Benefits of multilingual media monitoring

Ralf Steinberger

Ralf.Steinberger@jrc.ec.europa.eu

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

http://langtech.jrc.ec.europa.eu/
http://press.jrc.it/overview.html
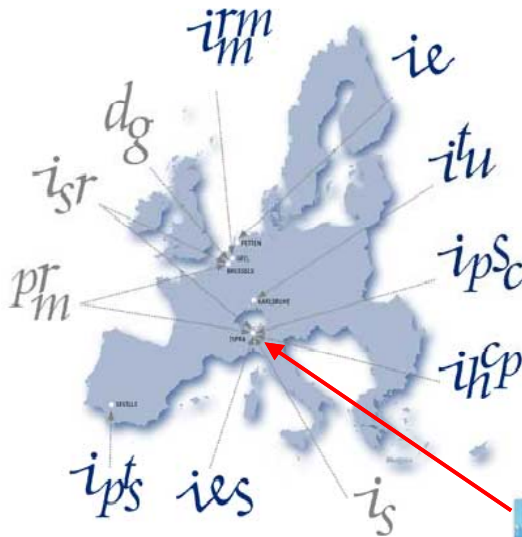

Agenda

- Who we are: European Commission – Joint Research Centre
- Who our users are
- Overview of what we do, what services we provide
  - Europe Media Monitor (EMM) applications   -   Functionality
  - Publicly accessible at http://press.jrc.it/overview.html

- Importance of multilinguality
- User requirements
- Our insights regarding needs

Joint Research Centre - Who we are

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
- Highly multi-disciplinary
- Serving the EU → multilingual
- Relatively small team working on Language Technology

Media Monitoring – Potential Users / Motivation

- ## Any large organisation wants to monitor
  - newest developments in their field of interest
  - their public image, people's views on certain issues, etc.
  - …

- ## Institutions and governments, e.g. EC, UN, National Public Health institutions
  - Security-related events (bombings, hostage takings, piracy, …)
  - Humanitarian events (displaced people, shortage of water/medicines, natural disasters…)
  - Threats to public health (contagious diseases, nuclear or chemical leaks, …)
  - Scientific and technological news
  - …

- ## The public (journalists, specialists, anybody)
  - General information / following the news
  - Meta-news sites are independent of individual (possibly biased) news sources

- **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. 30,000 anonymous users of publicly accessible EMM systems.
  - Combined between 1 and 2 Million hits per day
  - (http://press.jrc.it/overview.html)

Functionality of EMM systems. What we do

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.

- EMM news gathering engine
  - Monitors ~ 2,200 news sources
  - Gathers 80,000 – 100,000 news articles per day
  - In about 50 languages
  - Visits some sites every 5 minutes
  - Extracts text from the web page
  - Converts text into Unicode-encoded RSS
  - Feeds the news into the four publicly accessible media monitoring systems

---

- Group related news multi-monolingually

- Categorise news into hundreds of categories, including countries

- Extract information (entities and events)

- Track related news over time and across languages

- Collect multilingual information on entities

- Generate social networks based on info extracted from multilingual news

- Provide tools for human moderation and dissemination

- Public site: http://emm.newsbrief.eu/   (online since 2002)
- Categorises news into ~ 600 categories, using:
  - Boolean search word combinations
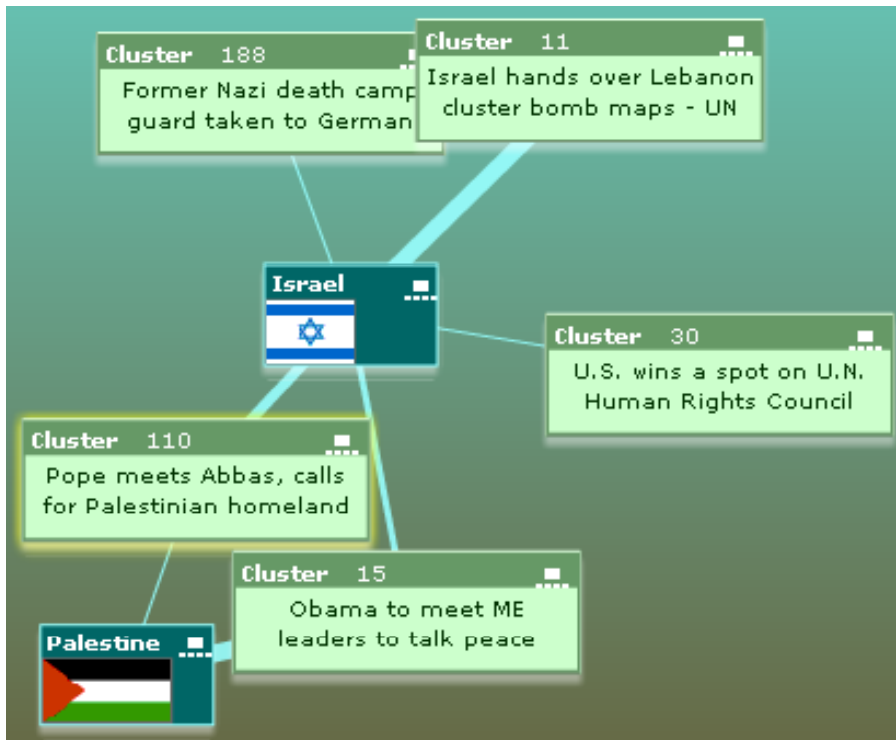 - vicinity operators
  - optional weights
  - regular expressions

