Tracking "Change": The Importance of Applying Version Control to Government Resources

Shannon N. Simmons
Center for Information Policy and E-Government
College of Information Studies, University of Maryland
4105 Hornbake Building
College Park, MD 20742-4343
simmons5@umd.edu

Justin M. Grimes
Center for Information Policy and E-Government
College of Information Studies, University of Maryland
4105 Hornbake Building
College Park, MD 20742-4343
jgrimes2@umd.edu

Elizabeth M. Bonsignore
Maryland Institute for Technology in the Humanities
Center for Advanced Study of Communities and Information
College of Information Studies, University of Maryland
4105 Hornbake Building
College Park, MD 20742-4343
ebonsign@umd.edu

ABSTRACT

Transparency and information access are crucial components of a successful democratic society. President-elect Obama has pledged to use all available technologies to make the government more transparent and accountable to its citizens. Version control provides a perfect example of one such technology which is useful for fostering accountability, increasing public confidence, and encouraging participation. However, as a new technology, the design and implementation of version control raises important issues of public policy.

Version control helps track and manage changes in the process of digital information. Version control, also referred to as revision control, exists to manage this in a structured, efficient manner through the collocation and management of the same unit of information in a digital document. It is in essence, the "art of managing changes to information" [6].

This idea is not a radically new concept and has existed in the software development community for years, invented to handle issues arising from multiple developers working on source code at the same time. Version control is not unique to software development; in fact, version control has been successfully extended across other industry and academic workflow applications, and now exists as a fairly common concept in the modern world (examples include Microsoft Word's "Track Changes" tool at the document level or a content management system at the networked repository level).

Applying open access version control techniques to government resources can provide several benefits: improved efficiencies, increased access to information, increased participation, and preservation of governance processes and artifacts. Resources for information production and exchange must be available, easily accessible and open for all to ensure the networked economy of the 21st century will flourish [3]. Many scholars have argued the importance of information access to society and the democratic process [2, 1, 4, 9, 11, 12, 17]. Version control techniques would enhance both information access and the democratic process.

Opening version control to the public sphere to support the management of government documents in general and e-government in particular can increase efficiency, improve government decision-making and dissemination processes, but most importantly, provide transparency. The federal government generates massive amounts of information and managing that government information is, in turn, a very complex process. Just as version control tools supported rapid development/test cycles, improved quality control, team dynamics and decision-making in the software engineering realm, parallel arguments can be made in favor of the use of version control to support participatory governance activities ranging from increased access to government documents to the collaborative creation of legal codes.

For the citizens and the general public, open access version control opens windows in terms of access to government information. Instead of seeing the final form, citizens would see the entire process. Version control would allow constituents to hold legislative body accountable; the entire legislative process of collaboratively writing and constructing documents could become more efficient, effective and productive. Researchers, librarians, citizens and other vested stakeholders have argued that version control would add value, as it would increase access to and the preservation of information [13, 19].

It has been said that "Laws are like sausages, it is better not to see them made". The irony of this famous Otto Von Bismark quote is that, much like sausages, it is not the theoretical process, but the transformation of the product itself, that disturbs people. This current Prussian approach is counterintuitive: if people find the process disturbing, we should not hide the process, but use the opportunity to make a more appealing sausage. If as U.S. Supreme Court Jus-

tice Louis Brandeis said "Sunlight is the best disinfectant", then version control would be a magnifying lens. However, two questions remain when discussing the implementation of a version control system: what is the level of information granularity and what is the depth of usage and breadth of usability.

Information granularity exists on a sliding scale. Version control can exist at the macro or document level, answering the question: what version of the document is this? While this approach can increase users' certainty that they are working on the most current version of a document, the real advantage of version control can be seen at the micro level. Micro level version control enables the tracking of atomic bits of the document including sentences, words and ideas, and overall document structure as these items are negotiated throughout the writing process. This information is just as essential, yet void if the granularity is set at the macro document level.

The usage problem, simply stated, is that version control only works if everyone uses it. The value is lost if changes are not handled by the system. Therefore, from a design perspective, version control should be required and automated for all users at every level. Furthermore, on the issue of usage, the ability to view and track changes must be intuitive and easy to discern on the part of the end user.

The case for the application of version control to government documents and e-government can be best illustrated through the following examples: Congressional legislation and President Elect Obama's *change.gov* website.

Congressional legislation

Version control for legislative information systems has been long argued for by government outsiders such as Karl Fogel and Tim O'Reilly and many nonprofit organizations [5, 15]. Applying version control systems and techniques to legislation would provide a wealth of advantages to a multitude of stakeholders. Version control would bring greater transparency, openness and accountability to the legislative process. Every piece of text from legislative documents would be trackable and traceable, leaving Congress no place to hide text anonymously. This type of system could have solved the problem of anonymous earmarks, which was only recently taken care of through legislative reforms.

