

Open Data: la filosofia e gli strumenti

Vincenzo Patruno

DATA
JOURNALISM 2
school

Un mondo di dati



I dati non solo sono la misura di qualcosa,
ma raccontano quello che facciamo



Il diluvio dei dati

The Data Deluge

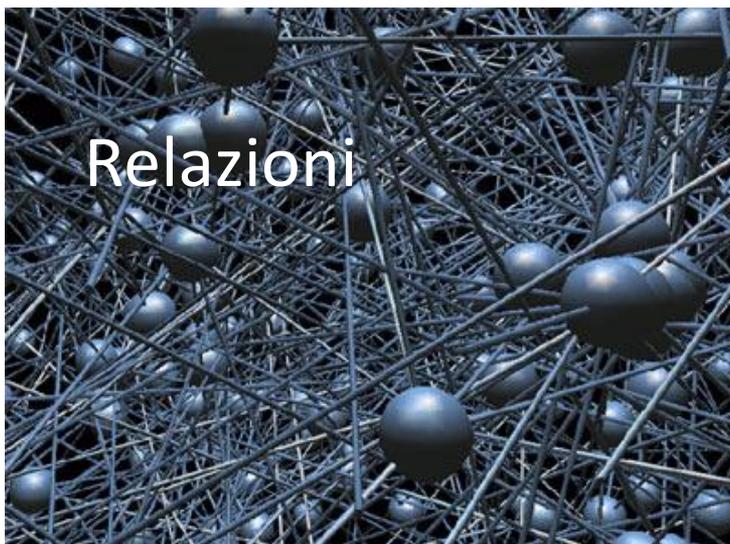
Un mondo di dati



Volume



Fonti



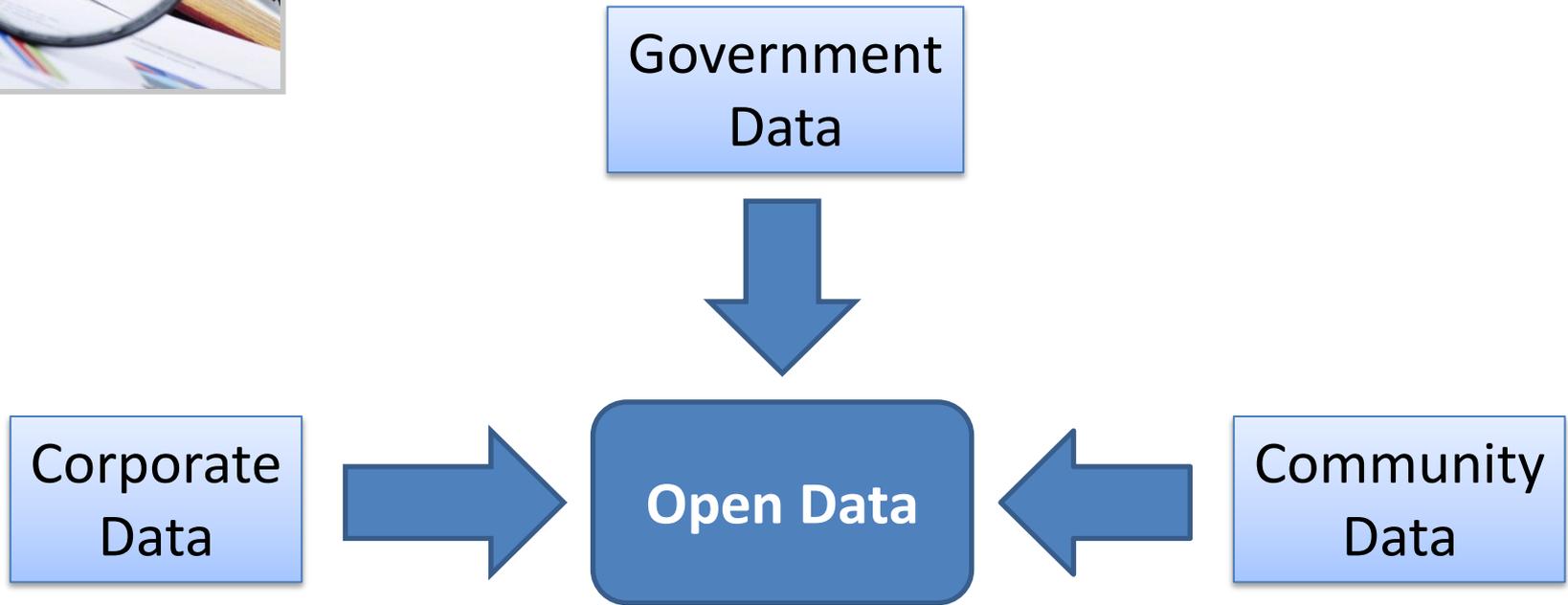
Relazioni



Contesto



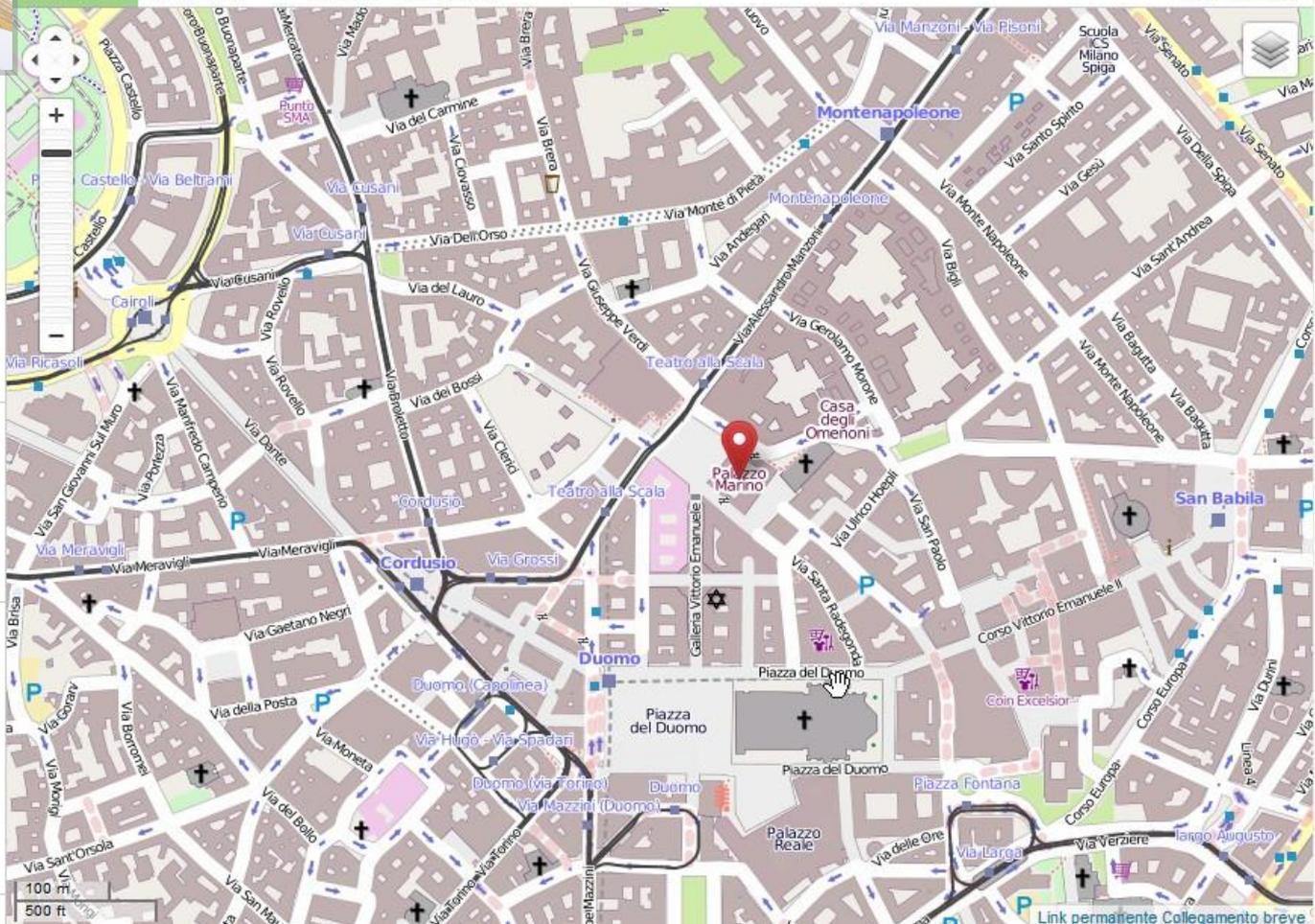
Open Data



Community Data



Visualizza Modifica Storico Esporta entra | iscriviti



OpenStreetMap
La wiki-mappa Libera del Mondo

milano

esempi: 'Trieste', 'Via Dante Alighieri, Trieste', 'CB2 5AQ', oppure 'post offices near Trieste' altri esempi... Dove sono?

OpenStreetMap è una mappa liberamente modificabile dell'intero pianeta. È realizzata da persone come te.

I dati possono essere scaricati liberamente e utilizzati in accordo alla **licenza libera**. Crea un account utente per aiutare a migliorare la mappa.

