



Language Technology

Complementarity of information found in media reports across different countries and languages



Ralf Steinberger

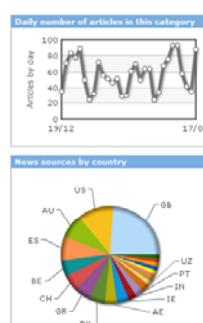
& the JRC's *OPTIMA* team – Open Source Text Information Mining and Analysis

Technical details and publications: <http://langtech.jrc.ec.europa.eu/>
Applications: <http://emm.newbrief.eu/overview.html>

Agenda

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

- JRC: Who we are – what we do – our customers.
- Europe Media Monitor (EMM) family of applications
 - Publicly accessible at <http://emm.newbrief.eu/overview.html>
- Motivation for multilingual text processing
- How to get access to this complementary information
 - Multilingual category definitions and alerts
 - Linking of related news across languages
 - Multilingual information gathering on named entities
 - Multilingual event scenario template filling
- Ongoing work & Summary



Latest News About - Climate Change

Page: 1 2 3 4 Next

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Medioambientales en obras de carreteras

on Monday, January 17, 2011 4:19:00 PM CET | info | Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Medioambientales, cuya elaboración contempla la ejecución de las obras con compromisos en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma...

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

on Monday, January 17, 2011 4:05:00 PM CET | info | With 1.8 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

on Monday, January 17, 2011 4:03:00 PM CET | info | ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Regulatory Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

Articles published more than 30 minutes ago

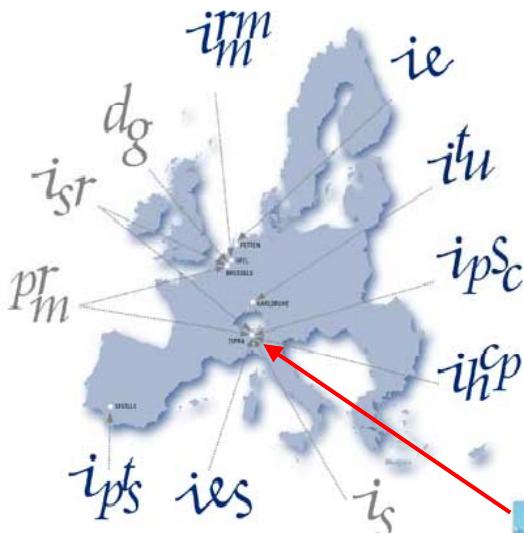
Gas drilling done 'by the book'

on Monday, January 17, 2011 4:02:00 PM CET | info | A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe...

At least 39 dead in South African floods

Johannesburg, Johannesburg Metropolitan Council, Gauteng, South Africa
Lat/Lon: -26.1441, 28.0562 Size: 39

Last update: 2011-01-17T16:00:00Z
JOHANNESBURG (Reuters) - Floods have killed or left missing at least 39 people in South Africa and destroyed thousands of homes in neighbouring Mozambique, officials said on Monday.
Event Type:Floods. Severity: 39 killed, 0 injured and 0 victims were At least 39 killed and not reported injured.



BRUSSELS (BE)
[The Directorate General \(DG\)](#)
[The Institutional and Scientific Relations Directorate \(ISR\)](#)
[The Programme and Resource Management Directorate \(PRM\)](#)

GEEL (BE)
[The Institute for Reference Materials and Measurements \(IRMM\)](#)

KARLSRUHE (DE)
[The Institute for Transuranium Elements \(ITU\)](#)

ISPRA (IT) Download the Ispra site Brochure ([English](#) - [Italian](#))
[The Institute for the Protection and Security of the Citizen \(IPSC\)](#)
[The Institute for Environment and Sustainability \(IES\)](#)
[The Institute for Health and Consumer Protection \(IHCP\)](#)
[The Ispra site Directorate \(IS\)](#)

PETTEN (NL)
[The Institute for Energy \(IE\)](#)

SEVILLE (E)
[The Institute for Prospective Technological Studies \(IPTS\)](#)



- European Commission (scientific-technical arm of public administration)
- Non-commercial
- Multi-disciplinary / multilingual
- Relatively small team working on Language Technology and media monitoring

EMM media monitoring users – wide coverage, world-wide



- European Commission (most DGs) and other EU Institutions
- EU Agencies:
 - e.g. Public Health (ECDC), Food Safety (EFSA), Chemicals Bureau (ECHA), etc.
- EU Member State organisations: e.g.
 - Public Health,
 - law enforcement authorities,
 - parliaments,
 - crisis management/humanitarian
- International and extra-European organisations: e.g.
 - various UN organisations
 - Centres for Disease Prevention and Control in the US, Canada, China, ...
- The public:
 - Ca. 20 - 30,000 anonymous internet users of publicly accessible EMM systems.
 - Combined between 1 and 2 Million hits per day



- ~ 2500 Sources (world-wide, with focus on Europe)
 - ~ 2300 news sources (web portals)
 - ~ 200 specialist medical sites
 - ~ 20 commercial newswires
 - Specialist pay-for sources (LexisMed)
 - 24/7, updated every 10 minutes
- ~ 100,000 articles / day in ~ 50 languages
- Converts dirty html with adverts, menus, html tags, 'related stories', etc. into clean and standardised UTF-8 encoded RSS format.
- Articles are fed into the various EMM applications:



The screenshot shows a news article titled "Mistanke om rabies hos svenska flagermus" (Suspicion of rabies in Swedish bats) from the website "www.dagensmedicin.se". The article discusses a suspicion of rabies in bats found in Sweden. The interface includes a sidebar with navigation links like "Hem", "Nyhet", "Sport", "Analys", "Politik", "Doktor", and "Om oss".

Steinberger Ralf, Bruno Pouliquen & Erik van der Goot (2009). **An introduction to the Europe Media Monitor Family of Applications**. In: Fredric Gey, Noriko Kando & Jussi Karlgren (eds.): Information Access in a Multilingual World - Proceedings of the SIGIR 2009 Workshop (SIGIR-CLIR'2009), pp. 1-8. Boston, USA, 23 July 2009.

Agenda

- JRC: Who we are – what we do – our customers.
- Europe Media Monitor (EMM) family of applications
 - Publicly accessible at <http://emm.newsbrief.eu/overview.html>
- **Motivation for multilingual text processing**
- How to get access to this complementary information
 - Multilingual category definitions and alerts
 - Linking of related news across languages
 - Multilingual information gathering on named entities
 - Multilingual event scenario template filling
- Ongoing work & Summary

The screenshots show the EMM NewsExplorer and MedSys interfaces. The NewsExplorer screen displays a world map with news items and a pie chart of news sources by country. The MedSys screen shows a bar chart of articles by category and a map of Europe with specific news items marked.

Latest News About - Climate Change

Page: 1 2 3 4 Next

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Medioambientales en obras de carreteras

Monday, January 17, 2011 4:17:00 PM CET | info | Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Medioambientales, cuya elaboración contempla la mejora de las condiciones ambientales en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma...

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

Tuesday, January 18, 2011 4:05:00 PM CET | info | With 1.8 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

Wednesday, January 19, 2011 4:05:00 PM CET | info | ISLAMABAD, Jan 19 (APP): The Competition Commission of Pakistan (CCP) has taken notice of exclusive rights granted by the Pakistan Electronic Media Rights Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

Articles published more than 30 minutes ago

Gas drilling done 'by drilling' to ensure clean energy for all

Wednesday, January 19, 2011 4:02:00 PM CET | info | With a million people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

Wednesday, January 19, 2011 4:05:00 PM CET | info | ISLAMABAD, Jan 19 (APP): The Competition Commission of Pakistan (CCP) has taken notice of exclusive rights granted by the Pakistan Electronic Media Rights Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