- Clusters and tracks news *live* (multi-monolingually)
- Detects breaking news
- Sends out email notifications for each category
- Also accessible as RSS, KML

Live display of latest geo-located news clusters
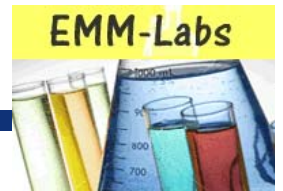
Display of relations between clusters and categories

Aggregation of multilingual information

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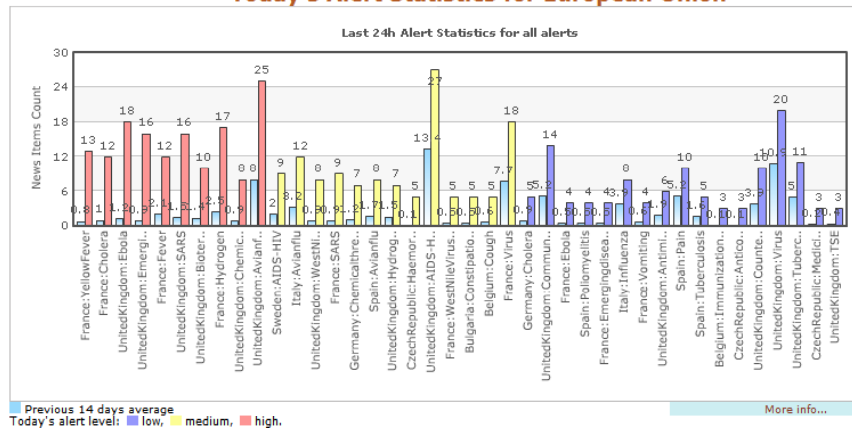
Ralf Steinberger, Flavio Fuart, Erik van der Goot, Clive Best, Peter von Etter & Roman Yangarber (2008). **Text Mining from the Web for Medical Intelligence.** In: Fogelman-Soulié Françoise, Domenico Perrotta, Jakub Piskorski & Ralf Steinberger (eds.): Mining Massive Data Sets for Security. pp. 295-310.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

Detection and visualisation of events (violence/disasters/humanitarian/...)

Objective: global crisis monitoring.   **Languages**: En, Fr, Es, It, Pt, Ru + (Ar)

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)

Detection and visualisation of events (violence/disasters/humanitarian/…)

EMM-Labs

В Турции террористы захватили 15 заложников

Turkey Lat/Lon: 39.9293,32.8533 Size: 18

Last update 2009-05-13T17:35+0200

Двое неизвестных в масках совершили в среду вооруженный налет на филиал банка в турецком курортном городе Кушадасы на побережье Эгейского...

Event Type:Kidnapping/Hostage Taking. Severity :0 Killed, 0 Injured and 15 .Victims were **no one** killed and **no one** injured.

На турецком курорте грабители взяли в заложники посетителей банка

Освобождены захваченные в турецком банке заложники
news-open 13.05.2009 17:40:00 *i*
Перевести:[ar] [bg] [zh] [hr] [cs] [da] [nl] [en] [fi] [fr] [de] [el] [hi] [it] [ja] [ko] [no] [pl] [pt] [ro] [es] [sv]
Турецкая полиция обезвредила налетчика, захватившего в заложники более 10 человек в банке в курортном городе Кушадасы на побережье Эгейского моря, передают местные СМИ. Полиция пошла на штурм банка....

Сдался полиции преступник, захвативший 10 заложников в банке Кушадасы
news-meta 13.05.2009 17:21:00 *i*
Турецкая полиция обезвредила налетчика, захватившего в заложники более 10 человек в банке в курортном городе Кушадасы на побережье Эгейского моря,......

Полиция Турции обезвредила налетчика, захватившего заложников в банке
rian 13.05.2009 16:54:00 *i*
Драма с заложниками продолжалась более шести часов - с 11.00 местного времени (12.00 мск). По сообщению телеканалов NTV и CNN-Turk, полиции удалось убедить налетчика сдаться. Подробности проведенной операции пока неизвестны....

Google™   This page was automatically translated from Russian.
View original web page or mouse over text to view original language.

// World news / / Wednesday, May 13, 2009
Archive | Search

In Turkey, an unknown masked seized 13 hostages in the bank

publication time: 14:29   Images print download submit
Last Updated: 19:08

The police stormed the bank in the western Turkish city of Kusadasi, where the hostages were 13 people. Robbers arrested, ITAR-TASS reported with reference to state television.

