

Sistemi per il recupero delle informazioni

Gabriele Pozzani

A.A. 2014/2015

**Corso di Laurea Magistrale in
Editoria e Giornalismo**

Non solo pagine web

Google Trends

<http://www.google.it/trends>

- È uno strumento di Google che analizza le ricerche effettuate (dal 2004) nel mondo e in alcune regioni
- Permette di vedere
 - Ricerche più frequenti (“hot trends”/”hot searches”) di ogni giorno (in alcuni Paesi)
 - In diverse categorie, classifica delle ricerche più frequenti o in maggiore ascesa di popolarità (solo USA)
 - Esplorare e confrontare l'andamento nel tempo e nello spazio delle ricerche (nel mondo o a livello (sub)nazionale)

3

Ricerche più frequenti

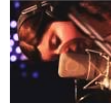
- Attivo solo per alcuni Paesi
- Ricerche maggiormente eseguite ogni giorno

4

Ricerche più frequenti

Incorpora

martedì 26 novembre 2013



Olivia Wise

10,000+ ricerche

Toronto teen Olivia Wise loses fight with brain cancer Toronto Sun
Olivia Wise — the terminally ill teen who captured the world's heart with a viral vi...

CTV News

1



Canucks

10,000+ ricerche

Kings catch, down Canucks in OT Vancouver Sun
A nudge from the L.A. Kings pushed the Canucks into the game emotionally but ...

Vancouver Sun

2

lunedì 25 novembre 2013



Brian Griffin

200,000+ ricerche

Ricerche correlate: **family guy, brian griffin dead, family guy brian, family guy ...**
Family Guy's Brian Griffin killed off CBC.ca
Brian has been a series staple since the show debuted in 1999. An articulate bl...

Metro

1



Bound 2

50,000+ ricerche

Ricerche correlate: **james franco, kanye west**
James Franco and Seth Rogen Recreate Kanye West's 'Bound 2' Video Bilboard
Less than a week after West premiered his buzzworthy naked-Kim Kardashian-o...

Daily Mail

2



Black Friday

20,000+ ricerche

Is Black Friday At Risk of Extinction? Huffington Post (blog)
For adventure seekers that want the thrill of getting the real deal, the controlled c...

3

5

Classifiche

- Attivo solo per gli USA
- Per diverse categorie (e.g., libri, attori, alimenti, atleti, aziende, ...) vengono mostrate le classifiche mensili
 - Per “volume di ricerca”
 - Per maggiore ascesa nelle ricerche
- Ogni categoria è esplorabile per vederne nel dettaglio la classifica

6

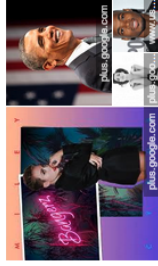
Paesi e regioni



1. United States	—	—
2. Mexico	↑	—
3. India	↓	—
4. China	—	—
5. Canada	—	—

Altro ▶

Persone



1. Miley Cyrus	—	—
2. Barack Obama	↑	—
3. Kim Kardashian	↑	—
4. Adrian Peterson	⚡	—
5. Drake	↓	—

Altro ▶

Politici



1. Barack Obama	—	—
2. John F. Kennedy	↑	—
3. Abraham Lincoln	↓	—
4. Ted Cruz	↑	—
5. George Washington	—	—

Altro ▶

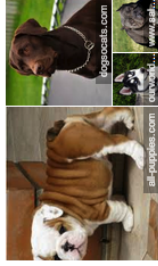
Programmi televisivi



1. The Walking Dead	↑	—
2. Breaking Bad	↓	—
3. American Horror Story	⚡	—
4. Saturday Night Live	↑	—
5. Glee	↑	—

Altro ▶

Razze canine



1. Bulldog	—	—
2. Labrador Retriever	—	—
3. Siberian Husky	↑	—
4. Pit Bull	↓	—
5. German Shepherd	↓	—

Altro ▶

Reality show



1. Keeping Up with the Kardashians	↑	—
2. The Voice	↑	—
3. Dancing with the Stars	↓	—
4. Bigg Boss	↑	—
5. Project Runway	↑	—

