WARWICK

CS918: LECTURE 18

Question Answering

Arkaitz Zubiaga, 3rd December, 2018

LECTURE 18: CONTENTS

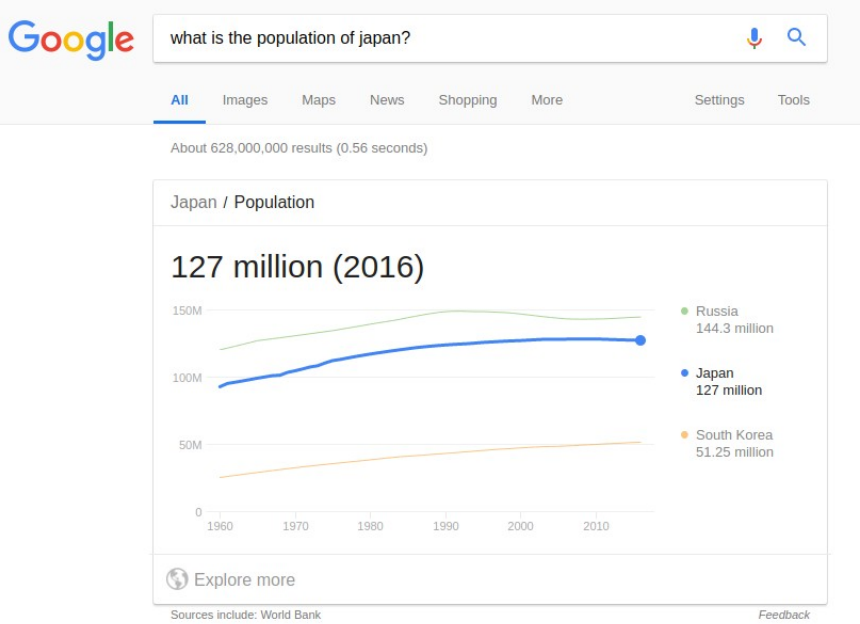
- What is Question Answering?
  - Paradigms for Question Answering.
- Answer Types and Query Formulation.
- Passage Retrieval and Answer Extraction.
- Using Knowledge in Question Answering.

WHAT IS QUESTION ANSWERING?

- NLP task, also relevant to Information Retrieval.
- **Question answering (QA):** systems that automatically answer questions posed by humans in a natural language.
- Questions can be of **many different types**.

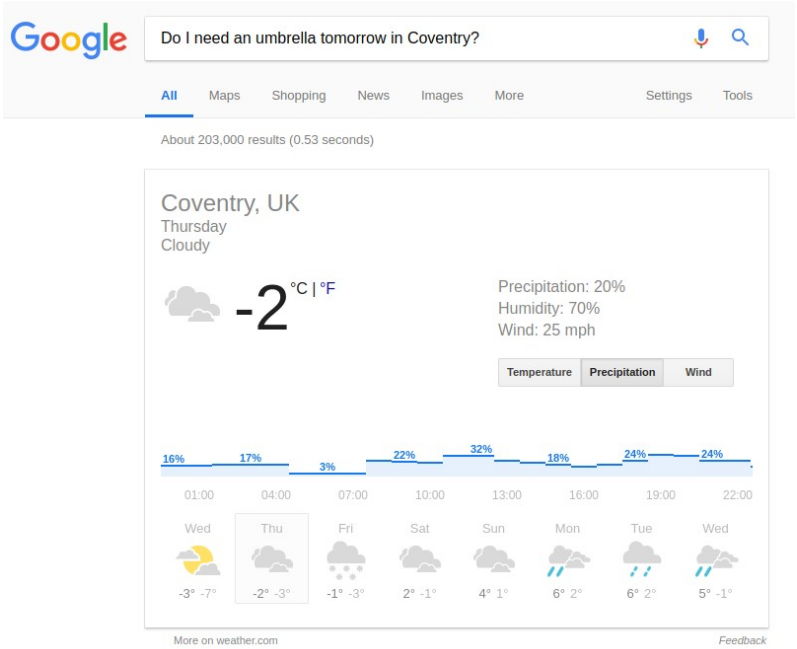
QA: GOOGLE

- We get **direct answers** sometimes.