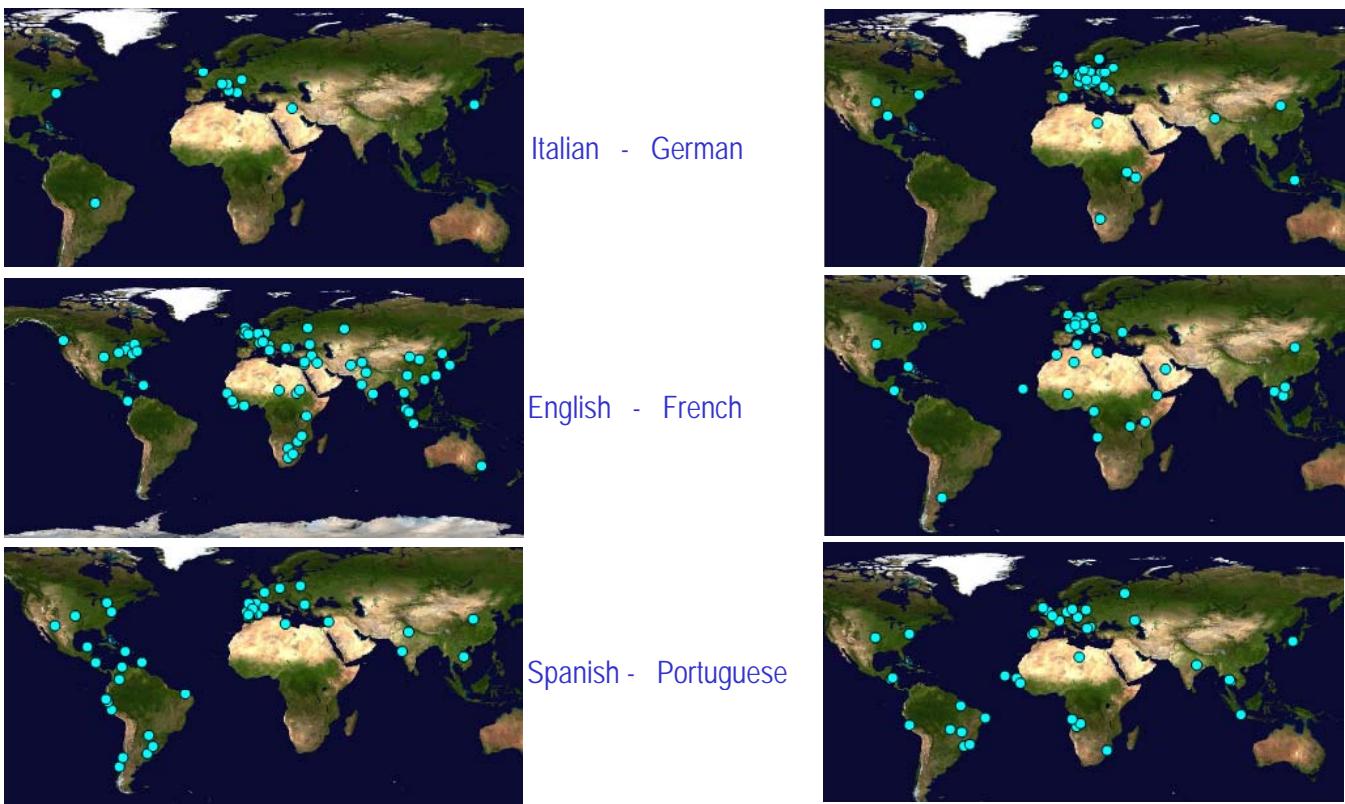
Articles published more than 30 minutes ago

Gas drilling done 'by drilling'

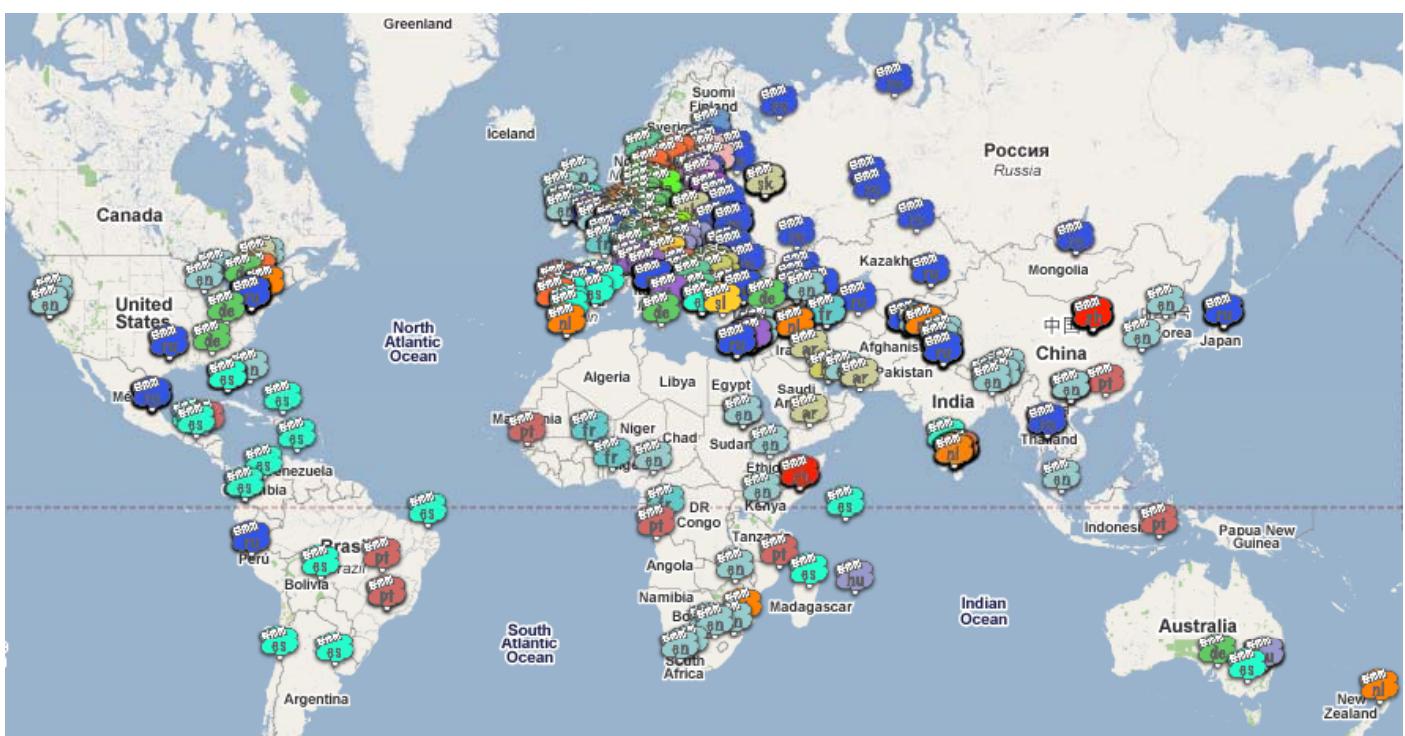
Wednesday, January 19, 2011 4:02:00 PM CET | info | A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe...

The screenshot shows a map of Africa and South America with several news items marked. One item in South Africa is highlighted with a callout box: "At least 39 dead in South African floods" from "Johannesburg.Johannesburg Metropolitan Council, Gauteng, South Africa" with coordinates "Lat/Lon: -26.1441,28.0562 Size: 39". The callout box also includes the text "Last update: 2011-01-17T16:00:00Z" and "JOHANNESBURG (Reuters) - Floods have killed or left missing at least 39 people in South Africa and destroyed thousands of homes in neighbouring Mozambique, officials said on Monday. Event Type:Floods. Severity: 39 Killed, 0 Injured and 0 Victims were At least 39 killed and not reported injured."

Locations mentioned in MediSys medical articles across languages – [complementary coverage](#)



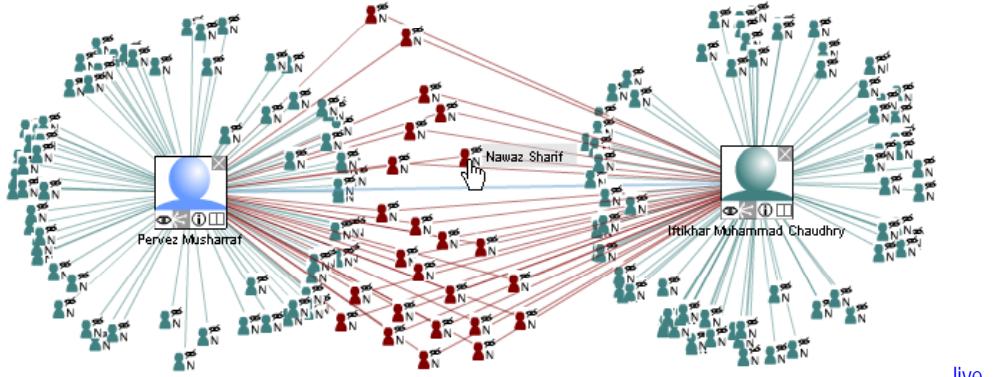
Display of latest geo-located news clusters





Co-occurrence relation between people produced on the basis of many languages is **less biased**.

