

WNV-BOT Portal Project Summary

Hsinchun Chen Daniel Zeng Chunju Tseng Catherine A. Larson
Department of Management Information Systems, University of Arizona, Tucson, Arizona

Millicent Eidson Ivan Gotham
New York State Department of Health; SUNY, Albany

Cecil Lynch
California Department of Health Services; UC Davis

Michael Ascher
Lawrence Livermore National Laboratory

The WNV-BOT Portal project is an NSF-funded infectious disease informatics project (grant number: EIA-9983304) aimed at demonstrating and assessing the technical feasibility and scalability of a cross-jurisdiction, cross-species infectious disease information sharing, alerting, and analysis framework. The related technical and system development research is summarized in a regular paper and a demonstration plan appeared in this proceedings. In this project summary, we briefly present the project accomplishments and describe the inter-agency, inter-disciplinary, and academia-government partnership that has been critical to the success of our project. We also share the lessons learned from our project on important issues faced by many digital government projects such as data sharing, academic research versus practical applications, funding mechanism, and project management.

Project Accomplishments

Our project started in October 2003. After an intensive 4-month research and system development effort, we have successfully completed a research prototype called the WNV-BOT Portal system. This system provides integrated, Web-enabled access to distributed data sources related to two prominent infectious diseases: West Nile Virus (WNV) and Botulism. These data sources encompass those available from all levels of government (federal, state, and local), in public health and other related domains (e.g., wildlife). As part of the project, cross-jurisdiction data sharing agreements have been developed. In addition to data access, the WNV-BOT Portal system also provides advanced information visualization and data modeling support.

Partnership

Our highly interdisciplinary research team mainly consists of three groups: (1) the Artificial Intelligence Laboratory at the University of Arizona, (2) the New York State Department of Health and its partner Health Research, Inc., and (3) the California State Department of Health Services and its partner PHFE Management Solutions. The National Biological Information Infrastructure/National Wildlife Health Center, as part of U.S. Geological Surveys, has also been an active research partner. This team composition is well-balanced, representing significant Information Technology research and system development experience, public health domain expertise, infectious disease data analysis experience and research competence. In addition, the providers of the key data sources and the targeted user community are well-represented in the team, reducing the barrier to data collection and sharing and helping ensure the practical relevance of the research and system development effort. We believe that this type of inter-agency, inter-disciplinary, academia-practitioner partnership is an important, if not indispensable, component of almost all large-scale, high-impact digital government projects.

Data Sharing and Protection

A significant portion of our initial project effort has been directed towards the development of data sharing agreements and a flexible data access control module as part of the Portal system. In public health, among other government operations, many applicable laws and regulations delineate the way through which data can or should be collected or disseminated. Sharing data across jurisdictions, in particular, poses many challenges. Some of these challenges are concerned with legal, regulatory, and policy requirements; others with the operational “norms” and other cultural and organizational factors. Our project experience attested to the effectiveness of an approach involving the explicit development of data sharing agreements between data providers and users, and a flexible data access control mechanism enforcing the agreed upon sharing rules. Such data sharing agreements need to cover in specific terms items such as data confidentiality (e.g., Can datasets be shared with parties outside of the project? Should the data be destroyed after a certain period of time?), ownership (e.g., Who owns the original data? Who owns the derivative data?), and the specific procedurals that need to be followed to bring in additional data providers and users. Typically, data sharing agreements can be made at the institutional level and are signed by the authorized institutional representatives. In cases where additional data protection is desired, individual user-level data sharing agreements can be created and all system users need to sign them. In addition to documentation of data sharing policies, it is important to design and implement the data access control mechanism to protect data and enforce the sharing policy at the system level. This mechanism should be an integral part of the system architectural design rather than an ad-hoc afterthought.

Research versus Application

Digital government projects are expected to accomplish both of the NSF mandates: advancing digital government and related technical and policy research, and making an operational and broad impact. Activities targeted at research and application, however, do not always align perfectly. In our WNV-BOT project, we have tried to deal with these alignment issues in a systematic manner through a staged approach. Specifically, in the initial stage of our research, we have focused on application-driven issues such as collecting data, identifying application needs, negotiating data sharing agreements, and building the initial computational infrastructure to support data access and sharing. These activities in themselves do not represent research innovations; however, they result in a working system useful for practitioners and provide the necessary technological and information infrastructure for follow-up research. In the second stage, we have mainly targeted at identifying specific technical research issues and developing innovative solutions. These solutions are in turn implemented in the prototype infrastructure and made available to practitioners for evaluation and operational use.

Funding Mechanism

Funding for our project is provided jointly by the NSF and Intelligence Technology Innovation Center through the NSF’s Digital Government Program. The University of Arizona team is the primary grantee with the other teams as the subcontractors. This funding model based on joint funding from the NSF and domain agencies has worked well. In addition, providing appropriate funding support for involved government agencies as data providers, domain experts, users and evaluators, is key to maintaining the partnership and promoting intense and productive academia-practitioner interactions.

Project Management

Extensive communications between team members, especially in the distributed, cross-jurisdiction context, have been critical to the success of our project. Our communication plan includes bi-weekly conference calls, joint presentations at disease informatics conferences, and milestones and the corresponding system demonstrations and face-to-face meetings. The WNV-BOT project has also greatly benefited from valuable feedback from two national committees. These committees are charged with the task of developing a national infectious disease infrastructure and consist of senior members of virtually all main federal agencies.