



Matteo Palmonari

matteo.palmonari@disco.unimib.it

Knowledge Graphs: Supporting Agents with Web-friendly Knowledge Bases



ITIS Lab – Innovative Technologies for Interaction and Services

Dipartimento di Informatica, Sistemistica e Comunicazione
Università degli Studi di Milano-Bicocca



Outline

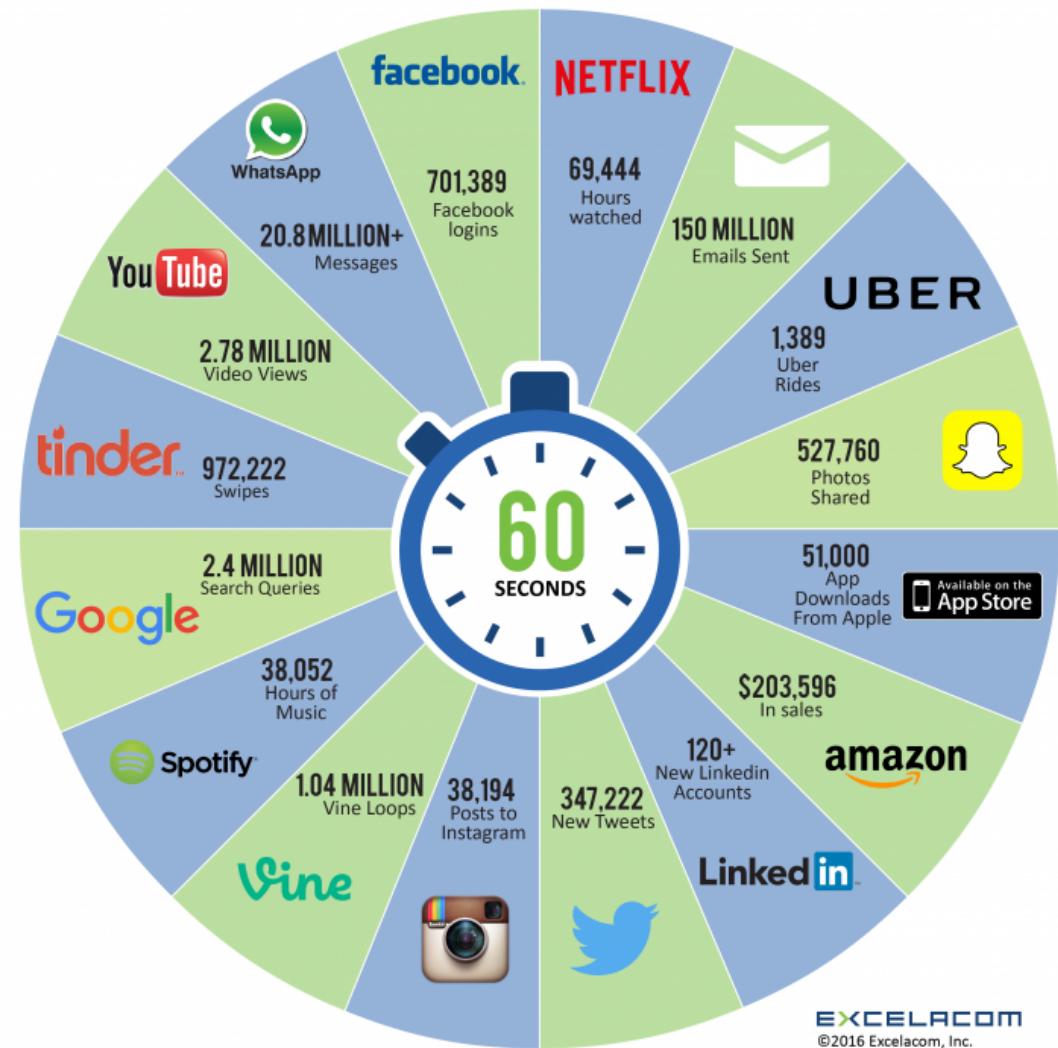
- Big Data & the Web
- Interpreting & Connecting Web Data
- Complex Questions & Decision Making
- Knowledge Bases on the Web (Examples of Linked Data)

Big Data

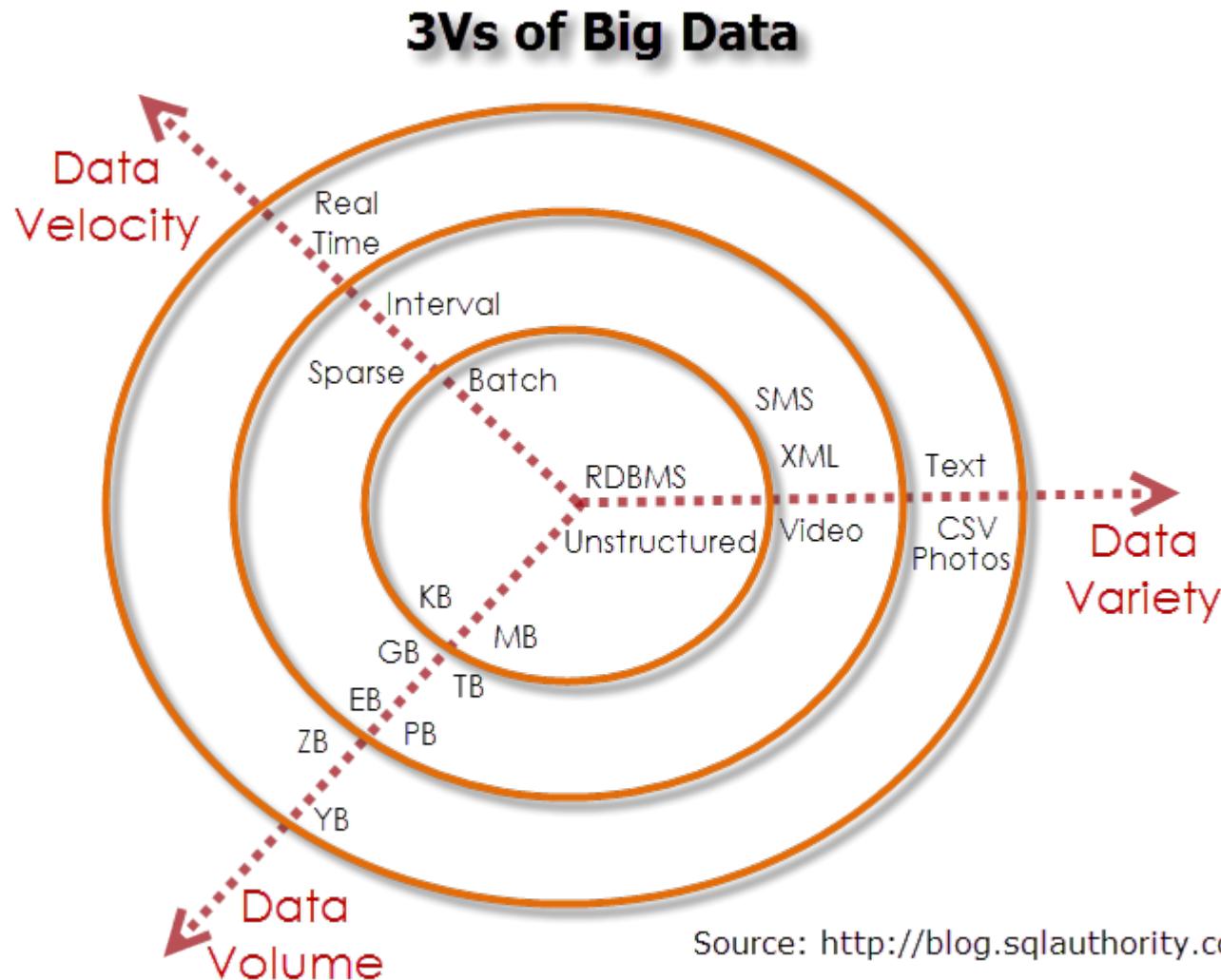
Volume Velocity Variety

Every 60 secs in 2016

2016 What happens in an INTERNET MINUTE?



Big Data: Volume Velocity Variety

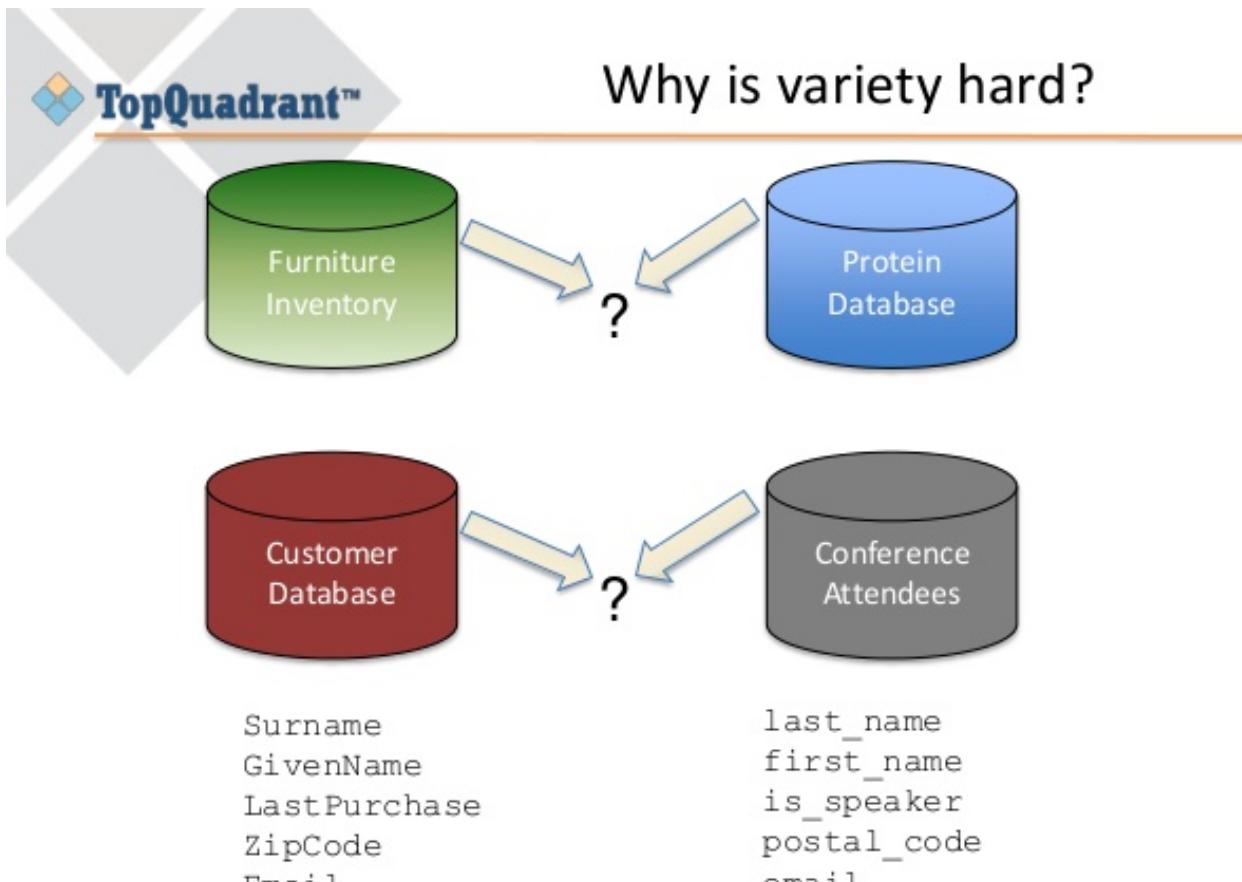


Variety



According to Kapow Software, a comprehensive list of typical Big Data sources comprises these nine categories: archives, docs, media, business apps, social media, public Web, data storage, machine log data, and sensor data. Yellow means external data (public Web), red is internal (archives and data storage), and orange designates both (docs, media, business apps, social media, sensor data).

Variety & Semantics (First Dive)



Web Data

Stadium	City
Eden Gardens	Kolkata
Feroz Shah Kotla Ground	Delhi
M. A. Chidambaram Stadium	Chepauk, Chennai
Wankhede Stadium	Mumbai
Green Park	Kanpur

2. Semi-Structured Data (Hard)

- Hierarchical organization
- May have no schema
e.g., HTML Tables and spreadsheets with nested headers

IBM's Watson Computer Made A BBQ Sauce, And It's Delicious

Watson, a cognitive computing system that can learn and process natural human language, has been one of IBM's most exciting projects of the last decade. Over the past few years, Watson has learned a variety of

1. Semi-Structured Data (Easy)

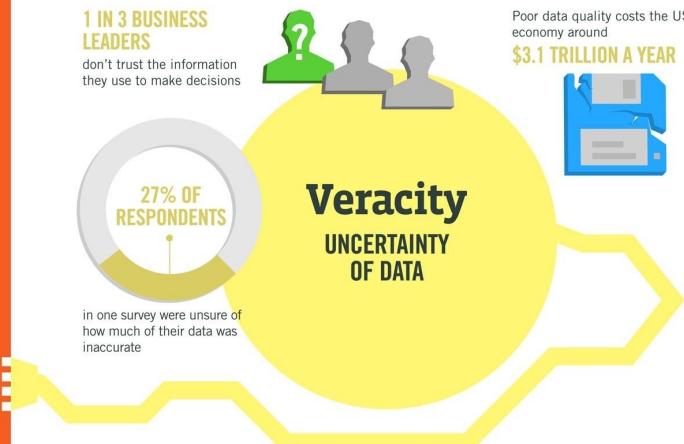
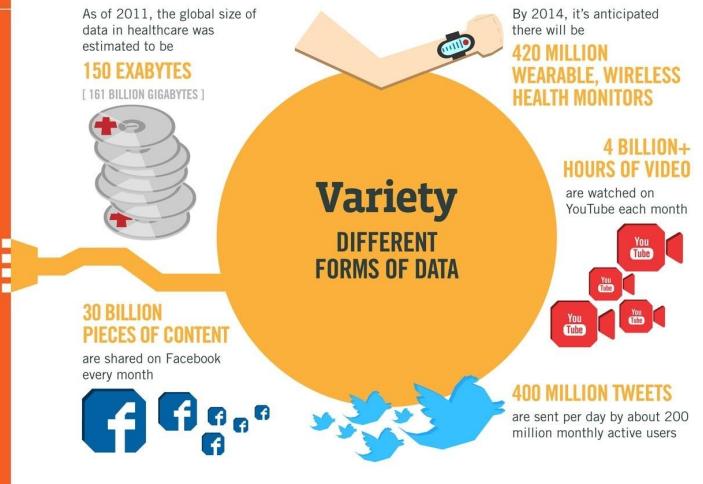
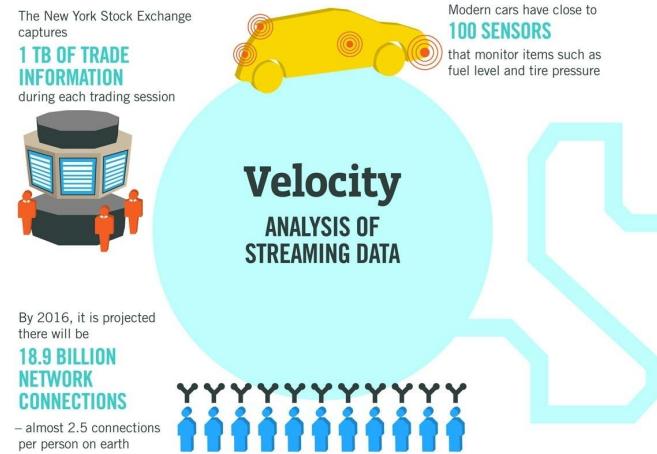
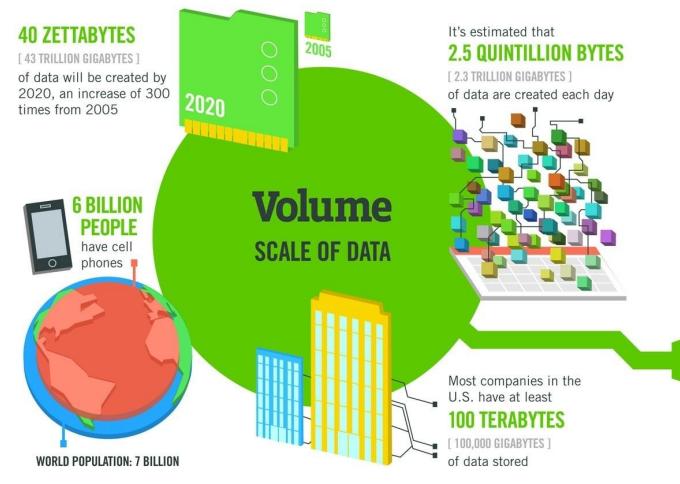
- Organized in attributes
- Well defined schema, follow some order
e.g., Relational tables, tables in Wikipedia

	Temperate climate		Semi-arid climate		Arid climate	
	%	mm	%	mm	%	mm
Total precipitation	100	500–1,500	100	200–500	100	0–200
Evaporation /Evapotranspiration	~ 33	160–500	~ 50	100–250	~ 70	0–140
Groundwater recharge	~ 33	160–500	~ 20	40–100	~ 1	0–2
Surface runoff	~ 33	160–500	~ 30	60–150	~ 29	0–60

