assessment and mapping, the paper presents the capabilities offered by satellite data and GIS techniques to manage flooding and related risks.

The satellite-based products will contribute to a preventive consideration of extreme flood events in Romania, by elaborating plans for flood mitigation, building infrastructure in flood-prone areas and by optimizing the distribution of spatial flood-related information to end-users. At the same time this information will provide decision-makers with updated maps of land cover/land-use, hydrological networks and with more accurate/comprehensive thematic maps at various spatial scales showing the extension of flooded areas and affected zones.

References

Brakenridge, GR, Anderson, E, Nghiem, SV, Caquard, S and Shabaneh, TB (2003), 'Flood Warnings, Flood Disaster Assessments, and Flood Hazard Reduction: The Roles of Orbital Remote Sensing', Proc. of the 30th Int. Symposium on Remote Sensing of Environment (Honolulu, Hawaii), 131–140.

Samuels, PG (2004), 'A European perspective on current challenges in the analysis of inland flood risks', Proc. NATO Advanced Research Workshop–Flood Risk Management Hazards, Vulnerability, Mitigation Measures (Ostrov u Tise, Czech Republic), 3–12.

Stancalie, G and Craciunescu, V (2005), 'Contribution of Earth Observation data supplied by the new satellite sensors to flood disaster assessment and hazard reduction', Geo-information for Disaster Management (ed. by P.van Oosterow, S. Zlatanova & E. M.Fendel), 1315–1332.

MedISys: A Multilingual Media Monitoring Tool for Medical Intelligence and Early Warning

Ralf Steinberger, Flavio Fuart, Bruno Pouliquen and Erik van der Goot

European Commission, Joint Research Centre, 21020 Ispra (VA), Italy,
email: ralf.steinberger@jrc.it

Keywords: Public health, CBRN, threat monitoring, early warning, multilinguality.

Introduction

Most industrialised countries have an institution that monitors their country's Public Health (PH) situation, by looking out for potential chemical, biological, radiological or nuclear threats (CBRN). One of the daily duties of these PH bodies normally includes monitoring the local, national and international media for reports on disease outbreaks, on the disappearance or the release of dangerous substances, etc. In order to facilitate this task, the European Commission's *Joint Research Centre* (JRC), in collaboration with the European Commission's *Directorate General for Health and Consumer Protection* (DG SANCO) has developed the automatic *Medical Intelligence System* MedISys, which takes away many of the repetitive and time-consuming tasks of this process and detects early warning signals that can be used as a starting point for the daily media review.

MedISys monitors an average of 50,000 news reports per day in 43 languages, from about 1,500 web portals and twenty commercial news providers. It filters documents of potential interest to PH officials, categorises them, monitors trends, and alerts users of an unexpected increase of articles in any of the fine-grained categories, separately for each country of the world. Articles are categorised using Boolean combinations of search words or lists of search words, with positive or negative weights, and the setting of an acceptance threshold. Users may require that

search words occur within a certain proximity (number of words), and may use wild cards.

MedISys allows to customise the view of reports to specific languages, subjects and news sources. It furthermore allows moderators to select, group and disseminate the information to further user groups, via email, SMS, web pages and PDF reports. For English media reports, MedISys displays automatically extracted structured information on disease outbreak events, including the disease name, time and place of the disease outbreak, the type of victims (people or animals), as well as their number and status (e.g. *dead* or *hospitalised*). The engine behind this event extraction process is Helsinki University's PULS system. PULS uses natural language pattern recognition in combination with an ontological knowledge base, which organises terms referring to diseases, viruses, drugs, locations, and more. For more details on MedISys and PULS, see Steinberger et al. (forthcoming).

A challenge faced by all PH organisations is the fact that a lot of relevant information is only available in foreign languages and that employing experts in all the languages is expensive and difficult. By aggregating the information found in many different news articles in all the available languages, MedISys can automatically issue warnings to users (through graphs, email alerts, etc.) as soon as relevant information appears in any of the languages covered, and often before the information is published in the user's own language. Fig. 1 shows a screenshot of MedISys.

MedISys provides access at three levels: (1) free public access with restricted functionality (http://medusa.jrc.it/), (2) restricted access for Public Health professionals outside the European Commission (EC), and (3) full access inside the EC. The EC version includes pay-for sources and newswires, which cannot be made available outside the Commission.