Altro ▶

Incorpora

Condividi

7

Stati Uniti ▼

ott 2013 ▼



Visualizzazione degli argomenti più cercati

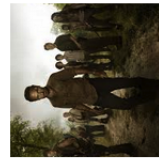


Intrattenimento > Programmi televisivi



I più cercati

Tendenza



1

The Walking Dead

100

Condividi

The Walking Dead is an American horror drama television series created and produced by Frank Darabont. It is based on the comic book series of the same name by Robert Kirkman, Tony Moore, and Charlie Adlard... Wikipedia

↑ +4 dal mese scorso
Ha raggiunto la posizione #1 (mensile)
24 mesi nei primi 10
Esplora



2

Breaking Bad

72

Breaking Bad is an American crime drama television series created and produced by Vince Gilligan. Set and produced in Albuquerque, New Mexico, Breaking Bad is the two-year-long story of Walter White,... Wikipedia

↓ -1 dal mese scorso
Ha raggiunto la posizione #1 (mensile)
9 mesi nei primi 10
Esplora



3

American Horror Story

45

American Horror Story is a horror television series created and produced by Ryan Murphy and Brad Falchuk... Wikipedia

⚡ Nuovo nella classifica
Ha raggiunto la posizione #3 (mensile)
6 mesi nei primi 10
Esplora



4

Saturday Night Live

42

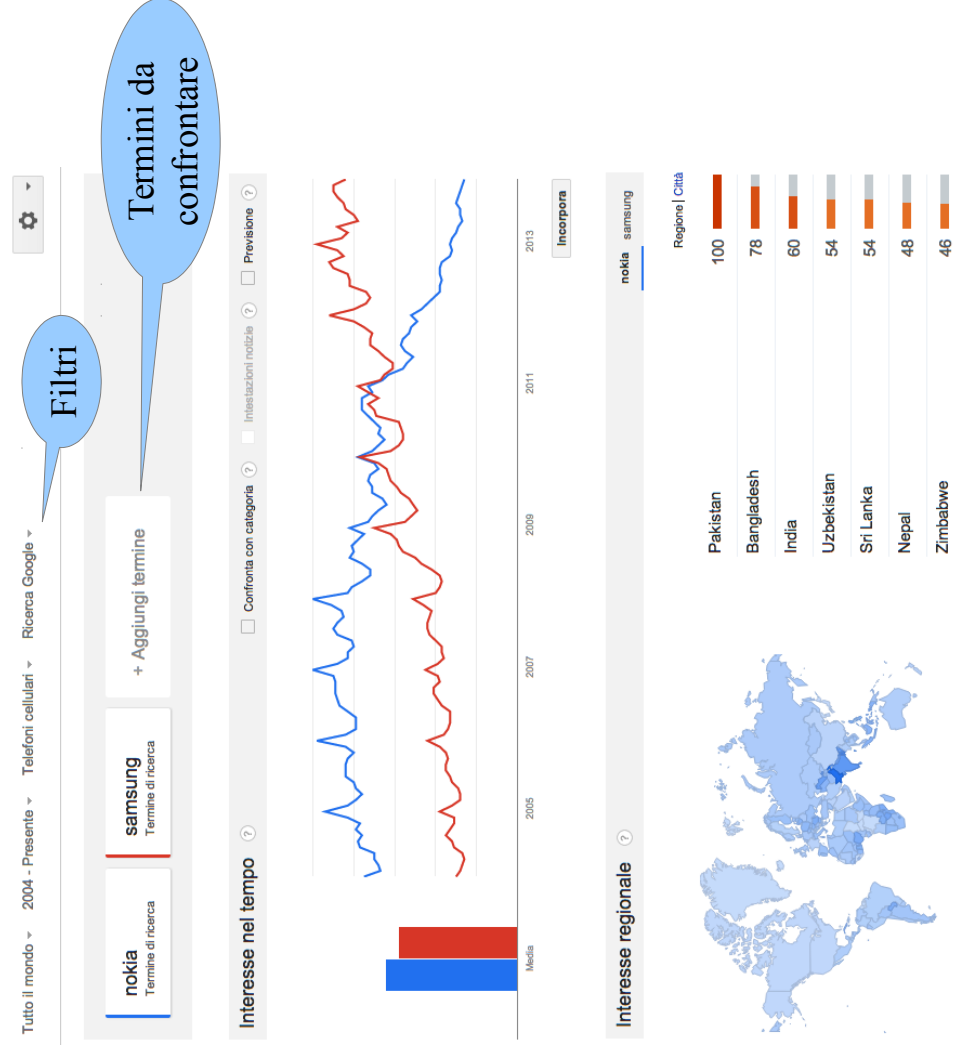
Saturday Night Live is an American late-night live television sketch comedy and variety show created by Lorne Michaels and developed by Dick Ebersol. The show premiered on NBC on October 11, 1975, under the original title, *Nighty, Saturday*.

