

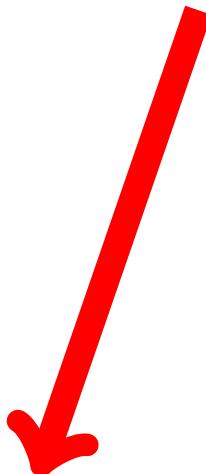
Information Extraction: Motivation

Fabian M. Suchanek

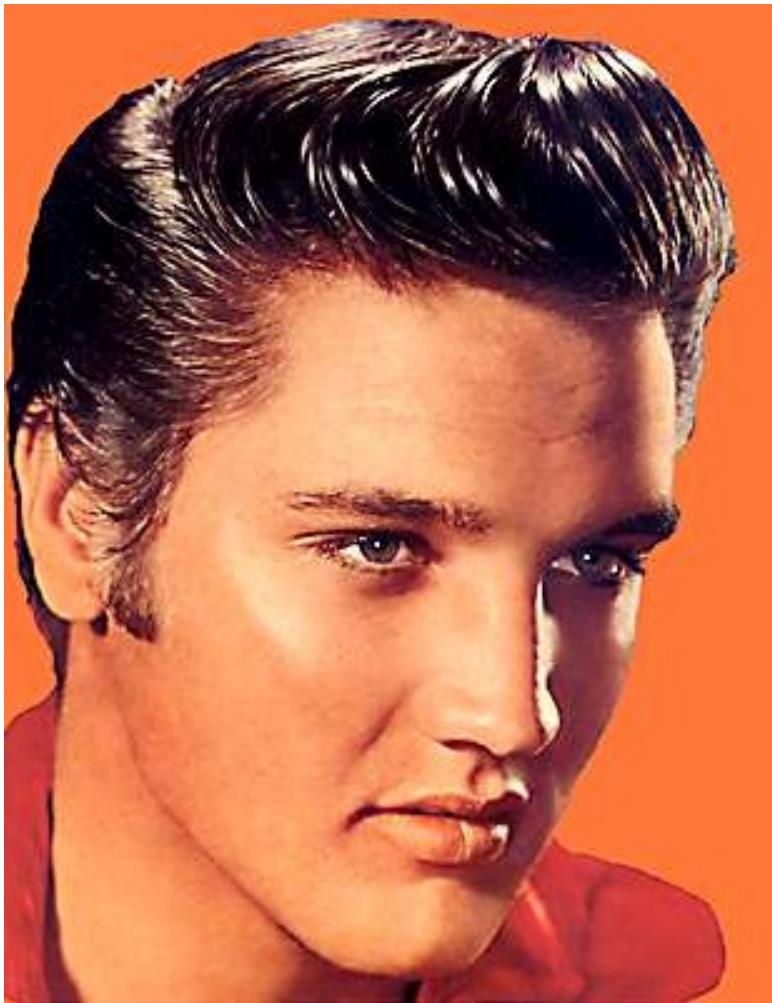
Lecturer: Fabian Suchanek



Professor at
Télécom ParisTech
DBWeb team, office C201-6
<https://suchanek.name>



Elvis Presley



Elvis,
when I need you,
I can hear you!

Is Elvis still alive?

Elvis Alive: Google



is elvis alive?

Web

Images

Maps

Shopping

Videos

More ▾

Search tools

About 21,700,000 results (0.22 seconds)

[WebSpawner - IS ELVIS DEAD OR ALIVE ???](#)

www.webspawner.com/users/elviselvis/

As it turns out, there are many concrete reasons to believe that the King is still **alive**.
The Gravesite. For Starters, **Elvis's** name is spelled wrong on his headstone.

Google

Google: undecided

Elvis Alive: Bing

 is elvis alive 

Web Images Videos Maps News Explore

45 500 000 RESULTS

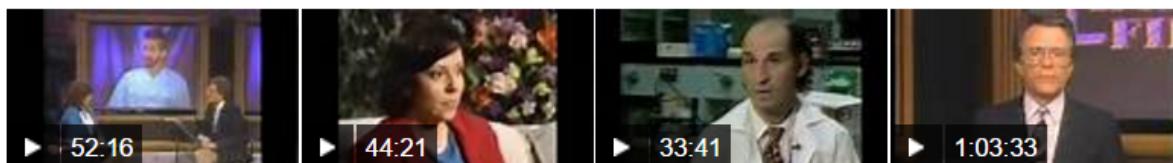
Date ▾

Language ▾

Region ▾

Videos of **is elvis alive**

bing.com/videos



Bing: yes

[See more videos of **is elvis alive**](#)

Elvis Presley still alive? (October 2016 update) - **Mediamass**

en.mediamass.net › People › **Elvis Presley** ▾

News of the King **Elvis Presley**'s death on August 16, 1977 stunned the world. Ever since, millions of fans worldwide have come to believe the singer's death was a ...

ELVIS PRESLEY - I'M ALIVE - YouTube

[Youtube](#) ▾

BY KNIGHTS OF THE KINGS REALM 6 MIN 1,9M VIEWS

Bing

Elvis Alive: Ask.com

Réponses

Images



is elvis alive?

[Proof Elvis is Alive by J. Parra - ...](#)

www.youtube.com/watch?v=TRBm6q-3SP0

7 Jan 2012 ... In this video, it will show that Elvis is still alive and a doctor who treated him claims that he is still alive.

Ask.com

Ask.com: yes

Elvis Alive: Evi



Is elvis alive?



You asked: Is elvis alive?

Elvis Presley died in Memphis on August 16th 1977. He was killed by a myocardial infarction. He was 42 years, 7 months and 8 days old when he died.



Elvis Presley



[website](#)



[wikipedia](#)

Evi

Evi: no

Elvis Alive: Wolfram Alpha



is elvis alive?



Ex

Input interpretation:

Elvis Presley alive?

Result:

No

Wolfram Alpha

Wolfram Alpha: no

How do they know?



Sources



- The New York Times Company. "Obituaries." *The New York Times*. »

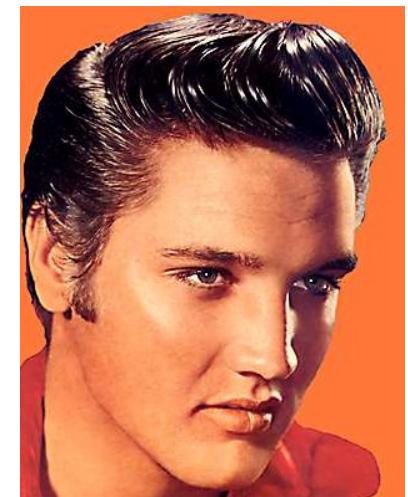


Elvis Presley, the first
and greatest American
rock-and-roll star,
died yesterday at
the age of 42.

Text understanding

To answer the question, Wolfram Alpha has to understand this text.

Elvis Presley, the first
and greatest American
rock-and-roll star,
died yesterday at
the age of 42.



Information Extraction

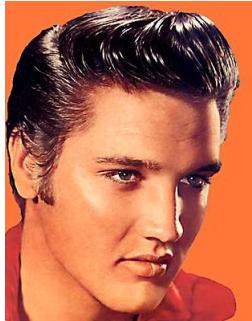
Information Extraction (IE) is the process of deriving structured factual information from digital text documents.

Elvis is a singer.

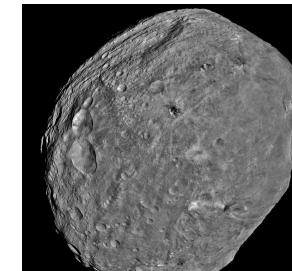
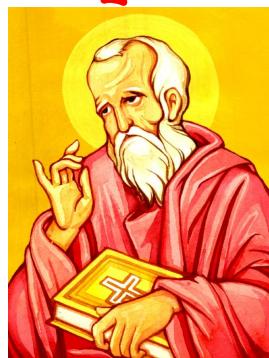
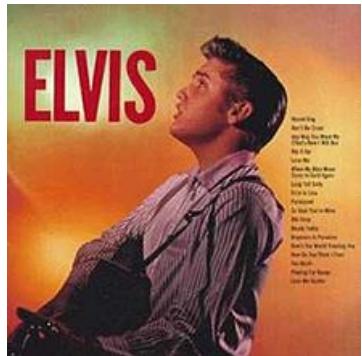


Person	Occupation
Elvis	singer

Difficulty:Ambiguity



Elvis is a singer.



Difficulty: Detect phrases

Elvis is a rock star.



Elvis is a rock star.



Elvis is a real rock star.



Difficulty:Verbal phrases

Elvis works as a rock star.

Elvis is famous as a rock star.

Elvis claims to be a rock star.

Elvis is not a rock star.

Now is he a rock star or what??

Difficulty: Parsing

Bell, a telecommunication company, which
is based in Los Angeles, makes and distributes
electronic, computer, and building products.

Difficulty: Parsing

Bell, a telecommunication company, which
is based in Los Angeles, makes and distributes
electronic, computer, and building products.

<u>Subject</u>	<u>Predicate</u>	<u>Object</u>
Bell	is	telecomm. company
Bell	based in	Los Angeles
Bell	makes	electronic products
Bell	distributes	electronic products
Bell	makes	computer products
...		