Aiuto
Aiuto
Documentazione
Copyright e licenza
Comunità
Blog della comunità
Fondazione
Diari degli utenti
Tracciati GPS
Legenda

100 m
500 ft

Link permanente Collegamento breve

Corporate Data



Enel Open Data

AAA | English | Italiano

Ricerca

[Open data](#)

[Dati Gruppo Enel](#)

[Dati Enel Spa](#)

[Dati di Sostenibilità](#)

[Dati Ambientali](#)

[Enel e il Territorio](#)

Enel Open Data

Aumentare la trasparenza ed il coinvolgimento da parte di tutti gli stakeholder

Migliorare il business: "better information makes better markets"

Favorire l'innovazione, i dati aperti stimolano lo sviluppo di nuove applicazioni



Dati aperti

Trasparenza, condivisione, libero uso: ecco la sfida raccolta da istituzioni pubbliche e private per aprire il proprio patrimonio di dati al mondo esterno



Enel Open Data

Un nuovo servizio di condivisione di alcune tipologie di dati attraverso modelli semplificati, che ha l'obiettivo di renderli liberamente accessibili a tutti.

[Dataset Gruppo Enel](#)

[Dataset Enel Spa](#)

[Dataset Sostenibilità](#)

[Tutti i dati Enel anche su Google Fusion Tables](#)

Strumenti

Condividi

Info [Contattaci](#)

Link esterni

- [The World Bank](#)
- [US Government](#)
- [UK Government](#)
- [Seguici su Twitter](#)

Bilanci Enel 2011

- [Consolidato](#)
- [Esercizio](#)
- [Sostenibilità](#)
- [Ambientale](#)

Government Data



L'Open Data si basa sulla constatazione che il dato pubblico è stato prodotto con denaro pubblico, quindi **della collettività**. Ed è a questa che il dato deve essere restituito.





Cosa sono gli Open Data



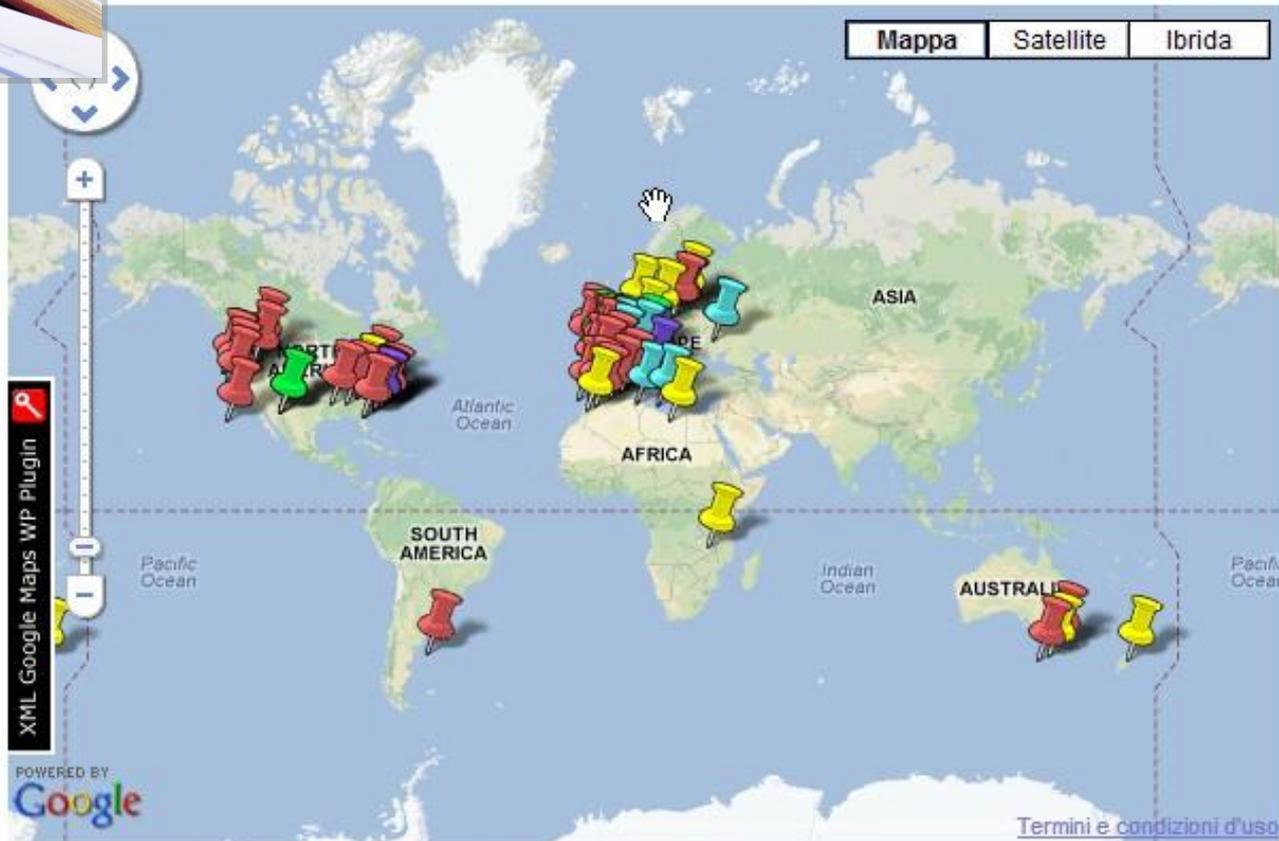
Dati liberamente accessibili a tutti in formato aperto senza restrizioni di copyright, brevetti o altre forme di controllo che ne limitino l'utilizzo.



