Emerging Health Threats Journal 2011, **4**:s103. **doi:** 10.3134/ehtj.10.103

www.eht-journal.org

ABSTRACT

Real-time, reusable, dynamic public health surveillance

P Mirhaji, R Vedam, and MM Vagnoni

School of Biomedical Informatics, University of Texas, Houston, TX, USA
E-mail: Parsa.Mirhaji@uth.tmc.edu

Objective
This paper describes use of semantic technologies in combination with Services Oriented Architecture (SOA) to construct dynamic public health surveillance systems[1] used for just-in-time monitoring of emerging infectious disease outbreaks. The system was used for surveillance of schools in the third largest population center, Harris County.

Introduction
The resources available in most public health departments are limited. Access to trained technical personnel and state-of-the-art computing resources are also lacking. Customizable off-the-shelf systems contribute only to creation of information silos, are expensive, and not affordable by the limited budget available to the departments of health (only growing worse with the recession). The one thing that has increased is the need for surveillance in more areas, from diseases to environmental exposures to unexpected disasters. One solution would be an adaptable system able to cope with changing requirements while reusing or eliminating infrastructure from both computing hardware and technical personnel.[2] We report in this paper an instance of such system as used to perform disease surveillance across the Harris County school system. The system is designed to be customizable for surveillance of any disease, while simultaneously accommodating other use cases like disaster response and registries.

Methods
The Survey-On-Demand System (SODS) enables epidemiologist to create all necessary data capture for an area to be surveilled. This eliminates the need for personnel to design a web page, desktop client, or the backend data store. In addition, the system ties directly into a surveillance dashboard that allows epidemiologist to view trends, set alerts, and monitor multiple surveillance projects simultaneously.

The system is designed to abstract the format of the data captured so that it can be seamlessly exported to other formats (that is, Excel or SAS). All the data captured is backed by controlled and/or colloquial terminologies, which enable integration across multiple surveillance projects (past, present, and future) or information exchange with remote collaborators. Data is either mapped manually or natural language processing suggests appropriate terms from formal terminologies.

In order to cope with partial Internet connection (online somewhere, offline most places), an interaction model enables offline use and then synchronization when Internet is available. The server manages the state and macroscopic provenance of the forms. The client software logs every user interaction providing audit trails and item level rollback. All data is strongly encrypted. The system effectively manages the dynamic changes in the structure of the survey; no data is lost. The client keeps past data coherent with current data, as data collection needs change.

Results
More than 200 schools from Harris County were surveilled for absenteeism due to influenza-like illness through the academic year 2009–2010. The system is being evaluated to be used in other surveillance projects, including food-borne illness, environmental services, and children learning assessment. The SOA architecture enables a cloud-based information processing and data storage.[3]

Conclusions
The SODS system has been successful at both disaster and disease surveillance. It is being applied to multi-institutional clinical research and learning assessment in primary schools. The dynamic, model driven, information processing back-end services support context independent, disparate use cases harnessing the same system and investment without creating another information/systems silo. The warehouse is a hybrid containing both relational tables for Online Analytics Processing (OLAP) and a semantic repository for flexible, extensible representation of highly variable data. In addition, the surveys can be shared across organizations and projects, enabling reuse and further minimizing resources required to do surveillance. There are many interesting possibilities to expand this surveillance toolkit.

Acknowledgements

This paper was presented as an oral presentation at the 2010 International Society for Disease Surveillance Conference, held in Park City, UT, USA, on 1–2 December 2010.

References

1 Mirhaji P, Casscells SW, Srinivasan A, Kunapareddy N, Byrne S, Richards DM, *et al*. Services oriented architectures and rapid deployment of *ad-hoc* health surveillance systems: lessons from Katrina relief efforts. *AMIA Annu Symp Proc* 2006; 569–573.
2 Mirhaji P, Zhu M, Vagnoni M, Bernstam EV, Zhang J, Smith JW, *et al*. Ontology driven integration platform for clinical and translational research. *BMC Bioinformatics* 2009;**10** (Suppl 2): S2.
3 Bechhofer SK, Stevens RD, Lord PW. GOHSE: ontology driven linking of biology resources. *Web Semantics: Science, Services and Agents on the World Wide Web* 2006;**4:**155–63.