



Rules and Facts

- Rules:
 - R1: $\text{Man}(x) \rightarrow \text{Mortal}(x)$
 - R2: $\text{Bird}(x) \rightarrow \text{CanFly}(x)$
- Facts:
 - F1: $\text{Man}(\text{Socrates})$
 - F2: $\text{Man}(\text{Plato})$
 - F3: $\text{Bird}(\text{Parrot}_A)$
- Typically used in expert systems.



QA Archive

- FAQ list
- Question Answering forums:
 - Yahoo! Answers
 - Stack Overflow
 - Quora
 - ...



Stack Overflow - Questions

Java - How can I disable a TLS cipher for only some protocols using JVM Config?

[Ask Question](#)

I've seen lots of examples of disabling TLS ciphers in java using `jdk.tls.disabledAlgorithms`, for example:

10

```
jdk.tls.disabledAlgorithms=MD2, RSA keySize < 1024, TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256
```

▼

But how can I disable a cipher for only certain protocols, using `jdk.tls.disabledAlgorithms` or a similar config?

4

For example, how can I disable `TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256` for `TLSv1.1` only?

It doesn't seem to support the `openssl` way of doing this, which is like so:

```
TLSv1.1:TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256
```

It doesn't cause any errors, but the cipher is still allowed.

EDIT: Note that I'm only really interested in JVM config based answers, as I don't control the code that's on lots of these servers, just the JVM and JVM configurations. Some are even 3rd party servers, so more of an ops level thing than anything.

EDIT 2: Note that you can run a java app and supply arguments that change which protocols and ciphers are used, e.g. `java -server -Djava.security.properties=/my/custom/java.security -jar myapp.jar` will do it - but it won't let you filter ciphers by protocol, only ciphers, or protocols, from what I can see. The file would contain a property entry like `jdk.tls.disabledAlgorithms`

`java` `java-8`

asked 7 days ago

viewed 149 times

active today

Looking for a job?

DevOps Engineer

M800 Limited 9 Kowloon, Hong Kong

automation devops

Front-end Developer

FITCH 9 Hong Kong

node.js reactjs

DevOps Engineer

Surevine 9 No office location

£40K - £50K 9 REMOTE

amazon-web-services ansible

Android Developer

Xapo 9 No office location



Stack Overflow - Answers

2 Answers

active oldest votes

2

[JSSE docs](#) say that the `https.protocols` property can store comma separated list of supported protocols in a given SSL context, however this property is used by current JSSE implementation, but could be disregarded by other vendors or future versions, so YMMV.

Programmatically you can achieve it like so:

```
SSLSocket socket = (SSLSocket) SSLSocketFactory.getDefault().createSocket();
socket.setEnabledCipherSuites(new String[] {
    CipherSuite.TLS_RSA_WITH_RC4_128_MDS.javaName,
    CipherSuite.TLS_RSA_WITH_RC4_128_SHA.javaName,
});

//allow TLS1.2 only
socket.setEnabledProtocols(new String[] {
    TlsVersion.TLS_1_2.javaName,
});
```

[share](#) [improve this answer](#)

[edited Oct 15 at 10:53](#)

[answered Oct 15 at 10:46](#)



diginoise

4,169 ⚡ 1 ⚡ 15 ⚡ 26

- 1 Thanks for the response... I'm mostly interested in config only options though, as I don't control the code that's on lots of these servers, just the JVM and JVM configurations. Some are even 3rd party servers, so more of an ops thing than anything. If I don't get any config based answers, I'll give it to you ;-) I've added this as a clarification to my question. Thanks! – [Brad Parks](#) Oct 16 at 10:43 ↗

[add a comment](#)





Knowledge Graph

*"The **Knowledge Graph** is a knowledge base used by Google to enhance its search engine's search results with semantic-search information gathered from a wide variety of sources."*

"A Knowledge graph (i) mainly describes real world entities and interrelations, organized in a graph (ii) defines possible classes and relations of entities in a schema" (iii) allows potentially interrelating arbitrary entities with each other... [Paulheim H.]

"We defines a Knowledge Graph as an RDF graph consists of a set of RDF triples where each RDF triple (s,p,o) is an ordered set of following RDF term" [Pujara J. al al.]

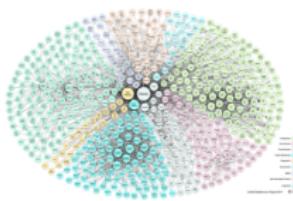
Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials



Knowledge Graph

LinkedIn
Knowledge Graph

 Freebase™



Amazon
Product Graph



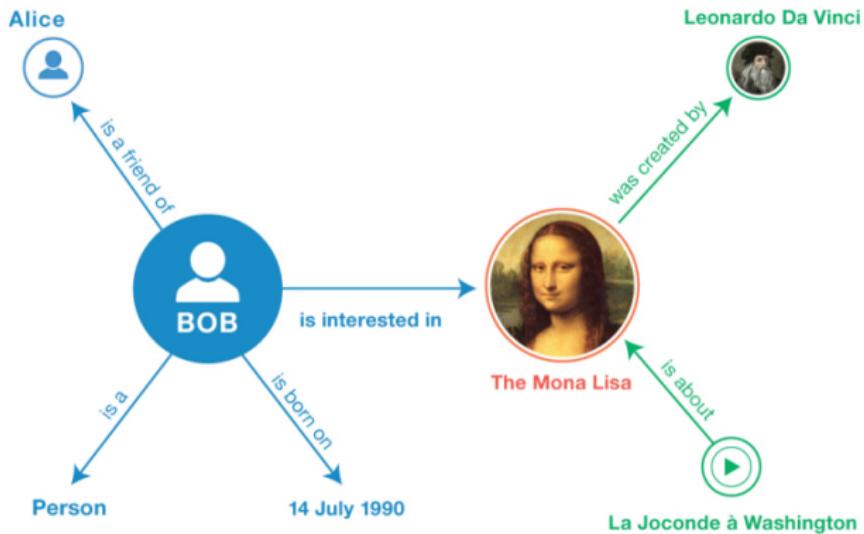
Facebook
Entity Graph

Microsoft
Satori

Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials



A (very small) Knowledge Graph



<http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/example-graph.jpg>



- The English version of the DBpedia knowledge base describes 4.58 million things, out of which 4.22 million are classified in a consistent [ontology](#), including 1,445,000 persons, 735,000 places (including 478,000 populated places), 411,000 creative works (including 123,000 music albums, 87,000 films and 19,000 video games), 241,000 organizations (including 58,000 companies and 49,000 educational institutions), 251,000 species and 6,000 diseases.
- In addition, we provide localized versions of DBpedia in 125 languages. All these versions together describe [38.3 million things](#), out of which 23.8 million are localized descriptions of things that also exist in the English version of DBpedia.