Change.gov

Throughout his campaign President-elect Barack Obama placed great emphasis on the importance of using cutting edge technologies to create a more transparent and connected democracy and even included such statements in his technology policy (see: `http://www.barackobama.com/issues/technology/transparent-democracy`). Two days after winning the presidential election, Obama launched *change.gov*, which serves as an official source of information on the on news, events and announcements regarding the transition to an Obama administration. While many took this as evidence that the administration would continue to rely on the bottom-up open communication principles which helped to drive the campaign, some were quick to point out that within days the site had taken its agenda down for revisions, replacing the content with "three broad sentences about what Obama hopes to do" [22]. The New York Times noted that "there was an almost instantaneous outcry from bloggers and other advocates of transparency in government who noticed the disappearance" [8]. Members of the blogosphere began tracking changes themselves, which one person described as being noticeable only to those who kept a very close eye on the site[22]. Bloggers' calls for a "recent changes module" or "website as a wiki" have recently been also been backed by Tim O'Reilly, who suggests that the Obama administration start the version control movement on *change.gov* to "demonstrate that the system works, that it has enormous benefits in transparency" [16], and use that to work toward what he calls the "holy grail" of providing revision control for all government documents.

Preservation of Information

Possibly most importantly, without the use of open-access version control, the evolution of government materials and the processes underlying their development will be lost to posterity. For the people, open access to the historical record, as it is being formed and archived, can engender in them a sense of individual agency that is the foundation of democratic ideals [14]. For scholars, open access, coupled with emerging tools of digital humanities scholarship, would enable analysis of the process and products of governance from multiple perspectives across time and space, across diverse disciplines, in more intimate detail than ever before [7, 10, 18, 20, 21].

This poster will be used to discuss the potential of versioning control and computer supported cooperative work research towards the application of government resources and what the next necessary steps are in terms of research.

Keywords

e-government, information policy, version control, preservation, information access

1. REFERENCES

[1] G. Barnum. Availability, access, authenticity, and persistence: Creating the environment for permanent public access to electronic government information. *Government Information Quarterly*, 19:37–43, 2002.

[2] G. Bass. Moving toward a 21st century right-to-know agenda: Recommendations to president-elect obama and congress., November 2008.

[3] Y. Benkler. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, New Haven, CT, 2006.

[4] J. Brito. Hack, mash & peer: Crowdsourcing government transparency. *SSRN eLibrary*, 2007.

[5] N. Cohen. Crossing out, for emphasis., July 2007.

[6] B. Collins-Sussman, B. W. Fitzpatrick, and C. M. Pilato. *Version control with Subversion*. O'Reilly Media, Sebastopol, CA, 2004.

[7] G. Crane, A. Babeu, D. Bamman, L. Cerrato, and R. Singhal. Tools for thinking: ephilology and cyberinfrastructure. Technical report, Council on Library and Information Resources National

Endowment for the Humanities, September 15, 2008 2008.

[8] M. Falcone. Changes at change.gov: Return of 'agenda.', November 2008.

[9] Z. Fang. E-government in digital era: Concept, practice and development. *International Journal of the Computer, the Internet and Management*, 10(2):1–22, 2002.

[10] E. Folsom. Database as genre: The epic transformation of archives. *PMLA: Publications of the Modern Language Association of America*, 122(5):1571–1579, 2007.

[11] P. T. Jaeger. Information policy, information access, and democratic participation: The national and international implications of the bush administrationâĂŹs information politics. *Government Information Quarterly*, 24(4):840–859, 2007.

[12] P. T. Jaeger and G. Burnett. Information access and exchange among small worlds in a democratic society the role of policy in shaping information behavior in the post-9/11 united states. *Library Quarterly*, 75(4):464–495, 2005.

[13] K. R. Murray and I. K. Hsieh. Archiving web-published materials: A needs assessment of librarians, researchers, and content providers. *Government Information Quarterly*, 25(1):66–89, 2008.

[14] M. A. Norcia. Out of the ivory tower endlessly rocking: Collaborating across disciplines and professions to promote student learning in the digital archive. *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture*, 8(1):91–114, 2008.

[15] T. O'Reilly. Why congress needs a version control system, July 2007.

[16] T. O'Reilly. Put change.gov under revision control!, November 2008.

[17] A. C. Quinn. Keeping the citizenry informed: early congressional printing and 21st century information policy. *Government Information Quarterly*, 20(3):281, 2003.

[18] S. Schreibman, A. Kumar, and J. McDonald. The versioning machine. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing*, 18(1):101–107, 2003.

[19] B. Selby. Age of aquariusâ the fdlp in the 21st century. *Government Information Quarterly*, 25(1):38–47, 2008.

[20] D. Smith. Textual variation and version control in the tei. *Computers and the Humanities*, 33(1-2):103–112, 1999.

[21] STOA. Stoa consortium for electronic publication in the humanities, 2008.

[22] M. Whitney. Change.gov pulls its agenda, November 2008.