



max planck institut
informatik

Advanced Topics in Information Retrieval

Semantic Search

Vinay Setty

`vsetty@mpi-inf.mpg.de`

Jannik Strötgen

`jannik.stroetgen@mpi-inf.mpg.de`

ATIR – June 9, 2016

What is Semantic Search?

Semantics

the study of meaning

Search

today is the sixth lecture ;-)

What is Semantic Search?

Semantic Search

Semantic Search is about

- going **beyond** documents and queries as **bag of words**
- having a **deeper understanding** of document contents by leveraging **world knowledge as structured data**
- going **beyond 10 blue links** and providing users with **direct answers** to their (natural language) questions







List Queries

Google who played in pulp fiction


Google albums by metallica



All Videos Shopping Images News More Search tools


Metallica / Albums



					
Kill 'Em All 1983	...And Justice for All 1988	Death Magnetic 2008	Metallica 1991	Load 1996	S&M 1999


Factoid Questions









[All](#)
[Images](#)
[News](#)
[Videos](#)
[Maps](#)
[More ▾](#)
[Search tools](#)

About 396,000 results (0,96 seconds)

Gerhard Weikum / Age

58 years

September 28, 1957



Feedback

Homepage: [Gerhard Weikum \(Max-Planck-Institut für Informatik\)](http://people.mpi-inf.mpg.de/~weikum/)

people.mpi-inf.mpg.de/~weikum/ ▾

Jan 19, 2009 - Gerhard Weikum is a Research Director at the Max-Planck Institute for ...
 Slides; Knowledge Harvesting in the Age of Big Data Analytics (jointly ...

Gerhard Weikum - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Gerhard_Weikum ▾

Gerhard Weikum

Computer science researcher



Gerhard Weikum is a Research Director at the Max Planck Institute for Computer Science in Saarbrücken, Germany, where he is leading the databases and information systems department. [Wikipedia](#)

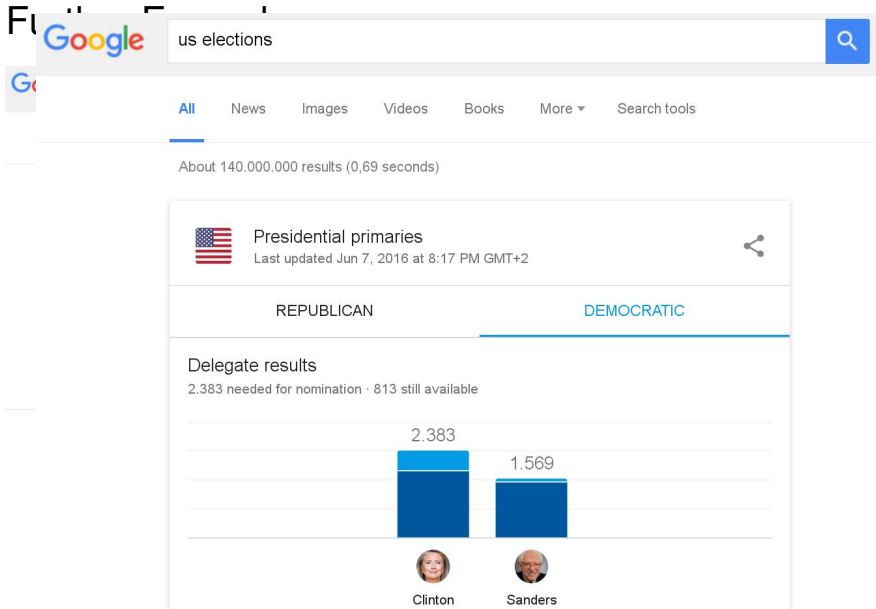
Born: September 28, 1957 (age 58), Frankfurt

Books: Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery

Feedback

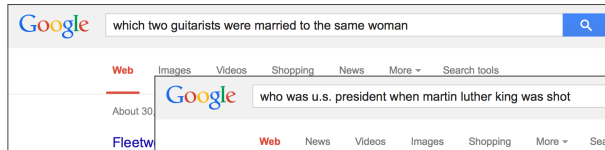


max planck institut
informatik



2014 Lecture


Semantic Search




Still Beyond
Hope



Still not solved ;-)





[All](#)
[Images](#)
[Videos](#)
[Shopping](#)
[News](#)
[More ▼](#)
[Search tools](#)

About 41.700.000 results (1,25 seconds)

[Lita Ford - Wikipedia, the free encyclopedia](#)

https://en.wikipedia.org/wiki/Lita_Ford ▼

Lita Rossana Ford (born September 19, 1958) is an English-born American rock guitarist, The couple were married after knowing each other for only two weeks. ... Runaways, in which she spoke candidly about her time in the all-girl band.

[Jim Gillette - Lita \(album\) - Out for Blood - Living Like a Runaway](#)

[Eric Clapton - Wikipedia, the free encyclopedia](#)

https://en.wikipedia.org/wiki/Eric_Clapton ▼

Eric Patrick Clapton, CBE (born 30 March 1945), is an English rock and blues guitarist, singer ... Two of his most popular recordings were "Layla", recorded with Derek and the Dominos; and Robert Johnson's "Crossroads", recorded with Cream. unavailable woman and who went crazy because he could not marry her.

[Bonnie Guitar - Wikipedia, the free encyclopedia](#)

https://en.wikipedia.org/wiki/Bonnie_Guitar ▼

Bonnie Guitar is an American singer, musician, producer and businesswoman. She is best remembered for her 1957 country-pop crossover hit "Dark Moon". She became one of the first female country music singers to have hit songs ... The final version consisted of just two guitars and a bass backing Bonnie. The song was ...



max planck institute
for informatics

© Jannik Strötgen – ATIR-06

Solved...



who was us president when martin luther king was shot



All

News

Videos

Images

Shopping

More ▾

Search tools

About 766.000 results (0,97 seconds)

United States of America / President (1968)

Lyndon B. Johnson



More about Lyndon B. Johnson

Feedback

[Assassination of Martin Luther King, Jr. - Wikipedia, the free ...](https://en.wikipedia.org/wiki/Assassination_of_Martin_Luther_King,_Jr.)

https://en.wikipedia.org/wiki/Assassination_of_Martin_Luther_King,_Jr.

President Lyndon B. Johnson declared April 7 a national day of mourning for Rev. King. A crowd of 300,000 attended his funeral two days later on April 9. Vice President Hubert

Solved, with explanation



who was us president when martin luther king was shot



Examples Random

Input interpretation:

United States President Martin Luther King Jr. assassinated date

Results:

Hide details

Lyndon B. Johnson (from 22. November 1963 to 20. January 1969)

Martin Luther King Jr. assassinated date 04. April 1968

Basic information:

official position	President (36 th)
country	United States
political affiliation	Democratic



max planck institut
informatik

© Jannik Strötgen – ATIR-06

12 / 64

But still things beyond hope



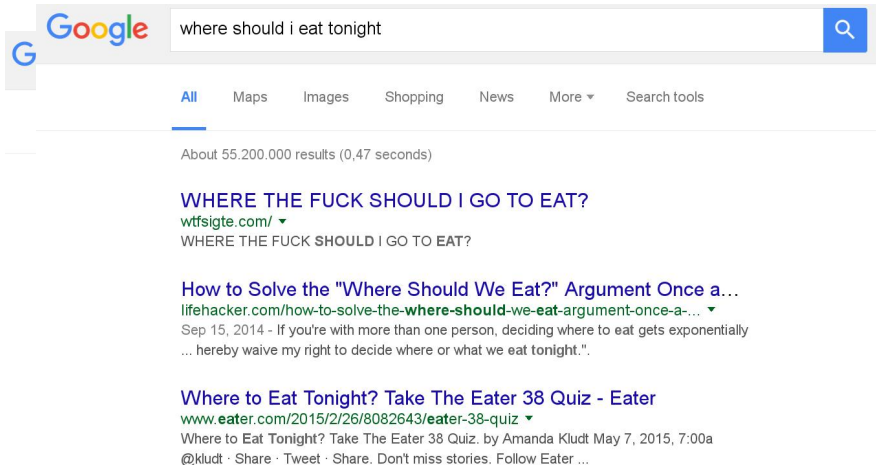
who was German soccer champion|when martin luther king was shot



≡ Examples ↔ Random

Wolfram|Alpha doesn't know how to interpret your input.

...



A screenshot of a Google search page. The search bar contains the text "where should i eat tonight". Below the search bar, the "All" tab is selected. The search results show three entries:

- WHERE THE FUCK SHOULD I GO TO EAT?**
[wtfsgite.com/](#) ▼
WHERE THE FUCK SHOULD I GO TO EAT?
- How to Solve the "Where Should We Eat?" Argument Once a...**
[lifehacker.com/how-to-solve-the-where-should-we-eat-argument-once-a-...](#) ▼
Sep 15, 2014 - If you're with more than one person, deciding where to eat gets exponentially ... hereby waive my right to decide where or what we eat tonight."
- Where to Eat Tonight? Take The Eater 38 Quiz - Eater**
[www.eater.com/2015/2/26/8082643/eater-38-quiz](#) ▼
Where to Eat Tonight? Take The Eater 38 Quiz. by Amanda Kludt May 7, 2015, 7:00a
@kludt · Share · Tweet · Share. Don't miss stories. Follow Eater ...

But not just for question-style queries

Google

pulp fiction

Q

Google

max planck institute for computer science

Q

All

Maps

Images

News

Videos

More ▾

Search tools

About 583,000 results (0,72 seconds)

Max Planck Institute for Informatics

https://www.mpi-inf.mpg.de/ ▾

Computer Science Department, Saarland University - Max Planck Institute for ... The Max-Planck Institute for Informatics is devoted to cutting-edge research in ...

Departments - Databases and Information ... - Algorithms & Complexity - English

Max-Planck-Institut für Informatik: Departments - Max Planck Institute ...

https://www.mpi-inf.mpg.de/departments/ ▾

Human-Computer Interaction and Interactive Technologies - Prof. Dr. Jürgen Steimle ... Formal Methods and Theoretical Computer Science, Prof. Dr. Viorica ...