DBpedia - Linked Data Fragments

DBpedia 2016-04

Query DBpedia 2016-04 by triple pattern



subject:

predicate:

object:

"1899-05-06"^^<http://www.w3.org/2001/XMLSchema#date>

Find matching triples

Matches in DBpedia 2016-04 for { ?s ?p "1899-05-06"^^<http://www.w3.org/2001/XMLSchema#date> }

Showing triples 1 to 74 of 74 with 100 triples per page.

Billy_Cotton birthDate "1899-05-06".
Charlie_Irvis birthDate "1899-05-06".
Edward_Grahame_Johnstone birthDate "1899-05-06".
Eiliv_Austlid birthDate "1899-05-06".
Julio_Vega_Batlle birthDate "1899-05-06".
Kenneth_Hand birthDate "1899-05-06".
Ralyn_M._Hill birthDate "1899-05-06".
Tommy_Magee birthDate "1899-05-06".
Wally_Sieb birthDate "1899-05-06".
Alfred_Pillinger deathDate "1899-05-06".
Earley_F._Poppleton deathDate "1899-05-06".
Edward_Butterfield deathDate "1899-05-06".



DBpedia: Knowledge extraction

New York City

From Wikipedia, the free encyclopedia

Coordinates: 40°42'45"N 74°00'21"W

New York, New York, New York, often called **New York City** or simply **New York**, is the most populous city in the United States.^[1] With an estimated 2016 population of 8,537,875^{[2][3]} distributed over a land area of about 303.6 square miles (786 km²),^{[2][3][4]} New York City is also the center of the New York metropolitan area, one of the most populous urban agglomerations in the world,^{[2][3][4]} with an estimated 23.7 million residents as of 2016.^[5] A global power city,^[18] New York City has been described as the cultural, financial, and media capital^{[6][7][8]} of the world,^[19] and exerts a significant impact upon commerce,^[20] entertainment, research, technology, education, politics, and sports. The city's fast pace^{[23][24]} defines the term "New York time".^[25] Home to the headquarters of the United Nations,^[26] New York is an important center for international diplomacy.^[27]

Situated on one of the world's largest natural harbors,^{[28][29]} New York City consists of five boroughs, each of which is a separate county of New York State.^[30] The five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island – were consolidated into a single city in 1898.^[31] The city and its metropolitan area constitute the premier gateway for legal immigration to the United States,^[32] and as many as 800 languages are spoken in New York,^{[33][34][35]} making it the most linguistically diverse city in the world.^{[34][35][37]} New York City is home to more than 3.2 million residents born outside the United States,^[38] the largest foreign-born population of any city in the world.^[36] In 2013, the tri-state New York Metropolitan Statistical Area (MSA) produced a gross metropolitan product (GMP) of nearly US\$1.4 trillion.^[40] If New York City were a country, it would have the 12th highest GDP in the world.^[41]

New York City traces its origins to a trading post founded by colonists of the Dutch Republic in 1624 on Lower Manhattan; the post was named New Amsterdam in 1624.^[42] The city and its surroundings came under English control in 1664^[43] and were renamed New York after King Charles II of England granted the lands to his brother, the Duke of York.^[43] New York served as the capital of the United States from 1789 until 1790.^[44] It has been the country's largest city since 1790.^[45] The Statue of Liberty greeted millions of immigrants as they came to America by ship in the late 19th and early 20th centuries,^[46] and is a world symbol of the United States and its ideals of liberty and peace.^[47] In the 21st century, New York has emerged as a global node of creativity and entrepreneurship,^[48] social media,^[49] and environmental sustainability.^{[50][51]} and as a symbol of freedom and cultural diversity.^[52]

Many districts and landmarks in New York City are well known around the world.^[53] In 2016,^[54] holding three of the world's ten most visited tourist attractions in 2016,^[55] several sources have ranked New York the most photographed city in the world.^[56] The Empire State Building, located at the city's "geographic center",^[57] is the fourth tallest building in the world.^[58] The Brooklyn Bridge, one of the world's busiest pedestrian crossings,^{[59][60]} and a major center of the world's intermodal shipping industry.^[61] The names of many of the city's landmarks, synagogues,^[62] and parks^[63] are known around the world. Anchored by Wall Street, the financial district of Lower Manhattan, the New York Stock Exchange and NASDAQ,^{[67][68]} Manhattan's most active market has been called both the most economically powerful city and the leading financial center of the world.^{[69][70]} The city is home to the world's largest stock exchanges by total listed capitalization, the New York Stock Exchange and NASDAQ.^{[67][68]} Manhattan's most active market is experiencing the most expansion in the world.^{[69][71]} Chinatown incorporates the highest concentration of Chinese people in the Western Hemisphere.^{[72][73]} With multiple signature attractions developing across the city,^{[73][74]} providing continuous 24/7 service,^[75] the New York City Subway is one of the most extensive metro systems worldwide, with 472 stations in operation.^{[76][77][78]} Over 120 colleges and universities are located in New York City, including Columbia University, New York University, and Rockefeller University, which have been ranked among the top universities in the world.^{[79][80]}

The City of New York, often called **New York City** or simply **New York**, is the most populous city in the United States.

<New York City>, <CityIn> <United States>.

<City Name>, <locatedIn> <Country Name>.

