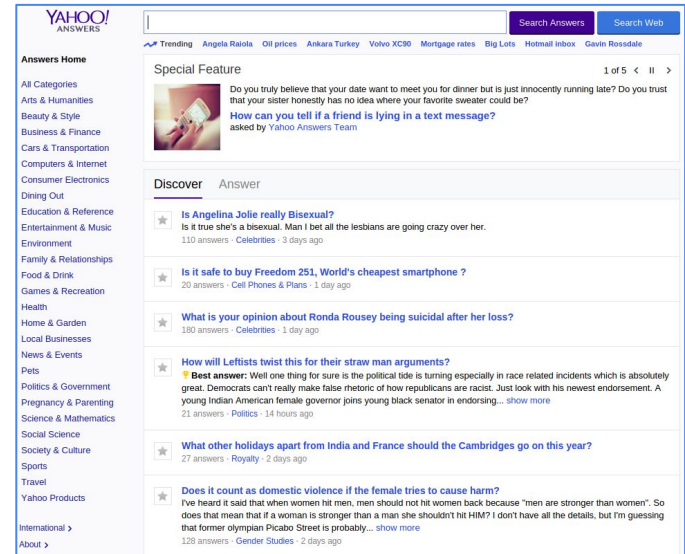
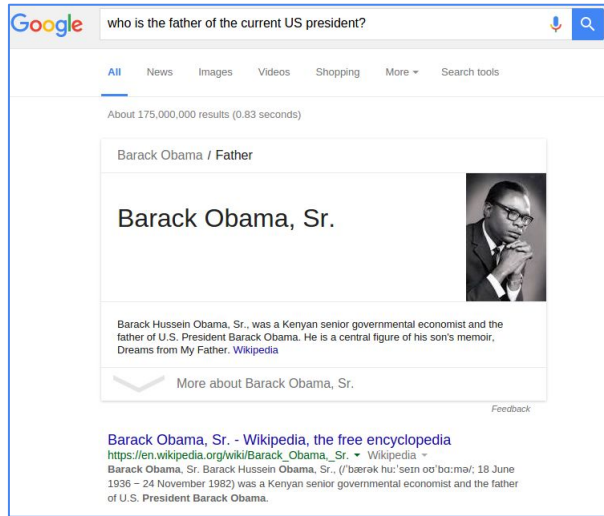
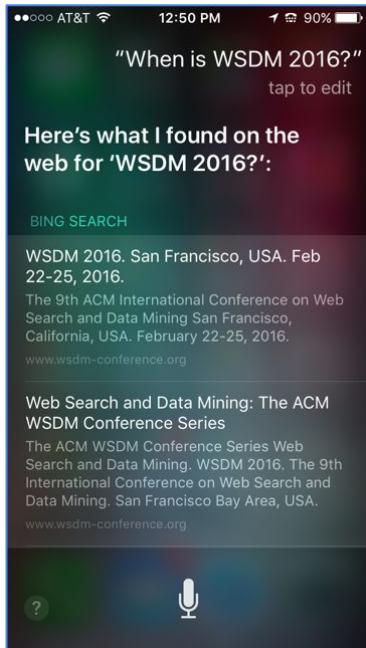


“Question answering using structured and unstructured data”

Denis Savenkov,
PhD student @ Emory University

Percentage of question search queries is growing^[1]



[1] “Questions vs. Queries in Informational Search Tasks”, Ryan W. White et al, WWW 2015

Factoid

WolframAlpha computational knowledge engine

what is the highest mountain in the usa?

Input interpretation: mountain with highest elevation in United States

Result: Denali (20 322 feet)

Properties:	
elevation	20 322 ft
country	United States
summit reached	1913 (103 years ago)

Image: who was the second person to walk on the moon?

Buzz Aldrin
Engineer

Buzz Aldrin is an American engineer and former astronaut, and the second person to walk on the Moon. He was the Lunar Module Pilot on in history. He set foot on the wing mission commander Neil officer and a Command ...

LinkedIn profile for Bradley Van Zile (1975 - 1978)

Jeopardy! game board with contestants Ken (\$300,000), Watson (\$1,000,000), and Brad (\$200,000).

(AP Photo/Jeopardy Productions, Inc.)

Non-factoid

Questions and Answers Natural Language Processing Asking Questions

Natural Language Processing: What is "Non-factoid question answering"?

Science & Mathematics > Astronomy & Space

How can I become a star?

I don't mean a Hollywood star.

Update: Yes, a glowing ball of hydrogen fusion. I will live for billions of years and provide life for cold, dead planets- thereby giving meaning to their otherwise pointless existence.