3. Unstructured Data

- Natural language text
e.g., News articles, Wikipedia content

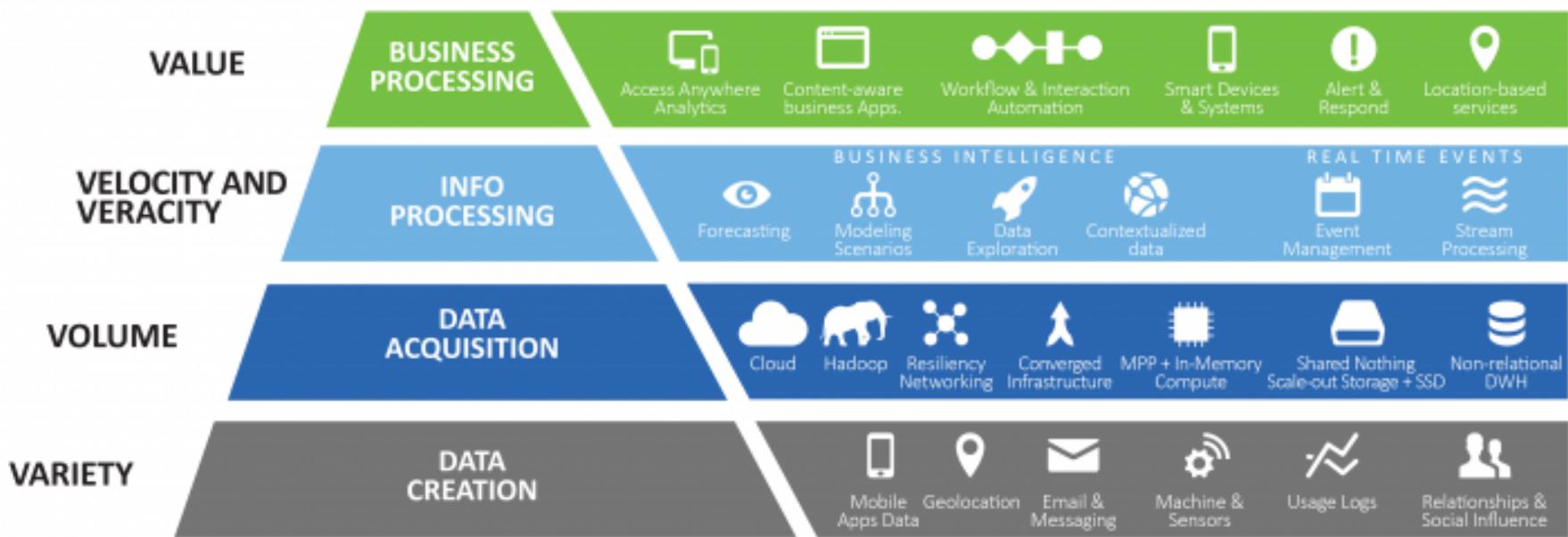
VVV + Veracity



Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

IBM

Big Data: VVVV + Value

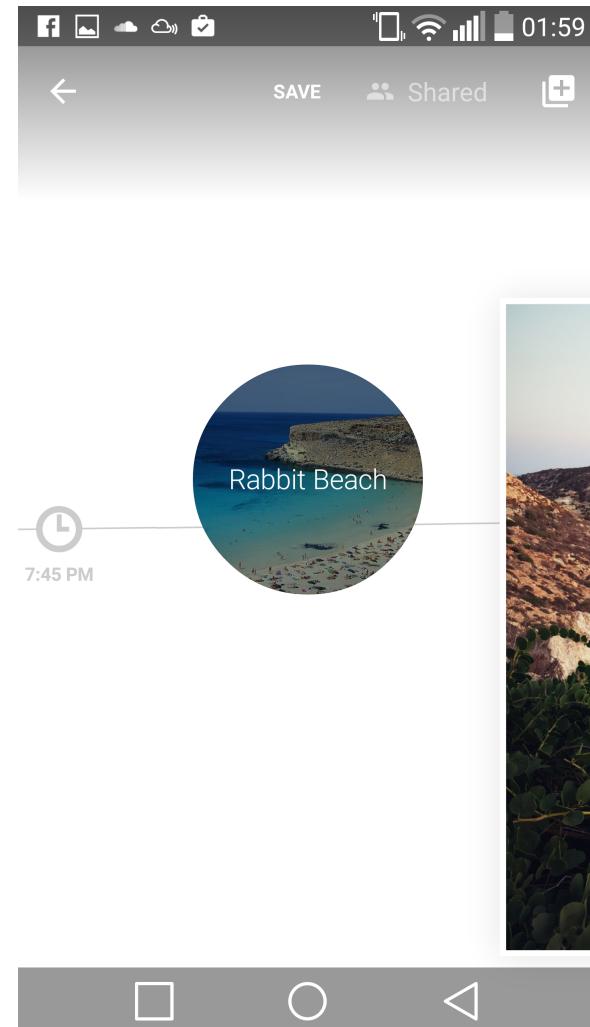
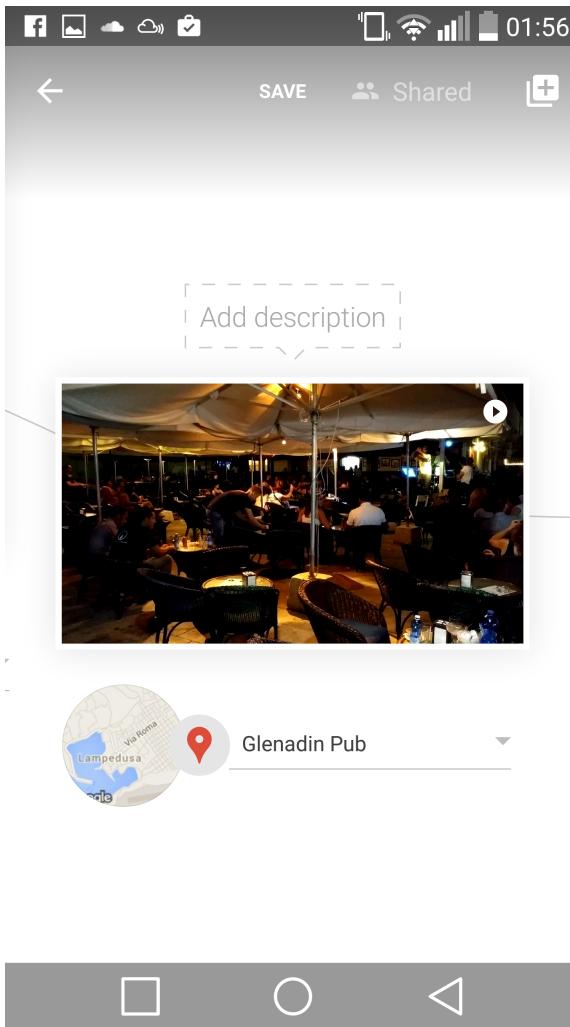


EXCELACOM
© 2015 Excelacom, Inc.

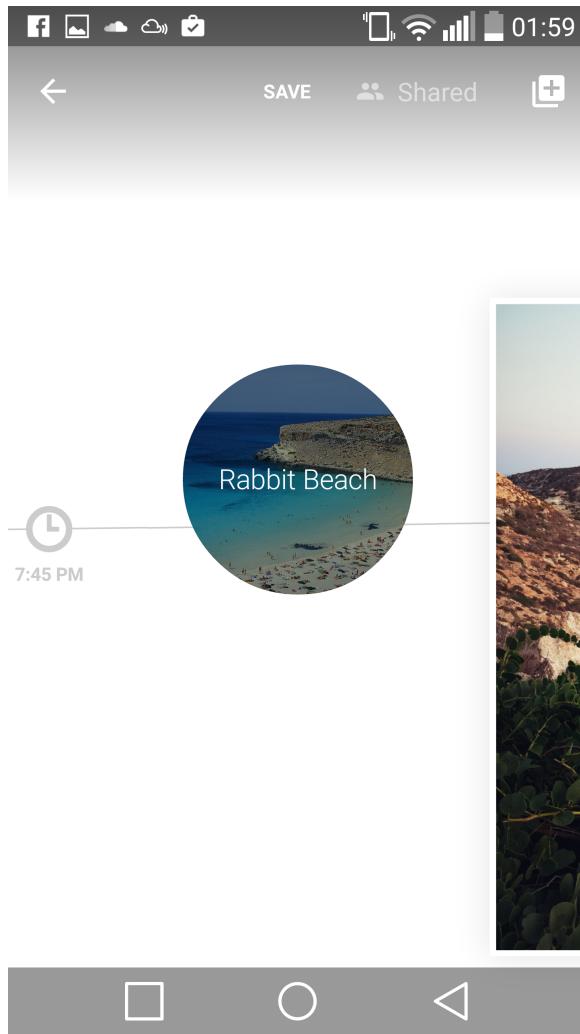
Interpreting and Connecting Web Data

From a Web of documents to
a Web of interconnected
facts

Connecting experiences to locations

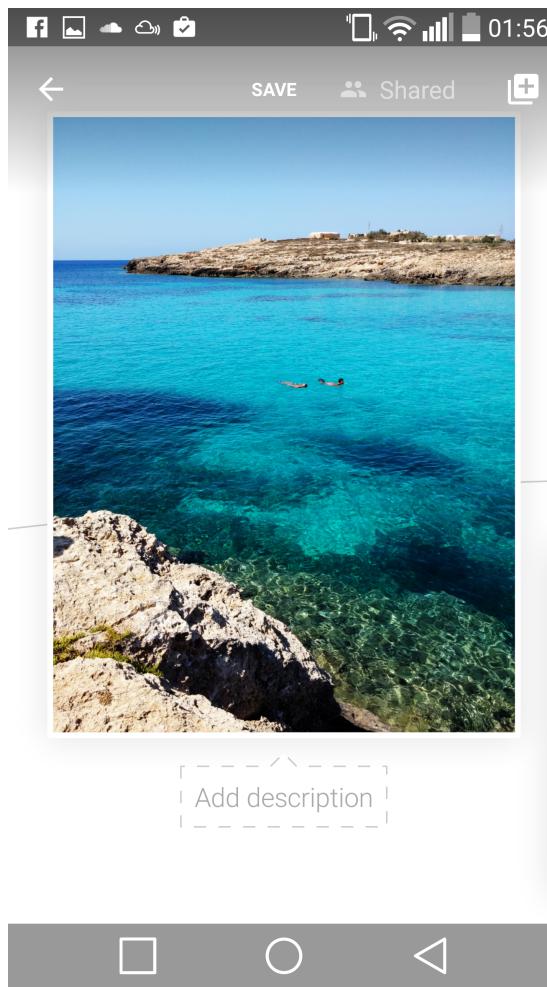


Connecting experiences to locations

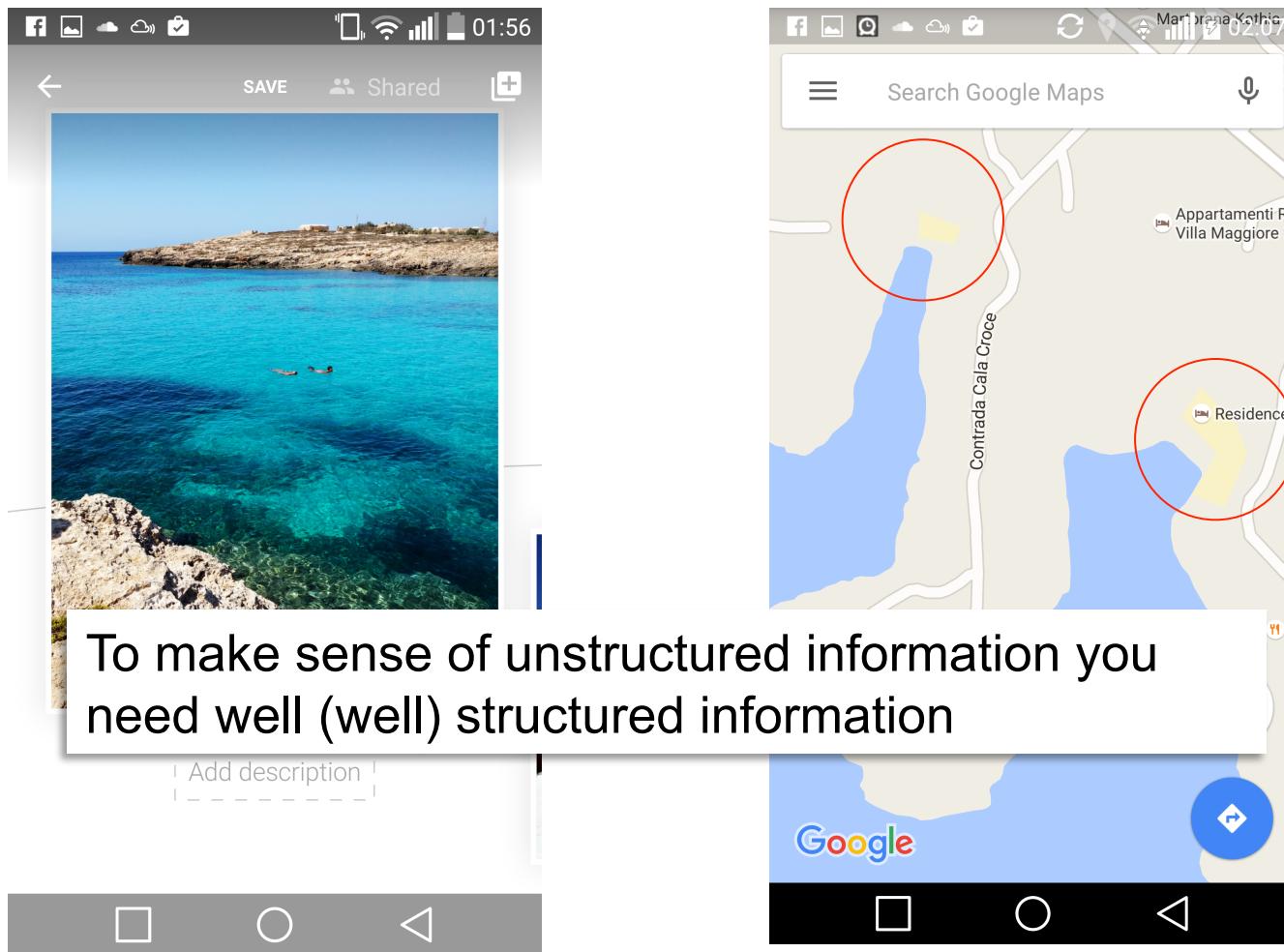


A screenshot of a mobile map application. The search bar at the top says "Lampedusa - Spiaggia dei Conigli". Below the search bar, there is a section for nearby services: "Cerca nelle vicinanze: hotel·ristoranti". The main result is a card for "Lampedusa - Spiaggia dei Conigli" located in "92010 Lampedusa E Linosa AG". It includes two buttons: "Indicazioni stradali" (with a road icon) and "Salva" (with a star icon). Below the card is a thumbnail image of the beach. Further down, it shows "49 Foto", a rating of "4,7 ★★★★☆ 26 recensioni", and links to "Attrazione turistica", "Scrivi una recensione · Aggiungi una foto", and "Suggerisci e modifica · Sei il proprietario?". To the right of the card, a map shows the location of the beach on the island of Lampedusa, marked with a red location pin and the text "Lampedusa - Spiaggia dei Conigli".

The value of reference data for apps



The value of reference data for apps



Documents vs data

Slide by

A. Bordes (Facebook)
E. Gabrilovich (Google)

The screenshot shows a Google search results page for the query "san francisco". A red circle highlights the search bar, and a large red arrow points from the search bar down to the detailed information panel on the right.

Search Bar: san francisco

Search Tools: Web Images Maps Shopping News More Search tools

50 personal results | 1,510,000,000 other results.

Welcome to SFGOV City and County of San Francisco Official site
www.ci.sf.ca.us/
SFGOV is the official website of the government of the City and County of San Francisco, providing information about departments, meetings, legislation, ...

SFO - San Francisco International Airport - Home Page
www.flysfo.com/
Guides on the airlines, concessions, general services, ground transportation and shopping can be found, including flight information, statistics, future ...

San Francisco - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/San_Francisco
San Francisco officially the City and County of San Francisco, is the leading financial and cultural center of Northern California and the San Francisco Bay Area. San Francisco Bay Area - History of San Francisco - 1906 San Francisco earthquake

San Francisco Travel Guide: Things to Do, Hotels, Events ...
www.sanfrancisco.travel/
The official travel and visitors guide for San Francisco. Only In San Francisco can you find San Francisco hotel reservations, tours, flights, maps, popular ...
10 Things Not to Miss in San - Visitor Information Center - San Francisco Events

San Francisco

©2012 Google
Map data ©2012 Google

San Francisco, officially the City and County of San Francisco, is the leading financial and cultural center of Northern California and the San Francisco Bay Area. Wikipedia

Area: 231.9 sq miles (600.6 km²)
Founded: June 29, 1776
Weather: 59°F (15°C), Wind NE at 4 mph (6 km/h), 46% Humidity
Local time: Sunday 3:55 PM PT
Population: 812,826 (2011)

Factual information in search results

Slide by
A. Bordes (Facebook)
E. Gabrilovich (Google)

Top Search Results (Left Column):

- Population, San Francisco, CA** (www.google.com/publicdata)
 - 812,826 - Jul 2011
 - Source: U.S. Census Bureau
- San Francisco** (www.ci.sf.ca.us)
 - Welcome to SFGOV City and County of San Francisco Official site
 - SFGOV is the official website of the government of the City and County of San Francisco, providing information about departments, meetings, legislation, ...
- SFO - San Francisco International Airport - Home Page** (www.flysfo.com)
 - Guides on the airlines, concessions, general services, ground transportation and shopping can be found, including flight information, statistics, future ...
- San Francisco - Wikipedia, the free encyclopedia** (en.wikipedia.org/wiki/San_Francisco)
 - San Francisco officially the City and County of San Francisco, is the leading financial and cultural center of Northern California and the San Francisco Bay Area.
 - San Francisco Bay Area - History of San Francisco - 1906 San Francisco earthquake
- San Francisco Travel Guide: Things to Do, Hotels, Events ...** (www.sanfrancisco.travel)
 - The official travel and visitors guide for San Francisco. Only In San Francisco can you find San Francisco hotel reservations, tours, flights, maps, popular ...
 - 10 Things Not to Miss in San - Visitor Information Center - San Francisco Events

Bottom Search Results (Right Column):

- San Francisco** (www.ci.sf.ca.us)
 - Map data ©2012 Google
 - San Francisco, officially the City and County of San Francisco, is the leading financial and cultural center of Northern California and the San Francisco Bay Area. Wikipedia
 - Area: 231.9 sq miles (600.6 km²)
 - Founded: June 29, 1776
 - Weather: 59°F (15°C), Wind NE at 4 mph (6 km/h), 46% Humidity
 - Local time: Sunday 3:55 PM PT
 - Population: 812,826 (2011)

Rapresenting information (and links) by means of a *knowledge graph*

Slide by

A. Bordes (Facebook)
E. Gabrilovich (Google)

Why (knowledge) graphs?