Difficulty:Table Interpretation

No.	Country	Capital
1	Afghanistan	Kabul
2	Albania	Tirane
3	Algeria	Algiers
4	Andorra	Andorra la Vella

Difficulty:Table Interpretation

Afghanistan hasCapital Kabul

No.	Country	Capital
1	Afghanistan	Kabul
2	Albania	Tirane
3	Algeria	Algiers
4	Andorra	Andorra la Vella

Difficulty:Table Interpretation

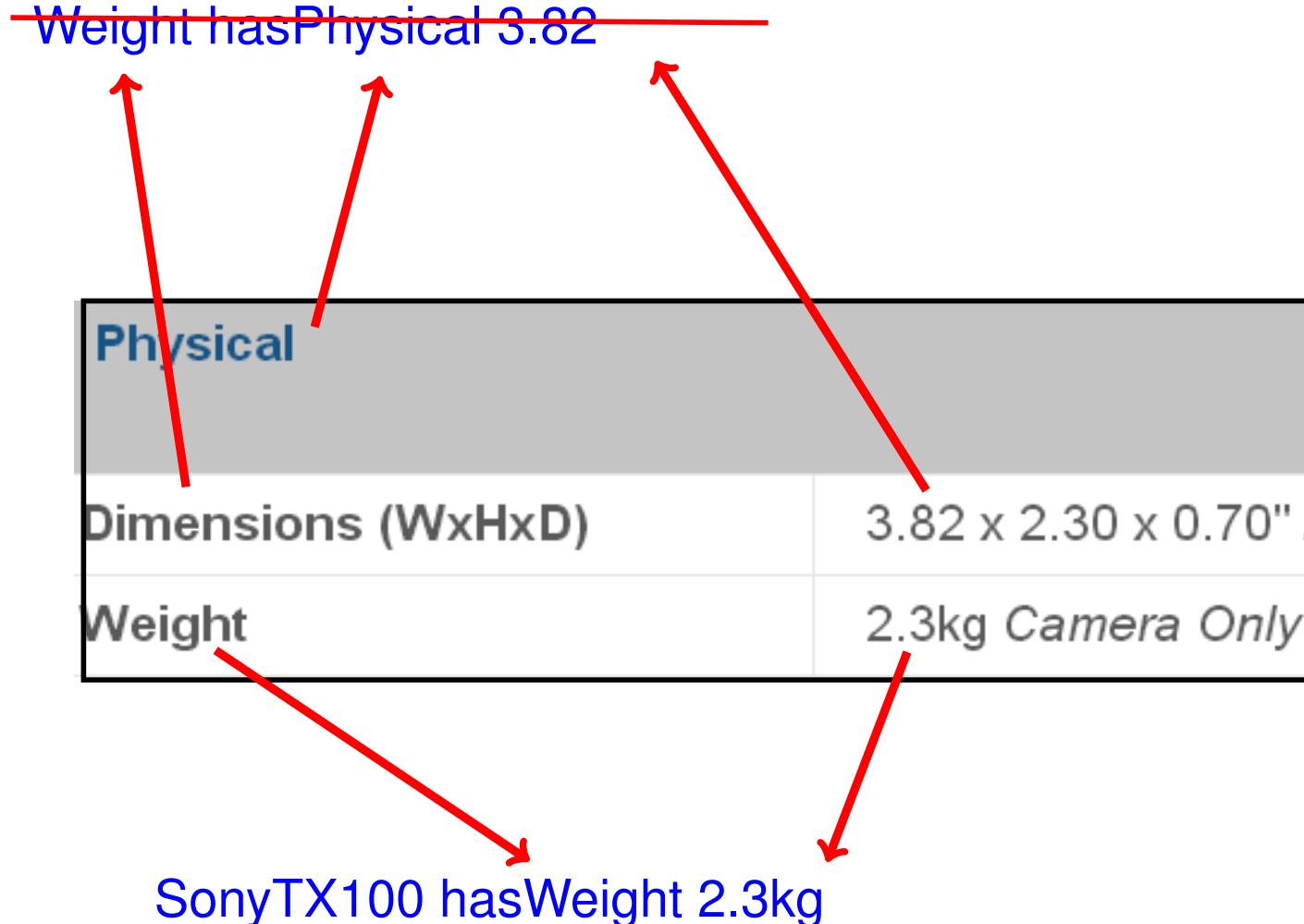
Physical	
Dimensions (WxHxD)	3.82 x 2.30 x 0.70"
Weight	2.3kg Camera Only

Difficulty:Table Interpretation

Weight hasPhysical 3.82

Physical	
Dimensions (WxHxD)	3.82 x 2.30 x 0.70"
Weight	2.3kg Camera Only

Difficulty:Table Interpretation



Difficulty: Parsing Literals

Background information

Birth name Elvis Aaron Presley

Born January 8, 1935

Tupelo, Mississippi, U.S.

Difficulty: Parsing Literals

Literals take many different forms:

8th of January 1935

1935-01-08

8/1/1935

8/Jan/1935

08-01-1935

b. 1935

1/8/1935

Jan 8, 1935

Difficulty: Segmentation

ADRESSES ET ACCÈS - CONTACTS

Adresse postale : 46 rue Barrault F-75634 Paris Cedex 13

Téléphone : +33 (0)1 45 81 77 77 - **Fax :** +33 (0)1 45 89 79

Difficulty: Segmentation

ADRESSES ET ACCÈS - CONTACTS

Adresse postale : 46 rue Barrault F-75634 Paris Cedex 13

Téléphone : +33 (0)1 45 81 77 77 - Fax : +33 (0)1 45 89 79

Difficulty: Segmentation

ADRESSES ET ACCÈS - CONTACTS

Adresse postale : 46 rue Barrault F-75634 Paris Cedex 13

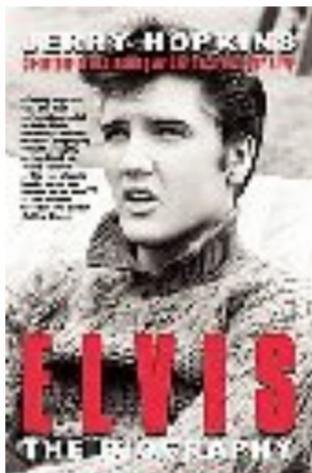
Téléphone : +33 (0)1 45 81 77 77 - Fax : +33 (0)1 45 89 79



No	Street	Country	ZIP	City
46	Barrault	France	75634	Paris

Difficulty:Bad HTML Code

```
<table>
```



```
<tr>
```

[View Larger Image](#)

Elvis : The Biography

Hopkins, Jerry

9780859653916

ISBN 10: **0859653919 / 0-85965-391-9**

ISBN 13: **9780859653916**

Publication Date: **2007**

```
</table>
```

- Tables are used for layout
- Lists do not have list tags
- HTML code is invalid

Why would anybody want to do Information Extraction?

The reason is that...

this is abundant...



...and this is useful.



Person	Occupation
Elvis	singer

Application: Emails

Pierre Senellart

An: Fabian M. Suchanek

Aw: Teaching activities for next year

Sicherheit: Signiert (pierre@senellart.com)

 Kontaktinformationen dieser E-Mail aktualisieren: Pierre Senellart Arb

OK, then we would like to take this offer of yours and book the 9x3 hours from SDP08 to d IMF102.

Added to the schedule of both classes, available on Ecole. Please check.

Best,

--

Pierre Senellart

pierre@senellart.com

<http://pierre.senellart.com/>

Tel. (work) : +33 1 45 81 77 59 / +65 6516 2864

Tel. (mobile) : +33 6 48 36 36 20 / +65 8385 9532

Fax : +33 9 57 71 19 06

Application: Emails

Pierre Senellart
An: Fabian M. Suchanek
Aw: Teaching activities for next year
Sicherheit: ☑ Signiert (pierre@senellart.com)

 Kontaktinformationen dieser E-Mail aktualisieren: Pierre Senellart Arbeit

OK, then we would like to take this offer of yours and book the 9:30 hours from SD2G0 to INF102.

Added to the schedule of both classes, available on Ecole. Please check.

Best,

Pierre Senellart
pierre@senellart.com
<http://pierre.senellart.com/>
Tel. (work) : +33 1 45 81 77 59 / +65 6516 2864
Tel. (mobile) : +33 6 48 36 36 20 / +65 8385 9532
Fax : +33 9 57 71 19 06

PS Pierre Senellart

Privat +33 9 52 71 19 06
Mobil +33 6 48 36 36 20
Privat +33 9 57 71 19 06
Arbeit +33 1 45 81 77 59 [\(i\)](#)
Telefon +65 6516 2864 [\(i\)](#)
Telefon +65 8385 9532 [\(i\)](#)

Anruf FaceTime  Audio

E-Mail pierre@senellart.com
E-Mail pierre.senellart@enst.fr
E-Mail pierre.senellart@gmail.com
E-Mail pierre.senellart@telecom-paris.fr

URL <http://pierre.senellart.com/>

[Ignorieren](#) [Kontakt aktualisieren](#)

Application: Emails

Hi I'm russian girl with 25 years old
looking for a faithful partner. Chat with
me on Monday 2pm at [http://wtf.ru!!!](http://wtf.ru)

Add to calendar?

Added Chat with Russian girl on Mon 2016-03-11 at 14:00. [Undo](#)

Application: Emails



Katerina Tzompanaki para mí, Marie, Antoine, Thomas, Fabian

We found these foulards online with JB and Thomas.

- 1) <http://www.galerieslafayette.com/p/foulard+en+coton+imprime>
- 2) <http://www.galerieslafayette.com/p/foulard+imprime+floral-desi...>
- 3) <http://www.galerieslafayette.com/p/foulard+far+wes-desigual/3>

Do you prefer any of these?

If we decide, I can go and buy it in the weekend.

Cheers



Responder a todos

I like the first two.

Either one is fine!

Thoughts?

Application: Emails

Screenshot of the Google Inbox interface showing a list of trips.