5 following 67 answers

1 Answer

Leonid Boytsov, PhD student at Carnegie Mellon University, Language Technologies Institute
225 Views

A non-factoid question answering (QA) is an umbrella term that covers all question-answering topics beyond factoid question answering. As a quick reminder: a factoid QA is about providing concise facts. For example, "who is the headmaster of Hogwarts?", "What is the population of Mars", and so on, so forth.

In contrast, a non-factoid question can be about anything. You can be asked to provide an answer to a math problem, to explain how to fix a specific model of a car, and so on, so forth. Answering multiple-choice questions also belongs to the area of non-factoid QA, though, there might be some overlap with factoid QA in this task.

An important sub-problem of non-factoid QA on community question answering sites such as

Written Jul 25, 2015 • View Upvotes

Upvote 2 Downvote Comment Share 1

In LITERATURE & LANGUAGE, DISSERTATIONS AND THESES, HUMAN RIGHTS

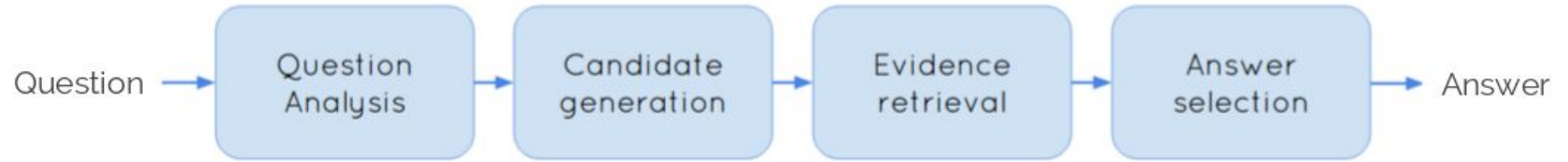
How do you choose PhD thesis topic?

Answered by The Community
Making the world better, one answer at a time.

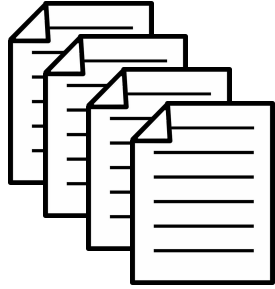
Choosing a PhD thesis topic can be difficult. The best way to approach selection of a topic is to select something within your specialized field of study that you would feel comfortable working with for several years and that you feel would be relevant several years down the line.

Was this answer useful?
YES SOMEWHAT NO

QA system architecture



Data Sources



2014 rank	City	State ^{RS}	2014 estimate	2010 Census	Change	2014 land area	2010 population density	Location
1	New York ^{RS}	New York	8,491,079	8,175,133	+3.86%	302.6 sq mi 783.8 km ²	27,012 per sq mi 10,430 km ⁻²	40.6643°N 73.9305°W
2	Los Angeles	California	3,928,864	3,792,621	+3.59%	468.7 sq mi 1,213.9 km ²	8,092 per sq mi 3,124 km ⁻²	34.0194°N 118.4108°W
3	Chicago	Illinois	2,722,389	2,695,598	+0.99%	227.6 sq mi 589.6 km ²	11,942 per sq mi 4,572 km ⁻²	41.8376°N 87.6313°W
4	Houston ^{TI}	Texas	2,239,958	2,100,263	+6.63%	599.6 sq mi 1,562.9 km ²	3,501 per sq mi 1,352 km ⁻²	29.7600°N 95.3663°W
5	Philadelphia ^{RI}	Pennsylvania	1,560,297	1,526,006	+2.25%	134.1 sq mi 347.3 km ²	11,379 per sq mi 4,384 km ⁻²	40.0094°N 75.1333°W
6	Phoenix	Arizona	1,537,058	1,445,632	+6.32%	516.7 sq mi 1,339.3 km ²	2,798 per sq mi 1,080 km ⁻²	33.5722°N 112.0880°W

San Francisco, California	
Consolidated city-county	
City and County of San Francisco	
Nickname(s): <i>The City by the Bay; Fog City; San Fran</i> ; ^[I] <i>Frisco</i> (locally disparaged); ^[I] ^[I] ^[I] <i>The City that Knows How</i> (past); ^[I] <i>Eighted by the Bay</i> (past); ^[I] <i>The Paris of the West</i> (past) ^[I] Motto: <i>Oro en Paz, Fierro en Guerra</i> (English: "Gold in Peace, Iron in War") Coordinates: 37°47'N 122°25'W	
Country	United States
State	California
CSA	San Jose–San Francisco–Oakland
Metro	San Francisco–Oakland–Hayward
Mission	June 29, 1776 ^[I]
Incorporated	April 15, 1850 ^[I]
Founded by	Jose Joaquin Moraga Francisco Palou St. Francis of Assisi
Named for	
Government	
• Type	Mayor-council
• Body	Board of Supervisors
• Mayor	Edwin M. Lee (D) ^[I] ^[I]
• Supervisors ^[I] ^[I]	List
• Assembly members ^[I] ^[I] ^[I]	David Chiu (D) Phil Ting (D)
• State senator	Mark Leno (D) ^[I] ^[I]
• United States Representatives ^[I] ^[I] ^[I]	Nancy Pelosi (D) Jackie Speier (D)

Food & Drink > Cooking & Recipes Next >

How do i cook lasagne?

