterizations of content illuminate the complexities involved in decisions about publishing content. The decision to publish is more complex than merely deciding whether or not to publish. Staff must decide what format to publish in, whether or not to change the text to improve usability, what rate and degree of change to build into the content, and whether or not to include sensitive information.

Results suggest that staff have different strategies for creating content, and these strategies affect the types of content created. “Basic” content strategies aim to create content with simple formatting, low rates of change, low degrees of change, and low degrees of sensitivity.⁷⁹ “Complex” content strategies allow for more sophisticated formatting, higher rates of change, higher degrees of change, and higher degrees of sensitivity. Fig. 1 suggests a continuum of content strategies.

Publication practices are guided by localized knowledge of the factors both promoting and discouraging content production. For example, in a low resource program, Web staff using a basic strategy would avoid not only controversial content, but also content that requires frequent or extensive updating. Web staff in different situations might choose different strategies, creating content with different characteristics.

These findings have important implications for future research on electronic government information. Because Web staff play a significant role in shaping the characteristics of content, research that seeks to understand government information publication decisions needs to consider the role of program level agency Web staff in shaping the characteristics of the content.

4.2. Agency differences

Data analysis identified nine factors affecting Web content production including the following:

Table 5

The characteristics of Web content from agency Web staff perspective

Characteristic	Definition
Style age	When content was created or to what extent it deviates from current stylistic and usability best practices.
Format	The end result of the conversion process including the file format (PDF or HTML) and the degree of complexity of the HTML conversion.
Degree of sensitivity	The degree to which content refers to potentially controversial issues or might attract scrutiny of the press or agency management.
The rate of change	How often staff expect to update a piece of content.
The degree of change	How much of the content will need to be changed as part of the updating process.
Owner	Who is responsible for content and how much content that person typically generates.

Basic content strategy: Leave up content of an older style age if it is still accurate. Create content with lower rates of change and lower degrees of change. Create more PDF files. Keep content as non-sensitive as possible.



Complex content strategy: Revise all content to comport with current best practices. Create more dynamic content with higher rate of change to keep site “fresh.” Convert more content to sophisticated HTML formats (especially large publications). Consider addition of content that might draw increased scrutiny.

Fig. 1. Content strategy continuum.

- information intensity;
- public education mission;
- public inquiry burden;
- top-down directives;
- existing maintenance burden;
- review and approval process;
- resources;
- management interest and goals; and
- support from other program staff.

The study was designed, in part, to compare the factors affecting content in agencies with large Web sites and small Web sites; and research began with the expectation that because of its smaller Web site, Transportation Web staff might have different characterizations of content, experience different factors, or experience the same factors differently. But overall, content production at Transportation was very similar to content production at the other agencies: content planning was done by programs, content conversion was done by program level Web staff with other responsibilities, and existing paper documents provided most of the fodder for textual content.

But the results suggest some differences. Transportation had a weaker public education mission as measured by references to educating or informing the public in its annual reports; it had fewer required legislative reports; it did not submit very many publications to the depository library; and it had a burdensome content approval process. Finally, Transportation had fewer FTE staff for content development as measured by span of service—though staff spent more of their work time on content activities.

One could argue that the results suggest that agencies with weaker public education missions or more burdensome approval processes will have smaller Web sites. But the

exploratory design of the study means that results can only suggest relationships; further research is needed to test them.

Further research should investigate relationships between the factors and program level content production. Contingency theory research suggests the relationship between management structures or methods of production and performance (e.g., the amount of content produced) is strongest at the subunit level.^{80–82} The researcher plans further inquiry at the program level to investigate these relationships.

4.3. Implications for usability, satisfaction, participation, and trust

The results of the study suggest that the nine factors affect not only the amount of information provided to the public, but also characteristics of the content; and, the inclusion of certain characteristics has important implications for usability, customer satisfaction, costs, government transparency, and citizen participation and trust in the government.

The types and characteristics of content created by Web staff have usability and user satisfaction implications. Web usability texts report that characteristics of content (such as writing style, use of bullets, short paragraphs, etc.) affect user satisfaction and use of information.^{83,84} But Web staff reported that due to resource constraints, they sometimes used basic content strategies that produced less user-friendly content. This suggests that higher resource programs maybe able to produce more usable content than others.