Associated People
Salman Bashir (1.9)
Джон Нэрронте (1.5)
Nawaz Sharif (1.2)
Раджа Омар Хатаб (1.2)
Мухоммад Али Джинны (1.2)
Зульфикара Али Бхutto (1.2)
Iftikhar Muhammad Chaudhry (1.1)
Christian College (1.1)
Tariq Azeeem (1.1)
Benazir Bhutto (1.1)
Malik Mohammad Qayyum (1.0)
Chaudhry Shujaat Hussain (1.0)
Furqan Bahadur (1.0)
Javed Cheema (0.9)
Shaukat Aziz (0.9)
Abdul Rashid Ghazi (0.9)
Amir Mir Lahore (0.9)
Amin Fahim (0.9)
Гордон Джонроу (0.9)
Wajihuddin Ahmed (0.9)
Mohammed Ali Durrani (0.9)
Rashid Qureshi (0.9)
Oazi Hussain Ahmed (0.9)

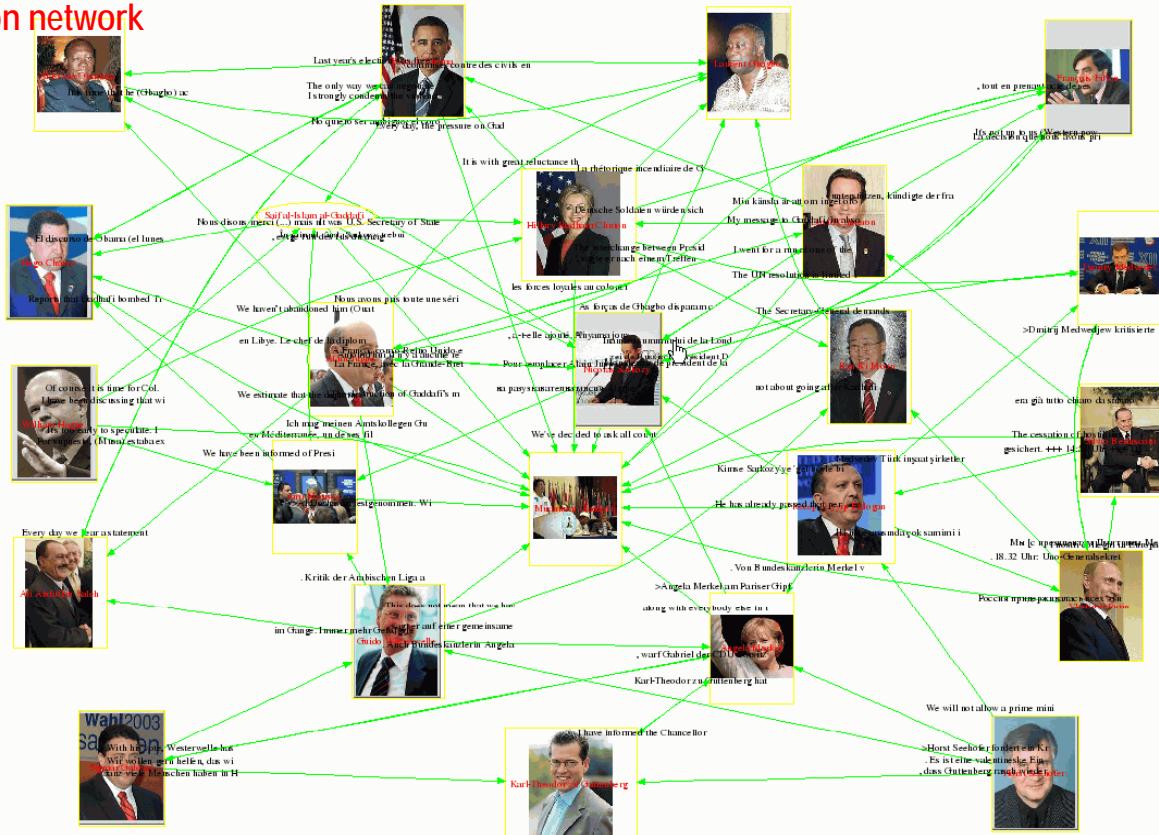

[live](#)

Pouliquen Bruno, Ralf Steinberger, Jenya Belyaeva (2007). **Multilingual multi-document continuously updated social networks**. Proceedings of the Workshop Multi-source Multilingual Information Extraction and Summarization (MMIES'2007) held at RANLP'2007, pp. 25-32. Borovets, Bulgaria, 26 September 2007.(PDF)

Hristo Tanev (2007). **Unsupervised Learning of Social Networks from a Multiple-Source News Corpus**. Proceedings of the Workshop Multi-source Multilingual Information Extraction and Summarization (MMIES'2007) held at RANLP'2007, pp. 33-40. Borovets, Bulgaria, 26 September 2007. (PDF)

Multilinguality: less-biased centrality in social networks

Quotation network


[live](#)

Alexander Litvinenko

Information about this person was last updated on Dienstag, 20. März 2007.

Names	Key Titles and Phrases	External resources
Alexander Litvinenko (Eu,nl)	russo (it,pt - 349)	 <p>Image obtained automatically from Wikipedia</p>
Alexander Litwinenko (de)	agent russe (fr - 134)	
Alexandre Litvinenko (fr)	ruso (es - 208)	
Aleksandr Litvinenko (fi,no)	agenten (de,sv - 134)	
Aleksander Litvinenko (nl,sv)	kritikers (de - 79)	
Александра Литвиненко (ru)	agent (en,sv - 130)	
Александр Литвиненко (ru)	russa (it,pt - 76)	
Alexander Litvinenko (it)	agent secret russe (fr - 39)	
Alexander V. Litvinenko (en)	russe (de,fr - 73)	
Alexandr Litvinenko (it)	former russian agent (en - 20) ←	
Alexander Litvineko (es)	morte di (it - 45)	
Alexandre Livinenko (fr)	rysk agenten (sv - 13) ←	
Alexander Litvenenko (en)	kritiker (de - 19) ←	
亞歷山大·利特維年科 (zh)	43 ans (fr - 17)	
Oleksandr Lytvynenko (en)	russi (it - 14)	
Olexandre Litvinenko (fr)	russian (en - 15)	
Aleksandar Litvinjenko (hr)	omicidio di (it - 11) ←	
Alexander Litvinen (it)	officer (en - 13) ←	
アレクサンダー・リトビネンコ (ja)	former (en - 16)	
Alexander Walterowitsch Litwinenko (de)		

[Read Wikipedia entry](#)

Steinberger Ralf & Bruno Pouliquen (2007). **Cross-lingual Named Entity Recognition**. In: Satoshi Sekine & Elisabete Ranchod (eds.), Journal *Linguisticae Investigaciones*, Special Issue on Named Entity Recognition and Categorisation, LI 30:1, pp. 135-162. John Benjamins Publishing Company. ISSN 0378-4169.

Agenda

Latest News About - Climate Change

[Page: 1 2 3 4 Next](#)

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Medioambientales en obras de carreteras

Monday, January 17, 2011 4:19:00 PM CET | [info](#) | Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Medioambientales, cuya elaboración contempla la mejora de las condiciones ambientales en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma...

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

Tuesday, January 18, 2011 4:05:00 PM CET | [info](#) | With 1.8 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

Wednesday, January 19, 2011 4:05:00 PM CET | [info](#) | ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Regulatory Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

Articles published more than 30 minutes ago

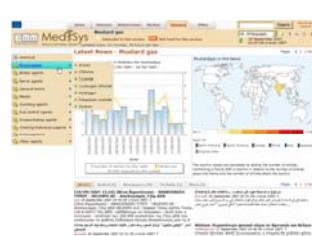
Gas drilling done 'by the book'

Wednesday, January 19, 2011 4:02:00 PM CET | [info](#) | A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe...

At least 39 dead in South African floods

Wednesday, January 19, 2011 4:01:00 PM CET | [info](#) |

A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe...



Multilingual Web Workshop, Pisa, Italy, 4 April 2011



- Public sites: <http://emm.newsbrief.eu/> & <http://medusa.jrc.it/>
- Categorises news into over 1000 categories, using:
 - Boolean search word combinations
 - vicinity operators
 - optional weights
 - regular expressions
- Clusters and tracks news live (multi-monolingually)
- Sends out email notifications for each category
- Detects breaking news
- Lookup of *known* entities
- Quotation recognition

The screenshot shows the EMM NewsBrief interface. On the left, there's a sidebar with a 'Main Menu' containing links like 'Top Stories', '24 hours Overview', 'Events Detection', 'Most Active Themes', 'Mobile Edition', 'Help', 'About EMM', 'Overview', 'Advanced Search', and 'Web Site Map'. Below that is the 'EU Focus' section with links to European Council, Hungarian Presidency, EC News, Council President, EU Foreign Affairs, Representative, Europe 2020, Financial Crisis, and Japan earthquake. The main content area features a 'Top Stories' chart showing article counts over the previous 4 hours. Below the chart are several news stories: 'Japan nuclear plant contaminates sea water' (articles: 139), 'Rebels say they control Gaddafi's hometown' (articles: 119), 'Libya: Nato in full command' (articles: 24), '20 killed in Afghan lorry explosion' (articles: 50), and 'Three suicide bombers kill 20 in Afghanistan'. A 'Country Watch' section highlights Japan as the most active country. The bottom right corner shows the European Commission logo.

