D5 Oberseminar

- Students will present their ongoing M.Sc. theses done @ D5
- When: November 24 from 14:00 to 16:00 (next Monday)
- <u>Where</u>: **E1.4 / R024** (this room)

- <u>You</u>:
 - should have time since there is no tutorial next week
 - get to see what M.Sc. thesis topics look like @ D5
 - are cordially invited to attend

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

									1	
Google	who play	yed in goo	od will hur	nting				Q		
	Web	Images	Videos	News	Shopping	g More –	Search tools		List	
	Good W	Vill Huntii	ng > Ca	ist					Ques	<u>stions</u>
	C	7	1	-	APR-		BAFTA			
		Goo	gle	albums b	y pink flo	byd				Q
	Matt Dam			Web \	/ideos	Images N	lews Shopping	g More - Sear	rch tools	
	Will Hunting			Pink Flo	yd > All	bums				
							PINK FLOYD THE WALL			
				The Endle	t	The Dark Side o he Moon 1973	of The Wall 1979	Wish You Were Here 1975	The Division Bell	Animals 1977

how tall is dirk nowitzki		☆ 🗖	
📟 🖸 🖽 🐬		≡ Examples 🛩 Random	
Input interpretation: Dirk Nowitzki height		Factoi	
Result:		Show non-metric	lons
2.13 meters	Google	how did janis joplin die	Q
Physical characteristics:	0		
height 2.13 meters		Web Images Videos News Shopping More - Sea	arch tools
weight 111 kg (kilogr		About 376,000 results (0.27 seconds)	
Unit conversions:			
6.99 feet		Heroin overdose	Jar
6′11.9″		Janis Joplin, Cause of death	Singer
83.9 inches			Janis I

Google	which t	two guitaris	sts were	married to the same woman م <u>Still Beyon</u> <u>Hope</u>
	Web About 30 Fleetw		videos	Shopping News More ~ Search tools who was u.s. president when martin luther king was shot Q Web News Videos Images Shopping More ~ Search tools About 6,050,000 results (0.45 seconds)
				Assassination of Martin Luther King, Jr. Martin Luther King, Jr. was an American clergyman, activist, and prominent leader of the African- American civil rights movement and Nobel Peace Prize laureate who became known for his advancement of civil rights by using civil disobedience. Assassination of Martin Luther King, Jr Wikipedia, the free en.wikipedia.org/wiki/Assassination_of_Martin_Luther_King,_Jr.
				Feedback
				Assassination of Martin Luther King, Jr Wikipedia, the free

Outline

- 3.1. Semantic Web
- 3.2. Knowledge Bases
- 3.2. (More) Semantic Keyword Search
- 3.3. Entity Search
- 3.4. Knowledge Search

1. Semantic Web

- Semantic Web is an extension of the World Wide Web, envisioned by Berners-Lee et al. [2], which aims at
 - giving **well-defined meaning** to information (in web pages)
 - making the Web interpretable for machines
 - facilitating exchange and reuse of data

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

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/pub/bands?query=FF#albums
 - scheme (e.g., http, ftp, urn) determines interpretation of URI
 - **authority** indicates who is responsible for the 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 the resource

- 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), predicate (URI), and an object (URI or literal)
- Example: Foo Fighters have member Dave Grohl

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 **(0)**



• RDF triples form a **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://allab	outmusic.org
1			a:Dave_Grohl a:Pat_Smear

• RDF triples form a **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://allab	outmusic.org
			a:Dave_Grohl a:Pat_Smear

• RDF triples form a **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://allab	outmusic.org
a:Foo_Figh	ters	a:member	a:Dave_Grohl
a:Foo_Figh	ters	a:member	a:Pat_Smear

SPARQL

- SPARQL Protocol and 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

<u>Query</u>

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

SPARQL

- SPARQL Protocol and 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



SPARQL

- SPARQL Protocol and 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



Linked Open Data Project

- <u>Problem</u>: 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://www.linkeddata.org



2. Knowledge Bases

- Knowledge bases (e.g., DBpedia, Freebase, 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 (e.g., WordNet, GeoNames, MusicBrainz, etc.)
 - form the core of the linked open data cloud

- Collaboratively edited encyclopedia that forms the common basis of other knowledge bases
 - 4.6 M articles / 34.2 M pages / 743.2 M edits (for English)
 - available in 287 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: 1977 births | Living people | People from Utica, New York | Lead guitarists | Slide guitarists | American blues guitarists | American rock guitarists | American rock guitarists | American blues singers | American rock singers | Blues rock musicians | Electric blues musicians | American people of Italian descent | Black Country Communion members | Epic Records artists | Musicians from New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | Resonator guitarists | Blue-eyed soul singers | People from New Hartford, New York | People from New Hartford, New York | People from New Hartford, New York | People from York | Pe







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

Wikimedia Shop

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 Article Talk

Eric Clapton

From Wikipedia, the free encyclopedia

Eric Patrick Clapton, CBE, (born 30 March 1945) is an English musician, singer-songwriter and guitarist. He is the only three-time inductee to the Rock and Roll Hall of Fame: once as a solo artist and separately as a member of the Yardbirds and Cream. Clapton has been referred to as one of the most important and influential guitarists of all time.^[1] Clapton ranked second in *Rolling Stone* magazine's list of the "100 Greatest Guitarists of All Time"^[2] and fourth in Gibson's "Top 50 Guitarists of All Time".^[3] He was also named number five in *Time* magazine's list of "The 10 Best Electric Guitar Players" in 2009 ^[4]

In the mid-1960s, Clapton left the Yardbirds to play blues with John Mayall & the Bluesbreakers. Immediately after leaving Mayall, Clapton joined Cream, a power trio with drummer Ginger Baker and bassist Jack Bruce in which Clapton played sustained blues improvisations and "arty, blues-based psychedelic pop".^[5] For most of the 1970s, Clapton's output bore the influence of the mellow style of JJ Cale and the reggae of Bob Marley. His version of Marley's "I Shot the Sheriff" helped reggae reach a mass market.^[6] Two of his most popular recordings were "Layla", recorded while he was a member of band Derek and the Dominos; and Robert Johnson's "Crossroads", recorded by band Cream. Following the death of his son Conor in 1991, Clapton's grief was expressed in the song "Tears in Heaven", which featured in his *Unplugged* album.