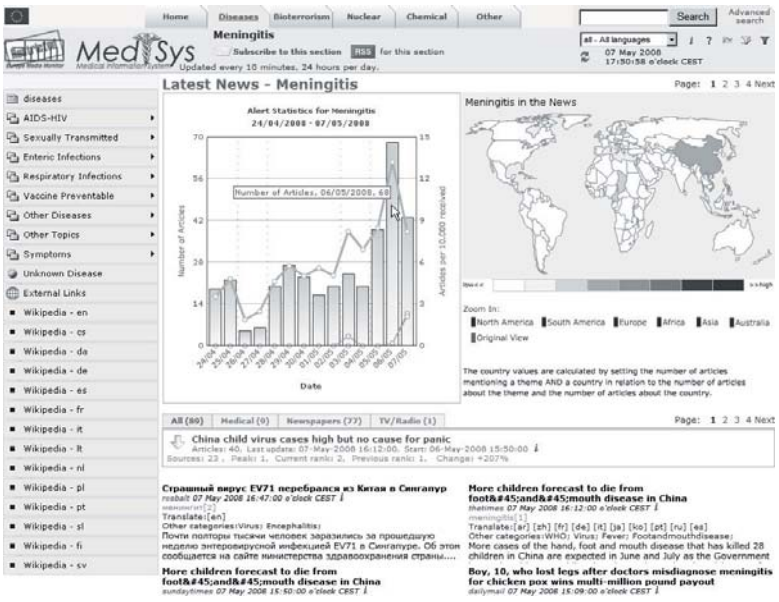

Fig. 1: View of the webpage on Meningitis in the Diseases section on the restricted MedISys site. The middle column shows various media report statistics regarding Meningitis. The map in the right column highlights the countries mentioned in the latest reports. The bottom section shows the latest articles plus extracted meta-information, and links to translate the articles automatically.

Related work

To our knowledge, there are two other systems that attempt to automatically identify information on communicable disease outbreaks and that display the results on maps: *Global Health Monitor* (Doan et al., 2008) and *HealthMap* (Brownstein et al., 2008). The former currently analyses English language news, the latter additionally offers French, Spanish and Russian. While *MedISys/PULS* and *Global Health Monitor* are fully automatic, *HealthMap* displays manually verified reports. The Canadian *Global Public Health Intelligence Network* GPHIN (Mawudeku et al., 2007) monitors news reports in eight languages purchased from two news aggregators. GPHIN does not extract events, but filters and ranks the articles for their CBRN relevance, and automatically translates them from and into English. A team of analysts additionally verifies part of the articles for relevance. All the mentioned systems also monitor the collaborative information-sharing portal ProMED-Mail (http://www.promedmail.org).

The major features that distinguish MedISys from the other systems are that MedISys (a) is fully automatic, (b) covers many more languages, (c) aggregates information across documents and languages, (d) calculates alert levels, (e) offers various ways to disseminate the information and (f) – in combination with the *Rapid News System* RNS - allows users to moderate MedISys input and disseminate it. MedISys goes beyond the functionality of *HealthMap* and *Global Health Monitor* in that it (g) monitors the whole range of CBRN and (h) automatically alerts users.

MedISys users

MedISys is actively used by international organisations such as DG SANCO, the *European Centre for Disease Prevention and Control* ECDC and the *World Health Organisation* WHO, by many PH bodies inside the European Union (e.g. the French *Institut de Veille Sanitaire* and the Spanish *Instituto de Salud Carlos III*), as well as by various bodies outside the EU (e.g. the US-American CDC and the Canadian *Global Public Health Intelligence Network* GPHIN). The public web site, available at http://medusa.jrc.it/, is visited by about 1,700 users per day.

Summary

MedISys is a multilingual media monitoring and alerting system. Access to the public version of MedISys is free. On request, Public Health bodies and other institutional users can get access to the restricted MedISys application, which offers more categories, and to the Rapid News System, which allows further customised views and tools for moderation and dissemination. MedISys is highly appreciated by the user community, but it can certainly be improved. Users are thus encouraged to give feedback on their user experiences.

References

S. Doan, Q. Hung-Ngo, A. Kawazoe, and N. Collier (2008), 'Global Health Monitor – a web-based system for detecting and mapping infectious diseases', *Proceedings of the International Joint Conference on Natural Language Processing* (IJCNLP), 951-956.

C. Freifeld, K. Mandl, B. Reis, and J. Brownstein (2008), 'HealthMap: Global infectious disease monitoring through automated classification and visualization of internet media reports', *Journal of the American Medical Informatics Association*, 15, 150–157.

A. Mawudeku, R. Lemay, D. Werker, R. Andraghetti and R. St. John (2007), 'The Global Public Health Intelligence Network'. In: N. M. M'ikanatha et al. *Infectious Disease Surveillance*, 304-317. Blackwell Publishing.

R. Steinberger, F. Fuart, C. Best, E. van der Goot, P. von Etter and R. Yangarber (in print), 'Text Mining from the Web for Medical Intelligence'. In: D. Perrotta, J. Piskorski, F. Soulié-Fogelman and R. Steinberger, *'Mining Massive Data Sets for Security'*, IOS Press.