The sidebar on the left includes:

- Inbox
- Snoozed
- Done
- Drafts
- Sent
- Reminders
- Trash
- Spam
- Contacts
- Trips** (selected)
- Bundled in the inbox
- Purchases
- Finance
- Skip the inbox
- ASLE
- Joomla!
- Publicidad
- Unbundled

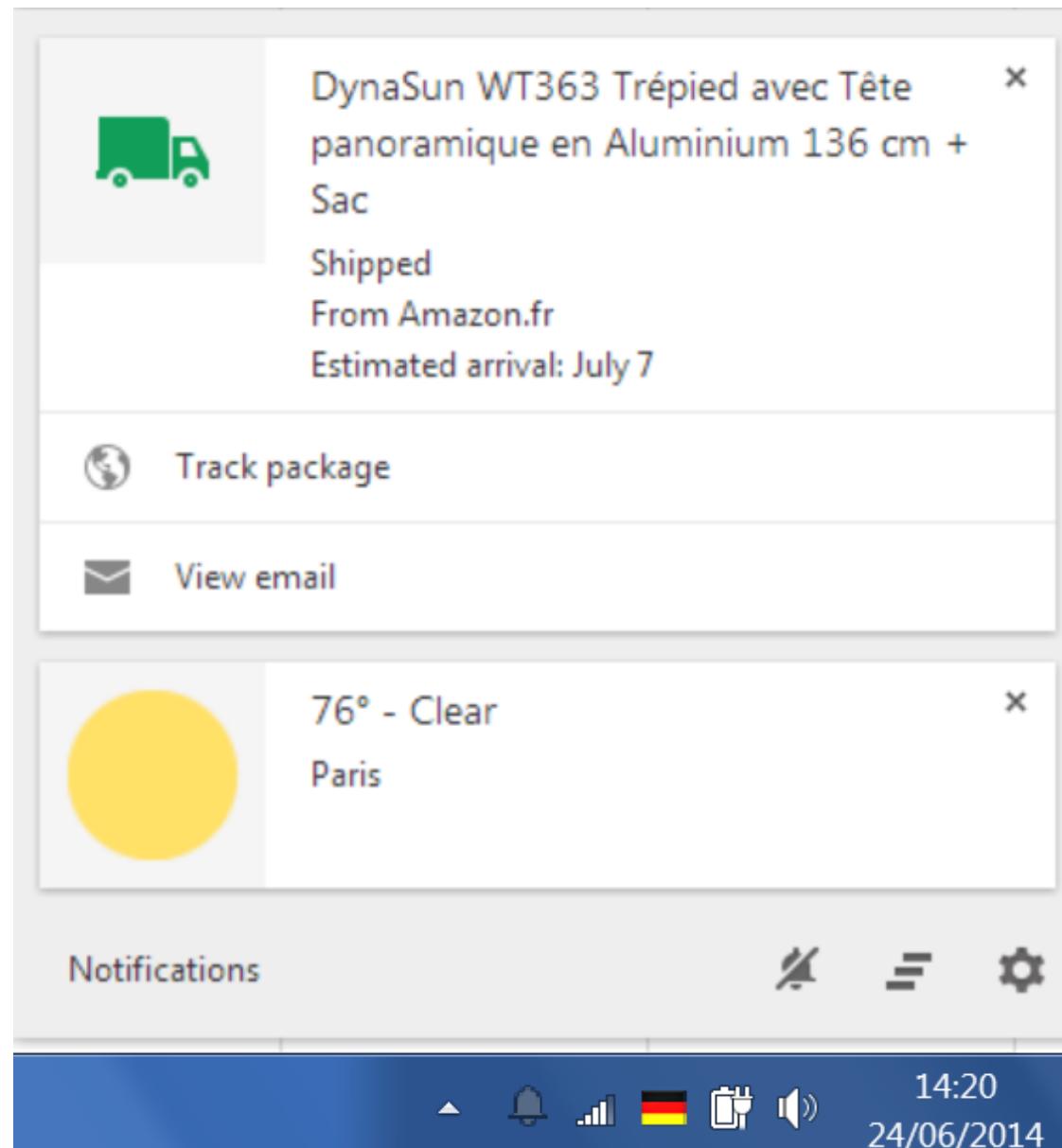
The main area shows a list of trips under the heading "Past". Each trip is represented by a card with a thumbnail image, a title, and a date range.

- Trip to Gloucester (May 26 – Jun 16)
- Trip from Gloucester to San Cristobal Island (May 29 – Jun 2)
- Trip to Florence (May 17 – 24)
- Trip to Montreal & New York City (Dec 4 – 15, 2014)
- Weekend in Lyon (Dec 6 – 8, 2014)
- Trip to Boston (Oct 29 – 31, 2014)
- Weekend in City of Brussels (Aug 16 – 17, 2014)

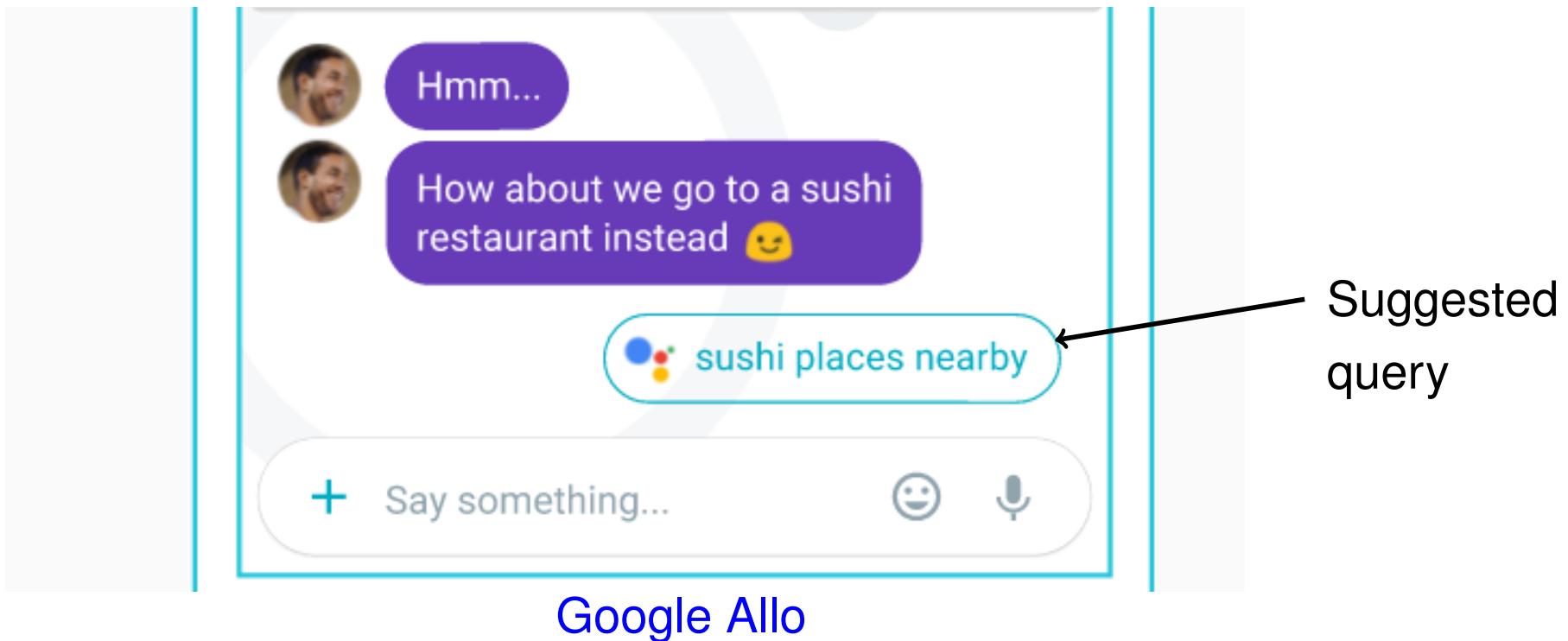
>assistant

More examples
Even more examples

Application: Intelligent Assistants



Application: Chat bots



Siri briefly thought Bulgaria's national anthem was 'Despacito'

>analysis

Application: Customer care

Dear Apple support,

my iPhone 7 charging cable is not interoperable
with my fridge. Can you please help me?

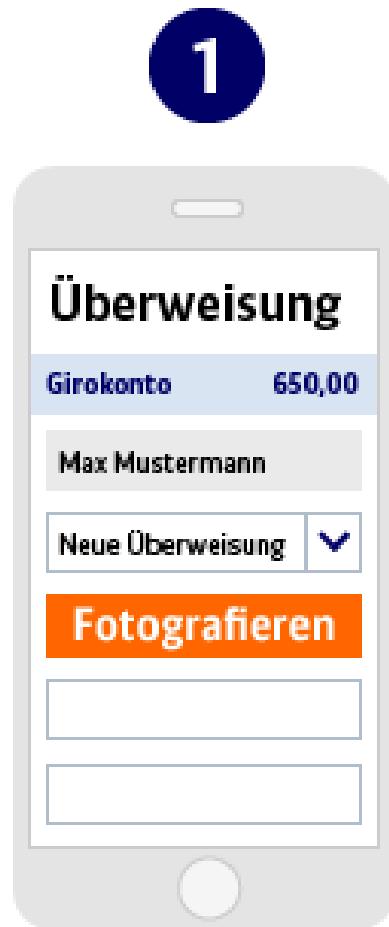


Product	Part	Problem
iPhone 7	Charger	Compat.



>Analysis

Application: Bank transfers



Überweisung
aufrufen

Rechnung
fotografieren

Prüfen und
freigeben

Application: Predictions

How two scientists are using the *New York Times* archives to predict the future

by Laura Hazard Owen FEB. 1, 2013 - 7:24 AM PST

7 Comments    +1 

A▼ A▲

SUMMARY: Researchers are creating software that analyzes 22 years of New York Times archives, Wikipedia and about 90 other web resources to predict future disease outbreaks, riots and deaths — and hopefully prevent them.

