

Question Answering

Debapriyo Majumdar
Information Retrieval
Indian Statistical Institute Kolkata

Adapted from slides by Dan Jurafsky (Stanford) and Tao Yang (UCSB)

Question Answering

One of the oldest NLP tasks (punched card systems in 1961)

Simmons, Klein, McConlogue. 1964. Indexing and Dependency Logic for Answering English Questions. American Documentation 15:30, 196-204

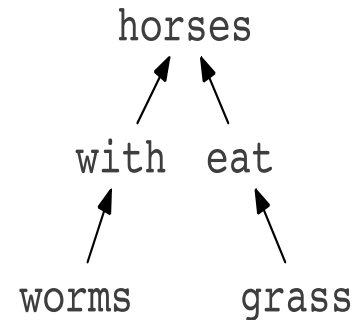
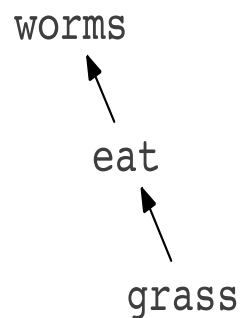
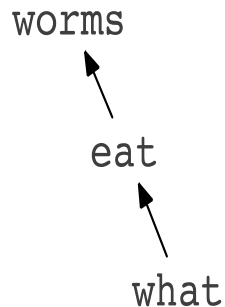
Question:

Potential Answers:

What do worms eat?

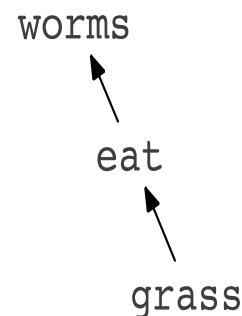
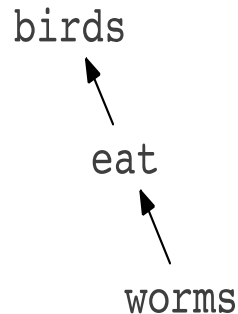
Worms eat grass

Horses with worms eat grass



Birds eat worms

Grass is eaten by worms



Question Answering: IBM's Watson

- Won Jeopardy on February 16, 2011!

Willian Wilkinson's "An account of the principalities of Wallachia and Moldovia" inspired **this** author's most famous novel.




Bram Stoker

Apple's Siri



- A seemingly “limited” set of few possible questions
- Answers based on contextual parameters

Wolfram Alpha, Google

 computational knowledge engine

calories in two slices of brown bread

Examples Random

Input interpretation:

boston brown bread	amount	2 slices	total calories
--------------------	--------	----------	----------------

Result:

176 Cal (dietary Calories)

Unit conversions: [More](#)

730 kJ (kilojoules)

0.73 MJ (megajoules)

730 000 J (joules)

7.3×10^{12} ergs
(unit officially deprecated)

180 kcal_{th} (thermochemical kilocalories)
(unit officially deprecated)

Comparison to reference daily intake:

	daily amount	% daily value
RDA	2000 Cal/day (dietary calories per day)	9%

(RDA: recommended daily allowance)

calories in two slices of brown bread

Web News Images Videos More Search tools

About 10,20,000 results (0.49 seconds)

146 calories

Type: Bread, wheat germ, toasted Quantity: 2 slice (50 g)

Sources include: USDA Feedback

Calories in Brown Bread - Weight Loss Resources
www.weightlossresources.co.uk/calories-in-food/.../Brown-Bread.htm
Brown Bread Calories and Nutrition per Serving (1 Serving=1 Med Slice/34g) ... It takes just 2 minutes to take a free trial and could be the breakthrough you've ...

Calories in Brown Toast - Weight Loss Resources
www.weightlossresources.co.uk/calories-in-food/.../Brown-Toast.htm
View calories and nutrition info per 1 Med Slice/24g of Brown Toast and see how many ... It takes just 2 minutes to take a free trial and could be the breakthrough ... Multi Seed Country Granary Bread · Hovis Original Wheatgerm Brown Bread ...

Calories in Brown Bread - Weight Loss For All
www.weightlossforall.com/calories-bread.htm
How many calories are in a slice of bread? The calories found in white bread are the most with 100 calories per slice. The calories in brown bread is 82 per slice ...


Calories in 2 slices of Whole Wheat Bread and Nutrition Facts
<https://www.falsest.com/calories-nutrition/.../whole-wheat-bread?...2...>
There are 256 calories in 2 slices of Whole Wheat Bread. Get full nutrition facts and other common serving sizes of Whole Wheat Bread including 1 oz and 1 ...

PDF Food Calories Food Calories Bread Vegetables Brown ...
www.ultimatehealthandfitness.co.uk/pdf/Calorie%20Food%20Index.pdf
Calories. Food. Calories. Bread. Vegetables. Brown Bread (1 slice). 52. Broccoli (2 tablespoons). 26. White Bread (1 slice). 76. Carrot (2 tablespoons). 35.

Calories in Wholemeal Bread - Nutracheck
www.nutracheck.co.uk/calories/calories.../calories_in_bread_wholemeal...
Calories in Bread, Wholemeal Bread, 25g Slice (small loaf), 54, 0.5. 36g Medium slice (large loaf), 78, 0.9. 44g Thick slice (large loaf), 95, 1.1. 100g, 217, 2.5 ...

Brown bread

Brown bread is a designation often given to breads made with significant amounts of whole grain flour, usually rye or wheat, and sometimes dark coloured ingredients such as molasses or coffee. [Wikipedia](#)



Nutrition Facts
Bread, wheat germ, toasted

Amount Per 2 slice (50 g)

Calories 146

	% Daily Value*
Total Fat 1.6 g	2%
Saturated fat 0.4 g	2%
Polyunsaturated fat 0.4 g	
Monounsaturated fat 0.8 g	
Cholesterol 0 mg	0%
Sodium 310 mg	12%
Potassium 142 mg	4%
Total Carbohydrate 28 g	9%
Dietary fiber 1.2 g	4%
Sugar 2 g	
Protein 5.4 g	10%
Vitamin A	0% Vitamin C
Calcium	5% Iron
Vitamin D	0% Vitamin B-6
Vitamin B-12	0% Magnesium

*Per cent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

Mathematical comparison: Wolfram Alpha



is 3 miles more than 5 kilometers?