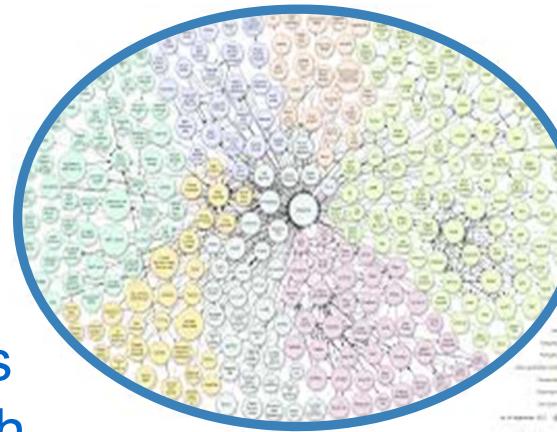
- We're surrounded by **entities**, which are connected by **relations**
- We need to store them somehow, e.g., using a **DB** or a **graph**
- **Graphs** can be processed **efficiently** and offer a convenient **abstraction**

Rapresenting information (and links) by means of a *knowledge graph*

Knowledge graphs



Facebook's Entity Graph



Microsoft's Satori

Slide by
A. Bordes (Facebook)
E. Gabrilovich (Google)



OpenIE
(Reverb, OLLIE)

Google's Knowledge Graph

Documents vs data

Google New York

Web Images News Maps Videos More Search tools

About 716,000,000 results (0.95 seconds)

New York - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/New_York ▾ Wikipedia

New York is a state in the Northeastern and Mid-Atlantic regions of the United States. New York is the 27th-most extensive, the third-most populous, and the ...
 New York City - Albany - List of cities in New York - New York metropolitan area

New York City - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/New_York_City ▾ Wikipedia

For other uses, see NYC (disambiguation) and New York, New ...
 Neighborhoods - History of New York City - Nicknames - Borough

The Official New York City Guide to NYC Attractions, Dining ...
www.nycgo.com/ ▾ New York City

Visit NYCgo for official NYC information on travel, hotels, deals and offers like Restaurant Week, and the best restaurants, shops, clubs and cultural events.
 Must-See NYC - Broadway Shows & Tickets - Events - Tours and Attractions

The New York Times - Breaking News, World News ...
www.nytimes.com/ ▾ The New York Times

There were no reports of survivors on a Malaysia Airlines flight that crashed on Thursday in eastern Ukraine near the Russian border, the scene of fighting ...
 Natalie Glance and one other person +1'd this

New York Magazine -- NYC Guide to Restaurants, Fashion ...
nymag.com/ ▾ New York Magazine

Daily coverage of New York's restaurants, nightlife, shopping, fashion, politics, and culture. NYMag.com is the online counterpart to New York Magazine.

News for new york

New York, Responding to Surge of Child Migrants, Forms ...
 New York Times - by Kirk Semple - 1 hour ago
 Opposition to sheltering a wave of young migrants has mounted in many communities across the country, but in New York City, the reaction has ...

More news for new york

NewYork.com - Your Official Site for Travelling To and Living ...
www.newyork.com/ ▾

Augmenting the presentation
with relevant facts



New York

US State

New York is a state in the Northeastern and Mid-Atlantic regions of the United States. New York is the 27th-most extensive, the third-most populous, and the seventh-most densely populated of the 50 United States. Wikipedia

Capital: Albany

Secretary of State: Cesar A. Perales

Minimum wage: 8.00 USD per hour (December 31, 2013)

Governor: Andrew Cuomo

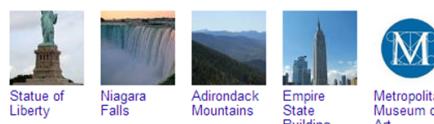
Colleges and Universities: Cornell University, More

Destinations



View 45+ more

Points of interest



View 40+ more

Feedback

Slide by
A. Bordes (Facebook)
E. Gabrilovich (Google)

Rich Snippets in Search Results

The screenshot displays three search results with rich snippets:

- Little Water Cantina - Eastlake - Seattle, WA**
www.yelp.com › Restaurants › Mexican
★★★☆☆ 90 reviews - Price range: \$\$
90 Reviews of Little Water Cantina. "Three things are on my list when I eat out: great food, atmosphere, and
- Vegetarian Vegan Pizza No Cheese) Recipe - Food.com - 248865**
www.food.com/recipe/vegetarian-vegan-pizza-no-c...
★★★★★ 2 reviews - 1 hr 32 mins - 242.9 cal
Aug 26, 2007 - This is from my dad, who developed some **vegan recipes** doesn't have any cheese, and you
- Leonard Cohen – Free listening, videos, concerts, stats, & pictures at ...**
www.last.fm/music/Leonard+Cohen
Watch videos & listen to Leonard Cohen: Suzanne, Hallelujah & more, plus 132 pictures. Leonard Cohen, (born September 21, 1934 in Montréal, Québec, ...
Track Duration
Suzanne 3:48
The Darkness 4:29
Going Home 3:51
Hallelujah 6:12

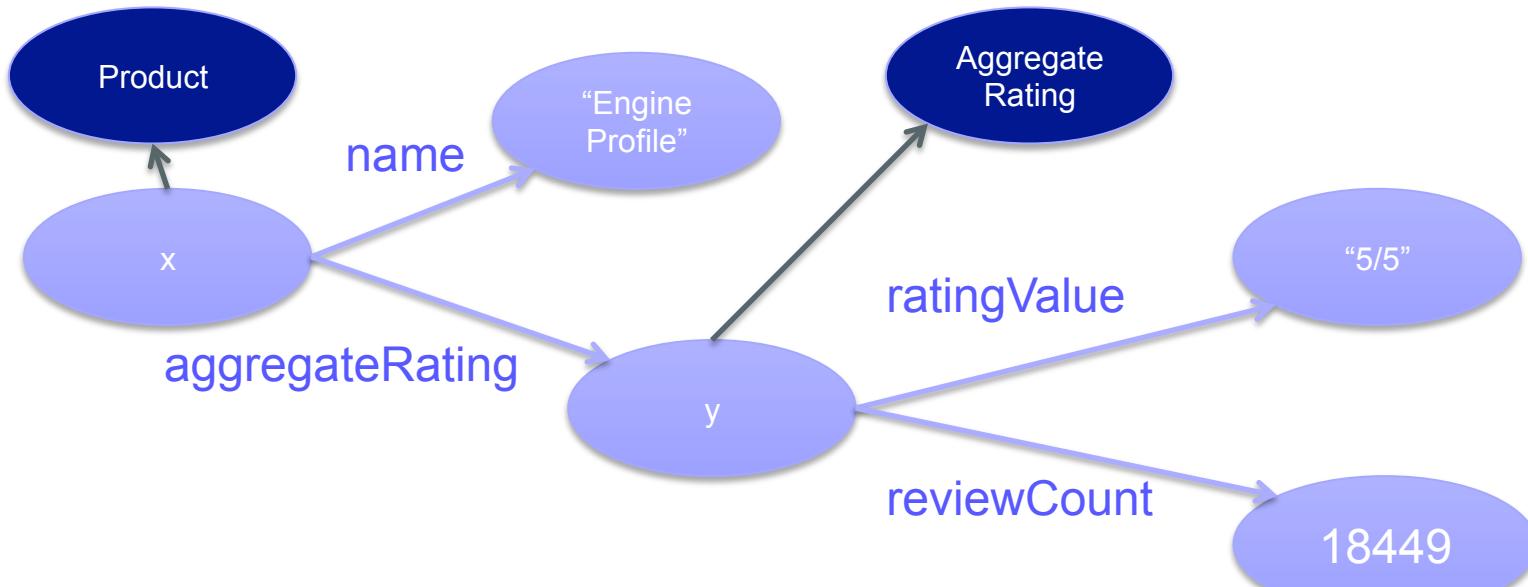
Also useful for
Search Engine Optimization

```
<span itemscope itemtype="http://schema.org/Product">  
<span itemprop="name"><strong>Engine Profiles</strong></span>  
<span itemprop="aggregateRating" itemscope  
      itemtype="http://schema.org/AggregateRating">  
    rated <span itemprop="ratingValue">5</span> / 5<br />  
    based on <span itemprop="reviewCount">18449</span></span>
```



Engine Profiles
rated **5 / 5**
based on **18449**

Rich Snippets in Search Results



```
<span itemscope itemtype="http://schema.org/Product">  
<span itemprop="name"><strong>Engine Profiles</strong></span>  
<span itemprop="aggregateRating" itemscope  
      itemtype="http://schema.org/AggregateRating">  
    rated <span itemprop="ratingValue">5</span> / 5<br />  
    based on <span itemprop="reviewCount">18449</span></span>
```

Engine Profiles
rated 5 / 5
based on 18449

Explorative Search

Sightseeing near New York, NY

- Gray Line New York (2.6 stars, 64 reviews)
- City Sightseeing New York (4.0 stars, 57 reviews)
- Circle Line Sightseeing Cruises (2.2 stars, 60 reviews)
- CitySights NY (3 stars, 15 reviews)
- NYC & Co (3.6 stars, 341 reviews)
- The New York Pass (4.4 stars, 725 reviews)
- Rockefeller Center (4.7 stars, 1,157 reviews)
- The Metropolitan Museum of Art (4.8 stars, 2,081 reviews)
- Empire State Building (4.5 stars, 1,157 reviews)
- Statue of Liberty (2.08 stars, 1,157 reviews)

Gray Line New York Sightseeing Tours, Cruises & Attractions
www.newyorksightseeing.com/ ▾ Gray Line New York ▾
 New York's famous Empire State Building, a New York City and a National Historic ... to Top of the Rock AND 1-hour Statue of Liberty New York Harbor Cruise! NYC double decker tours - Loops Tour Map - All Loops Tour Plus - Contact Us

New York attractions: The 50 best sights and attractions in ...
www.timeout.com/newyork/attractions/.new-york-attractions ▾ Time Out ▾
 by Amy Pitt - Apr 25, 2013 - Sights like the Empire State Building and the Statue of Liberty are perennial favorites, but we've also highlighted newcomers and lesser-known ...

BLDG 92 - Free attractions in New York - Brooklyn Flea - Chrysler Building

New York City Tours and Attractions - NYC Sightseeing ...
www.nycgo.com/toursandattractions/ ▾ New York City ▾
 ITINERARIES. Novel New York, by Jessica Allen. NYC's novelistic life comes alive with self-guided tours to landmarks from books like The Catcher in the Rye, ...

New York: Sightseeing in NYC - TripAdvisor
www.tripadvisor.com/.resonse?Action=GetTourList&TourID=1000000000000000000 ▾ New York (NY) ▾ Before You Go ▾ TripAdvisor ▾
 Inside New York: Sightseeing in NYC - Before you visit New York, visit ... the main sights of the New York City harbor, including the Statue of Liberty, Ellis Island, ...

NYC Sightseeing Tour | New York City Double Decker Tou...
skylinesightseeing.com/ ▾
 See NYC's landmarks your way with our Hop-on, ... The most historic neighborhood in New York City, downtown Manhattan features the Empire State Building, ...

City Sightseeing New York, Hop On - Hop Off Bus Tours
www.city-sightseeing.com/tours/united-states-of-./new-york.htm ▾
 Choose your own way in which you use your ticket according to your own itinerary. There are three tour routes to choose from, allowing you to explore.

Top 25 New York City Tours - New York Magazine
nymag.com/visitorsguide/sightseeing/citytours.htm ▾ New York Magazine ▾
 Aug 16, 2013 - Views From the Top The Big Apple Tour See the city streets from a

Slide by
A. Bordes (Facebook)
E. Gabrilovich (Google)

Connecting People to Locations

The diagram illustrates the interconnected nature of social media data. Three Facebook pages are shown as nodes, each with a blue double-headed arrow indicating bidirectional connectivity:

- Harvard University**: A page for the college/university with 2,703,624 likes.
- People who like Harvard University**: A page for users who have liked the Harvard University page.
- People who visited Harvard University**: A page for users who have visited the Harvard University location.

Below these pages, their respective feeds are displayed:

- Paul McDonald**: Likes Harvard University, iRunFar.com and 182 others. Studied Computer Science at Harvard University '03.
- Ekaterina Skorobogatova**: Likes Harvard University, Loves Company and 3,455 others. Studied Interactive Multimedia at New York University.
- Gary Johnson**: Likes Harvard University and 278 others. Studied at Wharton School, University of Pennsylvania '08.
- Greg Marra (马格雷)**: Likes Harvard University, Emmy's Spaghetti Shack and 693 others. Studied Electrical and Computer Engineering at Franklin W. Olin ...
- Florence Trouche**: Visited Harvard University, Marché Poncelet and 317 other places. Studied at Rouen Business School '90.
- Andrew Tulloch**: Likes Harvard University. Studied Machine Learning at University of Cambridge.
- Joseph Barillari (joeb)**: Visited Harvard University, Philz Coffee At Facebook and 990 other places. Studied Computer Science at Harvard University '07.
- Sheryl Sandberg**: Chief Operating Officer at Facebook. Visited Harvard University and 423 other places. Studied at Harvard Business School.

Slide by
A. Bordes (Facebook)
E. Gabrilovich (Google)

Complex Questions & Decision Making

Complex Questions over Social Networks

Structured search within the graph

The screenshot shows a Facebook search results page with the query "People who like Harvard University and Basketball and work at Facebook". The results list five profiles:

- Mike Vernal**: VP Engineering at Facebook. Likes Harvard University, Harvard Crimson and F@ceb00k Su... Studied Computer Science at Harvard University '02. 10 mutual friends including Keith Adams and Philip Bohannon.
- Jared Morgenstern**: Product Manager / Ninja - Games ... Likes Harvard University, F@ceb00k Summer Basketball Leag... Studied Computer Science at Harvard University. 5 mutual friends including Clodagh Chloe Takeuchi and Pierre ...
- Florin Rățiu**: Software Engineer at Facebook. Likes Harvard School of Public Health, Stanford 6th Man and ba... Studied Management Science and Engineering at Stanford Univ... 3 mutual friends including Alexey Spiridonov and Serkan Plantino
- Ning Zhang (张宁)**: Software Engineer at Facebook. Likes Harvard University, Basketball and 314 others. Studied Computer Science at University of Waterloo '06. 5 mutual friends including Tudor Bosman and Ves Stoyanov
- Zhongyuan Xu (徐重远)**: Software Engineer at Facebook. Likes Harvard University, Basketball and 364 others. Studied at Stony Brook University. 1 mutual friend: Ledell Wu

A blue arrow points from the text "Structured search within the graph" to the search bar at the top of the screenshot.

Question Answering

Slide by A. Bordes (Facebook) & E. Gabrilovich (Google)

Google search results for "Barack Obama place of birth". The results page shows a map of Honolulu, Hawaii, with various landmarks labeled. Below the map, a large image of the city skyline is displayed. The text "Honolulu, HI" is prominently shown, followed by "Barack Obama, Place of birth". Below this, there are two sections of text: one about "Barack Obama citizenship conspiracy theories" and another about "Barack Obama - Wikipedia, the free encyclopedia".

Google

EVI (Amazon) mobile interface. At the top, it shows the time as 10:39. A message bubble says "I thought you asked...". Below it, a question "Where is born Barack Obama" is followed by an answer card: "Barack Obama was born in Honolulu, Hawaii." An "attribution" link and a small image of the city are also present. At the bottom, there are "Good answer" and "Bad answer" buttons.

EVI

(Amazon)

Siri (Apple) mobile interface. At the top, it shows the time as 10:40. A message bubble says "Barack Obama was born in Honolulu." Below it, a question "Where is born Barack Obama" is followed by an answer card: "Barack Obama was born in Honolulu, Hawaii." An "attribution" link and a small image of the city are also present. At the bottom, there is a microphone icon and a text input field that says "Type your question".

Siri

(Apple)

Querying the WWW

- A cinema enthusiast wants to travel in Italy visiting locations of movies she loves...
- **Which movies have been filmed in Umbria?**



In questi giorni gira in rete un video caricato su Youtube contenente gli spezzi di film girati a Perugia, come *Fumo di Londra* (1966), con Alberto Sordi, o *I corpi presentano tracce di violenza carnale* (1973), con Luc Merenda, antesignano del sottogenere horror "Slasher", poi reso celebre dal serial cult Venerdì 13.

Querying the WWW

- A cinema enthusiast wants to travel in Italy visiting locations of movies she loves...
- **Which movies have been filmed in Umbria?**
- **Which movies have been filmed in Umbria outside Perugia?**
- **Which movies with Alberto Sordi have been filmed in Perugia between 1960 and 1979?**

Ask complex questions to the web



Which is the **most successful rock singer** as of today in the UK

Online Marketing: Digital Data Trends

Search, Social, & Content Fusion (by Andy Betts, April 10, 2014)

- “One of the biggest challenges marketers face is understanding what to do with mass data at hand. This was a key theme from the recent [ClickZ Live New York](#).”
- “Identifying what content engages with your digital audience is the key to driving efficient, and scalable, integrated marketing campaigns. The panel was in agreement that content is the catalyst that should be placed at the center of all digital campaigns.”
- key trends that focus on breaking down big data into small pieces to help marketers understand how content, search, and social (and subsequent measurement) are morphing together.