NewsExplorer – Functionality

NewsBrief   NewsExplorer   MedISys   EMM-Labs

http://emm.newsexplorer.eu/

Since April 2004

EMM NewsExplorer — News Analysis

Clustered news for Thursday, September 6, 2007

JRC
EUROPEAN COMMISSION

NewsExplorer – Multilingual daily news overview

Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009   17

EMM NewsBrief   EMM NewsExplorer   Name Search   Text Search

EMM NewsExplorer
Europe Media Monitor

News Analysis
RSS feed for the latest news summary
Daily News Analysis, across languages and over time

Clustered news for Thursday, November 22, 2007
Read more...

Main Menu
News Summary
About EMM NewsExplorer

News language and date
Language or country:
en - English

ar - Arabic
bg - български
es - Español
de - Deutsch
da - Dansk
en - English
et - Eesti keel
fa - Farsi
fr - Français
it - Italiano
nl - Nederlands
no - Norsk
pl - Polski
pt - Português
ro - Română
ru - Russian
sl - Slovenščina
sv - Svenska
tr - Türkçe
Countries
AT - Austria
BE - Belgium
DE - Germany
ES - Spain
FR - France
GB - United Kingdom
IT - Italy
NL - Netherlands
US - United States
AC - AFRICA

View with Google Earth

Philippines braced for typhoon
Tens of thousands of people evacuate south of capital Manilla as storm approaches.

Pakistan court dismisses final Musharraf challenge [32]
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.
euronews-en 2:31:00 PM CET

Iran heeding transparency pledge, more needed - IAEA [32] de es fr it tr
efforts on schedule, countering Western doubts, but Tehran must step up cooperation to resolve remaining questions this year. Mohamed ElBaradei summarised findings of an International Atomic Energy Agency report on Iran at a debate of the IAEA's governing board, where differences simmered over....
austrianews 6:23:00 PM CET

Chinese bluster on Tibet and Taiwan [29] de es fr it nl da pt sv
Contrary to expectations, China is not doing much to soften its image ahead of the Beijing Olympics by allowing its domestic critics to speak their minds or championing human rights in Sudan. Instead, Chinese leaders are defending authoritarian rule at home and abroad and waging aggressive diplomacy against those who

Countries
United States (492)
United Kingdom (196)
Russian Federation (83)
Zimbabwe (67)
France (62)
China (58)
Italy (58)
Afghanistan (56)
Uganda (54)
Angola (48)
South Africa (45)
Pakistan (42)
India (38)
Netherlands (38)
Canada (36)
Congo, The Democratic Republic Of The (36)
Philippines (34)
Australia (33)
Spain (33)
Saudi Arabia (31)
Turkey (30)
Korea, Republic Of (30)
Korea, Democratic People's Republic Of (29)
Greece (28)

Related People
Steve McClaren (42)
George W. Bush (29)
Pervez Musharraf (26)
Brian Barwick (20)
Nicolas Sarkozy (18)
Mohamed ElBaradei (14)
Kevin Rudd (14)
Martin O'Neill (14)
Vladimir Putin (14)
John Howard (13)
Geoff Thompson (12)
José Mourinho (12)
Gordon Brown (11)
Hugo Chavez (11)
Mahmoud Abbas (10)
Sven Goran Eriksson (10)
Jaan de Hoon Scheffer (10)

This Week's New Stories
Claimants warned of fraud fears
November 20, 2007 - November 22, 2007
Search resumes at house after double body find
November 15, 2007 - November 17, 2007
Mbeki seeks Mugabe deal for summit
November 15, 2007 - November 22, 2007
Ex-Rhodesia leader Ian Smith dies
November 20, 2007 - November 22, 2007
Tapas Nine witness says 'Mediterranean man took Madeleine'
November 17, 2007 - November 21, 2007
Queen and Duke mark anniversary
November 17, 2007 - November 20, 2007
Philippines in 'separatist deal'
Read more...

This Month's New Stories
Cyclone leaves 242 dead in Bangladesh
November 14, 2007 - November 22, 2007
Georgia declares state of emergency
October 31, 2007 - November 22, 2007
France gripped by massive strike
November 12, 2007 - November 22, 2007
Space crew fixes solar wing
October 23, 2007 - November 14, 2007
Vegas showdown: Clinton 'on the hot seat'
October 23, 2007 - November

JRC
EUROPEAN COMMISSION

NewsExplorer – Cross-lingual cluster linking

NewsExplorer

Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009

Castro quits as president, state-run paper reports [72] de es fr it nl ar bg da
et fa no pl pt ro ru sl sv tr

Fidel Castro announced his resignation as presid of Cuba and commander-in-chief of Cuba's milit on Tuesday, according to a letter published by state-run newspaper Granma.
cnn 9:23:00 AM CET

