

A new online resource to monitor new or emerging plant pests: MEDISYS media monitoring and the case of *Xylella fastidiosa*

Franco Ferilli¹, Giuseppe Stancanelli¹, Jens P. Linge² and Maria Rosaria Mannino¹

¹Animal and Plant Health Unit, European Food Safety Authority, Parma, Italy

² Joint Research Centre, European Commission, Ispra, Italy

Corresponding author: Maria Rosaria Mannino; E-mail: alpha@efsa.europa.eu

Abstract

The European Food Safety Authority has established a horizon scanning exercise for plant pests by automated monitoring of open-source media. The news sources are screened using the publicly accessible MEDISYS (Medical Information System) platform of the Joint Research Centre of the European Commission. Here, we report the example of monitoring for *Xylella fastidiosa*, a highly polyphagous plant pathogenic bacterium. Since its first occurrence in Europe, news sources have reported findings and latest developments. Media monitoring-related data can support surveillance or plant pests' management programs by early warning and can help understand the impacts of plant pests and the societal response to new plant health threats.

Media monitoring platform announcement

In collaboration between the European Food Safety Authority (EFSA) and the European Commission Directorate-General Joint Research Centre (JRC), a media monitoring system for the surveillance of plant pests was set up; FAO defines a pest as “any species, strain or biotype of plant, animal or pathogenic agent injurious to plants or plant

products” (FAO 2011). The system is based upon the fully automatic event-based surveillance system MEDISYS (Medical Information System).

MEDISYS is part of the Europe Media Monitor (EMM) system (Steinberger et al. 2013) and is a fully automatic public health surveillance system to monitor reporting on human and animal communicable diseases, chemical, biological, radiological and nuclear threats, and food and feed contaminations (Linge et al. 2009). The EMM system retrieves news items from official and unofficial media sites, general news media and selected blogs. Generally, news items are either retrieved from RSS feeds or by crawling HTML sites. RSS feeds, or rich site summary, are a type of web feed allowing the user to access and keep track of updated online content in a standard format.

Over 22,000 feeds of general news sites from over 107 countries are currently monitored. These sources comprise news websites at national, regional and local levels. In addition, feeds from specialist plant health sources, such as scientific journals, were added in the directory of monitored sources.

A Plant Health Threat Ontology with pest and disease names coming from multilingual sources such as UniProt Taxon (Consortium 2016), EPPO (EPPO 2018), and Wikipedia was built (see Alomar et al. (2015) and (2016)). Multilingual keywords definitions for 560 categories for plant health threats, or pests, were created. Scientific and common names comprising colloquial terms used by journalists and the general public were used.

The system categorizes all incoming items according to the ontology of plant pests, identifies known names such as organizations, persons and locations, extracts events,

clusters news items and calculates statistics to detect emerging threats. Each item thus comes with metadata (Alomar et al. 2016) that are added during the automated processing. The metadata include among others: (i) entity, or people or organisations automatically recognized by EMM; (ii) category, or plant health threats as defined in the multilingual ontology; (iii) geolocation, or geographic locations mentioned in the items; and (iv) quote, or extract of text recognized by the processing chain as quotes.

Here, we present the monitoring results for the highly polyphagous plant pathogenic bacterium *Xylella fastidiosa* (Wells et al. 1987) carried out in the period from February 2017 to March 2018.

The multilingual category definitions for *Xylella fastidiosa* keywords cover the scientific names, various associated disease names, and common names of the pathogen.

Raw data were filtered to show: title of the news item, URL of the item, date of publication, and country of publication. Pivot table were created in Microsoft Excel to visualize and analyze the different trends over time and to identify peaks. These peaks were subsequently matched with specific events in the different countries in the news history. For example, a peak in the number of news items from Spain in July 2017 (see Table 1) is due to the first findings on Spanish mainland of *X. fastidiosa* in almond orchards. The infected plants were identified in the province of Alicante, Valencian Community.

The data for the period from February 2017 to March 2018 shows a high number of news stories about *X. fastidiosa*, with a total of 5,299 news items retrieved by the corresponding MEDISYS category. The news items come from a diverse group of 630

sources from 62 countries. This underlines the multilingual ability of EMM and of the developed ontology to retrieve news items. Table 1 details the number of news items selected for the different sources from the ten most active countries. Most of the news items mentions geolocations in Italy (2,287 items), Spain (1,746), United States (137), Greece (117), France (93), Germany (86), Great Britain (44), Belgium (36), Luxembourg (30), and Netherlands (23).

