Processes for Data Gathering, Assessment and Disease Event Tracking

Cynthia Johnson, DVM, Judy Akkina, Ph.D., Camilla Kristensen, Victoria Bridges, DVM

Center for Emerging Issues, Centers for Epidemiology and Animal Health, U.S. Department of Agriculture

OBJECTIVE

This paper describes a system of automatic and semi-automatic processes for data gathering, assessment, and event tracking used by the Center for Emerging issues (CEI) to enhance monitoring of global animal health events and conditions.

BACKGROUND

Protecting U.S. animal populations requires constant monitoring of disease events and conditions which might lead to disease emergence, both domestically and globally. Since 1999, CEI has actively monitored global information sources to provide early detection impact assessments and increased awareness of emerging disease events and conditions. The importance of these activities was reinforced after September 11, 2001, and these processes are now part of the U.S. Department of Agriculture's response to Homeland Security Presidential Directive 9 [1]. Electronic information sources available through the Internet have recently changed the way animal health information is gathered, processed and shared. To respond to these changes, CEI developed a dynamic system containing automated and semi-automated components that process information from various sources to identify, track, and evaluate emerging disease situations.

METHODS

Automated and manual scanning was used to assess information from a wide variety of open, proprietary and confidential sources. Text mining software, utilizing complex keyword queries, was used to filter the majority of the open source data gathered. Items of potential interest were assessed by analysts focusing on animal health issues that were significant, or in some way unusual, relative to morbidity, mortality, clinical signs, location, or public health. Events of interest were stored and tracked in a database, and an interactive prioritization algorithm was used to assign a level of importance to events and direct a potential response. High priority events generated a response in the form of alert notices, short reports, impact assessments, emerging disease notices, and periodic summary reports. Lower priority events were monitored for changes and developments that would elevate the event to high priority status. Verification of information took place through domestic and international collaborators.

RESULTS

From July 1, 2006, to June 30, 2007, nearly 470,000 records were processed from open sources alone. Of these records, 647 were events of interest, and 130 were classified as high priority events and tracked by an analyst. Only 13 events were considered of special interest and resulted in a response (Figure 1). Due in part to improvements made to event assessment processes, the number of events of interest decreased by 68% from 950 in the 05–06 period to 647 in the 06–07 period (Figure 1), yet the number of low and high priority events and responses remained similar for both time periods.

Figure 1. The records of interest are assessed and prioritized using an interactive algorithm, and are classified as low or high priority events. An immediate response is generated only for events of special significance or interest.
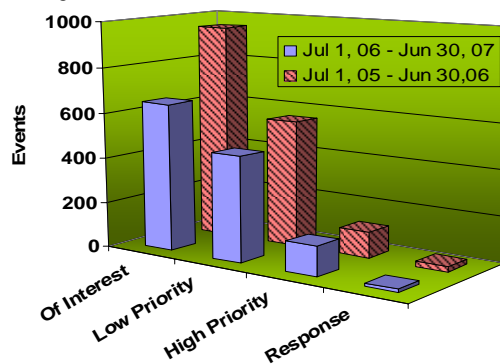
CONCLUSIONS

The dynamic nature of the processes used to monitor animal disease events and conditions has allowed individual components to be improved and added or removed independently to increase the effectiveness of the filtering and assessment processes used. In addition, the prioritization algorithm has provided consistency in the analysis of events and has minimized bias introduced by the subjectivity of individual human perspectives. To date, the open source information used has been limited to English language sources; however, the automated text mining process is currently under evaluation and is moving toward a more collaborative, integrated system that will include access to open and proprietary foreign language sources.

REFERENCES

[1]<http://www.whitehouse.gov/news/releases/2004/02/text/20040203-2.html>. Accessed July 20, 2007.