The types of content created by staff also have cost implications. Staff reported producing content mainly drawn from existing paper documents. Resource constraints and the need to put existing paper documents on the Web limited their ability to generate unique content for the Web. Moving beyond converting and posting existing paper documents would likely require more resources, which could raise agency-operating costs. This calls into question the extent to which e-government will really lead to reduction in agency operating costs.⁸⁵

Content characteristics may have important implications for citizen participation and government transparency. Results suggest that some programs avoid creating controversial content because it has higher costs in terms of approval requirements and potential scrutiny. Yet the goal of using e-government to enable political participation^{86,87} would seem to require that agencies actively disseminate information that may create controversy. While agencies seem enthusiastic about developing online comment submission systems for official policy making processes, results from this study and anecdotal evidence of Web site “scrubbing” suggest that agencies may be less willing to publish information that may promote critiques of established policies or current agency management.^{87,89,90} Yet access to this information may be an important part of citizen participation and open government.

This treatment of content also has significant public trust implications. Data from this study show that sensitive content is typically subject to extensive review by upper level agency managers, many of who are political appointees. Web staff typically agreed that these managers should review content for which they would be held accountable. But the finding that political appointees may review content, and the finding that top-down content changes do occur, raises serious concerns about the degree to which agency content may be shaped to cast the current administration in the best light possible, shield the agency and administration from criticism, or

assuage public concerns. Editing of government information, such as the health recommendations related to the World Trade Center bombings,⁹¹ or the removal of critical information in Justice Department diversity report⁹² may reduce public trust in government information if citizens begin to believe that agencies will not publish important, but potentially controversial content, or that high-level managers will “massage” content to keep administrative approval ratings high. This reinforces concerns raised in previous scholarship that describes how agencies have always manipulated public information by choosing which publications to actively disseminate⁹³ and what information to include in those documents.⁹⁴

This finding points to the conflict between views of government Web sites as neutral conduits of facts and government Web sites as tools to create a certain public perception of the agency and the administration. Some theorists argue that all information is socially shaped and that the interests and biases of powerful groups are embedded into information and technological structures such as Web sites.^{95–97} On the other hand, to say that all Web sites are equally “political” seems an oversimplification; more likely, Web sites vary in the degree to which they are politically shaped. For example, a portal or cross-agency site may be more subject to political influence than a program Web site. A Web site about a high profile or controversial topic (e.g., Mad Cow Disease, teen pregnancy) may be subject to greater political influence than Web sites about ski trails. Further research is needed to evaluate what types of Web sites are most likely to experience political shaping, what constitutes shaping, and the effect of shaping on public trust.

4.4. Limitations

The generalizability of the study is limited due to its exploratory design and the case selection criteria. Agencies of different sizes or different orientations may experience different factors affecting content production. Further, agencies with more centralized content management processes likely experience different factors. Further, content on different types of Web sites (e.g., state portals) may be managed in different ways and thus be affected by different factors. Further comparative case studies are needed to extend the categories and relationships suggested by the findings other agency types and other Web environments. Further, the study’s focus on static, text-based content also limits the generalizability of the findings. It is not clear to what extent the findings of this study will hold true for dynamic textual content or applications content.

Future research is also necessary to test the relationships between content types, factors affecting content production, and Web site size. As an exploratory case study, this work sought to identify important factors, generate theoretical tools, and develop working hypotheses. More research is necessary to test the circumstances under which the findings suggested by the data are valid.

5. Conclusion

This paper described an exploratory, multisite case study of agency staff characterizations of textual content for state agency Web sites. Study results suggest that staff distinguish between

content in terms of its format, its style age, its rate of change, its degree of change, its owner, and its sensitivity. Further, results suggest nine factors affecting content production including agency information intensity, public education mission, public inquiry burden, top-down directives, existing maintenance burden, review and approval process, resources, program management interest and goals, and support from other program staff.

