NBIC Biofeeds: A Digital Tool for Open Source Biosurveillance across Federal Agencies

Heather Baker*[1], Chandra Lesniak[1], Emily Iarocci[1], Gus Calapristi[2], Scott Dowson[2], Michelle Hart[2], Lauren Charles-Smith[2], Yi Huang[2] and Teresa Quitugua[1]

[1]National Biosurveillance Integration Center, Washington, DC, USA; [2]Pacific Northwest National Laboratory, Richland, WA, USA

Objective

The National Biosurveillance Integration Center (NBIC) is developing a scalable, flexible open source data collection, analysis, and dissemination tool to support biosurveillance operations by the U.S. Department of Homeland Security (DHS) and its federal interagency partners.

Introduction

The NBIC integrates, analyzes, and distributes key information about health and disease events to help ensure the nation's responses are well-informed, save lives, and minimize economic impact. NBIC serves as a bridge between Federal, State, Local, Territorial, and Tribal entities to conduct biosurveillance across human, animal, plant, and environmental domains. The integration of information enables early warning and shared situational awareness of biological events to inform critical decisions directing response and recovery efforts.

To meet its mission objectives, NBIC utilizes a variety of data sets, including open source information, to provide comprehensive coverage of biological events occurring across the globe. NBIC Biofeeds is a digital tool designed to improve the efficiency of reviewing and analyzing large volumes of open source reporting by biosurveillance analysts on a daily basis; moreover, the system provides a mechanism to disseminate tailored feeds allowing NBIC to better meet the specific information needs of individual, interagency partners. The tool is currently under development by the Department of Energy (DOE), Pacific Northwest National Laboratory (PNNL) and it is in a testing and evaluation phase supported by NBIC biosurveillance subject matter experts. Integration with the Defense Threat Reduction Agency (DTRA), Biosurveillance Ecosystem (BSVE) is also underway. NBIC Biofeeds Version 1 is expected to be fully operational in Fiscal Year 2017.

Methods

The PNNL is applying agile methodology to streamline the build of NBIC Biofeeds to specifications required for operational use by NBIC and its federal interagency partners. Biosurveillance, analytics, and system engineering subject matter experts provide guidance on the implementation of features in the tool to ensure functionality aligns with operational workflows and production support. PNNL is leveraging software from a previous government effort to repurpose the technology to meet NBIC needs. NBIC Biofeeds incorporates the open source, document-orientated MongoDB database to capture user- and system-generated metadata on hundreds of thousands of records, in part, to establish baselines to aid prospective and retrospective analysis on emerging biological events. NBIC Biofeeds integrates a biosurveillance taxonomy (uniquely developed by NBIC), which includes input from interagency partners to recognize critical characteristics of a biological event. In NBIC Biofeeds Version 1, metadata capture of reported events is done manually by NBIC analysts; however, moving forward in Version 2, the tool will be further automated to flag significant reporting on biological events with a human remaining in the loop to confirm the validity of the system-generated tags.

Results

To serve as a one-stop tool for open source biosurveillance, NBIC Biofeeds automatically harvests information from thousands of websites, utilizing third party aggregators, paid subscriptions to data feeds, and scraping of high priority sources. Users can develop desired queries for automatic updating, leverage a unique review and curation mechanism, and further analyze data from topical, geographic, and temporal visualization features in the tool. To meet NBIC's information sharing needs, the tool allows for design of tailored RSS feeds and electronic message-based delivery of analysis on biological events, intended for recipients in the government with unique missions around human, animal, plant, and environmental health.

Conclusions

Through current testing and evaluation – underway by biosurveillance subject matter experts – NBIC Biofeeds is demonstrating value in supporting open source biosurveillance by the Center for more rapid recognition and sharing of key event characteristics. Centralizing access and analysis of this dataset into a single system is increasing the efficiency of daily, global biosurveillance, while enhancing the value of information identified through use of the querying, curation, and production support features in the tool.

Keywords

Pacific Northwest National Laboratory; National Biosurveillance Integration Center; MongoDB; Biosurveillance Ecosystem; Defense Threat Reduction Agency

*Heather Baker
E-mail: heather.baker@Associates.HQ.DHS.GOV