Co-Evolution of Research and Practice:
The BLS and FedStats Case

Gary Marchionini,[1] Michael Levi,[2] and Stephanie Haas[1]
University of North Carolina at Chapel Hill[1], U.S. Bureau of Labor Statistics[2]
march@ils.unc.edu, levi_m@bls.gov, haas@ils.unc.edu

National Conference on Digital Government Research
Los Angeles, CA
May 21, 2002

The divide between research and practice arises in all complex enterprises. In industry, research and development units are necessarily at odds with production and marketing units; in academe, professors offer theory and students want skills they can adopt immediately; in medicine, researchers offer models of physiological processes and physicians want new diagnostic tools and treatments for patients. Thus, applying research to practical problems is a significant challenge in all areas of human endeavor. In rapidly developing fields such as information technology, the gaps may be temporarily narrow or the roles of research and practice may even be reversed. These natural and healthy differences lead to efforts to transfer technology from research to practice and to bring researchers and practitioners together to the mutual benefit of theory and application. This need is especially acute in government service, where so much of the business of government is information based (e.g., Department of Commerce, 2000). This need was one of the driving motivations for the NSF-led Digital Government program (http://www.diggov.org/) (See Schorr & Stalfo, 1997 http://www.isi.edu/nsf/prop.pdf for the rationale and http://www.isi.edu/nsf/ for associated white papers). It is especially sensible to find ways to help government agencies adopt the fruits of R&D that government agencies sponsor. Thus, important goals of digital government projects are to find ways to narrow the gap between basic research and practical application and to help researchers and government agency professionals learn from one another. To achieve these goals, one obvious path is to shorten the research horizon to focus on either low-hanging fruit or applied problems that are modules within long-term research agendas. This is the approach taken in the case reported here where information and computer scientists worked with the Bureau of Labor Statistics and other federal statistical agencies to broaden the populace's access to federal statistics. Over the course of five years of collaboration we claim that both research and practice have co-evolved toward this goal and we have laid a foundation for continued collaboration.

This paper first summarizes the specific ways that academic and government participants worked together toward their individual and shared goals, presents highlights of the results of these collaborations, and finally presents reflections on the factors that led to mutual benefits and those that continue to challenge us for the future.

1. Context for Collaboration

In 1996, two academic information scientists (Marchionini and Hert) began working with the Bureau of Labor Statistics (BLS) and the FedStats working group to improve outreach to constituent groups by means of the World Wide Web (WWW). One general goal of this work was to gain a better understanding of how non-specialists think about, access, and use statistical data. A second general goal was to understand and document how federal statistical agencies, the Bureau of Labor Statistics in particular, can adopt and adapt technologies to better serve the needs of diverse constituencies. Much of the work focused on a specific goal to design and test user interface tools that help citizens understand what federal statistics are available, access the

statistics that are most pertinent to their needs, and use these data to answer questions and make decisions. Over the years, additional academic and government partners joined the effort.

The research motivation arose from our long-term goals to understand human information-seeking behaviors and develop principles for human-computer interfaces that facilitate information-seeking. The practical motivation for BLS arose from the desire to improve dissemination of federal statistics by leveraging Internet technologies. The researchers were happy to have access to data (statistics) that had not been well-studied and the government participants were happy to have guidance on user-centered design principles. One very interesting new research interest related to organizational change emerged over the five years as the team worked together and reflected on data and experience. The collaboration has benefited both the original needs of the researchers and the government agencies as well as stimulated new research questions and interests, and serves as the basis of ongoing research and development.

In the first two years of collaboration, Marchionini and Hert conducted interviews with agency staff, intermediaries, and end users; analyzed the transaction logs of BLS systems; analyzed email requests; and created and usability tested prototype user interfaces. These efforts led to a user task and type taxonomy, methodologies for assessing user behavior, and user interface design guidelines. A user interface tool called the Relation Browser was developed for use in the Fedstats website and tested in laboratory studies. In the third year, Stephanie Haas joined the team and focused on understanding the linguistic mappings between end-user vocabularies and agency vocabularies. The Relation Browser tool was revised based on the usability tests and implemented as an alternative site map for Fedstats and field tested over a nine month period. This tool continues to be developed and expanded at UNC and has been applied to a patent database for carbon dioxide research and several digital library projects. A two year NSF Digital Government grant (1999-2001) brought Ben Shneiderman from the University of Maryland and Liz Liddy from Syracuse University together with Marchionini and Hert to focus on making statistical tables from federal agencies easier to find and use. The academic team worked with several government agency partners (BLS, EIA, Census, NCHS) to study how people find and understand statistical tables. Metadata issues for statistical information were investigated and mapped to interface tools. Dynamic query tools and natural language processing techniques were developed to facilitate finding statistical tables and a table browser tool was developed and extensively usability tested.