MediSys – Filtering and classification in up to 50 languages

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

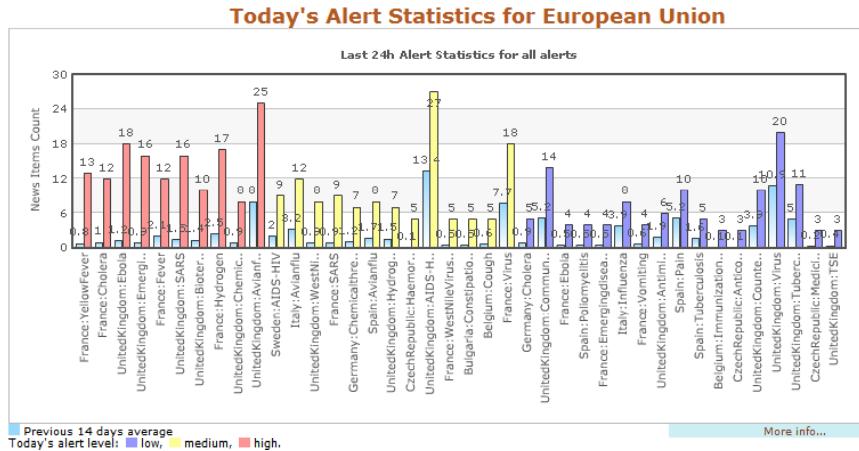


The screenshot shows the MediSys interface. At the top, there's a navigation bar with 'Home', 'Diseases', 'Bioterrorism', 'Nuclear', 'Chemical' (which is highlighted in blue), and 'Other'. Below that is a search bar with dropdowns for 'all - All languages' and a timestamp '18 September 2007 20:47:53 o'clock CEST'. The main content area features a 'Latest News - Mustard gas' section. On the left is a sidebar with a tree icon and a list of chemical agents: 'Blood agents' (highlighted in blue), Blister agents, Nerve agents, General terms, Metals, Vomiting agents, Riot control agents, Incapacitating agents, Choking-Pulmonaryagents, Anticoagulants, and Other agents. To the right is a chart titled 'Alert Statistics for Mustardgas 05/09/2007 - 18/09/2007' showing the number of articles per day. Below the chart is a world map titled 'Mustardgas in the News' with a color scale from light yellow (low) to dark red (high). A legend at the bottom right indicates regions: North America, South America, Europe, Africa, Asia, Australia, and Original View. A note states: 'The country values are calculated by setting the number of articles mentioning a theme AND a country in relation to the number of articles about the theme and the number of articles about the country.'

Access MediSys at
<http://medusa.jrc.it/>

The screenshot continues the MediSys interface. At the bottom left, there's a news alert for '18/09/2007 15:24 [Ξένη δημοσίευση] - ΑΝΑΚΟΙΝΩΣΗ ΤΥΠΟΥ - HELEXPO ΑΕ - Απολογισμός 72η ΔΕΘ' (Athens, 18 September 2007 14:43:00 o'clock CEST). The text discusses the 72nd DEΘ exhibition. At the bottom right, there's another news alert for 'WinGas: Περισσότερο φυσικό αέριο σε Βρετανία και Βέλγιο' (Athens, 17 September 2007 14:46:00 o'clock CEST) from WinGas, BASF, and others. The text mentions a new record for natural gas imports into the UK and Belgium. A note at the bottom right says: 'The country values are calculated by setting the number of articles mentioning a theme AND a country in relation to the number of articles about the theme and the number of articles about the country.'

- Documents from all languages get classified according to the same countries and categories.
 - An increase of the number of media reports on any country-category combination is detected, independently of the reporting language.
 - **Graphs and alerts may show events not yet reported in your own language.**



Ralf Steinberger, Flavio Fuart, Erik van der Goot, Clive Best, Peter von Etter & Roman Yangarber (2008). **Text Mining from the Web for Medical Intelligence**. In: Fogelman-Soulie Françoise, Domenico Perrotta, Jakub Piskorski & Ralf Steinberger (eds.): *Mining Massive Data Sets for Security*, pp. 295-310, IOS Press, Amsterdam, The Netherlands



Multilingual Web Workshop, Pisa, Italy, 4 April 2011

17

EMM NewsBrief

Explore our projects: EMM NewsExplorer | MediSys | EMM Labs

Ecology

UPDATED EVERY 10 MINUTES, 24 HOURS PER DAY.

Main Menu
Top Stories
24 Hours Overview
Event Selection
Most Active Themes
Mobile Edition
Help about EMM
Overview
Advanced Search
Web Site Map

EU Focus
European Council
Hungarian Presidency of the
Council of the European Union
EC News
Council President
EU Foreign Affairs
Regional Policy
Europe 2020
Financial Crisis
Cancun Summit

EU Policy Areas
Agriculture
Audiovisual and Media
Budget
Committee
Consumers
Culture
Customs
Development
Economic and Monetary Affairs
Environment
Health
Industry
Innovation
Justice
Trade

Extracted quotes

Fatih Birol said: "The main challenge from the economic crisis is over but full recovery is threatened by the lack of equilibrium between the cost and supply of oil." "High oil prices are a key risk to the fragile economic recovery among developed nations" khaleejtimes Friday, January 21, 2011 6:02:00 AM CET

Fatih Birol said: "Government subsidies of energy fuels leads to inefficiencies and waste through artificially high use." "Coupled with the current gas glut this poses a real threat to future energy security and renewables. Having said this, energy policy reforms being formulated and implemented by the UAE government need to be commended" khaleejtimes Friday, January 21, 2011 6:02:00 AM CET

Nicholas Stern said: "In order that climate change targets can be

Latest News About - Ecology

Page: 1 2 Next

Articles published more than 40 minutes ago

24 pilot whales die in New Zealand

net Friday, January 21, 2011 7:49:00 AM CET | info

WELLINGTON: New Zealand conservation officials on Friday euthanized 10 pilot whales, the only survivors of a 24-strong pod that became stranded in a mangrove swamp. The whales had been found earlier in the day trapped in the shallow water and mud in Parangarengae Harbor on North Island, with 14 already dead....

Articles published more than 1 hour ago

24 pilot whales die in stranding

NEWScomAU Friday, January 21, 2011 6:49:00 AM CET | info

NEW Zealand conservation officials euthanized 10 pilot whales, the only survivors of a 24-strong pod that became stranded in a mangrove swamp....

Muntinlupa turns green, bans plastic, styrofoam

abs-cbnnews Friday, January 21, 2011 6:44:00 AM CET | info

MANILA, Philippines - If you live in Muntinlupa City, you better bring your own bag the next time you shop....

Articles published more than 2 hours ago

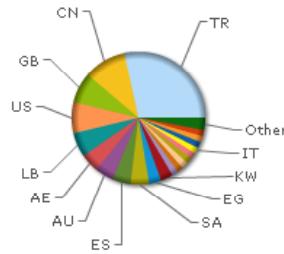
Calif. Plants Put a Wrinkle In Climate Change Plans

NPRnews Friday, January 21, 2011 6:10:00 AM CET | info

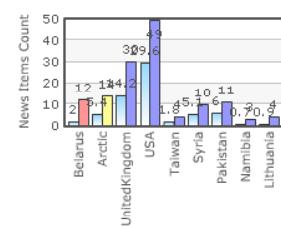
As global temperatures rise, many plants and animals are moving to higher elevations to keep their cool. But a new study finds plants in northern California are actually moving downhill, where it's wetter. "These plants are tracking water availability more so than temperature," one researcher says....

Renewable energy sector has lots of spark left, say Experts

News sources by country

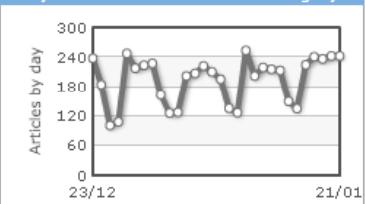


AlertLevel (24h)



Previous 14 days average
Alert level:
high medium low

Daily number of articles in this category



Most associated names

United Nations
Protection Agency
Environmental Protection Agency
Nature
Barack Obama
RIA Novosti
Europa Press
Dmitry Medvedev
World Bank
China Daily

Most reported countries (24h)



Most active sources

sondakika
haberaktuel
stargundem
haber7
yenisafak
webhaber
netgazete
stargazete
crisisgroup
millivet

Latest News About - Climate Change

Page: 1 2 3 4 Next

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Mediambientales en obras de carreteras

Monday, January 17, 2011 4:19:00 PM CET | info

Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Mediambientales, cuya elaboración contempla la necesidad de establecer las normas y procedimientos en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma....