Hi, i have never cooked lasagne before, i am house sitting and they have left me a fresh lasagne in a metal baking dish, do i put this in the oven? what temperature? how long should i cook it? i have no idea hey!

Follow 3 answers

Answers Relevance ▾

Best Answer: Hey! . Cook at 325 degrees in oven, on middle rack. If pan is really full, you may want to place on cookie sheet so it doesn't overflow as it heats up. How long you cook it depends on the depth of the pan. If the pan is three or so inches deep (like a "regular" lasagna pan), i would heat it for 45 minutes to an hour. You want the cheese to sort-of solidify. When it's nice and hot and bubbly and the ricotta is sort-of solid, take out of the oven and let "rest" (that is, let it sit for about 10 minutes or so) before you cut it, to let the cheese cool and set-up. Then, yum!

PS: If it's covered with plastic wrap, take that off first. If it's aluminum foil, you can leave it on

Sources^[I]
I've been making lasagna for 25+ years.....

C. L. H · 6 years ago



Unstructured data

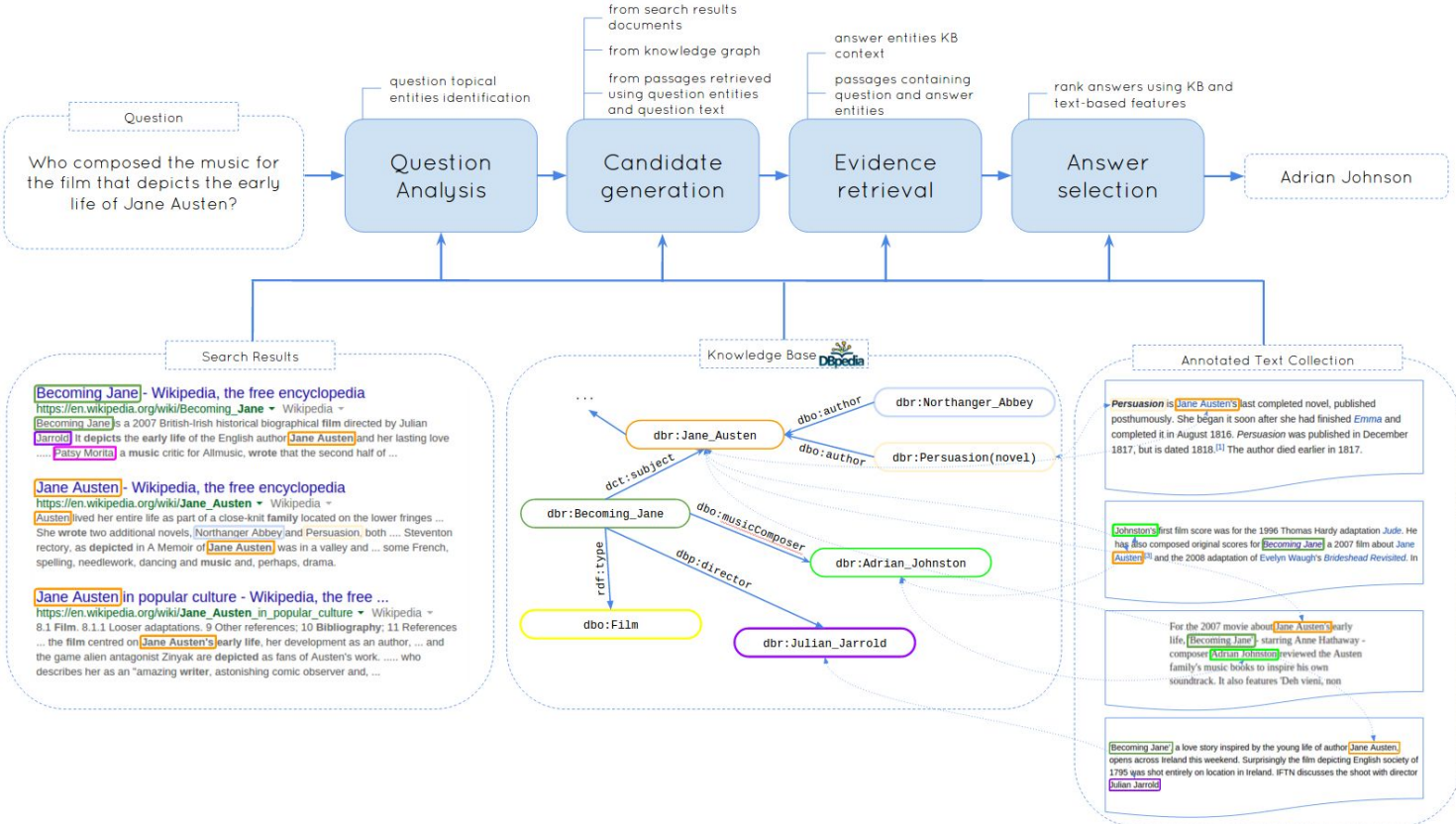
Semi-structured data

Structured data

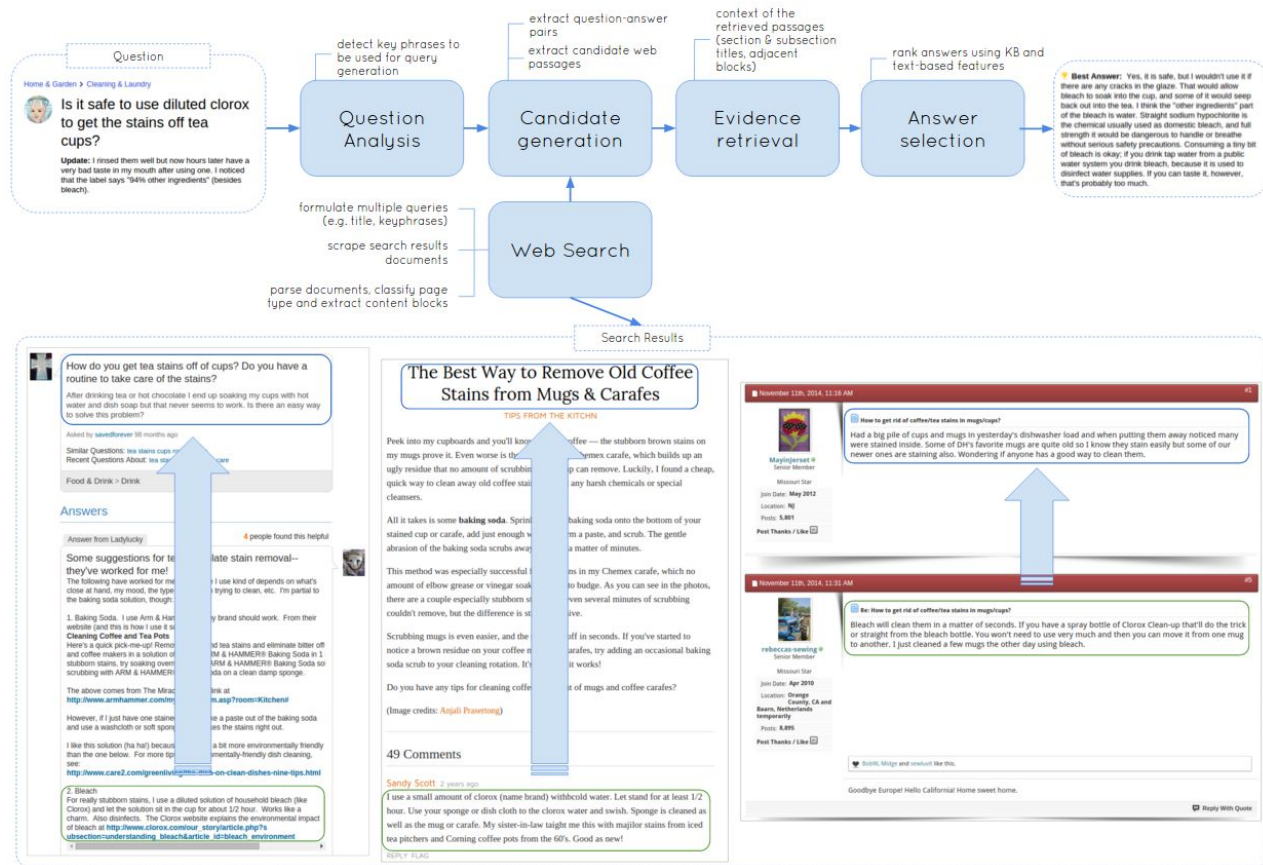
Structured and Unstructured data sources for QA