Online Marketing: Digital Data Trends

Search, Social, & Content Fusion (by Andy Betts, April 10, 2014)

- “top factors that impact conversion and how best to optimize campaigns around data.”

Key 2014 Digital Marketing Trends



AI for Data Analytics

- How to make sense of unstructured data for data analytics?
 - Intelligent data integration algorithms (data preparation cover 70% of data analytics projects)
 - Support integration of unknown data sources

Bruxelles, 13 Mar 2017



Horizon 2020 Big Data PPP



Supporting Event and Weather-based Data Analytics and
Marketing along the Shopper Journey

www.ew-shopp.eu

H2020-ICT-2016-2017- Topic: ICT-14-2016-2017 Project
id:732590

Starts on Jan 1st 2017, ends on Dec 31st 2019
Total cost ≈ 3.6M€ - Total funding ≈ 2.9M€

EW-Shopp partners



UNIVERSITY OF MILANO-BICOCCA (Coordinator)

Data Management, Service Science, Service Oriented Architecture



CENEJE (SI)

Leading comparison-shopping platforms in the Adriatic region



BROWSETEL (UK)

Cloud Service Provider offering a platform for a Real-time cross-channel Customer Engagement Platform



GFK EURISKO (IT)

Multinational market research company



BIG BANG (SI)

Biggest Slovenian Consumer Electronics Specialist

MEASURENCE (IE)

IoT platform to analyze consumer behavior in physical places

JOT INTERNET MEDIA (ES)

Specialized digital marketing agency that generates search qualified traffic to premium partners

ENGINEERING (IT)

Largest IT group in Italy

SINTEF (NO)

Scandinavia's largest independent non-profit research organisation

INSTITUT JOZEF STEFAN (SI)

Research institution for natural sciences in Slovenia

EW-SHOOPP



EW-Shopp

EW-Shopp aims at supporting companies operating in the fragmented European ecosystem of the eCommerce, Retail and Marketing industries to increase their efficiency and competitiveness by leveraging deep customer insights.

Improved insights will result from the ***analysis of large amount of data***, acquired from different sources and sectors, hence ***heterogeneous***.

EW-Shopp will provide a

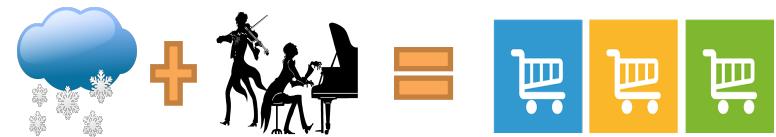
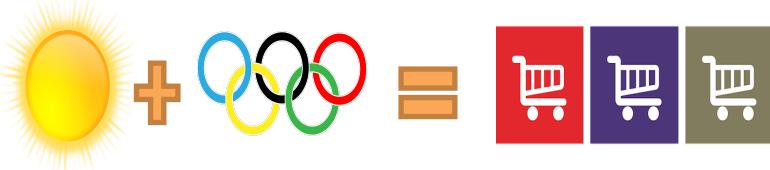
- ***integrate consumers*** across different channels and with different devices
- ***enrich them with information*** about their choices.

Simple example:

Advertise *bicycle equipment* harder on *sunny afternoons* during the *Tour de France*

activities across different countries, and
therefore reacting on consumer

EW-Shopp achievements will ***make up for the lack of knowledge, technology and resources*** that are needed to integrate and analyze large and diverse data in a timely manner ***to optimize digital marketing and resource management.***



Analytics on top of enriched data: examples from EW-Shopp

measurement



Collects data about consumer behavior in locations (e.g., passengers, receipts, etc.)

Find correlation between:

- half-hourly (or hourly) number of visitors, pedestrians and **weather** (i.e. temperature/precipitation)
- daily number of visitors, pedestrians, receipts (only for Kim Scout location) and **weather** (i.e. temperature/precipitation), **local events**
- weekly number of visitors, pedestrians in dependence on **weekly temperature**

Runs digital marketing campaigns, i.e., generates traffic by bidding for ads allocation on specific keywords

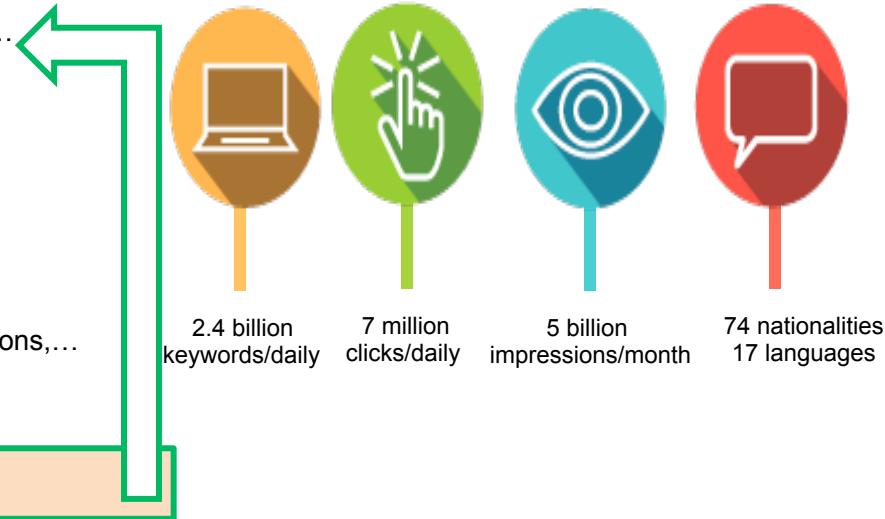
Questions to be answered with the analysis

- When should a campaign be launched, considering also **future events** and **weather forecasts**?
- Which keywords should go in that campaign and which ones should wait, considering also **future events** and **weather forecasts**?

Digital Marketing Campaign Optimization. CHALLENGES

Optimization of campaigns based on:

1. Involves **multiple variables**: CPC, CTR, CR, QS, AvgPos, Day/ Hour...
2. Requires analysis of **huge amount of information**
3. Influenced by **external factors (worldwide)**:
 - Other companies
 - Google/Bing policy and algorithms
 - Social behaviour
 - Environmental events: Weather, sports, concerts, strikes, elections,...



Making sense of digital marketing analytics ... with external events

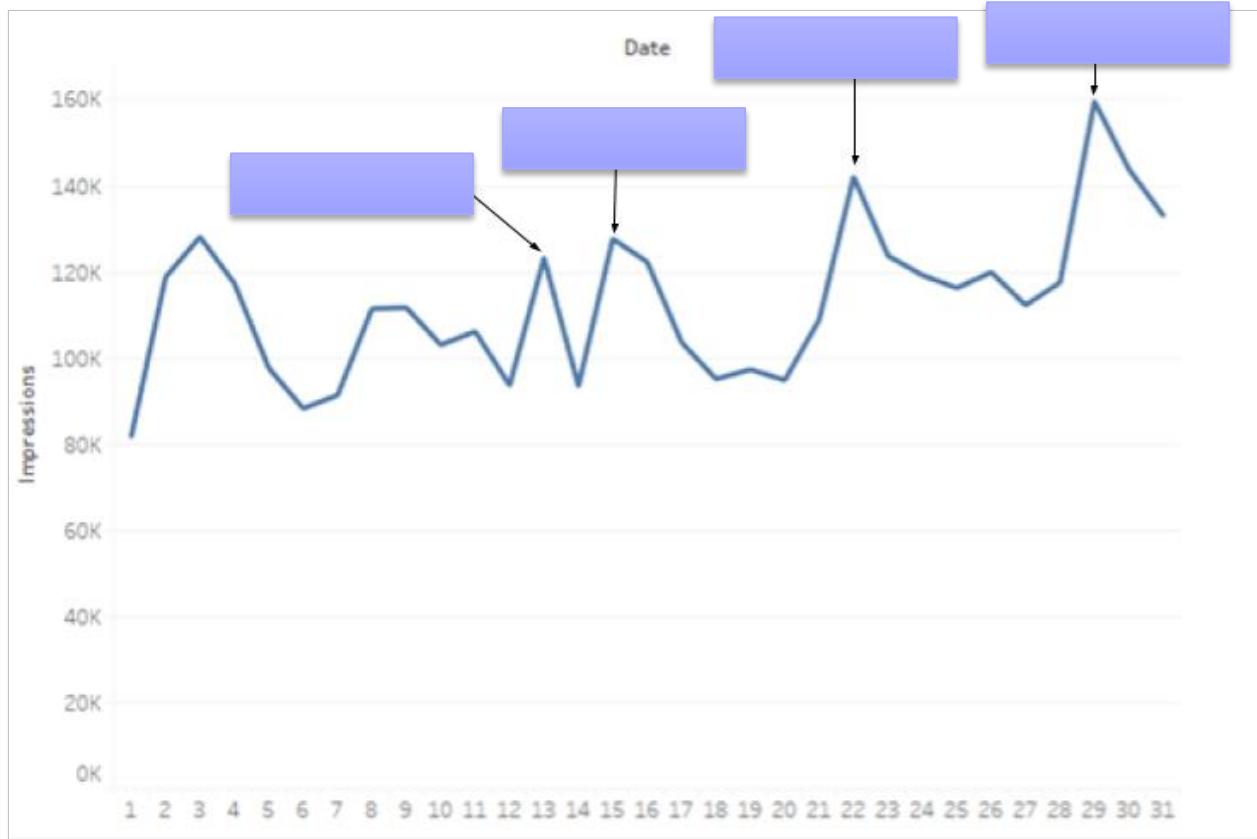
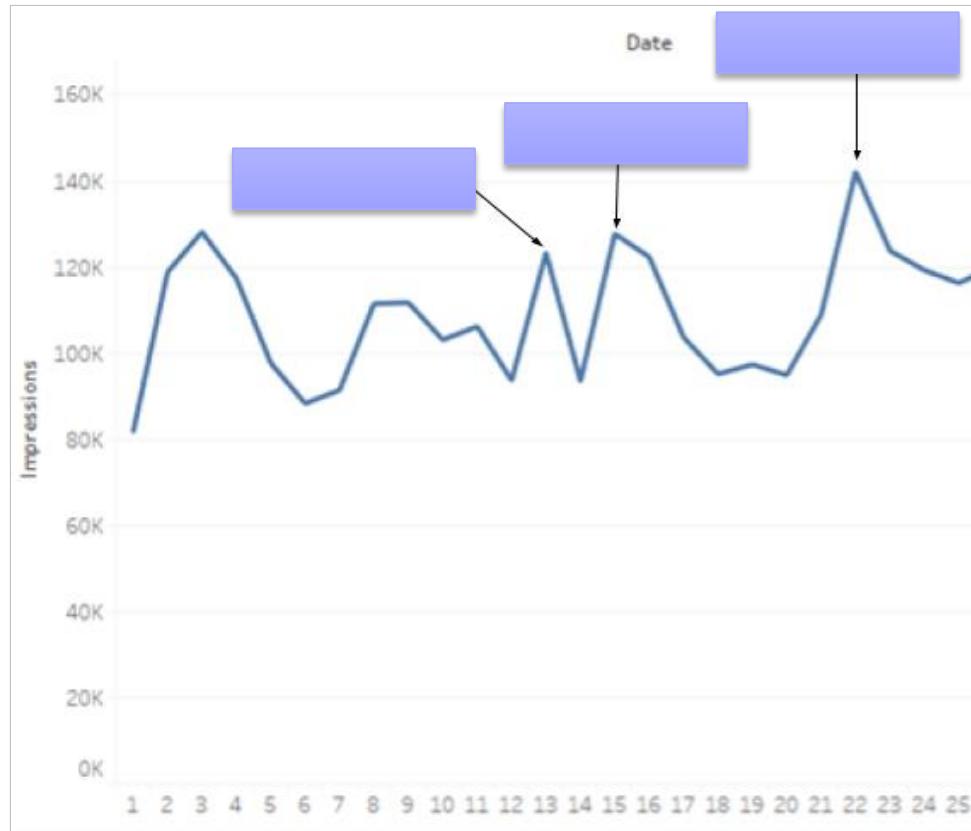


Figure 12: Correlation between **external events** and online digital indicators in the "SportFitness" category.

Making sense of digital marketing analytics ... with external events



germany january 2017

Tutti Immagini Notizie Video Maps Altro Impostazioni Strumenti

Circa 57.000.000 risultati (0,58 secondi)

Upcoming Events in Germany (January 2018), Events Happening in ...
<https://10times.com/germany?month=january> ▾ Traduci questa pagina
 Events in **Germany** in **January**. Follow 14,945 ... Januar 2018 at the ESTREL Convention Center in Berlin, Germany. This event ... PSI 2017 (PSI Dusseldorf).

Immagini relative a germany january 2017

→ Altre immagini per **germany january 2017** Segnala immagini non appropriate

A Month of Islam and Multiculturalism in Germany: January 2017
<https://www.gatestoneinstitute.org/9959/germany-islam-january> ▾ Traduci questa pagina
 19 feb 2017 - "If we are serious about the fight against Islamism and terrorism, then it must also be a cultural struggle." — German Vice Chancellor Sigmar ...

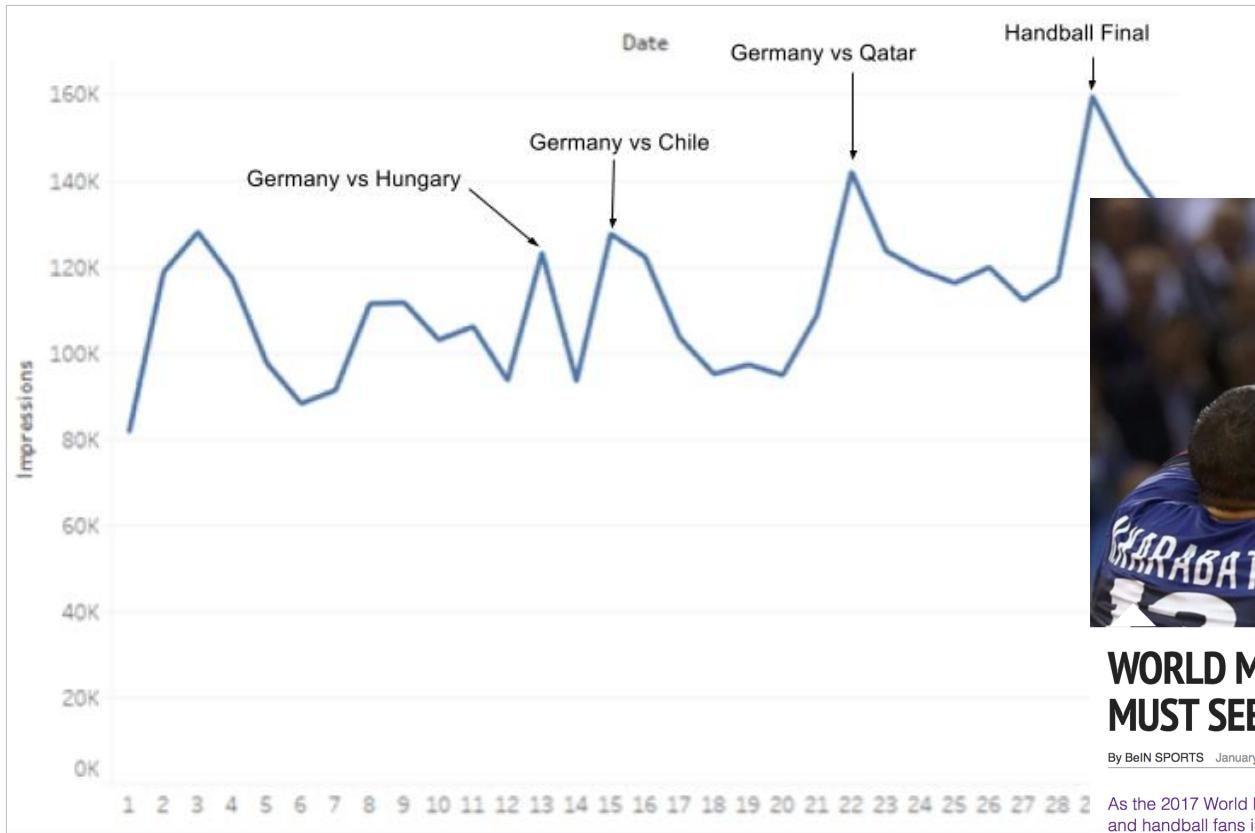
2017 in Germany - Wikipedia
https://en.wikipedia.org/wiki/2017_in_Germany ▾ Traduci questa pagina
 Passa a **January** - January[edit]. January 10: Roman Herzog, German politician, former president (born 1934); Januar 13: Udo Ulfkotte, German journalist ...
 Events · Deaths

Berlin January Weather 2017 - AccuWeather Forecast for Berlin ...
<https://www.accuweather.com/en/de/berlin/.../january.../178087> ▾ Traduci questa pagina
 Get Berlin, Germany typical January Weather including average and record temperatures from AccuWeather.com.

Public Holidays in Germany in 2017 | Office Holidays
www.officeholidays.com/countries/germany/index.php ▾ Traduci questa pagina
 List of National Public Holidays celebrated in Germany during 2017 with information on the ... 3 October 2017 This National holiday is in 14 days. ... Friday, January 06 Jan 6, Three King's Day, Baden-Württemberg, Bavaria, Saxony-Anhalt.

Figure 12: Correlation between Google Trends and online digital indicators in the "SportFitness" category.