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

UN Monday, January 17, 2011 4:11:00 PM CET | info

With 1.6 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges....

Environmental degradation needs immediate attention: Afridi

ago Monday, January 17, 2011 4:05:00 PM CET | info

ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Regulatory Authority (PEMRA) to Geo Super to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with PEMRA....

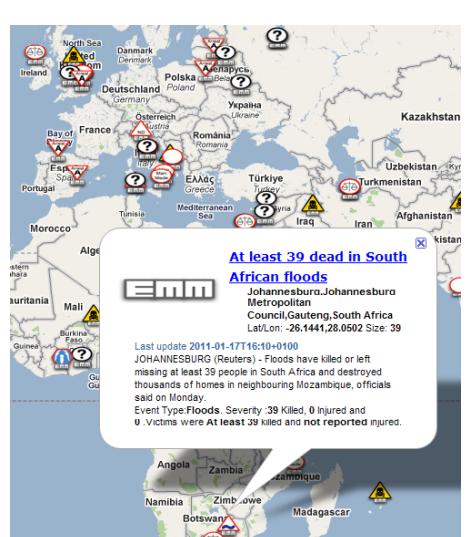
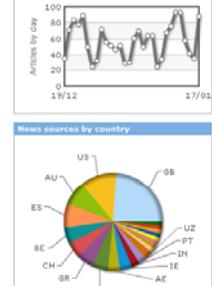
Articles published more than 30 minutes ago

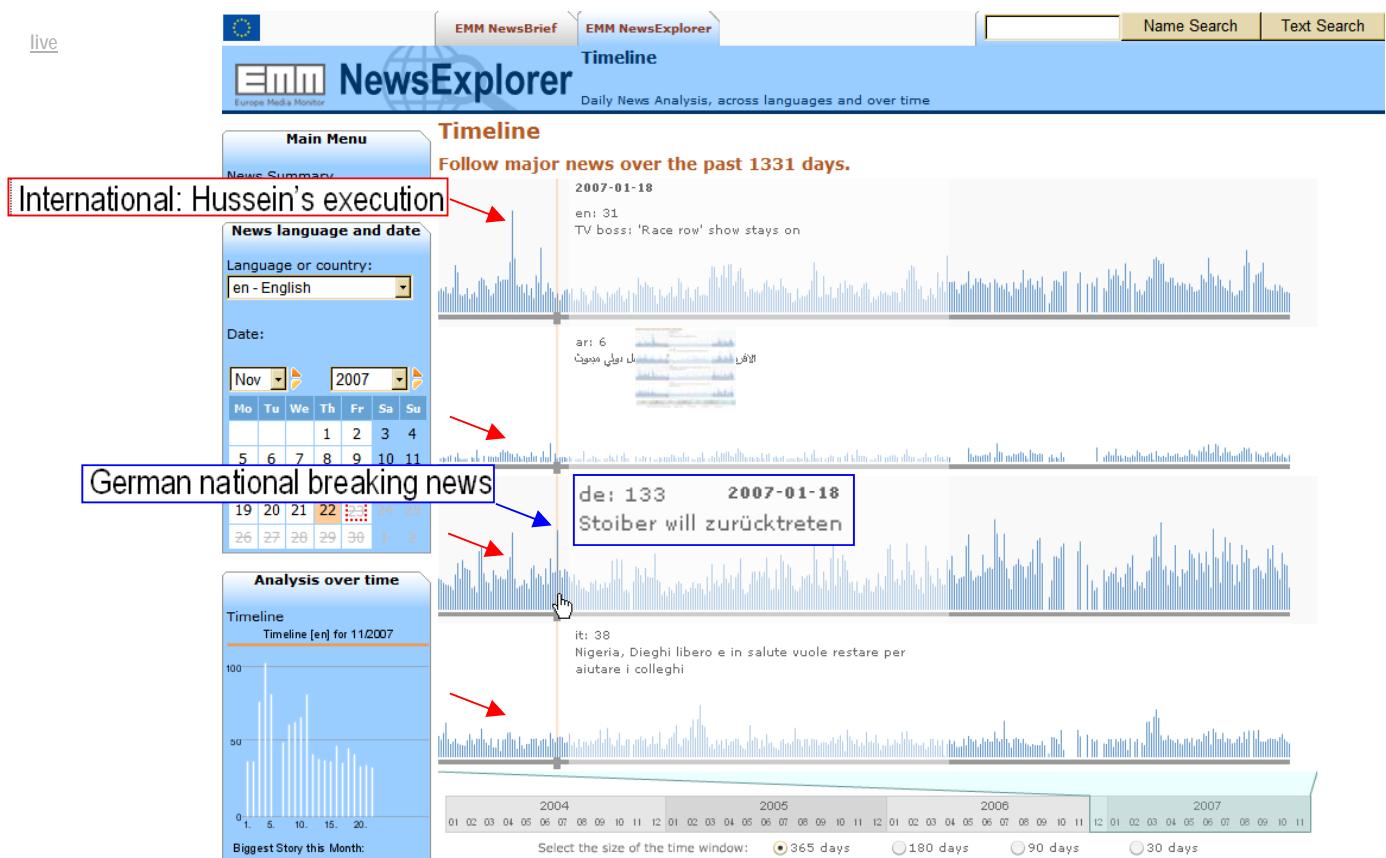
Gas drilling done by 'the book'

bbc Monday, January 17, 2011 4:02:00 PM CET | info

A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe....

Daily number of articles in this category





Thursday, November 22, 2007

Pakistan court dismisses final Musharraf challenge de es fr it nl ar bg da pl pt ru sl sv tr

Pakistan's new-look Supreme Court has, as expected, dismissed the last challenge to President Pervez Musharraf's re-election.
europress-en 2:31:00 PM CET

Did Pakistan Face Rule? A Little Test Says No

Story information

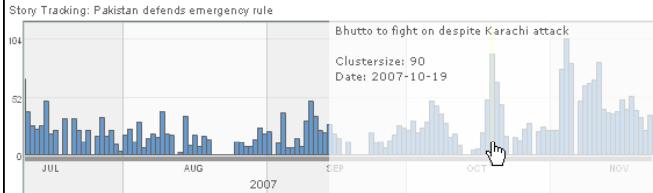
Stories consist of time-linked news clusters with overlapping keywords.

Keywords: Pakistan, Afghanistan, United States / Pervez Musharraf, Supreme Court / pakistani, islamabad, bhutto, military, government, army

Importance: 2833 articles in 125 clusters

Start date: Saturday, June 23, 2007 **End date:** Thursday, November 8, 2007

Timeline



Emergency rule 'destroying the judiciary'
BangkokPost 8:26:00 PM CET

Exiled ex-PM Sharif plots return to Pakistan
usaToday 10:58:00 PM CET

Pakistan Court Rules for Musharraf
ABCnews 2:43:00 PM CET

'Bush unembarrassed by Pakistan emergency'
dailymailUK 12:28:00 AM CET

PM in Uganda for Commonwealth summit
TorontoStar 5:40:00 PM CET

Commonwealth to drop Pakistan
guardian 1:03:00 PM CET

Pakistan's Commonwealth suspension in sorrow,
not anger: Britain (AFP)
news-yahoo 11:34:00 PM CET

Pakistan: Emergency and chaos

Countries

Pakistan (26)
United States (13)
Saudi Arabia (10)

Places

Islamabad(PK)
Rawalpindi(PK)
Karachi(PK)
Washington(US)
Ar Riyad(SA)
Jiddah(SA)

Related People

Pervez Musharraf (26)
Imran Khan (6)
George W. Bush (5)
Gordon Brown (4)
Don McKinnon (4)
Nazir Sharif (4)
Benazir Bhutto (3)
Malik Mohammad Qayum (3)
Iftikhar Muhammad Chaudhry (2)
Aitzaz Ahsan (2)
Ahsan Iqbal (2)
Thomas Jefferson (2)
Asma Jahangir (2)
David Cameron (2)
Louise Arbour (1)
Qazi Hussain Ahmed (1)
Stephen Harper (1)
Najam Sethi (1)
Khalid Hassan (1)
Hina Jilani (1)
Wajihuddin Ahmed (1)
John Negroponte (1)
Rashid Qureshi (1)
Jemima Khan (1)
Benazir Bhutto (1)

Other Names

Supreme Court (27)
Human Rights Watch (3)
Pakistan Muslim League (2)
High Court (2)
High Commission (2)
GEO Television (2)
Human Rights Commission (1)
Al Qaeda (1)

This cluster belongs to the following story: Pakistan defends emergency rule

islamabad, bhutto, military, government, army

Start date: Saturday, June 23, 2007

7 days before.