↑ +6 dal mese scorso
Ha raggiunto la posizione #1 (mensile)
89 mesi nei primi 10
Esplora

Esplora: confronta termini di ricerca

- Permette di inserire i termini di ricerca da confrontare
- Mostra e confronta
 - L'andamento temporale (dal 2004 ad oggi)
 - L'“interesse” nei diversi Paesi del mondo (e la sua evoluzione nel tempo)
 - L'andamento delle ricerche correlate

9

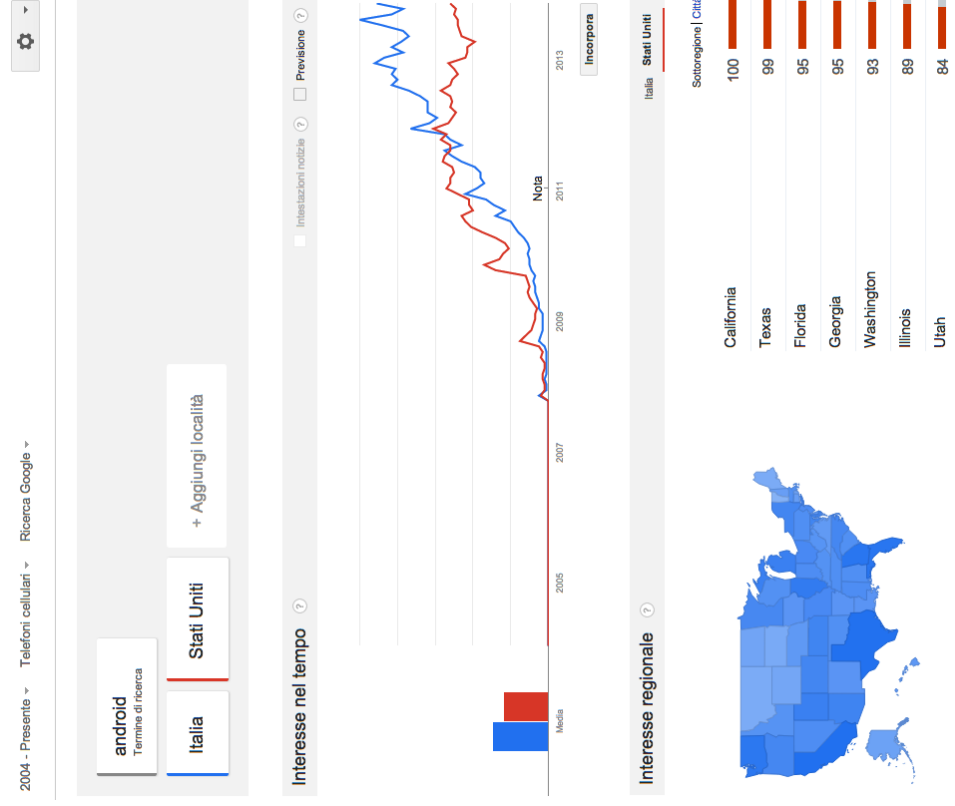


10

Esplora: confronto località

- Dato un termine di ricerca confronta la sua “importanza” in diversi Paesi del mondo
 - Andamento nel tempo
 - Andamento a livello (sub)nazionale (regione, città, ...), anche nel tempo
 - L'andamento delle ricerche correlate