Come trovare i dati

Google™
Google™

Le iniziative DataGov



<http://opengovernmentdata.org/files/2011/03/ogdmap.kml>



Le iniziative DataGov

data.gov.uk^{BETA}
Opening up government

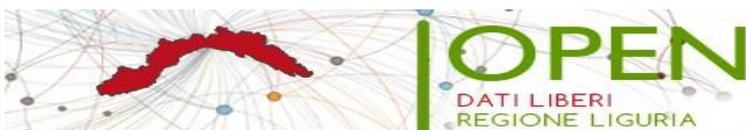


I cataloghi di dati



DATI.Piemonte.it

OpenData Lombardia



Open data Regione Toscana **BETA**

dati.emilia-romagna.it Beta 2.1.0

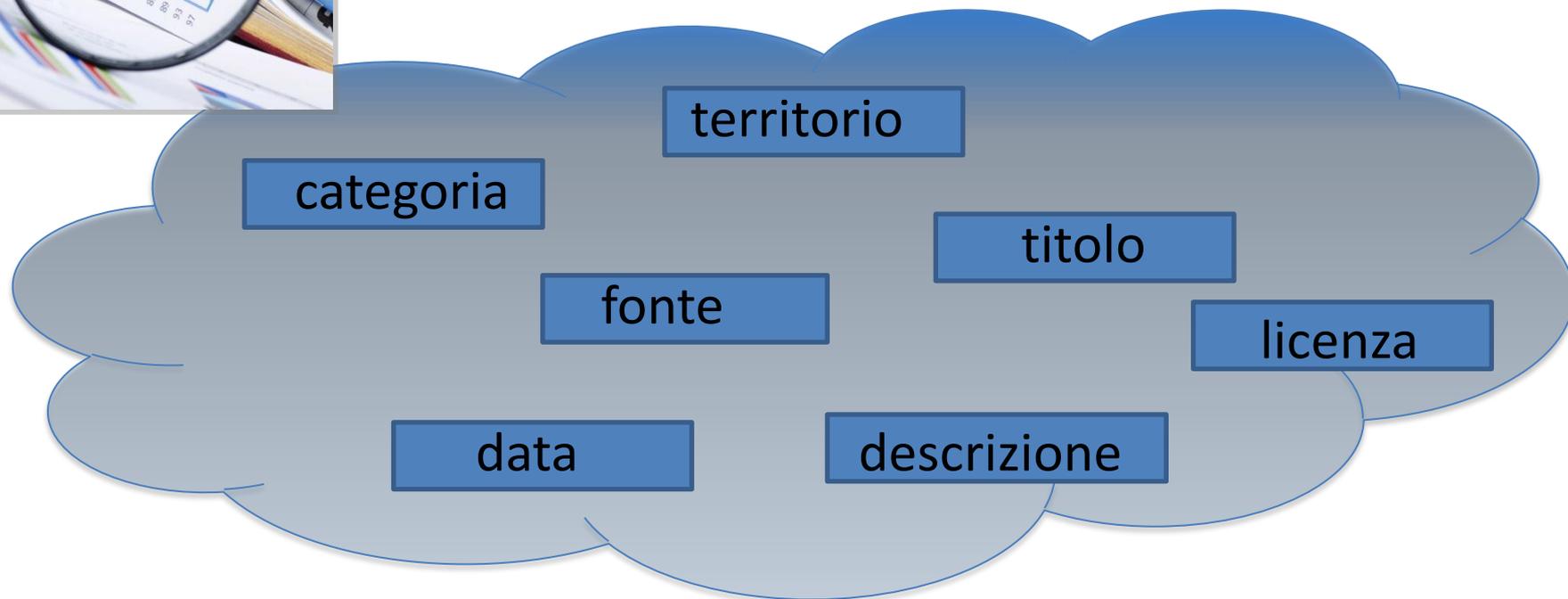


Il formato dei dati



Es. http://www.istat.it/it/files/2012/12/Tavole_XLS.zip

I cataloghi di dati



Metadati



url

Web Scraping



Scraper AGGIUNTO A CHROME

★★★★★ (63) | Strumenti per sviluppatori | da bit155.com | 20.399 Licenti

PANORAMICA | DETTAGLI | RECENSIONI | CORRELATI

Scraper - List of content management systems - Wikipedia, the free encyclopedia

| Name | Platform | Supported databases | Latest stable release | Extension | Latest update date |
|--------------------|----------|---------------------|-----------------------|-----------|--------------------|
| 1. AdaptCMS Lite | PHP | MySQL | 1.0 | GPL | 2009-07-25 |
| 2. ATutor | PHP | MySQL | 2.0.1 | GPL | 2010-09-09 |
| 3. AkaCMS | PHP | MySQL | 3.0.0 | GPL | 2009-12-15 |
| 4. Joomla! | PHP | MySQL | 3.0 RC | AGPL | 2009-06-17 |
| 5. Drupal | PHP | MySQL | 4.2.1 | GPL | 2008-01-15 |
| 6. WordPress | PHP | MySQL | 2.8.5 | GPL | 2008-05-02 |
| 7. CMS Made Simple | PHP | MySQL | 1.8.2 | GPL | 2010-08-12 |
| 8. Concrete5 | PHP | MySQL | 5.4.1 | MIT | 2010-08-04 |
| 9. Joomla! | PHP | MySQL | 2.5.11 | GPL | 2010-06-20 |