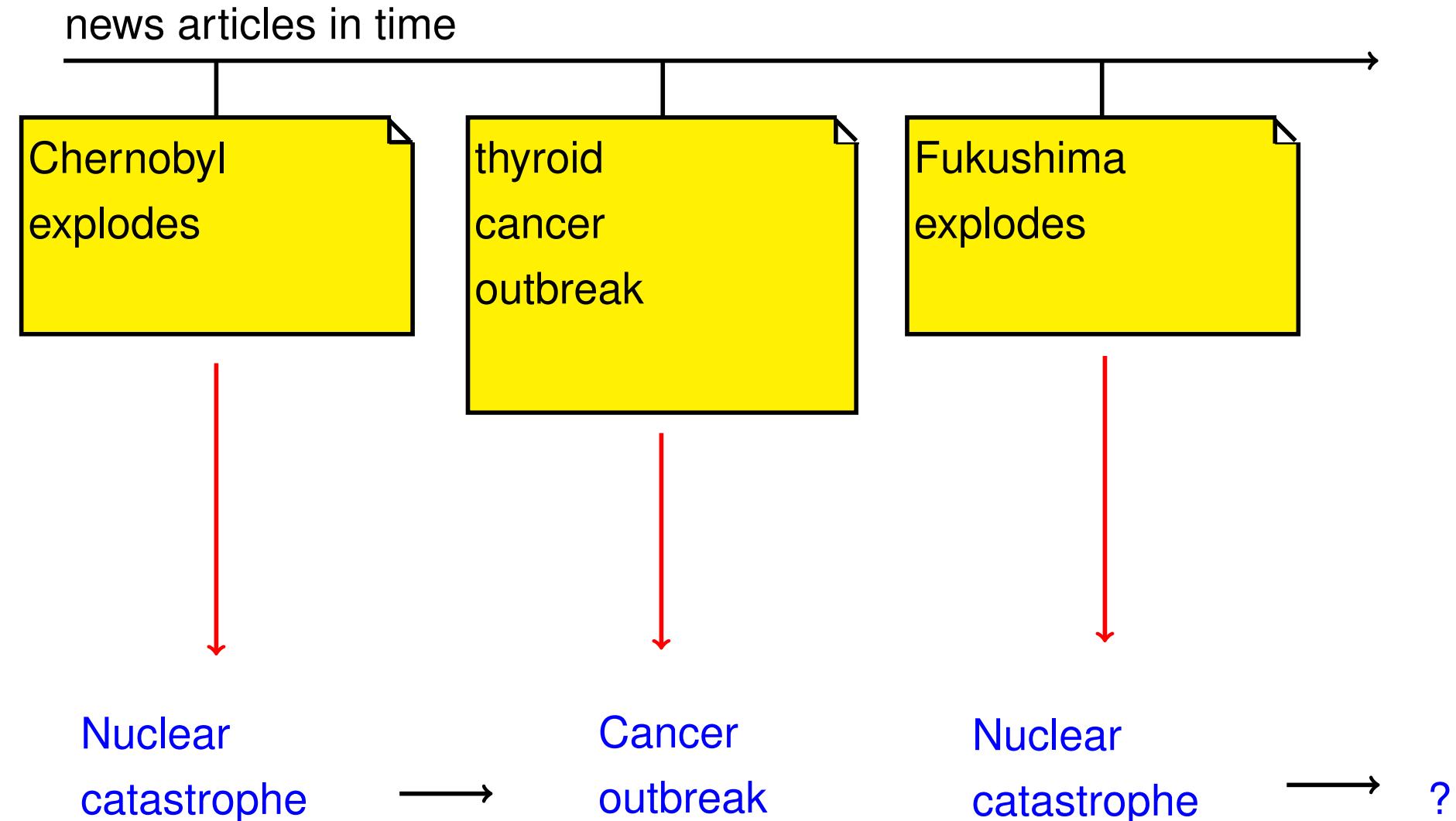
 tweet this



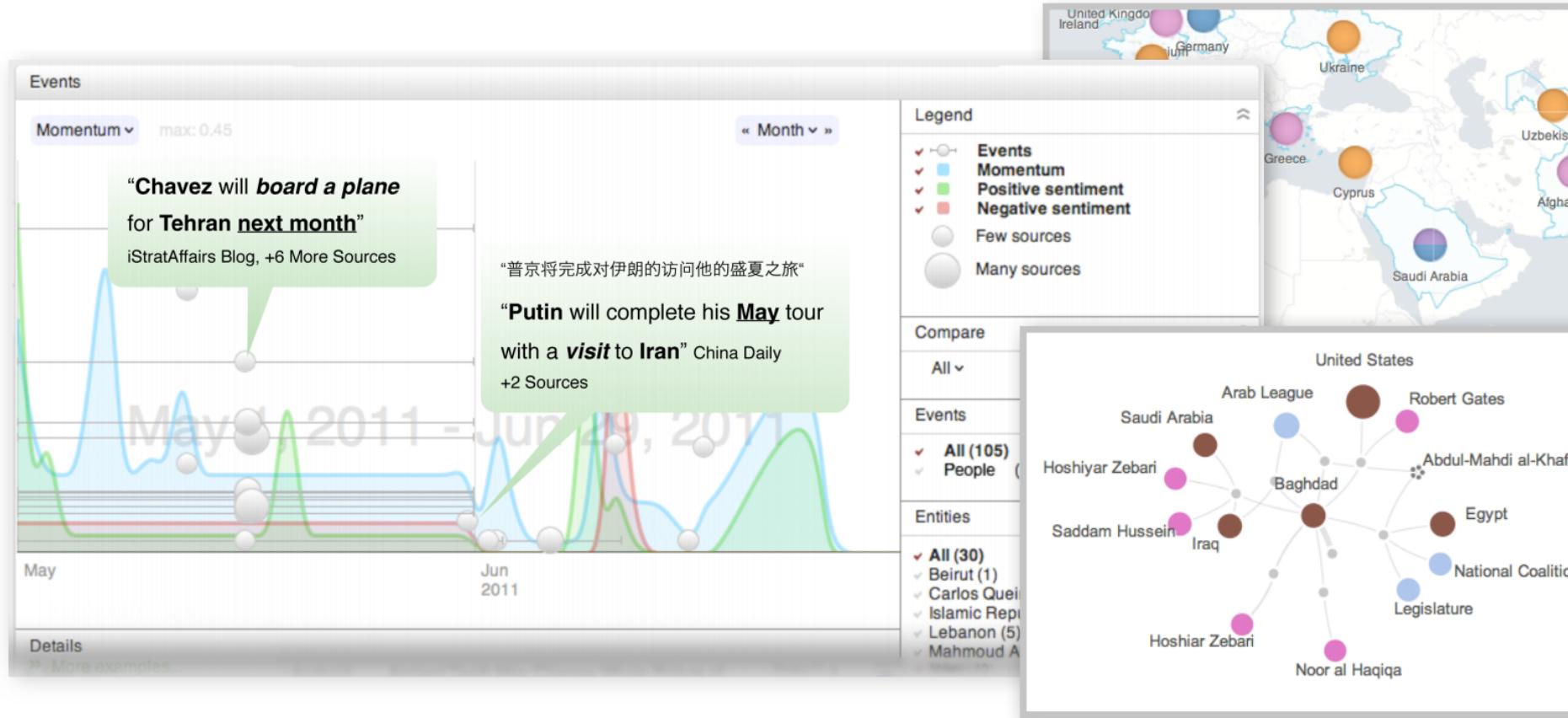
photo: Shutterstock / Ruggiero Scardigno