The results of the paper make an important contribution to the e-government and government information literatures by providing an internal view of Web content production and agency staff characterizations of content. A greater understanding of how agency staff perceive and treat content is important because agency Web staff influence what content is published and the characteristics of the content. Content production decisions take place within particular contexts where different factors may encourage or discourage content production and certain characteristics of content. These content decisions have important implications for information usability, costs, citizen participation in agency policymaking, government transparency, and public trust in government.

6. Uncited reference

88

Acknowledgments

Thanks to Alan Rubel, Sunny Kim, Ethelene Whitmire, Greg Downey, and Louise Robbins for comments on earlier versions of this manuscript. Thanks also to the anonymous GIQ reviewers for many helpful suggestions. Thanks also to the study participants for their time and candor.

References

1. Herson, P., & Relyea, H. (1995). Government publishing: Past to present. *Government Information Quarterly*, 12(3), 309–330.
2. Sprehe, J. (1999). Government information: From inaccessibility to your desktop and back again. *Journal of the American Society for Information Science*, 50(4), 340–345.
3. Fountain, J. (2002). *Developing a basic research program for digital government: Information, organizations and governance*. National Center for Digital Government, Kennedy School of Government, Harvard University.
4. Relyea, H. (2002). E-gov: Introduction and overview. *Government Information Quarterly*, 19, 9–35.
5. Connolly, C. (2002, October 22). Hill group faults HHS for ideology. *Washington Post*, A25.
6. OMB Watcher. (2002, October 28). Ideology trumps science at HHS, letter charges. Available: <http://www.ombwatch.org>. Accessed September 19, 2003.
7. EPA (Office of the Inspector General). (2003, August 21). Evaluation report: EPA's response to the World Trade Center Collapse: Challenges, Successes, and Areas for Improvement. In *Report No. 2003-P-00012* (pp. 14–17). Chapter 2 Council on Environmental Quality Influenced EPA Press Releases. U.S. Environmental Protection Agency.

8. GODORT. (2004). Task force on restriction on access to government information. Available: <http://www2.library.unr.edu/dept/bgic/Duncan/RAGI.html>. Accessed January 2004.
9. Halchin, L. (2002). Electronic government in the age of terrorism. *Government Information Quarterly*, 19, 243–254.
10. Fountain, J.E. (2001). *Building the virtual state: Information technology and institutional change*. Washington, DC: Brookings Institution Press.
11. Dawes, S., Pardo, T., & DiCaterino, A. (1999). Crossing the threshold: Practical foundations for government services on the World Wide Web. *Journal of the American Society for Information Science*, 50(4), 346–353.
12. Christensen, E., & Hughes, J. (2000). Putting your agency on the Web. In: G. Garson (Ed.), *Handbook of public information systems* (pp. 563–576). New York: Marcel Dekker.
13. The study did not include content related to agency intranet sites or restricted access extranet sites.
14. Yin, R. (2003). *Case study research: Design and methods*. Newbury Park: Sage.
15. Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.
16. Strauss, A., & Corbin, J. (1998). *Basics of qualitative research*. Thousand Oaks: Sage.
17. Moon, M.J. (2003). Can IT help government to restore public trust? Declining public trust and potential prospects of IT in the public sector. In: R. Sprague (Ed.), *Proceedings of the 36th Hawaii International Conference on Systems Sciences, Vol. HICSS 03*. IEEE Computer Society.
18. Cresswell, A., & Pardo, T. (2001). Implications of legal and organizational issues for urban digital government development. *Government Information Quarterly*, 18, 269–278.
19. Madden, M. (2003, December 22). America's online pursuits: The changing picture of who's online and what they do. *Pew Internet and American Life Project* [Online]. Pew Internet and American Life Project. Available: <http://www.pewinternet.org>. Accessed January 2004.
20. Thomas, J., & Streib, G. (2003). The new face of government: Citizen-initiated contacts in the era of e-government. *Journal of Public Administration Research and Theory*, 13(1), 83–102.
21. Gartner Research. (2002). *E-government: What are citizens really looking for?* (Andrea Di Maio No. COM-13-3960).
22. Moon, M. (2002). The evolution of e-government among municipalities: Rhetoric or reality? *Public Administration Review*, 62(4), 424.
23. Christensen and Hughes (2000).
24. Barnes, S., & Vidgen, R. (2003). Assessing the quality of a cross-national e-government Web site: A case study of the forum on strategic management knowledge exchange. In: R. Sprague (Ed.), *Proceedings of the 36th Hawaii International Conference on Systems Sciences, Vol. HICSS 03*. IEEE Computer Society.
25. Moon (2003).
26. Kakabadse, A., Kakabadse, N., & Kouzmin, A. (2003). Reinventing the democratic governance project through information technology? A growing agency for debate. *Public Administration Review*, 63(1), 44–57.
27. Welch, E., & Wong, W. (2001). Global information technology pressure and government accountability: The mediating effect of domestic context on Website openness. *Journal of Public Administration Research and Theory*, 11(4), 509–538.
28. Huang, C., & Chao, M. (2001). Managing WWW in public administration: Uses and misuses. *Government Information Quarterly*, 18, 357–373.
29. West, D. (2002). *State and federal e-government in the United States, 2002*. Brown University: Center for Public Policy.
30. Gant, J., & Gant, D. (2002). Web portal functionality and state government e-service. *Proceedings of the Hawaii International Conference on Systems Sciences, HICSS(36)*.
31. Andersen, D. (2002). *A usability analysis of selected federal government Websites*. Washington, DC: Andersen Office of Government Services.
32. McClure, C., Sprehe, T., & Eschenfelder, K. (2000). *Performance measures for agency Websites*. Washington, DC: GPO Access Federal Bulletin Board. Available: <http://fedbbs.access.gpo.gov/library/download/MEASURES/measure.pdf>. Accessed February 2004.