The media monitoring shows that *X. fastidiosa* is a highly discussed topic, with many items referring to detection, prevention and control of the pest in different European countries. The EFSA media monitoring results are regularly presented and discussed at the Plant Health section of the Committee on Plants, Animals, Food and Feed.

MEDISYS is publicly available; see <http://medisys.newsbrief.eu/medisys/alertedition/en/XylellaFastidiosa-PHT.html> for the latest news about *X. fastidiosa*. Articles can be filtered by language, news sources, country of the news sources, and country mentioned in the text. In addition, the content retrieved for all the 560 pest categories is available in the dedicated sections on the same webpage. A subscription to a customizable e-mail alert is available as well. Users can screen the categorized articles and display world maps highlighting event locations together with statistics on the reporting of plant health threats and countries.

A real-time map with event locations for articles retrieved by MEDISYS is available at <http://medisys.newsbrief.eu/emmMap/?type=category&language=all&category=XylellaFastidiosa-PHT>.

Media monitoring data can be used as a supplementary method to usual biosurveillance to help plant protection organizations with early warning, awareness and rapid response (Thomas et al. 2011). Monitoring public media can thus support surveillance or pest management programs and can help understand the impacts of plant pests and the societal response to new plant health threats.

DISCLAIMER

The positions and opinions presented in this article are those of the authors alone and are not intended to represent the views or any official positions of their institutions.

Literature Cited

- Alomar, O., et al. 2015. Development and testing of the media monitoring tool MedISys for early identification and reporting of existing and emerging plant health threats. *EPPO Bulletin* 45:288-293.
- Alomar, O., et al. 2016. Development and testing of the media monitoring tool MedISys for the monitoring, early identification and reporting of existing and emerging plant health threats. *EFSA Supporting Publications* 13.
- Consortium, U. 2016. UniProt: the universal protein knowledgebase. *Nucleic acids research* 45:D158-D169.
- EPPO. 2018. EPPO Global database (available online). <https://gd.eppo.int/>.
- FAO. 2011. ISPM 5: Glossary of Phytosanitary Terms.
- Linge, J. P., et al. 2009. Internet surveillance systems for early alerting of health threats. *Eurosurveillance* 14:19162.
- Steinberger, R., et al. 2013. An introduction to the europe media monitor family of applications. arXiv preprint arXiv:1309.5290.
- Thomas, C. S., et al. 2011. Use of media and public-domain Internet sources for detection and assessment of plant health threats. *Emerging health threats journal* 4:7157.
- Wells, J. M., et al. 1987. *Xylella fastidiosa* gen. nov., sp. nov: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *International Journal of Systematic and Evolutionary Microbiology* 37:136-143.

Table 1 Number of news items selected by the *Xylella fastidiosa* category from February 2017 to March 2018 for the ten most active countries.

Month and year of publication	Country of publication ^a										Grand Total
	IT	ES	US	GR	FR	DE	GB	BE	LU	NL	
February 2017	126	213	13	3	14		4	4	4	1	382
March 2017	401	125	6	1	7	1	2		3	2	548
April 2017	173	31	1	2	8	8	1		3	1	228
May 2017	138	44	10	2	2			6	4	1	207
June 2017	124	56	4	10	5	37	2	2	3	3	246
July 2017	137	254	7	8	11	9	4	9	1	1	441
August 2017	172	91	15	1	3	10	4	2	2	1	301
September 2017	263	190	12	18	11	1	7		2		504
October 2017	161	157	13	14	6	4	3	2	4	2	366
November 2017	126	161	8	10	10	5		4		7	331
December 2017	122	131	8	8	9		8	2	1	1	290
January 2018	54	111	10	5	2	1	4	1	1	3	192
February 2018	76	94	16	10	1	2	3	4	2		208
March 2018	214	88	14	25	4	8	2				355
Grand Total	2,287	1,746	137	117	93	86	44	36	30	23	4,599

^a Country of publication: IT=Italy, ES=Spain, US=United States, GR=Greece, FR=France, DE=Germany, GB=Great Britain, BE=Belgium, LU=Luxembourg, NL=Netherlands.