Editor: **Hsinchun Chen,** University of Arizona, hchen@eller.arizona.edu

AI, E-government, and Politics 2.0

Hsinchun Chen, *University of Arizona*

The advent of Web 2.0 has stirred much excitement and created abundant opportunities for reinventing businesses and governments. The recently concluded 2008 US House, Senate, and Presidential elections have provided the first signs of success for online campaigning and political participation. In what has been dubbed "politics 2.0," politicians are using highly participatory and multimedia Web platforms for successful policy discussion, campaign advertising, voter mobilization, event announcement, and even more importantly, online donations. The highly visible and successful Barack Obama campaign provides a testament to the power of Web 2.0 and the youth-based grassroots political movement.[1,2] For example, during the 2008 elections, the Obama campaign raised US$16.5 million in "micro" contributions. Obama averaged $1 million per day for major parts of his campaign.[3] After winning the presidential election, Obama has continued the "reinventing the government" theme in the White House. New "government 2.0" initiatives are under way, including presidential blogs; wikis for policy discussions; transparency in scheduling, meetings, and donations; communities of experts; and so on.

As the government and political process become more transparent, participatory, online, and multimedia rich, there is a great opportunity for adopting advanced AI and intelligent systems research in e-government and politics 2.0 applications. Selected techniques in data, text, Web, and opinion mining, social network analysis, visual analytics, multimedia analysis, ontological representations, and social media analysis can support online political participation, e-democracy, political blogs and forums, e-government service delivery, and transparency and accountability.[4]

Web 2.0 and the Wisdom of the Crowds

Unlike the basic publication, access, and content delivery model of Web 1.0, Web 2.0 promises to be more interactive, participatory, content rich, seamless, scalable, and service oriented. The term Web 2.0, coined by Tim O'Reilly, "is a trend in the use of the WWW technology and Web design that aims to facilitate creativity, information sharing, and collaboration among users. Web 2.0 is the business revolution in the computer industry caused by the move to the Internet as a platform, and an attempt to understand the rules for success on that new platform."[5] Web 2.0 applications have become pervasive and have appeared in highly visible systems such as Google AdSense, Flikr, Napster, Wikipedia, blog spaces, forums and bulletin boards, search engine optimization, Web services, tagging (folksonomy), content syndication, and so on.[5] Some of the key characteristics of Web 2.0 applications are

- architecture of participation and wisdom of the crowds;
- Web services and service-based architectures;
- cost-effective scalability, grid computing, and cloud computing;
- Ajax, JSON (JavaScript Object Notation), CSS, RSS, and dynamic content delivery;

1541-1672/09/$26.00 © 2009 IEEE
Published by the IEEE Computer Society

- remixable data sources, Web APIs, mashups, and the programmable Web;
- open source development, "some rights reserved," "copyleft"; and
- collective intelligence and user-generated content.

As explained by O'Reilly, the value of the software and application is only proportional to the scale and dynamism of the data it helps to manage. The Web service automatically gets better the more people use it—tapping into "the wisdom of the crowds." The network effects from user participation are the key to market dominance in the Web 2.0 era. By leveraging customer-generated content and algorithmic data management, we can reach out to the entire Web, to the edges and not just the center; to the long tail of the user, customer, and citizen curve, and not just the head. From the business to the government, and from Internet marketing to public health, Web 2.0 offers a tremendous transformational opportunity for many aspects of modern society.

Figure 1. My.BarakObama.com Web site, which serves as a one-stop shop for announcements, multimedia content, volunteer information, donations, and so on.

E-government, Politics 2.0, and the My.BarakObama.com Campaign

The US Government's painstaking process of "going electronic" is a good illustration of some of the unique challenges and issues facing government.[4] Legislation and regulations concerning information technologies and the Internet in particular were developed only recently. For example, the 1996 Information Technology Management Reform Act established the US Government's CIO position to manage IT resources. After seeing many successful Internet applications in the business sector, in August 1998 the US Government announced the WebGov portal project, which aimed to provide one-stop information dissemination for the government. The project failed and was replaced in 2000 by the FirstGov portal. In 2001, Congress passed the Health Insurance Portability and Accountability Act (HIPAA). This legislation, which requires all healthcare information to comply with privacy regulations, continues to have a significant impact on the US healthcare and insurance industries.

In addition to e-government modernization, the political campaign and citizen participation process has also undergone a major transformation recently. The popularity of the Web 2.0 media has created significant progress in politics 2.0. According to the Webby Awards' "Top Ten Web Political Moments" (www.webbyawards.com/press/top10political.php), the first political candidate Web site was established in 1994. In 1998, the Drudge Report broke the Lewinsky scandal, and the first Internet voting took place in 2000. Fast-forward: CNN hosted the first YouTube debates in 2007, and in 2008 Barack Obama launched a highly mobile and successful Internet-based presidential campaign.

Started with little support and funds, the young and relatively unknown Illinois Senator Barack Obama began his Internet campaign with a vision of leveraging the power of Web 2.0 to inspire for change, especially among Democrats and the young. As stated on the My.BarackObama.com Web site now (Figure 1), "When you create an account on My.BarackObama.com, you're joining the online community of organizers who helped elect the President and now are working to bring real change on critical issues, including healthcare, education and energy reform. Join millions of Americans calling for change using our online tools [to]: find an event near you, join a local organizing group, and get trained on community organizing." With the highly interactive and carefully designed Web site, users can create individualized accounts, watch candidate YouTube videos, read political blogs, help make campaign phone calls, participate in political events, join online and offline groups, send and receive campaign-related text messages, and make donations. As a result, Obama was able to raise US$45 million by February 2008. Many politicians now consider the Internet an indispensible platform for their future campaigns and for engaging in dialogues with their constituents.[1-3]

Trends . . .

Several trends have become more prevalent for e-government and politics 2.0.

Multiway Political Dialogues

Unlike the traditional one-way mass-media campaign (for example, via TV and radio), Web 2.0 supports rich interactions between politicians and the masses, and also within the masses. Instead of relying on traditional speeches, advertisements, town-hall or TV debates, newspaper editorials, and face-to-face campaigners, e-government and politics 2.0 can more effectively reach larger and more interactive and engaged audiences through webcasts, blogs, political forums, candidate videos, social bookmarking, and online debates and Q&A. However, computational analysis of such rich and diverse content will require advanced Web, text, and opinion mining techniques.

Viral Marketing and Narrowcasting

Instead of relying on the traditional mass broadcasting approach, Web 2.0 offers a unique opportunity for delivering content that is highly customized based on a Web citizen's interest and political leaning (that is, "narrowcasting"). Personalization techniques and recommendation systems have already shown significant results in Web 2.0 business applications for products such as books, video rentals, and movies. Through advanced social-network analysis, viral marketing methodology can be leveraged to conduct an effective political campaign through the word-of-mouth and network effect.

Rich Multimedia Content

The YouTube/CNN debates are a great example of the power of the multimedia social-networking sites. Through online channels, candidates can create cost-effective and timely campaign videos, engage in online debates, solicit citizen-submitted questions, and cultivate different online social groups. Web 2.0 service-based, lightweight technologies such as Ajax, JSON, CSS, and RSS make delivering rich political content possible. Advanced image and video analysis techniques will enable analysis of such multimedia content.

Online Fundraising

Web 2.0 technologies have also lowered the barriers to entry by allowing candidates to reach the masses and solicit contributions through "microdonations." The long tail of the Web 2.0 user curve is ideal for effective online political fundraising. Instead of targeting the affluent in a traditional campaign banquet format of $500 or $1,000 per plate, microdonations from the online masses on the scale of $10 and $20 per person (and with little transaction cost) quickly add up to a major campaign monetary pool.

From the Real World to the Virtual Worlds

Rich 3D virtual environments that can support realistic, interactive, avatar-based navigation, interaction, training, and collaboration have begun to draw attention from Web 2.0 researchers. Many perceive the virtual worlds as a likely "Web 3.0" future, which could have a significant impact on future e-government and politics 2.0 development. 3D worlds such as Second Life have already been used by politicians (both Obama and Hillary Clinton had presence in Second Life during their 2008 campaigns) and adopted by government and military organizations for customer service, product demonstration, military training, and healthcare education.[6] It remains an open question whether 3D worlds will be an effective platform for delivering government services or facilitating citizen participation.

. . . and Controversies

Despite these positive trends, several controversies present unique challenges for e-government and politics 2.0 researchers.

The Great Divide

Instead of acting as the "great equalizer," the Internet has oftentimes instead caused a "great divide," separating users from nonusers on the basis of income, social class, age, geography, access, race, and so on. Web 2.0 participation has clearly favored the young, affluent, and IT-savvy citizens. How to continue to lower the barriers to entry for citizen participation is still a major challenge.

Copyright

Copyrights to user-generated content continues to be controversial. Sites such as Facebook, TurnItIn.com, YouTube, and Digg have all faced legal threats or action regarding the copyright of content uploaded by users. For e-government and politics 2.0 sites, it is unclear how citizen-generated content at those sites can be used for other political or commercial purposes.

Security and Privacy

Another major obstacle for online political participation and citizen engagement relates to cybersecurity and privacy. How can the government ensure the proper usage and control of citizen-generated contents? How can the government safeguard online transaction and information sharing among citizens? As the Internet becomes more pervasive, the inherent danger of information leaks, information misuse, corruption and manipulation, and system intrusion has become more imminent. Careful cybersecurity and information assurance research must address these issues.

This issue's Trends and Controversies department includes five essays on e-government and politics 2.0 from distinguished experts. Each essay presents a unique, innovative research framework, computational methods, and selected results and examples.

In "Blogosphere Research: A Mixed-Methods Approach to Rapidly Changing Systems," David Karpf discusses the growth of the political blogosphere and presents a few lessons that may be applicable to IT and political research. Karpf suggests that the rapidly changing nature of information technologies and the sheer abundance of available data are two of the most pressing challenges for researchers. In "OntoCop: Constructing Ontologies for Public Comments," Hui Yang and Jamie Callan present a system, designed to help organize online public comments from citizens, that works interactively with a person to organize a set of concepts into an ontology. Experimental results show that interactive learning produces useful ontologies and saves time and human effort. In "Enabling the Dialogue—Scientist<>ResourceManager<>Stakeholder: Visual Analytics as Boundary Objects," Judith B. Cushing and colleagues report on scientists' use of visual analytics to repurpose ecology research for government resource managers. The authors believe that visual analytics can act as effective boundary objects to communicate and improve understanding when working across linguistic, conceptual, disciplinary, or interest group boundaries. In "Moving Toward 'Intelligent' Policy Development?" Ann Macintosh presents selected augmentation support tools for assisting effective policy development. Social network analysis, argumentation mining, argumentation analysis, and argumentation visualization are some of the promising areas of research she suggests. In the fifth and last essay, "E-Government 2.0 in Asia: Trends, Opportunities, and Challenges," Paul Jen-Hwa Hu and colleagues describe representative e-government projects in Asia and present several more in-depth case studies of e-government development in Taiwan. The authors suggest knowledge mapping, social media scanning, and social network analysis as promising research directions.

Acknowledgments

I greatly appreciate some of the materials presented and discussions that occurred in my recent Knowledge Management class at the University of Arizona. Special thanks to students Adam Aronoff, Ben Gray, Preston Troy, Pooja Bhandari, and Clayton Badeaux. I also acknowledge the previous funding support provided by the NSF Digital Government program for my COPLINK research ("COPLINK Center: Information and Knowledge Management for Law Enforcement," EAI-9983304, 2003–2008).