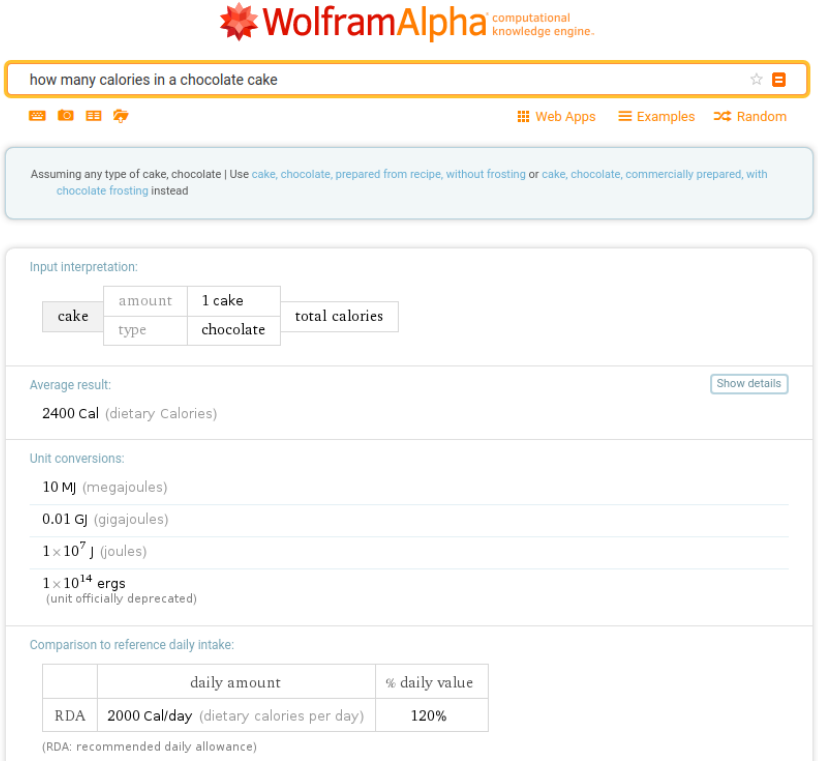


QA: GOOGLE

- Or just the **relevant information** sometimes.



QA: WOLFRAM ALPHA



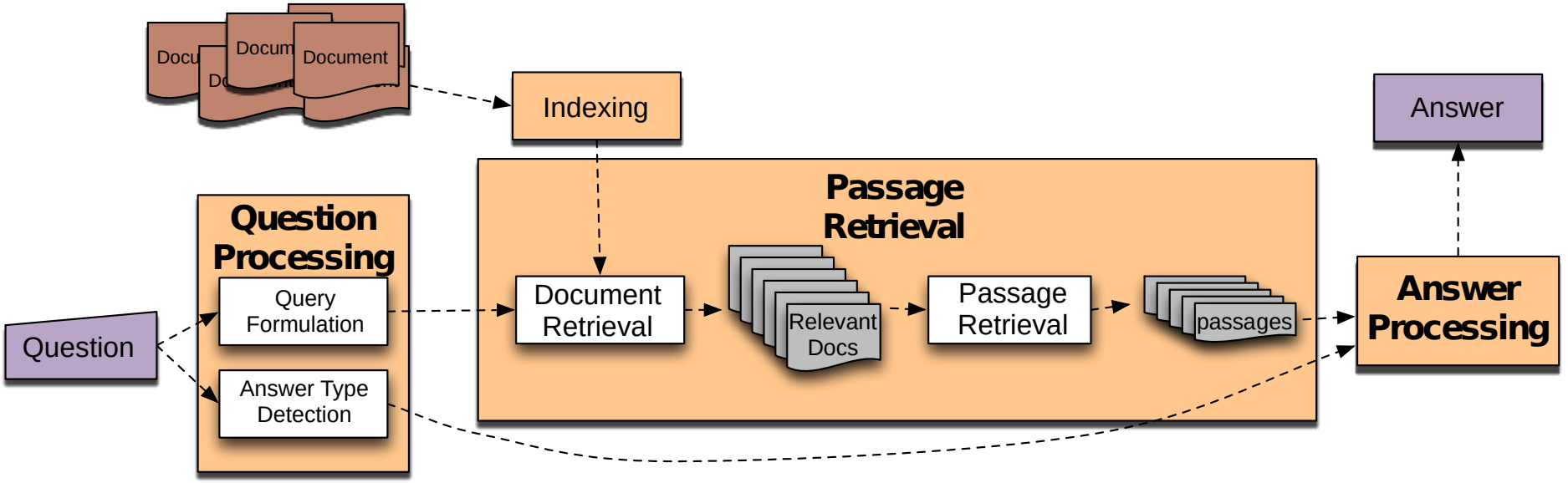
TYPES OF QUESTIONS

- **Factoid questions:**
  - Who wrote “The Universal Declaration of Human Rights”?
  - How many calories are there in two slices of apple pie?
  - Where is Apple Computer based?
- **Complex (narrative) questions:**
  - In children with an acute febrile illness, what is the efficacy of acetaminophen in reducing fever?
  - What do scholars think about Jefferson’s position on dealing with pirates?