11

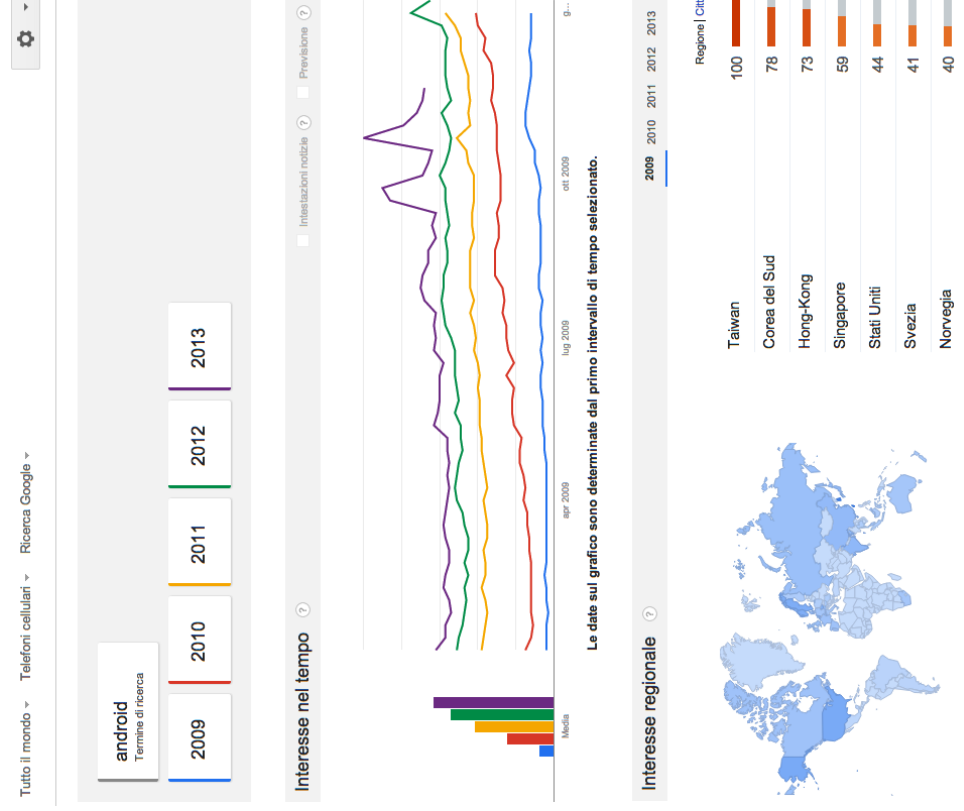


12

Esplora: confronto intervalli di tempo

- Immeso un termine di ricerca se ne può confrontare l'importanza in diversi intervalli temporali

13



14

Google Public Data (I)

<http://www.google.com/publicdata/>

- Introdotto nel marzo 2010
- Strumento che facilita l'esplorazione, la visualizzazione, il confronto e la comunicazione di grossi set di dati di pubblico interesse
- Raccoglie e permette l'analisi di diverse tipologie di dati prodotti originariamente da diverse istituzioni mondiali
 - Banca Mondiale
 - Eurostat, ISTAT e altri corrispettivi nazionali
 - Università
 - Istituti di ricerca statistica
- Permette di importare proprie statistiche e dati per poterli analizzare e visualizzare con i tool di Google