33. Eschenfelder, K.R., Beachboard, J., & McClure, C. (1997). Assessing U.S. federal government Websites. *Government Information Quarterly*, 14(2), 173–190.
34. Shi, W. (2002). The contribution of organizational factors in the success of electronic government commerce. *International Journal of Public Administration*, 25(5), 629–657.
35. Koehler, W. (2002). Web page change and persistence—A four-year longitudinal study. *Journal of the American Society for Information Science and Technology*, 53(2), 162–171.
36. Ryan, T., & Olfman, L. (2003). The evolution of U.S. state government home pages from 1997 to 2002. *International Journal of Human–Computer Studies*.
37. Musso, J., Weare, C., & Hale, M. (2000). Designing good Web technologies for local governance reform: Good management or good democracy? *Political Communication*, 17, 1–19.
38. Andersen (2002).
39. Birdsell, D., & Muzzio, D. (1999). *Government programs involving citizen access to Internet services*. (Markle Foundation). Markle Foundation.
40. Eschenfelder et al. (1997).
41. McClure et al. (2000).
42. Content management tools such as Web entry forms for database-driven sites may aid in the maintenance of content, but these tools do not negate the need for someone to actually rewrite the existing text.
43. Sprehe (1999).
44. Hert, C., Eschenfelder, K., McClure, C., Rubin, J., Taffet, M., Abend, J., & Pimentel, D. (1999). *Evaluation of selected Websites at the US Department of Education: Increasing access to Web-based resources*. Syracuse, NY: Information Institute of Syracuse.
45. Eschenfelder, K.R. (2004). How do government agencies manage the text content on their Web sites? A framework to compare Web content management practices. *Library and Information Science Research*, 26(4).
46. Ho, A. (2002, July/August). Reinventing local governments and the e-government initiative. *Public Administration Review*, 62(4), 434–444.
47. Fountain (2001).
48. Ibid.
49. Dawes et al. (1999).
50. Ibid.
51. Christensen and Hughes (2000).
52. Ho (2002).
53. Andersen (2002).
54. Coursey, D., & Killingsworth, J. (2000). Managing government Web services in Florida: Issues and lessons. In: G. David Garson (Ed.), *Handbook of public information systems* (pp. 331–344). New York: Marcel Dekker.
55. Fountain (2001).
56. Yin (2003).
57. The data described in this paper were drawn from a larger study of how state agencies manage the textual content on their Web sites and the processes by which content is produced and approved for publication (Eschenfelder, in press). This paper focuses primarily on data describing how agency staff members decide what types of content to publish.
58. Yin (2003).
59. Eisenhardt (1989).
60. Denzin, N., & Lincoln, Y. (1998). *Collecting and interpreting qualitative materials*. Thousand Oaks: Sage.
61. All the agency names used in this article are pseudonyms.
62. The figures in the table are rounded off to help preserve the confidentiality of the participating agencies.
63. This paper uses the term “program” to refer to all agency subunits (e.g., divisions, bureaus, and programs). Most of the subunits participating in the study were bureaus or programs within larger bureaus.
64. The vast majority of content at each agency was published on agency-controlled servers and identified with