TWO PARADIGMS FOR QA

- **IR-based approaches:** answers based on web search.
  - TREC; IBM Watson; Google;
- **Knowledge-based and Hybrid approaches:** answers based on databases, or hybrid.
  - IBM Watson; Apple Siri; Wolfram Alpha;

IR-BASED QA


Docu
Docum
Document
Indexing
Answer
Passage Retrieval
Question Processing
Query Formulation
Document Retrieval
Relevant Docs
Passage Retrieval
passages
Question
Answer Type Detection
Answer Processing

IR-BASED QA

- **Question processing:**
  - Detect question type, answer type, focus, relations.
  - Formulate queries to send to a search engine.
- **Passage retrieval:**
  - Retrieve ranked documents.
  - Break into suitable passages and rerank.
- **Answer processing:**
  - Extract candidate answers.
  - Rank candidates using evidence from the text and external sources.

KNOWLEDGE-BASED QA

- **Build a semantic representation of the query:**
  - Times, dates, locations, entities, numeric quantities.
- **Map from this semantics to query structured data or resources:**
  - Geospatial databases.
  - Ontologies (Wikipedia infoboxes, dbPedia, WordNet, Yago).
  - Restaurant review sources and reservation services.
  - Scientific databases.

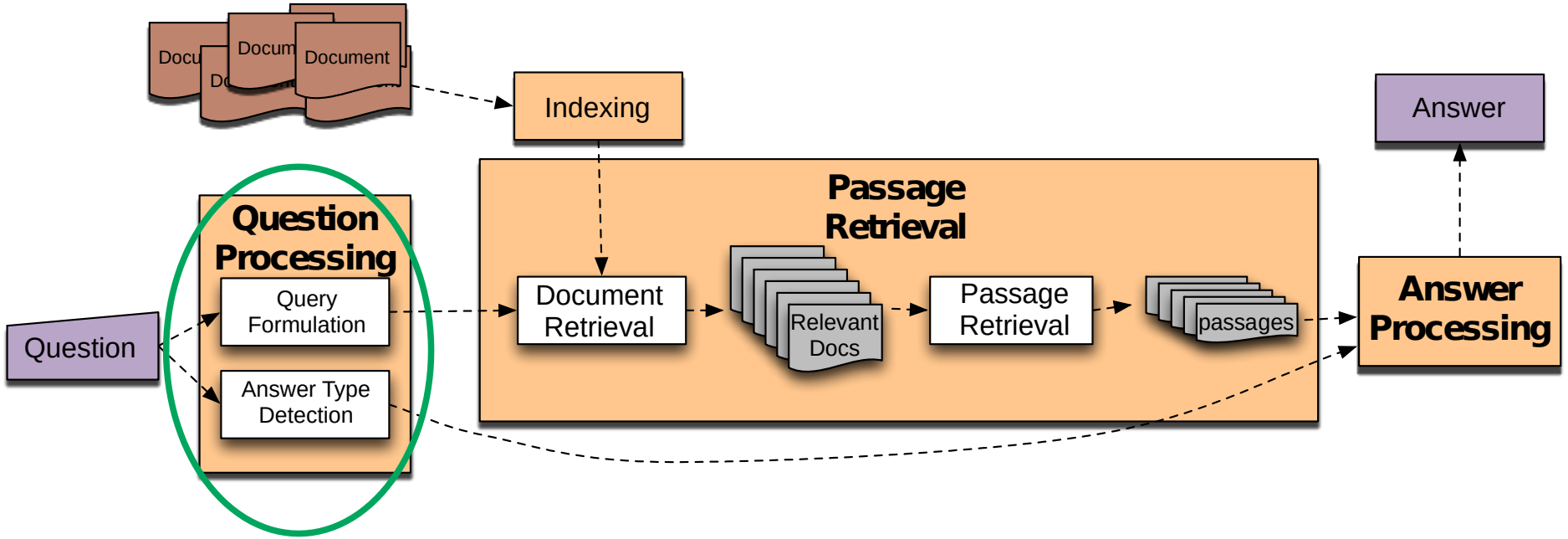
HYBRID APPROACHES TO QA

- **Build a shallow semantic representation of the query:**
- **Generate answer candidates using IR methods:**
  - Augmented with ontologies and semi-structured data.
- **Score each candidate using richer knowledge sources:**
  - Geospatial databases.
  - Temporal reasoning.
  - Taxonomical classification.