 Extended Keyboard

 Upload

 Examples

 Random

Input interpretation:

3 miles > 5 km (kilometers)

Result:

False

In this case, Google returned a standard list of document links (2016)

Mathematical comparison: Google (1)



is 3 miles more than 5 km?



[All](#)

[News](#)

[Images](#)

[Shopping](#)

[Videos](#)

[More](#)

[Tools](#)

About 23,60,00,000 results (0.66 seconds)

8 July 2021

A **5K** is 3.1 **miles**. On a standard outdoor track, a **5K** (5,000 meters) is 12.5 laps. 04-Jun-2021

<https://www.runnersworld.com/how-many-miles-is-a-5k>

[5K in Miles - How Far Are Common Race Distances?](#)

[About featured snippets](#) • [Feedback](#)

People also ask

Is 5K the same as 3 miles?



Which is bigger km or miles?



What is the difference between 1 mile and 1 kilometer?



How many miles are in a kilometer?



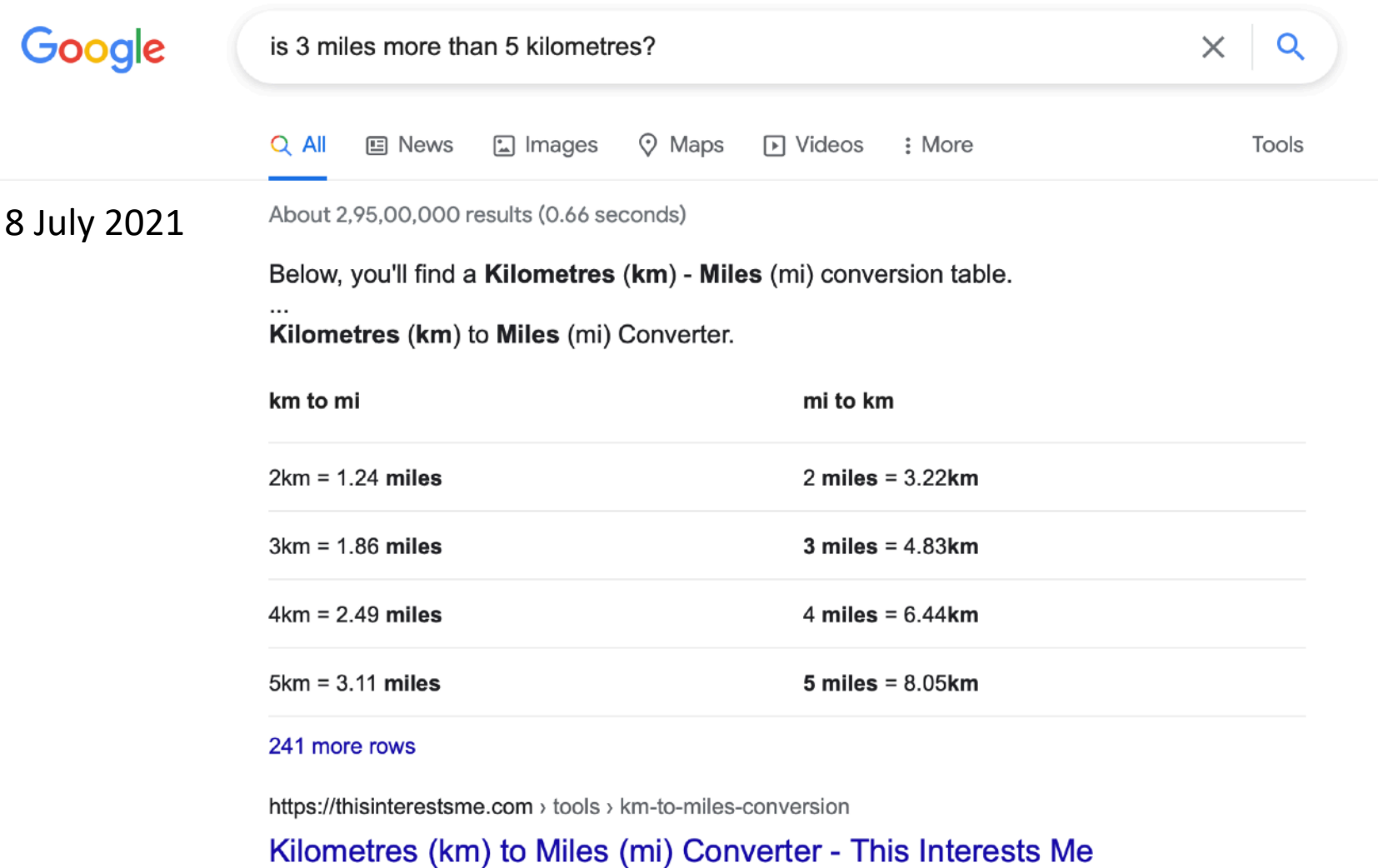
[Feedback](#)

https://www.answers.com/What_is_greater_-_5_miles...

[What is greater - 5 miles or 5km? - Answers](#)

23-May-2011 — 1 mile is slightly more than 1.6 km So 5 miles is a longer distance than 5 km.

Mathematical comparison: Google (1)



The screenshot shows a Google search interface. The search bar contains the text "is 3 miles more than 5 kilometres?". Below the search bar, the Google logo is on the left, and navigation links for "All", "News", "Images", "Maps", "Videos", and "More" are in the center. A "Tools" link is on the right. The search results show "About 2,95,00,000 results (0.66 seconds)". The main result is a link to a "Kilometres (km) to Miles (mi) Converter" from the website "thisinterestsme.com". Below the link is a table with two columns: "km to mi" and "mi to km". The table contains five rows of conversion data. A link "241 more rows" is shown below the table. The URL "https://thisinterestsme.com › tools › km-to-miles-conversion" is displayed below the table. The title "Kilometres (km) to Miles (mi) Converter - This Interests Me" is shown below the URL.