Musharraf unveils new interim PM Similarity: 0.87

6 days before.

Tendulkar sets up series win over Pakistan Similarity: 0.34

5 days before.

Musharraf swears in caretaker cabinet Similarity: 0.83

4 days before.

US envoy meets Musharraf Similarity: 0.85

Pakistan-India tennis series : Pakistan win second leg to level series 1-1 Similarity: 0.34

3 days before.

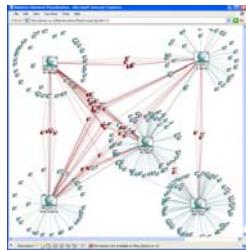
US tells Musharraf to step back Similarity: 0.86

NewsExplorer – Relation exploration



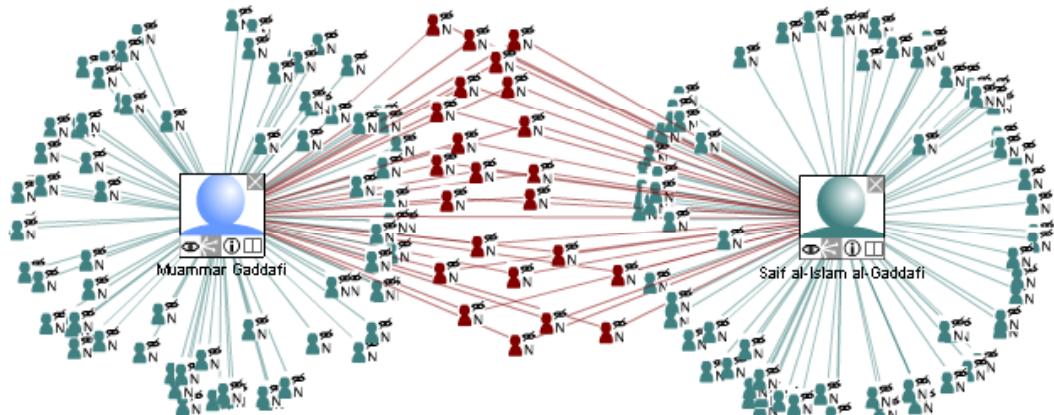
Associated People

- Maciej Krych (1.3)
- Hanibala Gadafija (1.1)
- Donatella Rovera (1.0)
- Saif al-Islam al-Gaddafi (1.0)**
- Richard Lebaron (1.0)
- Roberto Maronim (1.0)
- Devrim Komiteler (1.0)
- Maurizio Farina (1.0)
- Silvio Berlusconiego-oceni (0.9)
- Moustapha Abdel Jalil (0.9)
- Musa Kusa (0.9)
- Ben Jawad (0.8)
- Moustafa Gheriani (0.8)
- Abdel-Rahman Shalqam (0.8)
- Ibrahim Dabbashi (0.8)
- Al-Chuweildi al-Humaidi (0.8)
- Frederick Nyakabwa Atwoki (0.8)
- Bab Azizia (0.8)
- Rolf Porseryds (0.8)
- Musa Ibrahim (0.8)
- Mahmoud Jebril (0.8)
- Ali al-Essawi (0.8)
- Osten Libyens (0.8)
- Mustafa Mohamed Abud al-



live

Example:
Muammar Gaddafi &
son Saif al-Islam al-Gaddafi

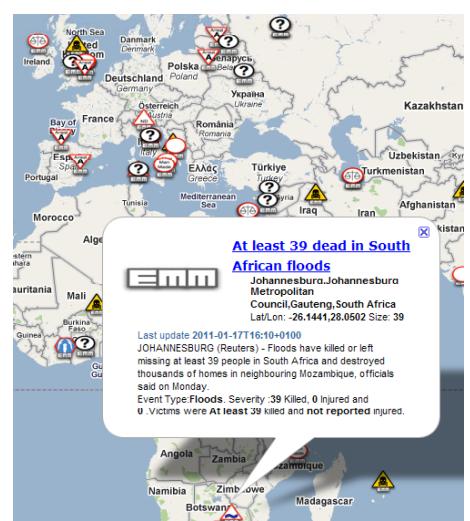
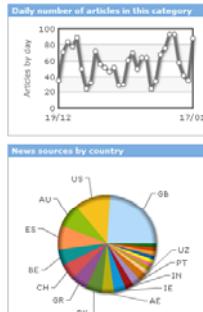


Saif al-Islam al-Gaddafi
Information about this person was last updated on Saturday, January 15, 2011.

Names	Key Titles and Phrases
Saif al-Islam (ar,sv)	son (en,sv - 125)
Saif al-Islam (da,sv)	fils (fr - 63)
سےف الدین (ar)	sohn (de - 46)
Saif al Islam (de,sl)	dirigente libyen (fr - 13)
Saif Al-Islam (en,ro)	zoon (nl - 26)
Cekib an Iscam (bg)	figlio (it - 26)
Saif al Islam (de,sl)	libyen (fr - 13)
سےف الدین (ar)	filho (pt - 22)
Saif al-Islam Gaddafi (da,sv)	hijo (es - 19)
Saif el-Islam (da,nl)	sin (sl - 12)
Saif al-Islam Kadafi (en,pt)	président libyen (fr - 6)
Sayf al-Islam (it,sv)	libyan leader (en - 7)
Saif el Islam (de,pt)	moammar gadhafi's son (en - 4)
Saif el-Islam Kadafi (fr,pt)	sina (sl - 7)
Saif el-Islam (fr,nl)	leader libico (it - 5)
Saif al-Islam Gaddafi (Eu,fr)	fils de #Moammar Kadafi (fr - 3)
Saif Al-Islam (fr)	

Agenda

- JRC: Who we are – what we do – our customers.
- Europe Media Monitor (EMM) family of applications
 - Publicly accessible at <http://emm.newsbrief.eu/overview.html>
- Motivation for multilingual text processing
- How to get access to this complementary information
 - Multilingual category definitions and alerts
 - Linking of related news across languages
 - Multilingual information gathering on named entities
 - **Multilingual event scenario template filling**
- Ongoing work & Summary



Latest News About - Climate Change

Page: 1 2 3 4 Next

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Medioambientales en obras de carreteras

Monday, January 17, 2011 4:19:00 PM CET | info |

Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Medioambientales, cuya elaboración contempla la mejora de las condiciones ambientales en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma...

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

Tue, Monday, January 17, 2011 4:05:00 PM CET | info |

With 1.8 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

Wed, Monday, January 17, 2011 4:05:00 PM CET | info |

ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Rights Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with PEMRA....

Articles published more than 30 minutes ago

Gas drilling done 'by the book'

Thu, Monday, January 17, 2011 4:02:00 PM CET | info |

A firm which is drilling for shale gas in Lancashire seeks to reassure residents and environmentalists that it is safe...



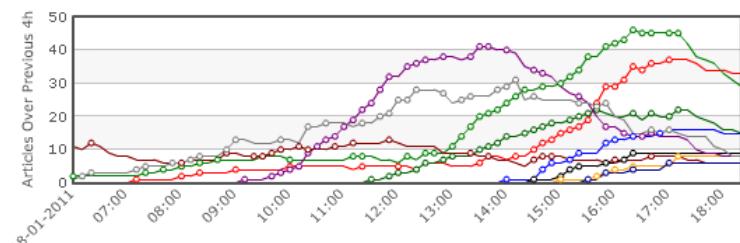
Access NEXUS at: <http://emm-labs.jrc.it/> or

<http://emm.newsbrief.eu/geo?type=event&format=html&language=all>

Atkinson Martin, Jakub Piskorski, Bruno Pouliquen, Ralf Steinberger, Hristo Tanev & Vanni Zavarella (2008). **Online-monitoring of security-related events**. In Proceedings of the 22nd International Conference on Computational Linguistics (CoLing'2008). Manchester, UK, 18-22 August 2008. (PDF)

- **NEXUS:**
Multilingual Information Extraction system
for the extraction of **structured event descriptions**
from online news referring to conflicts, crimes and disasters.
- **Currently 7 Languages:**
English, French, Portuguese, Arabic, Spanish, Italian, Russian (and Chinese).
- **Near real-time:** **every 10 minutes**, EMM clusters the latest articles
about the same event and NEXUS extracts structured information.
- **Objective:**
Global crisis monitoring
(Live situation or long-term trend).