Clapton has been the recipient of 18 Grammy Awards, and the Brit Award for Outstanding Contribution to Music. In 2004, he was awarded a CBE at Buckingham Palace for services to music.^{[7][8][9]} In 1998, Clapton, a recovering alcoholic and drug addict, founded the Crossroads Centre on Antigua, a medical facility for recovering substance abusers.^[10]

Contents [hide]

2 Career 2.1 Early career, breakthrough, and international success 2.1.1 The Yardbirds and the Bluesbreakers Read Edit View history

history Search

Eric Clapton

Create account Log in

Q



Clapton performing at Hyde Park, London in June 2008

Bac	kground information
Birth name	Eric Patrick Clapton
Also known	Slowhand
as	
Born	30 March 1945 (age 69) Ripley, Surrey, England
Genres	Blues rock, rock, psychedelic rock, hard rock, blues
Occupation(s)	Musician, guitarist, singer- songwriter, record producer, artist
Instruments	Guitar, vocals
Years active	1962-present

Advanced Topics in Information Retrieval / Semantic Search

1 Early life

Download as PDF Printable version	2.1 Early career, breakthrough, and international success 2.1.1 The Yardbirds and the Bluesbreakers	Instruments	Guitar, vocals
Languages 🔅 Afrikaans	2.1.1 The Yardbirds and the Bluesbreakers 2.1.2 Cream 2.1.3 Blind Faith, Delaney, and Bonnie and Friends	Years active Labels	1962-present Surfdog, Warner Bros., Reprise, Polydor, RSO, Atco, Apple
العربية Aragonés Azərbaycanca বाংলा Беларуская Беларуская (тарашкевіца) Български Bosanski Català Čeština Cymraeg Dansk Deutsch Eesti Eλληνικά ★ Español Esperanto	 2.2 "Layla" and solo career 2.2.1 Derek and the Dominos 2.2.2 Personal challenges and early solo success 2.3 Continued success 2.4 Resurgence and stardom 2.5 Collaboration albums 2.6 <i>Clapton, Old Sock</i> and recent events 3 Influences 4 Legacy 5 Guitars 6 Woman tone 7 Other media appearances 8 Personal life 8.1 Relationships and children 8.2 Political views 	Associated acts	The Yardbirds, John Mayall & the Bluesbreakers, Powerhouse, Cream, the Beatles, the Dirty Mac, Blind Faith, Free Creek, George Harrison, Plastic Ono Band, Delaney & Bonnie and Friends, Derek and the Dominos, the Rolling Stones, the Band, JJ Cale, Phil Collins, Robert Cray, Kate Bush, Luciano Pavarotti, Freddie King, B.B. King, Santana, Brian Wilson, John Mayer, Roger Waters, Dire Straits, Elton John, Steve Winwood, Stevie Ray Vaughan and Double Trouble, Sheryl Crow, Paul Wassif
Euskara	8.3 Controversy over remarks on immigration 8.4 Wealth and assets	Website	www.ericclapton.com
فارسی Français	8.5 Charitable work	N	otable instruments
Français Frysk Gaeilge Galego 한국어 Հայերեն हिन्दी Hrvatski Ido Bahasa Indonesia Íslenska	9 Awards and honours 10 Football 11 Clapton's music in film and TV 12 Discography 12.1 Solo studio albums 13 References 14 Further reading 15 External links	Fende Fende Gi	See: Guitars section er Stratocaster "Blackie" er Stratocaster "Brownie" fibson SG "The Fool" Gibson ES-335 Fender Telecaster Martin 000-42 acoustic
Italiano עברית	Early life [edit]		

Advanced Topics in Information Retrieval / Semantic Search

പത്തത്താ

Discography [edit]

Main article: Eric Clapton discography

Solo studio albums [edit]

- Eric Clapton (1970)
- 461 Ocean Boulevard (1974)
- There's One in Every Crowd (1975)
- No Reason to Cry (1976)
- Slowhand (1977)
- Backless (1978)
- Another Ticket (1981)
- Money and Cigarettes (1983)

References [edit]

- Behind the Sun (1985)
- August (1986)
- Journeyman (1989)
- Rush (soundtrack) (1992)
- From the Cradle (1994)
- Pilgrim (1998)
- Riding with the King (with B.B. King) (2000)
- Reptile (2001)

- Me and Mr. Johnson (2004)
- Sessions for Robert J (2004)
- Back Home (2005)
- The Road to Escondido (with JJ Cale) (2006)
- Clapton (2010)
- Old Sock (2013)^[163]
- The Breeze: An Appreciation of JJ Cale (2014)

- 1. ^ a b "55 Eric Clapton" @. Rolling Stone. Retrieved 11 November 2014.
- 2. ^ *a b* "100 Greatest Guitarists of All Time 2. Eric Clapton" &. Rolling Stone. Retrieved 30 November 2011.
- A ^{a b} "Top 50 Guitarists of All Time 10 to 1" ^b. Gibson Guitar Company. Retrieved 22 July 2011.
- 4. ^ Tyrangiel, Josh (2009-08-14). "The 10 Greatest Electric Guitar Players" ₽. *Time* (Time, Inc.). Retrieved 2011-04-26.
- 6. ^ "Inductee: Eric Clapton" & Rockhall.com. Retrieved 22 September 2014.
- [^] "Ex-rebel Clapton receives his CBE" ☑. *Telegraph.co.uk.* 4 November 2004. Retrieved 22 September 2014.
- 8. A "Eric Clapton, All Music: Grammy Awards" &. AllMusic. Retrieved 22 September 2014.
- 9. ^ a b "Brit Awards 1987" @. Brits.co.uk. Retrieved 22 September 2014.
- 10. ^ "Eric Clapton Biography" @. Rolling Stone. Retrieved 23 October 2011.
- 11. A Harry Shapiro (1992) Eric Clapton: Lost in the Blues pg. 29. Guinness,

Concert" " @. TIME.com. Retrieved 22 September 2014.

- 86. ▲ "Eric Clapton's T-Shirt Hit A Crossroad(s)" & Ticodo. 17 June 2013. Retrieved 27 June 2013.
- 87. ▲ "Eric Clapton Kicks Off 50th Anniversary Tour With Killer Setlist" Coolalbumreview.com. 15 March 2013. Retrieved 9 June 2013.
- * "Eric Clapton: 'When I'm 70, I'll Stop Touring' " RollingStone Magazine. Retrieved 11 June 2014.
- 89. ^ "More Signs Point Towards Eric Clapton Touring Retirement" & JamBase. Retrieved 11 June 2014.
- 90. ^ "Eric Clapton Says Touring Has Become 'Unbearable,' Confirms Retirement Plans" & Ultimate Classic Rock. Retrieved 27 June 2014.
- 91. A "Eric Clapton to headline Baloise Session in November" &. Where's Eric. Retrieved 8 April 2014.
- 92. [▲] "Eric Clapton Live from Basel on SRF3" ^B. Where's Eric. Retrieved 8 April 2014.
- 94. A "Eric Clapton and Friends Honor JJ Cale With New Tribute LP" 2.

4000

	Eric Clapton	[show]
	Associated bands	[show]
V*T*E	Rock and Roll Hall of Fame Class of 2000	[show]
Authority control	WorldCat &·VIAF: 46945873 &·LCCN: n84222859 &·ISNI: 0000 0001 0896 4267 &·GND: 119073102 &·SUDOC: 033662827 &·BNF: cb13892530m & (data) &·MusicBrainz: 618b6900-0618-4f1e-b835-bccb17f84294 &·NDL: 00620495 &	
Antonio de Este O	apton 1945 births Alumni of Kingston College (England) Alumni of Kingston University Atco Records artists	

Silver Clef Awards winners | Slide guitarists | Songwriters Hall of Fame inductees | The Yardbirds members | The Dirty Mac members Warner Bros. Records artists

This page was last modified on 10 November 2014 at 22:39.