قدّم تلفزيون فرانسه از كناره گيرى فيدل كاسترو de en fr it nl
شبكه بين المللى فرانس24 در برنامه ويژه اى به مناسبت كناره گيرى فيدل كاسترو از قدرت در كوبا با تحليلگران سياسى خود به گفتگو پرداخت. ژان برنار كاديه تحليلگر سياسى اين شبكه گفت دوره انتقالى پس از فيدل كاسترو در كوبا از مدتى پيش آغاز شده است. در 31 ژوئيه 2006 وى زمام قدرت را به برادرش راؤل كاسترو سپرد و....
iranpressnews 13:36:00 o'clock CET

Kuba: Fidel Castro gibt das Zepter ab en es fr it nl ar bg da et fa no pl pt ro ru sl sv tr
Der legendäre kubanische Staatschef verzichtet laut Online-Ausgabe der kommunistischen Parteizeitung a
kleinezeitung 10.

Fidel Castro går av de en es fr it nl
Partiav som pr styrtet stat på
VG

Fidel Castro renuncia a la Presidencia del Consejo de Estado de en es
et no pl pt ro sl sv tr

A Cuba, Fidel Castro renonce au pouvoir de en es it nl ar bg da c
ru sl sv tr

Cuba, Fidel Castro rinuncia alla presidenza de en es fr nl ar bg d
sl sv tr
L'A
htt

Cubaanse president Fidel Castro afgetreden de en es fr it nl
ro ru sl sv tr

كاسترو يستقيل وبوش يدعو للتحول الديمقراطى فى كوبا de en fr it nl
فى مؤتمر صحفى فى رواندا، إلى مساعدة كوبا على البدء بعملية "انتقال ديمقراطى"، وذلك إثر قرار الزعيم -- (CNN) هافانا، كوبا فيدل كاسترو، بالتنحى عن منصب كرئيس للبلاد. وقال بوش: "إن على المجتمع الدولى أن يعمل مع الشعب الكوبى لبناء بإقامة مؤسسات م
cnnarabic CET 01:21:00 م

Кастро се оттегли от президентския пост de en es fr it nl
Фидел Кастро обяви, че се отказва от президентския пост, съобщава АФП

Fogh vil ikke savne Castro de en es fr it nl
"Polit
berlin

REPLIIK: Castro-aja lõpu algus de en es fr it nl
Üks 20. sajandi menukam vabadus-võitleja ja tuntum diktaator Fidel C kui Nõukogude "gerondid", kelle jaoks tähendas lahkumine võimult ka võim kestab vähemalt esiotsa edasi, sest esimeseks asendajaks peeta
epl 23:17:00 CET

Fidel Castro zrezygnował! de en es fr it nl
Przywódca kubański Fidel Castro po 49 latach rządów zrezygnował we wtorek z funkcji przewodniczącego Rady Państwa Kuby.

Fidel Castro renunciou à presidência de Cuba de en es fr it nl
Anúncio no órgão oficial do Partido Comunista cubano Fidel Castro anunciou hoje que se retira da

Fidel Castro se retrage de la presedintia Cubei de en es fr it nl
Fidel Castro a anuntat, marti, ca renunta la presedintia Cubei, in editia electronica a cotidianului

Fidel Castro se je odpovedal položaju kubanskega predsednika de en es fr it nl
Kubanski voditelj Fidel Castro je danes sporočil, da se odpoveduje položaju predsednika države. Kot je Castro še zapisal v sporočilu, objavljenem v spletni izdaji uradnega glasila Granma, se ne poteguje in

Värmby skakade hand med Fidel Castro. de en es fr it nl
- Ja. Jag har inte tvättat högernäven sedan dess, 1983. Jag var på en stor sammandragning på Kuba med uppmaningen att USA skulle häva blockaden mot landet.
smp kl 19:51 CET