Google

is 3 miles more than 5 kilometres?

All News Images Maps Videos More Tools

8 July 2021

About 2,95,00,000 results (0.66 seconds)

Below, you'll find a **Kilometres (km) - Miles (mi)** conversion table.

...

Kilometres (km) to Miles (mi) Converter.

km to mi	mi to km
2km = 1.24 miles	2 miles = 3.22km
3km = 1.86 miles	3 miles = 4.83km
4km = 2.49 miles	4 miles = 6.44km
5km = 3.11 miles	5 miles = 8.05km

[241 more rows](#)

<https://thisinterestsme.com › tools › km-to-miles-conversion>

[Kilometres \(km\) to Miles \(mi\) Converter - This Interests Me](#)

Slight variation of queries (or even time) changes the answers

Types of Questions in Modern Systems

- Factoid questions

- Answers are short
- The question can be rephrased as “fill in the blanks” question

Examples:

- Who directed the movie Titanic?
- How many calories are there in two slices of apple pie?
- Where is Louvre museum located?

- Complex questions (answers may or may not be short):

- What precautionary measures should we take to be safe from swine flu?
- What do scholars think about Jefferson’s position on dealing with pirates?
- Is 3 miles more than 3 km?

Paradigms for QA

- Retrieval-based
 - TREC QA Track (e.g., AskMSR)
- Knowledge-base + IR
 - IBM Watson
 - Wolfram Alpha
 - Siri
 - True Knowledge Evi
- Passage retrieval + deep learning + knowledge base
 - Google
 - Other modern systems

The classification of commercial systems which do not publish how they work may be inaccurate

Common Evaluation Metrics

1. **Accuracy** (does answer match gold-labeled answer?)
 - Makes sense only for 1st answer
2. **Mean Reciprocal Rank**
 - For each query return a ranked list of M candidate answers.
 - Query score is $\frac{1}{\text{rank}}$ of the first correct answer
 - If first answer is correct: 1
 - else if second answer is correct: $\frac{1}{2}$
 - else if third answer is correct: $\frac{1}{3}$, etc.
 - Score is 0 if none of the M answers are correct
 - Take the mean over all N queries
 - $$\text{MRR} = \frac{1}{N} \sum_{i=1}^N \frac{1}{\text{rank}_i}$$

A Basic IR Based Approach

ASKMSR

AskMSR: Shallow approach

- *In what year did Abraham Lincoln die?*
- Ignore hard documents and find easy ones

Abraham Lincoln, 1809-1865

***LINCOLN, ABRAHAM** was born near Hodgenville, Kentucky, on February 12, 1809. In 1816, the Lincoln family moved to Indiana settling near Little Pigeon Creek in Perry (now Spencer) County. Two years later, Abraham Lincoln's mother died and his father married a woman who became his "angel" mother. Lincoln attended a formal school for only a few months but acquired knowledge through the reading of books. He moved to Illinois, in 1830 where he obtained a job as a store clerk and the local postmaster. He served without distinction in the Black Hawk War. He lost his attempt at the state legislature, but two years later he tried again, was successful, and was elected to the U.S. House in 1846, at which time he opposed the war with Mexico. By 1854, Lincoln was a member of the Republican party and gained national attention for his series of debates with Stephen A. Douglas for the U.S. senatorial seat. Although Lincoln lost the election he became a significant figure in his party. He was elected to the U.S. presidency and by the time of his inauguration on March 4, seven southern states had seceded. On April 4, 1865, Sumter was fired on by Confederate artillery. Lincoln called for 75,000 volunteers (approved for a total of 11). Lincoln immediately took action. The Emancipation Proclamation which expanded the purpose of the war. The dedication of a national cemetery in Gettysburg, Lincoln's famous Gettysburg Address explains with clarity and feeling the reason there was a war. Lincoln, a man of eminent, he spoke of "charity for all" and "malice to none." He was shot by General Ulysses S. Grant on April 9, 1865, at Appomattox Court House. He died at Ford's Theatre and was shot in the head by John Wilkes Booth.



Sixteenth President
1861-1865
Married to Mary Todd Lincoln



ABRAHAM LINCOLN

**Sixteenth President
of the United States**

Born in 1809 - Died in 1865

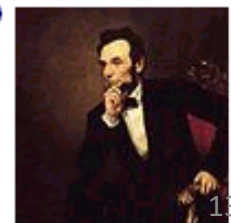
Abraham Lincoln

16th President of the United States (March 4, 1861 to April 15, 1865)