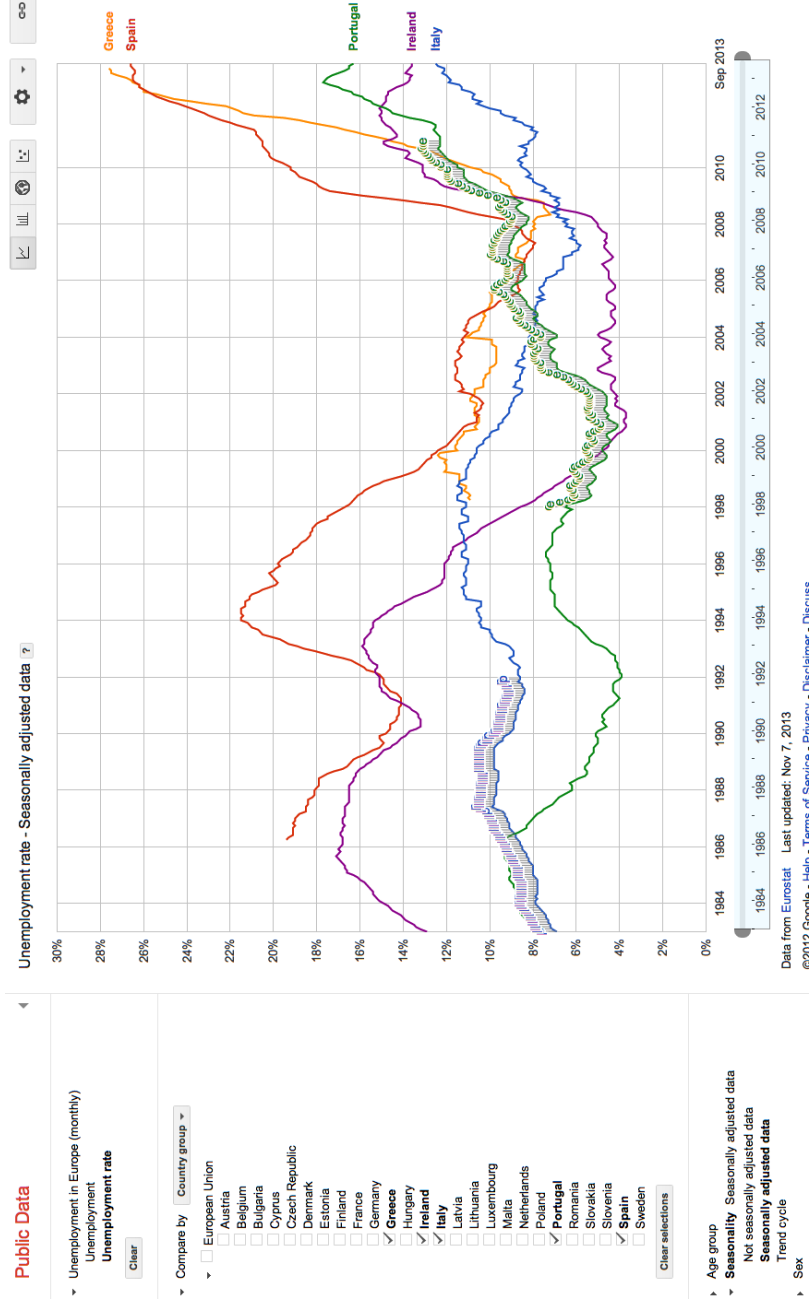
15

Google Public Data (II)

- Non fornisce dati “grezzi” che solo addetti ai lavori potrebbero capire
- Raggruppa e aggrega dati
- Presenta i dati in grafici di vario tipo, rendendone la lettura più semplice
- Permette di esplorare
 - Set di dati: una serie di statistiche, raggruppate in una unità omogenea da un unico fornitore di dati
 - Ad esempio, "Disoccupazione in Europa (mensile)", prodotto da Eurostat e "Indicatori di sviluppo mondiale", prodotto dalla Banca Mondiale
 - Metriche: singole metriche, ovvero statistiche, che sono componenti di set di dati più grandi
 - Ad esempio, "Popolazione", "Tasso di disoccupazione" e "Prodotto interno lordo"

16

Tasso di disoccupazione nei “PIIGS”



17

Applicazioni

- Google Trends e Google Public Data permettono di accedere ed analizzare una grande quantità di dati utili per diversi fini
- Un uso “avanzato” di Trends è lo studio dei trend influenzali

18

Flu Trends (I)

<http://www.google.org/flutrends/intl/it/>

- Determinati termini di ricerca sono validi indicatori dell'attività influenzale
- Utilizza dati di ricerca aggregati di Google per stimare l'attività influenzale
- Idea: più le persone cercano informazioni sull'influenza più questa si sta diffondendo
 - Analizzando la mole di ricerche si comprende la diffusione dell'influenza
 - L'andamento temporale delle ricerche è correlata alla sua evoluzione/diffusione nel tempo
 - Analizzando le zone del mondo dove avvengono tali ricerche se ne comprende l'evoluzione/diffusione spaziale

