

# A National Center of Excellence for Infectious Disease Informatics: Project Summary

Hsinchun Chen Daniel Zeng Chunju Tseng Catherine A. Larson  
Department of Management Information Systems, University of Arizona, Tucson, Arizona  
{hchen,zeng,cal}@eller.arizona.edu; chunju@u.arizona.edu

Millicent Eidson Ivan Gotham  
New York State Department of  
Health; SUNY, Albany  
{mxe04,ijg01}@health.state.ny.us

Cecil Lynch  
California Department of Health  
Services; UC Davis  
colynch@ucadvis.edu

Paul Hu  
School of Accounting and  
Information Systems, University  
of Utah  
actph@business.utah.edu

## ABSTRACT

In this project summary, we briefly present the technical objectives and accomplishments of our Infectious Disease Informatics project. We describe the inter-agency, interdisciplinary, and academia-government partnerships critical to project success and discuss the broader application context for this research.

## 1. INTRODUCTION

The National Center of Excellence for Infectious Disease Informatics is an NSF-funded project (grant number ITR 0428241; prior NSF funding received under grant number EIA-9983304) with the goals of: a) developing an integrated and scalable information sharing, monitoring and analysis environment across jurisdictions and species for major infectious diseases; b) developing novel data analysis, surveillance, and visualization techniques to meet the critical needs of infectious disease informatics (IDI); and c) gaining a systematic understanding of related policy, user evaluation and technology adoption issues. This effort represents a critical initial step towards the full-scale implementation of a national infectious disease information infrastructure (NIDII). A prototype system called BioPortal has been developed, implementing many of the functionalities required for a NIDII. BioPortal is available for demonstration at <http://www.biportal.org>. More information about BioPortal and related technical research is available in [1] and [2].

## 2. PROJECT ACCOMPLISHMENTS

Our project was initiated in October 2003. The BioPortal system initially provided integrated, Web-enabled access to distributed data sources related to two prominent infectious diseases: West Nile Virus (WNV) and Botulism. (The system was named the WNV-BOT Portal in its early stage of development.) Protocols and agreements for adding additional data sources have since been established, allowing new demonstration datasets for Foot-and-Mouth Disease, and from BioWatch (an early warning biosensor system operating in several major U.S. cities), to be integrated. The BioPortal's infectious disease data query and analysis functions are summarized below:

- **Data Access.** BioPortal provides customized query interfaces to the available infectious disease datasets. For each dataset, spatial and temporal coordinates of the disease cases and sightings/test results, among others,

are essential. Both coordinates can be presented at different granularities (e.g., for location, specific street address/county/state; for time, specific day and time/weekday/week/month/year). Users often need to perform queries and then aggregate the findings based on various location/time granularity requirements. BioPortal provides a flexible tabular tool that allows users to select a preferred granularity level and presents the summary data accordingly.

- **Analysis and Visualization.** BioPortal supports hotspot analysis using various methods for detecting *unusual* spatial and temporal clusters of events. Hot spot analysis facilitates disease outbreak detection and predictive modeling. As part of BioPortal research, we have developed a new hotspot analysis method based on support vector machines, which has been shown to outperform existing methods using both simulated and real-world datasets [1]. BioPortal also includes Spatial-Temporal Visualizer (STV), a visualization tool which allows users to effectively explore spatial and temporal patterns, based on an integrated tool set consisting of a GIS tool, a timeline tool, and a periodic pattern tool.

We recently completed a preliminary user evaluation study to collect feedback on various aspects of the BioPortal implementation [3]. Initial findings indicate that both user information satisfaction and end-user satisfaction are significantly higher with BioPortal than with the benchmark program (the spreadsheets traditionally used for infectious disease data analysis). As a system, BioPortal was considered more usable than the benchmark program by our subjects, who also perceived BioPortal to be more useful and easier to use than the benchmark program. Follow-up evaluation studies are underway with a larger pool of subjects and infectious disease analysis tasks of greater variety.

## 3. PARTNERSHIPS

In the initial phase of the project targeting at developing the WNV-BOT Portal system, our interdisciplinary research team included three groups: (1) the Artificial Intelligence Lab 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. Our current research team has been expanded to include the University of Utah, the University of California Davis, and additional state public health partners. This team composition represents a balance of significant Information Technology research and system development experience, public health domain expertise, infectious disease data analysis experience, and user evaluation competence. In addition, we have made contact with a number of potential collaborators in order to expand the number of sources from which we will receive data. Existing agreements are being revised in order to take HIPAA regulations into account.

#### **4. APPLICATION CONTEXT AND PROJECT MANAGEMENT**

The BioPortal project addresses the interdependencies among disparate and distributed disease data systems. Owing to the real-time sharing of data that BioPortal could provide, the system could benefit public health agencies in their infectious disease fighting activities (preventing, detecting, managing) across jurisdictions. We also see potential applications for law enforcement and national security concerning biological terror attacks, especially with regard to future alerting capabilities. The technological and policy issues explored in this project are also applicable to other digital government domains.

The University of Arizona team is the primary grantee with the other teams as subcontractors. Providing appropriate funding support to our data providers, domain experts, users and evaluators, is key to maintaining the partnership and promoting productive interactions. Extensive communications between team members, especially in this distributed, cross-jurisdictional context, are critical to the

success of our project. We communicate through regularly scheduled conference calls, and collaborate on papers and presentations for a wide variety of professional, academic and technical audiences. Project objectives are mutually agreed upon, and system demonstrations are examined carefully by all participants. Such close working relationships are key to achieving success and sustaining a multi-year research project. At the same time, they enable us to pursue additional funding and project opportunities that aim to operationalize BioPortal-enabled applications and have a potentially significant real-world impact.

#### **5. REFERENCES**

- [1] Zeng, D., Chang, W., and Chen, H. (2004). "A Comparative Study of Spatio-Temporal Data Analysis Techniques in Security Informatics," in *Proceedings of the 7<sup>th</sup> IEEE International Conference on Intelligent Transportation Systems*, pp. 106-111.
- [2] Zeng, D., Chen, H., Tseng, L., Larson, C., Eidson, M., Gotham, I., Lynch, C., and Ascher, M. (2004). "West Nile Virus and Botulism Portal: A Case Study in Infectious Disease Informatics," in *Intelligence and Security Informatics*, Proceedings of ISI-2004, Lecture Notes in Computer Science, pp. 28-41, Vol. 3073, Chen, H., Moore, R., Zeng, D., and Leavitt, J. (eds.), Springer.
- [3] Hu, P., Zeng, D., Chen, H., Larson, C., Chang, W., and Tseng, C. (2005). "Evaluating an Infectious Disease Information Sharing and Analysis Systems," in *Intelligence and Security Informatics*, Proceedings of ISI-2005, Lecture Notes in Computer Science, Vol. 3495, Kantor, P., Muresan, G., Roberts, F., Zeng, D., Wang, F.-Y., Chen, H., and Merkle, R. (eds.), Springer.