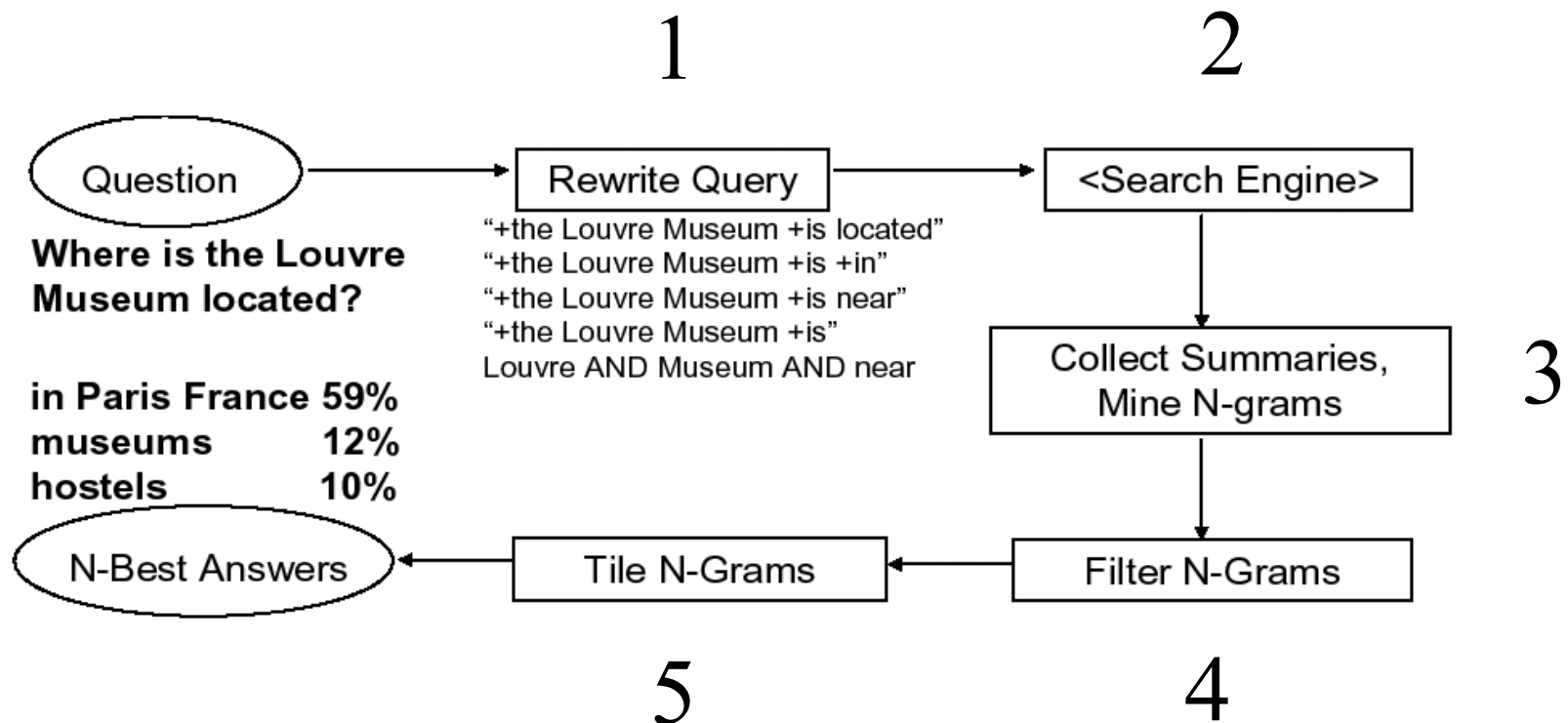
Born: February 12, 1809, in Hardin County, Kentucky

Died: April 15, 1865, at Petersen's Boarding House in Washington, D.C.

"I was born February 12, 1809, in Hardin County, Kentucky. My parents were both born in Virginia, of undistinguished families, perhaps I should say. My mother, who died in my tenth year, was of a family of the name of



AskMSR: Details



Step 1: Rewrite queries

- Intuition: The user's question is often syntactically quite close to sentences that contain the answer
 - Q: Where is the Louvre Museum located?
 - Hope: there would be a sentence of the form:
The Louvre Museum is located in ***Paris.***
 - Q: Who created the character of Scrooge?
 - Hope: there would be a sentence of the form:
Charles Dickens created the character of Scrooge.

Query rewriting

- Classify question into categories
 - **Who** is/was/are/were...?
 - **When** is/did/will/are/were ...?
 - **Where** is/are/were ...?
 - ...

a. Category-specific transformation rules

eg “For Where questions, move ‘is’ to all possible locations”

“Where is the Louvre Museum located”

→ “is the Louvre Museum located”

→ “the is Louvre Museum located”

→ “the Louvre is Museum located”

→ “the Louvre Museum is located”

→ “the Louvre Museum located is”

Some of these
are nonsense,
but who
cares? It's
only a few
more queries
to Google.

b. Expected answer “Datatype” (eg, Date, Person, Location, ...)

When was the French Revolution? → DATE

- Hand-crafted classification/rewrite/datatype rules
(Could they be automatically learned?)

Step 2: Query search engine

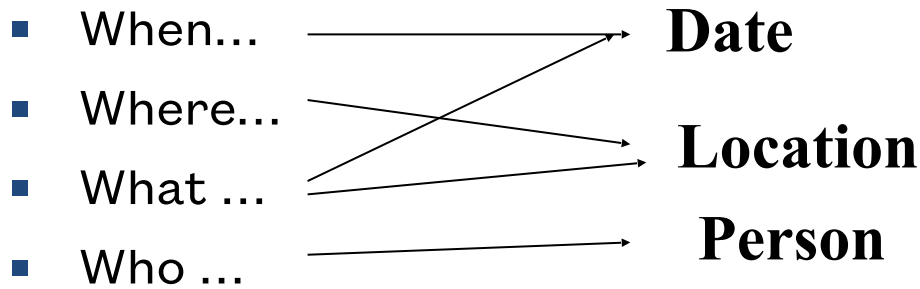
- Send all rewrites to a (Web) search engine
- Retrieve top N answers (may be 100)
- For speed, rely just on search engine's “snippets”, not the full text of the actual document

Step 3: Mining N-Grams

- Enumerate all N-grams (N=1,2,3 say) in all retrieved snippets
 - Use hash table and other fancy footwork to make this efficient
- Weight of an n-gram: occurrence count, each weighted by “reliability” (weight) of rewrite that fetched the document
- Example: “Who created the character of Scrooge?”
 - Dickens - 117
 - Christmas Carol - 78
 - Charles Dickens - 75
 - Disney - 72
 - Carl Banks - 54
 - A Christmas - 41
 - Christmas Carol - 45
 - Uncle - 31

Step 4: Filtering N-Grams

- Each question type is associated with one or more “**data-type filters**” = regular expression



- Boost score of n-grams that do match regexp
- Lower score of n-grams that don't match regexp

Step 5: Tiling the Answers

Scores

20

Charles Dickens

15

Dickens

10

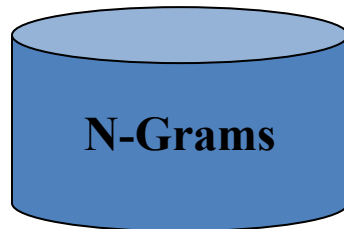
Mr Charles

merged, discard
old n-grams

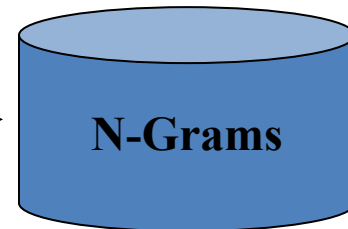
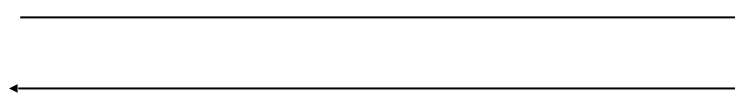