<http://www.nature.com/nature/journal/vaop/ncurrent/full/nature07634.html>

19

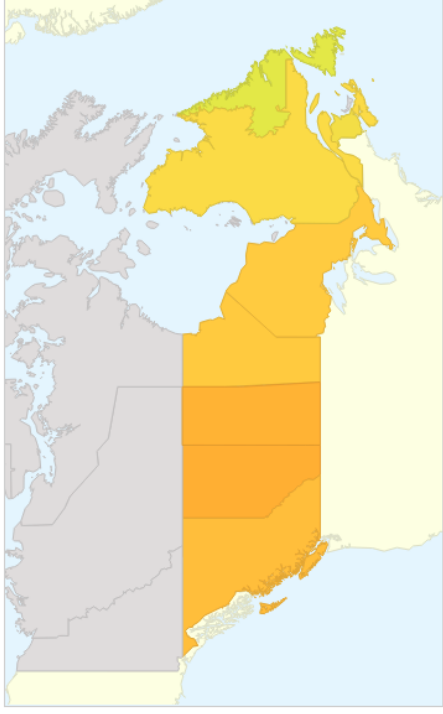
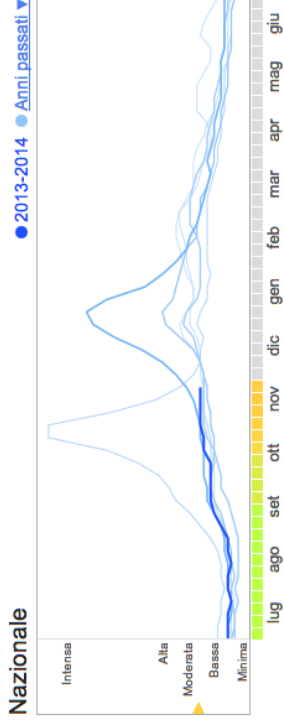
Flu Trends (II)

- Servizio utile a fini epidemiologici
- “La sorveglianza tradizionale dell'attività influenzale è molto importante, ma gran parte delle agenzie del settore sanitario si concentra su un singolo Paese o regione e aggiorna le proprie stime solo una volta alla settimana. Google Trend influenzali è attualmente disponibile per vari Paesi e regioni del mondo e viene aggiornato quotidianamente, quindi rappresenta uno strumento complementare a questi sistemi esistenti. [...] il rilevamento tempestivo di un'epidemia può ridurre il numero di persone colpite.”

20

Scopri i trend influenzali - Canada

Abbiamo scoperto che determinati termini di ricerca sono validi indicatori dell'attività influenzale. Google Trend influenzali utilizza dati di ricerca aggregati di Google per stimare l'attività influenzale. [Ulteriori informazioni »](#)



Le stime sono state calcolate utilizzando un modello che si è rivelato accurato se confrontato con i dati storici ufficiali sulle attività influenzali. Dati aggiornati al 25 novembre 2013.

21

Data Journalism (I)

- Più in generale strumenti come questi possono essere utili per il cosiddetto “data journalism”
- Il data journalism (anche “giornalismo basato sui dati” o “giornalismo di precisione”)
 - È un approccio a cavallo tra ricerca e inchiesta giornalistica
 - Fa uso intensivo di database, mappe digitali e software
 - Per analizzare, raccontare e visualizzare un fenomeno o una notizia
 - Spesso produce dei mashup di elementi diversi
- Il cronista adotta un metodo scientifico, e applica il rigore del metodo usato, per scrivere un articolo o scoprire un fatto
- La forza del metodo di analisi conferisce precisione, e quindi obiettività, alla notizia

22

Data Journalism (II)

- Richiede quindi competenze diverse:
 - Scienze sociali e comportamentali
 - Per l'approccio e la corretta contestualizzazione ed analisi dei dati
 - Utilizzo dei computer
 - Per il reperimento, elaborazione e visualizzazione
 - Matematica e statistica
 - Per il metodo scientifico di analisi
- I SRI (in particolare nel web) possono essere utili per il reperimento dei dati
 - Tramite classiche ricerche testuali
 - Tramite appositi servizi

23

Sorgenti dati (I)

- Il web contiene una sempre maggiore quantità di dati pubblici
- Molti siti, oltre a Google, li raccolgono e ne permettono l'analisi e la visualizzazione
 - Zanran <http://www.zanran.com/q/>
 - DataMarket <http://datamarket.com/>
 - Data360 <http://www.data360.org/>
 - Knoema <http://knoema.com/>

24

Sorgenti dati (II)

- Mentre liste di dataset pubblici possono essere trovati su vari siti:
 - <https://delicious.com/pskomoroch/dataset>
 - <http://www.kdnuggets.com/datasets/>
 - [http://www.w3.org/wiki/TaskForces/CommunityProjects/Linking OpenData/DataSets](http://www.w3.org/wiki/TaskForces/CommunityProjects/Linking_OpenData/DataSets)
 - <http://www.infochimps.com/datasets>
 - <http://www.quora.com/Data/Where-can-I-find-large-datasets-open-to-the-public>
 - <http://www.reddit.com/r/datasets/>
 - <http://datahub.io/dataset>
 - <http://datacatalogs.org/>

