Health GeoJunction: Tracking Infectious Disease Threats and Related Science

Michael Stryker Scott Pezanowski Ian Turton Alan MacEachren
Geography Department, North East Visual Analytics Center, Penn State University

Many global health issues are now confronted through an array of international, national, and local organizations that must coordinate to marshal resources in assisting more vulnerable regions of the world and to face threats like catastrophic pandemics heightened by increased global connectivity. Monitoring outbreak events, collaborating in research and public health practice, and making decisions about how to best use resources and structure these activities requires an understanding of information that may be buried within the text of numerous documents at different locations and in different formats. The goal of this research is to create a visual analytic environment for quickly making sense of a collection of documents through interactive visual representations of the geographic, thematic, and temporal references within the text.

We have developed the Health GeoJunction, a highly interactive visual analytic web portal, that extracts information from, integrates, and geographically contextualizes scientific literature, public health reports, and event data to enable situation assessment and surveillance for infectious disease and related threats. The current prototype retrieves outbreak incidents reported by the United Nations Office International des Epizooties (OIE), journal abstracts from the PubMed library, and daily Really Simple Syndication (RSS) news feeds from the World Health Organization (WHO). The goal of this environment is to support: (1) situation assessment enabling an analyst to quickly become apprised of the current and evolving state of science and public health policy related to an infectious disease threat (e.g., an avian flu pandemic threat); (2) situation awareness by providing access to key indicators of an evolving threat for an analyst with epidemiological expertise; (3) characterize the social network of researchers in this domain by topical coverage, geographic focus, and research facility locations.

The current GeoJunction web interface provides multiple coordinated views organized around an overview map and time series plot summarizing the number of publications and outbreak incidents by country. A separate *tag cloud* view displays the most common keywords found in journal abstracts with the more common terms displayed first and in a larger font the than the less common terms. As an example inquiry, a user may focus on a subset of recent science, narrowing the topic and then selecting for publications about a geographic region of interest (e.g. - research on the interaction between domestic and wild fowl in relation to human infection by avian influenza in South East Asia). Interacting with the map, time series plot and tag cloud provide the means for filtering a subset of documents and subsequently navigating associated document lists via links cross-referencing authors, places and keywords.

Future development will include both refinement of the text processing capabilities for extracting geographic entities and keywords and improvements to the interface that give the user greater control in defining the geographic and temporal constraints for selecting subsets of documents.