Score 45

Mr Charles Dickens



tile highest-scoring n-gram



Repeat, until no more overlap

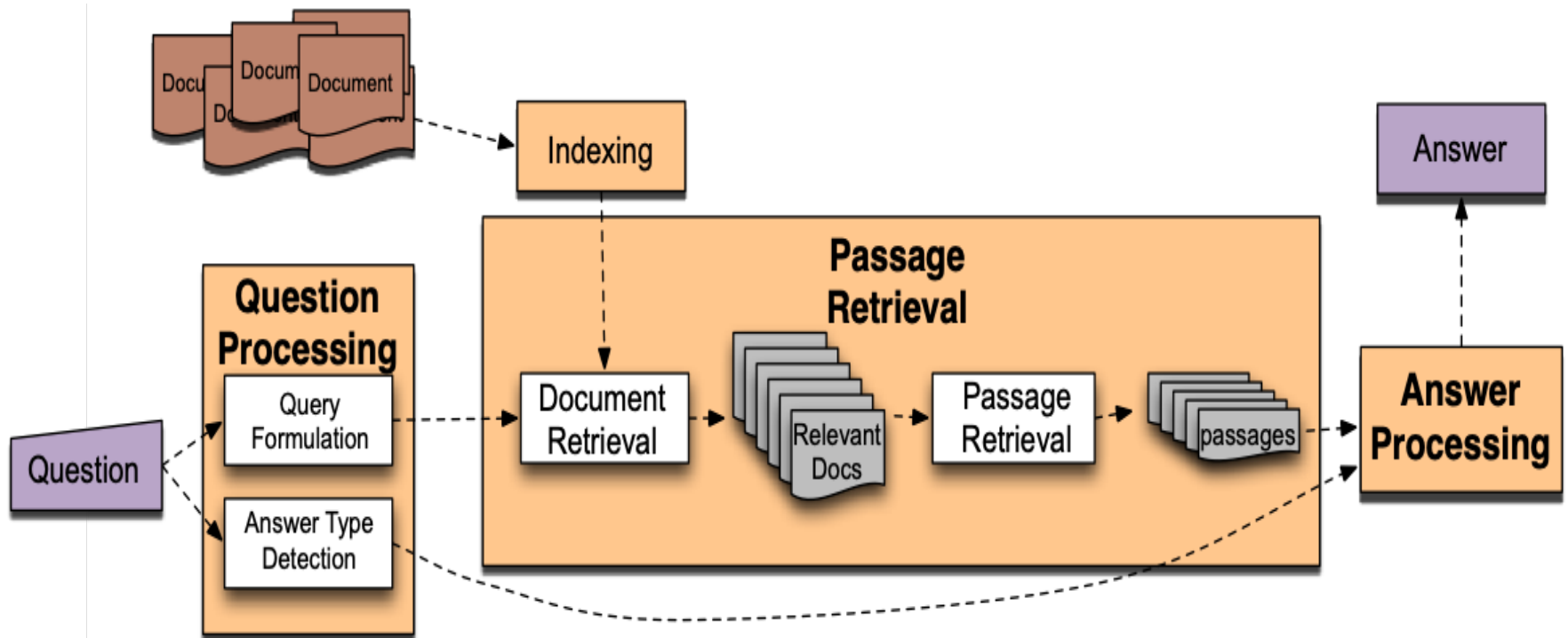
Results

- Standard TREC contest test-bed:
 - ~1M documents; 900 questions
- Technique doesn't do too well (though would have placed in top 9 of ~30 participants!)
 - $\text{MRR} = 0.262$ (ie, right answered ranked about #4-#5)
- Using the Web as a whole, not just TREC's 1M documents... $\text{MRR} = 0.42$ (ie, on average, right answer is ranked about #2-#3)
 - Why? Because it relies on the enormity of the Web!

Semi-modern QA: roughly IBM Watson (2011)