25

Riferimenti per il data journalism

- <http://datajournalismhandbook.org/>
- <http://www.theguardian.com/data>
- <http://journalistsresource.org/skills/research/understanding-data-journalism-overview-tools-topics>
- <https://fellows.knightscience.org/tutorials/resources-data-journalism-toolbox/>
- <http://datadrivenjournalism.net/>
- <http://clairemiller.net/blog/2012/01/huge-list-of-data-journalism-resources/>

26

Altri motori di ricerca

- Esistono diverse liste di motori di ricerca:
 - <http://l-lists.com/en/lists/qukoen.html>
 - http://en.wikipedia.org/wiki/List_of_search_engines
 - <http://www.thesearchenginelist.com/>

27

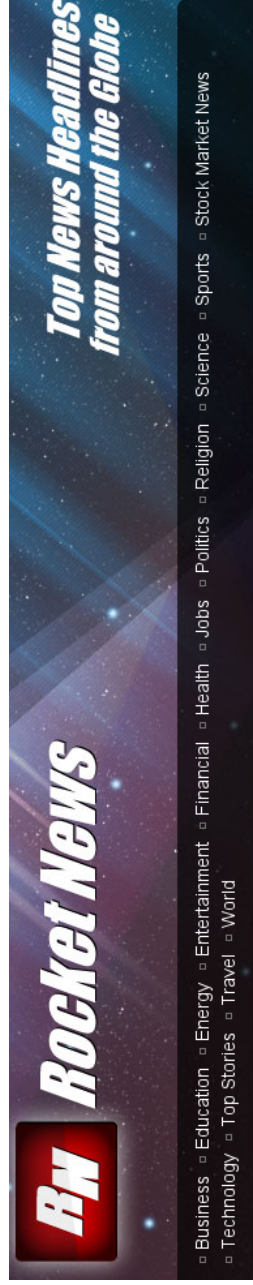
Rocket News

<http://www.rocketnews.com/>

- Aggregatore di news da tutto il mondo
- Raccoglie news da diverse fonti, tra cui BBC, Reuters, News Corporation, Yahoo
- Permette di cercare
 - Per categoria
 - Per termine

28

Rocket News



Search Results for: iran

Kerry delivers video on Iran deal to Congress

Posted on November 26, 2013 by Editor - World News

US Secretary of State John Kerry has taken the unusual measure of delivering a video message to Congress in an attempt to stop the body from collapsing last weekend's nuclear deal with Iran. Just back from his diplomatic triumph in ... [Continue reading →](#)

Share This:



Filed under [World](#) | Tagged [World News](#) |

Iraq says no success tracing killers of Iranian dissidents

Posted on November 26, 2013 by Editor - Top News Stories

BAGHDAD (Reuters) – Iraq is hunting militants, still unidentified, who led a deadly attack on an Iranian dissident camp near Baghdad and dismisses suggestions its own security forces were behind the violence, a senior government official said. More than 50 ... [Continue reading →](#)

Share This:



Filed under [Top Stories](#) | Tagged [Breaking News](#), [Current News](#), [Daily News](#), [Live News](#), [News](#), [Newspapers](#) |

RocketNews RSS News Feeds

RocketNews is a leader in live news stories, a resource for daily news feeds. RocketNews delivers the top stories and the **latest RSS news feeds** for:

Business
Entertainment
Financial
Health
Politics
Religion
Science
Sports
Stock Market News
Technology
Top Stories
Travel
World
RocketNews

[Subscribe](#)

Term Life Insurance Quotes

Term life insurance and no medical insurance can protect your family and provide key man business insurance. Use a **term life insurance quotes** calculator to get a best price quote.

[Website Awards](#)

socialmention

<http://socialmention.com/>

- Ricerca in real-time su diversi social media, inclusi Tweeter, Facebook e Google News di un termine immesso
- Calcola diverse misure
 - Sentiment
 - Passion
 - Principali parole chiave
 - Principali autori
 - Principali hashtag
 - ...

socialmention

Blogs Microblogs Bookmarks Comments Events Images News Video Audio Q&A Networks All
"silvio berlusconi" Search Advanced Search Preferences

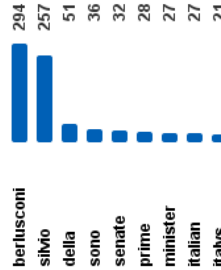
socialmention*



Sentiment



Top Keywords



Mentions about "silvio berlusconi"

Sort By: Date Results: Anytime Results 1 - 15 of 211 mentions.