Фидель Кастро отказался от поста председателя Госсовета Кубы de en fr it
ГАВАНА, 19 февраля. /ИТАР-ТАСС/. Фидель Кастро отказался от поста главы государства и правительства - председателя Государственного совета Кубы. Об этом он сообщил в обращении к народу
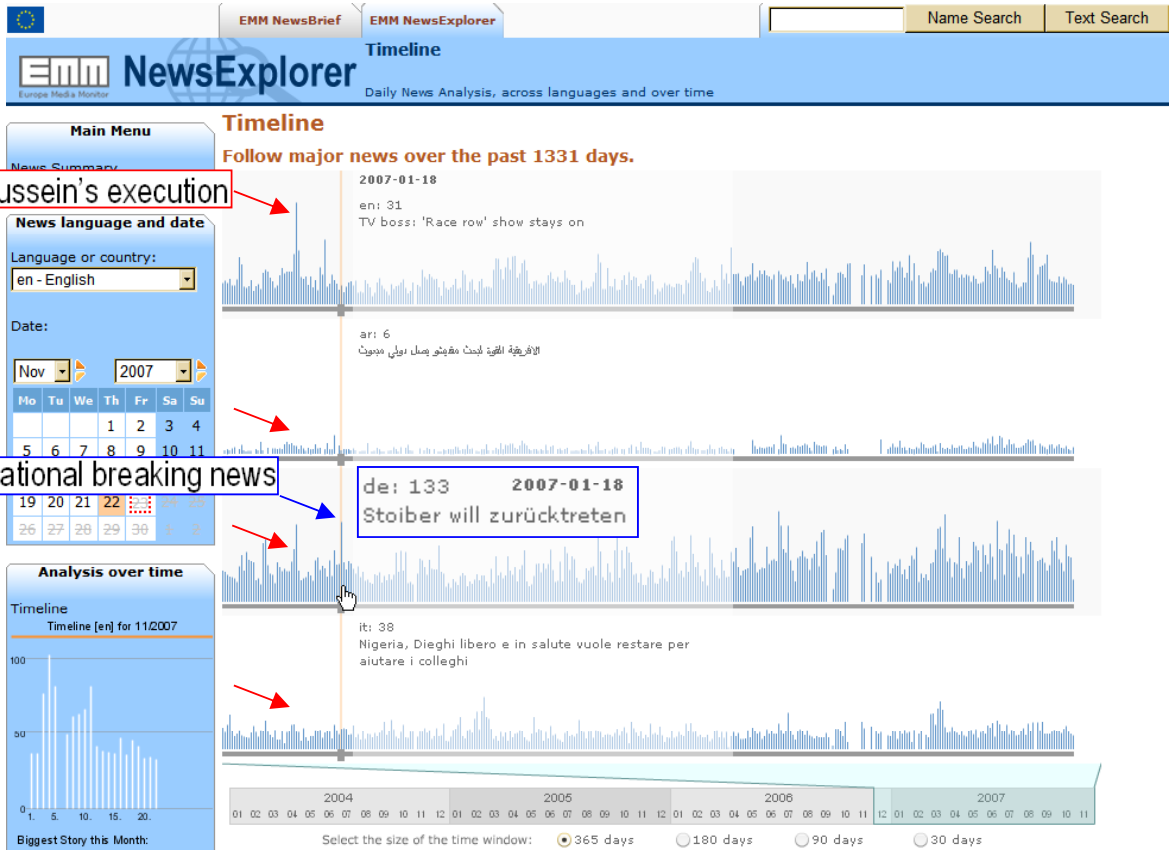
Bir dönemin sonu de en es fr it nl
Küba Komünist Partisi'nin yayın organı Granma'ya açıklama yapan Castro, devlet başkanlığına geri dönmeyeceğini belirtti. Fidel Castro 1959 yılından beri ülkeyi yönetiyordu. Ancak 2006'da geçirdiği ağır ameliyattan beri iktidar koltuğundan uzak kaldı. Ülke yönetimine, ağabeyi Fidel Castro'ya vekalet eden Raul Castro bakıyordu.
hurriyetim 10:15:00 CET

NewsExplorer – Time line: biggest clusters per day

International: Hussein's execution

German national breaking news

NewsExplorer – Aggregation of clusters into longer 'stories'

live

NewsExplorer –Information about people
collected from multiple languages and over time

JRC EUROPEAN COMMISSION

live

Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009

Pervez Musharraf
Information about this person was last updated on Friday, November 23, 2007.

Names
Pervez Musharraf (Eu,sv)
General Pervez Musharraf (da,sv)
Gen Musharraf (en)
Pervez Musharaf (sl)
Gen Pervez Musharraf (en)
Pervez Musharrafs (da,sv)
Pervez Musharaf (da,sv)
Первез Мушарраф (ru)
Perwez Musharraf (fr,sv)
Pervez Moucharraf (fr)
برويز مشرف (ar)
Perveza Mušarafa (sl)
Pervez Muscharraf (de)
Perves Muscharraf (de)

Key Titles and Phrases
pakistani president (en - 827)
president (de,sv - 3230)
president gen (en - 506)
président pakistanais (fr - 398)
pakistaanse president (nl - 235)
presidente paquistanês (pt - 277)
pakistani president gen (en - 181)
presidente paquistaní (es - 147)
präsident (de - 617)
general (en,sv - 450)
præsident (da - 210)
presidente (es,pt - 589)
president, gen (en - 62)

External resources

Related People
Benazir Bhutto (910)
Nawaz Sharif (552)
George W. Bush (537)
Iftikhar Muhammad Chaudhry (502)
Osama bin Laden (430)
Shaukat Aziz (395)
Tariq Azeem (251)

Other Names
Al Qaeda (855)
Supreme Court (505)
White House (312)
NATO (207)
GEO Television (197)
Lal Masjid (187)
Tautas Partija (176)
Daily Times (153)