Algorithms & Complexity - Max Planck Institute for Informatics

https://www.mpi-inf.mpg.de/departments/algorithms-complexity/ ▾

We study fundamental algorithmic problems that arise in various fields of mathematics and computer science. Historically we have focused on problems in graph ...

International Max Planck Research School for Computer Science

www.imprs-cs.de/ ▾

The International Max Planck Research School for Computer Science (IMPRS-CS) is a graduate program jointly run by the Max Planck Institute for Informatics ...

Max Planck Institute for Informatics - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Max_Planck_Institute_for_Informatics ▾

The Max Planck Institute for Informatics is a research institute in computer science with a focus on algorithms and their applications in a broad sense. It hosts ...

Max-Planck-Institut für Informatik – Wikipedia

https://de.wikipedia.org/wiki/Max-Planck-Institut_für_Informatik ▾

Translate this page

Jump to International Max Planck Research School for Computer Science - Sie ist die englischsprachige Graduiertenschule des MPI für Informatik ...

International Max Planck Research School for Computer Science ...

https://www.daad.de/deutschland/studienangebote/international_/_en/?p=d&s=kr... ▾

International Max Planck Research School for Software Systems (MPI-SWS) ...

Max Planck Institute for Informatics

Website

Directions

The Max Planck Institute for Informatics is a research institute in computer science with a focus on algorithms and their applications in a broad sense. It hosts fundamental research as well as research for various application domains. [Wikipedia](#)

Address: E1 5., Campus, 66123 Saarbrücken

Phone: 0681 93039100

Founded: 1988

[Suggest an edit](#)

Add missing information

[Add business hours](#)

[Add category](#)

Reviews

3 Google reviews

[Write a review](#) [Add a photo](#)

People also search for

[View 10+ more](#)

informatik

max planck research school for computer science

max planck research school for software systems (mpi-sws)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)

max planck research school for informatics (mpi-inf)


max planck research school for informatics (mpi-inf)


© Jannik Strötgen



AFIR-06

15 / 64

And of course: it's not just google







Deutsch Sign in  


Web Images Videos Maps News Explore



5.900.000 RESULTS Date Language Region

Pulp Fiction – Wikipedia Translate this page
https://de.wikipedia.org/wiki/Pulp_Fiction
Pulp Fiction ist ein Film von und mit Quentin Tarantino aus dem Jahr 1994. Der Film

Pulp Fiction
 IMDb
8,9/10 ★★★★★


















Sign in  1 


Web Images Videos Maps News Explore

Paris - Points of interest

 Eiffel Tower
 Louvre
 Arc de Triomphe
 Notre Dame de Paris
 Champs-Élysées
 Palace of Versailles
 Disneyland Paris
 Montmartre

Pulp Fiction (1994) - IMDb
www.imdb.com/title/tt0110912
CRIME/DRAMA JOHN TRAVOLTA/UMA THURMAN/SAMUEL L. JACKSON R
Information about the film from the Internet Movie Database (IMDb), including cast/credits, basic film outline and extra information.

Cast
 John Travolta
 Samuel L.
 Uma
 Bruce Willis
 Harvey


Pulp Fiction | Film 1994 | moviepilot.de Translate this page



But it is not perfect





Sign in



7



Web

Images

Videos

Maps

News

Explore

154,000 RESULTS

Any time ▾

Who was president when the Berlin Wall fell?

President Ronald Reagan

Jim Baker was secretary of state in the administration of **President Ronald Reagan** 20 years ago when the Berlin Wall fell.

James Baker on the Fall of the Wall - ABC News

abcnews.go.com/International/james-baker-fal...

Is this answer helpful?

Berlin Wall - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Berlin_Wall ▾

The fall of the Berlin Wall paved the way for German reunification. ... Khrushchev had been emboldened by **US President** John F. Kennedy's tacit indication ...

Background · Erection of the ... · Construction begins, 1961 · Structure and ...

fall of the Berlin Wall - US News & World Report

www.usnews.com > News > Ken Walsh's Washington ▾

Nov 07, 2014 · 25 years after the fall of the Berlin Wall, world leaders should learn from **President** ... More on the Berlin Wall: Bringing Down Walls, 25 Years Later; ...

Who was President when the Berlin Wall came down?

www.answers.com > ... > World War 2 > WW2 Axis Powers > Germany in WW2 ▾

Who was **President when the Berlin Wall** came down? ... and George H.W. Bush took over as **President**. The wall did not begin to come down ... The Berlin Wall 'fell' on ...



[See all images](#)



Berlin Wall

The Berlin Wall was a barrier that divided Berlin from 1961 to 1989. Constructed by the German Democratic Republic, starting on 13 August 1961, the Wall completely cut off West Berlin from surrounding East Germany and from East Berlin until government officials opened it in November 1989. Its demolition officially began on 13 June 1990 ... +

Wikipedia

Address: Bernauer Straße 111, Berlin · [Directions](#)

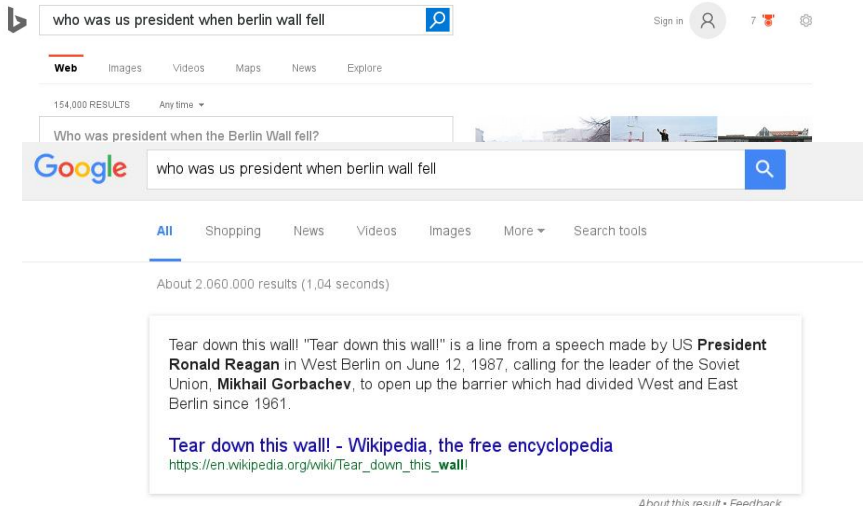
Construction started: Aug 13, 1961

Destroyed: Nov 09, 1989



max planck institut
informatik

But it is not perfect



The screenshot shows a Google search interface. At the top, there's a navigation bar with links for 'Web', 'Images', 'Videos', 'Maps', 'News', and 'Explore'. The search bar contains the query 'who was us president when berlin wall fell'. Below the search bar, it says '154,000 RESULTS' and 'Any time'. The search results show a snippet from Wikipedia: 'Tear down this wall! "Tear down this wall!" is a line from a speech made by US **President Ronald Reagan** in West Berlin on June 12, 1987, calling for the leader of the Soviet Union, **Mikhail Gorbachev**, to open up the barrier which had divided West and East Berlin since 1961.' Below this snippet is a link to the Wikipedia article: 'Tear down this wall! - Wikipedia, the free encyclopedia' with the URL 'https://en.wikipedia.org/wiki/Tear_down_this_wall!'. At the bottom of the snippet, there's a link 'About this result • Feedback'.

Fall of the Berlin Wall: 25 Years Later, Leaders Should Learn From ...

www.usnews.com/.../fall-of-the-berlin-wall-25-years-later-leaders-should-learn-from-...
 max plack ins this years after the fall of the Berlin Wall, Strotgen - ATIR-06
 informatik Soviet leader, Mikhail Gorbachev, former U.S. President ...



But it is not perfect


fall of the Berlin Wall

November 9, 1989

presidency of Ronald Reagan

January 20, 1981 – January 20, 1989

But it is not perfect

Google 

[All](#) [Images](#) [News](#) [Videos](#) [Shopping](#) [More](#) [Search tools](#)

About 579 000 results (0,77 seconds)

Ronald Reagan / Presidential term

January 20, 1981 – January 20, 1989

[Feedback](#)


Presidency of Ronald Reagan - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Presidency_of_Ronald_Reagan

The United States **presidency of Ronald Reagan**, also known as the **Reagan administration**, was a Republican administration headed by **Ronald Reagan** from January 20, 1981, to January 20, 1989.


Vice President: George H. W. Bush **Nationality:** American
Political party: Republican **Succeeded by:** George H. W. Bush

Domestic policy of the Ronald ...
 The domestic policy of the Ronald Reagan administration was the ...

Category:Presidency of ...
 Ronald Reagan was President of the United States from 1981 to ...

Ronald Reagan 

40th U.S. President





Ronald Wilson Reagan was an American politician and actor who was the 40th President of the United States from 1981 to 1989. [Wikipedia](#)


Born: February 6, 1911, Tampico, Illinois, United States
Died: June 5, 2004, Bel-Air, Los Angeles, California, United States
Party: Republican Party
Vice president: George H. W. Bush (1981–1989)
Nicknames: Ronnie, Dutch, The Jelly Bean Man, [more](#)

Quotes [View 7+ more](#)

But it is not perfect







All Images Videos News Maps More Search tools

About 6.740.000 results (0,64 seconds)

After several weeks of civil unrest, the East German government announced on 9 November 1989 that all GDR citizens could visit West Germany and West **Berlin**. Crowds of East Germans crossed and climbed onto the **Wall**, joined by West Germans on the other side in a celebratory atmosphere.