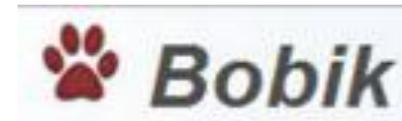
Scraper is a Google Chrome extension for getting data out of web pages and into spreadsheets.

Scraper is a simple data mining extension for Google Chrome™ that is useful for online research when you need to quickly analyze data in spreadsheet form.

To use it: highlight a part of the webpage you'd like to scrape, right-click and choose "Scrape similar...". Anything that's similar to what you highlighted will be rendered in a table ready for export, compatible with Google Docs™.

This is a work-in-progress (i.e. there are bugs), and is currently intended for intermediate to advanced users who are comfortable with XPath, though jQuery is also supported to an extent.

This extension is licensed under a BSD license that can be read here: <https://github.com/mmmlive/scraper/blob/>



<http://scrapy.org/>



ScraperWiki





Web scraping

http://www.comune.torino.it/ambiente/aria/qualita_aria/dati_aria/valori_annuali_pm10.shtml

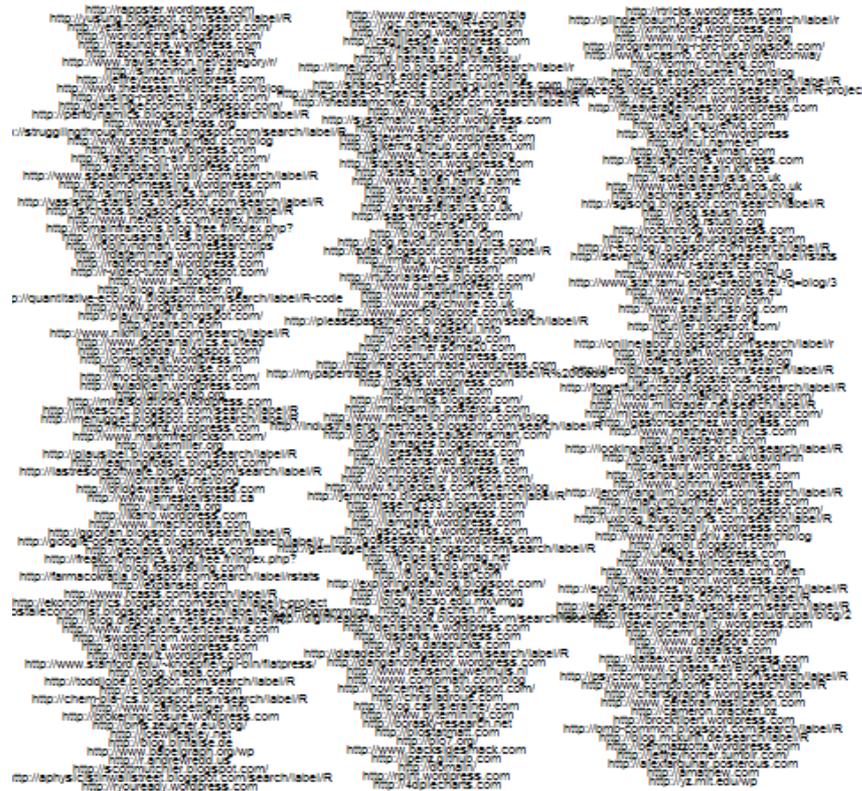
https://scraperwiki.com/scrapers/valori_pm10_in_comune_di_torino/