References

1. S.L. Stirland, "Propelled by Internet, Barack Obama Wins Presidency," *Wired,* 4 Nov. 2008; http://blog.wired.com/27bstroke6/2008/11/propelled-by-in.html.
2. N. Scola, "Obama and Politics 2.0," TechPresident, Personal Democracy Forum, 17 July 2008; http://techpresident.com/blog-entry/obama-and-politics-20-documenting-history-real-time.
3. J. Green, "The Amazing Money Machine," *The Atlantic,* vol. 301, no. 5, June 2008; www.theatlantic.com/doc/200806/obama-finance.
4. H. Chen et al., eds., *Digital Government: E-Government Research, Case Studies, and Implementation,* Springer, 2007.
5. T. O'Reilly, "What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software," presented at O'Reilly Media Web 2.0 Conf., 2004; http://oreilly.com/web2/archive/what-is-web-20.html.
6. C. O'Hara, "Virtual Learning Gets Second Wind from Second Life," *Federal Computer Week,* vol. 23, no. 10, 4 May 2009, pp. 26–31.

Hsinchun Chen is the McClelland Professor of Management Information Systems at the University of Arizona and director of the Artificial Intelligence Lab. He received his PhD in information systems from New York University. Contact him at hchen@eller.arizona.edu.

Blogosphere Research: A Mixed-Methods Approach to Rapidly Changing Systems

David Karpf, *Brown University*

"Politics 2.0" can be understood as the harnessing of the Internet's lowered transaction costs and condition of information abundance toward the goal of building more participatory, interactive political institutions. As with any such term, the definition's boundaries are fuzzy, and the topic lends itself both to technologically deterministic prognostication and hastily constructed rebuttal. Among the various elements of politics 2.0, the political blogosphere has attracted the greatest early scrutiny. This essay reviews political science research on the blogosphere, noting in particular how the medium itself has continued to change and evolve, undermining the assumptions in our research designs. Only through a combination of quantitative and qualitative methods have we been able to accurately depict the size, scope, and usage of political blogs in American politics today. A few lessons from the brief history of blog research may be applicable to the study of technology and politics 2.0 more generally.

Political Blogging, Take 1

The first generation of blog research emerged in 2003 and 2004. Although blog software had been widely available since Pyra Labs introduced its Blogger platform in 1999, blogging didn't receive attention as a significant tool for political engagement until 2002, when political bloggers played a key role in pressuring US Senator Trent Lott to step down from his position as Majority Leader (after making racially controversial statements at the birthday party of retiring Senator Strom Thurmond). Early bloggers were a mix of "citizen journalists" and amateur political strategists, relying upon the near-costless publishing platform to express their ideas and opinions, and creating nascent communities through blogrolls and within-text hyperlinks. Political bloggers were also influential in supporting Howard Dean's Presidential candidacy and in exposing forged documents presented by Dan Rather on *60 Minutes*.

Early academic research on blogging focused on the potential of the low-cost tool for radical increases in mass participation and as a challenge to elite media operations. Anecdotal examples of blogger effectiveness, coupled with empirical evidence of rapid blogosphere growth, suggested that the new medium could give every motivated individual the ability to reach millions and influence public policy.

By those measures, blogging proved to have a rather disappointing track record. As Matthew Hindman demonstrates in his 2008 book,[1] political blogs have attracted far less attention than humor and entertainment blogs, and the maturation of the blogosphere has featured the growth of a limited number of "hub" sites that attract exponentially more hyperlinks and site visits than the average blog. Blog software may give anyone a megaphone, but with the information abundance of the Internet, only the elite few can attract a large audience. Given that the lion's share of top political bloggers held advanced graduate degrees, the development of a new Internet elite seemed to mirror the same demographic disparities prevalent in offline society. Meanwhile, though blogs provided an alternate venue to the mainstream media, major media outlets began hiring bloggers and incorporating blogs into their set of offerings. Although newspapers are clearly experiencing financial distress today, blogs are hardly the main culprit. Rather, it was Craigslist that undermined the market for classified ads, and Google that altered how we search for and obtain new information. Blogs have served as an alternate venue for political reporting and engagement, but not one that necessarily replaces or challenges existing elite structures.

Blog software may give anyone a megaphone, but with the information abundance of the Internet, only the elite few can attract a large audience.

Political Blogging, Take 2

If the blogosphere failed to live up to the hopes of first-generation proponents and researchers, however, that isn't to say that it had no substantive impacts. Though the introduction of online self-publishing tools failed to radically recast elite political institutions, the blogosphere has nonetheless yielded a new set of elites and, arguably, increasingly porous networks of influence within those institutions. The introduction of these new technologies of political engagement has yielded opportunities for the formation of novel political associations and a host of expanded tactical repertoires that change how politically active citizens engage in politics. Indeed, one of the chief problems the initial generation of Internet researchers faced was that, as the blogosphere expanded, it was adopted in new and unexpected manners that didn't fit within the boundaries of our research program.

Consider DailyKos.com, the largest political blogging community in America today. Daily Kos was initially the personal blogging home of Markos "Kos" Moulitsas, an outspoken left-wing Democrat. In October 2003, Moulitsas adopted a new "community blogging" software platform operated by "Scoop." The community blogging platform allowed all registered users to post their own blog entries as "diaries" on the site. The site quickly soared in popularity as progressive bloggers settled upon it as a hub space for their community of interest. As several longtime bloggers revealed in interviews, “the difference between Kos and everyone else wasn’t that he was so much more talented. It was that he adopted the community platform first.”

The introduction of community blogging software is one dimension of the blogosphere’s shifting terrain. These platforms enable blogs to function as coordination points for online communities of interest, the largest of them functionally indistinguishable from political interest groups. The Daily Kos community establishes campaign priorities, endorses political candidates, fundraises and volunteers for them, and even holds an annual in-person convention. While the first generation of Internet researchers were busy exploring the degree to which individual elite bloggers were distinct from other political elites, the bloggers themselves were refashioning their sites to provide greater voice and mobility to their communities.

A second dimension of innovation involves the adoption of blogging platforms into the Web offerings of existing institutions. While early proponents of the blogosphere envisioned the small core of counter-institutional bloggers expanding ever outward and reshaping American politics, the actual diffusion process looked quite different. Once blogging gained enough notoriety to be taken seriously, existing media and political institutions adopted the technology themselves. Major media institutions hired bloggers to work full time for them, offering content on their sites. Political campaigns, businesses, and interest groups all added blogs to their own sites, and new content management systems made blogging a basic feature of Web site redesign. These “institutional blogs,” of course, adapt the technology to their own needs and goals. As technology writer Clay Shirky predicted in 2003, “At some point (probably one we’ve already passed), weblog technology will be seen as a platform for so many forms of publishing, filtering, aggregation, and syndication that blogging will stop referring to any particularly coherent activity.”[2] The central challenge for the research community, then, lies in accurately measuring and describing a phenomenon that is itself rapidly changing in unpredictable ways. I treat these two types of blog innovation as dimensions in a typology of blogspace, allowing for categorization and comparison of differing blog formats.[3]

Table 1. Blogosphere authority index overview.

Score	Definition
NCS	The *network centrality score* gathers applied sociometric data based on appearance in progressive blogrolls.
HAS	The *hyperlink authority score* rates a site’s authority among political bloggers using link patterns, as measured by Technorati.
STS	The *site traffic score* rates visits-per-day statistics as measured by a combination of Sitemeter and Alexa traffic rankings.
CAS	The *community activity score* rates interactive participation as measured by total blog comments.
BAI	The final *blogosphere authority index* score combines the raw NCS, HAS, STS, and CAS scores. The three best scores are added, and the fourth is dropped; the ranking of the cumulative scores gives the index.

An additional challenge in this regard comes with the issue of how we measure influence in the blogosphere. Hyperlink analysis and site traffic are the two most common measures, but each has its particular flaws. Hyperlinks give an accurate map of clustering and communities within the larger blogosphere, but the rise of “splogs” (spam blogs) muddies the waters. Also, as community sites like Daily Kos become more self-referential, a number of the outbound hyperlinks go to the traditional news organizations that employ full-time journalists. The relationship between links and site traffic heavily fluctuates. Site traffic, meanwhile, is notoriously difficult to measure, and the best data is kept behind proprietary firewalls. In my own research, I convert four independent measures of blog influence (network centrality, hyperlinks, site traffic, and total comment volume) into ordinal rankings, then merge those rankings to form a composite ranking system for the political blogosphere. Table 1 describes the individual measures I use. Critically, this methodology is the only currently available that produces rankings without relying solely on a single, flawed metric.[3]

Combining the 2D blogspace map with the composite rankings of the elite blogosphere yields several key findings about the partisan makeup of the American political blogosphere. The community blogging platform has been heavily used by progressive bloggers (to great financial benefit for their favored political candidates), while conservative attempts to replicate Daily Kos and the other large community sites have largely foundered. The most successful conservative bloggers operate closed-authorship individual or institutional blogs, limiting their capacity for collective action in the blogosphere. As blog traffic increased in the 2008 election season, the progressive blog network substantially increased its advantage over its conservative counterpart.[4] These findings only emerge when detailed qualitative analysis techniques are combined with large-scale data collection. The research literature ignored variability in blog type for years because qualitative insights on the use of blogs were left apart from quantitative research design choices.

Lessons for Future Researchers

The two greatest challenges for researchers interested in e-government and politics 2.0 are the rapidly changing nature of the technologies under study and the sheer abundance of available data, much of which is of questionable quality. Link-based analysis in particular has become highly popular recently, as complex data visualization packages claim to offer "maps" of Web communities and their most influential members. Although there is some substantial value in these tools, we should recall the longstanding slogan, "garbage in, garbage out": a data analysis program can never far exceed the quality of its inputs. And in particular, we lack an empirically based understanding of what a link actually signifies. Increased traffic? More or less. Ideological homophily? Occasionally. Avenues for diffusion? Possibly. What's more, it's entirely possible that a link between sites at time x, where x occurs in the lead-adopter phase of diffusion, will mean something quite different from a link between sites at time $x + 5$, where $x + 5$ occurs during the early-majority phase of adoption.

This points to the importance of the diffusion of innovation literature[5] for the research community. As with the blogosphere, it seems evident that most Web-based political tools are first used by a small core of highly tech-savvy "lead adopters" who cocreate and alter the medium. This group is demographically and ideologically distinct from the "early adopter" and "early majority" classes, each of which is a good deal larger than the initial class. As a technology scales up, it adapts to the interests of these larger segments of the populace, and in turn is modified and adopted by existing institutions that show little interest in the technology until it has moved beyond the tiny core of lead adopters. That these changes in the nature and scope of a social technology will occur is both predictable and model-able, but the resultant direction of such changes requires close observation of the emerging large-scale communities.

The condition of online data abundance is both a blessing and a curse to the Web sciences research community. Torrential data is available for analysis; what it signifies is far less clear. Large-scale data analysis must be wedded to close observation of the changing boundaries of the phenomena, communities, and platforms under study, so that we may remain attuned to changes in usage patterns, form realistic hypotheses, and separate high-quality data from the noise.

References

1. M. Hindman, *The Myth of Digital Democracy,* Princeton Univ. Press, 2008.
2. C. Shirky, "Power Laws, Weblogs, and Inequality," 2003; www.shirky.com/writings/powerlaw_weblog.html.
3. D. Karpf, "Understanding Blogspace," *J. Information Technology and Politics,* vol. 5, no. 4, 2008, pp. 369–385.
4. D. Karpf, "Stability and Change in the Political Blogosphere in the 2008 Election," paper presentation, Midwest Political Science Assoc. Ann. Meeting, Chicago, 2009.
5. E. Rogers, *Diffusion of Innovation,* 5th ed., Free Press, 2003.

David Karpf is a postdoctoral research associate at the Taubman Center for Public Policy at Brown University. He holds a PhD in political science from the University of Pennyslvania. Contact him at davekarpf@gmail.com.