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 Mobile view



- Collaboratively edited knowledge base intended to provide common source of structured data for other related projects (e.g., Wikipedia)
 - 12 M data items
 - started in **2012**



	Item	Discussion		Read	View history	А́ð English Search	Create account Log in	
WIKIDATA		Eric Clapton (Q48187)			[edit]			
Main page	-	English musician, singer, songwrite	er, and guitarist		[edit]			
Community portal Project chat		Also kilowii da.			[edit]			
Create a new item Item by title Recent changes			ikipedia pages linked to this item Wikinews pages linked to this Wikivoyage pages linked to this item Pages on other sites linke			linked to this item		
Random item Help Donate	Ι,	In other languages			[edit]			
Language select			Eric Clapton					
Select		Deutsch	englischer Blues- und Rock-Gitarrist und -Sänger					
What links here Related changes			Also known as: Eric Patrick Clapton Clapton Slowh	and				
Special pages Printable version Permanent link Page information			No label defined yet					
Concept URI Cite this page		Plattdüütsch	No description defined yet					
			Also known as:					
			No label defined yet					

Statements		
sex or gender	male	[edit]
	▶ 4 references	
		[add]
discography	Eric Clapton discography	[edit]
	▶ 1 reference	
		[add]
Commons category	Eric Clapton 🖗	[edit]
	▼ 0 references	[add reference]
		[add]
LCNAF identifier	an84222859 ₪	[edit]
	▶ 2 references	
		[add]
VIAF identifier	ੇ 46945873 🖗	[edit]
	▶ 3 references	

member of	Cream	[edit]
	- 0 references	[add reference]
	The Yardbirds	[edit]
	✓ 0 references	[add reference]
		[add]
NDL identifier	<mark>ੂ 00620495</mark> ਛੋ	[edit]
	✓ 0 references	[add reference]
		[add]
date of birth	30 March 1945	[edit]
	▶ 2 references	[odd]
topiala main actorony		[add]
topic's main category	Category:Eric Clapton • 0 references	[edit]
		[add reference]



- Knowledge base extracted from Wikipedia
 - **4.6 M things** (for English)
 - 80 M links to Wikipedia categories
 - 41 M links to YAGO categories
 - 50 M owl:sameAs links to other data sources
 - started in **2007**



http://www.dbpedia.org

About: Eric Clapton

An Entity of Type : musical artist, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Eric Patrick Clapton, CBE, (born 30 March 1945) is an English musician, singer and songwriter. He is the only three-time inductee to the Rock and Roll Hall of Fame: once as a solo artist and separately as a member of the Yardbirds and Cream. Clapton has been referred to as one of the most important and influential guitarists of all time.

Property	Value
dbpedia-owl:abstract	Eric Patrick Clapton, CBE, (born 30 March 1945) is an English musician, singer and songwriter. He is the only three-time inductee to the Rock and Roll Hall of Fame: once as a solo artist and separately as a member of the Yardbirds and Cream. Clapton has been referred to as one of the most important and influential guitarists of all time. Clapton ranked second in Rolling Stone magazine's list of the "100 Greatest Guitarists of All Time" and fourth in Gibson's "Top 50 Guitarists of All Time". In the mid-1960s, Clapton left the Yardbirds to play blues with John Mayall & the Bluesbreakers. Immediately after leaving Mayall, Clapton joined Cream, a power trio with drummer Ginger Baker and bassist Jack Bruce in which Clapton played sustained blues improvisations and "arty, blues-based psychedelic pop." For most of the 1970s, Clapton's output bore the influence of the mellow style of JJ Cale and the reggae of Bob Marley. His version of Marley's "I Shot the Sheriff" helped reggae reach a mass market. Two of his most popular recordings were "Layla", recorded by Derek and the Dominos, another band he formed, and Robert Johnson's "Crossroads", recorded by Cream. Following the death of his son Conor in 1991, Clapton's grief was expressed in the song "Tears in Heaven", which featured in his Unplugged album.Clapton has been the recipient of 17 Grammy Awards, and the Brit Award for Outstanding Contribution to Music. In 2004, he was awarded a CBE at Buckingham Palace for services to music. In 1998, Clapton, a recovering alcoholic and drug addict, founded the Crossroads Centre on Antigua, a medical facility for recovering substance abusers.
dbpedia-owl:activeYearsStartYear	 1962-01-01 (xsd:date)
dbpedia-owl: alias	 Clapton, Eric Patrick; Slowhand Slowhand
dbpedia-owl:associatedBand	 dbpedia:Cream_(band) dbpedia:Eric_Clapton_and_the_Powerhouse dbpedia:Delaney_&_Bonnie dbpedia:George_Harrison dbpedia:Kate_Bush dbpedia:SheryL_Crow dbpedia:The_Band dbpedia:The_Batles dbpedia:The_Rolling_Stones dbpedia:Elton_John dbpedia:Elton_John dbpedia:Dire_Straits dbpedia:Phil_Collins dbpedia:Phil_Collins dbpedia:Roge_Waters dbpedia:Derek_and_the_Dominos dbpedia:Derek_and_the_Dominos dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith dbpedia:Bhort_Faith



dbpedia-owl:background	 solo_singer
dbpedia-owl:birthDate	 1945-03-30 (xsd:date)
dbpedia-owl:birthPlace	 dbpedia:Surrey dbpedia:Ripley,_Surrey
dbpedia-owl:birthYear	 1945-01-01 (xsd:date)
dbpedia-owl:genre	 dbpedia:Blues dbpedia:Psychedelic_rock dbpedia:Rock_music dbpedia:Blues_rock dbpedia:Hard_rock
dbpedia-owl:individualisedGnd	• 119073102
dbpedia-owl:instrument	 dbpedia:Blackie_(guitar)
dbpedia-owl:lccnld	 n/84/222859
dbpedia-owl:recordLabel	 dbpedia:Warner_BrosRecords dbpedia:Apple_Records dbpedia:Polydor_Records dbpedia:Reprise_Records dbpedia:Atco_Records dbpedia:Surfdog_Records dbpedia:RSO_Records
dbpedia-owl:thumbnail	http://commons.wikimedia.org/wiki/Special:FilePath/Eric_Clapton_2.jpg?width=300
dbpedia-owl:viafld	 46945873
dbpedia-owl:wikiPageExternalLink	 http://crossroadsantigua.org/ http://guitarinternational.com/wpmu/2010/02/26/a-talk-with-lee-dickson-guitar-tech-to-eric-clapton/ http://muzica.acasa.ro/stiri-muzica/Clapton-si-Winwood-la-Bucuresti-sau-cum-am-ajuns-sa-am-din-nou-credinta-oarba-in-blues http://whereseric.com/home http://www.ericclapton.com/ http://www.free-covers.org/cd-covers/eric-clapton-backless-1978-cd-cover-1335.html http://www.gibson.com/en-us/Lifestyle/Features/how-to-get-claptons-classic/ http://www.spoonfed.co.uk/spooners/jhallam-2972/eric-clapton-at-floridita-461/
dbpedia-owl:wikiPageID	 10049 (xsd:integer)
dbpedia-owl:wikiPageRevisionID	 606135391 (xsd:integer)
dbpprop:alias	 Slowhand
dbpprop:alternativeNames	Clapton, Eric Patrick; Slowhand
dbpprop:associatedActs	 dbpedia:Cream_(band) dbpedia:Eric_Clapton_and_the_Powerhouse dbpedia:Delaney_&_Bonnie dbpedia:George_Harrison dbpedia:Kate_Bush dbpedia:Sheryl_Crow dbpedia:The_Band dbpedia:The_Beatles dbpedia:The_Rolling_Stones