Berlin Wall - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Berlin_Wall

About this result • Feedback

Berlin Wall - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Berlin_Wall


After several weeks of civil unrest, the East German government announced on 9 November 1989 that all GDR citizens could visit West Germany and West **Berlin**. Crowds of East Germans crossed and climbed onto the **Wall**, joined by West Germans on the other side in a celebratory atmosphere.
 Bed of nails - Conrad Schumann - Inner German border - East Berlin

Berlin Wall - Cold War - HISTORY.com
www.history.com/topics/cold-war/berlin-wall

The **Berlin Wall**. The **Fall** of the Wall. On November 9, 1989, as the Cold War began to thaw across Eastern Europe, the spokesman for East Berlin's Communist ...
 Videos - Kennedy Responds to Berlin - Ask HISTORY: Kennedy - Berlin Airlift

The Rise and Fall of the Berlin Wall | HISTORY
www.history.co.uk/shows/the-rise-and-fall-of-the-berlin-wall

9 November 1989 marks the infamous **fall of the Berlin Wall**. On midnight of that day, East Germany's Communist rulers gave permission for gates along the **Wall** to be opened as a result of days of mass protest.




Berlin Wall


The Berlin Wall was a barrier that divided Berlin from 1961 to 1989. Constructed by the German Democratic Republic, starting on 13 August 1961, the Wall completely cut off West Berlin from surrounding ...
 Wikipedia

Destruction date: November 9, 1989
Construction started: August 13, 1961
Height: 3.6 m
Opened: 1961
Destroyed by: Demolition


People also search for




Branden... Gate




Checkpoint Charlie



Reichstag building



East Side Gallery



Museum Island

Feedback

Outline

- 1 Semantic Web
- 2 Knowledge Bases
- 3 Entity Linking
- 4 Semantic Search

Outline

- 1 Semantic Web
- 2 Knowledge Bases
- 3 Entity Linking
- 4 Semantic Search

Semantic Web

Semantic Web is an extension of the **World Wide Web**,

Scientific American: Feature Article: The Semantic Web: May 2001



The Semantic Web

A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

[by TIM BERNERS-LEE, JAMES HENDLER and ORA LASSILA](#)

Semantic Web Standards

World Wide Web Consortium (W3C) Semantic Web standards

- **Unified Resource Identifier (URI)**
to uniquely identify an abstract of physical resource
- **Resource Description Framework (RDF)**
to describe properties of abstract or physical resources
- **Resource Description Framework Schema (RDF/S)**
to describe schemata of properties of abstract of physical resources
- **Web Ontology Language (OWL)**
to describe ontologies
- **SPARQL Protocol and Query Language (SPARQL)**
to formulate queries over properties of abstract or physical resources
- **schema.org [ASSIGNMENT]**

URI

Unified Resource Identifier (URI) is a string of characters that uniquely identifies an abstract or physical resource



http://en.wikipedia.org/wiki/Foo_Fighters
<http://www.bbc.co.uk/music/artists/67f66c07-6e61-4026-ade5-7e782fad3a5d>
<http://www.musicbrainz.org/artist/67f66c07-6e61-4026-ade5-7e782fad3a5d>

<http://www.host.org>www.host.org/pub/bands[pub/bands?query=Foo_Fighters](http://www.host.org/pub/bands?query=Foo_Fighters)

- **scheme** (e.g., http, ftp, urn) determines interpretation of URI
- **authority** indicates who is responsible for resource (e.g., a host)
- **path** provides hierarchical information for identifying the resource
- **query** provides non-hierarchical information for identifying the resource
- **fragment** refers to a specific part of resource

RDF

Resource description framework (RDF) provides a data model to describe properties of resources (identified by their URI)

- **RDF statements** are (S,P,O) **triples** consisting of a **subject** (URI), a **predicate** (URI), and an **object** (URI or literal)
- example

<http://dbtune.org/musicbrainz/page/artist/67f66c07-6e61-4026-ade5-7e782fad3a5d> (S)

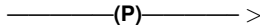
<http://xmlns.com/foaf/spec/20100809.html#member> (P)

<http://dbtune.org/musicbrainz/page/artist/4d5f891d-9bce-45ae-ad86-912dd27252fa> (O)



(S)

Foo Fighters



(P)

have member



(O)

Dave Grohl

RDF

RDF triples form an **RDF graph** (a labeled directed multigraph)

what's a

graph

directed graph

labeled directed graph

labeled directed multigraph

RDF

RDF triples form an **RDF graph** (a labeled directed multigraph)

- **Namespaces** represent common URI prefixes and allow for a more compact representation of RDF data
- RDF/N3 as one common text representation of RDF data

RDF

RDF triples form an **RDF graph** (a labeled directed multigraph)

- **Namespaces** represent common URI prefixes and allow for a more compact representation of RDF data
- RDF/N3 as one common text representation of RDF data

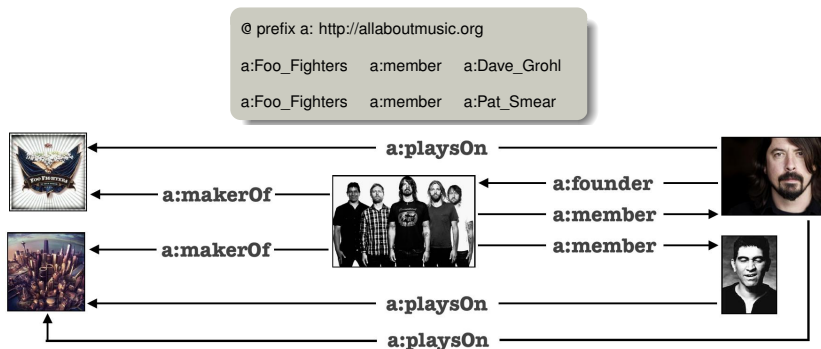
```
@ prefix a: http://allaboutmusic.org
```

```
a:Foo_Fighters    a:member    a:Dave_Grohl  
a:Foo_Fighters    a:member    a:Pat_Smear
```

RDF

RDF triples form an **RDF graph** (a labeled directed multigraph)

- **Namespaces** represent common URI prefixes and allow for a more compact representation of RDF data
- RDF/N3 as one common text representation of RDF data



SPARQL

SPARQL Protocol and RDF Query Language (SPARQL) is a query language for the Semantic Web standardized by W3C

SPARQL from the linguistic point of view:

a recursive acronym

- SPARQL has a **SQL-inspired syntax** to define **graph patterns** and retrieves **all matching subgraphs** as query answers

SPARQL

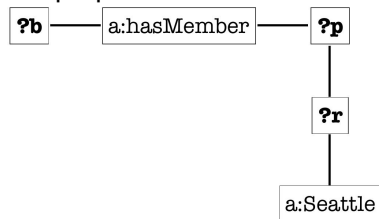
SPARQL Protocol and RDF Query Language (SPARQL) is a query language for the Semantic Web standardized by W3C

- SPARQL has a **SQL-inspired syntax** to define **graph patterns** and retrieves **all matching subgraphs** as query answers

Query:

```
PREFIX a:<http://allmusic.org/>
SELECT DISTINCT ?b, ?r, ?p WHERE {
  ?b a:hasMember ?p .
  ?p ?r a:Seattle .
}
ORDER BY ?p
```

Graph pattern:



Answer:



Linked Open Data Project

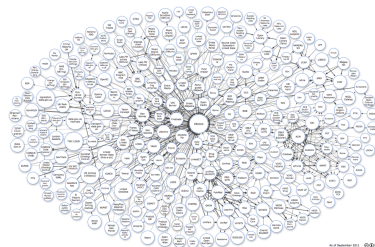
problem:

Resources can be referred to by
different URIs in different RDF datasets

- **Linked Open Data Project** creates a **Web of Linked Open Data** by establishing owl:sameAs links between RDF data sources

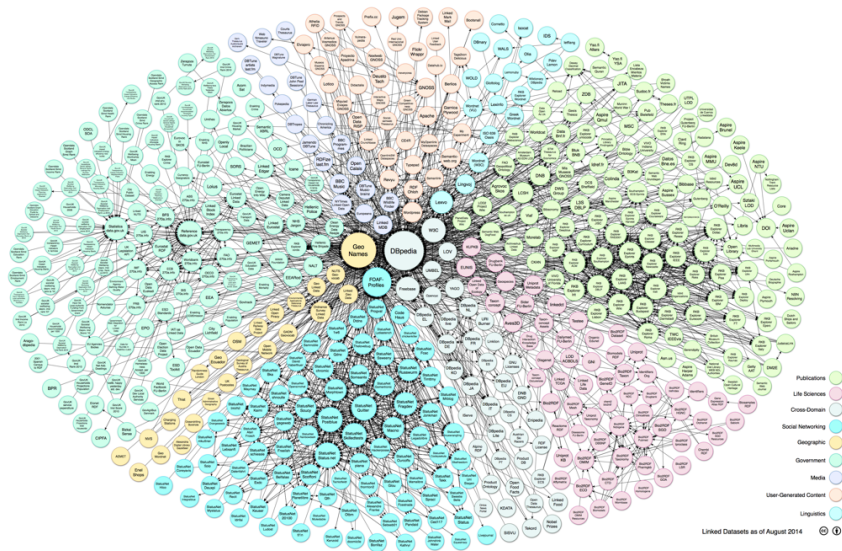
As of 2011 (!)