WARWICK

ANSWER TYPES AND QUERY FORMULATION

IR-BASED QA


Document
Indexing
Answer
Question Processing
Query Formulation
Answer Type Detection
Question
Passage Retrieval
Document Retrieval
Relevant Docs
Passage Retrieval
passages
Answer Processing

QUESTION PROCESSING

- **Answer Type Detection:**
  - Find the named entity type (person, place) of the answer.
- **Query Formulation:**
  - Choose query keywords for the IR system.
- **Focus Detection:**
  - Find the question words that are replaced by the answer.
- **Relation Extraction:**
  - Find relations between entities in the question.

QUESTION PROCESSING: EXAMPLE

- Example: What are the countries bordering France?

- Answer type: countries of the World
- Query: countries bordering france
- Focus: the countries
- Relations: border(France, ?x)

ANSWER TYPE DETECTION: NAMED ENTITIES

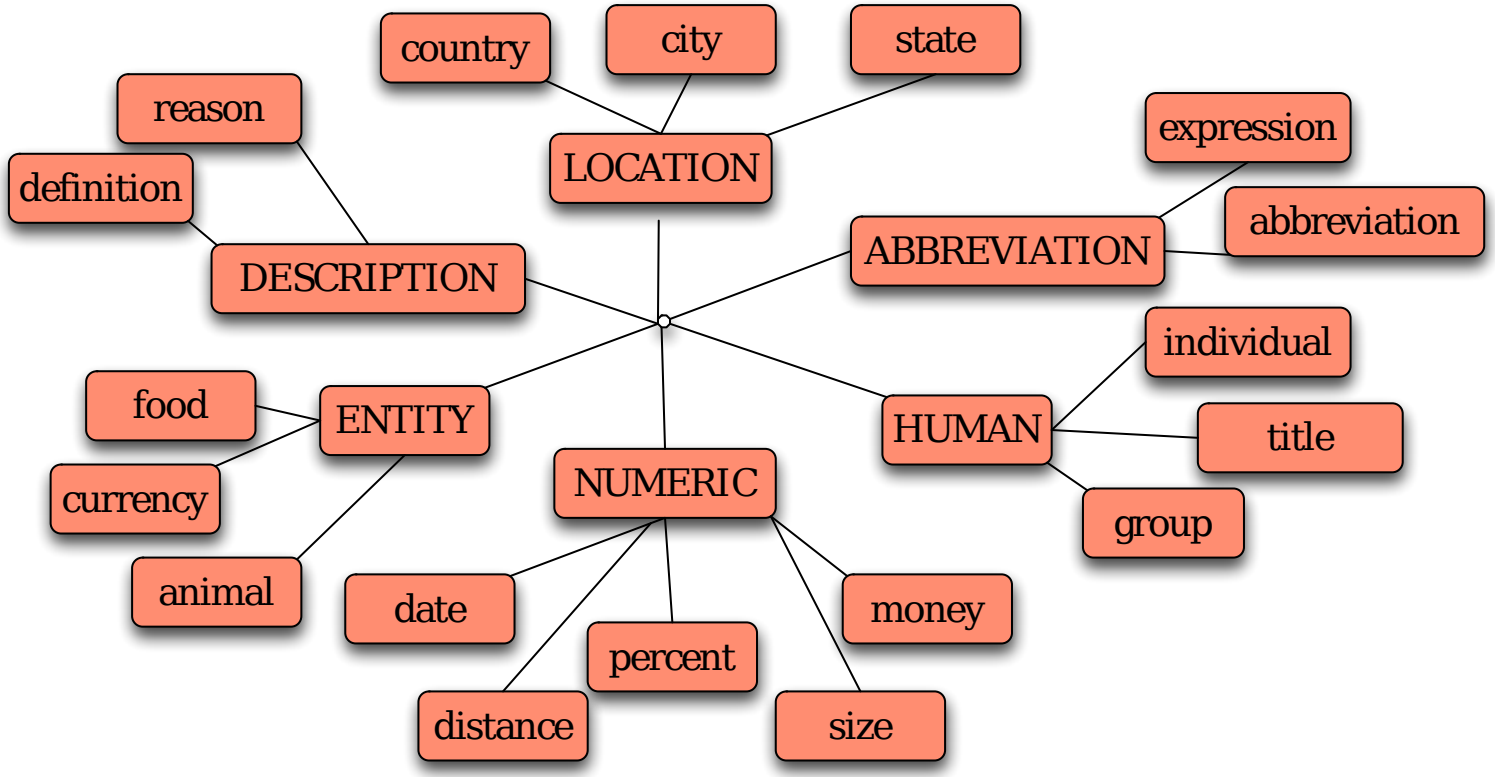
- Who founded Virgin Airlines?
  - PERSON
- What Canadian city has the largest population?
  - CITY
- When is St. Patrick’s Day?
  - DATE