Esempi



Web scraping

<http://thebiobucket.blogspot.it/2011/10/little-webscraping-exercise.html#more>



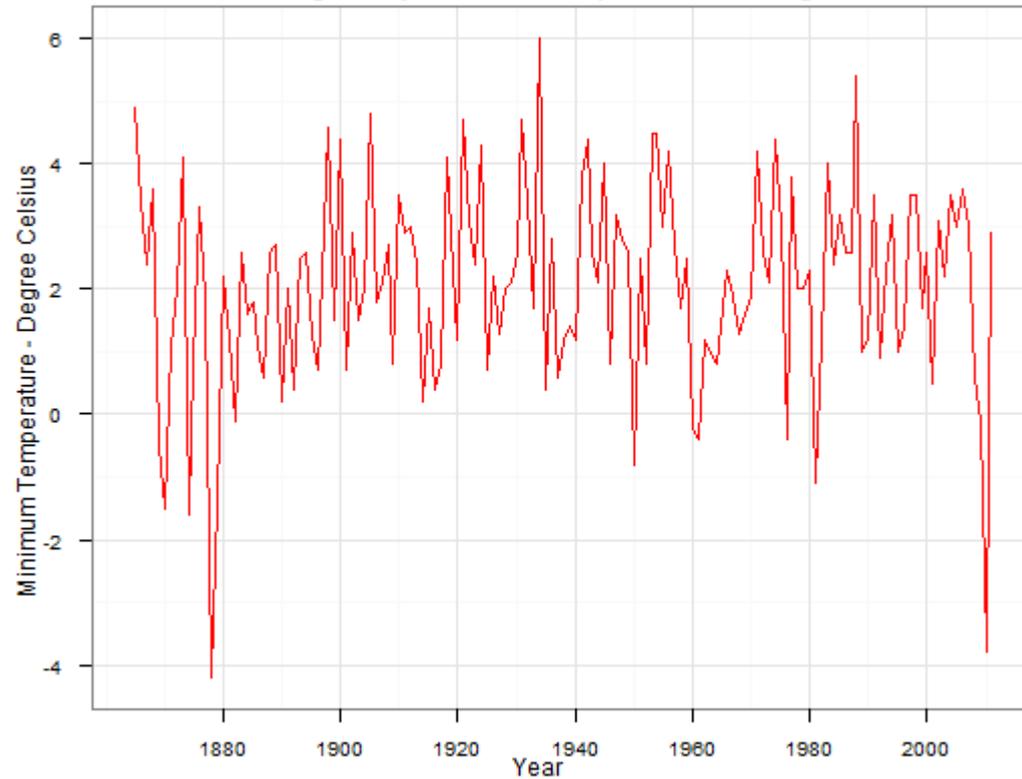
Esempi



Web scraping

<http://www.metoffice.gov.uk/climate/uk/stationdata/armaghdata.txt>

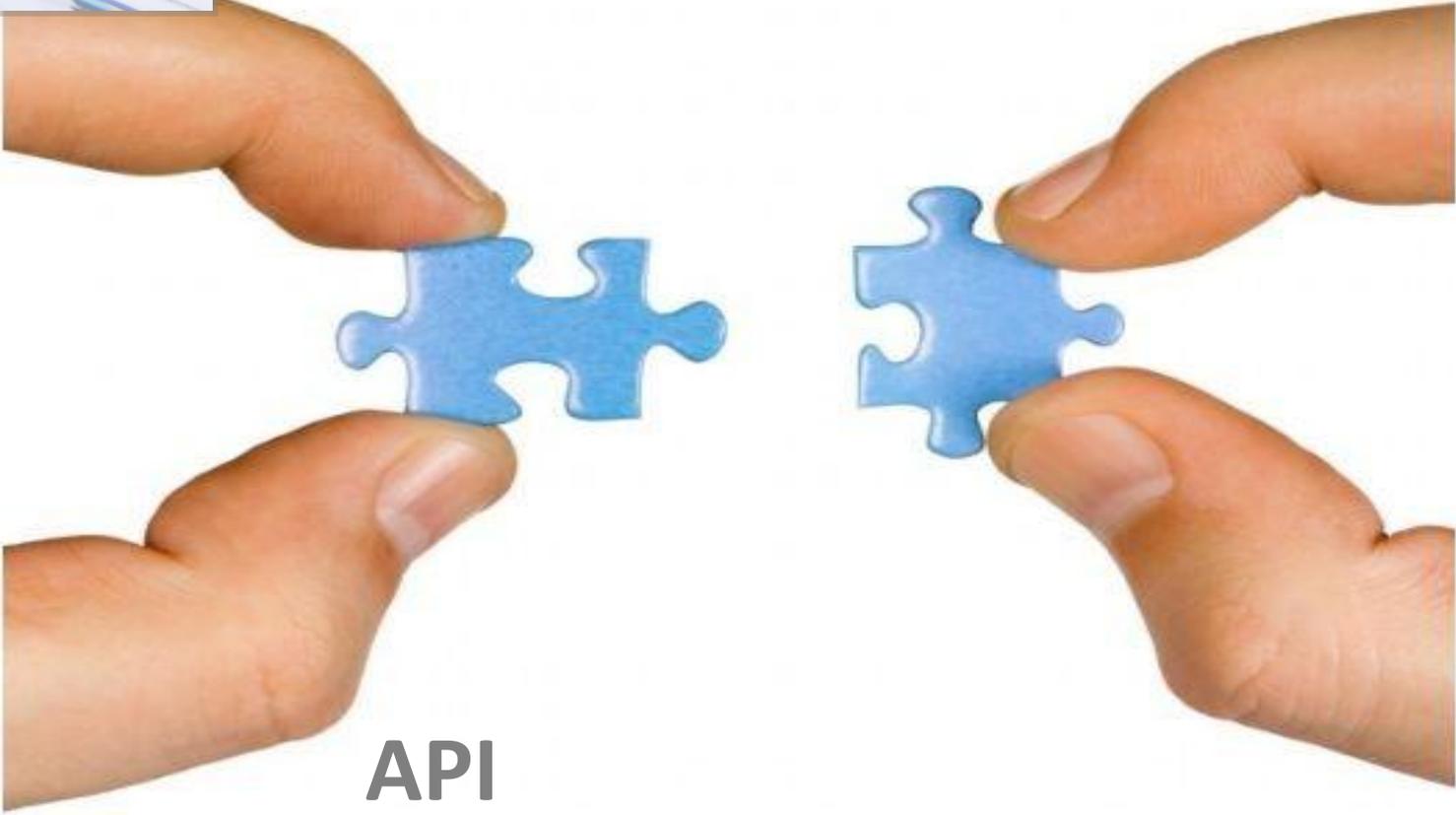
December Average Daily Minimum Temperature - Armagh 1865-2011