- 295 RDF data sources
- 32 billion RDF triples
- 504 million links



<http://linkeddata.org/>

Linked Open Data Project – 2014



Outline

- 1 Semantic Web
- 2 Knowledge Bases**
- 3 Entity Linking
- 4 Semantic Search

Knowledge Bases

Knowledge bases (e.g., DBpedia, Wikidata, YAGO)

- provide data about real-world **named entities**, their **properties** and **relations** between them
- have typically been **extracted from Wikipedia** and different other data sources (WordNet, GeoNames, MusicBrainz, etc.)
- form the core of the **linked open data cloud**

Wikipedia

Collaboratively edited encyclopedia that forms the common basis of other knowledge bases

- **5.1 M articles** / 3.4 M edits per month / 7,398 M views per month (for English, April 2016)
- in **293 languages** (with vastly different coverage)
- templates, **infoboxes**, and **categories** (quasi-structured)
- article contents (unstructured) with rich interlinkage
- revision history (who edited which page at what time)
- usage data (e.g., access statistics)
- started in 2001



Categories: 1963 births | 20th-century American singers | 21st-century American singers | American baritones | American heavy metal guitarists | American heavy metal singers | American male singer-songwriters | American people of English descent | American people of German descent | American people of Irish descent | American people of Scottish descent | American rock singers | American rock songwriters | American male songwriters | American skateboarders | Critics of Christian Science | Former Christian Scientists | Living people | Metallica members | Musicians from Downey, California | People from Brea, California | Rhythm guitarists | Rock and Roll Hall of Fame inductees | Singers from California | Songwriters from California | American hunters | American fishers | National Rifle Association people

Wikipedia



WIKIPEDIA
The Free Encyclopedia

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

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

In other projects

Article [Talk](#)

[Read](#)

[View source](#)

[View history](#)



James Hetfield

From Wikipedia, the free encyclopedia

Not to be confused with [James Hatfield](#).

James Alan Hetfield (born August 3, 1963) is an American musician, singer and songwriter known for being the co-founder, lead vocalist, rhythm guitarist and main songwriter for the American **heavy metal** band **Metallica**. Hetfield is mainly known for his intricate rhythm playing, but occasionally performs lead guitar duties and some solos, both live and in the studio. Hetfield co-founded Metallica in October 1981 after answering a classified advertisement by drummer **Lars Ulrich** in the *Los Angeles* newspaper *The Recycler*. Metallica has won nine **Grammy Awards** and released nine studio albums, three live albums, four extended plays and 24 singles.

In 2009, Hetfield was ranked at no. 8 in **Joel McIver**'s book *The 100 Greatest Metal Guitarists*,^[1] and ranked at no. 24 by *Hit Parader* on their list of the 100 Greatest Metal Vocalists of All Time. In *Guitar World*'s poll, Hetfield was placed as the 19th greatest guitarist of all time, as well as being placed second (along with Metallica lead guitarist **Kirk Hammett**) in *The 100 Greatest Metal Guitarists* poll of the same magazine. *Rolling Stone* placed Hetfield as the 87th greatest guitarist of all time.

Contents

- 1 Early life
- 2 Career
- 3 Personal life
- 4 Equipment
 - 4.1 Guitars
 - 4.2 Amplifiers and cabinets
 - 4.3 Effects
 - 4.4 Accessories
- 5 Discography
 - 5.1 Metallica
- 6 Guest appearances

James Hetfield



Hetfield performing with Metallica during Rock in Rio USA in Las Vegas, Nevada on May 9, 2015

Background information

Birth name	James Alan Hetfield
Born	August 3, 1963 (age 52) <div>Downey, California, U.S.</div>
Genres	Heavy metal, thrash metal, hard rock, speed metal
Occupation(s)	Musician, songwriter, producer



Wikipedia

In other projects

Wikimedia
Commons
Wikiquote

Languages



/English
 العربية
 Asturianu
 বাংলা
 Беларуская
 Български
 Bosanski
 Català
 Čeština
 Dansk
 Deutsch
 Eesti
 Ελληνικά
 Español
 فارسی
 Føroyskt
 Français
 Galego
 한국어
 Հայերեն
 Hrvatski
 Bahasa Indonesia
 Íslenska
 Italiano
 עברית
 ქართული
 Қазақша
 Лезги
 Latina
 Latviešu
 Lëtzebuergesch
 Lietuvių

6 Guest appearances

7 References

8 External links

Early life

James Alan Hetfield was born on August 3, 1963,^[2] in Downey, California, the son of Cynthia Bassett (née Nourse), a light opera singer, and Virgil Lee Hetfield, a truck driver.^[3] He is of German, English, Irish and Scottish descent.^[citation needed] He has two older half-brothers from his mother's first marriage and one younger sister. His parents divorced in 1976 when Hetfield was 13. Virgil and Cynthia were very strict **Christian Scientists**, and in accordance with their beliefs, Hetfield's parents strongly disapproved of medicine or any other medical treatment and remained loyal to their faith, even as Cynthia was dying from cancer. This upbringing became the inspiration for many of Hetfield's lyrics during his career with Metallica, most notably in the song **"The God That Failed"** from their eponymous 1991 album, *Metallica*. Cynthia Hetfield died of cancer in 1979 when James was 16 years old. After the death of his mother, Hetfield went to live with his older half-brother David. Virgil died in late 1996, during Metallica's *Load* tour.^[4] Hetfield attended **Downey High School** his freshman and sophomore years.^[5]

Hetfield was nine years old when he first began piano lessons, after which he took on his brother David's drums and finally, at the age of fourteen, he began to play guitar with Robert Okner.^[6] He was also in a few bands as a teenager – one being **Leather Charm** and another, Obsession. Hetfield identifies **Aerosmith** as having been his main musical influence as a child,^[7] and has said they were the reason he wanted to play guitar. He has also cited **Queen**, **Led Zeppelin**, **Black Sabbath**, **Misfits**, **AC/DC**, **Cheap Trick**, **Rush**, **Alice Cooper**, **Kiss**, **Ramones**, **Motörhead**, **Sex Pistols**, **Venom**, **Iron Maiden**, **Judas Priest**, **ZZ Top**, **Thin Lizzy**, **Lynyrd Skynyrd**, **Ted Nugent**, **Rainbow**, and **Deep Purple** as influences.^[citation needed]

Career

In the early days of the band, Metallica experimented with a few different vocals and guitar combinations, essentially creating a setup similar to that of British metal band **Diamond Head**, another major influence on Hetfield. Some of the options considered included adding another guitar player, having John Roads play lead guitar, as well as asking **John Bush** from **Armored Saint** (who later joined **Anthrax**) to sing for the band. The finalized line-up of the band became Hetfield (lead vocals and rhythm guitar), **Lars Ulrich** (drums), **Dave Mustaine** (lead guitar), and **Ron McGovney** (bass) who was soon replaced by **Cliff Burton**. Hetfield referred to their early sound as **power metal**. The term "thrash metal" was first used when *Kerrang* journalist **Malcolm Dome** described the Anthrax song **"Metal Thrashing Mad"** in an issue of *Kerrang* in February 1984.^[8]

From 1982 to 1983, Mustaine's alcoholism sparked heated altercations between himself and Hetfield.^[9] Mustaine also once poured beer onto McGovney's bass^[9] nearly causing serious damage. On April 1, 1983, the band recruited lead guitarist **Kirk Hammett** from the band **Exodus** and 10 days later Hetfield and Ulrich officially removed

Occupation(s) Musician, songwriter, producer

Instruments Guitar, vocals

Years active 1978–present

Labels Warner Bros., Elektra, Megaforce

Associated acts Metallica, Spastik Children, Leather Charm

Website metallica.com

Notable instruments

Gibson Explorer
 ESP Truckster
 Ken Lawrence custom models
 Gibson Les Paul Custom "Iron Cross"
 Gibson Flying V
 Fender Telecaster
 ESP Snakebyte



Wikipedia

Discography

Metallica

Main article: [Metallica discography](#)

- *Kill 'Em All* (1983)
- *Ride the Lightning* (1984)
- *Master of Puppets* (1986)
- *...And Justice For All* (1988)
- *Metallica* (1991)
- *Load* (1996)
- *ReLoad* (1997)
- *St. Anger* (2003)
- *Death Magnetic* (2008)

Guest appearances

- Hetfield sang "Stone Cold Crazy" at the [Freddie Mercury Tribute Concert](#), accompanied by the remaining members of [Queen](#) and [Tony Iommi](#) on rhythm guitar.
- He provided backing vocals on "Man or Ash" on the [Corrosion of Conformity](#) album *Wiseblood*.
- He sang backing vocals on "Twist of Cain" and "Possession" on [Danzig's](#) debut album, *Danzig*.
- He played guitar on "Eclectic Electric" from the [Primus](#) album *Antipop*.
- Though previously believed to have been sung by [Trey Parker](#), the song "Hell Isn't Good" from the film *South Park: Bigger, Longer and Uncut* was actually sung by Hetfield. He admitted to doing so in 2000, confirming what many Hetfield fans had already recognized as his trademark growling style.[The Playboy Interview: Metallica \(April 2001\)](#)[ⓘ].
- Hetfield sang on the [Gov't Mule](#) song "Drivin' Rain" with fellow guest [Les Claypool](#), which appears on the [compilation album](#) *Crank It Up with NASCAR*.
- He also performed at the Outlaws Concert in 2004, alongside [Hank Williams Jr.](#), [Cowboy Troy](#), [Big & Rich](#), [Gretchen Wilson](#), and [Kid Rock](#). Hetfield played one of close friend [Waylon Jennings's](#) songs, "Don't You Think This Outlaw Bit's Done Got Out of Hand," the very same he recorded on the tribute album to Jennings, *I've Always Been Crazy*. At this same event, he performed Metallica's "Mama Said" with [Jesse Colter](#), the widow of [Waylon Jennings](#), though this was cut from the television broadcast.
- Hetfield, as well as Metallica's lead guitarist [Kirk Hammett](#), voiced characters in a number of episodes of the [Adult Swim](#) cartoon *Metalocalypse*.
- Hetfield, along with Metallica drummer [Lars Ulrich](#), voiced a pair of teenage dragons in an episode of [Disney's](#) *Dave the Barbarian*, entitled "Here There Be Dragons".