Making sense of digital marketing analytics ... with external events



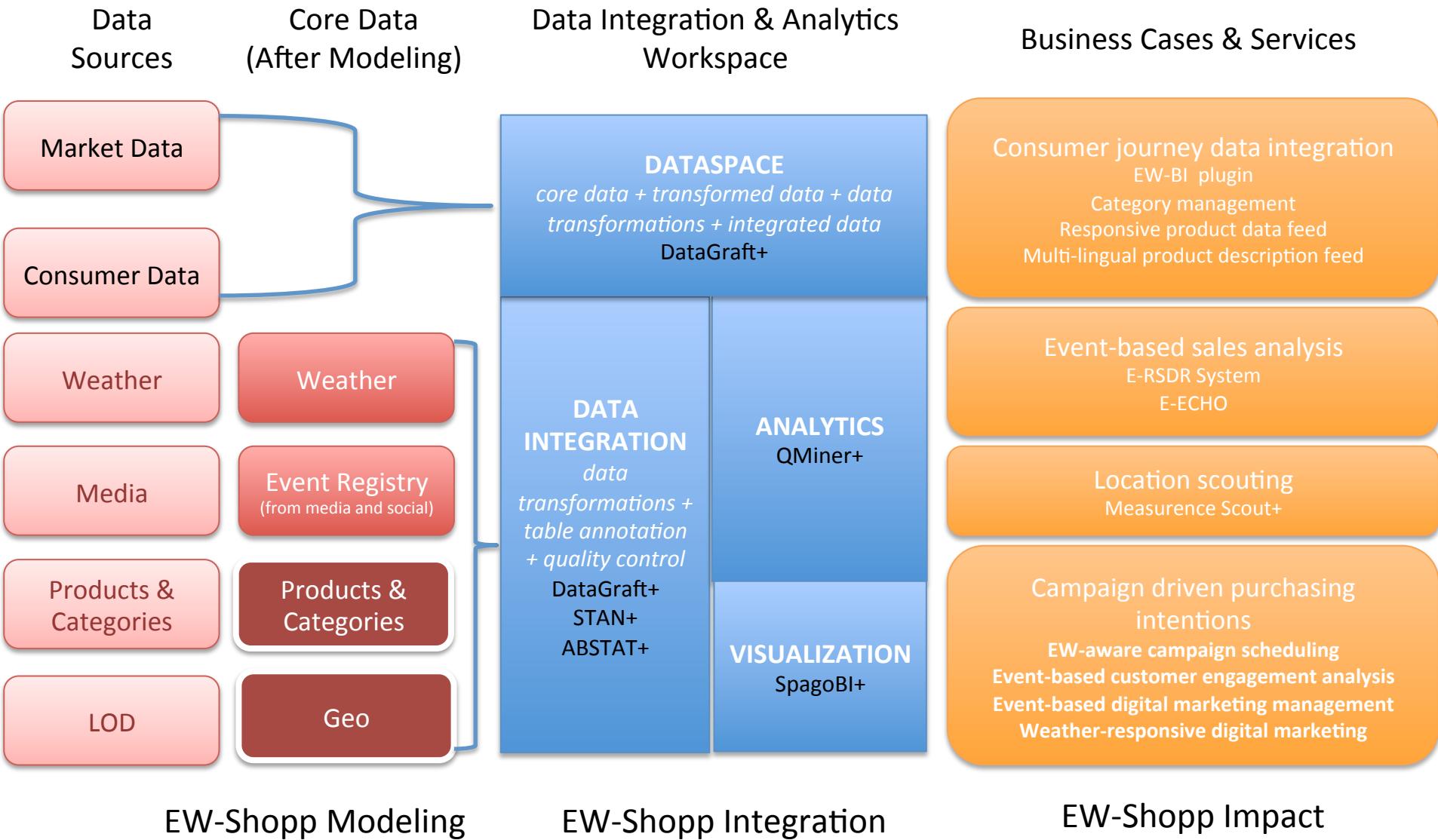
WORLD MEN'S HANDBALL CHAMPIONSHIP MUST SEE MATCHES

By BeIN SPORTS January 11, 2017 11:38 AM

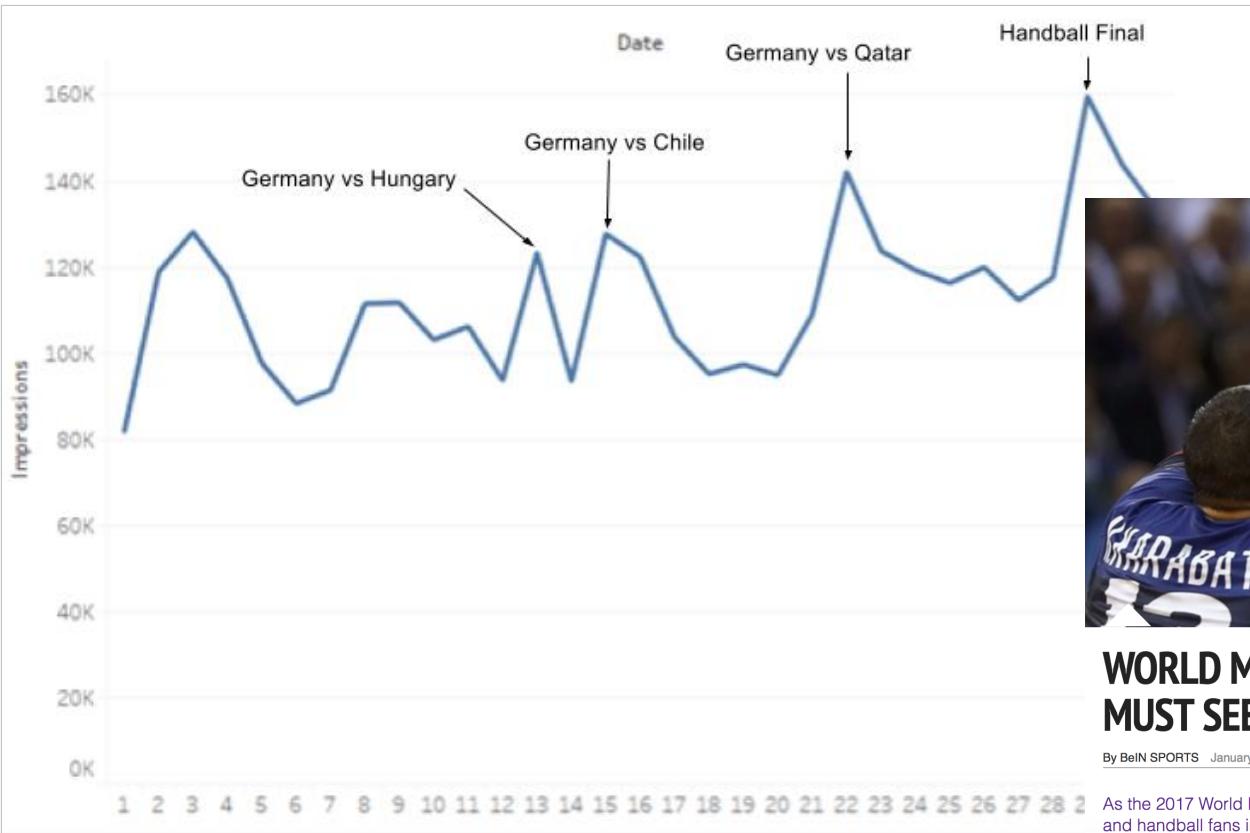
As the 2017 World Handball Championship in France draws near, Sports fans in general and handball fans in particular are looking forward to intense first round showdowns that rival knockout stages.

Figure 12: Correlation between German national handball team matches and online digital indicators in the "SportFitness" category.

EW-Shopp methodology: key components



Making sense of digital marketing analytics ... with external events



WORLD MEN'S HANDBALL CHAMPIONSHIP MUST SEE MATCHES

By BeIN SPORTS January 11, 2017 11:38 AM

As the 2017 World Handball Championship in France draws near, Sports fans in general and handball fans in particular are looking forward to intense first round showdowns that rival knockout stages.

Figure 12: Correlation between German national handball team matches and online digital indicators in the "SportFitness" category.

Event Data



- <http://eventregistry.org>
- Leading platform for large-scale global media monitoring platform
- Daily processing 300 k articles in 30 languages from all over the world
- Semantic annotation and cross-lingual linking

[eventregistry](#) [SEARCH](#) Aljaž Košmerlj

[EVENTS](#) [ARTICLES](#) Filter by: Locations Sources Categories Last month Any language Misc.

Sample queries

Looking for ideas? Try our sample queries:

[HIDE SAMPLES](#)

- Latest news about Bitcoin, Ethereum or Litecoin in English
- Renewable energy
- Events describing natural disasters that occurred in United States
- News about Apple Inc. reported by top business related news sources, sorted by relevance
- Spanish articles about football that had the highest number of shares on social media
- News about Refugees published by news sources located in Germany, sorted by source importance

Current articles and events

Any language [SEE LARGE MAP](#)

ARTICLES EVENTS

CPS drops all charges against former Hillsborough officer
Location: Liverpool, United Kingdom
Articles about event: 86

Bombshell report: Roosters chasing Munster
Location: (unknown)
Mentioned in 71 articles

Leadership spill: Australians 'putting out potatoes' for Peter Dutton
Location: Canberra, Australia
Mentioned in 348 articles

Jose Mourinho: Man Utd manager says Man City 'cannot buy class'
Location: (unknown)
Mentioned in 475 articles

'Definitely no joke': Why you shouldn't write Bolt off
Location: (unknown)
Mentioned in 6 articles

Follow current topics

Mergers and acquisitions

Donald Trump's actions as

Blockchain & Crypto News

Self-driving cars

Event Details

1. Time
2. Location
3. Article contents
4. Related concepts
5. Related events
6. ...

Trump says tariffs on another \$267 billion in China goods ready to go

WHEN
Thu, September 06, 2018

LOCATION
Beijing, China

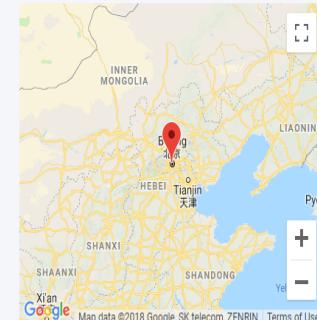
CATEGORIES
Business--International Business and Trade, Society--Issues--Economic, Business--International Business and Trade--Import and Export

President Donald Trump speaks during a press conference with China's President Xi Jinping at the Great Hall of the People in Beijing on November 9, 2017.

President Donald Trump told reporters on Air Force One Friday that he was "ready to go" on tariffs for another \$267 billion in Chinese goods, which would come on top of the \$200 billion in goods already targeted, according to Bloomberg and Reuters.

Already the administration was set to announce it would impose the tariffs on the \$200 billion

Title and summary extracted from: [CNBC](#)



Map data ©2018 Google, SK telecom, ZENRIN | Terms of Use

[f](#) [t](#) [in](#)

[Articles](#)

[Related articles](#)

[Top concepts](#)

[Tag cloud](#)

[Reporting timeline](#)

[Date Mentions](#)

[Article sources](#)

[Similar events](#)

Top concepts

Top concepts describing what the event is about.

Entities [Non-entities](#)

CONCEPT RELEVANCE	96%
China	96.0%
Donald Trump	80.0%
United States	74.0%
Beijing, China	41.0%
Office of the United States Trade...	26.0%
Washington (state), United States	17.0%
Canada	12.0%
European Union	11.0%
Mexico	11.0%
Apple Inc.	10.0%
Bloomberg Television	9.0%
United States Air Force	9.0%
CNBC	8.0%
Japan	8.0%
North American Free Trade Agreement	7.0%
Cisco Systems	7.0%
Federal government of the United States	7.0%
Xi Jinping	6.0%
Consumer electronics	6.0%
Bloomberg L.P.	5.0%

INPUT DATA – Event Data from the Event Registry

Access the Event Registry from the web application

List of events

The list contains the information about events that match the search criteria.

61 events match the search criteria

RESULT SORTING

**Smarter text:
Text + annotations
Organization/structure**

How Fiat Chrysler's diesel woes differ from VW scandal

WHEN: Thu, January 12, 2017

WHERE: Milan, Italy

CATEGORIES: Computers , Business , Science , Science→Environment , Science→Environment→Air Quality , Computers→Programming , Business→Telecommunications , Business→Telecommunications→Equipment , Computers→Programming→Software Testing

ABOUT: [Fiat Chrysler Automobiles](#) [United States of America](#) [United States Environmental Protection Agency](#) [Fiat Automobiles](#) [Software](#) [Volkswagen](#) [Diesel engine](#) [Pollution](#) [Automotive industry](#) [Automobile](#) [Air pollution](#) [Chrysler](#)

ARTICLES: 669 (386 English, 64 German, 81 Spanish, 31 Portuguese, 107 French)

After the U.S. Environmental Protection Agency took action, Fiat Chrysler Automobiles found itself with its own diesel emissions debacle on its hands Friday -- the possibility of fines, repairs and a publicity black eye that go with it. But even though EPA says the automaker is in violation of the Clean Air Act because of excessive diesel emissions in 104,000 Jeep Grand Cherokee SUVs and Ram 1500 EcoDiesel pickups, Fiat Chrysler insists there are vast differences between its case and the scandal ([USA Today](#))

UK seeks details from U.S. over Fiat Chrysler diesel emission accusations

WHEN: Thu, January 12, 2017

WHERE: Milan, Italy

CATEGORIES: Science , Science→Environment , Science→Environment→Air Quality , Science→Environment→Environmental Health

ABOUT: [United States of America](#) [Software](#) [Volkswagen](#) [Fiat Chrysler Automobiles](#) [United States dollar](#) [United States Environmental Protection Agency](#) [Automobile](#) [Fiat Automobiles](#) [Dodge Ram](#) [Jeep Grand Cherokee](#) [Diesel engine](#) [Nitrogen oxide](#)

ARTICLES: 144 (6 English, 64 German, 74 Italian)

"We are urgently seeking further information from the US Environmental Protection Agency... and will also be seeking information from the manufacturer regarding vehicles in the UK market," a spokesman at the Department for Transport said on Friday. Fiat Chrysler Chief Executive Sergio Marchionne angrily rejected the allegations on Thursday, saying there was no wrongdoing and the company never attempted to create software to cheat emissions rules by detecting when the vehicle was in test mode. ([FOCUS Online](#))

Union für Untersuchungsausschuss im Fall Amri

WHEN: Sat, January 14, 2017

WHERE: Milan, Italy

or through Python APIs

```
{
  "events": {
    "pages": 4,
    "results": [
      {
        "totalArticleCount": 669,
        "summary": { ... },
        "articleCounts": { ... },
        "eventDate": "2017-01-12",
        "categories": [ ... ],
        "concepts": [ ... ],
        "uri": "eng-2798668",
        "wgt": 47,
        "location": { ... },
        "id": 5557590,
        "title": { ... }
      },
      ...
    ],
    "page": 1,
    "count": 20,
    "totalResults": 61
  }
}
```

OUTPUT – Business data joined with events



Date	Rece ipts	visits _OH	walkby_ OH	Event Name	About	Articles	Snippet
2017-01-10	40	48	1094	### Morning note: Intesa Sanpaolo, la giornata dei mercati -2-	Milan, Metre, USA	20 (20 italian)	09:08 (Il Sole 24 Ore Radiocor Plus) - Milano, 13 gen - Stati Uniti - Il PPI a dicembre e' atteso in rialzo ...
2017-01-10	40	48	1094	Cina: -6,1% a/a esportazioni dicembre, +3,1% a/a importazioni - MilanoFinanza.it	Disc jockey, Миллиар д, China	20 (20 Italian)	MILANO (MF-DJ)--Le esportazioni cinesi (in dollarì) sono scese nettamente a ...
2017-01-10	40	48	1094	COMMENTO STAR: indice poco mosso, Panaria sale - MilanoFinanza.it	Italy Star, Disc jockey	5 (5 Italian)	MILANO (MF-DJ)-- L'indice MF Italy Star sale dello 0,09% in questi primi scambi di seduta ...

Event Data joined on *Date* – it is necessary to replicate each row as many times as many events are found for that date

OUTPUT – Business data joined with event stats



Date	Receipts	visits_OH	walkby_OH	eng_ac	recreation_cat	total_events	undefined_cat	deu_ac
2017-01-10	40	48	1094	7	0	12	11	14
2017-01-11	41	38	1420	0	0	13	13	0
2017-01-12	38	44	1258	435	1	20	13	128
2017-01-13	43	61	1470	18	0	12	9	9
2017-01-14	85	170	1579	0	0	4	4	47

Data joined on *Date* and grouped by Category, Articles count per Country: total number of events per category



euBusinessGraph

- **Predict the behaviour of a company**

Will company X buy new products, open a new office?

- **Recommendation systems**

If you are interested in company X you may also be interested in Y

- **Network analysis**

Building corporate network, influence of people in companies

- **Statistical models**

Computing credit score

- **Fact extraction**

Company X has launched a new product or service

Company X established a partnership with Company Y

- **Data quality problems and data profiling**

Combining data can surface inconsistencies and contradictions



SPAZIODATI

WHO WE ARE



**BIG DATA COMPANY
SPECIALIZED IN BUSINESS KNOWLEDGE**

2 locations in Italy: Trento, Pisa

25 Sw Eng. & Data Scientists (~30 yrs avg)
> 100 man-yrs in R&D



A single point of access to company data for lead generation / sales and marketing intelligence.

ATOKA WEB APPLICATION

The screenshot shows the Atoka web application's search results page. At the top, there are filters for 'AZIENDE' (Companies), 'PERSONE' (People), 'NOTE', 'LISTE', 'RICERCHE', and 'NOTIZIE'. A search bar shows 'Nome azienda o PIVA' with 'Sd - Luca Barsotti' and 'Enterprise' selected. Below the search bar, it says '509 aziende trovate' (509 companies found). A 'Salva ricerca' (Save search) button is visible. The main area displays search results for three companies: S.I.C.O.T., FRIULI INNOVAZIONE, and ENEL - SPA. Each result card includes basic information like sector (Settore ATECO), address (Sede legale), revenue (Ricavi), number of employees (Dipendenti), website (Sito web), and social media links (Social Networks). There are also buttons for 'TUTTI I DETTAGLI' (View all details) and 'Nelle tue liste' (Add to your lists).