Gigaom
>Analysis

Application: Predictions



Application: Analysis

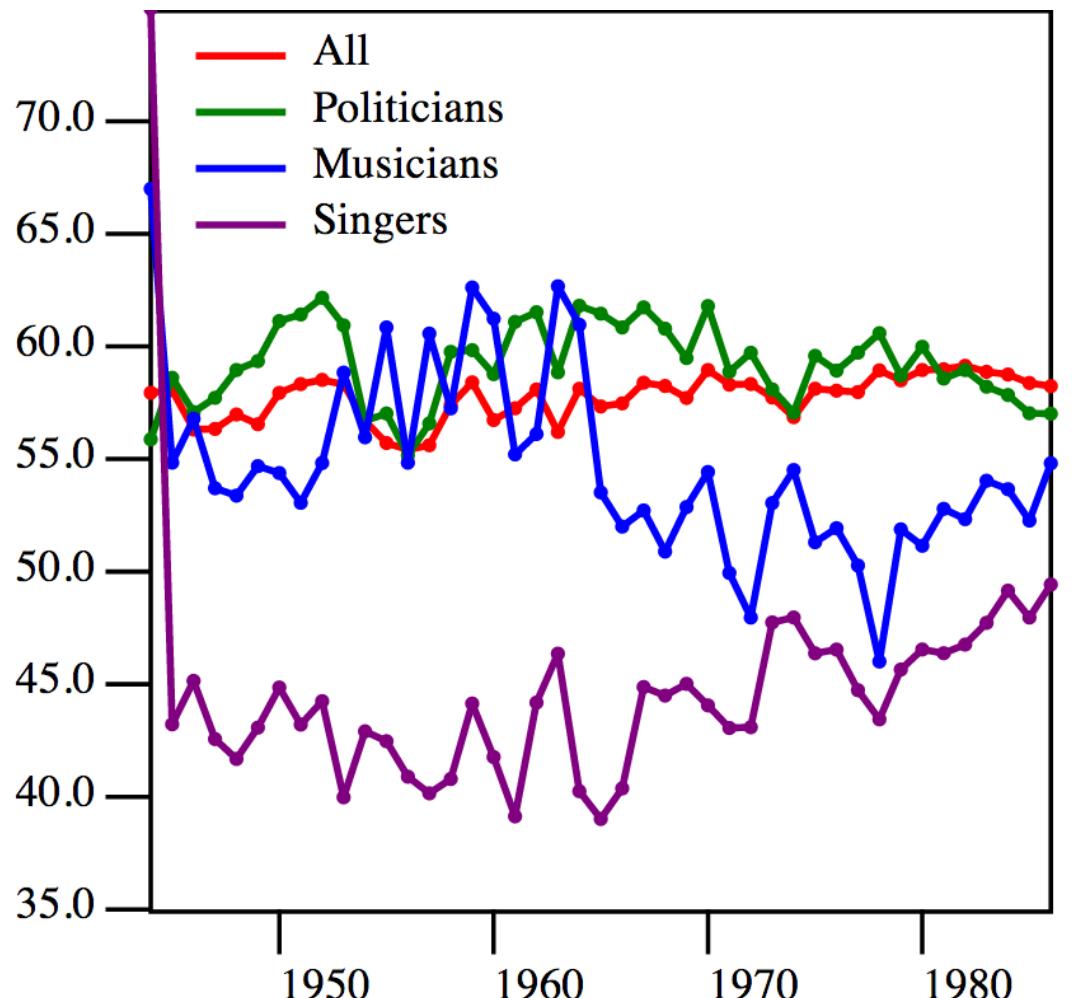


Information extracted from the news can help model and analyze political events.

Application: Analysis

Le Monde

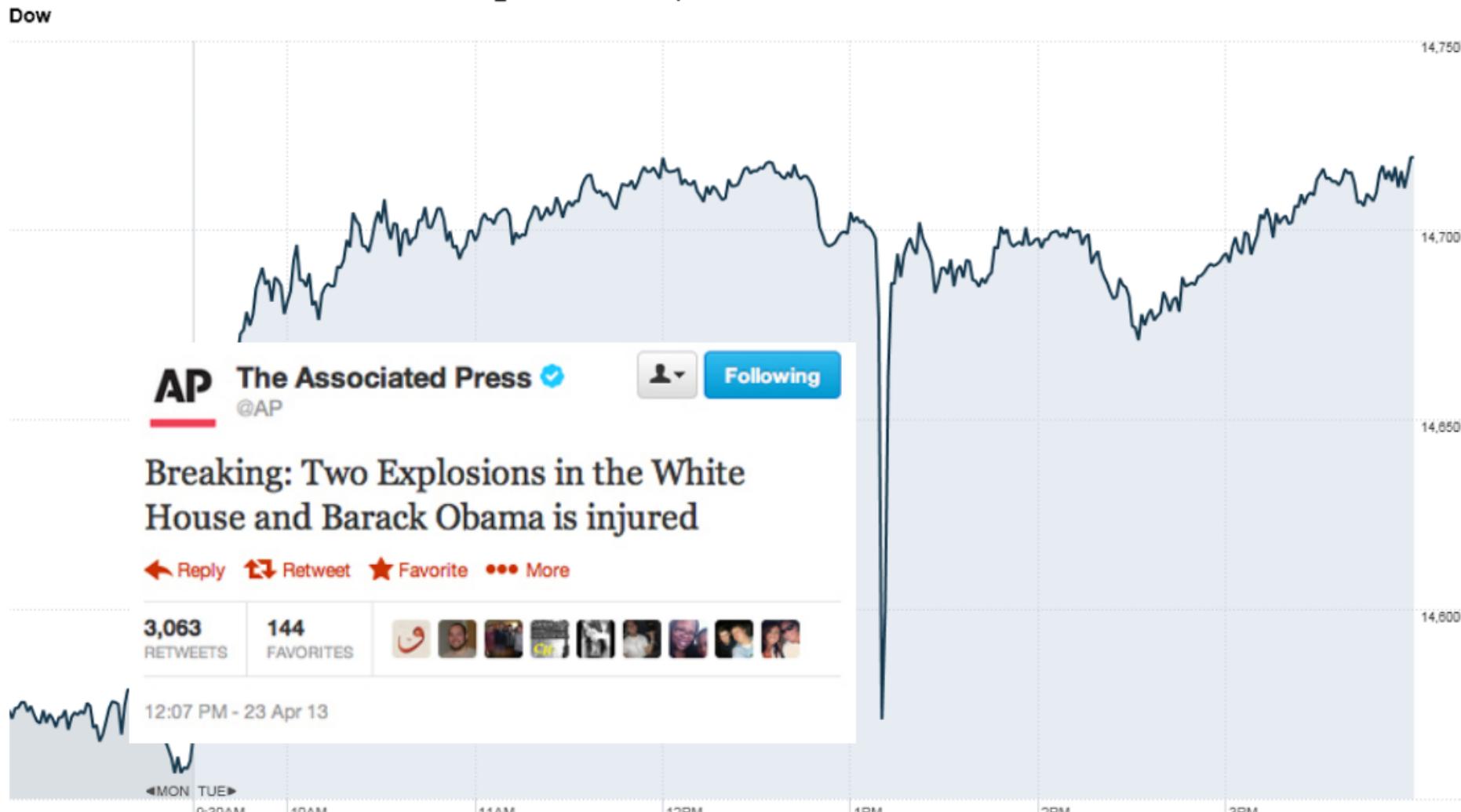
yago
select knowledge



Average
age of
people
mentioned

Application: Stock Market Selling

23 April 2013, Dow Jones (cnn.com)



AP account was hacked.

Automated trading systems react to news.

Slide from Pierre Senellart

>wipolo
42

Application: Assistants

Booking Confirmation

Guest Details

Name : Mr. Fabian Suchanek (ID#P5402)

Address : Campus E1.4
Saarbrücken, Saarland, 66111
Germany

4968193255014 4968193255014

f.m.suchanek@gmail.com

Room Type

Package/Rate Detail

Non-Refundable Promotion

Sleep Two(S)

Inclusions None

Amenities Wi-Fi High speed wireless Internet service

Requires

100% Deposit

forward email

to assistant



Wipolo

portal, bio > 26

Application: Biomedical



PubMed

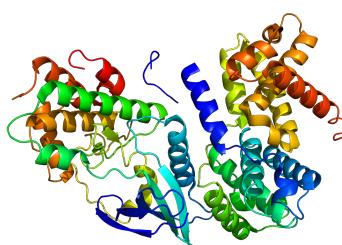
PubMed comprises more than 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

PubMed



tumor

- > suppressed by protein p53
- > switched off by kineasis proteins



In the last 30 years, scientists have uncovered 28 protein targets.

IBM's system discovered 6 in a month.



Time: IBM Watson's Startling Cancer Coup

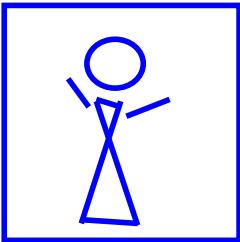
>go to Watson

Application: Portals



Extraction Engine

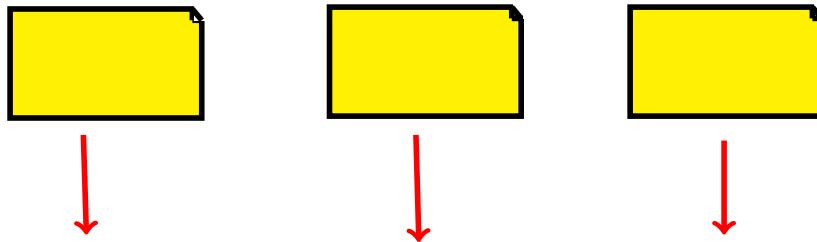
e.g. DBLife



Everything about
Angelina Jolie

Name, Birthdate, Lovers

Application: Portals



DBLife

Fabian Suchanek

Search

Fabian M. Suchanek

[Bing](#) [Citeseer](#) [DBLP](#) [Google](#) [Google Scholar](#) [Kosmix](#) [Wikipedia](#) [Yahoo!](#)



from Google Images

[more](#)

[News Archive](#)

Sorted by Year/Conf, [Year/Citation](#), [Citation](#)

[Community Statistics](#)

2011

31 PIKM 2011: the 4th ACM workshop for Ph.D. students in information and knowledge management. Anisoara Nica, Fabian M. Suchanek. CIKM 2011, 2633-2634. Web Search BibTeX [Download](#)

Related People

- [Gerhard Weikum](#)
- [Mauro Sozio](#)
- [Gjergji Kasneci](#)
- [Jian Pei](#)

[more](#)

Related Topics

- [location](#)
- [ranking](#)
- [privacy](#)
- [query processing](#)

[more](#)

Related Organizations

- [Microsoft Research](#)
- [Yahoo! Research](#)
- [University of Hong Kong](#)
- [Microsoft Statistic](#)

Application: Shopping



**Original Elvis Presley Kostüm
Elviskostüm Gr. 48/50 (M), 52/54 (L)**

Verfügbarkeit: sofort ■■■

Artikel-Nr: SM29142

Stückzahl: 1 Facebook Gefällt mir 0 Twitter Google+

44,90 EUR

inklusive 19% Mehrwertsteuer zzgl. Versand

Bitte wählen Sie eine Größe aus:
Größe bitte wählen... ▾

Kaufen MERKZETTEL WUNSCHLISTE

Application: Shopping

The screenshot shows a product listing for an "Original Elvis Presley Kostüm". The product image features a person dressed as Elvis Presley in a white jumpsuit with sequins and sunglasses, holding a microphone. The title is "Original Elvis Presley Kostüm" and "Elviskostüm Gr. 48/50 (M), 52/54 (L)". Availability is listed as "sofort versandfertig" with a green progress bar. The article number is SM29142. A quantity selector is present, along with social sharing buttons for Facebook, Twitter, and Google+. The price is prominently displayed as "44,90 EUR" (including VAT and shipping). A dropdown menu for selecting size is shown with the placeholder "bitte wählen...". Below the price are two buttons: a large red "Kauf(en)" button and smaller black buttons for "MERKZETTE" and "WUNSCHLISTE". Three red arrows point from the bottom of the slide towards the "Kauf(en)" button, the "MERKZETTE" button, and the "WUNSCHLISTE" button.