Wikipedia

References

1. ^ "I'm Over My Metallica Demons,' Says Dave Mustaine". *Classic Rock*.
2. ^ "James Hetfield Biography, Bio, Profile, pictures, photos from". *Netlimse.com*. Retrieved 2011-11-04.
3. ^ "Family Tree Maker's Genealogy Site: Genealogy Report: Ancestors of James Alan Hetfield". *Familytreemaker.genealogy.com*. 1963-08-03. Retrieved 2014-07-29.
4. ^ *Playboy Interview*. *likethat.com*. Retrieved on 2010-11-05.
5. ^ "Exclusive: Metallica's James Hetfield Inducted Into Downey High Hall Of Fame « CBS Los Angeles". *Losangeles.cbslocal.com*. 2011-05-27. Retrieved 2014-07-29.
6. ^ "Biography". *Web.archive.org*. October 27, 2009. Archived from the original on October 27, 2009. Retrieved 2010-11-11.
7. ^ Search Articles, Artists, Reviews, Videos, Music and Movies. *Rolling Stone*. Retrieved on 2010-11-05.
8. ^ Dome, Malcolm (February 23, 1984). "Anthrax 'Fistful of Metal'", *Kerrang!*. London, UK: Spotlight Publications Ltd. p. 8.
9. ^ a b c "Behind the Music: Megadeth". *Behind the Music*. 2001. VH1.
10. ^ Metallica.com. Metallica.com. Retrieved on 2010-11-05.
11. ^ Peisner, David (February 2007). "Rock Stars Who've Caught Fire Onstage!". *Blender Magazine Online*. Retrieved 2007-02-27.
12. ^ Music Fan Qubs Metallica - Metallica > The Band > James Hetfield. *Musicfanclubs.org* (August 3, 1963). Retrieved on 2010-11-05.
13. ^ a b Metallica.com. Metallica.com (January 25, 2005). Retrieved on 2010-11-05.
14. ^ Metallica - "My Apocalypse" Review - Music Review at IGN. *Uk.music.ign.com* (August 28, 2008). Retrieved on 2010-11-05.
15. ^ "Metallica: Inducted in 2009". *Rock And Roll Hall of Fame*. February 2007. Retrieved 2010-02-24.
16. ^ a b Metallica.com. Metallica.com (August 3, 1963). Retrieved on 2010-11-05.
17. ^ "IMDb.com - Biography for James Hetfield". Retrieved 2013-05-01.
18. ^ "Metallica News". *blabbermouth.net*. October 23, 2003. Retrieved 2009-08-21.
19. ^ S&M DVD special features interview
20. ^ a b Interview with James Hetfield about his new straight edge tattoo in *So What!* magazine
21. ^ "Metallica's James Hetfield on Rock and Therapy". January 28, 2005.
22. ^ Wood, Mkael. "Metallica: The Billboard Cover Story". *Billboard*.
23. ^ RMM: James Hetfield featured in ABSENTE documentary, exclusive interview with director Justin Hunt. *Rockvillmusicmagazine.com* (August 25, 2010). Retrieved on 2011-04-21.
24. ^ RMM: James Hetfield Interview with Filmmaker Justin Hunt: The Fight For Fatherhood. *Rockvillmusicmagazine.com*. Retrieved on 2011-04-21.
25. ^ a b RMM: James Hetfield to Appear on FOX and Friends with Filmmaker Justin Hunt. *Rockvillmusicmagazine.com* (March 28, 2011). Retrieved on 2011-04-21.
26. ^ Harris, Chris. (August 4, 2008) Metallica Tackle Forgiveness, Resentment In 'The Day That Never Comes' Clip - Music, Celebrity, Artist News. *MTV*. Retrieved on 2011-04-21.
27. ^ "JAMES HETFIELD Says METALLICA Is Focusing on Feeling Around 'Master of Puppets' For New CD - July 14, 2007".
28. ^ Smith, Clive Stafford (2008-05-18). "Welcome to 'the disco'". *The Guardian*. ISSN 0261-3077.
29. ^ "HETFIELD On Use Of METALLICA's Music To Torture Guantanamo Bay Prisoners: 'Part Of Me Is Proud'". *Blabbermouth*.
30. ^ "Metallica Urged U.S. to Stop Using Their Music for Torture, Says SEAL Who Killed Bin Laden | SPIN". *Spin*. Retrieved 2016-04-25.
31. ^ "METALLICA - Encyclopedia Metallica - Projects - James Hetfield". *Encycrmet.com*. Retrieved 2014-07-29.

External links

- James Hetfield** at the *Internet Movie Database*
- NPR Interview With Metallica Guitarist and Vocalist James Hetfield (2004, audio)



Wikiquote has quotations related to: **James Hetfield**



Wikipedia

	'I Disappear' • "St. Anger" • "Frantic" • "The Unnamed Feeling" • "Some Kind of Monster" • "The Day That Never Comes" • "My Apocalypse" • "Cyanide" • "The Judas Kiss" • "All Nightmare Long" • "Broken, Beat & Scarred" • "The View"
Promotional singles	"Fade to Black" • "For Whom the Bell Tolls" • "Don't Tread on Me" • "Ain't My Bitch" • "Bleeding Me" • "Better than You"
Other songs	"Seek & Destroy" • "Am I Evil?" • "Blitzkrieg" • "Breadfan" • "The God That Failed" • "Astronomy" • "Tuesday's Gone" • "Last Caress/Green Hell" • "Stone Cold Crazy" • "So What?" • "Overkill" • "The Unforgiven II"
Video albums	<i>Cliff 'Em All</i> • <i>2 of One</i> • <i>A Year and a Half in the Life of Metallica</i> • <i>For Those About to Rock</i> • <i>Cunning Stunts</i> • <i>S&M</i> • <i>Classic Albums: Metallica - Metallica</i> • <i>Some Kind of Monster</i> • <i>The Videos 1989–2004</i> • <i>Français Pour une Nuit</i> • <i>Orgullo, Pasión, y Gloria: Tres Noches en la Ciudad de México</i> • <i>The Big Four: Live from Sofia, Bulgaria</i> (with Slayer, Megadeth, and Anthrax) • <i>Quebec Magnetic</i> • <i>Metallica: Through the Never</i>
Tours	Kill 'Em All Tour • Ride the Lightning Tour • Damage, Inc. Tour • Monsters of Rock Tour 1987 • Monsters of Rock Tour 1988 • Damaged Justice Tour • Wherever We May Roam Tour • Guns N' Roses/Metallica Stadium Tour • Nowhere Else to Roam Tour • Shit Hits the Sheds Tour • Poor Touring Me • Poor Re-Touring Me Tour • Garage Remains the Same Tour • M2K Mini Tour • Summer Sanitarium Tour • Summer Sanitarium 2003 Tour • Madly in Anger with the World Tour • Escape from the Studio '06 • Sick of the Studio '07 • 2008 European Vacation Tour • World Magnetic Tour • 2012 European Black Album Tour
Related articles	Awards • Discography • Demos • Songs • Apocalyptica • Beatallica • Big Mick • Rick Rubin • Orion Music + More • Echobrain • Exodus • Flotsam and Jetsam • Metallica v. Napster, Inc. • Megadeth • Bob Rock • A Tribute to the Four Horsemen • The Blackest Album: An Industrial Tribute to Metallica • Guitar Hero: Metallica (song) • <i>Metallic Assault: A Tribute to Metallica</i> • <i>Metallic Attack: The Ultimate Tribute</i> • Newsted • Dane DeHaan • Flemming Rasmussen
	 Book: Metallica • Category: Metallica • Category: Metallica audio samples • Portal: Heavy metal
Authority control	WorldCat Identifies • VIAF: 84962297 • LCCN: nb2002016265 • ISN: 0000 0000 7884 8854 • GND: 136212735 • BNF: cb14048890g (data) • MusicBrainz: 68c34082-c1d5-4399-af9b-3921ca0a6a1f

Categories: [1963 births](#) | [20th-century American singers](#) | [21st-century American singers](#) | [American baritones](#) | [American heavy metal guitarists](#) | [American heavy metal singers](#) | [American male singer-songwriters](#) | [American people of English descent](#) | [American people of German descent](#) | [American people of Irish descent](#) | [American people of Scottish descent](#) | [American rock singers](#) | [American rock songwriters](#) | [American male songwriters](#) | [American skateboarders](#) | [Critics of Christian Science](#) | [Former Christian Scientists](#) | [Living people](#) | [Metallica members](#) | [Musicians from Downey, California](#) | [People from Brea, California](#) | [Rhythm guitarists](#) | [Rock and Roll Hall of Fame inductees](#) | [Singers from California](#) | [Songwriters from California](#) | [American hunters](#) | [American fishers](#) | [National Rifle Association people](#)

This page was last modified on 6 June 2016, at 01:06.

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.

[Privacy policy](#) [About Wikipedia](#) [Disclaimers](#) [Contact Wikipedia](#) [Developers](#) [Cookie statement](#) [Mobile view](#)



Wikidata

Collaboratively edited knowledge base

- intended to provide **common source of structured data** for other related projects (e.g., Wikipedia)
- **17 M data items**, 345 M edits
- started in 2012

Wikidata



[Main page](#)
[Community portal](#)
[Project chat](#)
[Create a new item](#)
[Item by title](#)
[Recent changes](#)
[Random item](#)
[Query Service](#)
[Nearby](#)
[Help](#)
[Donate](#)

[Print/export](#)
[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Tools
[What links here](#)
[Related changes](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Concept URI](#)
[Cite this page](#)

[English](#) [Not logged in](#) [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Item [Discussion](#)

Read

[View history](#)

Search



James Hetfield (Q484302)

American musician, songwriter, record producer and table

[edit](#)

James Alan Hetfield

[In more languages](#) [Configure](#)

Language	Label	Description	Also known as
English	James Hetfield	American musician, songwriter, record producer and table	James Alan Hetfield
German	James Hetfield	US-amerikanischer Sänger und Gitarrist	Hetfield James Alan Hetfield
French	James Hetfield	chanteur et guitariste de hard-rock	
Bavarian	No label defined	No description defined	

[More languages](#)

Statements

[sex or gender](#)



male

[edit](#)

[5 references](#)















[+ add](#)



Wikidata












[What links here](#)
[Related changes](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Concept URI](#)
[Cite this page](#)