ANSWER TYPE TAXONOMY

- **6 coarse classes:**
  - ABBREVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION, NUMERIC
- **50 finer classes:**
  - LOCATION: city, country, mountain…
  - HUMAN: group, individual, title, description
  - ENTITY: animal, body, colour, currency…

Xin Li, Dan Roth. 2002. Learning Question Classifiers. COLING'02

PART OF LI & ROTH'S TAXONOMY



ANSWER TYPES (I)

ENTITY	
animal	What are the names of Odin's ravens?
body	What part of your body contains the corpus callosum?
color	What colors make up a rainbow ?
creative	In what book can I find the story of Aladdin?
currency	What currency is used in China?
disease/medicine	What does Salk vaccine prevent?
event	What war involved the battle of Chapultepec?
food	What kind of nuts are used in marzipan?
instrument	What instrument does Max Roach play?
lang	What's the official language of Algeria?
letter	What letter appears on the cold-water tap in Spain?
other	What is the name of King Arthur's sword?
plant	What are some fragrant white climbing roses?
product	What is the fastest computer?
religion	What religion has the most members?
sport	What was the name of the ball game played by the Mayans?
substance	What fuel do airplanes use?
symbol	What is the chemical symbol for nitrogen?
technique	What is the best way to remove wallpaper?
term	How do you say " Grandma " in Irish?
vehicle	What was the name of Captain Bligh's ship?
word	What's the singular of dice?

ANSWER TYPES (II)

HUMAN	
description	Who was Confucius?
group	What are the major companies that are part of Dow Jones?
ind	Who was the first Russian astronaut to do a spacewalk?
title	What was Queen Victoria's title regarding India?
LOCATION	
city	What's the oldest capital city in the Americas?
country	What country borders the most others?
mountain	What is the highest peak in Africa?
other	What river runs through Liverpool?
state	What states do not have state income tax?
NUMERIC	
code	What is the telephone number for the University of Colorado?
count	About how many soldiers died in World War II?
date	What is the date of Boxing Day?
distance	How long was Mao's 1930s Long March?
money	How much did a McDonald's hamburger cost in 1963?
order	Where does Shanghai rank among world cities in population?
other	What is the population of Mexico?
period	What was the average life expectancy during the Stone Age?
percent	What fraction of a beaver's life is spent swimming?
speed	What is the speed of the Mississippi River?
temp	How fast must a spacecraft travel to escape Earth's gravity?
size	What is the size of Argentina?
weight	How many pounds are there in a stone?

ANSWER TYPE DETECTION

WARWICK

- Different approaches:
  - **Hand-written rules:**
    As for many other tasks, leads to high precision, low recall. Time-consuming and difficult to maintain.
  - **Machine Learning:**
    Feature extraction + supervised classification.
  - **Hybrids.**

RULES FOR ANSWER TYPE DETECTION

- **Regular expression**-based rules can get some cases:
  - Who {is|was|are|were} PERSON
- Other rules use the **question headword**:

  (the headword of the first noun phrase after the wh-word)
  - Which **city** in China has the largest number of foreign financial companies?

- Preferably (or complementarily), we treat it as machine learning classification problem.
  - **Define a taxonomy** of question types.
  - **Annotate training data** for each question type.
  - **Extract features** to train the classifier.
    - features may include those hand-written rules!

FEATURES FOR ANSWER TYPE DETECTION

- Question words and phrases.
- Part-of-speech tags.
- Parse features (headwords).
- Named Entities.
- Semantically related words.