dcterms:subject	 category:1945_births
	 category:Alumni_of_Kingston_College_(England)
	 category:Alumni_of_Kingston_University
	 category:Atco_Records_artists
	 category:Blind_Faith_members
	 category:Blues_rock_musicians
	 category:Blues_singer-songwriters
	 category:Brit_Award_winners
	 category:British_rhythm_and_blues_boom_musicians
	 category:Car_collectors
	 category:Commanders_of_the_Order_of_the_British_Empire
	 category:Cream_(band)_members
	 category:Delaney_&_Bonnie_&_Friends_members
	 category:Derek_and_the_Dominos_members
	 category:Electric_blues_musicians
	 category:English_Christians
	 category:English_autobiographers
	 category:English_blues_guitarists
	 category:English_blues_musicians
	 category:English_blues_singers
	 category:English_buskers
	 category:English_male_singers
	 category:English_people_of_Canadian_descent
	 category:English_rock_guitarists
	 category:English_rock_singers
	 category:English_singer-songwriters
	 category:Eric_Clapton
	 category:Grammy_Award-winning_artists
	 category:Guitar_performance_techniques
	 category:lvor_Novello_Award_winners
	 category:Jazz-blues_guitarists
	 category:John_Mayall_&_the_Bluesbreakers_members
	 category:Lead_guitarists
	 category:Living_people
	 category:People_from_Guildford_(district)
	 category:Plastic_Ono_Band_members
	 category:Polydor_Records_artists
	 category:RSO_Records_artists
	 category:Reprise_Records_artists
	 category:Resonator_guitarists
	 category:Rock_and_Roll_Hall_of_Fame_inductees
	 category:Silver_Clef_Awards_winners
	 category:Slide_guitarists
	 category:Songwriters_Hall_of_Fame_inductees
	 category:The_Dirty_Mac_members
	 category:The_Yardbirds_members
	 category:Warner_BrosRecords_artists

rdf:type	foaf:Person
	http://schema.org/Person
	 yago:WarnerBros.RecordsArtists
	 http://wikidata.dbpedia.org/resource/Q215627
	 owl:Thing
	 http://schema.org/MusicGroup
	 http://wikidata.dbpedia.org/resource/Q5
	 dul:Agent dul:NaturalPerson
	dbpedia-owl:Agent
	dbpedia-owl:Artist
	dbpedia-owl:MusicalArtist
	 dbpedia-owl:Person
	http://umbel.org/umbel/rc/Artist
	 http://umbel.org/umbel/rc/MusicalPerformer
	 yago:AlumniOfKingstonUniversity
	 yago:Alumnus109786338
	 yago:Artist109812338
	 yago:Autobiographer109825296
	 yago:Biographer109855433
	 yago:BluesRockMusicians
	 yago:BritishBluesGuitarists
	yago:Busker109883174
	 yago:CarCollectors
	 yago:CausalAgent100007347
	 yago:Collector109936620
	 yago:CommandingOfficer109941964
	 yago:Communicator109610660
	 yago:Creator109614315
	 yago:Entertainer109616922
	 yago:Guitarist110151760
	 yago:Intellectual109621545
	 yago:LivingThing100004258
	 yago:MilitaryOfficer110317007
	 yago:Musician110339966
	 yago:Musician110340312
	 yago:Object100002684
	 yago:Organism100004475
	 yago:Performer110415638
	 yago:Person100007846 yago:Person100007846
	 yago:Scholar110557854 yago:Scholar110557854
	 yago:Serviceman110582746
	 yago:Singer110599806
	 yago:SkilledWorker110605985
	 yago:Whole100003553
	 yago:Worker109632518
	 yago:Writer110794014
DBpedia

owl:sameAs fbase:Eric Clapton http://fr.dbpedia.org/resource/Eric_Clapton http://wikidata.dbpedia.org/resource/Clapton http://wikidata.dbpedia.org/resource/Clapton http://el.dbpedia.org/resource/Eric_Clapton http://el.dbpedia.org/resource/Eric_Clapton http://el.dbpedia.org/resource/Eric_Clapton http://el.dbpedia.org/resource/Eric_Clapton http://id.dbpedia.org/resource/Eric_Clapton
 http://kc.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http://kd.dbpedia.org/resource/Eric_Clapton http:/kd.dbpedia.org/resource/Eric_Cl

DBpedia

is dbpedia-owl:composer of	 dbpedia:Edge_of_Darkness
is dbpedia-owl:formerBandMember of	 dbpedia:Cream_(band) dbpedia:Eric_Clapton_and_the_Powerhouse dbpedia:Delaney_&_Bonnie dbpedia:Derek_and_the_Dominos dbpedia:Blind_Faith dbpedia:Plastic_Ono_Band dbpedia:The_Dirty_Mac
is dbpedia-owl:musicComposer of	 dbpedia:The_Story_of_Us_(film) dbpedia:Water_(1985_film) dbpedia:Lethal_Weapon_(film_series) dbpedia:Communion_(1989_film) dbpedia:Rush_(1991_film) dbpedia:Nil_by_Mouth_(film) dbpedia:Wonderwall_(film) dbpedia:The_Van_(1996_film) dbpedia:Lethal_Weapon dbpedia:Homeboy_(film) dbpedia:Homeboy_(film) dbpedia:Concert_for_George_(film) dbpedia:Concert_for_George_(film) dbpedia:Lethal_Weapon_2 dbpedia:Lethal_Weapon_3 dbpedia:Lethal_Weapon_4
is dbpedia-owl:musicalArtist of	 dbpedia:After_Midnight_(song) dbpedia:Lay_Down_Sally dbpedia:Willie_and_the_Hand_Jive dbpedia:Wonderful_Tonight dbpedia:Beat_Love_(Eric_Clapton_song) dbpedia:Bad_Love_(Eric_Clapton_song) dbpedia:Shot_the_Sheriff dbpedia:Cocaine_(song)Cocaine1 dbpedia:Change_the_World dbpedia:Tearing_Us_Apart dbpedia:Tearing_Us_Apart dbpedia:Tearing_bs_(Yellow_Magic_Orchestra_song)Behind_the_Mask3
is dbpedia-owl:musicalBand of	 dbpedia:After_Midnight_(song) dbpedia:Lay_Down_Sally dbpedia:I've_Got_a_Rock_'n'_Roll_Heart dbpedia:Willie_and_the_Hand_Jive