IR + ML + KNOWLEDGE BASED FACTOID QA

IR-based Factoid QA



IR-based Factoid 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

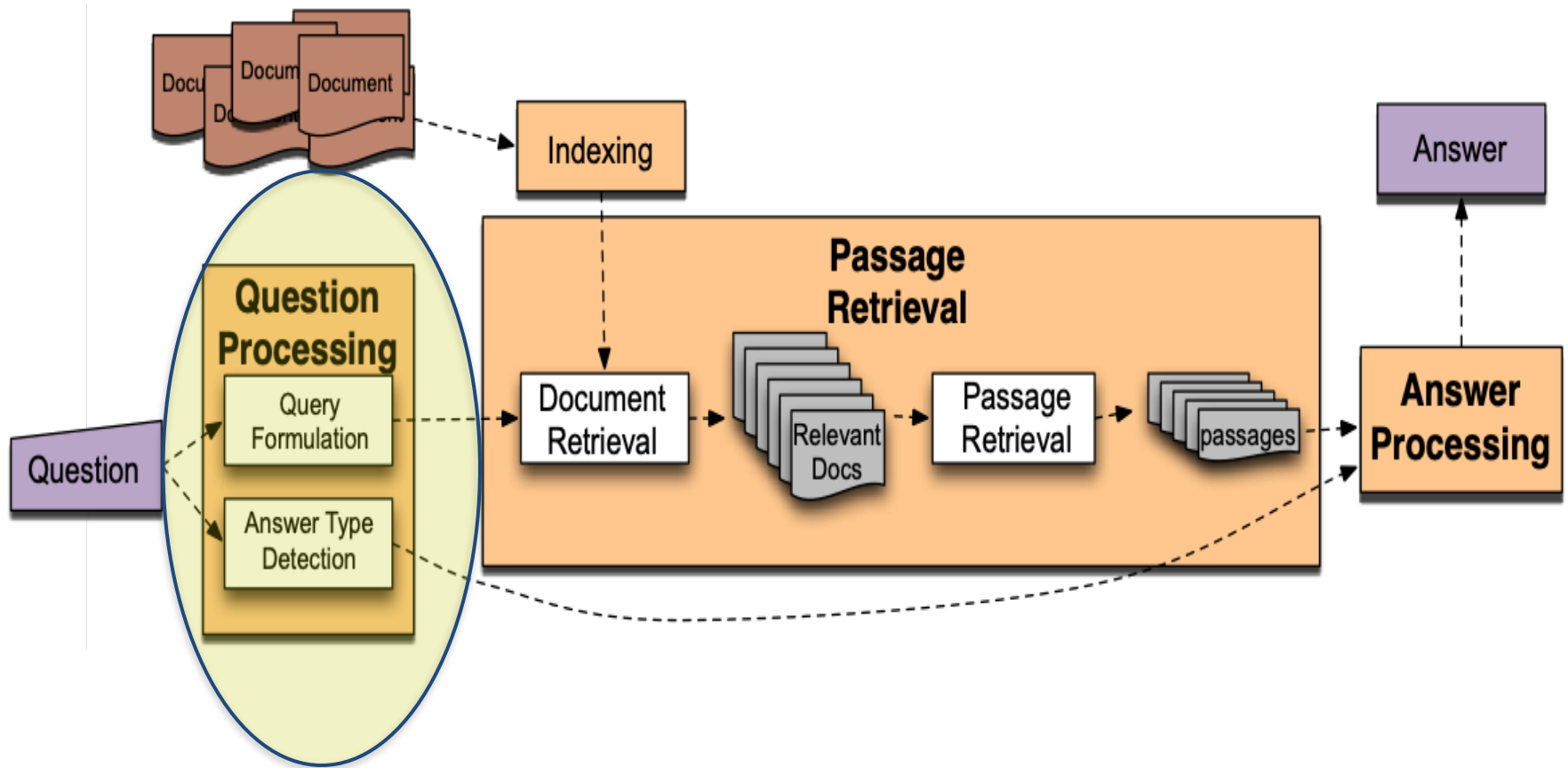
Knowledge-based approaches (Siri)

- 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 (IBM Watson)

- 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

IR-based Factoid QA



Question processing: Extraction from the question

- Answer Type Detection
 - Decide the **named entity type** (person, place) of the answer
- Query Formulation
 - Choose **query keywords** for the IR system
- Question Type classification
 - Is this a definition question, a math question, a list question?
- Focus Detection
 - Find the question words that are replaced by the answer
- Relation Extraction
 - Find relations between entities in the question

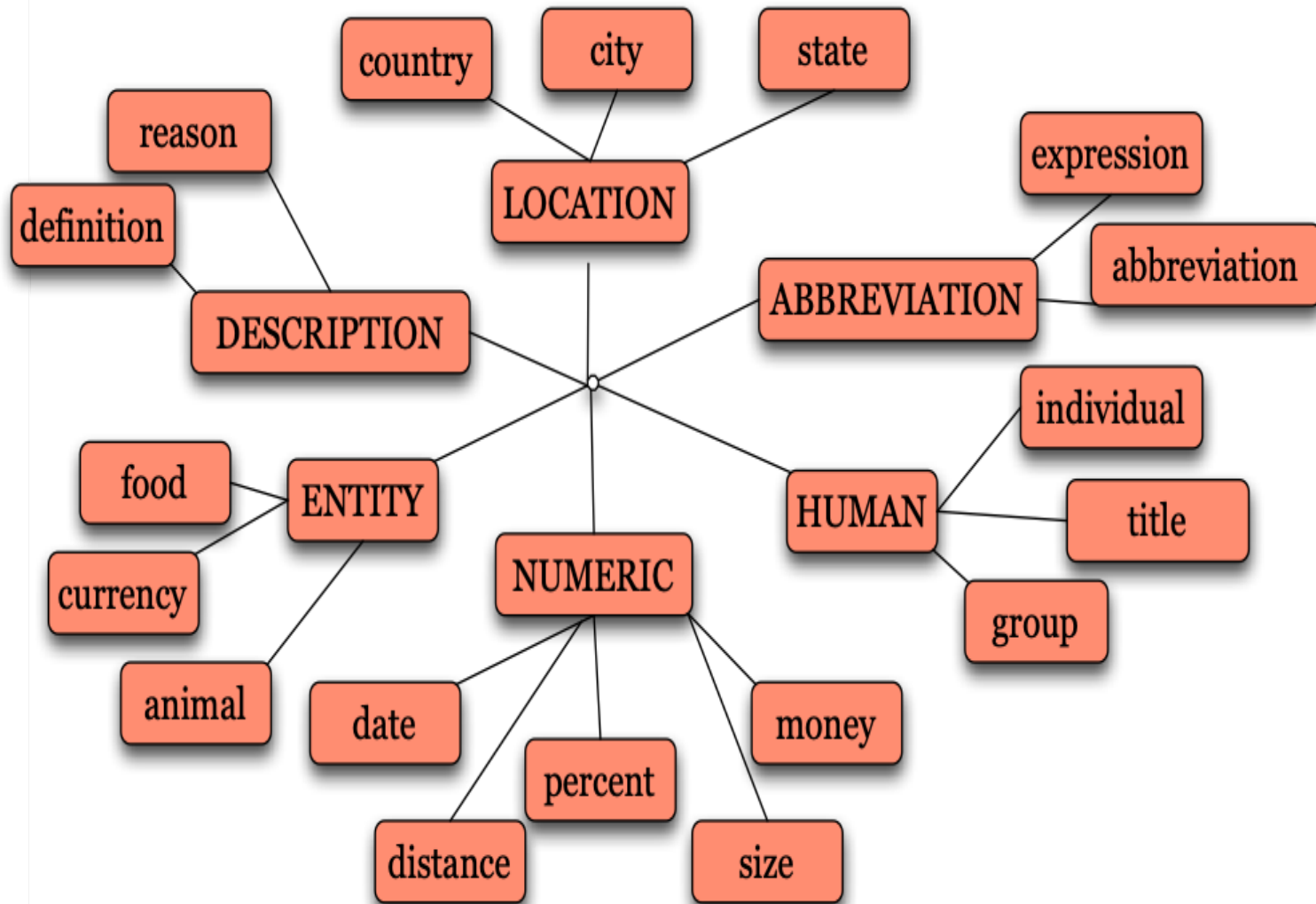
Answer Type Detection: Named Entities

- Who founded Virgin Airlines?
 - PERSON
- What Canadian city has the largest population?
 - CITY.