Associated People
Salman Bashir (1.9)
Джон Негропонте (1.5)
Nawaz Sharif (1.2)
Раджа Омар Хатаб (1.2)
Мухоммада Али Джинны (1.2)
Зульфикара Али Бхутто (1.2)
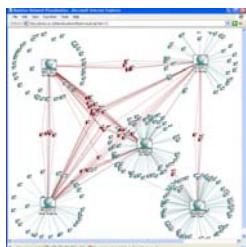Iftikhar Muhammad Chaudhry (1.1)

Related Stories
Pakistan defends emergency rule
June 23, 2007 - November 22, 2007
Pakistan's president urges calm
May 5, 2007 - June 23, 2007
Troops launch hostage rescue bid
June 23, 2007 - October 25, 2007
Protests in Pakistan take aim at President Musharraf
March 10, 2007 - April 3, 2007
42 die in bomb attacks on 'lucky' weekend
July 15, 2007 - October 25, 2007
Demonstrations at UK embassy in Iran

Latest Clusters - English
[de] [pt] [es] [nl] [fr] [ar] [sv] [it] [da] [pl] [sl] [ro] [ru] [no] [bg]
Pakistan court dismisses final Musharraf challenge Struggle to help cyclone surivors in Bangladesh
euronews 22-NOV-07   cnn 21-NOV-07
Musharraf 't
cnn 21-NOV-

Quotes from - English
[es] [ru] [sv] [pt] [nl] [de] [fr] [no] [bg] [it]
[-has said]: Where it fails to live up to those values, it needs to act with credibility and consistency. I think Pakistan is a test in that respect.
bday 22-NOV-0
[-has repeated]: The (presidential) oath can be taken ... by the weekend or immediately thereafter.

Quotes about - English
[es] [bg] [de] [pt] [ro] [fr] [no] [nl]
[said]: The m
to improve the
instead,
Rice [-said]: And look, a lot of that was done by (Pervez) Musharraf himself. And so for him at this point to help put his country back on the road to democratic reform is important. We're looking for him to take off his uniform,
expressindia 22-NOV-07
Brad Adams [said]: It's disgraceful that Musharraf is punishing Chief Justice Cha who challenged his power-grab, by keep judge's family under house arrest,
HumanRightsWatch 22-NOV-07
Brad Adams [said]: Rather than making
Brown [-said]: He (Musharraf) has assured

NewsExplorer – Relation exploration

JRC EUROPEAN COMMISSION

Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009

Example: Pervez Musharraf & Iftikhar Chaudhry

live

Associated People
Salman Bashir (1.9)
Джон Негропонте (1.5)
Nawaz Sharif (1.2)
Раджа Омар Хатаб (1.2)
Мухоммада Али Джинны (1.2)
Зульфикара Али Бхутто (1.2)
Iftikhar Muhammad Chaudhry (1.1)
Christian College (1.1)
Tariq Azeem (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)

Pervez Musharraf
Nawaz Sharif
Iftikhar Muhammad Chaudhry

- Document **clustering**

- **Named entity recognition** (persons, organisations)

- Name **variant matching**, including across scripts (transliteration, string similarity calculation)

- **Geo-tagging** (recognition and disambiguation of locations)

- Multi-label document **categorisation**

- **Quotation** recognition (and reference resolution for name parts)

- **Social network generation** (based on co-occurrence, quotations, relation types)

- **Topic detection and tracking** (TDT) (Multi-monolingual cluster similarity calculation)

- **Cross-lingual cluster similarity** calculation (for cross-lingual TDT).

Importance of multilinguality

Locations mentioned in MedISys medical articles across languages – **complementary coverage**
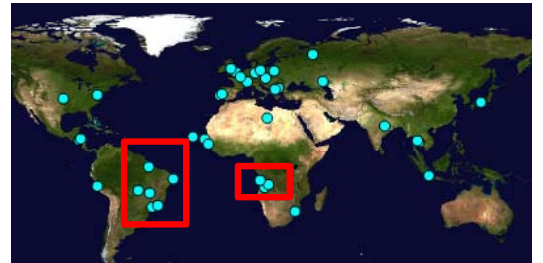

Italian - German

English - French
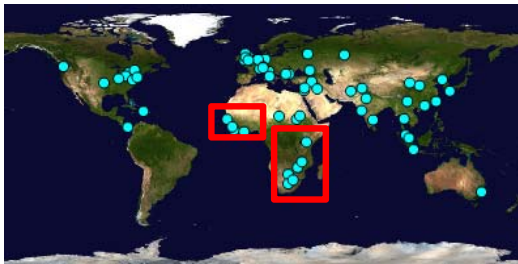
Spanish -   Portuguese

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)	
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 Livtinenko (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)	ryske 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 Litvinenk (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)		

Image obtained automatically from *Wikipedia*

Read Wikipedia entry

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

Multilinguality: less news bias and more transparency

Castro quits as president, state-run paper reports [72] de es fr it nl ar bg da et fa no pl pt ro ru sl sv tr