OntoCop: Constructing Ontologies for Public Comments

Hui Yang and Jamie Callan, *Carnegie Mellon University*

US law defines a process known as *Notice and Comment Rulemaking* that requires regulatory agencies to seek comment from the public before establishing new regulations. Regulatory agencies are also expected to demonstrate that the final regulation addresses all substantive issues raised by the public. Most proposed regulations attract few comments from the public, but each year a few regulations attract tens of thousands or hundreds of thousands of comments. When comment volume is high, most comments are form letters and modified form letters, which are not difficult to process,[1] but there are still tens of thousands of unique comments that address a variety of issues. There may also be political or public pressure for the agency to issue regulations quickly. In these cases, regulatory agencies and other interested parties need tools that help them to quickly make sense of public comments.

Browsing hierarchies such as the Yahoo Directory are a popular method of quickly discovering the "lay of the land" in a large text corpus. By associating documents with topics and concepts in a hierarchy, we can structure and partition the information space into smaller spaces that are easy to understand and navigate. Regulation-specific browsing hierarchies make it easier to understand the range of issues that the public has raised; they also let agencies and policymakers drill down into comments that discuss particular issues, which lets them be more responsive to the public's concerns. Information analysis tools that allow quick and efficient analysis of comments also increase the likelihood that agency staff will process public comments within the agency, rather than subcontracting the job to outside companies that have less topic expertise.

There is much literature on automatic formation of concept hierarchies (also called *taxonomies* and *ontologies*), most of it focused on the creation of general-purpose concept hierarchies. Most prior research exploited simple lexical-syntactic patterns[2] or contextual information,[3] combined with a lexical resource such as WordNet.[4] However, when the corpus is small and focused on a narrow set of issues, important concepts and relations may be rare or missing in general-purpose lexical resources or corpora. For example, in public comments submitted about a proposed Environmental Protection Agency regulation (USEPA-TRI-2005-0073), RHOM is a chemical company; however, this relation is neither available in WordNet, nor is it part of a valid hypernym pattern in the full set of public comments. Cases such as this require an approach that combines multiple technologies.

OntoCop is new software that works interactively with a person to organize a set of concepts into an ontology. Ontology construction consists of two subtasks: *concept extraction* and *relation formation*. OntoCop takes a conventional approach to concept extraction and a more novel approach to relation formation.

Concept Extraction

Concept extraction identifies the concepts mentioned in a corpus. Concepts are nouns or noun phrases. Concept extraction first extracts all possible concept candidates, such as nominal unigrams, bigrams, trigrams, and named entities (NEs), using part-of-speech (POS) tagging results. Since POS taggers make mistakes, these initial concept candidates contain false-positive noun phrases. For example, a POS tagger incorrectly tagged "protect polar bear" as three nouns, and hence incorrectly considered it a nominal trigram. A Web-based frequency test filters out false-positive concept candidates. OntoCop sends queries formed by each concept candidate to the Google search engine. Among the first 10 returned snippets, if a concept candidate appears more than a threshold number of times (set to four), OntoCop considers it a commonly used phrase, and hence a good concept candidate; otherwise, OntoCop discards it.

Relation Formation

Relation formation discovers relations among concepts and builds ontologies based on these relations. Given a set of concepts, OntoCop clusters the concepts and presents an initial ontology. The human then teaches OntoCop by providing manual guidance. OntoCop learns from this manual guidance and adjusts the clustering process accordingly to modify the ontology. Teaching and learning alternate until the human is satisfied.

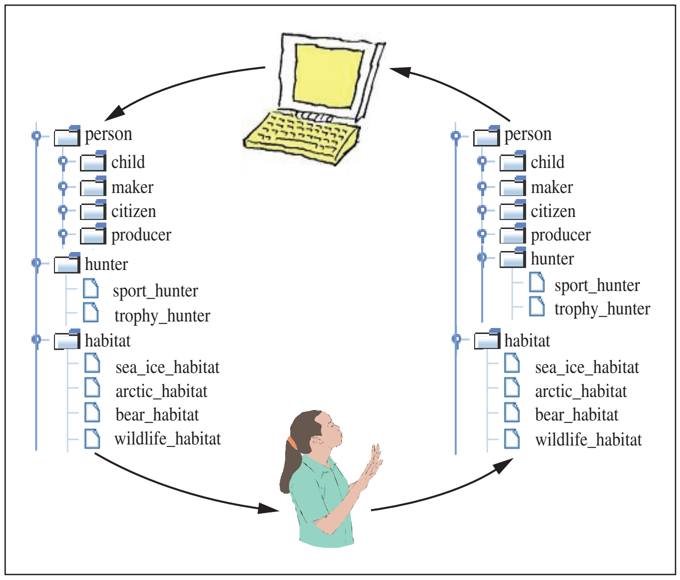

Figure 2. OntoCop human-computer interaction cycle (Polar Bear comment set). OntoCop presents an initial ontology, the human modifies it, OntoCop adjusts its algorithm based on the change and presents an updated ontology to the human. This cycle continues until the user is satisfied with the ontology.

Figure 2 illustrates the human-computer interaction cycle for a fragment of the ontology created for the *Polar Bear* comment set. The cycle starts when OntoCop presents an initial ontology that consists of three concept groups: *person, hunter,* and *habitat*. The human makes a modification by dragging and dropping the *hunter* group to be under the *person* group, which makes *hunter* a child concept of *person*. OntoCop recognizes the change, adjusts the clustering algorithm, and shows an improved ontology to the human. The human-computer interaction cycle continues until the human is satisfied with the ontology.

The initial ontology is based on head-noun matching, WordNet hypernyms,[4] and lexico-syntactic patterns.[5]

If two bigrams share the same head noun, OntoCop assigns the two bigrams to the same group; it elects their head noun as the parent of the group. For example, *pollution* becomes the parent of *water pollution* and *air pollution*. The parent concept and the child concepts form an ontology fragment. OntoCop also applies head-noun matching to unigrams, trigrams, and named entities. The head-noun matching technique effectively creates the initial ontology fragments.

OntoCop uses WordNet hypernyms to identify parent-child relations among sibling concepts within an ontology fragment. Given two sibling concepts x and y, if x is y's hypernym in WordNet, OntoCop promotes x as the parent of y in their ontology fragment. OntoCop can also apply refinement by WordNet hypernyms among ontology fragments. OntoCop examines all roots of the ontology fragments in WordNet, and connects them as one fragment if they are in the same WordNet hypernym chain.

OntoCop also uses lexico-syntactic patterns to identify parent-child relations for the roots of the ontology fragments. It puts each pair of roots into the patterns to form a text segment. For example, OntoCop puts *heavy metals* and *toxins* into the pattern "NP_B and other NP_A" to form "heavy metals and other toxins." OntoCop then searches for the text segment in the corpus that supplied the concepts. If OntoCop finds the text segment in the corpus, then the pair of roots has a parent-child relation and is connected.

Head-noun matching, WordNet hypernyms, and lexico-syntactic patterns create an initial forest of ontology fragments, which is the initial version of the ontology. OntoCop then waits for human guidance.

During each human-computer interaction, the human modifies the ontology. OntoCop represents the grouping of concepts before this set of modifications as a *before matrix*; likewise, it represents the new grouping of the concepts as an *after matrix*. Formally, an ontology with n concepts can be represented by an $n \times n$ *ontology matrix*. The (i, j)th entry of an ontology matrix indicates whether c_i, the concept at the ith row, is a sibling of c_j, the concept at the jth column. The value of the (i, j)th entry $v_{ij} = 1$ if $i = j$ or c_i is a sibling of c_j; 0, otherwise. The manual guidance **M** is a submatrix that consists of some entries of the after matrix; at these entries, there exists a difference between the before matrix and the after matrix. Formally, if **A** is the before matrix and **B** is the after matrix, then $\mathbf{M} = \mathbf{B}[r; c]$, where $r = \{i : b_{ij} - a_{ij} \neq 0\}$, $c = \{j : b_{ij} - a_{ij} \neq 0\}$, a_{ij} is the (i, j)th entry in **A**, and b_{ij} is the (i, j)th entry in **B**. For example, we obtain the following manual guidance by dragging and dropping concept *hunter* to be under the concept *person*:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

=	*Person*	*Hunter*	*Producer*
Person	1	0	0
Hunter	0	1	1
Producer	0	1	1

OntoCop modifies the ontology once at each human-computer interaction by using the manual guidance **M** as training data. A supervised distance-learning algorithm learns a distance function between concepts in **M**. The distance function used here is a Mahalanobis distance, which measures the correlation between two concepts by assigning adaptive weights to different underlying feature functions.[6] The feature functions include contextual features, co-occurrence, syntactic dependency, lexical-syntactic patterns, word length difference, and definition overlap. These heterogonous features evaluate the semantic relation between two concepts from different aspects and aim to capture a wide range of characteristics of semantic relations.

The training data consists of a set of concepts $\mathbf{x}^{(i)}$, the set of concepts in $\mathbf{M}^{(i)}$, the manual guidance obtained at the ith iteration of human-computer interaction, and its corresponding distance matrix $\mathbf{y}^{(i)}$. The entry of $\mathbf{y}^{(i)}$ that corresponds to concept $x_j^{(i)}$ and $x_k^{(i)}$ is $y_{jk}^{(i)} \in \{0,1\}$, where $y_{jk}^{(i)} = 0$, if $x_j^{(i)}$ and $x_k^{(i)}$ are in the same cluster; 1, otherwise. We use mean squared error as the loss function and define the optimization function as

$$\min_S \sum_{j=1}^{|\mathbf{x}^{(i)}|} \sum_{k=1}^{|\mathbf{x}^{(i)}|} \times \left(y_{jk}^{(i)} - \sqrt{\Phi\left(x_j^{(i)}, x_k^{(i)}\right)^T \mathbf{S}^{-1} \Phi\left(x_j^{(i)}, x_k^{(i)}\right)} \right)^2,$$
$$\text{subject to } \mathbf{S} \geq 0$$

where $\Phi(x_j^{(i)}, x_k^{(i)})$ represents a set of pairwise underlying feature functions; **S** is the parameter matrix, which weighs the feature functions.

Given the learned parameter matrix **S**, OntoCop then applies the newly learned distance function to obtain distance scores for the ungrouped concepts in the ontology. The *K*-medoids clustering algorithm further clusters the ungrouped concepts on the basis of these scores and produces the next version of the ontology. The ungrouped concepts are the top-level concepts in the existing ontology. OntoCop clusters concepts at one level during each iteration, building the entire concept hierarchy in a bottom-up manner. OntoCop then presents the modified ontology to the human and waits for the next round of manual guidance.