- Collaboratively edited knowledge base which includes information extracted from other data sources (e.g., Wikipedia, MusicBrainz)
 - 44 M topics (organized into domains)
 - 2.4 B facts
 - focused on entertainment
 - many topics without textual data
 - started in 2007, acquired by Google in 2010
 - forms the core of the Google Knowledge Graph

http://www.freebase.com



Freebase F	ïnd		Browse	Query	Help		Sign In or Sign Up	English 👻
Topic	Eric Patrick Clapton, separately as a mem Stone magazine's lis 10 Best Electric Guit joined Cream, a pow	able type: /music/artist of CBE, is an English musici- ber of the Yardbirds and C t of the "100 Greatest Guit ar Players" in 2009 In the r er trio with drummer Ginge	an, singer-songwi Cream. Clapton ha tarists of All Time' mid-1960s, Clapto er Baker and bass	riter and guit is been refer ' and fourth i on left the Ya sist Jack Bru	arist. He is the only three-ti red to as one of the most in n Gibson's "Top 50 Guitaris rdbirds to play blues with J ce in which Clapton played	Cr me inductee to the Rock and Roll Hall hportant and influential guitarists of all sts of All Time". He was also named nu ohn Mayali & the Bluesbreakers. Imme sustained blues improvisations and "a Bob Marley. His version of Marley's "I	time. Clapton ranked second umber five in Time magazine's ediately after leaving Mayall, C arty, blues-based psychedelic	t and I in Rolling I list of "The Clapton I pop". For
Properties	l18n	Keys	Links					Ö.
View and edit specifi							Types:	
Common /commo							Common Topic	
Also known as /co Also known as * Clapton, Eric Eric Claption							Film Film music contr Film actor Person or entity in film	
Eric Patrick Clapton eric_clapton Slow Hand God Derek and the Domino Blind Faith	es						Music Musical Artist Musician Record Producer Composer Guitarist Songwriter	r
The Yardbirds							Lyricist	



mage /common/topic/image		TV Personality
		Broadcast
		Broadcast Artist
Official website /common/topic/official_website		People
ttp://www.ericclapton.com/		Person Measured person
opic equivalent webpage /common/topic/topic_equivalent_webpage		Measured person
opic equivalent webpage *		Internet
ttp://www.nytimes.com/top/reference/timestopics/people/c/eric_clapton		Social network user
ttp://www.nndb.com/people/670/000024598/		
ttp://www.imdb.com/name/nm0002008/		Organization
tp://www.myspace.com/ericclapton		Organization founder
ttp://www.tvrage.com/person/id-64649		Media
tp://www.netflix.com/RoleDisplay/17288		Author
tp://it.wikipedia.org/wiki/index.html?curid=102202		
ttp://es.wikipedia.org/wiki/index.html?curid=49785		Awards
ttp://de.wikipedia.org/wiki/index.html?curid=92491		Award Nominee
ttp://pt.wikipedia.org/wiki/index.html?curid=73685		Award Winner
4 values total »		Hall of fame inductee
ocial media presence /common/topic/social_media_presence	✓ 2014-09-04	Ranked item
ocial media presence		Influence
tp://www.facebook.com/ericclapton		Influence Node
tp://www.twitter.com/EricClaptonNews		0. L L W
opical webpage /common/topic/topical_webpage		Celebrities
ppical webpage		Celebrity
tp://www.whereseric.com/		
tp://www.discogs.com/artist/Eric+Clapton		
ttp://musicmoz.org/Bands_and_Artists/C/Clapton,_Eric/		
ttp://whereseric.com/		
ttp://lyrics.wikia.com/Eric_Clapton		
ttp://decoda.com/eric-clapton-lyrics		

Film /film	
Film music contributor /film/music_contributor	
Film music credits /film/music_contributor/film	
Film music credits Nil by Mouth	
Rush	
Water	
Lethal Weapon 2	
Wonderwall	
Lethal Weapon 4	
Lethal Weapon	
Lethal Weapon 3	
Homeboy	
The Van	
14 values total »	

Film actor /film/acto

Film performances /film/actor/film			
Film 🗵	Character	Special Performance Type	Character note
Sweet Toronto			
The Cream of Eric Clapton			
The Very Best of Cher: The Video Hits Collection			
Tommy	The Preacher		
Hail! Hail! Rock 'n' Roll			
Water			
The Rolling Stones Rock and Roll Circus			
Concert for George			
Blues Brothers 2000	The Louisiana Gator Boys member #4		
Cucumber Castle			
52 values total »			

Music /music	
Musical Artist /music/artist	
Musical Genres /music/artist/genre	
Musical Genres	
Blues	
Rock music	
Blues rock	
Pop rock	
Hard rock	
Psychedelic rock	
Reggae	
Soft rock	
Place Musical Career Began /music/artist/origin	
London	
Albums /music/artist/album	
Albums 🗠	
The Early Clapton Collection	
Eric Clapton	
Eric Clapton's Rainbow Concert	
461 Ocean Boulevard	
There's One in Every Crowd	
E.C. Was Here	
No Reason to Cry	
Slowhand	
Backless	
224 values total -	
Record Labels /music/artist/label	
Record Labels	
Apple Records	

fember of /music/group_member/membership				
roup 🗠	Roles V	Period (start)	Period (end)	
erek and the Dominos	Vocals	1970	1971	
	Lead guitar			
ohn Mayall & the Bluesbreakers	Guitar	1965 April	1966 July	
lind Faith	Guitar	1968	1969 October	
	Vocals			
he Yardbirds	Guitar	1963 October	1965	
ream	Guitar	1966	1968	
	Vocals			
he Plastic Ono Band	Lead guitar	1969	1969	
elaney & Bonnie and Friends				
ting with Eric Clapton				
he Dirty Mac				
ric Clapton & The Powerhouse				
I values total »				
rtists supported /music/group_member/artists_supp				
nstruments Played /music/group_member/instrume				
struments Played				
iuitar				
lide guitar				
iano				
landolin				
ocal Range /music/group_member/vocal_range				



People /people	
Person /people/person	
Date of birth /people/person/date_of_birth	✓ 2013-12-05
3/30/1945	
Place of birth /people/person/place_of_birth	✓ 2013-12-12
Ripley	
Country of nationality /people/person/nationality	✓ 2014-02-10
England	
Gender /people/person/gender	✓ 2013-12-05
Male	
Profession /people/person/profession	
Profession	
Guitarist	
Singer-songwriter	
Musician	
Composer	
Artist	
Actor Record producer	
Film Producer	
Television producer	
Film Score Composer	
11 values total »	
Religion /people/person/religion	
Ethnicity /people/person/ethnicity	
Parents /people/person/parents	