TYPE
PLACE
TIME
DEAD COUNT
DEAD DESCRIPTION
WOUNDED COUNT/DESC
DISPLACED COUNT/DESC
HOMELESS COUNT/DESC
ARRESTED COUNT/DESC
PERPETRATOR
WEAPONS



Baghdad car bombs kill at least 127

Event Type: Terrorist Attack
 Severity: 127 killed 448 injured
 Weapons: car bomb
 Place: Baghdad

Johannesburg: cinq suspects arrêtés pour le meurtre du curé français

Event Type: Arrest
 Severity: 1 killed 0 injured
 Victims: prêtre français/ Louis Blondel killed
 Place: Johannesburg

Police search for killer bus driver

Event Type: Man-Made Disaster
 Severity: 1 killed 6 injured
 Victims: passenger killed
 Place: London

Timor-Leste: Indonésios estão a fazer "cortina de fumo" sobre morte dos "5 de Balibó" - viúva (C/AUDIO)

Severity: 5 killed, 0 injured
 Victims: jornalistas killed
 Place: Timor-Leste.

Car bomber strikes north Pakistan

ech-chorouk-en Tuesday, November 10, 2009 2:23:00 PM CET

A car bomb has exploded in Pakistani's northwestern town of Charsadda killing at least 10 people....

Bomb explodes in northwestern Pakistani town

yediotaharonot Tuesday, November 10, 2009 1:58:00 PM CET

A bomb exploded in the northwestern Pakistani town of Charsadda on Tuesday causing an unknown number of casualties, police said. "It was a bomb blast....

10 killed in Pakistan bomb

RTERadio Tuesday, November 10, 2009 1:57:00 PM CET

A bomb has exploded in the north-western Pakistani town of Charsadda, killing 10 people....

TYPE	Bombing
PLACE	Charsadda, Pakistan
TIME	Tuesday, November 10, 2009
DEAD COUNT	10
DEAD DESCRIPTION	people
WOUNDED COUNT/DESC	
DISPLACED COUNT/DESC	
HOMELESS COUNT/DESC	
ARRESTED COUNT/DESC	
PERPETRATOR	
WEAPONS	Bomb

Event extraction – Text Version

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

31



Main Menu

- [Top Stories](#)
- [24 Hours Overview](#)
- Events Detection**
- [Most Active Themes](#)
- [Mobile Edition](#)
- [Help about EMM](#)
- [Overview](#)
- [Advanced Search](#)
- [Web Site Map](#)

[live](#)

Explore our projects: EMM NewsExplorer | MediSys | EMM Labs

Events Detection

UPDATED EVERY 10 MINUTES, 24 HOURS PER DAY.

Three die as Egypt protests erupt
 Articles: , Last update: , Start:
 Event Type: **Disorder/Protest/Mutiny**, Severity: 3 killed injured and kidnapped. Victims were **Three** killed and injured. Perpetrators were and the weapons used were .
 Place: Cairo. [i](#)

Three die as Egypt protests erupt
 Wednesday, January 26, 2011 6:30:00 AM CET | [info](#) 
 Egyptian police fired tear gas and rubber bullets and beat protesters to clear thousands of people from a central Cairo square after the biggest demonstrations in years against President Hosni Mubarak's authoritarian rule.... [More articles...](#)

Five dead as blast rips through bus
 Articles: , Last update: , Start:
 Event Type: **Bombing**, Severity: 5 killed 13 injured and kidnapped. Victims were **Five** killed and 13 injured. Perpetrators were and the weapons used were .
 Place: Manila. [i](#)

Five dead as blast rips through bus
 Wednesday, January 26, 2011 6:30:00 AM CET | [info](#) 
 The death toll from a bus bombing in the Philippine capital has risen to five with 13 wounded as authorities suspect terrorists are to blame and step up security.... [More articles...](#)

Hamas takes back authorization for Abbas, PLO to negotiate with Israel
 Articles: , Last update: , Start:
 Event Type: **Other**, Severity: killed injured and kidnapped. Victims were killed and injured. Perpetrators were and the weapons used were .
 Place: Jerusalem. [i](#)

Hamas takes back authorization for Abbas, PLO to negotiate with Israel
 Wednesday, January 26, 2011 6:30:00 AM CET | [info](#) 
 Hamas movement on Tuesday withdrew its authorization for Palestinian President Mahmoud Abbas and Palestine Liberation Organization (PLO) to negotiate with Israel on behalf of the Palestinians. "The authorization that Abbas got in the National Agreement Document is no longer existed and we will not....." [More articles...](#)

Hawker suspect publishes memoir
 Articles: , Last update: , Start:
 Event Type: **Arrest**, Severity: 1 killed injured and kidnapped. Victims were **British woman** killed and injured. Perpetrators were and the weapons used were .

Hawker suspect publishes memoir
 Wednesday, January 26, 2011 6:30:00 AM CET | [info](#) 