Esempi



API – Application Programming Interface





API – Application Programming Interface

`http://apistat.istat.it/?q=gettable&dataset=DCIS_POPORESBIL&dim=82,0,0,0&lang=0&tr=&te=`



query string

API



API – Application Programming Interface

<http://cs.croakun.com/index.php?do=/kafechew/blog/twitter-listening-apps/>



API

Lavorare in Team



Gli sviluppatori SW

Data Integration



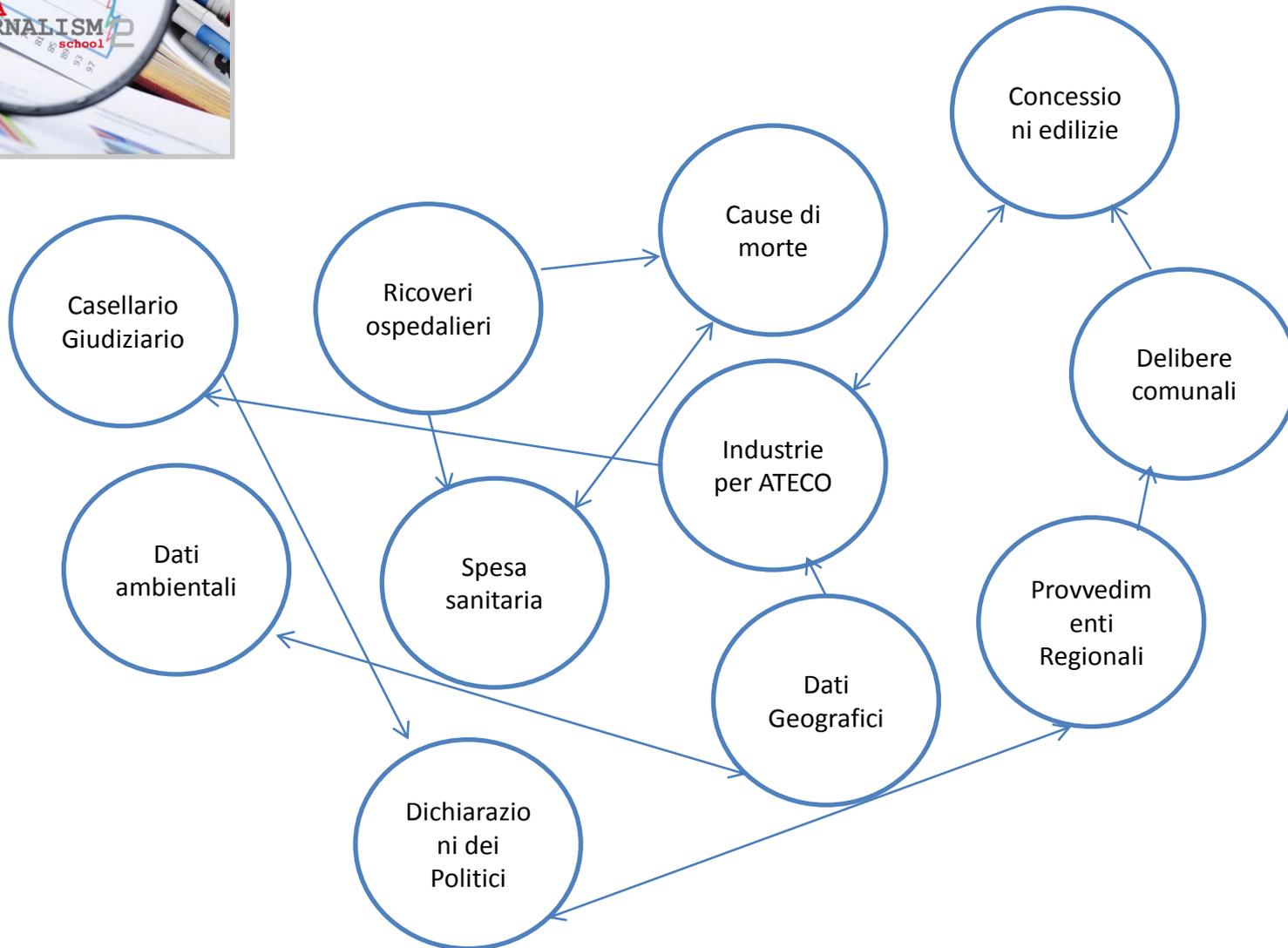
Ricoveri
ospedalieri



Data Integration



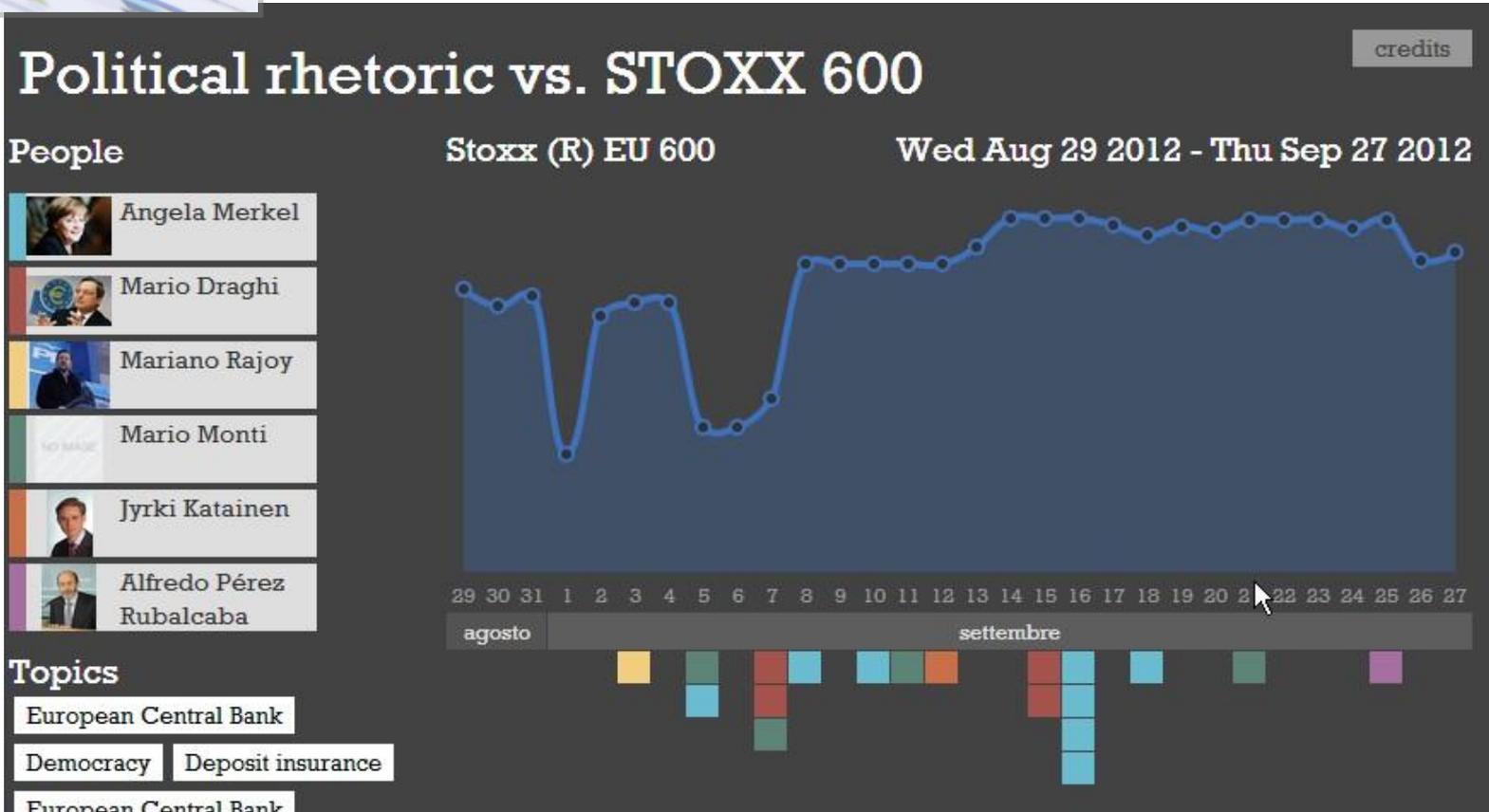
Data Integration





Esempio

<http://ask.thepund.it/?#/timeline/31951d93/20120927>



<http://ask.thepund.it/?#/timeline/55fe01de>



www.thepund.it

Pundit

<http://thepund.it/journalism-how-to.php>



Grazie dell'attenzione!



vincenzo.patruno@istat.it



[@vincpatruno](https://twitter.com/vincpatruno)