KEYWORD SELECTION FOR IR QUERY

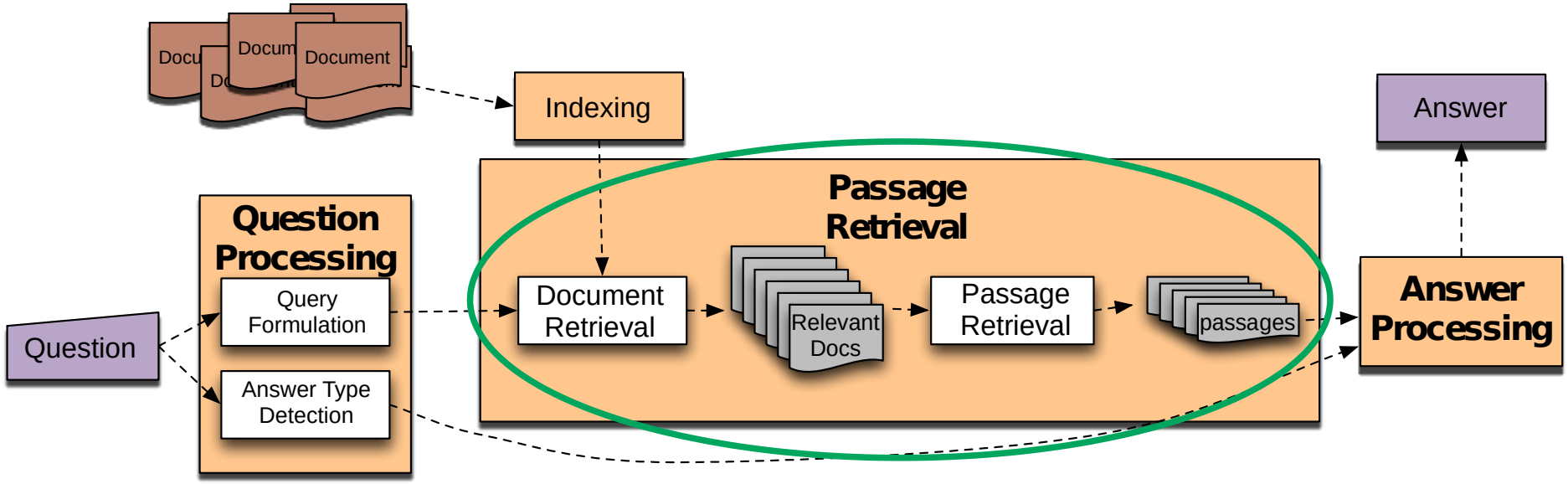
- Get rid of low-IDF keywords.

  ~~Who~~ coined ~~the~~ term “cyberspace” ~~in his~~ novel “Neuromancer”?

WARWICK

PASSAGE RETRIEVAL AND ANSWER EXTRACTION

IR-BASED QA


Docu
Docum
Document
Indexing
Answer
Question Processing
Query Formulation
Answer Type Detection
Question
Passage Retrieval
Document Retrieval
Relevant Docs
Passage Retrieval
passages
Answer Processing

PASSAGE RETRIEVAL

- **Step 1:** IR engine retrieves documents using query terms.
- **Step 2:** Segment the documents into shorter units.
  - Something like paragraphs.
- **Step 3:** Passage ranking.
  - Use answer type to help rerank passages.

FEATURES FOR PASSAGE RANKING

- Number of Named Entities of the right type in passage.
- Number of query words in passage.
- Number of question N-grams also in passage.
- Proximity of query keywords to each other in passage.
- Longest sequence of question words.
- Rank of the document containing passage.

ANSWER EXTRACTION

- Run an **answer-type named-entity tagger on the passages**.
  - Each answer type requires a named-entity tagger that detects it.
  - If answer type is CITY, tagger has to tag CITY.
    - Can be **full NER**, simple **regular expressions**, or **hybrid**.
- Return the **string with the right type**:

How tall is Mt. Everest? (LENGTH)
The official height of Mount Everest is 29035 feet

RANKING CANDIDATE ANSWERS

WARWICK

- But there can be multiple candidate answers of the right type!

- **Q:** Who was Queen Victoria's second son?
- **Answer Type: Person**
- **Passage:**
  The Marie biscuit is named after **Marie Alexandrovna**, the daughter of **Czar Alexander II of Russia** and wife of **Alfred**, the second son of **Queen Victoria** and **Prince Albert**.