Title	Type	Price
Elvis Kostüm	Costume	44.90

Application: Shopping

pricegrabber
Smart Shopping Anytime Anywhere

Shop by Department ▾ elvis costumes

Women's Boots | Men's Wallets | Women's Handbags | Women's Dresses | north face jackets

Home > Clothing

Featured Brand

See Products from:



Filter By :

Price Range: \$1 - \$50

[Clear All](#)

Elvis Costumes Clothing - 325 Matches

[Quick Compare](#)

[Top 3 Products](#)

Category Matches

[Adult Costumes \(116\)](#)

[Kids' Costumes \(127\)](#)

[Adult Costume Accessories \(74\)](#)

[Men's Tees \(3\)](#)

[Women's Headwear \(2\)](#)

[Sunglasses \(1\)](#)

Price Range

\$1 - \$50

\$ 1 to \$ 50

[Update](#)



Rubie's Costume Co. Adult Elvis Costume

[Write a review](#)

Sold by
[HalloweenCostumes.com](#)
Seller Not Rated

As low as
\$47.99

[SEE IT](#)



Elvis Presley Eagle Jumpsuit Rock Gold...

[Write a review](#)

Seller Not Rated

As low as
\$24.00

[SEE IT](#)

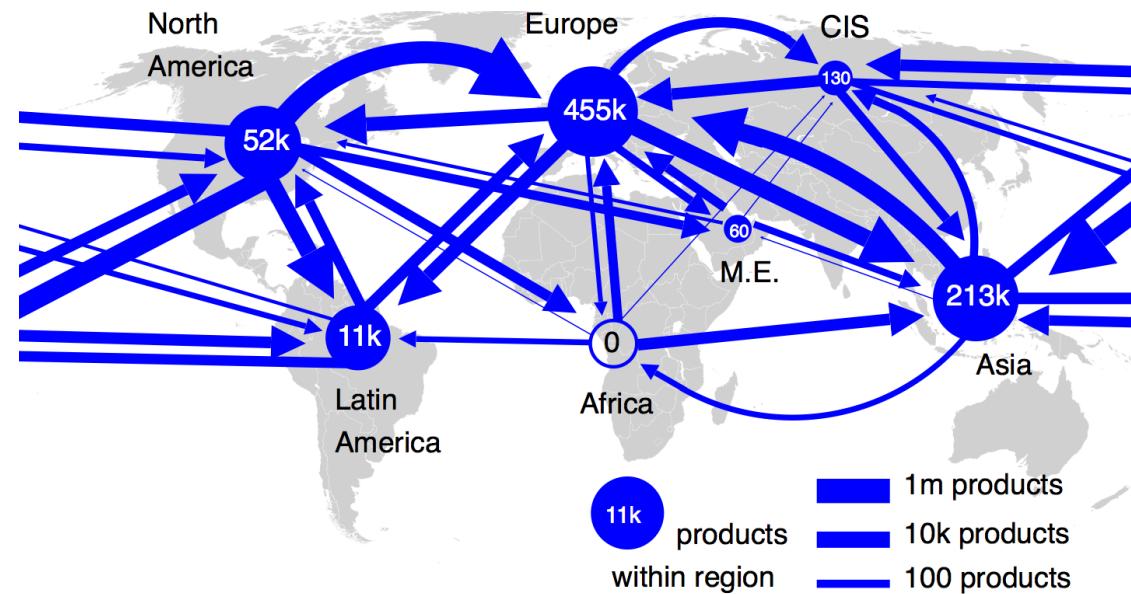
Price Grabber

[>Statistics](#)

Application: Entity statistics

Extracted

- 13m email addresses with their name
- 235K chemical products
- 1.4m books
- 1.1m products



Analyzed

- Global trade flow
- frequent email providers
- frequent people names and more

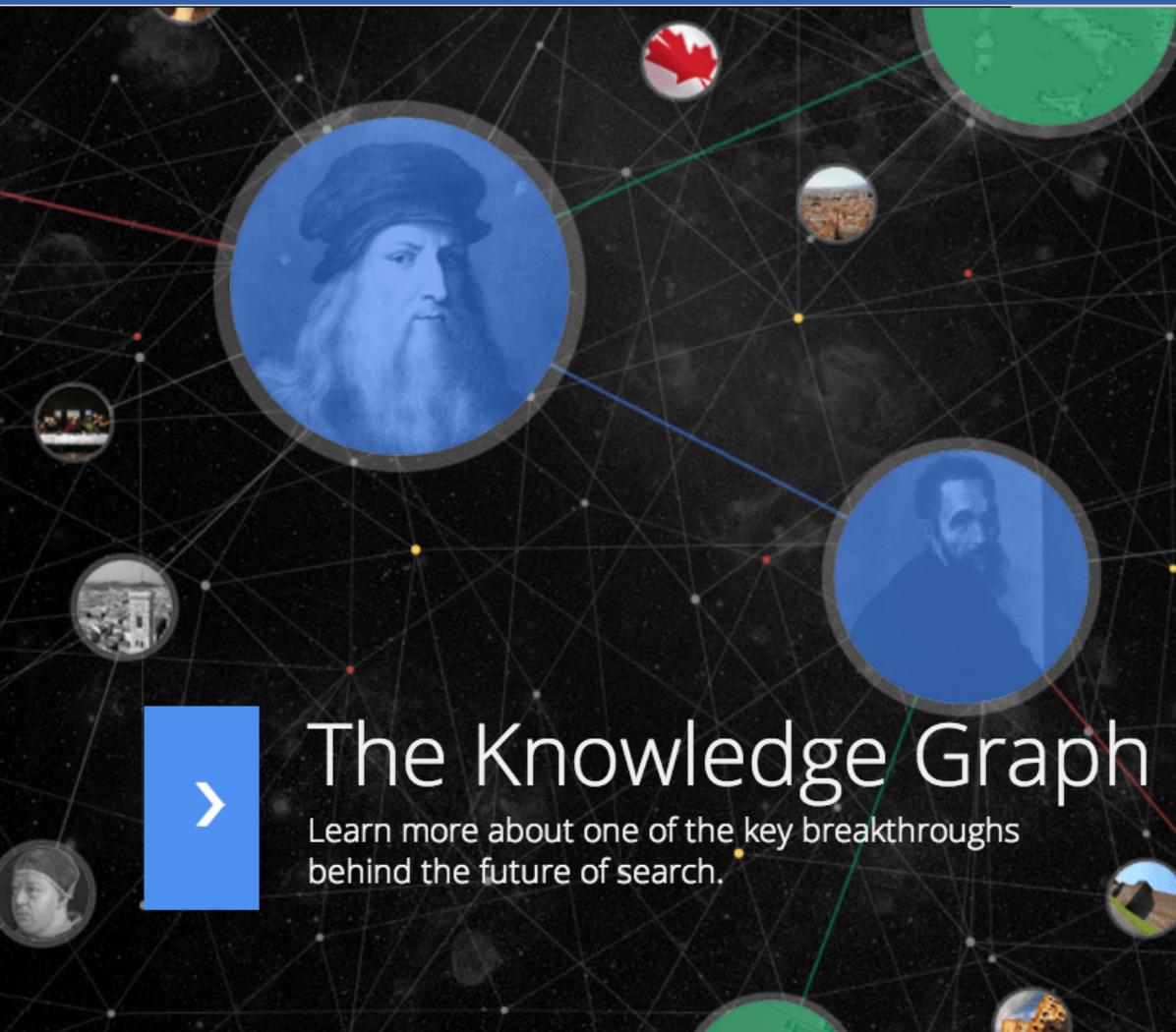
All data available online at

<http://resources.mpi-inf.mpg.de/d5/ibex/>

Application: Search

Google | Inside Search

Home Tips & Tricks Features Search Stories Playground Blog Help



Google Knowledge Graph

Google built its “knowledge graph”- a collection of factual knowledge

Application: Search

Google when was elvis born? 1 + Share

Web Images Maps Shopping Videos More ▾ Search tools SafeSearch ▾

About 33,500,000 results (0.23 seconds)

January 8, 1935

Elvis Presley, Date of birth

[Elvis Presley - Wikipedia, the free encyclopedia](#)
en.wikipedia.org/wiki/Elvis_Presley

Elvis Presley was born on January 8, 1935, in Tupelo, Mississippi, to 18-year-old Vernon Elvis and 22-year-old Gladys Love Presley, in the two-room shotgun ...

[Elvis Presley discography - Personal relationships of Elvis - Elvis Presley \(album\)](#)

Elvis Presley



Elvis Aaron Presley was an American singer and actor. A cultural icon, he is commonly known by the single name Elvis. One of the most popular musicians of the 20th century, he is often referred to as the "King of Rock and Roll" or "the King". [Wikipedia](#)

Born: January 8, 1935, [Tupelo](#)

Died: August 16, 1977, [Memphis](#)

Height: 1.83 m

Albums: [Elvis Presley](#), [Elvis' Christmas Album](#), [Christmas Duets](#), [More](#)

Songs

Jailhouse Rock	1957	Billboard Top Ro...
Suspicious Minds	1969	From Elvis in Me...
Always on My Mind	1973	Separate Ways

Factual information
helps Google
answer user queries.