ATOKA ENTERPRISE EDITION

ATOKA API

<https://atoka.io/>

FEATURES:

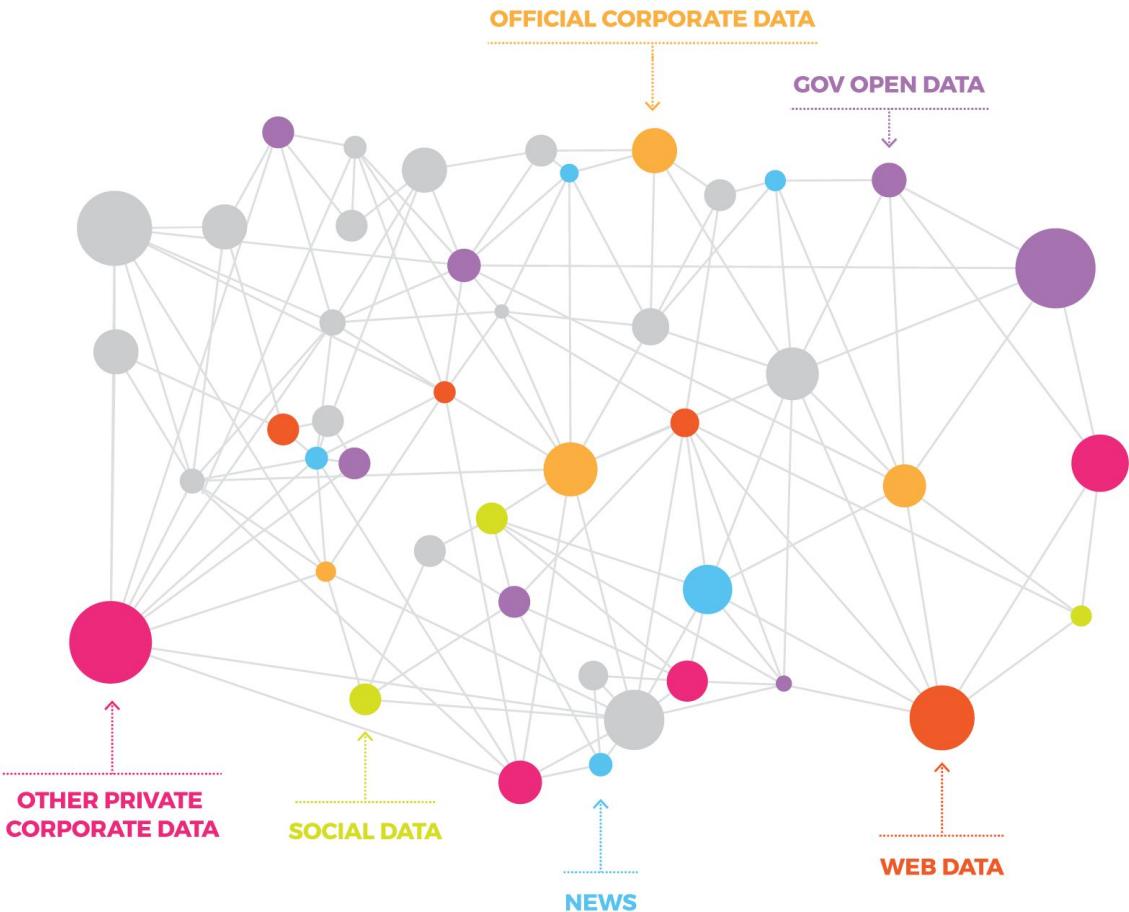
- COMPANY SEARCH & DETAILS
 - JOB ADS
 - GOVERNMENT CONTRACTS
- CONTACTS
- LISTS
- SIMILARITY
- NEWS SEARCH
- MONITORING

POSSIBLE CUSTOMIZATIONS:

- BYOD / SOFT INTEGRATION

(programmatic access to all of the above)

atoka DATA SOURCES



MORE AND MORE DATA

We collect and connect an ever growing number of business information sources.

Just in Italy we crunch:

- 6M companies
- 12M key people
- 1M company websites
- 70k daily news

Beyond Italy, we have UK companies data and 2 other countries coming in 2017.



SEMANTIC TECHNOLOGY

Thanks to our **Semantic text analysis**, we can make sense of raw data by identifying:



COMPANIES



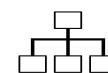
DESCRIPTIONS



CONTACTS



KEY PEOPLE



BIO & ROLES



BUSINESS METRICS



KEY ENTITIES
(Activities & products)



SOCIAL
ACCOUNTS



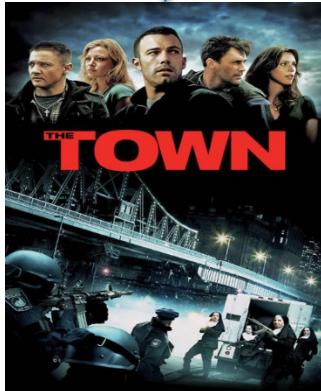
COMPANY AND KEY PEOPLE
MENTIONS (news)

Making sense of social media content

- How has the film “The Town” been received?
- What people think about Ben Affleck?
- How has the TV series “The Town” been received?

Linking Tweets to Real-world Entity Descriptions (in a Knowledge Base)

The Town might be one of the best movies I have seen all year. So, so good. And don't worry Ben, we already forgave you for Gigli. Really.



http://dbpedia.org/page/Ben_Affleck

[https://en.wikipedia.org/wiki/The_Town_\(2010_film\)](https://en.wikipedia.org/wiki/The_Town_(2010_film))



<http://dbpedia.org/page/Gigli>

Linking Tweets to Real-world Entity Descriptions (in a Knowledge Base)

The **Town** might be one of the best movies I have seen all year. So, so good. And don't worry **Ben**, we already forgave you for **Gigli**. Really.



http://dbpedia.org/page/Ben_Affleck

[https://en.wikipedia.org/wiki/The_Town_\(2012_TV_series\)](https://en.wikipedia.org/wiki/The_Town_(2012_TV_series))



<http://dbpedia.org/page/Gigli>

“The” linking problem:
Natural language is ambiguous!

Open Data: Tabelle

Agenzie Turistiche Umbria

Query?

Integration?

	A	B	C	E	F	G	H	I
1	tipo	lingua	id contenuto	denominazione	indirizzi	telefoni	fax	email
2	Agenzia viaggio	it_IT	35362	YELLOW RABBIT TRAVEL	[06012 CITTA' DI CASTELLO PG Via E. Kant 29/G]	075/8511766	075/8511486	turismo@autonoleggivaltiberina.it
3	Agenzia viaggio	it_IT	35372	HERMITAGE TRAVEL FILIALE	[05015 FABRO 55011 TR Via del Casermone,54] [80763/832149 081/8706966	0763/832149 08	contabilita@hermitagetravelsrl.it booking@h	
4	Agenzia viaggio	it_IT	35382	ADMIRE	[06122 PERUGIA 54039 PG Via del Sole, 6]	075/5730808 3356247240	075/5730303	admire.italy@libero.it
5	Agenzia viaggio	it_IT	35392	SPAZIO E TEMPO LIBERO VIAGGI	[05100 TERNI 55032 TR Via I Maggio, 13]	0744/59679	0744/406943	info@spaziotempolibero.it
6	Agenzia viaggio	it_IT	35402	STRAVAGARIO VIAGGI E TURISMO	[05018 ORVIETO 55023 TR C. Cavour, 97]	0763/340890	0763/343422	stravagario@tiscali.it
7	Agenzia viaggio	it_IT	35412	INVIAGGI	[05100 TERNI 55032 TR V.le della Stazione 27]	0744/803711	0744/817404	contabilita@inviaggi.it info@inviaggi.it
8	Agenzia viaggio	it_IT	35422	AMBASSADOR TRAVEL SERVICE FILIALE	[05021 ACQUASPARTA 55001 TR Via Roma, 9]	0744/930831 0541/605509	0744/930831 05	ambassador@ambassadortravel.it
9	Agenzia viaggio	it_IT	35432	MELPIDEA VIAGGI	[05015 FABRO 55011 TR P.le C.Levi, 17]	0761/556566 3493818827	0761/1760298	melpidea@sensidiaggi.it
10	Agenzia viaggio	it_IT	34522	IMPAEKT	[06062 CITTA' DELLA PIEVE PG Via G. Marconi, 1]	0578/298446 3461380007		haret@impaekt.com info@impaekt.com
11	Agenzia viaggio	it_IT	34532	LOCCI VIAGGI	[06049 SPOLETO 54051 PG Piazza Garibaldi, 2]	0743/54415	0743/202677	fabio@locchiaggi.it
12	Agenzia viaggio	it_IT	34542	ITALSPRING	[06121 PERUGIA 54039 PG Via Baglioni, 12]	075/5731732	075/5090929	info@italspring.it
13	Agenzia viaggio	it_IT	34552	VENTO DI VACANZE	[05022 AMELIA 55004 TR Via Europa, 48]	0744/978673	0744/060064	info@ventodivacanze.it
14	Agenzia viaggio	it_IT	34562	ANTHOS VIAGGI	[06126 PERUGIA 54039 PG Via della Pallotta, 18/c]	075/30465	075/36194	anthos@anthosviaggi.it
15	Agenzia viaggio	it_IT	34572	I VIAGGI DI MACOS	[06121 PERUGIA 54039 PG Via dei Filosofi, 25/C]	075/36396	075/34015	macosviaggi@tiscali.it
16	Agenzia viaggio	it_IT	34582	MENIGATTI	[06132 PERUGIA 54039 PG Via del Pettiroso, 4]	075/5149707	075/774611	amministrazione@menigattiviaggi.it

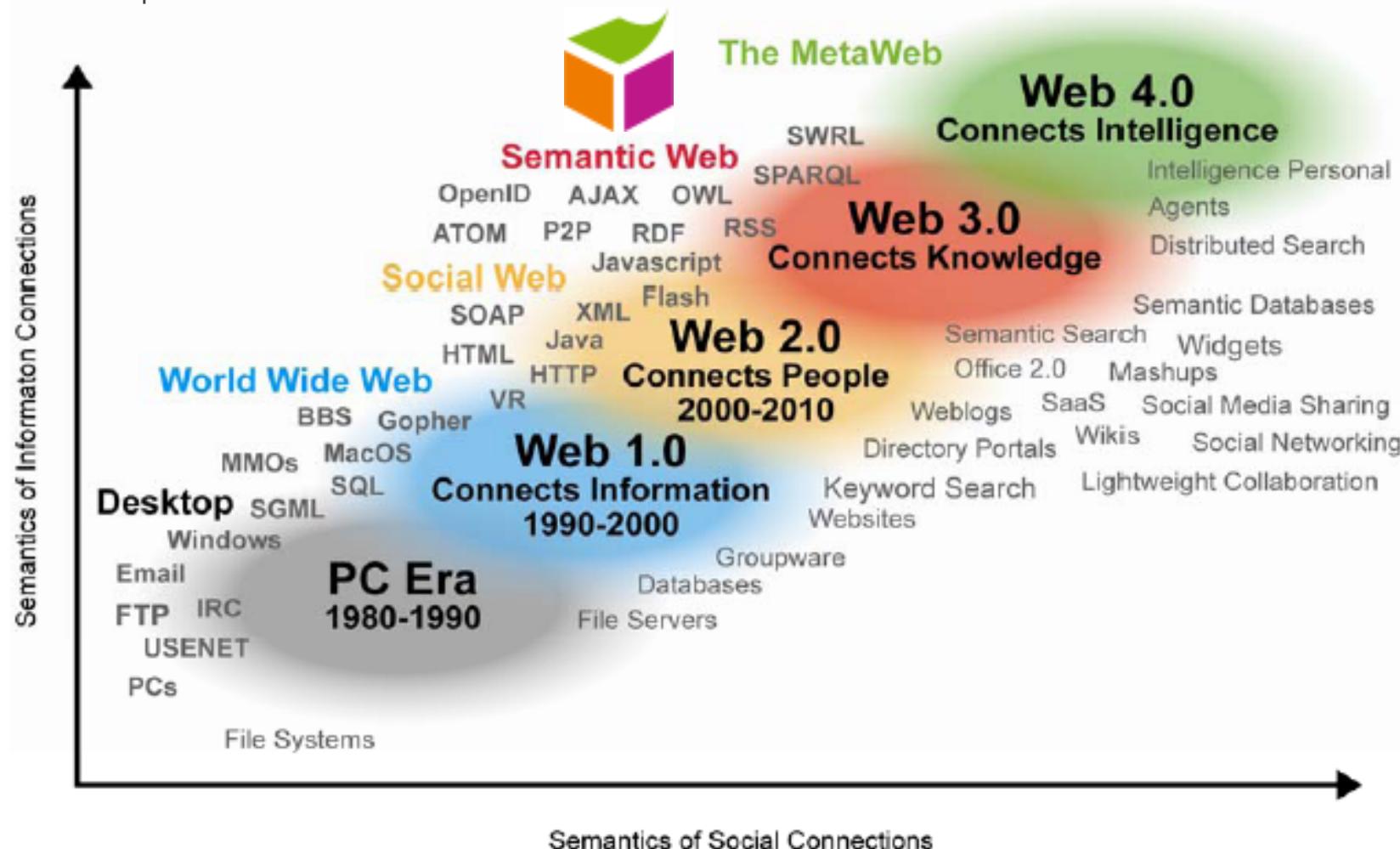
Eventi

What is the meaning of these data?

	A	B	C	D	E	F	G	H	I	J	K	L
1	tipo	lingua	id contenuto	id contenuto	url risorsa	titolo	descrizione					
2	Evento	it_IT	2916883	90470	http://www.umbriatourism.it/-/it-fits	IT FITS	Il 25 settembre, ad Assisi, la terza edizione del Forum Italiano sul Turismo e la Sostenibilità.	25/09/15		26/09/15		
3	Evento	it_IT	2032108	90422	http://www.umbriatourism.it/-/il-lusso-del-s	Il Lusso del Sonno	Fino al 31 agosto all'Antiquarium di Corciano la mostra dedicata ai letti etruschi in bronzo rinvenuti	21/08/15		31/08/16		
4	Evento	it_IT	2032245		http://www.umbriatourism.it/-/ecobike	Ecobike	Fino a ottobre escursioni e visite guidate attraverso l'Umbria in sella a una bicicletta elettrica.	21/08/15		31/10/15		
	M	P	Q	T	U	V	W	X	Y	Z	AA	AB
1	titolo testo	caregorie evento	immagini evento	immagine copertina evento	testo alternativo immagi	latitudine	longitudine	comune	codice IS	via o piazza	localitv†	cap
2	Forum Italiano sul Turismo e la Sostenibilità	Mostre e rassegne Assisi	http://www.umbriatourism.it	http://www.umbriatourism.it	IT FITS	430.707.017	12.619.596.600.000.000	Assisi	54001			
3	I letti in bronzo di Strozziacapponi	Mostre e rassegne Corciano	http://www.umbriatourism.it	http://www.umbriatourism.it	Il lusso del Sonno	431.237.872	12.289.259.600.000.000	Corciano	54015			
4	Alla scoperta dell'Umbria con la bici elettrica	Eventi sportivi	http://www.umbriatourism.it	http://www.umbriatourism.it	Ecobike							

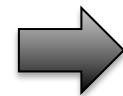
The Web: Evolution

Source: <http://www.radarnetworks.com>

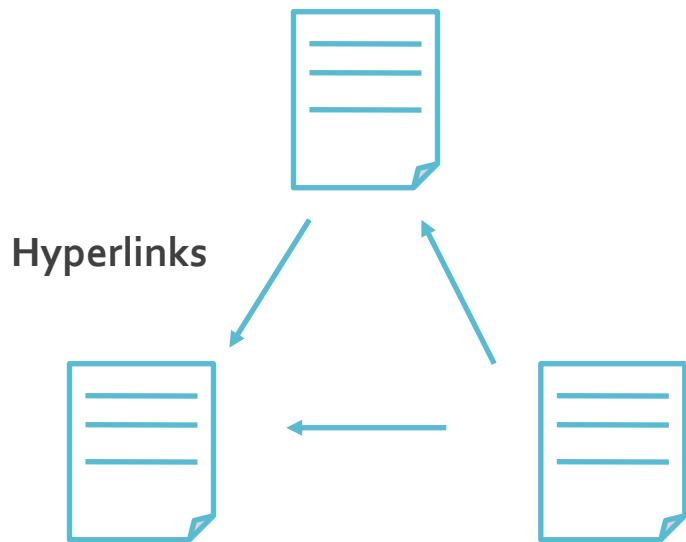


The Web: Evolution

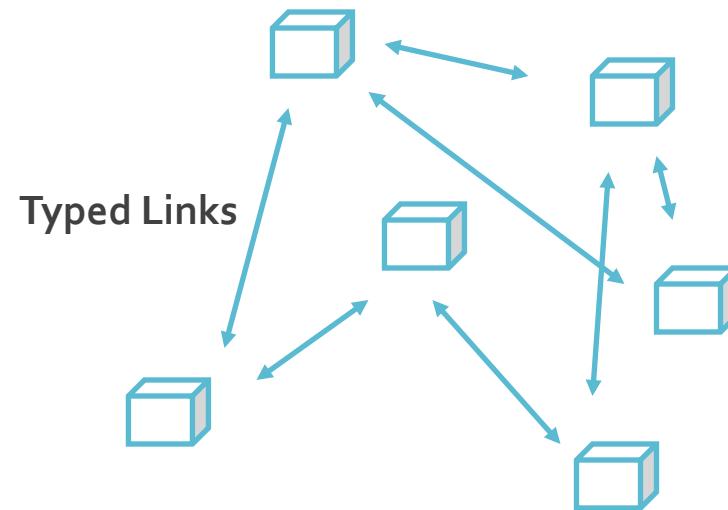
Web of Documents



Web of Data



"Documents"



"Things"

From Documents to ~~data~~ Knowledge Graphs

- Entities and links between entities
- Entities and their properties
- Machine-processable: support for complex queries
- Integration of texts, images, video with structured data (→ knowledge)
- Significant investments to build KGs

From Documents to ~~data~~ Facts

- **Facts** about relations between two entities
- **Facts** about entities
- Machine-processable: support for complex queries **over these facts**
- Integration of texts, images, video with structured data (→ **facts**)
- Significant investments to build **fact-based knowledge bases**

Knowledge Graphs in the Industry

DOI:10.1145/3331166

 Article development led by **ACMqueue**
queue.acm.org

**Five diverse technology companies
show how it's done.**

BY NATASHA NOY, YUQING GAO, ANSHU JAIN,
ANANT NARAYANAN, ALAN PATTERSON, AND JAMIE TAYLOR

Industry-Scale Knowledge Graphs: Lessons and Challenges

Natasha Noy, Yuqing Gao, Anshu Jain, Anant Narayanan, Alan Patterson, Jamie Taylor. Communications of the ACM, August 2019, Vol. 62 No. 8, Pages 36-43

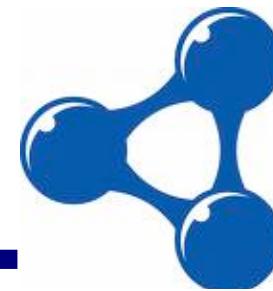
Common characteristics of the knowledge graphs.