Knowledge base extracted from Wikipedia, WordNet, and GeoNames

- 10 M named entities
- 120 M facts

 $oldsymbol{O}$

- 72 relation types with manually assessed accuracy (95%)
- started in **2007**

YAGO



http://www.yago-knowledge.org



YAGO

Search: eng 🗘	<eric_clapton></eric_clapton>	
← <pattie_boyd> <aashish_khan> <abbey_road> <abe_laboriel,_jr.> <a_boxful_of_treasures> <accafellas> <acce_(band)> <ace_(band)> <ace_frehley> <achy_breaky_heart> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_queen> <acid_still_required></acid_still_required></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></acid_queen></achy_breaky_heart></ace_frehley></ace_(band)></acce_(band)></accafellas></a_boxful_of_treasures></abe_laboriel,_jr.></abbey_road></aashish_khan></pattie_boyd>	<ismarriedto></ismarriedto>	rdfs:label "Eric Clapton"@afr "Eric Clapton"@tric Clapton"@tric Clapton"@cat "Eric Clapton"@cas "Eric Clapton"@cas "Eric Clapton"@dan "Eric Clapton"@deu "Eric Clapton"@deu "Eric Clapton"@eta "Eric Clapton"@eta "Eric Clapton"@eta "Eric Clapton"@eta "Eric Clapton"@fin "Eric Clapton"@glg "Eric Clapton"@glg "Eric Clapton"@hun "Eric Clapton"@hun "Eric Clapton"@hun "Eric Clapton"@ido "Eric Clapton"@hun
<ain't_nobody's_business> </ain't_nobody's_business>	linksTo>	<ahoy_rotterdam> <aid_still_required> <air_canada_centre> <a_knight's_tale> <albert_king> <alex_lifeson> <alice_ormsby-gore> <alicia_witt> <allmusic> <all_things_must_pass> <almost_famous></almost_famous></all_things_must_pass></allmusic></alicia_witt></alice_ormsby-gore></alex_lifeson></albert_king></a_knight's_tale></air_canada_centre></aid_still_required></ahoy_rotterdam>

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers

- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers

- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers





- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers









- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers





- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

 Entity linking (aka. "Wikification") is the process of spotting mentions of named entities in a text document, resolving their ambiguity, and linking them to known identifiers





- Examples:
 - Wikipedia Miner (<u>http://wikipedia-miner.cms.waikato.ac.nz</u>)
 - AIDA (<u>https://gate.d5.mpi-inf.mpg.de/webaida/</u>)

Google Freebase Annotations

 ClueWeb09/12 are large-scale collections of web documents (1 B web pages) often used in Information Retrieval research



 Google has annotated these collections with Freebase topics (<u>http://googleresearch.blogspot.de/2013/03/learning-....html</u>)











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

3. (More) Semantic Keyword Search

 (More) Semantic Keyword Search improves ad-hoc information retrieval by leveraging knowledge bases (input: keywords, <u>output</u>: list of documents)

$$a \times I$$
 Keywords
 \blacksquare
 Text
 \blacksquare
 Documents

Explicit Semantic Analysis

- Synonymy (e.g., car/automobile) and polysemy (e.g., jaguar as cat, car, guitar) have negative effects on retrieval effectiveness
- Concept-based Information Retrieval represents queries and documents in terms of (semantic) concepts instead of terms
 - WordNet synsets after performing word sense disambigutation
 - Latent concepts identified based on word co-occurrences (e.g., latent semantic analysis applies SVD to term-document matrix)
- Egozi et al. [4] represent queries and documents in terms of knowledge base concepts (in their case from Wikipedia)

Words in Terms of Concepts

- Explicit Semantic Analysis (ESA) represents words (e.g., jaguar) in terms of the Wikipedia concepts that they describe (e.g., JAGUAR, JAGUAR_CARS, JAGUAR_XK, FENDER_JAGUAR)
- <u>Step 1</u>: Construct tf.idf vectors for Wikipedia concepts based on the text of their corresponding article

Fender_Jaguar	guitar:0	.8 fender :0	.6 strin	igs :0.3	amp:0.5 cc	bain :0.1
JAGUAR_XK	car :0.2	engine:0.3	v8 :0.7	fuel :0.4	british:0.2	racing:0.1

Words in Terms of Concepts

• <u>Step 2</u>: **Invert** the **concept vectors** to obtain **word vectors**

Car JAGUAR_XK:0.2 PORSCHE_911:0.3 NISSAN_MICRA:0.1 MINI_COOPE	ER:0.4					
fuel SHELL:0.2 JAGUAR_XK:0.4 BP:0.4 PORSCHE_911:0.4						
fender Fender_Jaguar:0.6 Eric_Clapton:0.2 Fender_Stratocaster:0.8						

 Similarity of word vectors (e.g., in terms of their cosine similarity) can be used to find out about semantic relatedness of words

Texts in Terms of Concepts

- Texts (e.g., documents, passages, queries) can now be represented in the space of concepts by combining the vectors of contained words
- For a sequence of words d = < v₁, v₂, ..., v₃ > with w(w_i, d) as the tf.idf weight of word v_i in d and w(c, w_i) as the tf.idf weight of concept c for word v_i construct a vector for d in terms of concepts as

$$w(c,d) = \sum_{v_i \in d} w(c,v_i) \cdot w(v_i,d)$$

ESA-Based Retrieval

- Egozi et al. [4] represent and index documents as a whole and their contained passages in terms of concepts; the obtained vectors are truncated, setting all but the s largest components (concepts) to zero
- Queries are also represented in terms of concepts; the obtained query vectors are truncated, setting all but the s largest components (concepts) to zero
- For a document d containing passages $\langle p_1, p_2, ..., p_3 \rangle$ the retrieval score for query q is determined as

$$\operatorname{score}(q,d) = \operatorname{sim}(q,d) + \max_{i} \operatorname{sim}(q,p_{i})$$

ESA-Based Retrieval

- Unfortunately, the approach is prone to topic drift and performs worse (MAP 0.1760) than a word-based retrieval method (MAP 0.2481) on TREC-8 benchmark
- <u>Example</u>: Concepts with high weight for query estonia economy ESTONIA, ECONOMY OF ESTONIA, ESTONIA AT THE 2000 SUMMER OLYMPICS, ESTONIA AT THE 2004 SUMMER OLYMPICS, ESTONIA NATIONAL FOOTBALL TEAM, etc.