- file names using the agency URL (e.g., <http://www.agencyname.state.us/programarea/program1>). A few programs maintained applications on outsourced servers. This content would not have appeared in the results.
65. The state statutes described the process by which agencies should submit reports to the state legislature (i.e., who to give the report to, how many copies). When the legislature required a report from an agency, it typically made reference to the section of the statutes describing the process. It is possible that some report requirements did not reference this section.
 66. Strauss and Corbin (1998).
 67. Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks: Sage.
 68. Shi (2002).
 69. Web staff did disagree with upper level management on other aspects of the content review process. Readers interested in the review process should refer to Eschenfelder, K. R. (2003). The importance of articulation work to agency content management: Balancing publication and control. Proceedings of the 36th Annual Hawaii International Conference on Systems Sciences (HICSS-36) IEEE Computing Society. Available: <http://www.hicss.hawaii.edu/>. Accessed January 2003.
 70. Ho (2002).
 71. Dawes et al. (1999).
 72. Andersen (2002).
 73. Federal reporting requirements (in terms of publications or data) are substantial. For example, a search of the U.S. Department of Transportation site for documents, which contained the state name in the title of the document, generated 406 results. This suggests that the case site is providing a great deal of information to the U.S. DOT.
 74. Not all agencies reported the number of statutory changes per year.
 75. Dawes et al. (1999).
 76. This paper provides only a summary description of the content production, review, and approval processes at each case site. Readers interested in these processes and agency and program level differences between these processes should see Eschenfelder.
 77. Percent of time spent on Web work varied by programs within agencies. While the numbers in the table represent the most common response for each agency, at least one program within each agency had full-time Web staff.
 78. Fountain (2001) and Ho (2002).
 79. Not all PDF files produced are done as part of a “basic” content strategy. In some cases, PDF may be preferred—regardless of the availability of resources or other factors. For example, agencies posting legal documents may choose a PDF format in order to best retain the original formatting and appearance of the document.
 80. Gerwin, D. (1979). The comparative analysis of structure and technology: A critical appraisal. *Academy of Management Review*, 4(1), 41–51.
 81. Gillespie, D., & Mileti, D. (1977). Technology and the study of organizations: An overview and appraisal. *Academy of Management Review*, 2(1), 7–16.
 82. Alexander, J., & Randolph, W. (1985). The fit between technology and structure as a predictor of performance in nursing subunits. *Academy of Management Journal*, 28(4), 844–859.
 83. National Cancer Institute. (2004). *Research-based Web design and usability guidelines*. Available: <http://usability.gov/guidelines/index.html>. Accessed February 12, 2004.
 84. Nielsen, J. (2000). *Designing Web usability: The practice of simplicity*. Indianapolis, IN: New Riders Publishing.
 85. Moon (2002).
 86. Kakabadse et al. (2003).
 87. Musso et al. (2000).
 88. OMB Watcher (2002).
 89. Connolly (2002).

90. EPA (2003).
91. Ibid.
92. Johnston, D., & Lichtblau, E. (2003, October 31). Critical study minus criticism of justice dept. *New York Times*, 1 (section A, column 4).
93. Sprehe (1999).
94. Hernon and Relyea (1995).
95. Bowker, G., & Star, S. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.
96. Sawyer, S., & Eschenfelder, K. (2002). Social informatics in the information sciences: Perspectives, trends and research opportunities. In: B. Cronin, & D. Shaw (Eds.), *Annual Review of Information Science and Technology*. Silver Spring, MD: American Society for Information Science and Technology.
97. Eschenfelder, K.R. (2003). The customer is always right, but whose customer is more important? Conflict and Web site classification schemes. *Information Technology and People*, 16(4), 419–439.