USE MACHINE LEARNING TO RANK CANDIDATES

WARWICK

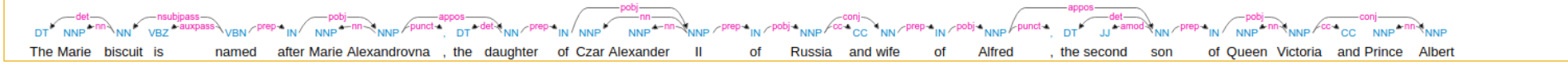
- **Features:**
  - **Answer type match:** candidate is of the right type.
  - **Pattern match:** Regex pattern matches the candidate.
  - **Keyword distance:** Words between candidate and query keywords.
  - **Punctuation:** Candidate is followed by comma, period, quotation marks…

USE MACHINE LEARNING TO RANK CANDIDATES

- **More features:**
  - **Dependency parsing:**

Syntactic Dependency Parse from TurboParser [0.6s]

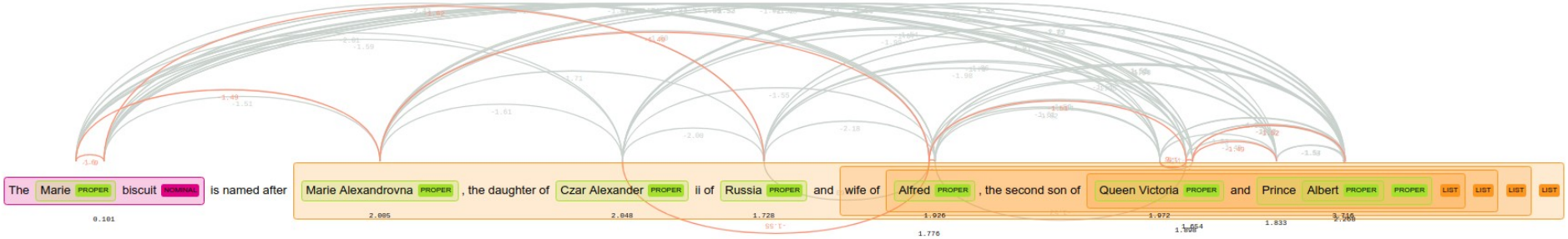
Tree View (explanation)



CoNLL Format »

USE MACHINE LEARNING TO RANK CANDIDATES

- **More features:**
  - **Co-reference resolution:**



USING KNOWLEDGE IN QA

RELATION EXTRACTION

- **Answers:** Databases of Relations, e.g.
  - born-in(“Emma Goldman”, “June 27 1869”)
  - author-of(“Cao Xue Qin”, “Dream of the Red Chamber”)
  - Draw from Wikipedia infoboxes, DBpedia, FreeBase, etc.

RELATION EXTRACTION

- **Questions:** Extracting Relations in Questions, e.g.
  - Whose granddaughter starred in E.T.?
    - (acted-in ?x “E.T.”)
    - (granddaughter-of ?x ?y)

TEMPORAL REASONING

WARWICK

- Again we can use databases of relations, as well as obituaries, biographies, etc.

- **Q:** ”In 1594 he took a job as a tax collector in Andalusia”
- **Candidates:**
  - Thoreau is a bad answer (born in 1817)
  - Cervantes is possible (was alive in 1594)

GEOSPATIAL KNOWLEDGE

WARWICK

- Using gazeteers like geonames.org:
  Coventry is a possible answer of type: UK CITY

coventry | all countries

search | show on map | [advanced search]

426 records found for "coventry"

	Name	Country	Feature class	Latitude	Longitude
1	Coventry CVT,Coventria,Coventry,Gorad Koventry,Koventri,Koventrija,Koventris,Kovuntri,kao wen chui,kawntry,kh...	United Kingdom, England Coventry	seat of a second-order administrative division population 359,262	N 52° 24' 23"	W 1° 30' 43"
2	Coventry COV,Coventry	United Kingdom, England Coventry	second-order administrative division population 352,911	N 52° 25' 0"	W 1° 30' 0"
3	Coventry Coventry,Koventri,South Coventry,kao wen chui,kwfntry,Ковентри,کنیکٹیکٹ ،کووینٹری,کوفنتري,考文垂	United States, Connecticut Tolland County > Town of Coventry	populated place elevation 160m	N 41° 46' 12"	W 72° 18' 18"
4	Coventry Airport CVT,EGBE,Koventri,Port lotniczy Coventry,Ковентри	United Kingdom, England Warwickshire > Warwick District	airport elevation 81m	N 52° 22' 11"	W 1° 28' 47"
5	Coventry Cathedral Ardeaglais Coventry,Catedral de Coventry,Cathedrale Saint-Michel de Coventry,Cathédrale Saint-Michel...	United Kingdom, England Coventry > GB.ENG.C7.44UF	church	N 52° 24' 30"	W 1° 30' 25"

CLARIFICATION QUESTIONS

WARWICK

With chatbots, e.g. Siri:

Q: Where can I buy Thai food?