1	History
1.1	Dynasty
1.2	Earl
1.3	Duke
1.4	English rule
1.5	American Revolution
1.6	Nineteenth century
1.7	Modern history
2	Geography
2.1	Cityscape
2.2	Architecture
2.3	Boroughs
2.4	Climate
2.5	Parks
2.5.1	National parks
2.5.2	State parks
2.5.3	City parks
2.6	Military installations
3	Demographics
3.1	Population density
3.2	Race and ethnicity
3.3	Sexual orientation and gender identity
3.3.1	Transgender contribution
3.4	Religion
3.5	Income
4	Economy
4.1	City economic overview
4.2	Wall Street
4.3	Global cities

New York City	
City	
City top	Midtown Manhattan, Times Square, the Ushers in Queens, the Brooklyn Bridge, the Empire State Building, the One World Trade Center, Central Park, the headquarters of the United Nations, and the Statue of Liberty
Flag	
Nickname(s)	See Nicknames of New York City
Location within the U.S.	
Map	
Show map of the U.S.	Show map of the U.S.
Show all	Show all
Location in the contiguous United States and New York	Location in the contiguous United States and New York
Coordinates:	40°42'45"N 74°00'21"W ^[1]
Country	United States
State	New York
Counties / Boroughs	Bronx, Kings (Brooklyn), New York (Manhattan), Queens, Richmond (Staten Island)
Historic colonies	New Netherland, Province of New York
Settled	1624
Consolidated	1898
Named for	James, Duke of York
Governance ^[2]	Mayor–Council New York City Council Boroughs of the City
Area ^[3]	Total: 469.464 sq mi (1,213.07 km ²) Land: 302.643 sq mi (783.84 km ²) Water: 165.841 sq mi (430.23 km ²) Metro: 13,318 sq mi (34,490 km ²)

Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials





DBpedia: Knowledge extraction

IBM

From Wikipedia, the free encyclopedia
(Redirected from [bm](#))

For other uses of IBM, see [IBM \(disambiguation\)](#). "Big Blue" redirects here. It is not to be confused with [New York Giants](#). For other uses of Big Blue, see [Big Blue \(disambiguation\)](#).

IBM ([International Business Machines Corporation](#)) is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries. The company originated in 1911 as the [Computing-Tabulating-Recording Company](#) (CTR) and was renamed "International Business Machines" in 1924.

IBM manufactures and markets computers hardware, middleware and software, and offers hosting and consulting services in areas ranging from mainframe computers to nanotechnology. IBM is also a major research organization, holding the record for most patents generated by a business (as of 2017) for 24 consecutive years.^[2] Inventions by IBM include the automated teller machine (ATM), the PC, the floppy disk, the hard disk drive, the magnetic stripe card, the relational database, the SQL programming language, the UPC barcode, and dynamic random-access memory (DRAM). The IBM mainframe, exemplified by the System/360, was the dominant computing platform during the 1960s and 1970s.

IBM has continually shifted its business mix by commoditizing markets focusing on higher-value, more profitable markets. This includes spinning off printer manufacturer Lexmark in 1991 and selling off its personal computer (ThinkPad/ThinkCentre) and x86-based server businesses to Lenovo (2005 and 2014, respectively), and acquiring companies such as PwC Consulting (2002), SPSS (2009), and The Weather Company (2016). Also in 2014 IBM announced that it would go "fabless", continuing to design semiconductors, but offloading manufacturing to GlobalFoundries.

Nicknamed **Big Blue**, IBM is one of 30 companies included in the Dow Jones Industrial Average and one of the world's largest employers, with (as of 2018) nearly 380,000 employees. Known as "IBMers", IBM employees have been awarded six Nobel Prizes, six Turing Awards, ten National Medals of Technology and five National Medals of Science.

Contents [edit]	
1	History
2	Headquarters and offices
3	Products and services
4	Research
5	Brand and reputation
6	People and culture
6.1	Employees
6.1.1	IBM alumni
6.2	Board and shareholders
7	See also
8	References
9	Further reading
10	External links

History [edit]

[Main article: History of IBM](#)

In the 1880s, technologies emerged that would ultimately form the core of International Business Machines (IBM). Julius E. Pirnat patented the computing scales in 1886.^[6] Alexander Dey invented the dial recorder (1888).^[7] Herman Hollerith patented the [Electric Tabulating Machine](#),^[8] and Willard Bundy invented a time clock to record a worker's arrival and departure time on a paper tape in 1889.^[9] On June 16, 1911, their four companies were amalgamated in New York State by Charles Flint forming a fifth company, the Computing-Tabulating-Recording Company (CTR) based in Endicott, New York.^{[10][11]} The five companies had 1,300 employees and offices and plants in Endicott and Binghamton, New York; Dayton, Ohio; Detroit, Michigan; Washington, D.C.; and Toronto. They manufactured machinery for sale and lease, ranging from commercial scales and industrial time recorders, meat and cheese slicers, to tabulators and punched cards. Thomas J. Watson, Sr., fired from the National Cash Register Company by John Henry Patterson, called on Flint and, in 1914, was offered CTR.^[12] Watson joined CTR as General Manager then, 11 months later, was made President when court cases relating to his time at NCR were resolved.^[13] Having learned Patterson's pioneering business practices, Watson proceeded to put the stamp of NCR onto CTR's companies.^[13] He implemented sales conventions, "generous sales incentives, a focus on customer service, an insistence on well-groomed, dark-suited salesmen and had an evangelical fervor for instilling company pride and loyalty in every worker".^{[14][15]} His favorite slogan, "THINK", became a mantra for the company's employees.^[14] During Watson's first four years, revenues reached \$9 million and the company's operations expanded to Europe, South America, Asia and Australia.^[14] "Watson had never liked the clumsy hyphenated title of the CTR and in 1924 chose to replace it with the more expansive

Wikipedia Infobox

International Business Machines Corporation



IBM Watson system in 2011

Type	Public
Traded as	NYSE: IBM ^[2] DAX Component FTSE 100 Component S&P 500 Component US\$450B(2010)14
Industry	Cloud computing - Cognitive computing
Founded	June 16, 1911; 108 years ago
Founder	Computing-Tabulating-Recording Company Endicott, New York, U.S. ^[1] Charles Ranlett Flint
Headquarters	America, New York, U.S.
Area served	177 countries ^[2]
Key people	Ginn Rosette (Chairwoman, President and CEO) See IBM products
Revenue	▼ US\$ 79.919 billion (2016) ^[2]
Operating income	▼ US\$ 13.031 billion (2016) ^[2]
Net income	▼ US\$ 11.872 billion (2016) ^[2]
Total assets	▲ US\$ 117.47 billion (2016) ^[2]
Total equity	▲ US\$ 18.362 billion (2016) ^[2]
Number of employees	380,000 (2016) ^[4]
Website	www.ibm.com/

Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials



DBpedia: Knowledge extraction

Ontology Classes

- owl:Thing
 - Activity (edit)
 - BoardGame (edit)
 - CardGame (edit)
 - Sales (edit)
 - Sport (edit)
 - Athletics (edit)
 - Boxing (edit)
 - BoxingCategory (edit)
 - BoxingStyle (edit)
 - HorseRiding (edit)
 - TeamSport (edit)
 - Soccer (edit)
 - Agent (edit)
 - Deity (edit)
 - Employee (edit)
 - Family (edit)
 - NobleFamily (edit)
 - FictionalCharacter (edit)
 - ComicsCharacter (edit)
 - AnimangaCharacter (edit)
 - DisneyCharacter (edit)
 - MythologicalFigure (edit)
 - NarutoCharacter (edit)
 - SoapCharacter (edit)
 - Organisation (edit)
 - Broadcaster (edit)
 - BroadcastNetwork (edit)
 - RadioStation (edit)
 - TelevisionStation (edit)

Organisation (edit)

- Broadcaster (edit)
 - BroadcastNetwork (edit)
 - RadioStation (edit)
 - TelevisionStation (edit)
- Company (edit)
 - Bank (edit)
 - Brewery (edit)
 - Caterer (edit)
 - LawFirm (edit)
 - PublicTransitSystem (edit)
 - Airline (edit)
 - BusCompany (edit)
 - Publisher (edit)
 - RecordLabel (edit)
 - Winery (edit)
- EducationalInstitution (edit)
 - College (edit)
 - Library (edit)
 - School (edit)
 - University (edit)
- EmployersOrganisation (edit)
- GeopoliticalOrganisation (edit)
- GovernmentAgency (edit)
 - GovernmentCabinet (edit)
- Group (edit)
 - Band (edit)
 - ComedyGroup (edit)

Person (edit)

- Archeologist (edit)
- Architect (edit)
- Aristocrat (edit)
- Artist (edit)
 - Actor (edit)
 - AdultActor (edit)
 - VoiceActor (edit)
- Comedian (edit)
- ComicsCreator (edit)
- Dancer (edit)
- FashionDesigner (edit)
- Humorist (edit)
- MusicalArtist (edit)
 - BackScene (edit)
 - ClassicalMusicArtist (edit)
 - Instrumentalist (edit)
 - Guitarist (edit)
 - MusicDirector (edit)
 - Singer (edit)
 - Painter (edit)
 - Photographer (edit)
 - Sculptor (edit)
- Astronaut (edit)
- Athlete
 - ArcherPlayer (edit)
 - AthleticsPlayer (edit)

Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials



DBpedia: Knowledge extraction

Properties on Actor:

Name	Label	Domain	Range
academyAward (edit)	Academy Award	Artist	Award
afiAward (edit)	AFI Award	Artist	Award
arielAward (edit)	Ariel Award	Actor	Award
associatedAct (edit)	associated act	Artist	Artist
baftaAward (edit)	BAFTA Award	Artist	Award
cesarAward (edit)	Cesar Award	Artist	Award
disciple (edit)	disciple	Artist	Artist
dutchRKDCode (edit)	Dutch RKD code	Artist	xsd:string
emmyAward (edit)	Emmy Award	Artist	Award
field (edit)	field	Artist	owl:Thing
filmFareAward (edit)	Film Fare Award	Artist	Award
gaudiAward (edit)	Gaudí Award	Artist	Award
geminiAward (edit)	Gemini Award	Actor	Award
goldenCalfAward (edit)	Golden Calf Award	Actor	Award
goldenGlobeAward (edit)	Golden Globe Award	Artist	Award
goldenRaspberryAward (edit)	Golden Raspberry Award	Actor	Award
goyaAward (edit)	Goya Award	Artist	Award
grammyAward (edit)	Grammy Award	Artist	Award
iftaAward (edit)	IFTA Award	Actor	Award

Nitish Aggarwal, et al., Knowledge Graphs: In Theory and Practice, CIKM17 Tutorials



Content

3

Question Answering

- System structure
- Knowledge sources
- Techniques: knowledge representation
- Techniques: question understanding
- Techniques: answer generation
- Open domain QA
- Machine reading comprehension (MRC)



Knowledge representation

- => Linguistic Annotations
- => Statistical Representation
- => Neural Representation

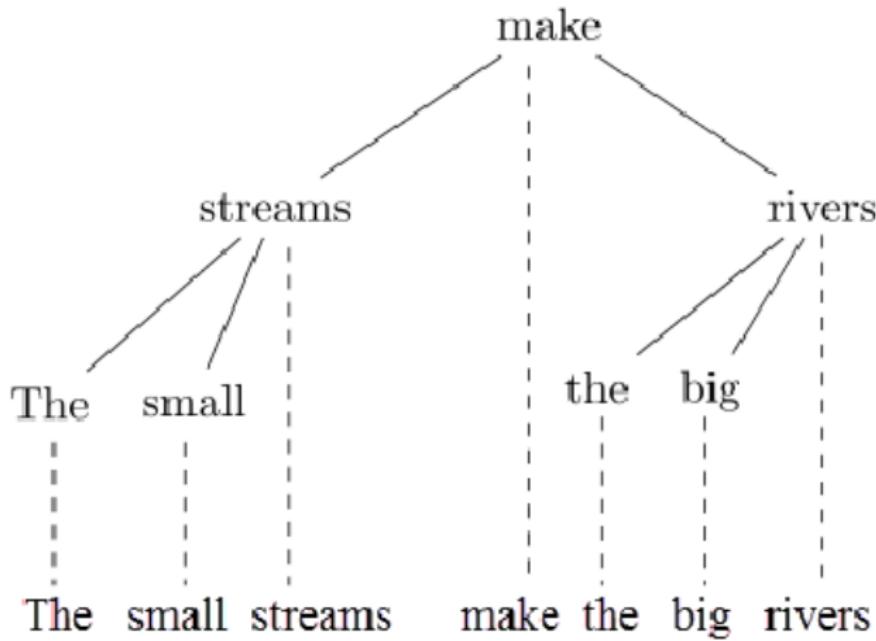


Linguistic annotations

- Part-of-Speech Tagging
- Syntactic Parsing: Constituent or Dependency
- Named Entities
- Entity Relations
- Events
- Anaphora, Co-Reference and Mentions
- Rhetorical Structures
- Semantic Role Labelling
- Word Senses
- Sentiments



