

Multilingual news gathering, information mining and analysis

Europe Media Monitor - EMM

EMM is a live media monitoring system that gathers reports from news portals world-wide in 35 languages, aggregates and classifies the articles, analyses the news texts by extracting information from them and produces intuitive visual presentations of the data found.

User-driven multilingual news gathering and aggregation

Vital statistics

- Monitors ~1,200 key news portals world-wide plus 20 commercial newsfeeds and, for some applications, also specialist sites.
- Retrieves over 40,000 reports per day in 35 languages.
- Classifies all news according to hundreds of subjects and countries.
- Access on the web, via email and by RSS.
- Runs 24 hours per day, 7 days a week.
- Developed, maintained and run by the JRC.

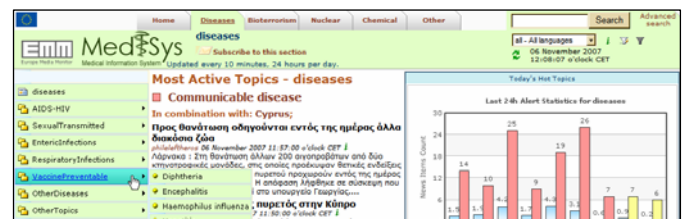
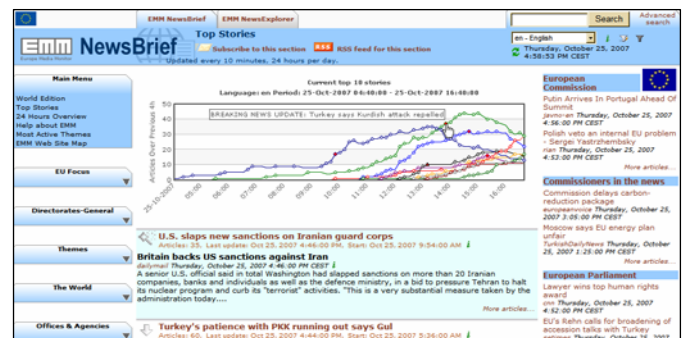
EMM was initially developed as an in-house application for the European Commission's **Directorate General Communication** (DG COMM) to replace their tedious and expensive manual media monitoring and press cutting services. Since then, EMM has become a crucial instrument in the daily work of the Commission and other public organisations. EMM is the news gathering engine behind a number of applications. These include the three **public web portals** *NewsBrief*, *NewsExplorer* and *MedSys*, which attract up to 1.2 Million hits per day.

Breaking news from around the world

The public web application **NewsBrief** (<http://press.jrc.it/>) shows the hottest topics discussed across multiple news sources and languages during the last few minutes and hours. Updated every ten minutes, the site groups related news and displays the biggest stories first. Users can opt to see only news falling into subjects of their choice.

The Medical Information System **MedSys** (<http://medusa.jrc.it/>) displays only articles with interest to Public Health, grouped by disease or disease type.

Both applications offer to send email alerts for breaking news or daily news updates.



Collecting reports from correspondents

The European Commission has representations in many countries where staff read the local press daily and report new events of interest back to Brussels. The web-based **PressReview** software allows representations to submit their reviews and attached newspaper clippings to a central system. Local staff can then combine them easily into general or thematic reports for distribution.

Customising the view of the news and tools to produce newsletters

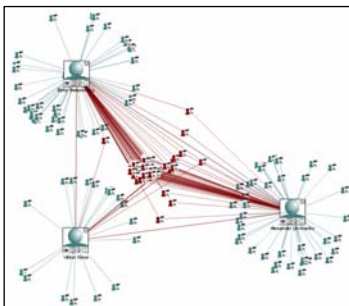
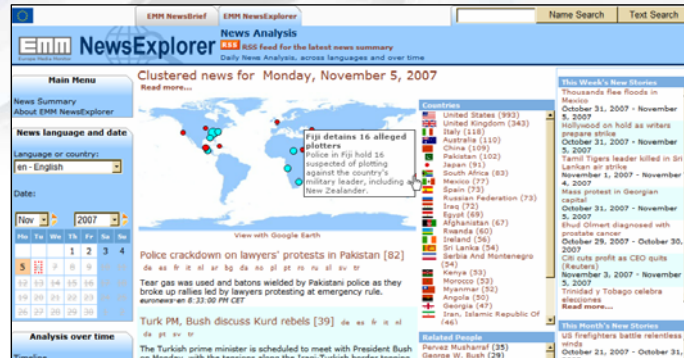
Registered organisations can use the **Rapid News Service** RNS to select news items by language, news source and subject and to produce their own newsletters by grouping news items and by dragging and dropping them in an easy-to-use interface. With a simple click, email or PDF versions can be sent to defined user groups.

News analysis across languages and over time

The public **NewsExplorer** application (<http://press.jrc.it/NewsExplorer/>) further analyses the EMM news articles in currently nineteen languages, extracts information about people, organisations and locations, and links related news items over time and across languages. The news from all sources and all countries are aggregated into one view, thus offering an impartial view of the world news, independent of regional or linguistic bias.

Daily news summary

For each language, all news articles talking about the same event or subject are grouped into **clusters** and displayed by cluster size. Users can see the people, organisations and locations mentioned most in today's news. A map shows where today's events happened. The calendar allows to read about past events.

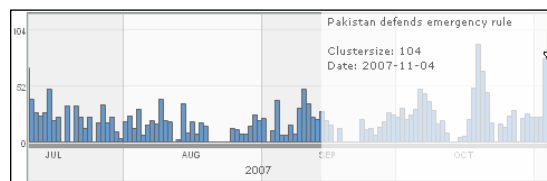


Collecting information about people, organisations and countries

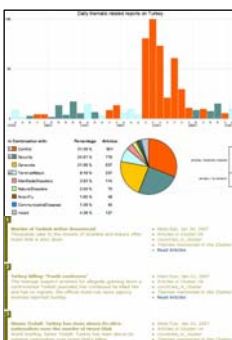
The application keeps track of who gets mentioned where and how often, who gets mentioned with which other persons or countries, etc. Up to 170 spelling variants for the same person have been detected automatically so that users can search for information about persons independently of the name spelling. Quotations by or about people are also displayed. The database contains close to 700,000 names.

Trends over time

Each day's news cluster is automatically linked to the news of previous days and months. Timelines show developments over time and allow users to jump to the origin or to peaks of each story.



Cutting edge research results



EMM-Labs gives an insight into the team's innovative research activities, showing how a unique combination of advanced text mining and visualisation tools can produce concrete analysis results. Prototype applications include: monthly **Information Sheets** for countries and themes; automated **Social Network** construction; extraction of **Violent Events** including cause, victims and geo-location from live news reports; **Networks** of news themes and real time news clusters; **Media Impact** by theme, country or person; and more.



CONTACTS

Dr. Delilah Al Khudhairy, Head of Unit 'Support to External Security'
Tel.: +39-0332-78 5696
Fax: +39-0332-78 5154

E-mail: Delilah.AI-Khudhairy@ec.europa.eu
Websites: <http://press.jrc.it>, <http://press.jrc.it/NewsExplorer>