- [f Gioirò solo quando potrò gettare un fiore sulla bara di Silvio Berlusconi... Il fatto che sia ancora vivo è un pericolo immane per il nos...](#)

[www.facebook.com/permalink.php?id=1000021799417633&v=wall&story_fbid=551807001568626](#)
14 seconds ago - by [Tullio Facciolini](#) on [facebook](#)

- [RT @Rosariotres: Silvio Berlusconi fue expulsado del Senado italiano <http://t.co/9q3DLxT5jV>](#)

[twitter.com/AdrianTebes17/status/405832422509404162](#)
37 seconds ago - by [@AdrianTebes17](#) on [twitter](#)

- [Berlusconi expelled from Senate: Love him or hate him, Silvio Berlusconi is a divisive but still dominant figu... <http://t.co/OcwrdM3i1N>](#)

[twitter.com/NewsCrazynet/status/405832422832345088](#)
37 seconds ago - by [@NewsCrazynet](#) on [twitter](#)

- [f Silvio Berlusconi e espulso do senado italiano por ter sido condenado, agora poderá responder por outros crimes já que perdeu a imunidade ...](#)

[www.facebook.com/permalink.php?id=1000016182016233&v=wall&story_fbid=618151454915456](#)
42 seconds ago - by [Pedro Ruth Miranda](#) on [facebook](#)

- [f Senado deve à silvio berlusconi?](#)

31

Topsy

<http://topsy.com/>

- Tweet search engine
- Social search
 - Permette di ricercare i tweet con un determinato termine, hashtag, persona, ...
 - Calcola anche il “sentiment”
- Social analytics
 - Permette di confrontare il volume di tweet con diversi termini, hashtag, ...
- Esiste una versione “pro” con maggiori funzionalità

32

Topsy: social search

TOPSY

m5s

Sort by relevance

⚙️

Latest Results

Past 1 Hour159

Past 21 Hours3K

Past 1 Day3.1K

Past 7 Days13K

Past 30 Days59K

All Time

Specific Range

Everything

Links

Tweets

Photos

Videos

Influencers

All Languages

English

中文

日本語

...

Past 21 Hours

3,001 TWEETS

Topsy Sentiment Score: 47

Tweets Over Past 30 Days

View trends on Topsy Analytics

Want to see more?

Topsy Pro provides complete social insight

FREE TRIAL

PAOLA TAVERNA (M5S) umilia BERLUSCONI il PDL sbroccia

youtube.com/watch?v=23hpQq5xkQ0

beppe_grillo Beppe Grillo

RT @MartinaVerrilli: Paola #Taverna (M5S) umilia #Berlusconi youtu.be/23hpQq5xkQ0 via @youtube #decadenza

6 hours ago

Reply

Retweet

Favorite

345 more

Beppe Grillo @beppe_grillo

Sulla #decadenza di Berlusconi il gruppo M5S al Senato ci ha messo la firma: goo.gl/chLx3 @Senato5Stelle

pic.twitter.com/TeTYRhApoA

9 hours ago

Reply

Retweet

Favorite

133 more

Michele Emiliano @micheleemiliano

Bellissimo intervento della Senatrice Taverna del M5S scondiamente interrotta da figure travestiti da senatori immeritevoli del loro ruolo

8 hours ago

Reply

Retweet

Favorite

144 more

Vittorio Zucconi @vittoriozucconi

Silvio e Beppe insieme all'opposizione. Fl e M5S appaiati nei sondaggi. Silvio ha le reti, Beppe ha la Rete. Le solite convergenze parallele

4.6 hours ago

Reply

Retweet

Favorite

44.9 more

Topsy: social analytics

TOPSY

Share

Get the full picture with trends, geo, sentiment, and more with Topsy Pro.

Get Topsy Pro

"forza italia"

"nuovo centrodestra"

starbucks

Search

Advanced search

October 28th — November 27th

Tweets per day: "forza italia" and "nuovo centrodestra"

010/2810/3111/0311/0611/0911/1211/1511/1811/2111/2411/26

7K6K5K4K3K2K1K0

"forza italia"36,363

"nuovo centrodestra"5,392

ANALYTICS BY TOPSY

34