Event extraction – Display on a map

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

32



Event extraction – Display on a map – click on one event

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

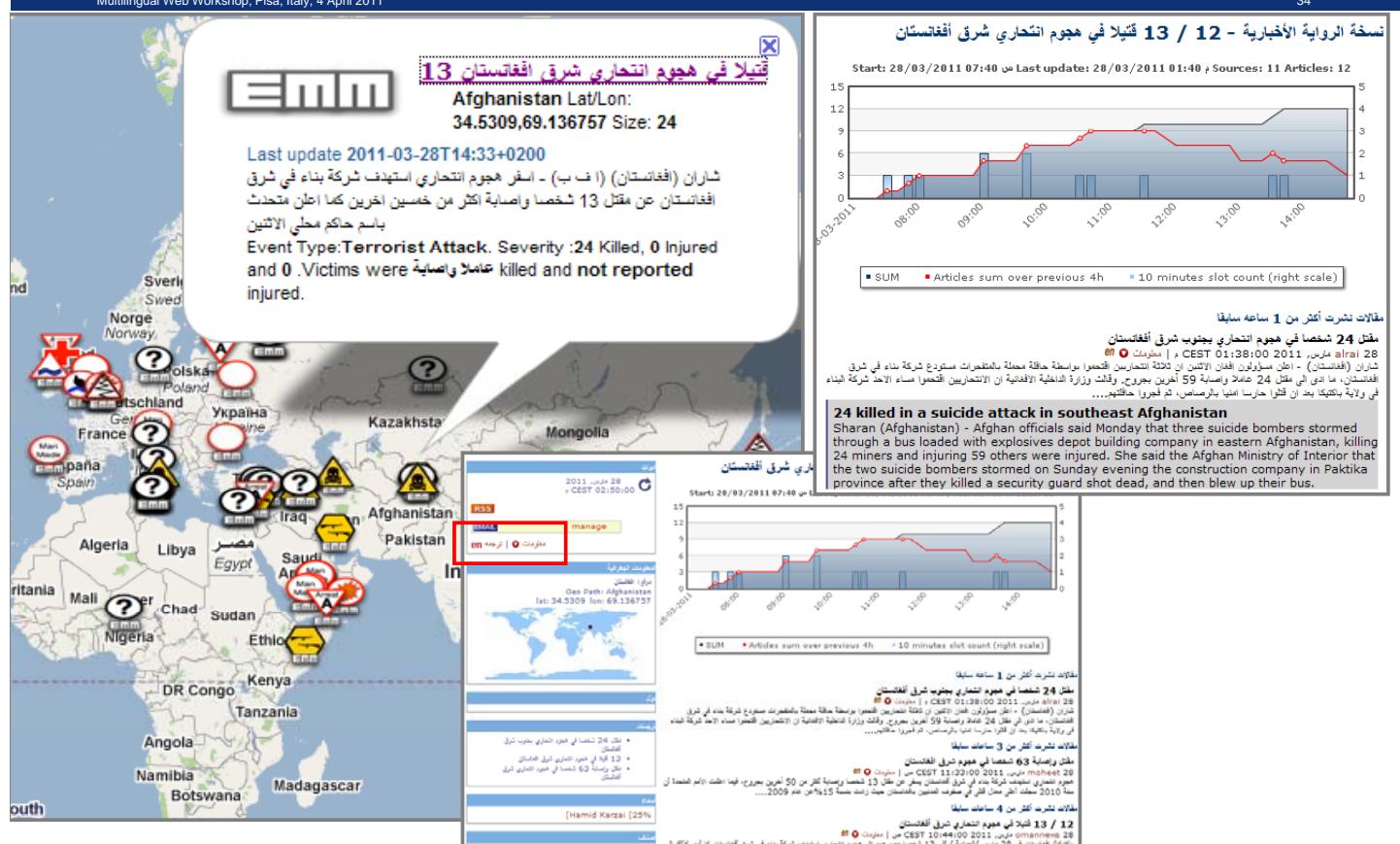
33



Event extraction – View news cluster and translation

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

34

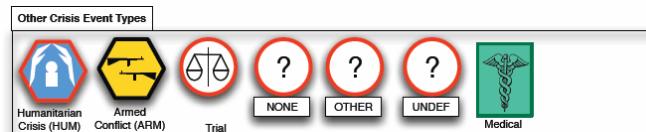
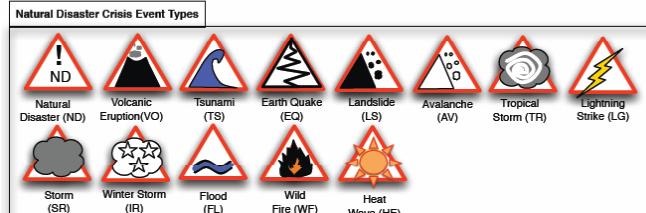
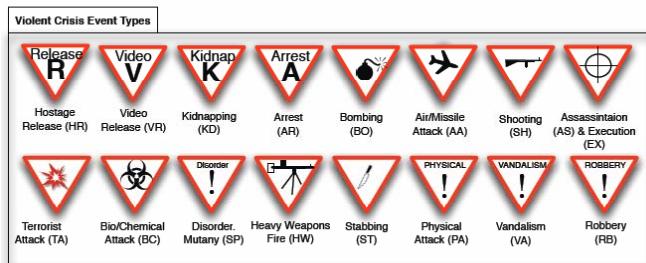


Event types currently recognised

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

35

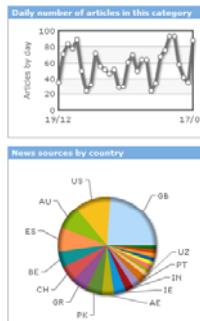
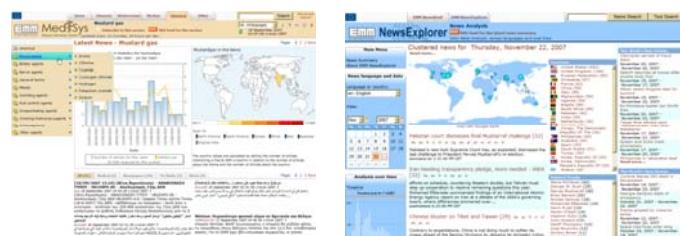
NEXUS Crisis Events



Agenda

Multilingual Web Workshop, Pisa, Italy, 4 April 2011

- JRC: Who we are – what we do – our customers.
- Europe Media Monitor (EMM) family of applications
 - Publicly accessible at <http://emm.newsbrief.eu/overview.html>
- Motivation for multilingual text processing
- How to get access to this complementary information
 - Multilingual category definitions and alerts
 - Linking of related news across languages
 - Multilingual information gathering on named entities
 - Multilingual event scenario template filling
- **Ongoing work & Summary**



Latest News About - Climate Change

Page: 1 2 3 4 Next

Articles published more than 10 minutes ago

El Gobierno de Aragón pone en marcha el Manual de Buenas Prácticas Medioambientales en obras de carreteras

Los Departamentos de Obras Públicas, Urbanismo y Transportes y de Medio Ambiente del Gobierno de Aragón han puesto en marcha el Manual de Buenas Prácticas Medioambientales, cuya elaboración contempla la mejora de las condiciones ambientales en las obras relacionadas con las carreteras de titularidad de la Comunidad autónoma...

Articles published more than 20 minutes ago

Ban calls for global revolution to ensure clean energy for all

With 1.8 billion people in developing countries still lacking access to electricity, Secretary-General Ban Ki-moon today called for a global revolution to address the world's energy challenges...

Environmental degradation needs immediate attention: Afridi

ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Regulatory Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

Articles published more than 30 minutes ago

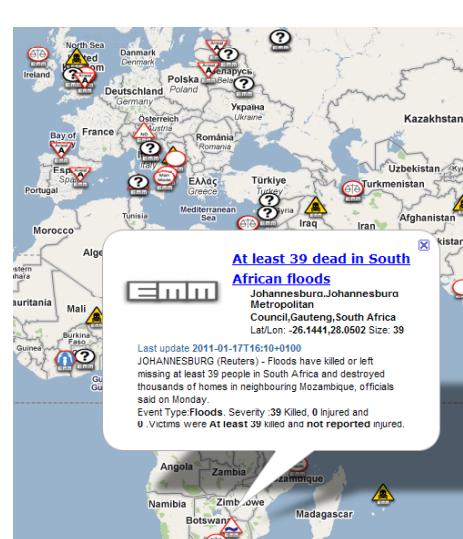
Gas drilling done 'by the book'

ISLAMABAD, Jan 17 (APP): The Competition Commission of Pakistan (CCP) has taken notice of the exclusive rights granted by the Pakistan Electronic Media Regulatory Authority (PEMRA) to "Geo Super" to telecast all Cricket World Cup 2011 matches on cable and satellite under an exclusive agreement with ESPN....

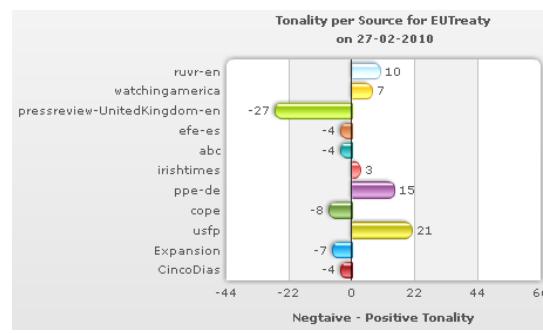
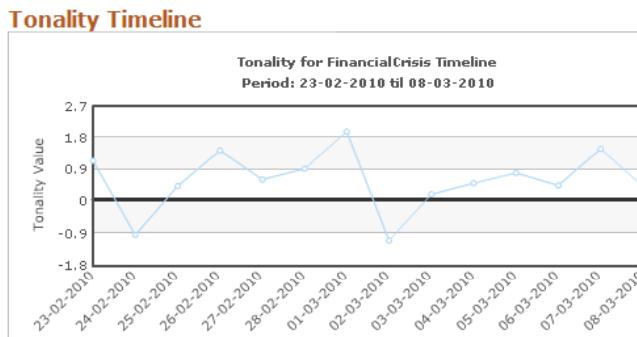
At least 39 dead in South African floods

Johannesburg, Johannesburg, South Africa

Last update: 2011-01-17 16:00:00 Size: 39



- E.g. Detect opinions on
 - European Constitution; EU press releases;
 - Entities (persons, organisations, EU programmes and initiatives);
 - Detect and display **opinion differences** across sources and across countries;
 - Follow trends over time.



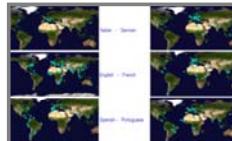
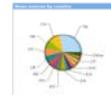
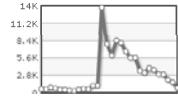
- **Facebook:**
Keyword searches on publicly available posts
e.g. search for *Chikungunya* on openbook.org

extract publicly available friend networks.
 - **Twitter:**
Keyword searches on publicly available tweets
e.g. search for *Chikungunya* on twitter.com
 - **Blogs**

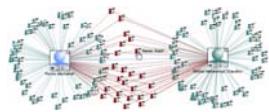


A screenshot of a Facebook page titled "Open Book". The page has a blue header with the title and a green "Share Your Update" button. Below the header, there's a post from "Facebook" itself encouraging users to connect with people in their life. A "More Information" button is also present. The main content area shows several user posts. One post from "Recent Searches" (294 per hour) shows a photo of a man with a red t-shirt and a white camera, with the caption: "Vivek Parolia got tested for Malaria, Dengue, Chikungunya, UTI..... n still wondering what made me so sick" and "2 hours ago". Another post from "Daku Kaul" shows a photo of a man with dark hair, with the caption: "Daku Kaul my dear one effected with "Chikungunya" :-(ms wishes for fast recovery !! 4 hours ago".

- News content (and internet content in general) is complementary across languages.
- EMM gathers and processes multilingual news, etc.
 - Multilingual category definitions and alerts → alert and produce statistics



- Linking of related news across languages
- Multilingual information gathering on named entities



Libyan rebels continue westward push [67] de es fr it nl ar bg da et fa no pl pt ro ru sl sv sw tr

- Multilingual event scenario template filling

TYPE	Bombing
PLACE	Charsadda, Pakistan
TIME	Tuesday, November 10, 2009
DEAD COUNT	10 people
WOUNDED COUNT/DESC	
DISPLACED COUNT/DESC	
HOMELESS COUNT/DESC	
ARRESTED COUNT/DESC	
PERPETRATOR	
WEAPONS	Bomb