Statements

sex or gender	 male 
	<div> <div>» 5 references</div> <div>+ add</div> </div>
Commons category	 James Hetfield 
	<div> <div>» 0 references</div> <div>+ add reference</div> </div> <div>+ add</div>
occupation	 guitarist 
	<div> <div>» 1 reference</div> </div>
	 singer-songwriter 
	<div> <div>» 1 reference</div> </div>
	 singer 
	<div> <div>» 1 reference</div> </div>
	 composer 
	<div> <div>» 1 reference</div> </div>
	 musician 








Wikidata











ISNI	 0000 0000 7884 8854   edit ▶ 1 reference + add
place of birth	 Downey  edit ▶ 1 reference + add
member of	 Metallica  edit ▼ 0 references + add reference + add
date of birth	 3 August 1963  edit ▶ 2 references + add
instance of	 human  edit ▶ 2 references + add



Wikidata

image	 edit 1 reference + add
Commons gallery	 edit 1 reference + add
voice type	 edit 0 references + add reference + add
given name	 edit 0 references + add reference + add
instrument	 edit 0 references + add reference

Wikidata

official website	 http://www.metallica.com/  edit
	▶ 1 reference
	+ add
	+ add
Identifiers	
VIAF ID	 84962297  edit
	▶ 2 references
	+ add
LCAuth ID	 nb2002016265  edit
	▶ 1 reference
	+ add
GND ID	 136212735  edit
	▶ 1 reference
	+ add
MusicBrainz artist ID	 68c34082-c1d5-4399-af9b-3921ca0a6a1f  edit

Wikidata

Wikipedia (61 entries) [edit](#)

ang	James Hetfield
ar	جيمس هتفيلد
ast	James Hetfield
be	Джэймс Алан Хэтфілд
bg	Джеймс Хетфийлд
bn	জেমস হেটফিল্ড
bs	James Hetfield
ca	James Hetfield
cs	James Hetfield
da	James Hetfield
de	James Hetfield
el	Τζέιμς Χέτφιλντ
en	James Hetfield
es	James Hetfield
et	James Hetfield
fa	جيمس هتفيلد
fi	James Hetfield
fo	James Hetfield
fr	James Hetfield
gl	James Hetfield
he	דזשימס העטפילד
hr	James Hetfield
hu	James Hetfield
hy	Ջեյմս Հեթֆիլդ
id	James Hetfield
is	James Hetfield
it	James Hetfield

^

Wikibooks (0 entries) [edit](#)

Wikinews (0 entries) [edit](#)

Wikiquote (10 entries) [edit](#)

az	Ceyms Hetfild
bg	Джеймс Хетфийлд
en	James Hetfield
es	James Hetfield
it	James Hetfield
pl	James Hetfield
pt	James Hetfield
ru	Джеймс Хэтфилд
tr	James Hetfield
uk	Джеймс Гетфілд

^

Wikisource (0 entries) [edit](#)

Wikiversity (0 entries) [edit](#)

Wikivoyage (0 entries) [edit](#)

Other sites (1 entry) [edit](#)



DBPedia

Knowledge base **extracted from Wikipedia**

- **4.6 M things** (for English, as of 2014)
- 80 M links to Wikipedia categories
- 41 M links to YAGO categories
- **50 M owl:sameAs links** to other data sources
- started in 2007

DBPedia


[Browse using](#)
[Formats](#)
[Faceted Browser](#)
[Sparql Endpoint](#)

About: James Hetfield

An Entity of Type : [musical artist](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)

James Alan Hetfield (born August 3, 1963) is the co-founder, lead vocalist, rhythm guitarist, main songwriter, and lyricist for the American heavy metal band Metallica. Hetfield is mainly known for his intricate rhythm playing, but has also performed occasional lead guitar duties both in the studio and live. Hetfield co-founded Metallica in October 1981 after answering a classified advertisement by drummer Lars Ulrich in the Los Angeles newspaper The Recycler.

Property	Value
dbo:abstract	<ul style="list-style-type: none"> James Alan Hetfield (born August 3, 1963) is the co-founder, lead vocalist, rhythm guitarist, main songwriter, and lyricist for the American heavy metal band Metallica. Hetfield is mainly known for his intricate rhythm playing, but has also performed occasional lead guitar duties both in the studio and live. Hetfield co-founded Metallica in October 1981 after answering a classified advertisement by drummer Lars Ulrich in the Los Angeles newspaper The Recycler. Metallica has won nine Grammy Awards and released nine studio albums, three live albums, four extended plays and 24 singles. In 2009, Hetfield was ranked number 8 in Joel McIver's book The 100 Greatest Metal Guitarists, and ranked 24th by Hit Parader on their list of the 100 Greatest Metal Vocalists of All Time. In Guitar World's poll, Hetfield was placed as the 19th greatest guitarist of all time, as well as being placed 2nd (along with Kirk Hammett) in The 100 Greatest Metal Guitarists poll of the same magazine, only behind Tony Iommi. Rolling Stone placed Hetfield as the 87th greatest guitarist of all time. ^(en)
dbo:activeYearsStartYear	<ul style="list-style-type: none"> 1978-01-01 ^(xsd:date)
dbo:associatedBand	<ul style="list-style-type: none"> dbr:Spastik_Children dbr:Leather_Charm dbr:Metallica



DBpedia


[Browse using](#)
[Formats](#)
[Faceted Browser](#)
[Sparql Endpoint](#)

rdf:type

- yago:AmericanBaritones
- yago:AmericanHeavyMetalGuitarists
- yago:AmericanHeavyMetalSingers
- yago:AmericanMaleSingers
- yago:AmericanMusiciansOfGermanDescent
- yago:AmericanMusiciansOfIrishDescent
- yago:AmericanPeopleOfEnglishDescent
- yago:AmericanPeopleOfScottishDescent
- yago:AmericanSkateboarders
- yago:Artist109812338
- yago:Athlete109820263
- yago:Baritone109839436
- yago:CausalAgent100007347
- yago:Contestant109613191
- yago:Creator109614315
- yago:Entertainer109616922
- yago:Guitarist110151760
- yago:LivingThing100004258
- yago:Musician110339966
- yago:Musician110340312
- yago:Object100002684
- yago:Organism100004475
- yago:Performer110415638
- yago:Person100007846
- yago:ReligiousPerson109628382
- yago:Singer110599806



DBpedia


[Browse using ▾](#)
[Formats ▾](#)
[Faceted Browser](#)
[Sparql Endpoint](#)
[rdfs:comment](#)

■ James Alan Hetfield (born August 3, 1963) is the co-founder, lead vocalist, rhythm guitarist, main songwriter, and lyricist for the American heavy metal band Metallica. Hetfield is mainly known for his intricate rhythm playing, but has also performed occasional lead guitar duties both in the studio and live. Hetfield co-founded Metallica in October 1981 after answering a classified advertisement by drummer Lars Ulrich in the Los Angeles newspaper The Recycler. ^(en)

[rdfs:label](#)

■ James Hetfield ^(en)

[owl:differentFrom](#)

■ [dbr:James_Hatfield](#)

[owl:sameAs](#)

- <http://zitgist.com/music/artist/68c34082-c1d5-4399-af9b-3921ca0a6a1f>
- [freebase:James Hetfield](#)
- [dbpedia-fr:James Hetfield](#)
- [dbpedia-es:James Hetfield](#)
- [dbpedia-de:James Hetfield](#)
- [dbpedia-el:James Hetfield](#)
- [dbpedia-es:James Hetfield](#)
- [dbpedia-id:James Hetfield](#)
- [dbpedia-it:James Hetfield](#)
- [dbpedia-ja:James Hetfield](#)
- [dbpedia-ko:James Hetfield](#)
- [dbpedia-nl:James Hetfield](#)
- [dbpedia-pl:James Hetfield](#)
- [dbpedia-pt:James Hetfield](#)
- [dbpedia-wikidata:James Hetfield](#)
- [openccy:James Hetfield](#)
- [wikidata:James Hetfield](#)



Freebase

Collaboratively edited knowledge base

- which includes information **extracted from other data sources** (e.g., Wikipedia, MusicBrainz)
- **44 M topics** (organized into domains, as of 2014)
- **2.4 B facts**
- more than 60 M deleted facts (as of 2013)
- focused on entertainment
- many topics without textual data
- started in 2007, acquired by Google in 2010
- forms the **core of the Google Knowledge Graph**
- now: **freebase data “moved” to wikidata** [4]

attention

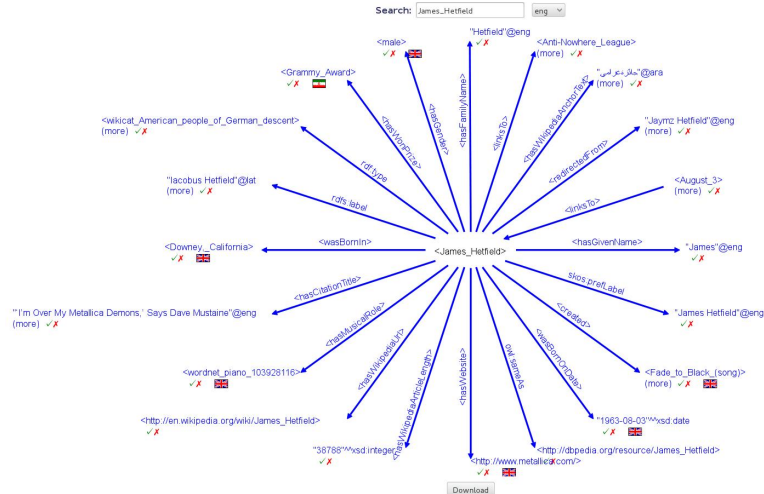
resources evolve, change, disappear

Yago

Knowledge base extracted from **Wikipedia, WordNet, and GeoNames**

- 10 M named entities
- **120 M facts**
- 72 relation types with **manually assessed accuracy (95%)**
- started in 2007

Svg Browse YAG03



From <http://yago-knowledge.org>
Give us feedback:
✓ = fact is correct
✗ = fact is false

Entity Linking

Entity linking (aka. “Wikification”):

- process of spotting mentions of named entities in documents, resolving their ambiguity, and linking them to known identifiers

Entity Linking

Entity linking (aka. “Wikification”):

- process of spotting mentions of named entities in documents, resolving their ambiguity, and linking them to known identifiers

Page played his Gibson on Kashmir.