>QA

Application: Search



best places to visit in new zealand

Web

Images

Maps

News

Videos

More

New Zealand > Points of interest



Milford Sound



Bay of Islands



Lake Taupo

Application: Search



who was Elvis' wife?

Web

Images

Videos

Shopping

News

More ▾

Search tools

About 29,600,000 results (0.33 seconds)

Priscilla Presley (m. 1967–1973)

Elvis Presley, Spouse



>QA

Knowledge Graph

Person	Profession
Elvis	Singer
Bob	Painter



The knowledge graph is based on

- different structured data sources
- IE from the Web (“Knowledge Vault”)

[See Wikipedia](#)

Appl.: Question Answering



Appl.: Question Answering



Where do babies come from?

Input interpretation:

Where do babies come from?

Result:

From their mothers.

(through the processes of fertilization, prenatal development, and birth)

Computed by **Wolfram Mathematica**

Appl.: Question Answering

- When is Madonna's next concert in Europe?
- Which protein inhibits atherosclerosis?
- Who was king of England when Napoleon I was Emperor of France?
- Has any scientist ever won the Nobel Prize in Literature?
- Which countries have a HDI comparable to Sweden's?
- Which scientific papers have led to patents?

Question Answering

Questions of this kind can be answered with structured information.

Person	Occupation	Awards
Russell	Mathematician	NP Literature
Bohr	Physicist	NP Physics
Elvis	Singer, Hero	Grammy

Question Answering

Structured information can be extracted from digital documents.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export

Languages
Afrikaans
አማርኛ
العربية
Aragonés^{aragonés}
অসমীয়া

Article Talk Read Edit View history Search

Bertrand Russell

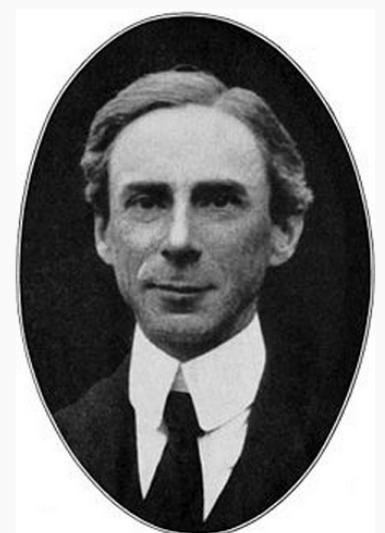
From Wikipedia, the free encyclopedia

Bertrand Arthur William Russell, 3rd Earl Russell, OM, FRS^[1] (18 May 1872 – 2 February 1970) was a British philosopher, logician, mathematician, historian, and social critic.^[2] At various points in his life he considered himself a liberal, a socialist, and a pacifist, but he also admitted that he had never been any of these in any profound sense.^[3] He was born in Monmouthshire, into one of the most prominent aristocratic families in Britain.^[4]

Russell led the British "revolt against idealism" in the early 20th century.^[5] He is considered one of the founders of analytic philosophy along with his predecessor Gottlob Frege and his protégé Ludwig Wittgenstein. He is widely held to be one of the 20th century's premier logicians.^[2] He co-authored, with A. N. Whitehead, *Principia Mathematica*, an attempt to ground mathematics on logic. His philosophical essay "On Denoting" has been considered a "paradigm of philosophy".^[6] His work has had a considerable influence on logic, mathematics, set theory, linguistics, computer science (see type theory and type system), and philosophy, especially philosophy of language, epistemology, and metaphysics.

Russell was a prominent anti-war activist; he championed anti-imperialism^{[7][8]} and went to prison for his pacifism during World War I.^[9] Later, he campaigned against Adolf Hitler, then criticised Stalinist totalitarianism, attacked the United States of America's involvement in the Vietnam War, and was an outspoken proponent of nuclear disarmament.^[10] In 1950 Russell was awarded the Nobel Prize in Literature, "in recognition of his varied and significant writings in which he champions humanitarian ideals and freedom of thought."^[11]

Bertrand Russell



Born 8 May 1872
Trellech, Monmouthshire, UK

Died 2 February 1970 (aged 97)
Penrhyneddraeth, Wales, UK

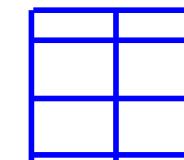
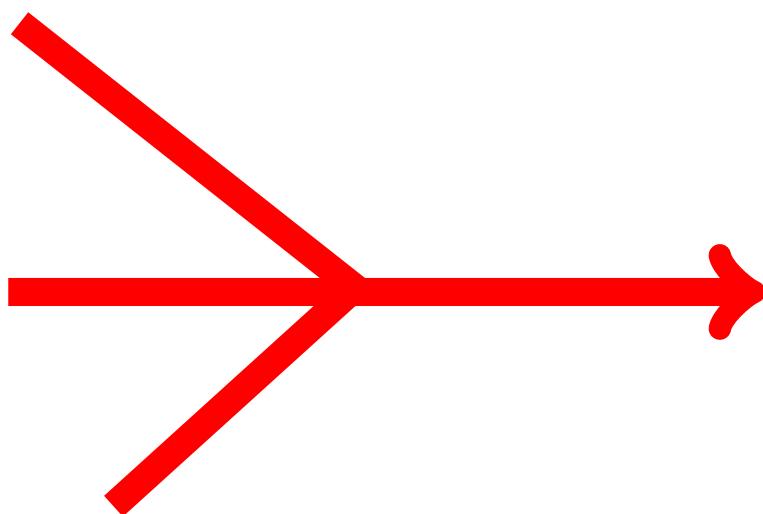
YAGO

The YAGO project extracts information from Wikipedia and other sources.



WordNet

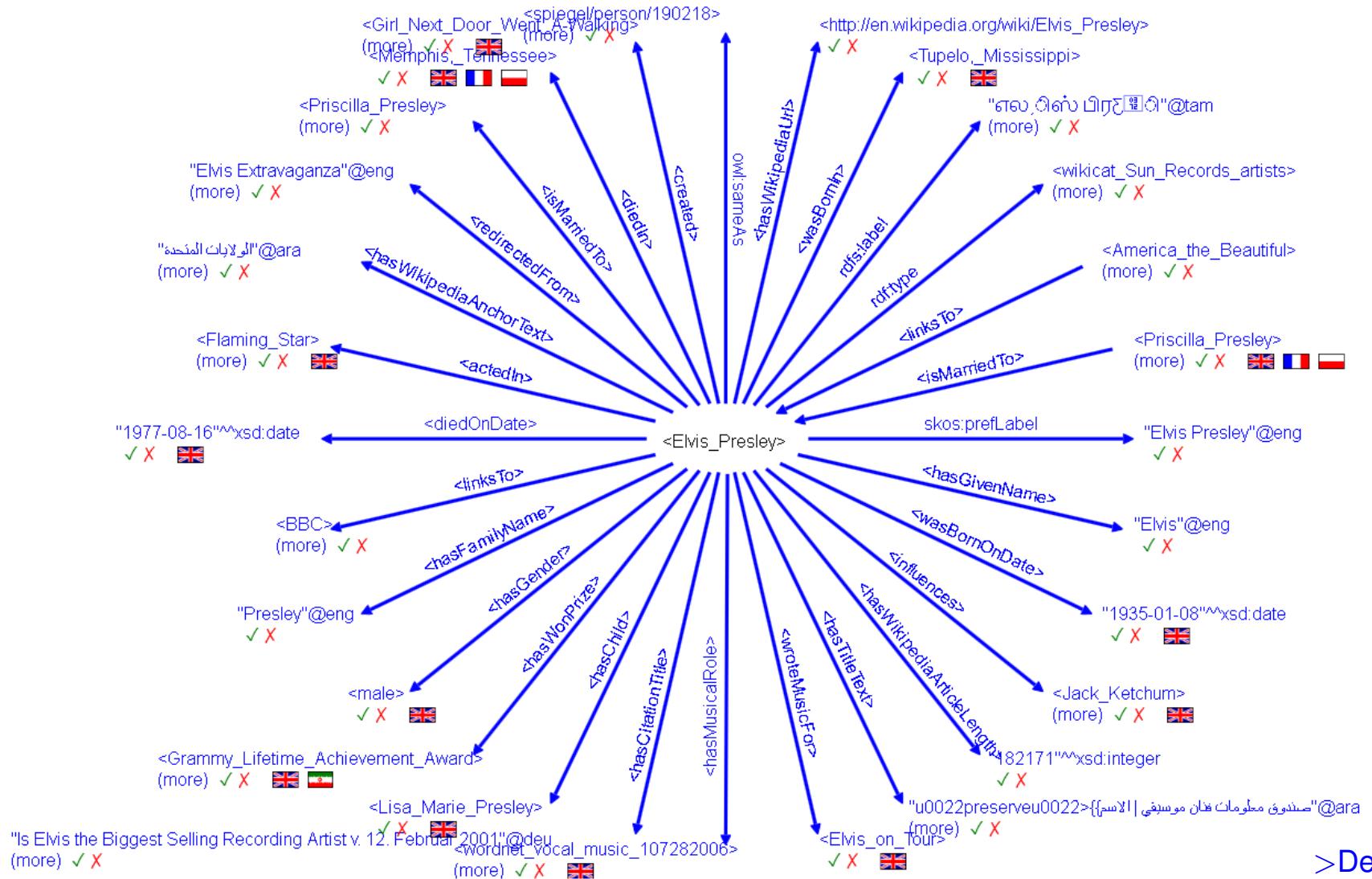
GeoNames



YAGO

YAGO

YAGO is a large knowledge graph built from Wikipedia.



DEANNA

Which mathematician won the Nobel Prize in Literature?