Answer type taxonomy

- 6 coarse classes
 - ABBEVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION, NUMERIC
- 50 finer classes
 - LOCATION: city, country, mountain...
 - HUMAN: group, individual, title, description
 - ENTITY: animal, body, color, currency...

Part of Li & Roth's Answer Type Taxonomy



Answer Types

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?

More Answer Types

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 types in Jeopardy

- 2500 answer types in 20,000 Jeopardy question sample
- The most frequent 200 answer types cover < 50% of data
- The 40 most frequent Jeopardy answer types
he, country, city, man, film, state, she, author, group, here,
company, president, capital, star, novel, character, woman, river,
island, king, song, part, series, sport, singer, actor, play, team,
show, actress, animal, presidential, composer, musical, nation,
book, title, leader, game

Ferrucci et al. 2010. Building Watson: An Overview of the DeepQA Project. AI Magazine. Fall 2010. 59-79.

Answer Type Detection

- Hand-written rules
- Machine Learning
- Hybrids

Answer Type Detection: Rules

- Regular expression-based rules can get some cases:
 - Who {is|was|are|were} PERSON
 - PERSON (YEAR – YEAR)
- 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?
 - What is the state **flower** of California?

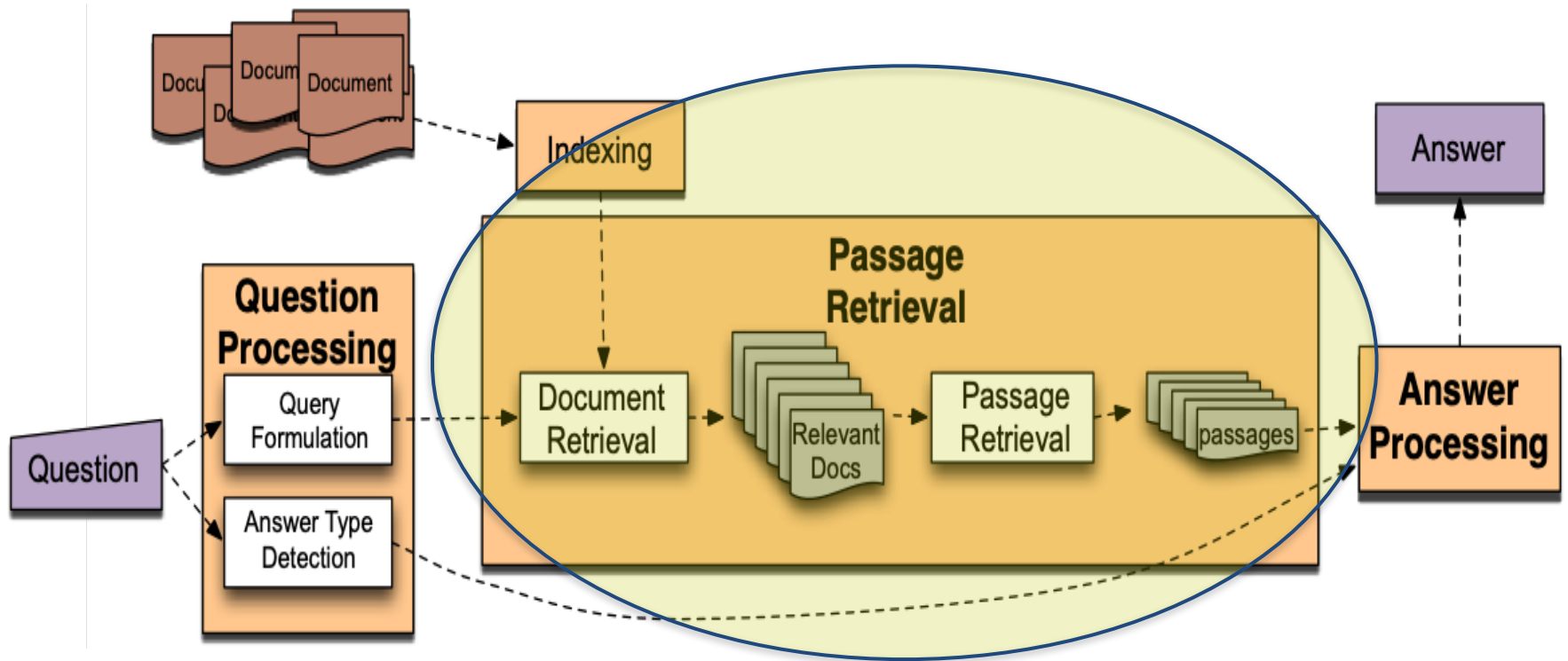
Answer Type Detection: ML

- Treat the problem as machine learning classification
 - **Define** a taxonomy of question types
 - **Annotate** training data for each question type
 - **Train** classifiers for each question class using a rich set of features.
 - features include those hand-written rules!
- Features
 - Question words and phrases
 - Part-of-speech tags
 - Parse features (headwords)
 - Named Entities
 - Semantically related words

Query formulation: keyword selection

1. Select all non-stop words in quotations
2. Select all NNP words in recognized named entities
3. Select all complex nominals with their adjectival modifiers
4. Select all other complex nominals
5. Select all nouns with their adjectival modifiers
6. Select all other nouns
7. Select all verbs
8. Select all adverbs
9. Select the QFW word (skipped in all previous steps)
10. Select all other words