Through concept extraction and relation formation, OntoCop builds ontologies interactively according to manual guidance. Figure 3 shows part of an ontology for the *Mercury* comments, and how to use it for navigation and drilldown. A policymaker viewing the ontology (Figure 3, left pane) quickly sees the issues raised by the public. She may be familiar with many of these issues, but perhaps "asthma attack" is unexpected. A click on "asthma attack" leads to a list of comments that discuss this topic (middle pane). A click on any comment displays its text with the concept conveniently highlighted (right pane). This combination of navigation and drilldown helps a policymaker quickly understand what

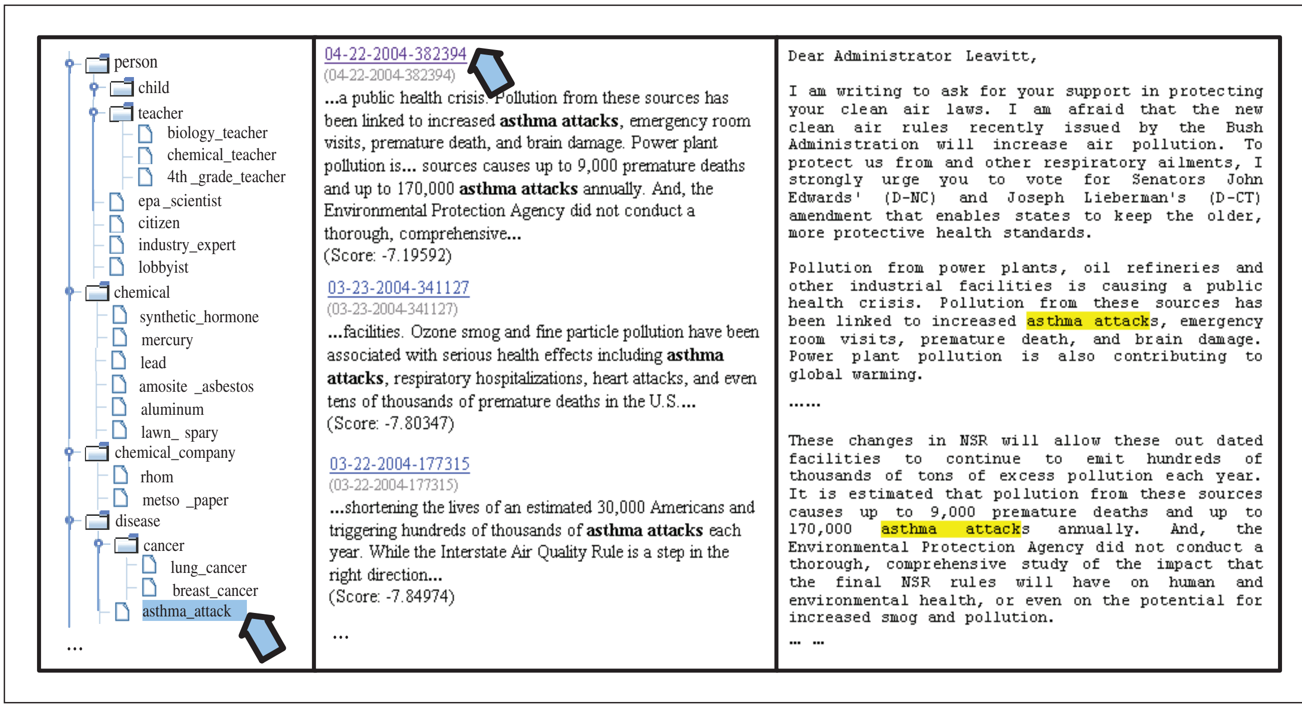

Figure 3. Using an ontology for drilldown (Mercury comment set). A policymaker viewing the ontology (left pane) quickly sees the issues raised in the comment set. Her click on a concept in the ontology leads to a list of comments that discuss this topic (middle pane). Another click on one of these comments displays the comment's content (right pane).

issues the public raised and what was said about each issue, thus improving the responsiveness of regulatory agencies during the rulemaking process.

Experiments

We collaborated with the Qualitative Data Analysis Program (QDAP) at the University of Pittsburgh's University Center for Social and Urban Research (UCSUR) to conduct a user evaluation. Participating in the experiment were 12 professional coders familiar with the problem domain, divided into two groups: four in the manual group and eight in the interactive group. We asked both groups to construct an ontology with the same concept candidates until they felt satisfied with their work or reached a 90-minute limit. These experiments used three public comment data sets: Wolf (RIN-1018-AU53), Polar Bear (RIN-1018-AV19), and Mercury (OAR-2002-0056)—all available at http://erulemaking.cs.cmu.edu/data.php. Table 2 gives an overview of these three data sets.

Quality Comparison: Interactive vs. Manual Ontology Construction

Our quality comparison experiment investigated whether OntoCop can produce ontologies with the same quality as those built manually. We compared the intercoder agreement between two manual runs with that between one manual and one interactive run in this experiment, using Cohen's Kappa statistic to assess the intercoder agreement. Table 3 shows the averaged intercoder agreement for parent-child pairs in three public

Table 2. Overview of three public comment data sets.

Data set	No. of comments	No. of unique comments	Duration of comment period	Topic
Wolf	282,992	59,109	February 2007–August 2007	Delisting the northern Rocky Mountain population of gray wolves from the federal list of endangered and threatened wildlife.
Polar Bear	624,947	73,989	January 2007–October 2007	Listing the polar bear as threatened throughout its range under the Act (72 FR 1064).
Mercury	536,975	104,146	February 2004–June 2004	Proposing national emission standards for hazardous air pollutants.

Table 3. Average intercoder agreement on parent-child pairs.

Data set	Manual-manual	Manual-interactive	t	p
Wolf	0.55	0.55	0	0.5
Polar Bear	0.44	0.46	0.21	0.42
Mercury	0.61	0.51	1.89	0.03

Table 4. Average manual editing costs.

Construction method	No. of edits performed					
	Add	Delete	Move	Rename	Undo	Total
Manual	57	200	2,807	71	19	3,153
Interactive	21	129	1,694	40	8	1,890

Table 5. Ontology construction duration.

Construction method	Ontology construction time (hours:minutes)			
	Wolf	Polar Bear	Mercury	Average
Manual	1:24	1:22	1:33	1:27
Interactive				
Human contribution	0:33	0:29	0:30	0:31
Computer contribution	0:33	0:05	0:35	0:24

comment data sets. Both the intercoder agreement between manually built ontologies and that between manual-interactive runs are within the range of 0.44 to 0.61, which indicates moderate agreement. We also observe that manual-interactive intercoder agreement is comparable with manual-manual intercoder agreement, which indicates that the guided machine-learning approach can produce the same quality ontologies as humans do. A series of one-tailed t-tests also confirmed it. The significant test results are in the range of $t < 2$ and $p > 0.01$, indicating no statistically significant differences between pairs of manually built ontologies and interactively built ontologies.

Cost Comparison: Interactive vs. Manual Ontology Construction

We compared users' construction logs from both the manual and interactive groups. Table 4 shows the number of manual edits performed during ontology construction for the three data sets. The edits include adding a concept, moving a concept by drag and drop, deleting a concept, changing a concept name, and undoing previous actions. In total, the interactive users performed 40 percent fewer editing actions to produce ontologies of the same quality. A one-tailed t-test shows a significant reduction, $t = 10$ and $p < 0.001$, in the interactive runs in terms of editing costs as compared to the manual runs. This demonstrates that OntoCop is significantly more cost-effective than the manual work.

Table 5 shows the actual time needed to construct an ontology for both manual and interactive runs. For the interactive runs, it also shows the amount of time the human spent and the amount of time the machine spent. In general, the interactive runs save 30 to 60 minutes for building one ontology. Within an interactive run, a human user needs to spend an average of only 31 minutes to construct an ontology, which is 64 percent less than the 1 hour and 27 minutes required for a manual run. Thus, OntoCop greatly saves the time a human user spends on ontology construction.

OntoCop helps a person construct an ontology that provides a way of understanding the lay of the land in a corpus and then drilling down into specific topic areas. Unlike most prior research on ontology construction, OntoCop is designed for corpora that have a very task-specific focus, such as public comments submitted to regulatory agencies during Notice and Comment Rulemaking. Such domains contain many similar concepts, rare concepts, and rare relationships. OntoCop relies on human guidance and a set of feature functions that give it a stronger ability to disambiguate the subtle differences among similar concepts. Experimental results on three corpora of public comments show that interactive learning not only produces useful, domain-specific ontologies, but also saves time and human effort, thus accelerating the process of developing task-specific ontologies.

Acknowledgments

This research was supported by NSF grant IIS-0704210. Any opinions, findings, conclusions, or recommendations expressed in this essay are of the authors, and do not necessarily reflect those of the sponsor.

References

1. H. Yang and J. Callan, "Near-Duplicate Detection for eRulemaking," *Proc. 15th Nat'l Conf. Digital Government Research* (dg.o 05), Digital Government Soc. of North America, 2005, pp. 78–86.
2. D. Ravichandran and E. Hovy, "Learning Surface Text Patterns for a Question Answering System," *Proc. 40th Ann. Meeting Assoc. for Computational Linguistics* (ACL 02), Assoc. for Computational Linguistics, 2002, pp. 41–47.

3. P. Pantel and D. Ravichandran, “Automatically Labeling Semantic Classes,” *Proc. Human Language Technology Conf./North American Chapter of the Assoc. for Computational Linguistics Ann. Meeting* (HLT/NAACL 04), Assoc. for Computational Linguistics, 2004, pp. 321–328.
4. C. Fellbaum, *WordNet: An Electronic Lexical Database,* MIT Press, 1998.
5. M.A. Hearst, “Automatic Acquisition of Hyponyms from Large Text Corpora,” *Proc. 14th Int’l Conf. Computational Linguistics* (COLING), Assoc. for Computational Linguistics, 1992, pp. 539–545.
6. P.C. Mahalanobis, “On the Generalised Distance in Statistics,” *Proc. Nat’l Inst. of Sciences of India,* vol. 2, no. 1, 1936, pp. 49–55.

Hui Yang is a PhD student at Carnegie Mellon University’s Language Technologies Institute. Contact her at huiyang@cs.cmu.edu.

Jamie Callan is a professor at Carnegie Mellon University’s Language Technologies Institute and Heinz School of Public Policy and Management. Contact him at callan@cs.cmu.edu.

Enabling the Dialogue—Scientist<>Resource-Manager<>Stakeholder: Visual Analytics as Boundary Objects

Judith B. Cushing, Natalie Kopytko, Nik Stevenson-Molnar, and Lee Zeman, *The Evergreen State College*
Susan Stafford, *University of Minnesota*
John Bolte, Barbara Bond, and Denise Lach, *Oregon State University*
Robert McKane, *Environmental Protection Agency*

One active area of e-government research is exploring how information technology can improve policy, including how science informs policy making. Once policy is made, however, management strategies must be evaluated, adopted, and implemented. Applying ecological research to policy implementation—that is, natural resource management—has received less attention than science-based policy making, even though both have similar challenges, including transparency, communication across disciplines, and uncertainty. Similarly, data reuse within the sciences is an active research area (see, for example, *Proceedings of the 6th Int'l Conf. Data Integration in the Life Sciences;* www.cs.manchester.ac.uk/DILS09), but little work has been done on reusing or repurposing scientific data for resource management, with some notable exceptions.[3,4] Using informatics techniques to help make research available to resource management is thus a relatively new e-government research area. A similarly new e-government research area is visual analytics. Though it is already actively pursued for Homeland Security applications, we see many fruitful opportunities for visual analytics in resource management.

We believe that increasing the dialogue among scientists, resource managers, and other stakeholders will improve the application of scientific research results to resource management; thus, we focus some of our own research on strategies for communicating scientific results to natural resource managers and the public. We report here our use of visual analytics to transform, transfer, and repurpose ecological research data for resource managers’ use in day-to-day resource decisions. Our own preliminary work suggests that visual analytics can help resource managers more effectively use and interpret scientific data and models, but that more work must be done to assess what kinds of visualizations and technologies best accomplish this goal, and which collaborative approaches work best for codeveloping visual analytic tools. Because of the importance of making sound decisions to manage public lands and then making information about those decisions freely available to all stakeholders, developing visual analytics to improve government decisions and communication is a critical e-government research area.

Intelligent systems research relevant to this work includes

1. developing methodology and software to identify, integrate, and mine disparate data and models, especially when data are incomplete, uncertain, or probabilistic, and when the data crosses spatial or temporal scales;
2. conducting fundamental cognitive science research on how people from different disciplines and cultures understand and interpret scientific results;
3. conducting studies of how ecologists, resource managers, computer scientists, and software engineers work together to develop a shared understanding of how research results apply to resource management; and
4. developing methodology and software for end-user teams to themselves integrate data and models to create visualizations that improve their own and others’ understanding of the results.