	Data model	Size of the graph	Development stage
Microsoft	The types of entities, relations, and attributes in the graph are defined in an ontology.	~2 billion primary entities, ~55 billion facts	Actively used in products
Google	Strongly typed entities, relations with domain and range inference	1 billion entities, 70 billion assertions	Actively used in products
Facebook	All of the attributes and relations are structured and strongly typed, and optionally indexed to enable efficient retrieval, search, and traversal.	~50 million primary entities, ~500 million assertions	Actively used in products
eBay	Entities and relation, well-structured and strongly typed	Expect around 100 million products, >1 billion triples	Early stages of development and deployment
IBM	Entities and relations with evidence information associated with them.	Various sizes. Proven on scales documents >100 million, relationships >5 billion, entities >100 million	Actively used in products and by clients

Minimal Knowledge Graph Definition

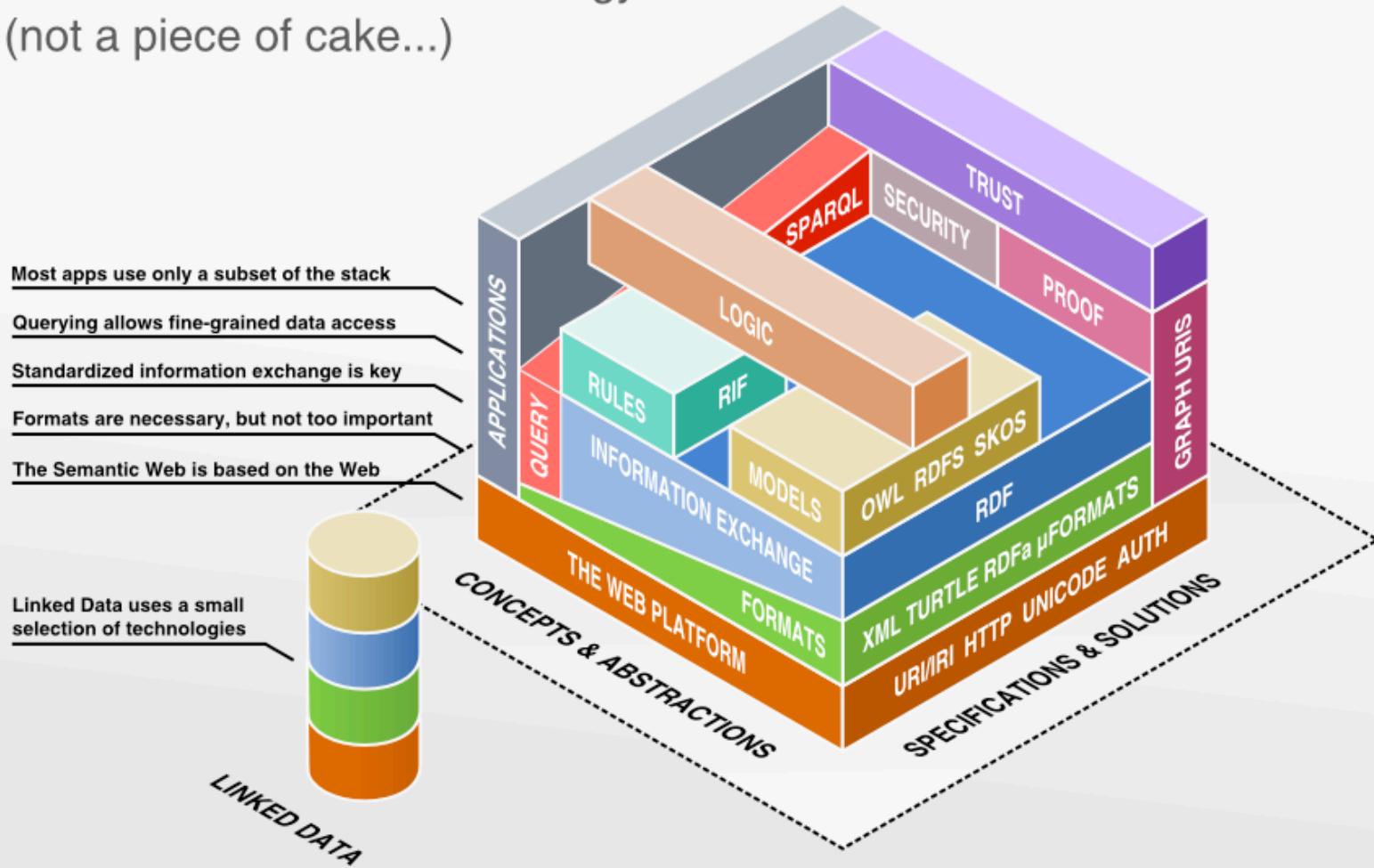
- A quintuple: $\langle E, L, T, P, A \rangle$
 - E : a set of entity id symbols (= constants)
 - L : a set of literals, i.e., strings, numbers, etc. (= constants)
 - T : a set of type symbols (= unary predicate symbols)
 - P : a set of relation symbols (= binary predicate symbols)
 - A : a set of axioms in a language
- A : a set of **axioms** in a language (\approx subset of FOL)
 - A-facts:
 - Types of entities and literals: $t(e) \mid t(l)$ - with $e \in E$ and $l \in L$
 - Relations between entities: $r(e_1, e_2) \mid r(e, l)$ – with $e_i \in E$ and $l \in L$
 - A-general-axioms:
 - Depends on the language....
 - Minimal requirements: axioms with the form $\forall x (t_1(x) \rightarrow t_2(x))$ that specify subtype dependencies that define a partial order over T

RDF: Resource Description Framework



Semantic Web Stack Revised

The Semantic Web Technology Stack
(not a piece of cake...)



RDF: Basic Principles

- RDF is a data model for representing data on the Web
 - based on:
 - Triples – basic unit to organize the information
 - Directed (labeled) graphs – sets of triples
 - URLs – unique resource identifiers

RDF: Basic Principles

- RDF is a data model for representing data on the Web
 - based on:
 - **Triples – basic unit to organize the information**
 - Directed (labeled) graphs – sets of triples
 - URLs – unique resource identifiers

RDF Triples // Stetements

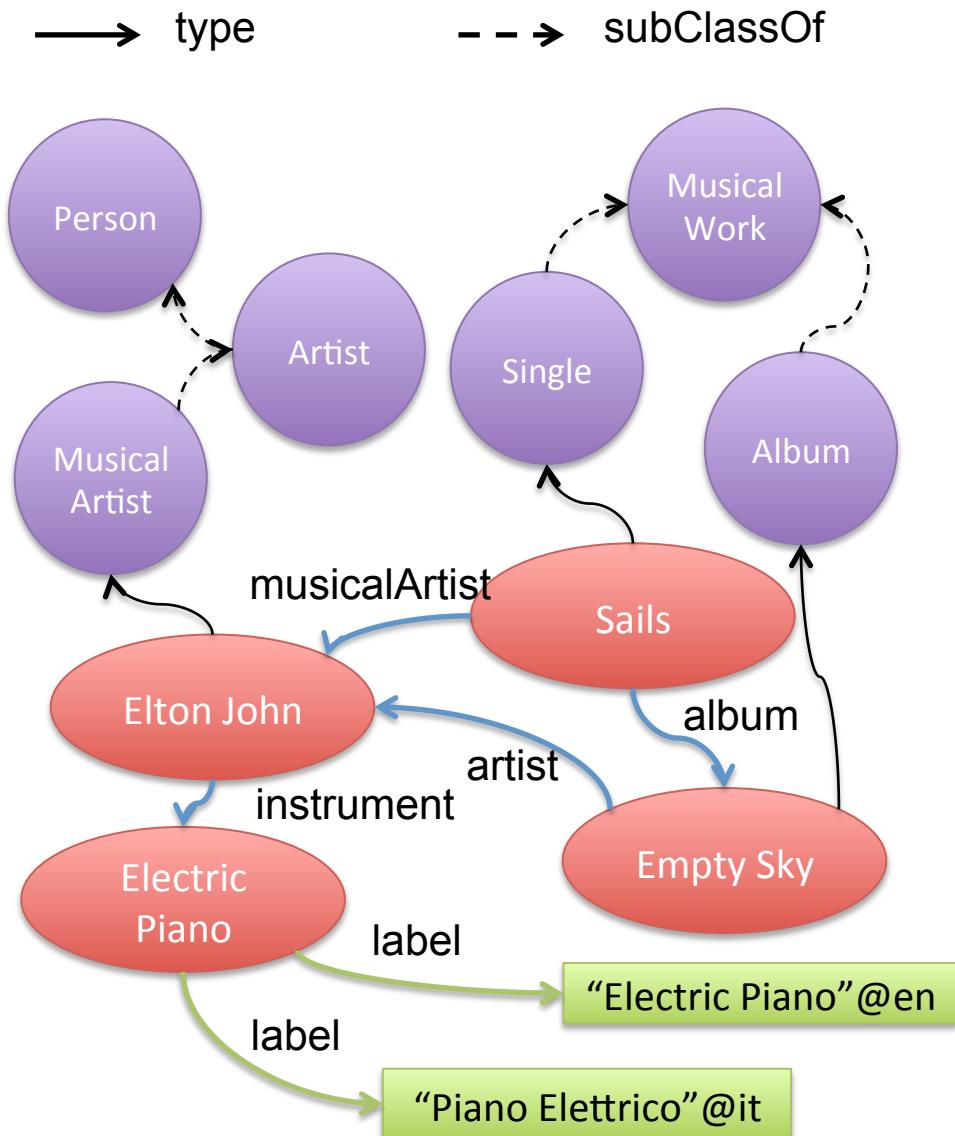
```
<Electric Piano, label, "Electric Piano"@en>
<Electric Piano, label, "Piano Elettrico"@it>
<Elton John, instrument, Electric Piano>
<Empty Sky, artist, Elton John>
<Sails, musicalArtist, Elton John>
<Sails, album, Empty Sky>
<Elton John, instrument, Electric Piano>
<Elton John, type, Musical Artist>
<Sails, type, Single>
<Empty Sky, type, Album>
<Musical Artist, subClassOf, Artist>
<Artist, subClassOf, Person>
<Album, subClassOf, Musical Work>
<Single, subClassOf, Musical Work>
```

RDF: Basic Principles

- RDF is a data model for representing data on the Web
 - based on:
 - Triples – basic unit to organize the information
 - **Directed (labeled) graphs – sets of triples**
 - URLs – unique resource identifiers

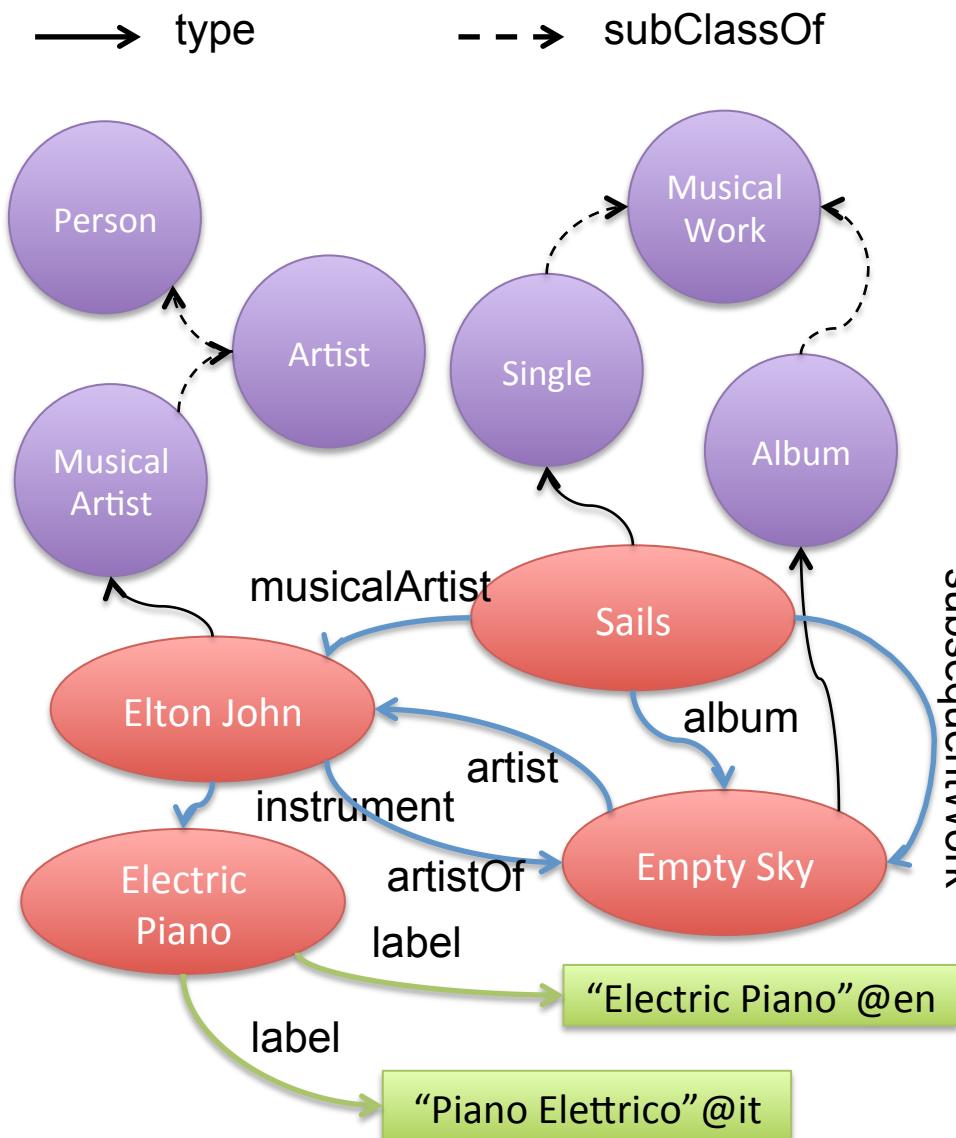
RDF: Graphs (vs. Triples)

<Electric Piano, label, "Electric Piano"@en>
 <Electric Piano, label, "Piano Elettrico"@it>
 <Elton John, instrument, Electric Piano>
 <Empty Sky, artist, Elton John>
 <Sails, musicalArtist, Elton John>
 <Sails, album, Empty Sky>
 <Elton John, instrument, Electric Piano>
 <Elton John, type, Musical Artist>
 <Sails, type, Single>
 <Empty Sky, type, Album>
 <Musical Artist, subClassOf, Artist>
 <Artist, subClassOf, Person>
 <Album, subClassOf, Musical Work>
 <Single, subClassOf, Musical Work>



RDF: Graphs (vs. Triples)

<Electric Piano, label, "Electric Piano"@en>
 <Electric Piano, label, "Piano Elettrico"@it>
 <Elton John, instrument, Electric Piano>
 <Empty Sky, artist, Elton John>
 <Sails, musicalArtist, Elton John>
 <Sails, album, Empty Sky>
 <Elton John, instrument, Electric Piano>
 <Elton John, type, Musical Artist>
 <Sails, type, Single>
 <Empty Sky, type, Album>
 <Musical Artist, subClassOf, Artist>
 <Artist, subClassOf, Person>
 <Album, subClassOf, Musical Work>
 <Single, subClassOf, Musical Work>
 <Elton John, artistOf, Empty Sky>
 <Sails, subsequentWork, Empty Sky>