These efforts led to several technical reports and refereed papers that address theoretical issues related to human information seeking, electronic statistical tables (e-tables) as a distinct medium, dynamic user interface design and testing, and the interactions among information flows and organizational cultures (see www.ils.unc.edu/govstat for various papers and reports). The long-term goals of the various academic partners were advanced in incremental but important ways—especially through extension to statistical data and the specialized needs and requirements numeric data bring to information seeking and the technology that supports it.

From the government agency point of view, these collaborations have served to guide the development of specific websites and helped establish procedures and policies for long-term planning. Agencies benefited in three significant ways:

1) Education. Though the principles of Human-Computer Interaction were not entirely new to BLS at the inception of the Digital Government collaborations, the partnership brought Federal staff a deeper and more grounded understanding of usability engineering in practice. Researchers explicitly and implicitly demonstrated concrete models of user-

centered design and evaluation, models that BLS staff have internalized and are now using and expanding in other areas of work

2) Functional requirements. The various reports and less formal communications with researchers provided agency staff with specific recommendations for site improvement, many of which were adopted. Researchers' empirical test results lent credibility to agency redesign decisions and were often quite useful in marketing changes to higher levels of management.

3) Motivation. Researchers' "inspirational prototypes" stimulated new ideas and gave agency staff the impetus for more imaginative thinking than might have occurred otherwise. In particular the iterative "generate one or more candidate designs, test them to ascertain which elements succeed and which fail, use these results to create the next generation of designs, and repeat" approach to systems analysis freed agency staff from the fear of making mistakes and encouraged creative risk-taking.

Over the course of these projects, the academic and government teams held formal face to face meetings at least twice a year, communicated via email, email lists, and phone conferences, and collaborated on presentations at various professional meetings (e.g., ASA, ICES, NHSC, ASIST, JCDL, DG01).

2. Status of Selected Collaborations

These various collaborations have yielded a range of benefits. We focus on two here: the BLS website and the organizational interface of BLS. The BLS website in 1996, like most academic and government websites of that day, was very organization-centric. There were nine graphic buttons that mainly reflected how BLS is organized and how data was distributed in the print medium (data, economy at a glance, keyword search, surveys and programs, publications and research papers, regional information, about BLS, other statistical sites, and what's new). The vocabulary used was predominantly agency-oriented throughout the website. Transaction log analyses demonstrated massive numbers of requests for the Occupational Outlook Handbook and key values such as the consumer price index and the unemployment rate. Recommendations for making those resources and data available directly from the home page and moving less-often requested data deeper in the site architecture were made early and these changes were undertaken over several iterations. Interview data as well as usability tests demonstrated a disconnect between ordinary language vocabulary and agency-specific vocabulary. Over time, terms such as 'inflation' were included even though official BLS reports do not use the term to reflect concepts of consumer or producer price increases over time. User testing of early user interface (UI) prototypes based on the user task taxonomy demonstrated the efficacy of clustering user options according to common tasks. Interviews and usability tests also demonstrated that people are concerned with localizing data to their geographic region. Various prototypes and, more importantly, the usability tests conducted to assess them, demonstrated the efficacy of designs that minimize user clicks and aim to give people 'look aheads' to make decisions about clicking on hyperlinks. Many of our recommendations to focus on getting highly organized information on the home page and providing this information in layers that provide increasing details without page reloading time penalties have been adopted in the current BLS website. The current page has a business-like look with a large number of highly organized links to BLS data labeled with terms that reflect non-specialist vocabularies. A prominently displayed set of commonly requested data is provided with additional details available via mouseover events. Common currently asked questions are available as well as a dynamic US map to allow selections by regions and states. Overall, the site is much more oriented to wide ranges of non-specialist users with diverse labor and economic statistical information needs. This evolution was facilitated by the research results as BLS leaders were able to justify design changes with empirical results.