Selective ESA-Based Retrieval

- Pseudo-relevant and pseudo-irrelevant documents identified using a word-based retrieval method yield a better query vector
- Let D_R and D_{NR} be the sets of top-k and bottom-k documents returned; assume D_R as relevant, D_{NR} as irrelevant
- Use Rocchio's relevance-feedback method to estimate a better query vector considering pseudo-(ir)rrelevant documents

$$w(c,q) = \left(\sum_{v \in q} w(c,v) \cdot w(v,q)\right)$$
$$+ \frac{1}{|D_R|} \left(\sum_{d \in D_R} w(c,v) \cdot w(v,d)\right) - \frac{1}{|D_{NR}|} \left(\sum_{d \in D_{NR}} w(c,v) \cdot w(v,d)\right)$$

Fused ESA-Based Retrieval

- Pseudo-relevance feedback helps (MAP 0.2200) but the approach still performs worse than the baseline (MAP 0.2481)
- Fusion of ESA-based results and word-based results beneficial
 - Let RESA be the results returned by the ESA-based retrieval method
 - Let \mathbf{R}_W be the results returned by the word-based retrieval method
 - Combine their retrieval scores as

 $\operatorname{score}(d) = \alpha \cdot \frac{\operatorname{score}_{ESA}(d) - \min_{ESA}}{\max_{ESA} - \min_{ESA}} + (1 - \alpha) \cdot \frac{\operatorname{score}_{W}(d) - \min_{W}}{\max_{W} - \min_{W}}$

• Fused approach (MAP 0.2888) improves (+16.4%) over baseline

4. Entity Search

 Entity Search retrieves entities based on keyword queries which may include explicit/implicit cues about the target type (input: keywords + types, <u>output</u>: 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 a **common in e-commerce**



Language Models for Entity Search

- Balog et al. [1] develop a language modeling approach, assuming that entities come with ready-to-use representation:
 - a textual description (e.g., Wikipedia article, IMDB profile, etc.)
 - a set of categories with meaningful names (e.g., AMERICAN_FILM)



AMERICAN_FILM, 2002 FILMS, FILMS ABOUT DRUGS, ...



TOOLS, POCKET KNIVES, POWER & HAND TOOLS, ...

Language Models for Entity 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

$$\mathbf{P}\left[q \mid e\right] = \lambda \cdot \mathbf{P}\left[\theta_{q}^{T} \mid \theta_{e}^{T}\right] + (1 - \lambda) \cdot \mathbf{P}\left[\theta_{q}^{C} \mid \theta_{e}^{C}\right]$$

with probabilities estimated from Kullback-Leibler divergences
Language Models for Query Components

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

$$P\left[\theta_{q} \mid \theta_{e}\right] = \frac{\max_{e^{*}} KL(\theta_{q} \| \theta_{e^{*}}) - KL(\theta_{q} \| \theta_{e})}{\sum_{e'} \max_{e^{*}} KL(\theta_{q} \| \theta_{e^{*}}) - KL(\theta_{q} \| \theta_{e'})}$$

- Language modeling framework with various instantiations, e.g.:
 - estimate θ_q^T from T and C but ignore θ_q^T (M2)
 - estimate both θ_q^T and θ_q^T from T (M3)
 - estimate θ_q^T only from T and θ_q^C only from C (M4)
 - estimate θ_q^T only from T and C and θ_q^C only from C (M5)

Language Models for Query Components

• Estimate both θ_q^T and θ_q^C from T (M3)

$$\begin{split} \mathbf{P}\left[\left. v \mid \boldsymbol{\theta}_{q}^{T} \right. \right] &= \frac{tf(v,T)}{\sum_{v'} tf(v',T)} \\ \mathbf{P}\left[\left. c \mid \boldsymbol{\theta}_{q}^{C} \right. \right] &= \begin{cases} \frac{\mathbf{P}\left[T \mid c \right]}{\sum_{c' \in \mathrm{tc}(k,T)} \mathbf{P}\left[T \mid c' \right]} &: c \in \mathrm{tc}(k,T) \\ 0 &: \text{otherwise} \end{cases} \end{split}$$

with P[T|c] as query likelihood of generating T from terms in c and tc(k,T) as the set of k categories with highest query likelihood

• Estimate θ_q^T only from T and θ_q^C only from C (M4)

$$\mathbf{P}\left[v \mid \theta_q^T\right] = \frac{tf(v, T)}{\sum_{v'} tf(v', T)} \qquad \mathbf{P}\left[c \mid \theta_q^C\right] = \frac{tf(c, C)}{\sum_{c'} tf(c', C)}$$

Language Models for Entity Components

• Language models θ_e^T and θ_e^C for entity components based on maximum-likelihood estimation with **Dirichlet smoothing**

$$\mathbf{P}\left[v \mid \theta_{e}^{T}\right] = \frac{tf(v, e) + \mu^{T} \cdot \frac{tf(v, E)}{\sum_{v'} tf(v', E)}}{\sum_{v'} tf(v', e) + \mu^{T}}$$

$$\mathbf{P}\left[c \mid \theta_{e}^{C}\right] = \frac{tf(c, e) + \mu^{C} \cdot \frac{tf(c, E)}{\sum_{c'} tf(c', E)}}{\sum_{c'} tf(c', e) + \mu^{C}}$$

with μ^{T} and μ^{C} as component-specific smoothing parameters

Language Models for Entity Search

• Example: Query $T = \{Paul Auster novels\} C = \{NOVELS\}$

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	

Language Models for Entity Search

• Example: Query T = {Paul Auster novels} C = {NOVELS}

		-	
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	



Categories: 1987 novels | Literary trilogies | Metafictional works | Mystery novels | Novel series | Novels by Paul Auster | Novels set in New York City | Faber and Faber books

Language Models for Entity Search

• Example: Query $T = \{Paul Auster novels\} C = \{NOVELS\}$

	-	-	
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	



Categories: 1958 novels | American historical novels | Novels by Anya Seton | Novels set in Massachusetts | Novels set in the American colonial era

5. Knowledge Search

 Knowledge Search retrieves subgraphs as answers to structured queries which may be augmented with keywords (input: structured + keywords, <u>output</u>: list of subgraphs)



- **SPARQL** as a query language for RDF has **some shortcomings**
 - it provides **no meaningful ranking** of query answer subgraphs
 - it assumes that users know what the data looks like (i.e., its schema)

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

<u>Query</u>

?p directed ?m .
?p actedIn ?m .

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

<u>Query</u>

?p directed ?m .
?p actedIn ?m .

<u>Answers</u>

Ben_Affleck directed Argo . Ben_Affleck actedIn Argo .

Til_Schweiger directed Der_Eisbär . Til_Schweiger actedIn Der_Eisbär .

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

<u>Query</u>

?p directed **?m**. **?p** actedIn **?m**.

Answers

- Ben_Affleck directed Argo .
 Ben_Affleck actedIn Argo .
- **Zil_Schweiger** directed **Der_Eisbär**.
 Til_Schweiger actedIn **Der_Eisbär**.

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

<u>Query</u>

Quentin_Tarantino ?p ?m { movie, vampires, criminals }.

 Elbassuoni et al. [5] develop a language modeling approach for search in RDF graphs, which provides a ranking of query answers, and augments SPARQL/RDF with keywords

<u>Query</u>

Quentin_Tarantino ?p ?m { movie, vampires, criminals }.

<u>Answers</u>

Quentin_Tarantino actedIn From_Dusk_till_Dawn.