Entity Linking

Entity linking (aka. “Wikification”):

- process of spotting mentions of named entities in documents, resolving their ambiguity, and linking them to known identifiers

Page played his Gibson on Kashmir.

Entity Linking

Entity linking (aka. “Wikification”):

- process of spotting mentions of named entities in documents, resolving their ambiguity, and linking them to known identifiers



Page played his Gibson on Kashmir.



Entity Linking

Entity linking (aka. “Wikification”):

- process of spotting mentions of named entities in documents, resolving their ambiguity, and linking them to known identifiers



Page played his Gibson on Kashmir.



Outline

- 1 Semantic Web
- 2 Knowledge Bases
- 3 Entity Linking**
- 4 Semantic Search

Remember: Named Entities

entity

anything you can refer to with a name

- location, person, organization
- facilities, vehicles, songs, movies, products
(and domain-dependent ones: genes & proteins, ...)
- sometimes: numbers, dates

relevant in IR

entities are popular and extremely frequent in queries

names are highly ambiguous

- Washington → place(s), person(s), (government)
- Springfield

Remember: Named Entity Recognition

tasks

- extraction → determine the boundaries
- classification → assign class (PER, LOC, ORG, ...)

systems

- rule-based → with gazetteers, context-based rules (Mr.), ...
- machine learning → features: mixed case (eBay), ends in digit (A9), all caps (BMW), ...
- several tools available (e.g., Stanford NER)

extraction is good, but normalization is better

Remember: Named Entity Normalization

same task, many names

- normalization
- linking
- resolution
- grounding

example: Washington

- /wiki/Washington,_D.C.
- /wiki/Washington_%28state%29
- /wiki/Washington_Irving
- /wiki/Washington_Redskins
- /wiki/George_Washington

tools

- several tools available (AIDA, ...)

Assignment

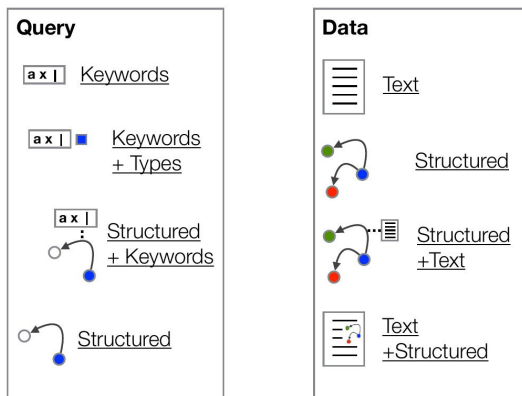
just **recognition**, no normalization

Outline

- 1 Semantic Web
- 2 Knowledge Bases
- 3 Entity Linking
- 4 Semantic Search**

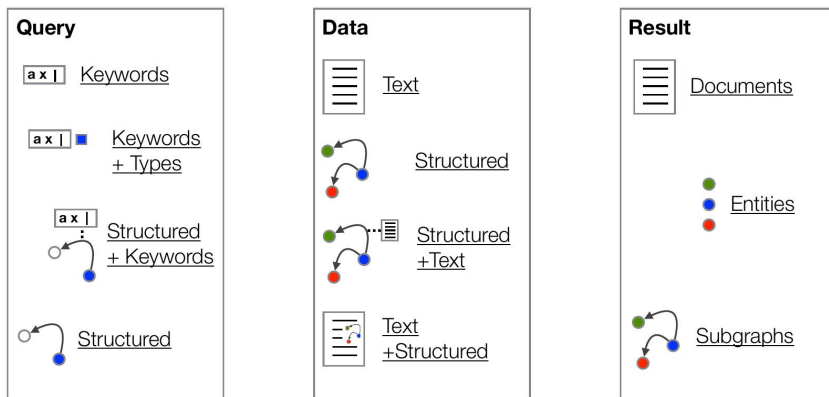
Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Tranh and Mika [5] for a comprehensive overview)



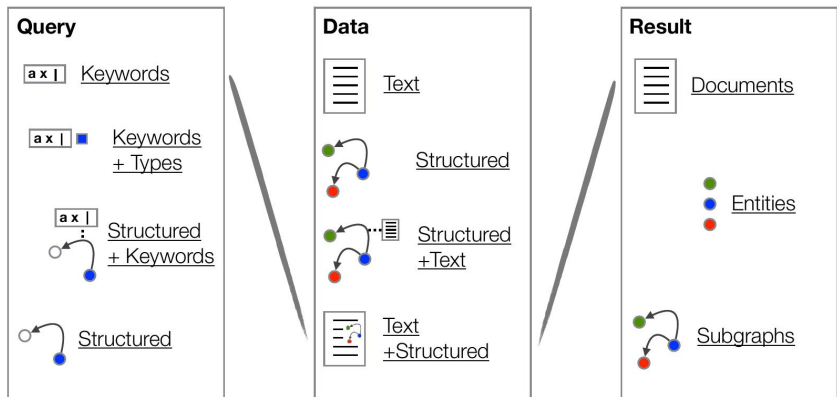
Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Tranh and Mika [5] for a comprehensive overview)



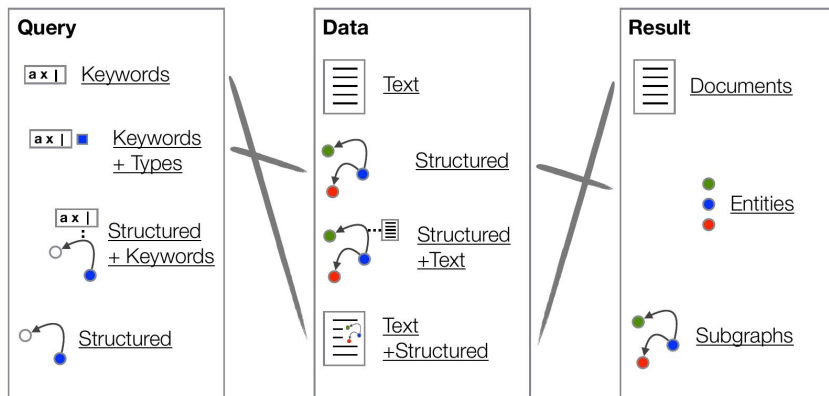
Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Trinh and Mika [5] for a comprehensive overview)



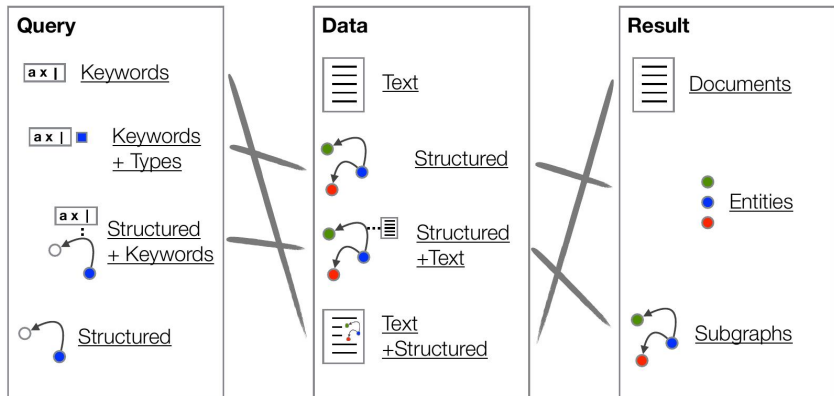
Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Tranh and Mika [5] for a comprehensive overview)



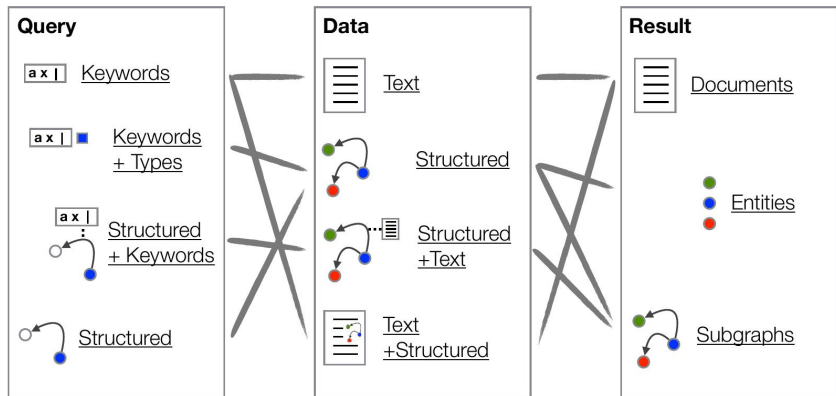
Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Tranh and Mika [5] for a comprehensive overview)



Semantic Search

- **Semantic search** includes a **wide variety of approaches** (see Tranh and Mika [5] for a comprehensive overview)



Semantic Search Examples

- **(More) Semantic Keyword Search** improves **ad-hoc information retrieval** by leveraging knowledge bases (input: keywords, output: list of documents)
- **Entity Search** retrieves **entities** based on **keyword queries** which may include explicit/implicit **cues about the target type** (input: keywords + types, output: list of entities)
- **Knowledge Search** retrieves **subgraphs** as answers to **structured queries** which may be **augmented with keywords** (input: structured + keywords, output: list of subgraphs)

Semantic Keyword Search

- **Synonymy** (e.g., car/automobile) and **polysemy** (e.g., jaguar as cat, car, guitar): **negative effects** on retrieval effectiveness
- Concept-based IR represents **queries and documents in terms of (semantic) concepts** instead of terms
 - **WordNet** synsets after **word sense disambiguation**
 - **Latent concepts** identified based on word co-occurrences
- **Explicit Semantic Analysis** (ESA) [3]: words (e.g., jaguar) in terms of the Wikipedia concepts that they describe (e.g., JAGUAR, JAGUAR_CARS, JAGUAR_XK, FENDER_JAGUAR)

do you remember last week?