Fidel Castro announced his resignation as presid of Cuba and commander-in-chief of Cuba's milit on Tuesday, according to a letter published by state-run newspaper Granma.
cnn 9:23:00 AM CET

شبكة بين المللي فرانس24 در برنامه ويژه اى به مناسبت كناره گيري فيدل كاسترو از قدرت در كوبا با تحليلگر سياسي خود به گفتگو پرداخت. ژان برنار كاديه   de en fr it nl
تحليلگر سياسي اين شبكه گفت دوره انتقالي پس از فيدل كاسترو در كوبا از مدتى پيش آغاز شده است. در 31 ژوئيه 2006 وي زمام قدرت را به برادرش راؤول ....كاسترو سپرد و
iranpressnews 13:36:00 o'clock CET

Fidel Castro går av de en es fr it nl

Kuba: Fidel Castro gibt das Zepter ab en es fr it nl ar bg da et fa no pl pt ro ru sl sv tr
Der legendäre kubanische Staatschef verzichtet laut Online-Ausgabe der kommunistischen Parteizeitung a
kleinezeitung 10.

Fidel Castro renuncia a la Presidencia del Consejo de Estado de en fr it et no pl pt ro sl sv tr

Fidel Castro zrezygnował! de en es fr it nl
Przywódca kubański Fidel Castro po 49 latach rządów zrezygnował we wtorek z funkcji przewodniczącego Rady Państwa Kuby

A Cuba, Fidel Castro renonce au pouvoir de en es it nl ar bg da et ru sl sv tr

Fidel Castro renunciou à presidência de Cuba de en es fr it nl
Anúncio no órgão oficial do Partido Comunista cubano Fidel Castro anunciou hoje que se retira da

Cuba, Fidel Castro rinuncia alla presidenza de en es fr nl ar bg sl sv tr

Fidel Castro se retrage de la presedintia Cubei de en es fr it nl
Fidel Castro a anuntat, marti, ca renunta la presedintia Cubei, in editia electronica a cotidianului

Cubaanse president Fidel Castro afgetreden de en es fr it ar ro ru sl sv tr

Fidel Castro se je odpovedal položaju kubanskega predsednika de en es fr it nl
Kubanski voditelj Fidel Castro je danes sporočil, da se odpoveduje položaju predsednika države. Kot je Castro še zapisal v sporočilu, objavljenem v spletni izdaji uradnega glasila Granma, se ne poteguje in

كاسترو يستقيل ويوش يدعو للتحول الديمقراطي في كوبا de en fr it nl
في مؤتمر صحفي في رواندا، إلى مساعدة كوبا على البدء بعملية "انتقال ديمقراطية"، وذلك إثر قرار الزعيم -- (CNN)هافانا، كوبا
فيدل كاسترو، بالتنحي عن منصبه كرئيس للبلاد. وقال بوش: "إن على المجتمع الدولي أن يعمل مع الشعب الكوبي للبدء بإقامة مؤسسات
cnnarabic CET 01:21:00 م

Värmby skakade hand med Fidel Castro.
– Ja. Jag har inte tvättat högernäven sedan dess, 1983. Jag var på en stor sammandragning på Kuba med uppmaningen att USA skulle häva blockaden mot landet.
smp kl 19:51 CET

Кастро се оттегли от президентския пост de en es fr it nl
Фидел Кастро обяви, че се отказва от президентския пост, съобщава АФП

Фидель Кастро отказался от поста председателя Госсовета Кубы de en fr it
ГАВАНА, 19 февраля. /ИТАР-ТАСС/. Фидель Кастро отказался от поста главы государства и правительства - председателя Государственного совета Кубы. Об этом он сообщил в обращении к

Fogh vil ikke savne Castro de en es fr it nl
"Polit

Bir dönemin sonu de en es fr it nl
Küba Komünist Partisi'nin yayın organı Granma'ya açıklama yapan Castro, devlet başkanlığına geri dönmeyeceğini belirtti. Fidel Castro 1959 yılından beri ülkeyi yönetiyordu. Ancak 2006'da geçirdiği ağır ameliyattan beri iktidar koltuğundan uzak kaldı. Ülke yönetimine, ağabeyi Fidel Castro'ya vekalet eden Raul Castro bakıyordu.
hurriyetim 10:15:00 CET