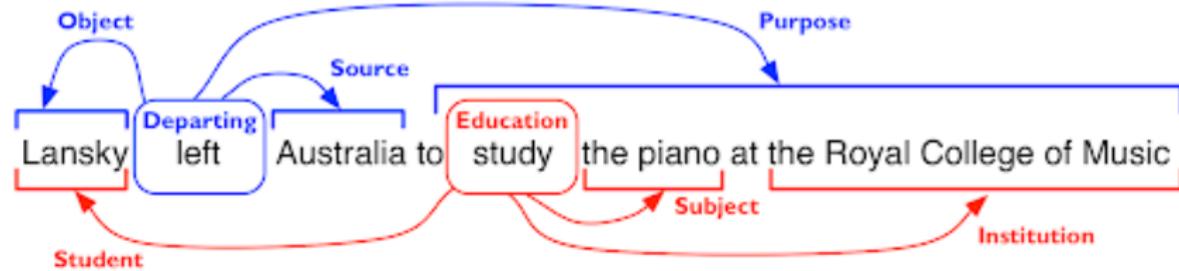
Dependency trees



Haitao Liu, et al. "Dependency distance: a new perspective on syntactic patterns in natural languages." Physics of life reviews 21 2017



Semantic role labeling



<http://ivan-titov.org/topics/semantics.html>

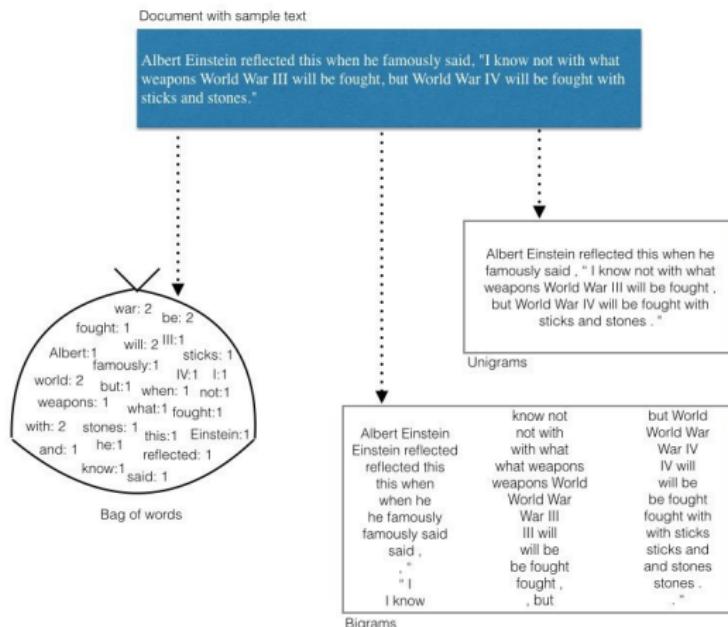


Knowledge representation

- => Linguistic Annotations
- => Statistical Representation
- => Neural Representation



Bag-of-Words and Bag-of-Ngrams



<https://qph.fs.quoracdn.net/main-qimg-c47060d2f02439a44795e2fbef2ca347.webp>



Vector Space Model and TF-IDF

- Vector space model represent a document as a high dimensional vector rather than a set (bag).
- TF-IDF is the most commonly used weighing schema in VSM in information retrieval
 - TF: term frequency
 - IDF: inverse document frequency

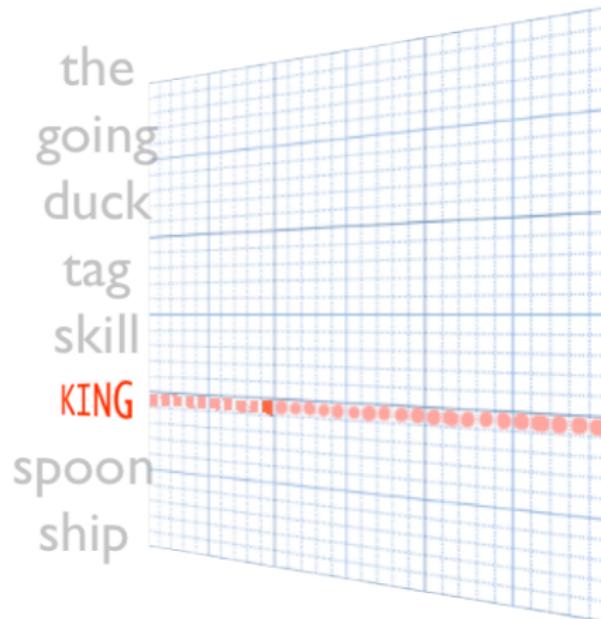


Knowledge representation

- => Linguistic Annotations
- => Statistical Representation
- => Neural Representation

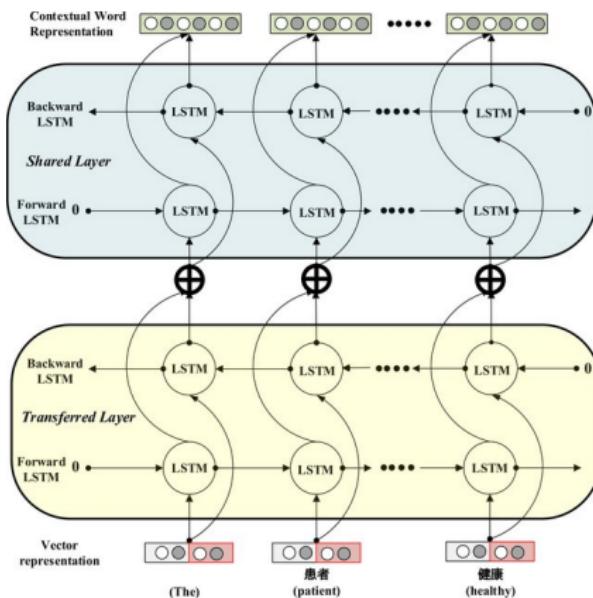


Word embeddings





Contextual word embeddings



Dong et al., Deep learning for named entity recognition on Chinese electronic medical records: Combining deep transfer learning with multitask bi-directional LSTM RNN, PLoS ONE 14(5):e0216046, 2019