From the academic point of view, these changes have come about very slowly and not all of the recommendations have been adopted. For example, Java applets allow more innovative designs that provide more layers of information on a single screen. However, a government website that serves the entire population must aim for modest technical environments and user population expertise levels, thus Java applets and the associated technical challenges they bring in today's environment can not entirely replace functionality that depends only on lower-end capabilities. There is a significant tension between a demonstrably useful UI feature in the laboratory and implementing it in a production environment that includes strong competing service perspectives from diverse groups of distinct stakeholders (various government and political groups, consumer and business groups, etc.). Facing this tension directly over time in ongoing relationships is helpful to researchers to develop richer theoretical models of human or system behavior that include the constraints of real-world systems. The tension is also helpful to government technicians as they invent ways to incorporate useful features by pressing existing tools and techniques more fully and include research results in design rationales and long-term planning.

Equally important to the UI evolution of the BLS website was the development of a user-oriented culture within the BLS organization itself. The data systems, policies and procedures, corporate culture, and public face of an agency or institution taken together make up its *organizational interface* and encompasses the intersection of people, data, tools, and policies specific to an organization. The evolution of BLS' organizational interface is evidenced in several ways beyond the user-orientation of the website. Interviews with staff demonstrate that more time and effort is devoted each year to interacting with the public via email. New attention to quality control abounds as formal procedures for timely releases and error corrections/updates are developed. There is a realization that posting values in the WWW environment means that corrections are extremely difficult to propagate since distribution is wide and instantaneous (compared to traditional settings where a few large, well-known users accessed the data and could be directly alerted to errors or updates). In some cases, email responses to public questions are treated as official documents, thus making them subject to the same review requirements as press releases. Some interviewees say that this attention to quality has actually raised the morale of the staff by reinforcing the importance of their work. These changes have been an interesting new research thread to follow and we have begun to develop a theory of organizational interface that includes bidirectional information flows within the information life cycle.

Additionally, the collaboration has influenced the evolution of the BLS website from its origins as an add-on service to a crucial part of the Bureau's information infrastructure. Most importantly, usage continues to expand as millions of requests are handled each month and this broad-based usage has strong political influence. It is hard to imagine BLS operations today without Internet-based outreach. In addition to the public aspects of the organizational interface, there are several internal manifestations of this infrastructure. The staff and systems (LABSTAT) within BLS have assumed more prominence and resources in the organization. BLS operates a usability testing laboratory that tests UIs for the various web services as well as the survey data collection interfaces. The overall website development is seen as an ongoing process with planned redesigns and adoptions of new technical components. The staff has expanded transaction log analysis as a technique for monitoring and improving website quality as well as reporting impact data to oversight agencies. In our face-to-face meetings, the team has repeatedly discussed the increased volume of email requests that accompany the rising web traffic at BLS and ways that these requests can be better managed without new personnel.

Finally, the collaborations have brought new partnerships for ongoing work. We are beginning a new effort as part of the NSF Digital Government Program to study issues of integration requirements needed for a national statistical knowledge network (NSF Grant # 0131824). We

will work with Fedstats agencies and state agencies to make government statistical data more accessible and understandable to non-specialists who need data to make better decisions. We will address horizontal integration across federal agencies (e.g., interoperation for data and metadata vocabulary, interfaces, and the overall user experience), and vertical integration across levels of government (e.g., linking state and local data to federal sources). This new effort will build upon existing partnerships and extend them to new levels of government.

3. Reflections on Collaborations

From the research point of view, these collaborations have been highly beneficial. Our long-term research agenda to understand human information seeking and the influences of technologies on those behaviors has been enriched by adding statistical data as a new facet. The bulk of the work in information retrieval focuses on textual information. Our own work has focused on text retrieval interfaces and more recently on video data. The addition of statistical data adds an important dimension to our own research as well as the field of information retrieval overall. Unlike video, which is highly expressive with multiple channels of information flow requiring massive human bandwidth, statistics stand at the other end of the representation spectrum. Statistics are highly compressed representations—a single number that requires a few bits of storage and bandwidth to transmit can represent a complex amalgam of raw data and underlying assumptions. The CPI, for example represents an often-debated condensation of change in prices for 80,000 items in 200 consumer categories for two overlapping samples of the US population.

