ABSTRACT

Mitigating data collection challenges with adaptive frameworks

A Poku[1], RA Wojcik[1], J Skora[1], S Lewis[1], and DL Blazes[2]

[1]The Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA; and [2]Division of GEIS Operations, Armed Forces Health Surveillance Center, Silver Spring, MD, USA
E-mail: adjoa.poku@jhuapl.edu

Objective

This paper describes the common challenges of data collection and presents a variety of adaptable frameworks that succeed in overcoming obstacles in applications of public health and electronic disease surveillance systems and/or processes, particularly in resource-limited settings.

Introduction

Electronic disease surveillance systems can be extremely valuable tools; however, a critical step in system implementation is collection of data. Without accurate and complete data, statistical anomalies that are detected hold little meaning. Many people who have established successful surveillance systems acknowledge the initial data collection process to be one of the most challenging aspects of system implementation.[1] These challenges manifest from varying degrees of economical, infrastructural, environmental, cultural, and political factors. Although some factors are not controllable, selecting a suitable collection framework can mitigate many of these obstacles. JHU/APL, with support from the Armed Forces Health Surveillance Center, has developed a suite of tools, Suite for Automated Global bioSurveillance (SAGES), that is adaptable for a particular deployment's environment and takes the above factors into account. These subsystems span communication systems such as telephone lines, mobile devices, internet applications, and desktop solutions—each has compelling advantages and disadvantages depending on the environment in which they are deployed. When these subsystems are appropriately configured and implemented, the data are collected with more accuracy and timeliness.

Methods

With SAGES we piloted multiple data collection methods against disease surveillance use cases. These pilots allowed real-time assessment of initial requirements and evaluation of performance. The discussion describes various challenges encountered by users and implementers of data collection tools for disease surveillance systems: availability, privacy regulations, timeliness, latency, coverage, digitalization, automaticity, reliability, centralization, usability, sustainability, and cost.[2] Solutions for these obstacles are presented along with a comparative analysis of the advantages and disadvantages of the various frameworks. Field notes from data collection initiatives with collaborators in Peru (US Naval Medical Research Center Detachment) and the Republic of the Philippines (Philippines-Armed Forces Research Institute of Medical Services Virology Research Unit), as well as relevant analysis of popular frameworks, supplement the discussion and provide real-life examples. Emphasis is placed on a process of understanding the targeted locale and then adapting the collection framework to operate within that locale's environment and unique impacting factors to capture relevant, accurate, and timely information.

Results

The success of any data collection initiative is dependent on the type of framework put in place. The negative impact and long-term effects resulting from using inappropriate frameworks are explored during the discussion. Furthermore, the discussion defines a process for mitigating data collection challenges by leveraging adaptable tools (Figure 1).

Conclusions

With understanding of the targeted environment, the success of surveillance systems is achievable even with inherent challenges being present. The discussion offers a method for decomposing the environment, identifying potential obstacles, and comparing suitable solutions

Figure 1 Data collection quadrant of APL's SAGES suite is customizable for a deployment's needs.

for those challenges to define the data collection framework that will best meet the needs of stakeholders and end users.

References
1 Lewis SH, Wojcik R. Methodologies for data collection. *BMC Proc* 2008;**2** (Suppl 3):SS. http://www.biomedcentral.com/1753-6561/2/S3/S5.
2 Ashar R, Lewis S, Blazes DL, Chretien JP. Applying information and communications technologies to collect health data from remote settings: a systematic assessment of current technologies. *J Biomed Inform* 2010;**43**:332–41.