Sentence embeddings

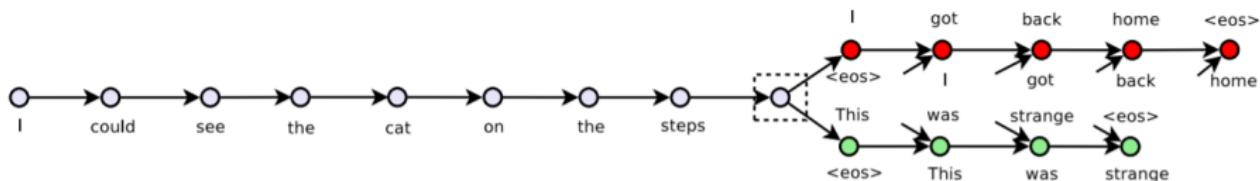


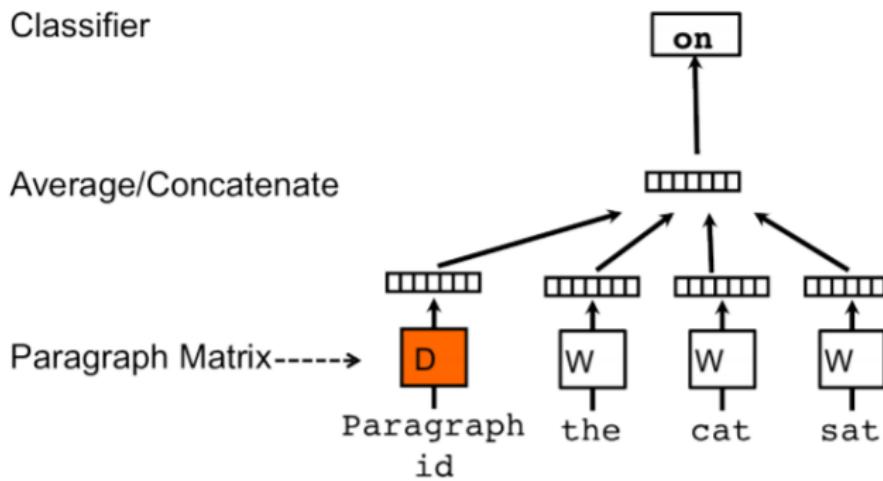
Figure 1: The skip-thoughts model. Given a tuple (s_{i-1}, s_i, s_{i+1}) of contiguous sentences, with s_i the i -th sentence of a book, the sentence s_i is encoded and tries to reconstruct the previous sentence s_{i-1} and next sentence s_{i+1} . In this example, the input is the sentence triplet *I got back home. I could see the cat on the steps. This was strange*. Unattached arrows are connected to the encoder output. Colors indicate which components share parameters. $\langle \text{eos} \rangle$ is the end of sentence token.

Kiros, Ryan, et al. "Skip-thought vectors." Advances in neural information processing systems. 2015



Document embeddings

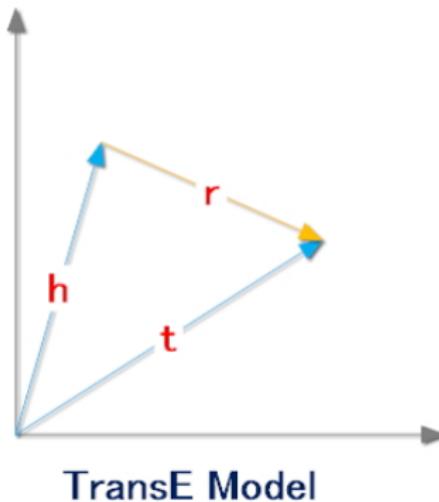
Doc2Vec



A gentle introduction to Doc2Vec



Knowledge embeddings



Bordes, et al., Translating embeddings for modeling multi-relational data, Adv. Neural Inf. Process. Syst., 2013



Content

3

Question Answering

- System structure
- Knowledge sources
- Techniques: knowledge representation
- Techniques: question understanding**
- Techniques: answer generation
- Open domain QA
- Machine reading comprehension (MRC)



Question understanding

- Pattern Matching
- Similar Question Retrieval
- Question Paraphrasing
- Question Classification and Slot Filling
- Semantic Parsing



Pattern matching

- Example: Who is X of Y?
 - Who is the first president of US?
 - Who is the second husband of Wendi Deng?
- Accurate Understanding
- Low Coverage
- Only suitable for very small domain

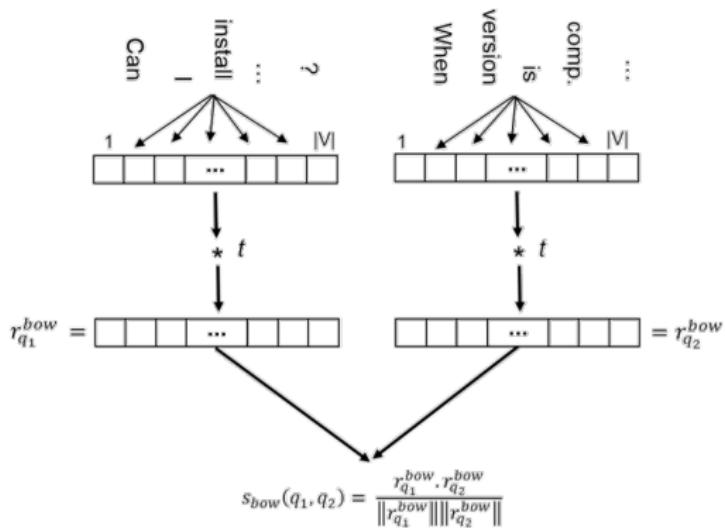


Equivalent Question Retrieval

- Basic ideas:
 - Retrieve existing questions with the query using an IR technologies;
 - Select the best one from the top ranked question returned by the IR system, using some similarity metrics;
- Can be used in scenarios like community QA or open-domain QA.



Equivalent Question Retrieval

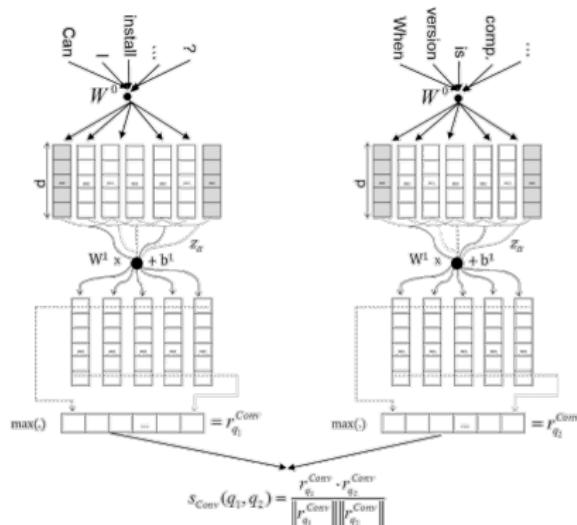


Question retrieval with bag-of-words similarity

Dos Santos, et al., Learning hybrid representations to retrieve semantically equivalent questions. ACL-IJCNLP2015



Equivalent Question Retrieval



Question retrieval with convolutional similarity

Dos Santos, et al., Learning hybrid representations to retrieve semantically equivalent questions. ACL-IJCNLP2015



Question paraphrasing

Question Paraphrasing is effective for most QA tasks.