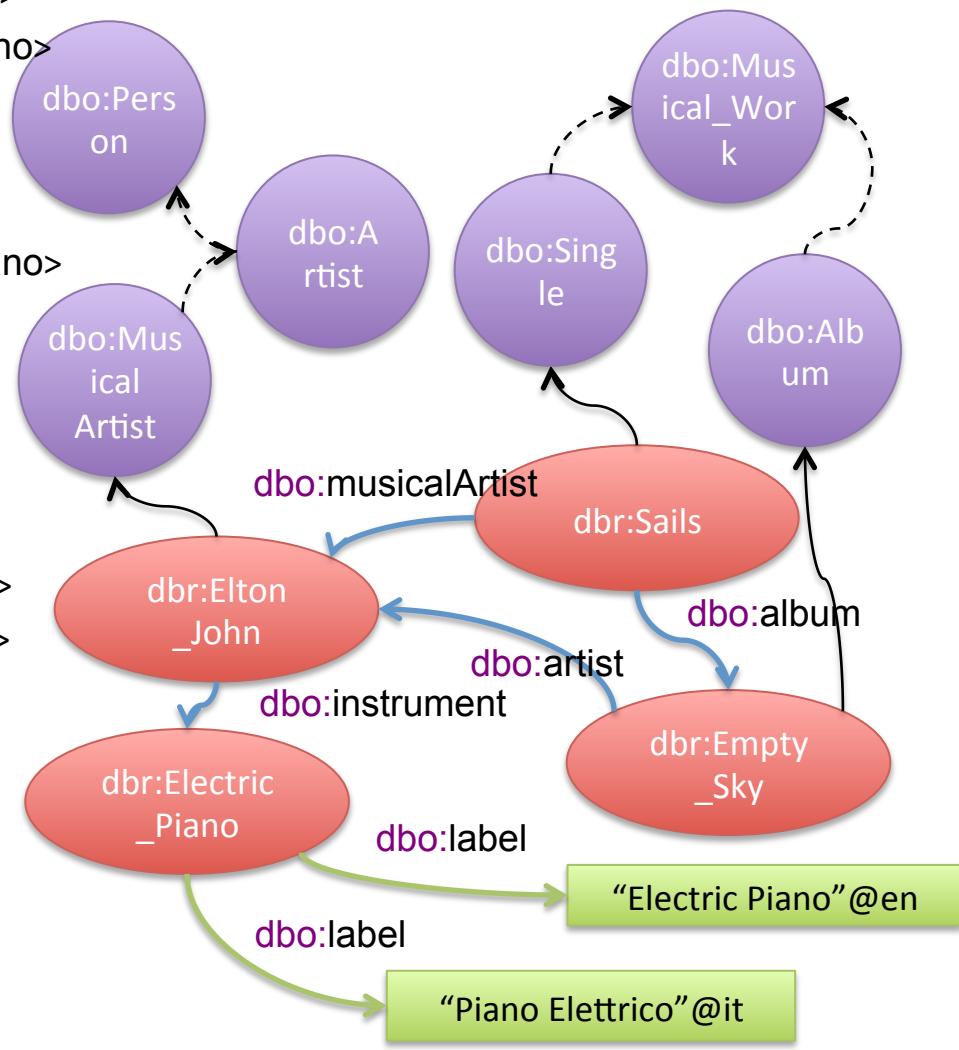


RDF: Basic Principles

- RDF is a data model for representing data on the Web
 - based on:
 - Triples – basic unit to organize the information
 - Directed (labeled) graphs – sets of triples
 - **URIs – unique resource identifiers**

RDF: URLified Graphs and Triples

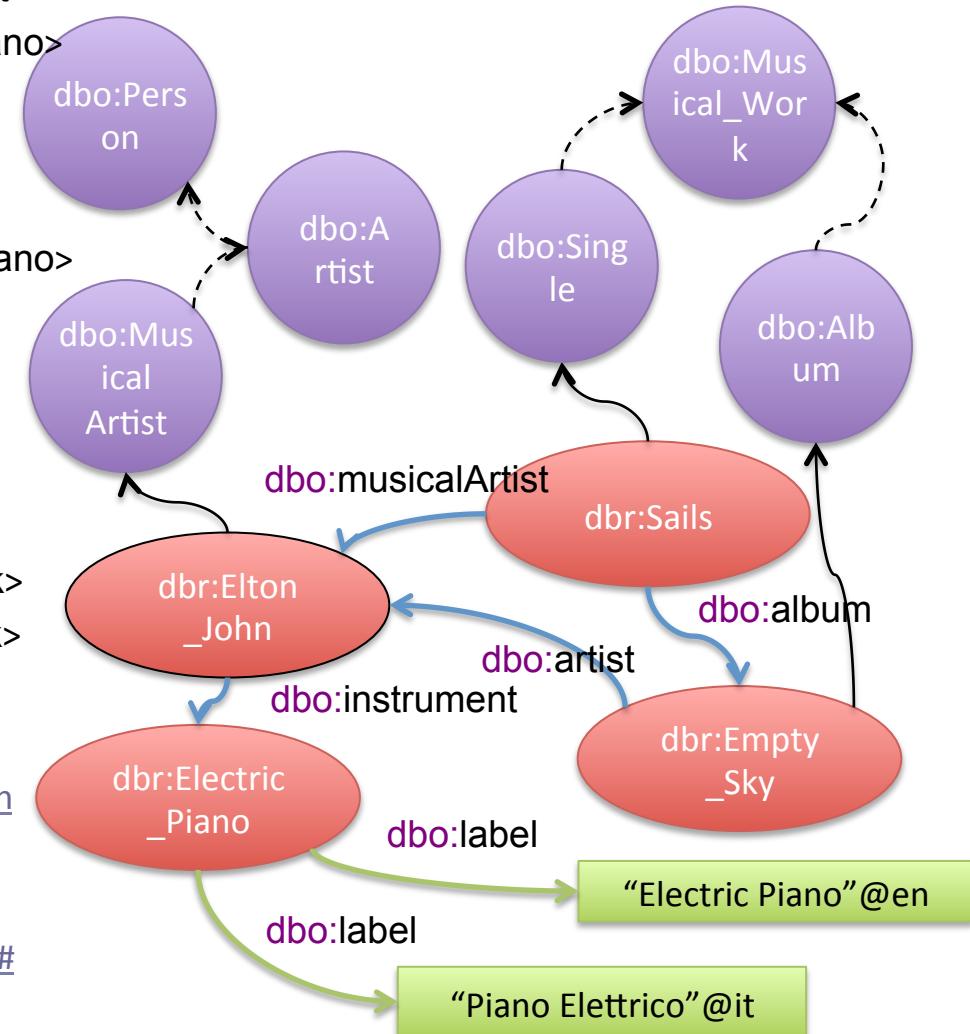
<dbr:Electric_Piano, rdfs:label, "Electric Piano"@en>
 <dbr:Electric_Piano, rdfs:label, "Piano Elettrico"@it>
 <dbr:Elton_John, dbo:instrument, dbr:Electric_Piano>
 <dbr:Empty_Sky, dbo:artist, dbr:Elton_John>
 <dbr:Sails, dbo:musicalArtist, dbr:Elton_John>
 <dbr:Sails, dbo:album, dbr:Empty_Sky>
 <dbr:Elton_John, dbo:instrument, dbr:Electric_Piano>
 <dbr:Elton_John, rdf:type, dbr:Musical_Artist>
 <dbr:Sails, rdf:type, dbo:Single>
 <dbr:Empty_Sky, rdf:type, dbo:Album>
 <dbo:Musical_Artist, rdfs:subClassOf, dbo:Artist>
 <dbo:Artist, rdfs:subClassOf, dbo:Person>
 <dbo:Album, rdfs:subClassOf, dbo:Musical_Work>
 <dbo:Single, rdfs:subClassOf, dbo:Musical_Work>



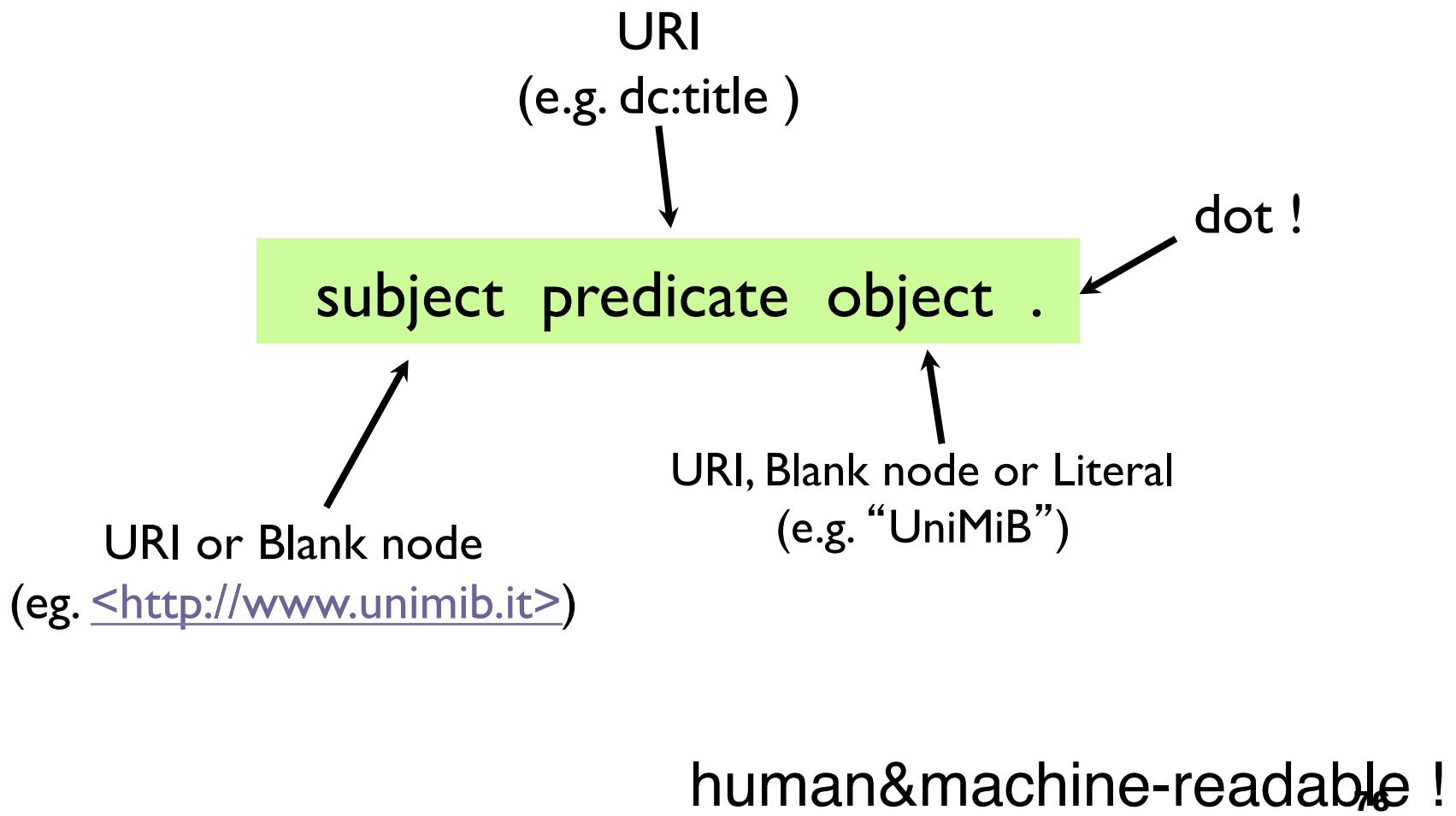
RDF: URLified Graphs and Triples

```
<dbr:Electric_Piano, rdfs:label, "Electric Piano"@en>
<dbr:Electric_Piano, rdfs:label, "Piano Elettrico"@it>
<dbr:Elton_John, dbo:instrument, dbr:Electric_Piano>
<dbr:Empty_Sky, dbo:artist, dbr:Elton_John>
<dbr:Sails, dbo:musicalArtist, dbr:Elton_John>
<dbr:Sails, dbo:album, dbr:Empty_Sky>
<dbr:Elton_John, dbo:instrument, dbr:Electric_Piano>
<dbr:Elton_John, rdf:type, dbr:Musical_Artist>
<dbr:Sails, rdf:type, dbo:Single>
<dbr:Empty_Sky, rdf:type, dbo:Album>
<dbo:Musical_Artist, rdfs:subClassOf, dbo:Artist>
<dbo:Artist, rdfs:subClassOf, dbo:Person>
<dbo:Album, rdfs:subClassOf, dbo:Musical_Work>
<dbo:Single, rdfs:subClassOf, dbo:Musical_Work>

dbr: for http://dbpedia.org/resource/
e.g.: http://dbpedia.org/resource/Elton\_John
dbo: for http://dbpedia.org/ontology/
rdfs for https://www.w3.org/2000/01/rdf-schema#
rdf for http://www.w3.org/1999/02/22-rdf-syntax-ns#
```



RDF Turtle/N3 Triples



RDF Turtle/N3 Terms

URIs as identifiers for individuals, properties, classes, and datatypes

`<http://www.unimib.it>`

`dc:title`

`archeoserver:material`



Blank node

`_:name`
[]

Literal

Strings, numbers, dates, ...

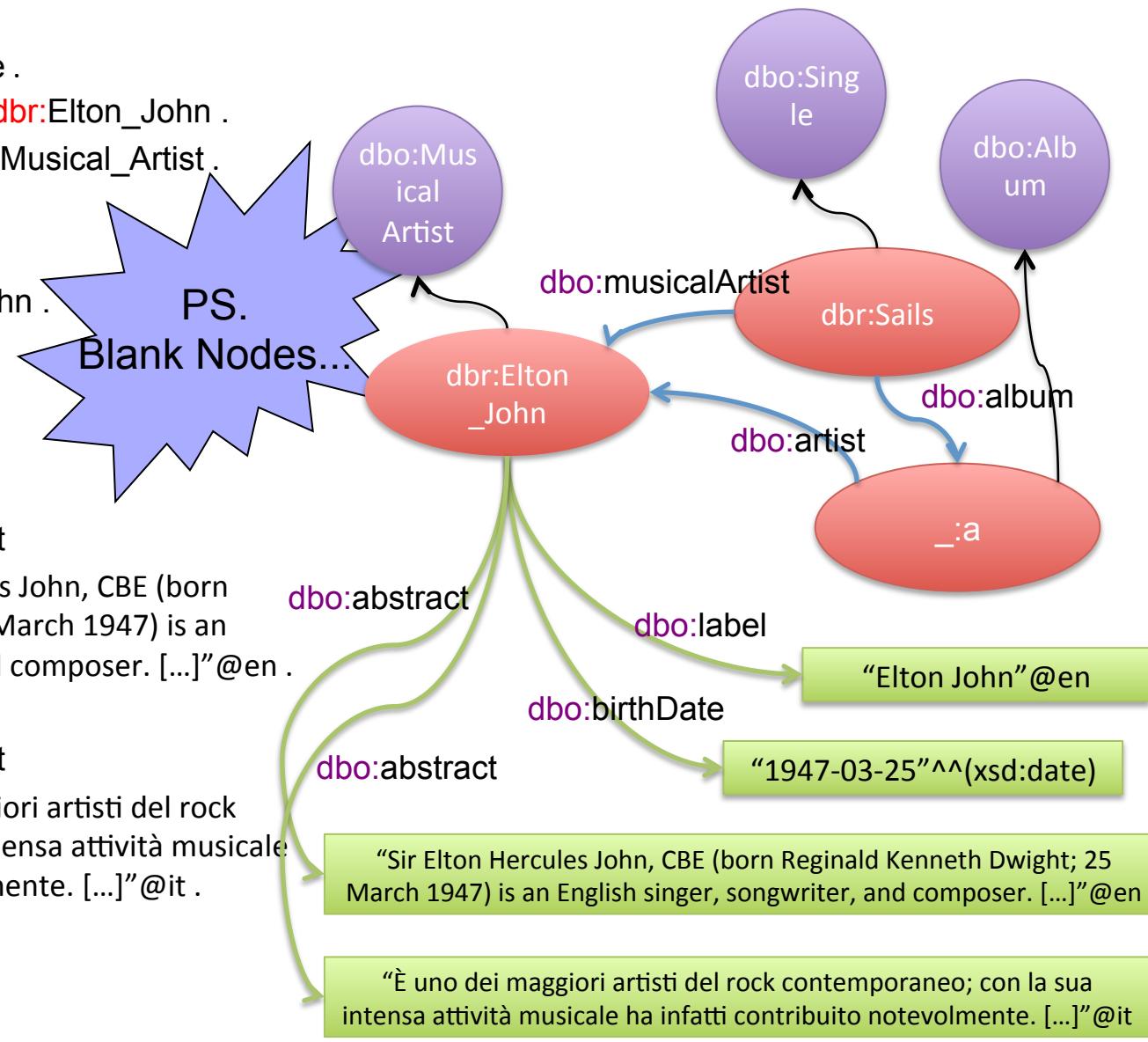
- Plain literals
“2006”
“Bronze Age”
- Typed literals
“2006”^{^^}xsd:integer
“Bronze Age”^{^^}xsd:string

Datatypes

xsd:boolean (true, false)
xsd:integer (10, -1293)
xsd:decimal (12.13)
xsd:double (2.78e-2)
xsd:string (“Bronze Age”)

RDF: URLified Graphs + BN + Literals

dbr:Sails rdf:type dbo:Single .
 dbr:Sails dbo:musicalArtist dbr:Elton_John .
 dbr:Elton_John rdf:type dbr:Musical_Artist .
 dbr:Sails dbo:album _:a .
 _:a rdf:type dbo:Album .
 _:a, dbo:artist, dbr:Elton_John .



Exercise... match the minimal KG definition to RDF language ...

- A quintuple: $\langle E, L, T, P, A \rangle$
 - E : a set of entity id symbols (= constants)
 - L : a set of literals, i.e., strings, numbers, etc. (= constants)
 - T : a set of type symbols (= unary predicate symbols)
 - P : a set of relation symbols (= binary predicate symbols)
 - A : a set of axioms in a language
- A : a set of **axioms** in a language (\approx subset of FOL)
 - A-facts:
 - Types of entities and literals: $t(e) \mid t(l)$ - with $e \in E$ and $l \in L$
 - Relations between entities: $r(e_1, e_2) \mid r(e, l)$ – with $e_i \in E$ and $l \in L$
 - A-general-axioms:
 - Depends on the language....
 - Minimal requirements: axioms with the form $\forall x (t_1(x) \rightarrow t_2(x))$ that specify subtype dependencies that define a partial order over T