Our own work and this essay focus on areas 1, 3, and 4.

Our Approach: Visual Analytics as Boundary Objects to Communicate Scientific Results

The new science of visual analytics is defined as “analytical reasoning facilitated by interactive visual interfaces” used “to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data." (For more information, see *Research & Development Agenda for Visual Analytics,* http://nvac.pnl.gov/agenda.stm.) We use visual analytics as *boundary objects* to transform scientific data into forms more useful to resource managers and other end users. Boundary objects are communication mechanisms used to improve understanding when working across linguistic, conceptual, disciplinary, or interest-group boundaries.[7] The communication of ecological research results to resource managers and the public is a particularly challenging type of boundary work, given the diversity of stakeholders and alternative points of view on common areas of interest. Here, interpretive differences in what a word, measurement, or outcome means can limit effective communication of knowledge across boundaries.[1] Furthermore, different experiences and presumptions about what constitutes a persuasive argument often hinder mutual understanding between experts and decision makers. We reason that visual analytics can be used as boundary objects to effectively transform scientific data for use by natural resource managers.

A Case Study in Visual Analytics: Which Trees to Leave?

Our work is driven by the hypothesis that visual analytics can make complex research data more accessible to resource managers. As a first step in testing this hypothesis, we conducted a proof-of-concept project that used legacy research data to help resource managers make decisions about which trees to leave when harvesting timber stands. This work involved

- creating visualizations of scientific measurements of natural phenomena—thus transforming the research data for new end users;
- providing simple summary statistics as an adjunct to the visualizations;
- determining which terminology used in policy statements needed clarification for resource managers; and
- asking scientists to use the visualizations to more precisely define those terms.

We developed most of these visualizations (3D and 2D) using specialized software for visualizing ecology research data that we wrote in Python and Java, using the Visualization Toolkit (VTK, described at www.vtk.org); a sample of the visual analytics for this work is publicly available. (Our visual analytics for tree and forest structure is available at http://acdrupal.evergreen.edu/dnr, http://acdrupal.evergreen.edu/canopyview, and http://acdrupal.evergreen.edu/canopyview2d.) The project suggested that, where research involves real-world phenomena relevant to natural resource management, the visualization of those data along with summary statistics about the phenomena can effectively communicate research results across disciplines.[5]

> Interpretive differences in what a word, measurement, or outcome means can limit effective communication of knowledge across boundaries.

Our visualizations of raw research data enabled researchers and managers to make their own conclusions about the extent to which data collected for a specific study (how tree and forest structure varies in differently aged stands) could be applied to a new problem (how tree and forest structure can inform which trees provide wildlife value). The project addressed the challenges of transparency (how policy language is used in practice) and communication across boundaries (how researcher definitions of habitat can be made relevant to managers; how managers can communicate policy to field foresters who actually select "trees to leave"; and how managers might better explain to stakeholders how decisions were made about which trees, and how many, to leave). In the future, we will continue to work with resource managers on combining the new definitions with visualizations to help managers communicate policy to field foresters and stakeholders.

Using Visual Analytics to Scale up in Time and Space

Communicating and handling uncertainty in scientific data pose significant challenges in developing and implementing policy. For example, extrapolation of plot-level research across landscapes or regions can accumulate errors and thereby add uncertainty for scientific, resource management, and policy purposes. Uncertainty also surrounds questions of how ecosystems change through time. Because these types of uncertainty are extremely difficult to quantify, clearly communicating the complexities of uncertainty is a critical issue for long-term resource management as well as for policy formulation.[4]

We hypothesize that over the longer term, appropriate visual analytics can

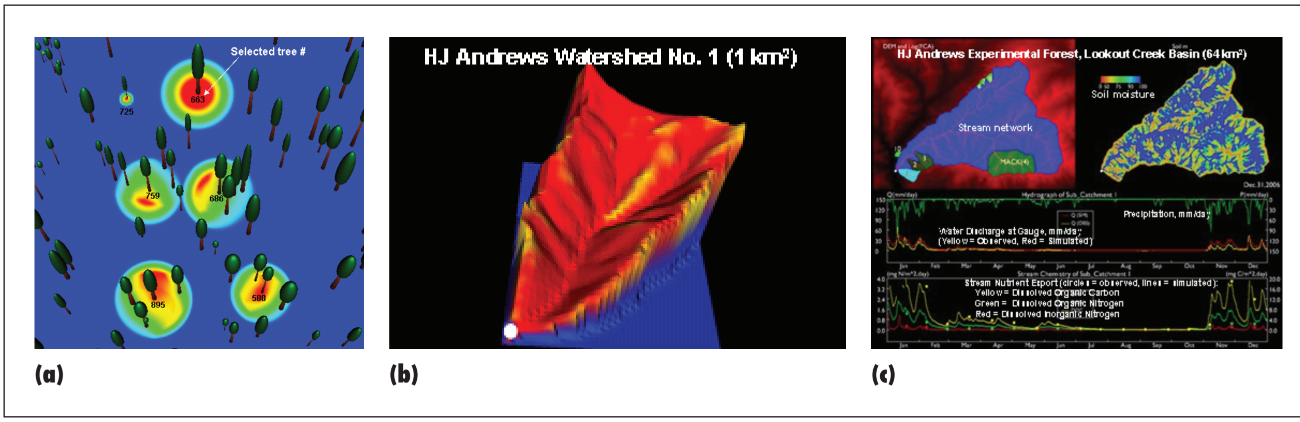

Figure 4. Our software helps scientists display and interpret very large data sets: (a) Effect of tree size and competitors on nitrogen uptake in a 400-year forest. Visualized spatial patterns not evident in the raw data provide new insight into forest habitat structure. (b) Soil moisture patterns for a headwater catchment visualized on 3D topographic data help scientists calibrate simulation models used for scaling up experimental data in space and time. (c) Visualizing ecosystem modeled data at the basin scale helps scientists and land managers understand and communicate effects of climate change and forest harvest: stream network, soil moisture, and stream water quality and quantity. Visualizations scale up from (a) 0.1-km² forest stand, to (b) 1-km² catchment, to (c) a 64-km² basin. Images in (b) and (c) can be animated to show change over time.

help address uncertainty issues, especially as they apply to scaling up or down in space and time, by allowing the display and interpretation of very large data sets, including modeled data. We also hypothesize that scientists can use visual analytics to improve their ability to parameterize models and to explain their work to other scientists and resource managers. We also think better visualizations will eventually help resource managers not only to make decisions themselves, but to work with the public on management decisions.

To test these hypotheses, and to determine what kinds of visual analytics are effective and how to codevelop them, we have launched a new project to develop visual analytics that complement and enhance resource managers' use of new and existing research data. The project involves data from the National Science Foundation's Long-Term Ecological Research (LTER) site at H.J. Andrews Experimental Forest (HJA) and from environmental science models. The HJA is situated on the western slope of the Cascade Mountains in Oregon and is a center for stream and forest ecology research. It is one of 26 sites in the LTER program funded by the National Science Foundation. The environmental science models we used were GTHM-PSM (http://adsabs.harvard.edu/abs/2006AGUFM.B22E..05S) and Envision (http://envision.bioe.orst.edu).

Figure 4 shows how our visualization software accommodates the rather large data sets generated by models such as the GTHM-PSM,[6] overlaying those data sets with field and sensor data, and creating visualizations at increasing spatial scales. This ability to scale up also implies that a user could zoom in from a larger spatial scale to examine data at lower scale. To illustrate time, we are writing software that creates animations. The visualization work to help scientists themselves better understand their data and models and better communicate with each other is well under way. A key intelligent systems topic of interest is indexing and caching the large data sets for real-time display. These visualizations, and the software that scientists use to produce them, have involved a close working relationship between developers and end users.

Because our ultimate research goal is to extend our tools for resource managers who work with the public to articulate and examine scenarios for natural resource management, we are collaborating with scientists at HJA working to scale up ecosystem services models to make projections at the regional level for alternative management scenarios. The computational challenges involved in accomplishing this are considerable, but we contend that equally formidable are challenges involved with interpreting and explaining the model results to a wide audience. Figure 5b shows how Envision, a decision support tool that integrates stressor scenarios, decision rules, ecological models, and evaluation indices within a geographical information system (GIS) framework, currently shows results of alternative land use scenarios. Although Envision's map-based visualizations and tabular representations (Figure 5a, 5b) communicate results well to scientists, this is less true for other stakeholders.[2] Future visual analytics research with Envision will involve extending and complementing its current output mechanisms to help managers better understand the effects on the landscape of different policies and share those with stakeholders prior to implementation. We anticipate that this research will involve a complex codevelopment relationship among computer scientists, programmers, ecologists, model

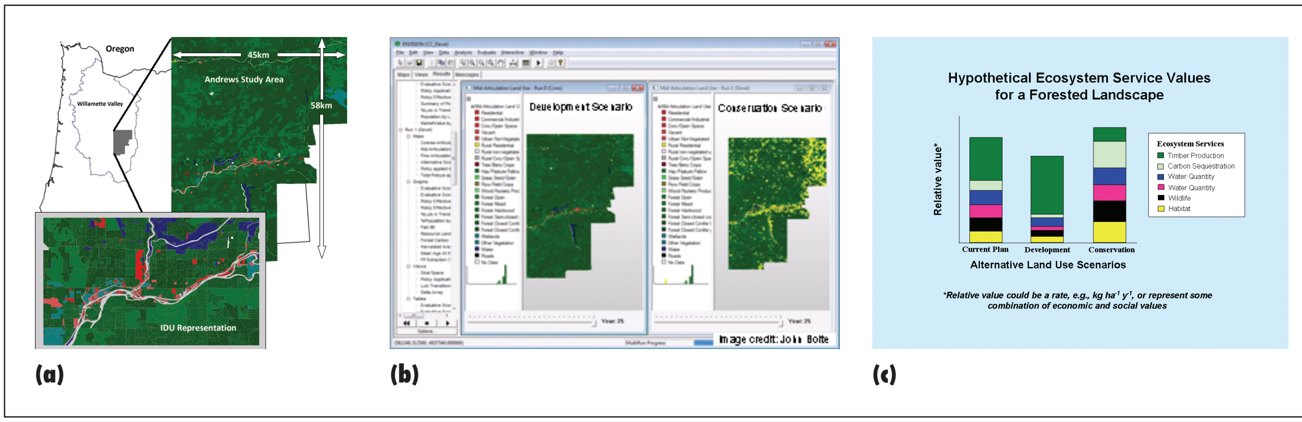

Figure 5. Visualizing models and data for decision support. From the 64 km² at the basin level, we here scale up to 2,000 km². One objective is to overlay 2D model output of features onto a 3D, topographically complex landscape. Another is to complement those still images and animations with statistical information in ways that a wide variety of stakeholders will understand. (a) Visualizing sensor and modeled data for regional landscapes: land cover type for Envision Andrews. (b) Envision provides animated 2D map-based visualizations to summarize user scenarios generated by ecological process models. (c) Simple analytics could complement more complex animated 2D and 3D visualizations such as those in Figures 4, 5a, and 5b.

developers, and stakeholders, and that the sociological observation of this relationship is as valuable an outcome of the work as the visualizations and software systems.

We generated Figures 4a and 4b with specialized software we are currently developing in our lab using C++ and VTK to display 3D and topological ecology data for researchers. Figures 4c, and 5a and 5b were generated using GTHM-PSM and Envision (respectively); we hope in the future to extend that software to handle 3D topographic visualizations. Figure 5c is a mockup of a possible future visual analytic display.