Show Sample Questions • Show Advanced Options • Time: 19006 ms

Structured Query

```
?x type wordnet_mathematician_110301261 ;  
?x hasWonPrize Nobel_Prize_in_Literature
```

Try it out 



```
?x type wordnet_mathematician_110301261 ; ?x hasWonPrize Nobel_Prize_in_Literature
```

Submit

IBM WATSON

IBM Watson is a question answering system based on many sources (i.a. YAGO).



Watson

New York Times article

It outperformed the 74-fold human winner of the Jeopardy! quizz show.

->knowledge-representation

Are Computers Intelligent?

- Computers can answer factual questions
- ... in natural language
- Computers can analyze events
- Computers beat humans in QA
- Oh, and computers drive cars, airplanes and space shuttles, and play better chess

Are Computers Intelligent?

- Computers can answer factual questions
 - Yes, but this is just from a database
- ... in natural language
 - Yes, but this is just database+parsing
- Computers can analyze events
 - Yes, but this is just joining databases
- Computers beat humans in QA
 - Yes, but this is just more of the above
- Oh, and computers drive cars, airplanes and space shuttles, and play better chess
 - Yes, but after all, it's all computation

AI Effect

Once you have done it, it all seems obvious.

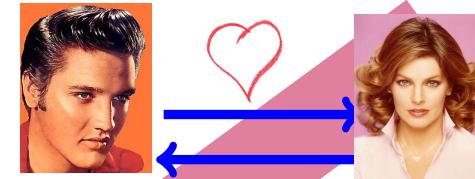
Or:

“Artificial Intelligence
is always that which
has not been done yet.”

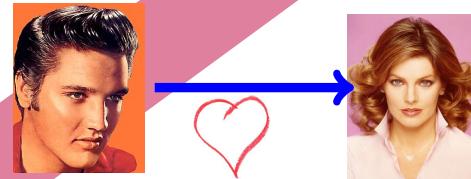
->knowledge-representation

Semantic IE

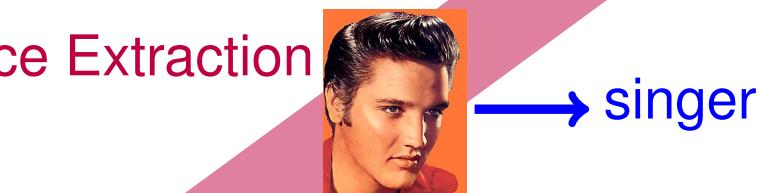
IE by Reasoning



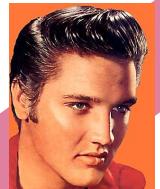
Fact Extraction



Instance Extraction

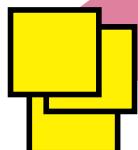


Entity Disambiguation



singer **Elvis**

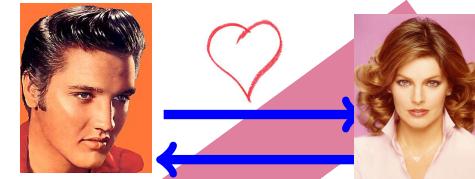
Entity Recognition



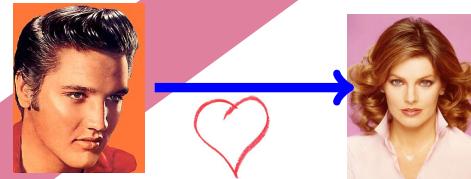
Preparation

Semantic IE

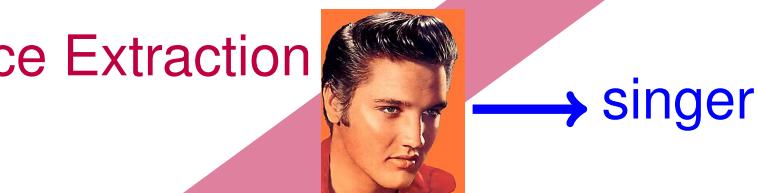
IE by Reasoning



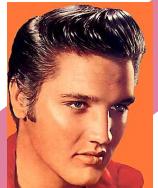
Fact Extraction



Instance Extraction



Entity Disambiguation



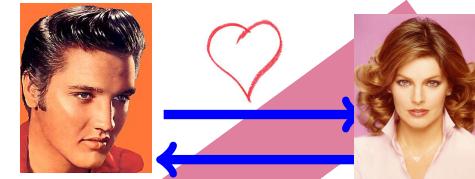
singer **Elvis**

Preparation

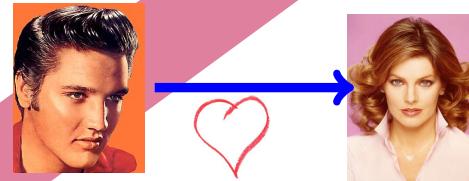
How do we represent world knowledge?
Where do we find the documents?

Semantic IE

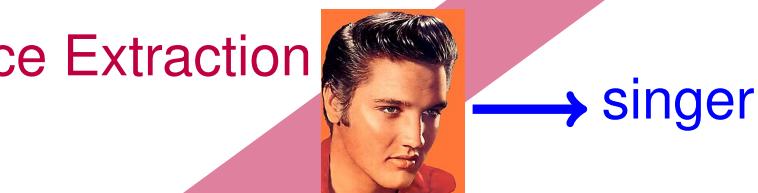
IE by Reasoning



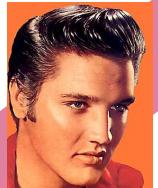
Fact Extraction



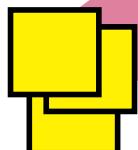
Instance Extraction



Entity Disambiguation



singer **Elvis**

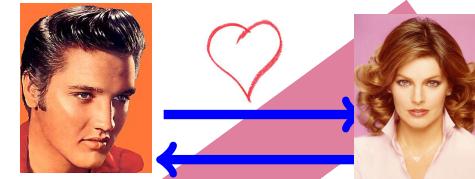


Preparation

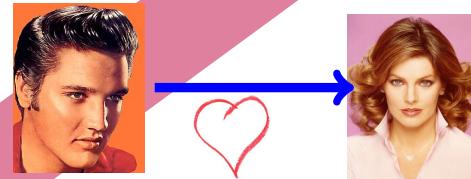
Named Entity Recognition (Tries)
Named Entity Annotation (Rule-based
or by machine learning)

Semantic IE

IE by Reasoning



Fact Extraction

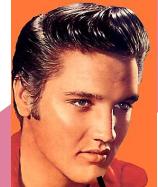


Instance Extraction



Entity Disambiguation

singer **Elvis**

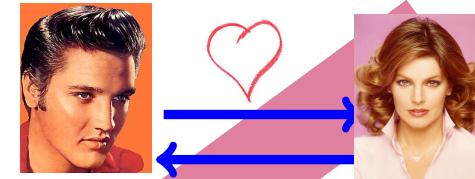


Preparation

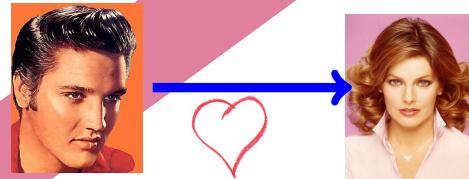
Finding which entity a text refers to,
for example with **AIDA**

Semantic IE

IE by Reasoning



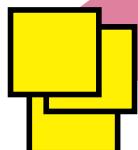
Fact Extraction



Instance Extraction



singer **Elvis**

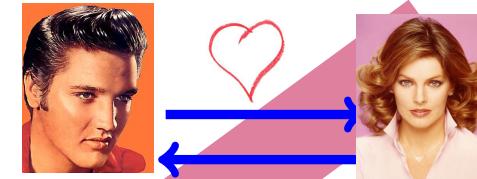


Preparation

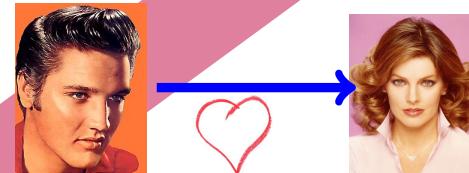
Extracting which class an entity belongs to (from the Web and from text documents), for example as in NELL

Semantic IE

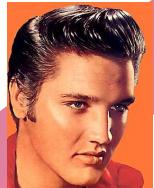
IE by Reasoning



Fact Extraction

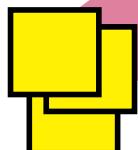


Instance Extraction



singer

singer **Elvis**

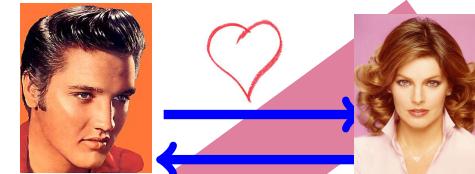


Preparation

DIPRE Algorithm, POS tagging,
Dependency Parsing.
Example: YAGO

Semantic IE

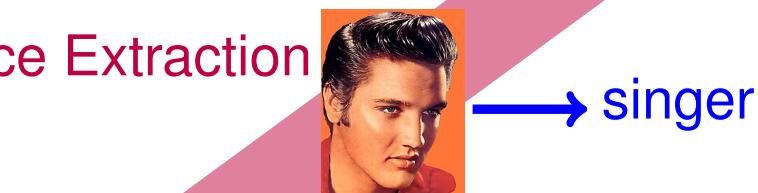
IE by Reasoning



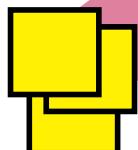
Fact Extraction



Instance Extraction



singer **Elvis**



Preparation

Taking into account
background knowledge.
Rule mining in KBs.

References

Sunita Sarawagi: Information Extraction

Serge Abiteboul et al: Web data management

Staab and Studer: Ontology Handbook

W3C: The Semantic Web

->knowledge-representation