A: Are you looking for a Thai restaurant or a supermarket that sells Thai food?

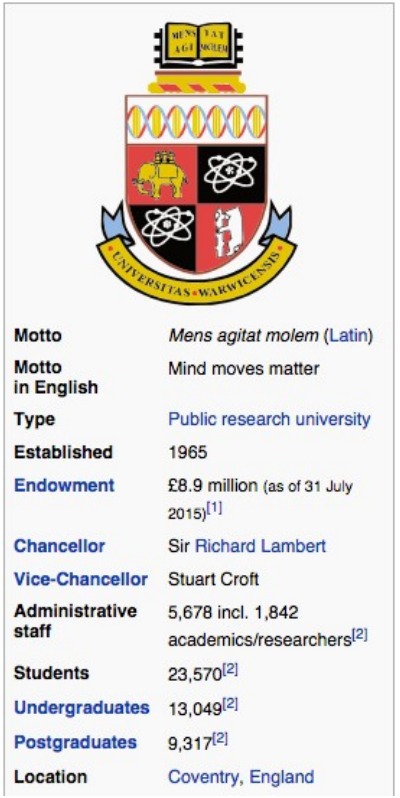
WARWICK

DATASETS AND EVALUATION

COLLECTING DATASETS

- We can build a dataset **from online data**:
  - Get question-answer pairs from Wikipedia infoboxes:

    Where is the University of Warwick located?



Motto	*Mens agitat molem* (Latin)
Motto in English	Mind moves matter
Type	Public research university
Established	1965
Endowment	£8.9 million (as of 31 July 2015)[1]
Chancellor	Sir Richard Lambert
Vice-Chancellor	Stuart Croft
Administrative staff	5,678 incl. 1,842 academics/researchers[2]
Students	23,570[2]
Undergraduates	13,049[2]
Postgraduates	9,317[2]
Location	Coventry, England

COLLECTING DATASETS

- We can build a dataset **from trivia questions** too.
  - Then try to answer the questions from online data.

1.	What number is on Lightning McQueen in Cars movie?	95
2.	Name 2 liquids a tanker truck could be carrying.	Petrol Milk
3.	Who is the leader in blue of the Teenage Mutant Ninja Turtles?	Leonardo
4.	What is a flying dinosaur called?	Pterodactyl
5.	How many days in a fortnight?	14
6.	Who sings "Can't Stop The Feeling!"?	Justin Timberlake
7.	How many letters are there in the alphabet?	26
8.	What is the name of Donatello's weapon in Ninja Turtles?	Bo Staff
9.	What is a labradoodle a mix of?	Labrador & Poodle
10.	What is the teacher's name in Peppa Pig?	Madame Gazelle
11.	Name 2 people that work with Fireman Sam.	Elvis, Penny, Station Officer Steele etc.
12.	How many minutes in an hour?	60

OR USE EXISTING DATASETS

- The Stanford Question Answering Dataset: https://rajpurkar.github.io/SQuAD-explorer/

- Quasar: https://github.com/bdhingra/quasar

EVALUATION METRICS

- **Accuracy:** ratio of correct predictions.
- **Mean Reciprocal Rank (MRR):**
  - For each query return a **ranked list of M candidate answers**.
  - Query score **is 1/Rank of the first correct answer**.
    - If first answer is correct: 1
    - else if n^{th} answer is correct: 1/n
    - score is 0 if none of the M answers are correct
  - Take the **mean over all N queries**.

$$MRR = \frac{\sum_{i=1}^{N} \frac{1}{rank_i}}{N}$$

QA: SUMMARY

- Pipeline:
  - Question Processing.
  - Passage Retrieval.
  - Answer Processing.
- Types:
  - IR-based QA.
  - Knowledge-based QA.
- A smart strategy is to combine sources for our answers:
  - Easy questions like “what is the population of X?” → extract from knowledge base.
  - More difficult questions like “can X cure Y?” → IR-based QA.
  - Opinionated questions, e.g. “Is NLTK better than spacy?” → show web results.

ASSOCIATED READING

- Jurafsky, Daniel, and James H. Martin. 2009. Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics. 3rd edition. **Chapter 23.**