The scope of the 'story' behind this single numeric value is astounding. A single value packs an enormous amount of effort and has strong influence on the thinking and behavior of experienced observers. From an information theoretic perspective, the single value that takes little screen real estate or ink on paper is associated with a huge volume of additional information—there is a network of knowledge behind it. Finding ways to link or reveal these associations at a useful point in the user's interaction is an important design challenge. For example, some background information is necessary so that information seekers can be guided during search and to help them recognize that this background information itself has value in contextualizing the statistical values they will encounter during search. Other information aids in the use and understanding of retrieved statistical values. One of our happy discoveries is how rich the agency's information holdings are in this type of supporting information; the development task is therefore how and when to "make visible the invisible," rather than to create masses of new information resources.

Many of these challenges became clear as we worked to understand e-tables and how people find and understand them. At the level of general information-seeking activity, users' support needs are the same regardless of the format of the information; they require an understanding of what information they can expect to find on the site, how it is organized and structured, and how to synthesize it for their own purposes (Tenopir, 1999; Young & Diaz, 1999). Exploring how best to provide support for different granularities of user needs as well as types of data, including statistical data, is an important extension of information seeking research. Additionally, the partnerships with statistical agencies provided real data with all the special case constraints a production system entails. This was valuable to our theoretical understandings as we begin to incorporate statistical data into our overall theory of information architecture and information seeking. In addition to the theoretical advances, associations with important government agencies also brought credibility as we worked with people and forged partnerships with others.

The partnerships have been very beneficial from the agency point of view, as well. BLS, in particular, took advantage of several key products of the collaboration. The first was the user task

and type taxonomy created by Hert and Marchionini early on in the project. Prior to this work BLS had a fairly solid intuitive sense of the diversity of users' goals, statistical sophistication, and computer expertise, but had little or no detailed information and certainly no empirical evidence on which to base system specifications. Not only did the first Hert and Marchionini report portray a nuanced target population for future designs, it also specified the population that had to be sampled for end-user testing.

The second significant work product was the sequence of prototype user interfaces and the accompanying usability evaluations. These gave the LABSTAT design and implementation team a set of heuristics and concrete functional recommendations that were introduced piecemeal from 1997 to 2000, and then played a central role in the full site redesign of 2001.

The linguistic analysis conducted by Haas has been used to develop the "Topics A-Z" section of the current BLS public Web site as well as the site glossary. It is typically quite difficult for domain experts to "forget" their specialized vocabularies and use language that is intelligible to the lay public. Vocabulary mappings such as those prepared by Hass are invaluable both as a point of departure and as a pointed reminder that help systems and other explanatory materials must be framed in a context that makes sense to the intended user populations.

The table browser, though still not implemented on the BLS site, may ultimately prove to be the work product with the most significant long-term impact. Using the Web as a convenient mechanism to deliver traditional statistical artifacts rapidly to a large group of end users is one thing; taking advantage of the Web to fundamentally transform the most basic statistical artifact that exists is quite another. BLS is currently working hard to create the basic operational infrastructure required to implement table browser functionality. Building new applications on top of this infrastructure will be an exciting next step.

Finally, the importance of inspiration cannot be overemphasized. It is easy for a development organization embedded within a large bureaucracy to become so mired in day-to-day operations and maintenance activities that it loses sight of the broader horizons. Having the opportunity to meet and exchange ideas regularly with researchers who are not bound by the same constraints frees the imagination and creativity of in-house staff. Some of these new ideas are based on specific communications with the research community, others may have no obvious connection to the topics discussed but nevertheless would not have developed without such interaction.