Thus far, our work to develop software for visualizing ecological research data has shown that visual analytics can be effective as boundary objects, but confirms other observations that development, evaluation, and interpretation of the visualizations is an iterative process that requires multiway communication. Visualizations provide an effective means for synthesizing and transforming scientific research data for specific purposes and end users. Our work confirms that visual analytics can only be most effective when user needs are incorporated in design, and that communication between developers and users is essential to develop successful systems and visualizations. The effectiveness of efforts to mobilize science and technology suffers when communication is largely one-way, is infrequent, or occurs only at the outset of a project. Our current work in visual analytics aims to determine what kinds of visual analytics increase communication and mutual understanding between scientists, resource managers, and other stakeholders; to build software for developing visualizations to that end; and to establish best practices for the codevelopment of visual analytics by computer scientists, ecology researchers, and resource managers.

Acknowledgments

We gratefully acknowledge the work of Marc Stiegliz and Feifei Pan of the Georgia Institute of Technology on the Georgia Tech Hydrology Model (GTHM), and their permission for our using that model. The visual analytics work is funded by the US National Science Foundation (grants BIO-0417311, CISE-0417311, and CISE-0639588). We also acknowledge the work of Fred Martin and Edward Keeley of the Washington State Department of Natural Resources in our prior study of visual analytics to repurpose ecological research data for managers. The information in this document has been funded in part by the US Environmental Protection Agency. It has been subjected to the Agency's peer and administrative review and approved for publication. The conclusions and opinions are solely those of the authors and are not necessarily the views of the Agency. Data and ideas contained in this essay were also provided in part by the HJ Andrews Experimental Forest research program, funded by the National Science Foundation's Long-Term Ecological Research Program (DEB 08-23380), US Forest Service Pacific Northwest Research Station, and Oregon State University.

References

1. P.R. Carlile, "Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge across Boundaries," *Organization Science,* vol. 15, no. 5, 2004, pp. 555–568.
2. S. Duncan and D. Lach, "Privileged Knowledge and Social Change: Effects on Different Participants of Using Geographic Information Systems Technology in Natural Resource Management," *Environmental Management,* vol. 38, no. 2, 2006, pp. 267–285.
3. P.N. Edwards et al., "Understanding Infrastructure: Dynamics, Tensions, and Design," workshop report, "History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures," Ann Arbor, Mich., 2007; http://deepblue.lib.umich.edu/bitstream/2027.42/49353/3/UnderstandingInfrastructure2007.pdf.
4. D. Lach et al., "Advocacy and Credibility of Ecological Scientists in Resource

Decision-Making: A Regional Study," *BioScience*, vol. 53, no. 2, 2003, pp. 170–178.
5. N. Kopytko et al., "Making Ecology Research Results Useful for Resource Management: A Case Study in Visual Analytics," *Proc. 10th Ann. Conf. Digital Government Research* (dg. o 09), Digital Government Soc. North America, 2009, pp. 206–215.
6. M. Stieglitz, R. McKane, and C. Klausmeier, "A Simple Model for Analyzing Climatic Effects on Terrestrial Carbon and Nitrogen Dynamics: An Arctic Case Study," *Global Biogeochemical Cycles*, vol. 20, no. 3, 25 Aug. 2006, GB3016.
7. S.L. Star and J.R. Griesemer, "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39," *Social Studies of Science,* vol. 19, no. 3, 1989, pp. 387–420.

Judith B. Cushing is a senior member of the faculty in computer science at the Evergreen State College. Her research involves developing information technology for scientists, particularly for environmental scientists. Contact her at judyc@evergreen.edu.

Natalie Kopytko received a master's degree in environmental studies from the Evergreen State College. Her thesis involved adaptation problems that arise when nuclear power is used in climate change mitigation. Contact her at kopnat10@evergreen.edu.

Nik Stevenson-Molnar works as a scientific programmer and research assistant with Judith Cushing. He received his BA in computer science and media studies from the Evergreen State College. Contact him at stevensn@evergreen.edu.

Lee Zeman works as a scientific programmer and research assistant with Judith Cushing at the Evergreen State College. She has BS in computer science from the University of Washington. Contact her at zemanl@evergreen.edu.

Susan Stafford is a professor of applied statistics and information management at the University of Minnesota. Her research interests are facilitating and understanding interdisciplinary research and education, and using visual analytics to aid in the flow and understanding of environmental data and information among researchers, practitioners, and stakeholders. Contact her at stafford@umn.edu.

John Bolte is an associate professor in the Department of Bioengineering at Oregon State University. His primary areas of interest are systems analysis and simulation, intelligent systems design, complex systems analysis, and ecological applications. Contact him at john.bolte@oregonstate.edu.

Barbara Bond is a professor of forest tree physiology in the Department of Forest Ecosystems and Society at Oregon State University. Her research interests are physiological processes at whole tree and forest ecosystem scales, and the development of new tools and techniques for measuring and monitoring physiological processes in forests. Contact her at barbara.bond@oregonstate.edu.

Denise Lach is an associate professor in the Department of Sociology at Oregon State University. Her primary areas of interest are environmental natural resource sociology and applied sociology, including social impact assessment, program evaluation, and organizational development. Contact her at denise.lach@oregonstate.edu.

Robert B. McKane is an ecologist at the Environmental Protection Agency. His research interests include ecosystem and plant community ecology, biogeochemistry, and modeling. He holds a PhD in soil science from the University of Minnesota. Contact him at mckane.bob@epa.gov.

Moving Toward "Intelligent" Policy Development?

Ann Macintosh, *University of Leeds*

Most problems faced by government are not as straightforward as voting in an election. They are, typically, what Horst Rittel and Melvin M. Webber termed "wicked problems." Examples of these are, "How do you tackle urban deprivation?" "How do you end international drug trafficking?" and "What do you do about climate change?" Participatory policy development around such issues must engage a range of actors in rational discussion to construct and deliberate on the complex policy options. This is typically done through a range of processes that, conventionally, include the provision of expert reports in which various experts present their choice of policy option on the basis of their specific expertise, and through consultation processes in which various stakeholders are asked to provide their views and debate options. With the development of Web 2.0, dominated by wikis, blogs, and a range of social-networking sites such as MySpace, Facebook, and YouTube, we see a further process emerging for policy discussion. These tools are appearing as important forums for civil society to debate and talk about policy issues such as global warming and terrorism, but also about local issues such as transport and urban development. On Facebook and YouTube, we see civil society discussing policy issues independently from government.

The evolving information, derived from these very different parallel processes, must be analyzed, managed, and integrated appropriately. Added to this, the resulting knowledge must be provided to the range of actors, who will need to make sense of it from numerous perspectives. In summary, policy development involves a large amount of knowledge that must be made explicit in different formats. This includes knowledge from many different sources and channels. Policy development represents one of the fundamental problems of information and knowledge management, that of abstracting meaningful messages from large volumes of heterogeneous data and information produced by multiple stakeholders.

Putting to one side, but definitely not ignoring, the social and political research required, and focusing on the technological options available, the following questions summarize the problems posed:

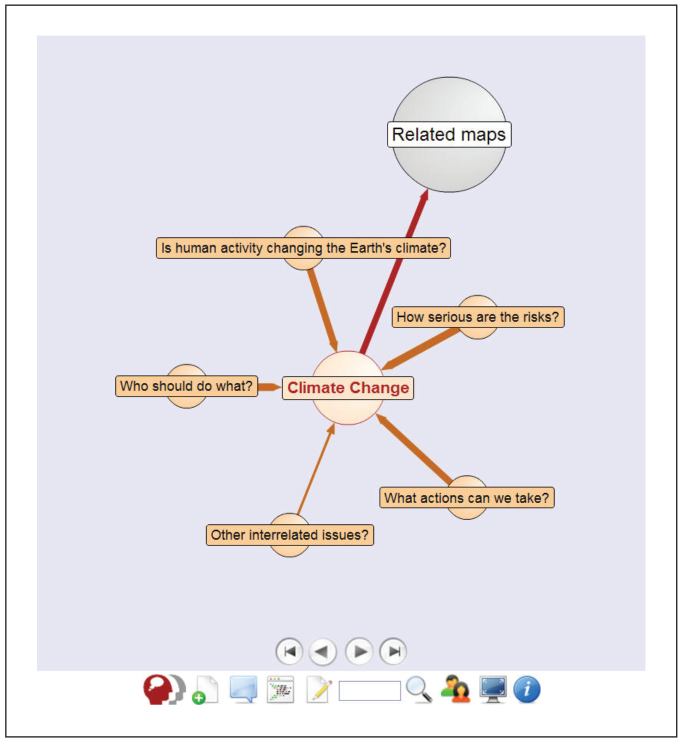

Figure 6. Debategraph perspective on a public debate over climate change: a lightweight, shareable flash view.

- How can the various actors determine the relationships between contributions to policy development—whether taken from expert papers, consultations, or public forum discourse—and appreciate how these contributions are taken through to decisions?
- How can the unstructured text from the various information sources be analyzed to enable the reconstruction of formal arguments?
- How can the actors understand better what critical questions to ask to determine the validity of the information put forward?
- Given the large, dynamic nature of the information base, how can the actors identify which issues are of importance to them, and how can they be supported to make reasoned contributions to policy development?

Having stated the problem, let's now consider some work that has already been undertaken in an attempt to address these questions, focusing on argumentation support tools. Floris Bex and colleagues divide such tools into two distinct types.[1] First, those that contain knowledge about a problem domain and can perform reasoning to suggest solutions to the problem—for example, dialogue and mediation tools. Second, those they term "sense-making" systems,[2] which typically do not support reasoning but rather structure the problem using visualization techniques.

In another publication, Thomas F. Gordon, Alastair Renton, and I offer detailed descriptions of existing argument tools and the use of argumentation systems to support political debate.[3] Alexandra Okada, Simon Buckingham Shum, and Tony Sherborne provide further useful descriptions on the application of argumentation systems including the use of Web 2.0.[4] (These two publications provide further details of all the systems I describe in this essay.)

Some existing systems include the Hermes argumentation tool, developed under the European Commission Intelligent Collaboration and Transaction Environments in Public Administration Networks (ICTE-PAN) project, which has a discussion forum with support for argumentation. Hermes allows for the construction of a diagram of the discourse that is composed of the ideas so far expressed during the discussion. The Parmenides system from the University of Liverpool uses a computational model of an argumentation scheme for practical reasoning to guide and focus deliberation dialogues by helping users to systematically address appropriate critical questions.