IR-based Factoid QA



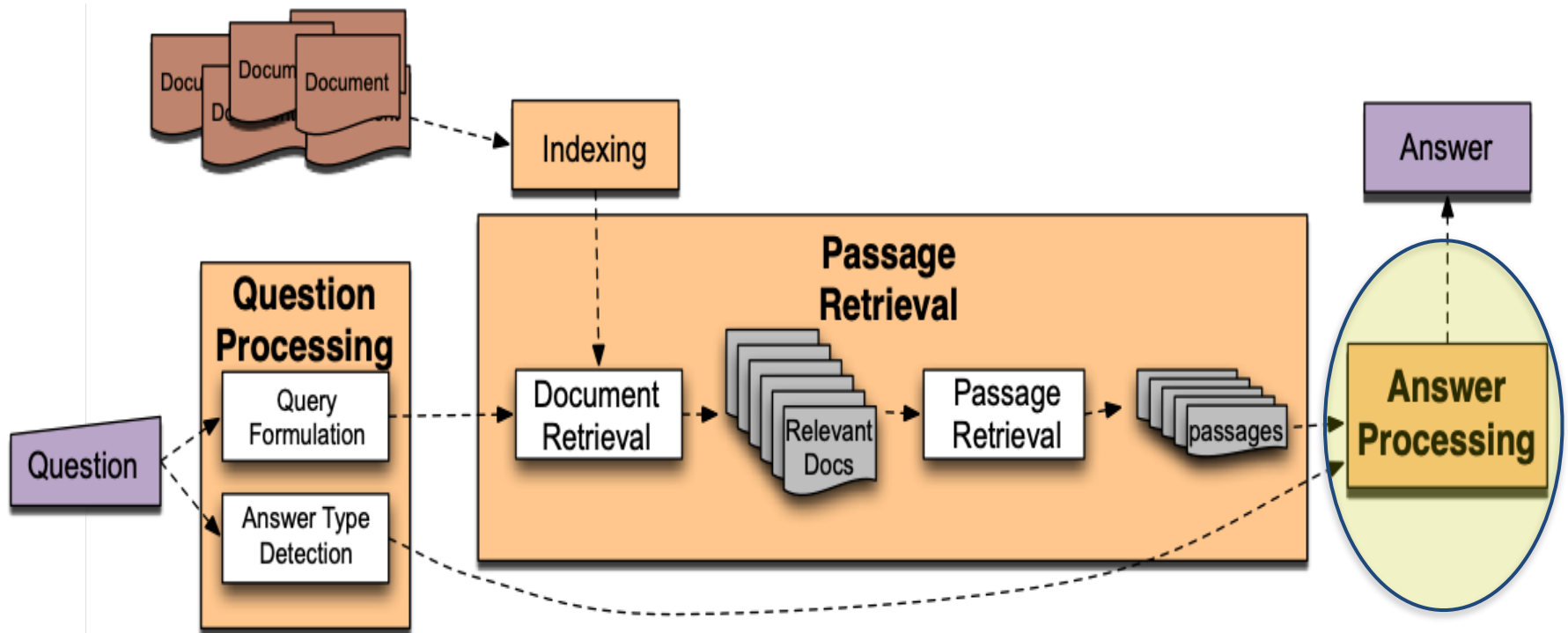
Passage Retrieval

- Intuition: the answer to a question is usually not the full document. Either a passage, or a part of it.
- Step 1: IR engine retrieves documents using query terms
- Step 2: Segment the documents into shorter units
 - Paragraphs, ...
- Step 3: Passage ranking
 - Use answer type to help re-rank 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

IR-based Factoid QA



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:
 - Who is the prime minister of India (PERSON)
Manmohan Singh, Prime Minister of India, had told left leaders that the deal would not be renegotiated.
 - How tall is Mt. Everest? (LENGTH)
The official height of Mount Everest is **29035 feet**

Ranking Candidate Answers

- But what if there are multiple candidate answers!

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

Ranking Candidate Answers

- But what if there are multiple candidate answers!

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

Ranking candidate answers

Use Machine learning

Features for ranking candidate answers:

- **Answer type match:** Candidate contains a phrase with the correct answer type
- **Pattern match:** Regular expression pattern matches the candidate.
- **Question keywords:** # of question keywords in the candidate.
- **Keyword distance:** Distance in words between the candidate and query keywords
- **Novelty factor:** A word in the candidate is not in the query.
- **Apposition features:** The candidate is an appositive to question terms
- **Punctuation location:** The candidate is immediately followed by a comma, period, quotation marks, semicolon, or exclamation mark.
- **Sequences of question terms:** The length of the longest sequence of question terms that occurs in the candidate answer.

Knowledge base: Relation Extraction

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

Whose granddaughter starred in E.T.?

(acted-in ?x “E.T.”)

(granddaughter-of ?x ?y)

Knowledge base: Temporal Reasoning

- Relation databases

- (and obituaries, biographical dictionaries, etc.)

- IBM Watson

“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)

Knowledge base: Geospatial knowledge

- **Beijing** is a good answer for "Asian city"
- **California** is "southwest of Montana"
- **geonames.org**:

The screenshot shows the geonames.org search interface. The search bar contains 'palo alto' and the country dropdown is set to 'all countries'. The search results show 459 records found for 'palo alto'. The first three results are listed in a table below.

	Name	Country	Feature class	Latitude	Longitude
1	Palo Alto Palo Al'to, Palo Alto, pa luo ao duo, paroaruto, Пало Алто, Пало Альто, פאלו אלטו, パロアルト, 帕羅奧多	United States , California Santa Clara County	populated place population 64,403, elevation 9m	N 37° 26' 30"	W 122° 8' 34"
2	Palo Alto Township Palo Alto Township	United States , Iowa Jasper County	administrative division elevation 256m	N 41° 38' 15"	W 93° 2' 57"
3	Borough of Palo Alto	United States , Pennsylvania Schuylkill County	administrative division population 1,032, elevation 210m	N 40° 41' 21"	W 76° 10' 2"

Context and conversation (Siri)

- Coreference helps resolve ambiguities

U: “Book a table at Il Fornaio at 7:00 with **my mom**”

U: “Also send **her** an email reminder”

- Clarification questions:

U: “Chicago pizza”

S: “Did you mean pizza restaurants in Chicago or Chicago-style pizza?”