Reflecting on these collaborations, it seems clear now that there was a co-evolutionary process underway as the academic researchers interacted with the agency personnel. It is certain that BLS and Fedstats would have continued to evolve their Internet-based dissemination and that many organizational changes would have occurred without any research partner involvement, although the self-reflection and user-orientation may not have been as prominent. Likewise, the academic researchers would have continued to develop their research agendas, although the statistical data facet would have not been as prominent. Reflecting on why these collaborations were fruitful and are the basis for continued work, there appear to be two primary factors of influence. Foremost, participants from both communities have been willing to be flexible in their goals and specific activities for the partnerships. The flexibility on the research side was related first to undertaking theoretical work with an eye to practical applications. This flexibility is possible if researchers see the specific studies as incremental contributors to a long-term research agenda rather than ends in themselves. A second kind of flexibility was related to remaining patient as recommendations rooted in empirical work were slowly digested and only sometimes heeded. On the practice side, participants were willing to look beyond immediate operational needs in a large, complex organization and consider alternative viewpoints about end users' needs that sometimes

conflicted with well-established practice. This give and take on both sides made co-evolution rather than parallel evolution possible.

Second, participants made real commitments to the collaboration. This was demonstrated most directly in the time invested on both sides and in the funding sought and provided to support the work (both in the form of grants to the researchers and in the use of BLS resources to implement the recommendations.) These commitments were facilitated by strong and sustained communication channels. In the early years, there were formal semiannual meetings at BLS and in later years, these were augmented by additional formal semiannual meetings with the larger Fedstats community. Email and telephone conference calls maintained momentum between the formal meetings and listservs allowed the entire community and subgroups to sustain discussion and exchange. There were more than 600 email messages on the BLS aspects of the project in years four and five; 1200 email messages logged for the two year NSF project, plus 100 project listserv (with 46 members) messages and several hundred more in specific folders for studies. In addition to these project-based communication venues, participants from academe and government collaborated on talks and papers at professional conferences and used those opportunities to continue discussion and planning. Presentations and papers were given at venues as diverse as the American Statistical Association, National Health Statistics Conference, Digital Government 2001, Joint ACM/IEEE Conference on Digital Libraries, American Society for Information Science and Technology, ACM Conference on Universal Usability, National Association of Counties Conference and Exposition, and the International Conference on Establishment Surveys. In addition, there were a number of workshops and meetings sponsored by NSF or BLS that involved multiple partners in presentations.

From a funding point of view, BLS invested each year in the work with contracts to academic researchers, NSF provided a two-year grant and in both cases, researchers invested time in preparing proposals and work plans. Another important illustration of time and resource commitments was Hert's ASA Fellowship to BLS for a year in which she interacted on a daily basis with the federal statistics community. Although grants and contracts flow from government agencies to researchers, these investments are truly a two-way street as substantial time, effort, and travel are required on the part of the researchers to carry out the work and insure that communication continues.

Others wishing to build ongoing partnerships are encouraged to adopt the following guidelines.
- Create or build upon shared interests and clearly state goals for the collaboration. From the shared goals, create specific objectives that unite theoretical frameworks with concrete products and outcomes.
- Hold regular meetings in small working groups plus annual conferences with the entire stakeholder community; augment these face to face meetings with email and listservs.
- Align individual goals (e.g. tenure/promotion) of participants with products and outcomes.
- Provide sufficient budget and time to accomplish the shared goals.

It is only through a series of quid pro quo exchanges of flexibility and commitments that trust builds up over time. Ultimately, it is this trust that causes co-evolution. These two factors of flexibility and commitment are of course strongly influenced by the personalities of the participants and one unifying characteristic in this case is a strong commitment by all participants to improving the human experience. This commitment is manifested in curiosity and willingness to try different approaches and a belief in sharing the results as broadly as possible.

References

US. Department of Commerce (2000). *Tech Transfer 2000: Making Partnerships Work*
http://www.ta.doc.gov/PRel/cd110a.pdf

Schorr, H., & Stalfo, S. (1997). *Towards the digital government of the 21st Century: A report from the workshop on research and development in federal information services.*
http://www.isi.edu/nsf/prop.pdf

Tenopir, C. (1999). Electronic reference and reference librarians: a look through the 1990s. *Reference Services Review*, 27, 33, 276-279.

Young, C. & Diaz, K. (1999). E-reference: incorporating electronic publications into reference. *Library Hi Tech*, 17, 1, 55-62.

For specific papers related to this project see www.ils.unc.edu/govstat