Tarantino plays criminal in vampire movie From Dusk till Dawn

Ranking for RDF Graphs

• Consider triple pattern query Q and matching answer A

$$Q = \langle q_1, \dots, q_n \rangle \qquad \qquad A = \langle a_1, \dots, a_n \rangle$$

Estimate language models over n-tuples of triples for Q and A

$$T = \langle t_1, \dots, t_n \rangle$$

Assume that a witness count c(t_i) is known for every triple t_i which reflects how often the fact has been mentioned
 (e.g., on the Web or in another document collection)

Ben_Affleck directed Argo		122,789
Til_Schw	reiger directed Der_Eisbär	2,014

Query Language Model

• Query language model θ_Q generates T with probability

$$\mathbf{P}\left[T \mid \theta_Q\right] = \prod_{i=1}^n \mathbf{P}\left[t_i \mid q_i\right]$$

assuming independence for generation of individual triples t_i

 $\circ~$ Triple t_i is generated from triple pattern q_i with probability

$$\mathbf{P}\left[t_{i} \mid q_{i}\right] = \begin{cases} \frac{c(t_{i})}{\sum_{t \in m(q_{i})} c(t)} & : & t_{i} \in m(q_{i})\\ 0 & : & \text{otherwise} \end{cases}$$

with m(qi) as the set of triples matching the triple pattern qi

Answer Language Model

• Answer language model θ_A generates T with probability

$$\mathbf{P}\left[T \mid \theta_A\right] = \beta \cdot \left(\prod_{i=1}^{n} \mathbf{P}\left[t_i \mid a_i\right]\right) + (1-\beta) \cdot \left(\prod_{i=1}^{n} \mathbf{P}\left[t_i \mid \theta_{KB}\right]\right)$$

assuming independence for generation of individual triples t_i and smoothing with a knowledge base language model θ_{KB}

- Triple **t**_i is generated from **answer triple** a_i with probability $P[t_i | a_i] = \mathbb{1}(t_i = a_i)$
- Triple ti is generated from knowledge base with probability

$$\mathbf{P}\left[t_i \mid \theta_{KB}\right] = \frac{c(t_i)}{\sum_{t \in KB} c(t)}$$

Ranking of Query Answers

• Divergence-based ranking of query answers according to

$$KL(\theta_Q \| \theta_A) = \log P[T | \theta_Q] \cdot \frac{P[T | \theta_Q]}{P[T | \theta_A]}$$

Augmenting SPARQL/RDF with Keywords

 Users may not know what the data looks like, but they can provide structure, and are familiar with keyword querying

"New movies by this person who made that movie with criminals and vampires"

- Assume that a witness count c(t_i, w) is known for every triple t_i which reflects how often the corresponding fact has been mentioned together with the keyword v
- Augment query triple patterns with sets of keywords with

$$Q = \langle q_1 \, w_1, \dots, q_n \, w_n \rangle$$

Augmenting SPARQL/RDF with Keywords

 Users may not know what the data looks like, but they can provide structure, and are familiar with keyword querying

"New movies by this person who made that movie with criminals and vampires"

?qt ?m ?om { vampires, criminals }. ?qt ?m **?nm** { }.

- Assume that a witness count c(t_i, w) is known for every triple t_i which reflects how often the corresponding fact has been mentioned together with the keyword v
- Augment query triple patterns with sets of keywords with

$$Q = \langle q_1 \, w_1, \dots, q_n \, w_n \rangle$$

Augmented Query Language Model

• Keyword-augmented query model θ_Q generates T with

$$\mathbf{P}\left[T \mid \theta_Q\right] = \prod_{i=1}^n \mathbf{P}\left[t_i \mid q_i \, w_i\right]$$

• Triple \mathbf{t}_i is generated from triple pattern \mathbf{q}_i and keywords \mathbf{w}_i with

$$P[t_i | q_i w_i] = \prod_{v \in w_i} (\alpha \cdot P[t_i | q_i v] + (1 - \alpha) \cdot P[t_i])$$

smoothing with uniform probability $P[t_i]$ and $P[t_i|q_i v]$ defined as

$$\mathbf{P}\left[t_{i} \mid q_{i} v\right] = \begin{cases} \frac{c(t_{i}, v)}{\sum_{t \in m(q_{i})} c(t, v)} & : & t_{i} \in m(q_{i}) \\ 0 & : & \text{otherwise} \end{cases}$$

Ranking of Query Answers

• Answer language model θ_A generates T with probability

$$\mathbf{P}\left[T \mid \theta_A\right] = \beta \cdot \left(\prod_{i=1}^{n} \mathbf{P}\left[t_i \mid a_i\right]\right) + (1 - \beta) \cdot \left(\prod_{i=1}^{n} \mathbf{P}\left[t_i \mid \theta_{KB}\right]\right)$$

and thus **remains unchanged**

• Divergence-based ranking of query answers according to

$$KL(\theta_Q \| \theta_A) = \log P[T | \theta_Q] \cdot \frac{P[T | \theta_Q]}{P[T | \theta_A]}$$

- Example results obtained on Internet Movie Database (IMDb)
- <u>Query 1</u>: A thriller movie and its director (structure only)
 - 1. Batman_Begins, Christopher_Nolan
 - **2.** Murder!, Alfred_Hitchcock
 - **3.** Spider-Man_3, Sam_Raimi
 - 4. Eyes_Wide_Shut, Stanley_Kubrick
- Query 2: An award winning director who directed a movie that is based on a {true story} (structured + keywords)
 - **1.** Martin_Scorsese, Good_fellas
 - **2.** Steven_Spielberg, Schindler's_List
 - **3.** Peter_Jackson, Heavenly_Creatures
 - 4. Bernando_Bertolucci, Little_Buddha

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
- Explicit Semantic Analysis represents queries and documents in terms of concepts; can improve ad-hoc information retrieval
- Entity Search retrieves entities based on keyword queries which may include type cues; applications in web search and business
- Knowledge Search on RDF graphs adds ranking to SPARQL and provides user-friendly querying through keywords

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] **J. Dalton, L. Dietz, J. Allan:** *Entity Query Feature Expansion using Knowledge Base Links*, SIGIR 2014
- [4] **O. Egozi, S. Markovitch, E. Gabrilovich:** Concept-Based Information Retrieval Using Explicit Semantic Analysis, ACM TOIS 29(2), 2011
- [5] **S. Elbassuoni, M. Ramanath, R. Schenkel, M. Sydow, G. Weikum:** *Language-model-based Ranking for Queries on RDF-Graphs*, CIKM 2009
- [6] **R. Neumayer, K. Balog, K. Nørvåg:** On the Modeling of Entities for Ad-Hoc Entity Search in the Web of Data, ECIR 2012

References

- [7] **U. Sawant and S. Chakrabarti:** *Learning Joint Query Interpretation and Response Ranking*, WWW 2013
- [8] **T. Tranh and P. Mika:** Semantic Search Systems, Concepts, Methods, and the Communities behind it, 2013
- [9] **M. Yahya, K. Berberich, S. Elbassuoni, M. Ramanath, V. Tresp, G. Weikum:** *Natural Language Questions for the Web of Data*, EMNLP 2012