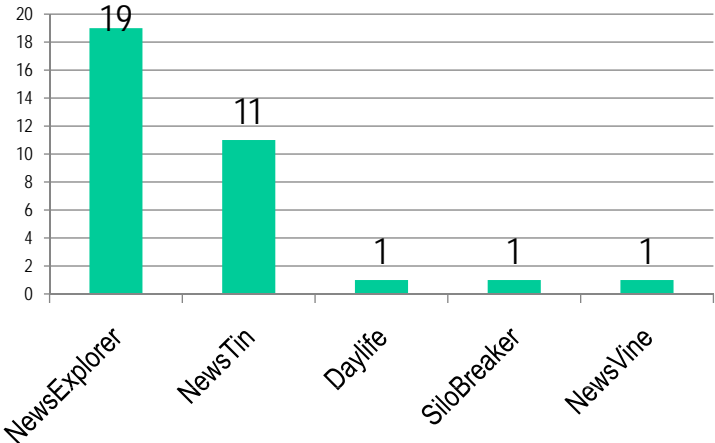REPLIIK: Castro-aja lõpu algus de en es fr it nl
Üks 20. sajandi menukam vabadus-võitleja ja tuntum diktaator Fidel C kui Nõukogude "gerondid", kelle jaoks tähendas lahkumine võimult ka võim kestab vähemalt esiotsa edasi, sest esimeseks asendajaks peeta
epl 23:17:00 CET

Language coverage of various media analysis tools (August 2009)

News aggregators

- Google News: 45
- Yahoo News: 35
- Europe Media Monitor: 50

News analysis systems

- NewsExplorer: 19
- NewsTin: 11
- Daylife: 1
- SiloBreaker: 1
- NewsVine: 1

in early 2008: 34, 17, 43 languages

in early 2008: the same

Social networks produced on the basis of many languages are
more complete and **less biased**.

Associated People

Salman Bashir (1.9)
Джон Негропонте (1.5)
Nawaz Sharif (1.2)
Раджа Омар Хатаб (1.2)
Мухоммада Али Джинны (1.2)
Зульфикара Али Бхутто (1.2)
Iftikhar Muhammad Chaudhry (1.1)
Christian College (1.1)
Tariq Azeem (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)
Qazi 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**)

EMM user requirements
EMM insights

JRC EUROPEAN COMMISSION
EMM user requirements: many languages; summarisation
Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009   31

- **Many languages, more languages**, depending on user group
  - Majority of EU institutions: the 23 official **EU** languages
  - Public Health: **Asian, African**
  - Nuclear Safeguards: **countries with nuclear facilities** or aspirations
  - Humanitarian Aid: **crisis areas**, e.g. in Africa, Asia or Latin America

- **Multilingual, multi-document summarisation,** update summarisation. But: how?

JRC EUROPEAN COMMISSION
EMM user requirements: Machine Translation
Multilingual Information Access Technology Transfer Day, Berlin, Germany, 8 December 2009   32

- **Machine Translation**, including for less widely spoken languages

- Although even other simple CLIA methods are useful

- MT not on demand, but pre-translated (expensive)

14:16 WHO warns of the resurgence of bird flu virus
The World Health Organization warned "the likelihood of renewed the emergence of bird flu after the registration of new cases of the virus in poultry in Egypt, Indonesia, Thailand and Vietnam," referring to the supervision of the seriousness of "the occurrence of a mixture between bird flu virus and avian pigs, which will lead to the emergence of a new strain of the disease more lethal."

- · **Information fusion** before presentation to the users would save time and effort

---

- · Translating everything into one target language is good, but has drawbacks.
  - · Important specific terms and names are often not translated, or translated badly, e.g.
    - · Person and organisation names
    - · Locations
    - · Names of diseases and toxins

- · Our insight: source language information extraction is better for information filtering than searching machine-translated text
  - · MT as a second step

- **Keep applications simple!**
    - Not only from the user's point of view, but also from the developer's point of view, e.g. multilingual EMM category definitions
        - Manually produced Boolean combinations or cumulative weights, using wild cards
        - Faster and easier than learning categorisers for hundreds of classes in 50 languages

    - Ideally language-independent rules, to which language-specific resources can be plugged in

    - If possible, avoid language pair-specific resources
        - N*(N-1) / 2   language pair combinations
        - NewsExplorer's 19 languages: 171 language pairs!

Steinberger Ralf, Bruno Pouliquen & Camelia Ignat (2008). **Using language-independent rules to achieve high multilinguality in Text Mining.** In: Fogelman-Soulié Françoise, Domenico Perrotta, Jakub Piskorski & Ralf Steinberger (eds.): Mining Massive Data Sets for Security. pp. 217-240. **IOS Press**, Amsterdam, The Netherlands. (Overview article explaining the **design principles to achieve highly multilingual applications** such as NewsExplorer) (PDF)

→ Need for **uniform highly multilingual resources** for information extraction, etc.

Similar to Multext and Multext-East, EuroWordNet, JRC-Acquis, …

- Multilingual **gazetteers** for geo-coding
- Multilingual **dictionaries** using the same features and format
- Multilingual **taggers and parsers** producing the same output type and output format.
- Multilingual annotated text **corpora**
- **Single access point for licensing** issues
- Negative example: dependency parsers to be used for relation and event extraction
- …

Summary

- Europe Media Monitor applications (EMM; http://press.jrc.it/overview.html):
  - Users
  - Functionality

- Importance of multilinguality
- EMM user requirements
- Our insights