Two projects currently funded by the European Commission are the Vidi Project (www.vidi-project.eu) and the Wave Project. Vidi (short for "visualising the impact of legislation by analyzing public discussions using statistical means") aims to apply methods for visualizing statistical information to conventional discussion forums—for example, to visualize which topics, or which participants, have received the most attention. The Wave Project uses a publicly available tool, Debategraph (www.debategraph.org), to make the impact of complex EU environmental legislation on climate change easier for citizens to understand. The strategy is to integrate argument mapping with Web 2.0 techniques to attract users and make the maps viral. Figures 6 and 7 are screen captures from

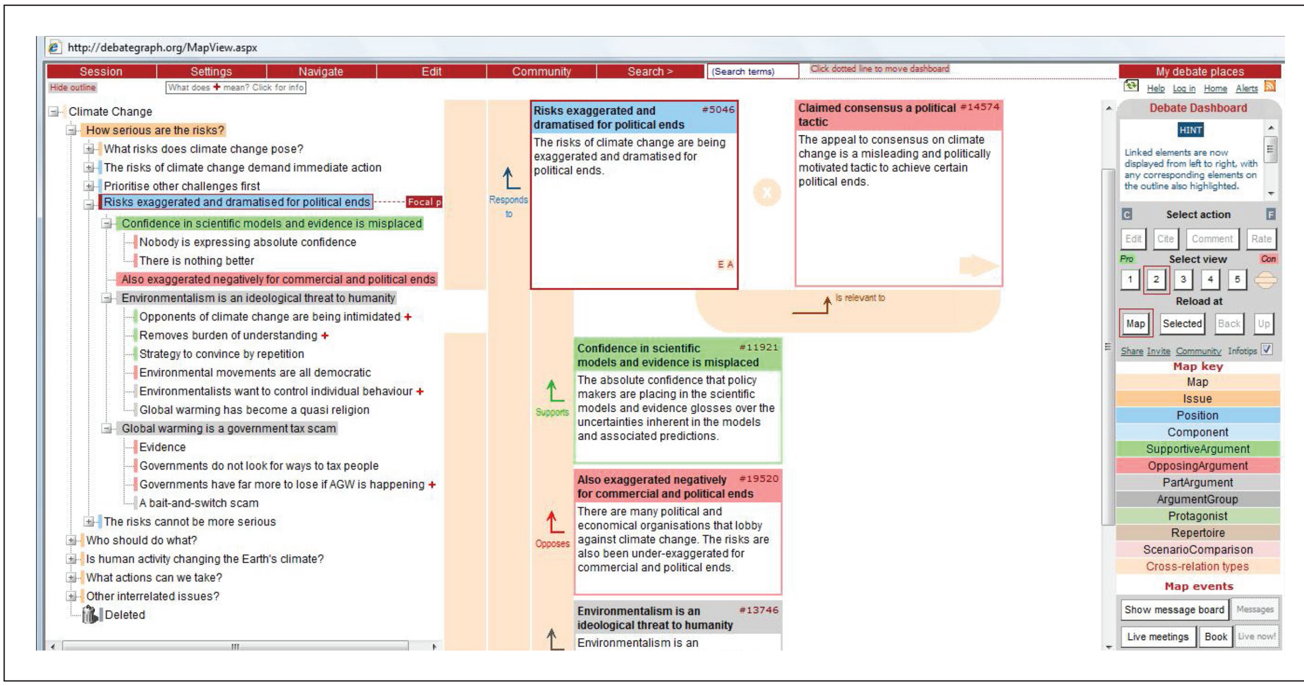

Figure 7. A second Debategraph view of the climate change debate, showing detailed editing and commenting.

Debategraph, showing the same debate on climate change from different perspectives.

Other researchers have used the Compendium tool to investigate how such systems can be used within a political context to support participatory policy development. Their work considered four possible uses for the Compendium tool: to provide information; to support consultation by considering an alternative way of setting out the responses to an online consultation over a published draft policy document; to support deliberation by setting out the consultation responses in the form of an inverted tree designed to let users see how their convictions on one issue may conflict with other beliefs; and, finally, to analyze a discussion forum to establish whether or not individual contributors had remained consistent throughout the debate—this could be used to support the analysis of the consultation process.

An interesting pilot version of a Semantic Web-based argumentation system, ArgDF (www.ArgDF.org), lets users construct arguments, then query them using a Semantic Web query language. The ArgDF Ontology is based on the concepts of the unified Argumentation Interchange Format (AIF), extended to include Doug Walton's argumentation schemes.[5] The pilot still requires more development work, particularly with regard to the interface and argument visualization.

Notwithstanding the more recent work on ArgDF, it should be very apparent that several constraints and limitations prevent existing systems from adequately addressing the four problems I posed earlier. For example, the argument-mapping tools typically require arguments to be entered manually; this is not only time consuming but also raises issues of objectivity. These tools are also usually based only on an informal model of argument, similar to Rittel's issue-based information system (IBIS) model, rather than using the state of the art in the field of computational models of argument. Moreover, some systems, such as Debategraph, let participants in the debate edit the argument maps directly; this raises concerns about the actors' ability to understand the argumentation model, given that even the simplest models have proved too complex for lay persons to edit and manipulate.

So how do we take things forward? What research is being undertaken, or should be undertaken, to address our four questions? The answers lie not in any single technological research area but in various directions. We need not only to progress the state of the art in specific areas but innovatively to combine research outputs to make substantial steps toward "intelligent" policy development. Among others, these technological areas include

- social-network analysis techniques to understand the network of connections between contributions to policy development (whether taken from expert papers, consultations, or public-forum discourse), and how ultimately they trace through to decisions;

- argument mining, based on machine-learning and data-mining algorithms, to support the reconstruction of arguments from the various information sources, including unstructured text;
- argument analysis, based on computational models of argument methods, to identify the argumentation schemes used and reveal implicit premises, so that actors can critically appraise the information;
- argument visualization, based on argument-mapping methods, to provide a variety of interactive perspectives of argument maps, at different levels of abstraction and detail to enable the actors to appreciate the complexity of the policy issues in their entirety, to zoom in on the issues that are of interest to them, and then to be in a position to articulate a reasoned contribution to the policy development.

By progressing the research in these areas and combining techniques appropriately, we will create platforms that have the potential to accumulate an evolving policy description and so move toward "intelligent" policy development.

References

1. F. Bex et al., "Towards a Formal Account of Reasoning with Evidence: Argumentation Schemes and Generalizations," *Artificial Intelligence and Law,* vol. 11, nos. 2–3, 2003, pp. 125–165.
2. P.A. Kirschner, S.J. Buckingham Shum, and C.S. Carr, *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making,* Springer, 2003.
3. A. Macintosh, T.F. Gordon, and A. Renton, "Providing Argument Support for eParticipation," *J. Information Technology and Politics,* vol. 6, no. 1, 2009, pp. 43–59.
4. A. Okada, S. Buckingham Shum, and T. Sherborne, eds., *Knowledge Cartography: Software Tools and Mapping Techniques,* Springer, 2008.
5. D. Walton, *Fundamentals of Critical Argumentation,* Cambridge Univ. Press, 2006.

Ann Macintosh is professor of digital governance in the Institute of Communications Studies at the University of Leeds. Contact her at a.macintosh@leeds.ac.uk.

E-Government 2.0 in Asia: Trends, Opportunities, and Challenges

Paul Jen-Hwa Hu and Han-fen Hu, *University of Utah*
Eric Liluan Chu, *Governor's Office, Taoyuan County, Taiwan*
Fang-Ming Hsu, *National Dong Hwa University*

E-government keeps citizens at arm's length and enhances government accessibility, efficiency, coordination, and service quality. To achieve full maturity, many countries are exploring e-government 2.0, which leverages salient Web 2.0 technology and social media tools to facilitate citizens' participation in public affairs and policy making.[1] E-government is evolving toward citizen-centered e-participation and e-inclusion, which aim at increased engagement by wider swaths of citizens and constituencies.[2]

In addition to delivering ubiquitous, performance-driven services through choice-based channels, e-government 2.0 emphasizes social networking and embraces e-democracy.[3] It empowers citizens and constituencies through network-based connections and discussions with government agencies characterized by enhanced openness, transparency, and experience sharing. A striking delineation is the changing role of government—from dominant to participative.

Take the United States, for example. The federal government has implemented assorted e-government 2.0 services, including the Gov Gab Blog, USA.gov on Twitter, and podcasts. Similar developments are under way in Asia, a region known for prolific e-government developments that consistently top global rankings. In this essay, we report several representative e-government 2.0 initiatives in Asia and highlight some important trends, opportunities, and challenges.

E-Government 2.0 Initiatives in Asia

E-government developments in Asia appear to progress in stages: infrastructure establishments, followed by primitive informational services, online transactions, digital transformation through vertical and horizontal integration, and finally citizens' e-participation and e-engagement. Moving forward with novel Web 2.0 applications, many Asian governments are experimenting with e-government 2.0; the following sections describe some representative initiatives.

Taiwan

U-Taoyuan, a large-scale e-government initiative undertaken by Taoyuan County, Taiwan, represents an important point of departure between e-government 1.0 and e-government 2.0. Figure 8 shows an English-language screenshot; for the actual Web site, see www.tycg.gov.tw/site/site_index.aspx?site_id=035&site_content_sn=9907. U-Taoyuan is the only Asian community to gain recognition as a Smart21 Community of 2009 by the Intelligent Community Forum (www.intelligentcommunity.org), a think tank devoted to studying the economic and social development of 21st-century communities and sharing best practices to foster effective adaptations to the broadband economy that will be crucial to communities' sustainable renewal and growth. The ICF has also awarded a Founders' Award to U-Aerotropolis, a part of U-Taoyuan, recognizing it as among the most innovative intelligent city development plans globally.

U-Taoyuan aims at advancing e-government toward full maturity by providing ubiquitous, comprehensive government services, with intelligence and seamless integration, to support the county's aviation and logistics industry. Citizen participation and community building are critical. Blogs facilitate open discussions and in-depth exchanges among citizens about policy and public-administration issues. The county government uses RSS feeds to disseminate instant updates and promote new services; mashups offer essential information to the general public and targeted communities. To increase e-participation, the county government facilitates citizens' ability to comment on and debate policies before their finalization.

The central government also is exploring novel e-government 2.0 services. Blogs seem to be a focal interest (see, for example, http://blog.www.gov.tw, launched January 2009). A common goal is to create a community-owned online platform on which citizens publish, discuss, and debate politics and public affairs. Blogs can foster e-democracy by facilitating citizen participation; they also enable government agencies to deliver accurate, up-to-date information or fact-based clarifications in real-time. For example, using the "track-back" function, a citizen can find relevant information from different sources (such as interest groups or government agencies) and review posted comments about a specific topic.

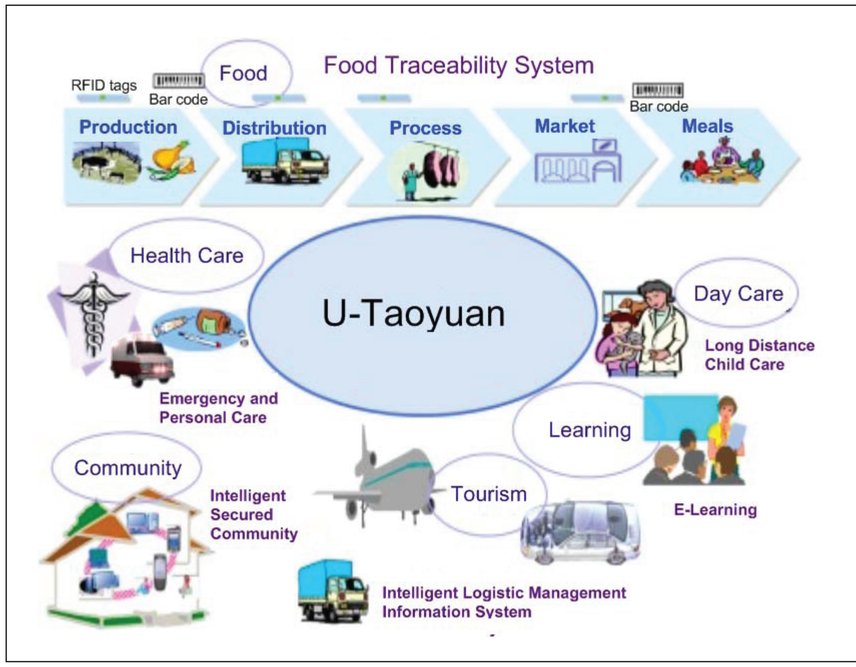

Figure 8. U-Taoyuan screenshot.

Governments can benefit from blogs as well. For example, the National Tax Administration uses blogs to deliver up-to-date tax information, filing requirements, and recommended techniques. Using these blogs, citizens can share tax filing tips, knowledge, and experiences.