Why Novelty & Diversity?

- ▶ **Ambiguity of query** needs to be reflected in the returned results to account for **uncertainty about the user's information need**



- ▶ **Query ambiguity** comes in **different forms**
 - ▶ **topic** (e.g., **jaguar**, **eclipse**, **defender**, **cookies**)
 - ▶ **intent** (e.g., **java 8** – **download** (transactional), **features** (informational))
 - ▶ **time** (e.g., **olympic games** – 2012, 2014, 2016)

- ▶ **Diversity:** How well do returned results reflect query ambiguity?

Implicit vs. Explicit Diversification

- ▶ **Implicit diversification methods** do not represent query aspects explicitly and instead operate directly on **document contents and their (dis)similarity**
 - ▶ Maximum Marginal Relevance [3]
 - ▶ BIR (Beyond Independent Relevance) [11]
- ▶ **Explicit diversification methods** represent query aspects explicitly (e.g., as categories, subqueries, or key phrases) and consider **which query aspects individual documents relate to**
 - ▶ IA-Diversify [1]
 - ▶ xQuad [10]
 - ▶ PM [7,8]

5.4.2. eXplicit Query Aspect Diversification

- ▶ Santos et al. [10] use **query suggestions** from a web search engine as **query aspects**
- ▶ **Greedy algorithm**, inspired by IA-SELECT, iteratively builds up a set **S** of size **k** by selecting document having highest probability

$$(1 - \lambda) P[d | q] + \lambda P[d, \neg S | q]$$

where $P[d|q]$ is the document likelihood and captures **relevance** and $P[d, \neg S|q]$ is the probability that **d** covers a query aspect not yet covered by documents in **S** and captures **diversity**

Searches related to jaguar

jaguar xj	jaguar animal
audi	jaguar price
jaguar xf	jaguar fittings
jaguar mining	jaguar india

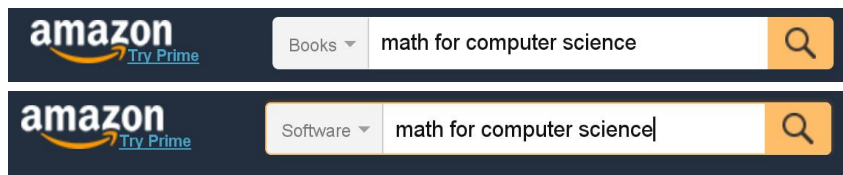
jaguar

- jaguar
- jaguar xe
- jaguar de
- jaguar f type
- jaguar xf
- jaguar xe 2015
- jaguar forum
- jaguar e type

Entity Oriented Search

Entity Search retrieves **entities** based on **keyword queries**

- queries may include explicit/implicit **cues about target type** (input: keywords + types, output: list of entities)
- Queries for specific entities are **common in web search** (e.g., museums with van goghs, bayern munich players, etc.)
- Product search with type cues are **common in e-commerce**



LM for Entity Oriented Search

Balog et al. [1] develop a **language modeling approach**, assuming that entities come with ready-to-use representation:

- a **textual description** (Wikipedia article, IMDB profile, etc.)
- a **set of categories** with meaningful names (e.g., American film)

The screenshot shows the Wikipedia article for "25th Hour". The title is "25th Hour". The text describes the film as a 2002 American drama film directed by Spike Lee and starring Edward Norton. It is based on the novel "The 25th Hour" by David Benioff, who also wrote the screenplay. It tells the story of a man's last 24 hours of freedom as he prepares to go to prison for seven years for shooting drugs. The article mentions positive reviews and several critics have named it one of the best films of the decade.

Contents (hide)

- Plot
- Cast
- Production
- Reception
- Music
- On popular culture
- See also
- References
- External links

Plot

A century yellow vintage Super Bee pulls up short on a New York City street, and Montgomery "Monty" Brogan gets out with his buddy Koryto to look at a dog lying in the road. The animal was mauling in a daylight, so Monty intends to shoot him, but changes his mind after he looks it in the eye. Monty decides to take the dog for a nearby clinic instead.

Fast forward to late 2002. Monty is about to begin serving a seven-year prison sentence for shooting drugs. He sits in a park with Doyle, the dog he rescued, on his last day of freedom. He plans to meet childhood friends Frank (Christopher) and Jack (Ensign) at a club with his girlfriend Natashia. Frank is a lobbyist trade on

The screenshot shows the Amazon product page for the "Wenger 16999 Swiss Army Knife Giant". The product is currently unavailable. The page features a large image of the knife, which is a multi-tool with various functions. The text indicates that the knife is currently unavailable and that the seller does not know when or if it will be back in stock. The page also shows customer reviews and a "Currently unavailable" message. The product is described as a "Giant" and is made of stainless steel. It has a red handle and a silver blade. The product is sold by "Father's Day deals" and is available in the "Tools & Home Improvement" category.

American film, 2002 film,
films about drugs, ...

tools, pocket knives, power
and hand tools, ...

LM for Entity Oriented Search

Query $q = (T, C)$ consists of **textual terms** T and **categories** C

- Estimate **language models** θ_q^T and θ_q^C for **query components**
- Estimate **language models** θ_e^T and θ_e^C for **entity components**
- Rank entities for query q according to **query likelihood**

$$P[q|e] = \lambda \cdot P[\theta_q^T | \theta_e^T] + (1 - \lambda) \cdot P[\theta_q^C | \theta_e^C]$$

with probabilities estimated from **Kullback-Leibler divergences**

LM for Entity Oriented Search

Heuristic estimation of probabilities $P[\theta_q|\theta_e]$ from Kullback-Leibler divergences to make them comparable and combinable

$$P[\theta_q|\theta_e] = \frac{\max_{e^*} KL(\theta_q||\theta_{e^*}) - KL(\theta_q||\theta_e)}{\sum_{e'} \max_{e^*} KL(\theta_q||\theta_{e^*}) - KL(\theta_q||\theta_{e'})}$$

Language modeling framework with **various instantiations** evaluated and compared, e.g.,

- estimate θ_q^T from T and C but ignore θ_q^C (M2)
- estimate both θ_q^T and θ_q^C from T (M3)
- estimate θ_q^T only from T and θ_q^C only from C (M4)
- estimate θ_q^T from T and C and θ_q^C only from C (M5)

LM for Entity Oriented Search

- dataset: INEX entity ranking: given a query, return entities
- Example: Query $T = \{\text{Paul Auster novels}\}$ $C = \{\text{NOVELS}\}$

what performs better M3 or M4?

Table VI. The Top 10 Ranked Entities for Topic #91

rank	M3 qT-Q	M4 qT-QC
1	The New York Trilogy	In the Country of Last Things
2	Moon Palace	The Winthrop Woman (novel)
3	Brooklyn Follies	Extremely Loud and Incredibly Close
4	In the Country of Last Things	Novel sequence
5	The Book of Illusions	The City and the Pillar
6	Oracle Night	Going After Cacciato
7	The Music of Chance	Raj Quartet
8	Paul Auster	Paul Clifford
9	Fanshawe (novel)	The Age of Reason (Sartre)
10	French literature of the 19th century	The Notebooks of Malte Laurids Brigge

Boldface Indicates Relevant Entities.

LM for Entity Oriented Search

Table VI. The Top 10 Ranked Entities for Topic #91

rank	M3 qT-Q	M4 qT-QC
1	The New York Trilogy	In the Country of Last Things
2	Moon Palace	The Winthrop Woman (novel)
3	Brooklyn Follies	Extremely Loud and Incredibly Close
4	In the Country of Last Things	Novel sequence
5	The Book of Illusions	The City and the Pillar
6	Oracle Night	Going After Cacciato
7	The Music of Chance	Raj Quartet
8	Paul Auster	Paul Clifford
9	Fanshawe (novel)	The Age of Reason (Sartre)
10	French literature of the 19th century	The Notebooks of Malte Laurids Brigge

Boldface Indicates Relevant Entities.

- (M4) uses a general target category (“novel”)
- “The Winthrop Woman (novel)” Wikipedia page has many categories with “novel”
- one of the categories found for (M3) is “Novels by Paul Auster”
- “The New York Trilogy” Wikipedia page has category “Novels by Paul Auster”

Summary

- **Semantic Web** aims at making the Web interpretable for machines; underlying standards include **RDF** and **SPARQL**
- **Knowledge Bases** provide **world knowledge in structured form** have been harvested from Wikipedia and other data sources
- **Entity Linking** is crucial (if we want to return text documents)
- **Entity Search** retrieves entities based on keyword queries which may include type cues; applications in web search and business

Thank you for your attention!

Summary

References

- 1 K. Balog, M. Bron, M. de Rijke: Query Modeling for Entity Search Based on Terms, Categories, and Examples, ACM TOIS 29(4), 2011
- 2 T. Berners-Lee, J. Hendler, O. Lassila: The Semantic Web, Scientific American, 2001
- 3 O. Egozi, S. Markovitch, E. Gabrilovich: Concept-Based Information Retrieval, Using Explicit Semantic Analysis, ACM TOIS 29(2), 2011
- 4 T. Pellissier Tanon, D. Vrandecic, S. Schaffert, T. Steiner, L. Pintscher: From Freebase to Wikidata: The Great Migration, WWW 2016
- 5 T. Tranh and P. Mika: Semantic Search â Systems, Concepts, Methods, and the Communities behind it, 2013

Summary

Thank you for your attention!

Furthermore:

- U. Sawant and S. Chakrabarti: Learning Joint Query Interpretation and Response Ranking, WWW 2013
- M. Yahya, K. Berberich, S. Elbassuoni, M. Ramanath, V. Tresp, G. Weikum: Natural Language Questions for the Web of Data, EMNLP 2012

Thanks

some slides / examples are taken from / similar to those of:

- Klaus Berberich, Saarland University, previous ATIR lecture