Question

q: *who created microsoft?*



Paraphrases

q₁: *who founded microsoft?*

q₂: *who is the founder of microsoft?*

q₃: *who is the creator of microsoft?*

...

q_m: *who designed microsoft?*

Li Dong, et al., Learning to Paraphrase for Question Answering, arXiv:1708.06022, 2017



Question classification

- Classification by answer types:
 - Factual questions: What is the largest city ... ?
 - Opinions: What is the authors' attitude ... ?
 - Summaries: What are the arguments for and against ... ?
- Classification by speech acts:
 - Yes/No questions: Is it true that ... ?
 - WH questions: Who was the first president ... ?
 - Indirect request: I would like you to list
 - Commands: Name all presidents
- Complex / difficult questions:
 - Why/How questions: How to build a house?
 - What questions: What did they do?



Question classification

Class	#	Class	#
ABBREV.	9	description	7
abb	1	manner	2
exp	8	reason	6
ENTITY	94	HUMAN	65
animal	16	group	6
body	2	individual	55
color	10	title	1
creative	0	description	3
currency	6	LOCATION	81
dis.med.	2	city	18
event	2	country	3
food	4	mountain	3
instrument	1	other	50
lang	2	state	7
letter	0	NUMERIC	113
other	12	code	0
plant	5	count	9
product	4	date	47
religion	0	distance	16
sport	1	money	3
substance	15	order	0
symbol	0	other	12
technique	1	period	8
term	7	percent	3
vehicle	4	speed	6
word	0	temp	5
DESCRIPTION	138	size	0
definition	123	weight	4

The distribution of 500 TREC 10 questions over the question hierarchy.

Xin Li and Dan Roth. "Learning question classifiers." COLING 2002.



Slot filling

- Question classification gives a fine grained understanding of user's intent.
- Give a class of questions, there are slots to be filled to give the details of the question.
- Example:
 - Input: When did the Second World War break out?
 - Output:
 - Classification: TIME
 - Slots:
 - Events: BREAK_OUT
 - Subjects: the Second World War



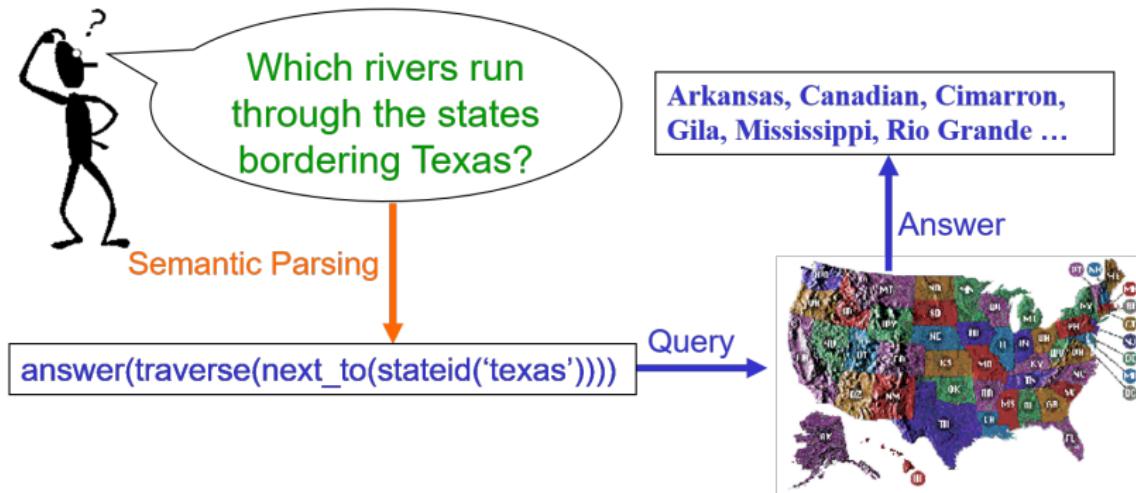
Semantic parsing

- Convert a natural language sentence to a semantic representation, for example:
 - Logic forms
 - Lambda Calculus
 - SQL
 - SPARQL
 - Programming language scripts: Prolog, Lisp, Python, etc.



Semantic parsing: An example

- Query application for U.S. geography database containing about 800 facts [Zelle & Mooney, 1996]



Raymond J. Mooney, Semantic Parsing for Question Answering (slides).



Content

3

Question Answering

- System structure
- Knowledge sources
- Techniques: knowledge representation
- Techniques: question understanding
- **Techniques: answer generation**
- Open domain QA
- Machine reading comprehension (MRC)



Answer generation

- Answer retrieval
- Answer extraction
- Answer ranking
- Database querying
- Knowledgebase query
- Rule-based inference engine
- Inference over knowledge graphs
- Answer Generation



Answer retrieval

- Find the paragraphs which are related to the question, from a large collection text
- Approaches:
 - Build a paragraph-based search engine.
 - Use an existing search engine to find a list of documents, then find the most relevant paragraphs
- A subtask for Web-based QA (Open domain QA)
- Technologies: IR



Answer ranking

- When multiple answer candidates are obtained / generated, a ranking algorithm will be used to find the best answer.
- Technology: Learning to rank