Yunlin County uses blogs for community building, connecting agricultural producers, consumers, and government agencies. Consumers exchange assessments of different products, share their knowledge and experiences, and offer product selection tips or competitive price information. The farming community uses the blogs to share farming techniques and solutions to common or emerging problems. The county government uses blogs to disseminate regulatory requirements, clarify concerns about policy, promote farming education, and foster best practices.

Other social networking media help keep citizens and government agencies abreast of new topics or events. Vision 2020 (www.vision2020.tw), a nongovernmental platform, engages citizens to articulate and share their vision of the Taiwanese government of the future, via online voting and Webminars.

Japan

Open-Gorotto (http://open-gorotto.jp), a project initiated by an IT professional in the city government of Yatsushiro, Japan, and launched in 2004, represents an important e-government 2.0 development. It targets social-networking services at the municipal level to support community building, citizen participation, and disaster management. Open-Gorotto in Yacchiro attracts substantial media attention and seemingly emerges as a model practice. It has inspired national and local government agencies to explore social-networking services. Recognized by the Ministry of Internal Affairs and Communications as an iconic development, Open-Gorotto is implemented with open source software (including Free BSD, PostgreSQL, and PHP), and its social-networking platforms are hosted on government servers. For enhanced open discussions and experience sharing among citizens, Open-Gorotto leverages personal networks as multipliers to create sustainable online channels for community building and then deliver government services in a citizen-centered, facilitative, intelligent manner. This e-government 2.0 project combines blogs, networking, personal profiles, multimedia libraries, calendars, and newsgroups; it has demonstrated the feasibility and desirability of government using social-networking services to connect and interact with citizens for sustainable, mutually beneficial citizen relationship management.

Another representative project is Ococo-Nagaoka, launched in December 2005 to promote social networking services for increased citizen engagement (www.sns.ococo.jp). For neutrality, a nonprofit organization hosts its social-networking platform, though several government agencies provide financial support for site development and service marketing. Disaster management is a key objective, as Nagaoka is in an area that suffers frequent earthquakes, snow hazards, and floods. In the event of a natural hazard, Ococo-Nagaoka can generate real-time alerts and quickly disseminate them to citizens and various communities, together with contextual and geographic information. Although current usage is not high, Ococo-Nagaoka demonstrates the practical value of e-government 2.0 networking services for citizens.

South Korea

The national government of South Korea has proclaimed its commitment to comprehensive e-government 2.0 services, with distinct focuses on enhanced participation, intelligent administrative services, agile public-safety information networks, and state-of-the-art infrastructures for preserving privacy and security. According to an internal evaluation, online public participation (at www.epeople.go.kr) represents a best practice, facilitating citizens' engagement in the democratic process by enabling them to express opinions and engage in open discussions and debates with other citizens, political analysts, and policymakers.

A host of networking services, supported by RSS, tagging, and mashups, also facilitate citizen participation through intelligent keyword sharing, point-of-contact services, open search API, multicontent uploads, and privacy protection. Working with experienced external partners and software developers (such as Posdaq), the government has positioned e-government 2.0 as a de facto platform for delivering more than 10,000 services.

Trends

We observe several emerging trends. First, e-government 2.0 in Asia, still in its infancy but already attracting substantial attention, is gaining momentum. Many governments are exploring the use of social networking to connect citizens, build communities, and encourage engagement in public affairs and thereby gain collective wisdom from citizens' opinions, discussions, and debates. E-participation gives citizens opportunities to ponder incumbent policies, reflect on alternative views, exchange viewpoints, and reach consensus. For instance, Reach (www.reach.gov.sg), a blog-based platform in Singapore, connects citizens and gathers opinions and assessments in the form of forum discussions and e-polling.

Second, current service provision and utilization is relatively low but likely will increase quickly. For example, blogs hosted by the Taiwanese government (http://blog.www.gov.tw/blog) average only a few postings a day, but service usage appears likely to grow considerably in the near future.

Third, citizen connections and community building appear to be the responsibility of nongovernmental organizations. That is, governments actively participate in social networks and online communities but do not manage them directly. Such peer participation can create desirable neutrality that further propels citizen engagement; blogs hosted on neutral platforms (such as http://dignews.udn.com/forum and http://tw.forum.news.yahoo.com) appear to attract more participation by citizens than sites administered by government agencies directly.

Fourth, although current e-government 2.0 projects use blogs extensively, interest in other social-networking media appears to be rising. A multichannel approach can mitigate information asymmetry and overcome channel constraints, connecting more citizens and constituencies through their preferred channels. The governments of both Singapore (www.gov.sg) and Hong Kong (www.gov.hk/en) employ RSS feeds and mashups to deliver instant updates.

Furthermore, many governments are targeting intelligent, personalized services enabled by sophisticated analyses and information systems (such as advanced search engines, RSS feeds, video streaming, grid computing, and virtual reality). The Taiwanese government employs advanced analytical methods and computational algorithms to discover key patterns, phenomena, or knowledge from its digital documents and archives. Appropriate text-mining techniques could comb digitized archives to produce important knowledge maps—about such things as the 921 Earthquake in Taiwan, for example. By combining intelligent analyses and social-networking services, governments can reveal important relations, coherent tendencies, and unmet expectations, drawing on the "wisdom of the crowds."

Opportunities

Asian governments are paying more attention to long-term citizen relationship management and can benefit from data, text, or Web mining that reveals citizens' needs, preferences, and interaction patterns, which are crucial to delivering intelligent, personalized services.[4] For example, in the U-Taoyuan initiative, the Social Services Department could better assist senior and disabled citizens by identifying information and services they need and actively pushing those services through appropriate channels.

Governments also should offer effective knowledge mapping, beyond existing keyword-based search support.[5] For example, in the U-Taoyuan initiative, citizens, business communities, and government agencies might ask questions that require sophisticated analysis support and data from different sources—for example, "Which scenic spots in Taoyuan County attracted at least 1 million visitors each year between 2003 and 2008 and created more than 2,500 jobs within 10 kilometers?" or "Which sectors generated more than 10,000 jobs in the last five years, demanded no additional public services for education or health care, and received lower tax incentives than the top three revenue-generating sectors in Taoyuan County?" Knowledge mapping involves advanced modeling, ontology, and natural language processing; it also demands sophisticated data and text mining across heterogeneous data, documents, and systems, with stringent computational efficiency and scalability requirements.

Finally, through Web 2.0 social media tools, governments can connect people, enrich their experiences, and harness collective intelligence without claiming ownership or authoritarian dominance. Constant scanning is crucial and allows governments to gain a comprehensive understanding of citizens' concerns and viewpoints in an indirect, unintimidating way. Governments must perform timely sentiment analyses to identify underlying causes and present fact-based clarifications to network members through the major dissemination paths along which misperceptions might propagate and escalate. This effort demands timely analyses of citizen-generated content (such as opinions and comments) to discover subgroups of actors in a network. By uncovering main communication paths or patterns, governments can recognize the network structure and information flows holistically, identify crucial members in a network and their roles (such as opinion leaders and hubs across subnetworks), and engage them in open communication.

A fundamental shift in government agencies' mindsets is critical; they should select social networks to join rather than administering networks directly.

Challenges

To implement e-government 2.0 effectively, governments will also have to face and find solutions to several challenges. These issues are equally important for e-government in Asia and in the rest of the world.

Citizen Outreach and Trust

A central challenge of e-government 2.0 is reaching out effectively to citizens, making them aware of social-networking services, and facilitating their participation and service usage. Trust is absolutely critical and requires open communication about key benefits, service access, and privacy issues. Governments can cultivate and solidify citizens' trust through increased participation and repeated service usage, which then promote positive experiences with the network. Governments must preserve individuals' anonymity and ensure that no participants, including government officials, violate commonly accepted norms of online behavior (social contracts) and jeopardize citizens' trust.

Contributions to Social Networks

A social network should be owned and regulated collectively by its members. Governments should identify each major social network devoted to public affairs and administration, become immersed in that network, and contribute to the network by adding value to the discussions and sharing. Unlike e-government 1.0—with its user-centered service design philosophy but a rather inward-looking view of service provision, reflecting existing workflows and agency operations—e-government 2.0 is community-oriented, and the boundary between service provision and consumption is fuzzy. A fundamental shift in government agencies' mindsets is critical; they should select social networks to join rather than administering networks directly, and they should participate as valued network members rather than becoming the network owners.

Agile Response

Opinions and comments can propagate at amazing speed in a social network, which creates enormous pressure on governments to stay abreast of citizen-generated content, identify misperceptions or misrepresentations, and quickly respond with detailed, fact-based clarifications. The traditional response management model that many agencies use—full data gathering, thorough analyses, response preparation, dedication, and then execution—is far too slow to meet the time-compression challenge of vibrant social networks. Governments need new response models that offer the information velocity necessary for agile responses.

Identity and Information Privacy

To provide intelligent, personalized services, governments must understand and "recognize" targeted service recipients, but not in a personally identifiable manner. Such identity recognition should be practical, acceptable, and private. E-government 2.0 services embrace open discussions among citizens, so privacy becomes a top priority and must be ensured through strong evidence-based commitment and appropriate technological safeguards. To deliver intelligent, personalized services, governments need to gather data about various entities and thoroughly analyze it to reveal essential preferences, patterns, or interactions; however, governments must also properly preserve and protect each citizen's information privacy.

Data Source, Ownership, and Access Rights

E-government 2.0 services involve information from autonomous, heterogeneous sources that often vary in system platform, storage organization, and structure. Data ownership and access rights represent an important challenge that demands practical data governance at the enterprise level. Issues pertaining to data ownership and access rights could become political, subject to incumbent government bureaucracies. Addressing these issues requires strong senior leadership and cultural changes.

Advances in social-networking services require governments to examine issues of sociopolitical equality and address each threat that could lead to inequalities.

Digital Divide

E-government 2.0 aims at greater participation by citizens and communities, but agencies cannot overlook underprivileged constituencies that cannot afford or use the technology, or that lack network access because of their geographic locations, economic conditions, or educational backgrounds. Advances in social-networking services require governments to examine issues of sociopolitical equality and address each threat that could lead to inequalities. ■

References

1. W.D. Eggers, *Government 2.0: Using Technology to Improve Education, Cut Red Tape, Reduce Gridlock, and Enhance Democracy*, Rowman & Littlefield Publishers, 2005.
2. T. Obi, "Waseda University 2008 World e-Government Ranking Released," *Digest of Electronic Government Policy and Regulation,* vol. 31, no. 2, 2008, pp. 51–55.
3. J. Padget, "E-government and E-democracy in Latin America," *IEEE Intelligent Systems,* vol. 20, no. 1, 2005, pp. 94–96.
4. B. Medjahed and A. Bouguettaya, "Customized Delivery of E-government Web Services," *IEEE Intelligent Systems,* vol. 20, no. 6, 2005, pp. 77–84.
5. M.J. Eppler, "Making Knowledge Visible through Intranet Knowledge Maps: Concepts, Elements, Cases," *Proc. 34th Hawaii Int'l Conf. System Sciences* (HICSS 01), vol. 4, IEEE CS Press, 2001, p. 4030.

Paul Jen-Hwa Hu is a professor in the David Eccles School of Business at the University of Utah. Contact him at paul.hu@business.utah.edu.

Han-fen Hu is a PhD student in the David Eccles School of Business at the University of Utah. Contact her at han-fen.hu@business.utah.edu.

Eric Liluan Chu is the mayor of Taoyuan County, Taiwan. Contact him at llchu@llchu.org.tw.

Fang-Ming Hsu is an associate professor in the Department of Information Management at National Dong Hwa University, Hualien, Taiwan. Contact him at fmhsu@mail.ndhu.edu.tw.