	Unstructured data	Structured data
factoid questions	<p><u>Text collections</u></p> <ul style="list-style-type: none">+ easy to match against question text+ cover a variety of different information types- each text phrase encodes a limited amount of information about mentioned entities	<p><u>Knowledge Bases (KB)</u></p> <ul style="list-style-type: none">+ aggregate all the information about entities+ allow complex queries over this data using special languages (e.g. SPARQL)- hard to translate natural language questions into special query languages- KBs are incomplete (missing entities, facts and properties)
non-factoid questions	<p><u>Text collections</u></p> <ul style="list-style-type: none">+ contain relevant information to a big chunk of user needs- hard to extract semantic meaning of a paragraph to match against the question (lexical gap)	<p><u>Question-Answer pairs</u></p> <ul style="list-style-type: none">+ easy to find a relevant answer by matching the corresponding questions- cover a smaller subset of user information needs

Combining text and KB for factoid QA



Using the structure of web documents for non-factoid QA



Research Questions

1. What types of questions can be answered using text, KB or a combination of both?
2. How does semantic annotation of unstructured data compare to information extraction for question answering?
 - information extraction for KB construction vs open information extraction vs unstructured data annotation
3. How does a combination of structured and unstructured data sources improve each of the main QA system components: question analysis, candidate generation, evidence extraction and answer selection?

Experiments: factoid QA

→ Datasets

- TREC QA
- WebQuestions
- new factoid QA dataset
 - ✓ derived from Yahoo! Answers CQA archive
 - ✓ filter factoid non-subjective questions (heuristics + mechanical turk)
 - ✓ mturk to label answers (KB entities, mentioned in CQA user answer)
 - ✓ Examples:
 - What is the name of Sponge Bob's pet snail?
 - What is the boot-shaped country found in Southern Europe?
 - Who was the first model to ever appear on the cover of Sports Illustrated?

→ Comparison

- state-of-the-art KBQA systems (WebQuestions benchmark)
- Open QA approach based on OpenIE extractions
- collection-based and hybrid techniques: AskMSR+, QuASE

Experiments: non-factoid QA

- TREC LiveQA shared task (<http://trec-liveqa.org/>): real user question answering in real-time
 - systems need to respond to questions posted to Yahoo! Answers
- Comparison:
 - baseline system developed for last year shared task
 - participants of TREC LiveQA 2016

[Beauty & Style](#) > [Fashion & Accessories](#)

[Next](#) >



White after Labor Day?

Is wearing no white after Labor Day still a thing? I have talked to many people who said they don't go by that and I also spoke to many people who said to not do it. What do you think?



Follow



4 answers

[Beauty & Style](#) > [Fashion & Accessories](#)

[Next](#) >



What are the price differences at outlet malls?

I live in Florida and I'm going to st augustine outlet mall because I really need clothes and I was wondering what is the price difference from a regular mall ?



Follow



1 answer

[Computers & Internet](#) > [Hardware](#) > [Desktops](#)

[Next](#) >



No first boot after PC opens?

This happened when my brother was using the PC. After a few hours, the PC suddenly hanged. Restarted the PC and heard 3 beeps. So, what we did was we moved the RAM to another slot. But after opening, no more first boot sound. What could be the possible problem? I already did the following: - re-attached hard disk.... [show more](#)



Follow



1 answer



Expected results

1. new factoid question answering dataset
2. new approach to combining unstructured text and structured KB data, that:
 - ✓ improves QA precision
 - ✓ improves recall, including questions that could only be answered with a combination of text and KB data
3. new system for non-factoid question answering with improved performance due to better utilization of the information provided in web documents

Open Questions

- How to incorporate other available sources of factual information, including semi-structured, e.g. tables, diagrams, etc?
- Non-factoid questions often include unique context information, that makes reusing information non effective. How a system can generate the answer given all potentially useful extracted information?
- How to construct KBs for non-factoid information needs?

Thank you!

Society & Culture > Languages

Next >



How to say thank you in different languages (German, Italian,.....)?

☆ 2 following ✉ 31 answers



Answers

Relevance ▾



👑 **Best Answer:** spanish - gracias

french - merci
german - danke
italian - grazi
portugese - obrigado
norwegian - takk
mandarin chinese - xie xie
finnish - kiitos
danish - tak
hebrew - toda
russian - spasibo
swahili - asante
swedish - tack
polish - dziekuje
korean - kamsa hamnida
japanese - arigato
indonesian - termi kasih
greek - efharisto
hindi - sukria

Krysti · 10 years ago



Comment