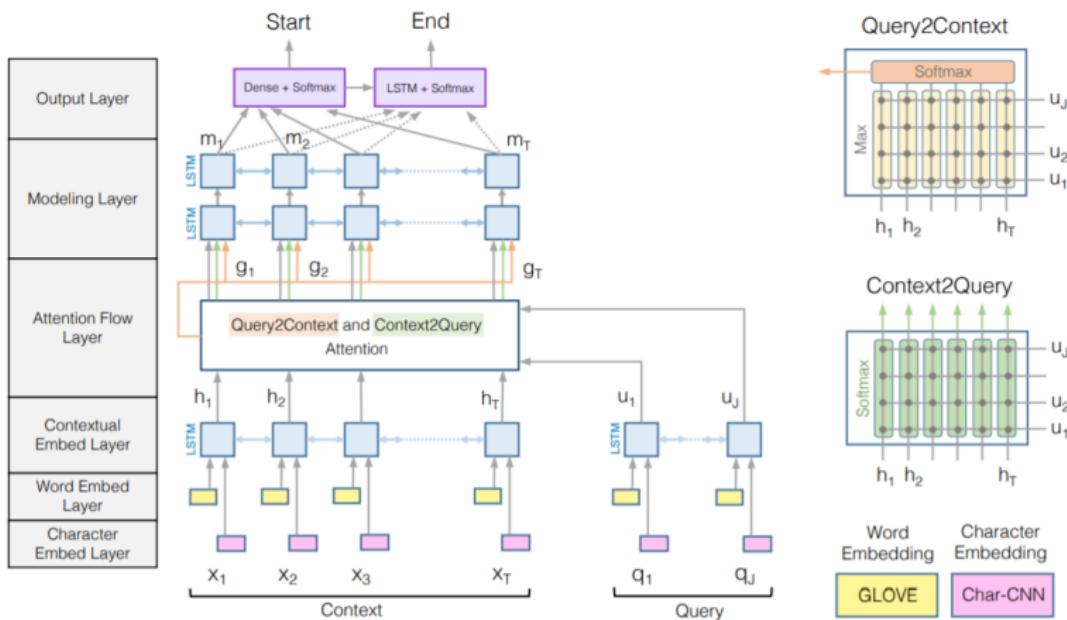


Answer extraction

- From a given piece of text, extract the word or phrase which can answer the question exactly.
- The task could be implemented by finding the start word and the end word of the answer in the text.
- Answer extraction is used in Open Domain QA and Machine Reading Comprehension.



Answer extraction



Seo, Min Joon et al. Bidirectional Attention Flow for Machine Comprehension. CoRR abs/1611.01603, 2016



Database querying

- For NLIDB:

- The user's question is converted to a database query (for example, a SQL statement) by the question understanding module.
- The query is then sent to the database which returns a result.
- The result is sent back to the user as the answer, possibly after post-processing by an answer generation module.



Knowledgebase querying

- For KBQA or KGQA:

- The user's question is converted to a knowledge base (or knowledge graph) query (for example, a SparQL statement) by the question understanding module.
- The query is then sent to the knowledge base (or knowledge graph) which returns a result.
- The result is sent back to the user as the answer, possibly after post-processing by an answer generation module.



SparQL Query

SPARQL Query

Query:

```
17 PREFIX ocrer: <http://purl.org/net/OCRe/research.owl#>
18 PREFIX ocrestd: <http://purl.org/net/OCRe/study_design.owl#>
19 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
20 PREFIX vcard: <http://www.w3.org/2006/vcard/ns#>
21 PREFIX vitro-public: <http://vitro.mannlib.cornell.edu/ns/vitro/public#>
22 PREFIX vivo: <http://vivoweb.org/ontology/core#>
23 PREFIX scires: <http://vivoweb.org/ontology/scientific-research#>
24 PREFIX core: <http://vivoweb.org/ontology/core#>
25
26 Select ?s ?p ?o
27 where {
28   ?s a vivo:Relationship .
29 }
30 ORDER BY ?s
31
32
```



Format for SELECT and ASK query results:

- RS_TEXT CSV TSV RS_XML RS_JSON

Format for CONSTRUCT and DESCRIBE query results:

- N-Triples RDF/XML N3 Turtle JSON-LD

Run Query

href`https://wiki.lyrasis.org/display/VIVODOC110x/SPARQL+Queries`VIVO 1.10.x Documentation: SPARQL Queries



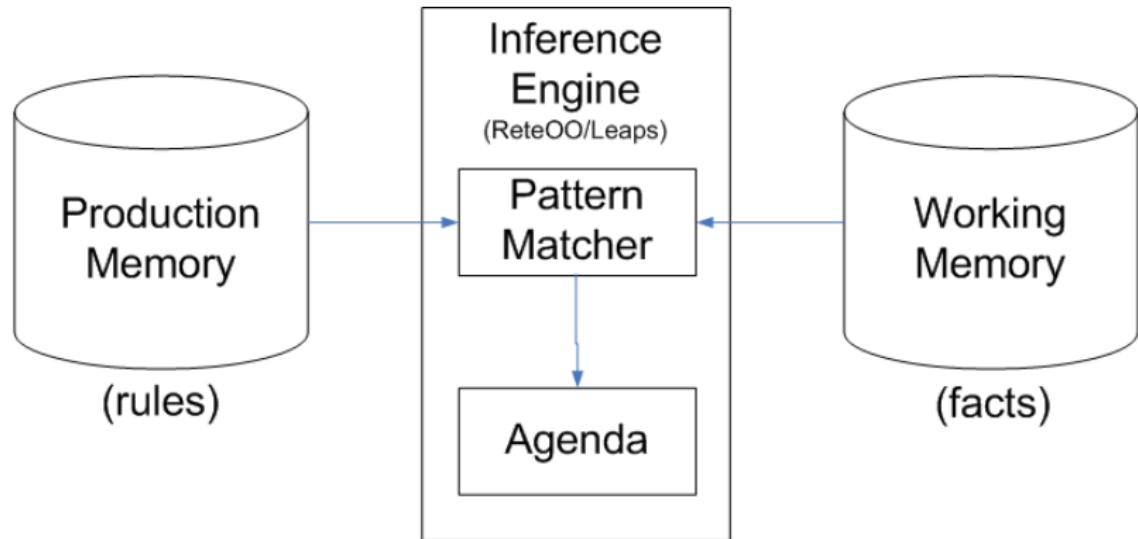
Rule-based inference engine

- Inference engine is a component of the system that applies logical rules to the knowledge base to deduce new information.
- The first inference engines were components of expert systems. The typical expert system consisted of a knowledge base and an inference engine.
- The knowledge base stored facts about the world.
- The inference engine applies logical rules to the knowledge base and deduced new knowledge.
- This process would iterate as each new fact in the knowledge base could trigger additional rules in the inference engine.

Wikipedia: Inference Engine



Rule-based inference engine



Fernandes et al., A rule-based system proposal to aid in the evaluation..., arXiv:1811.12454