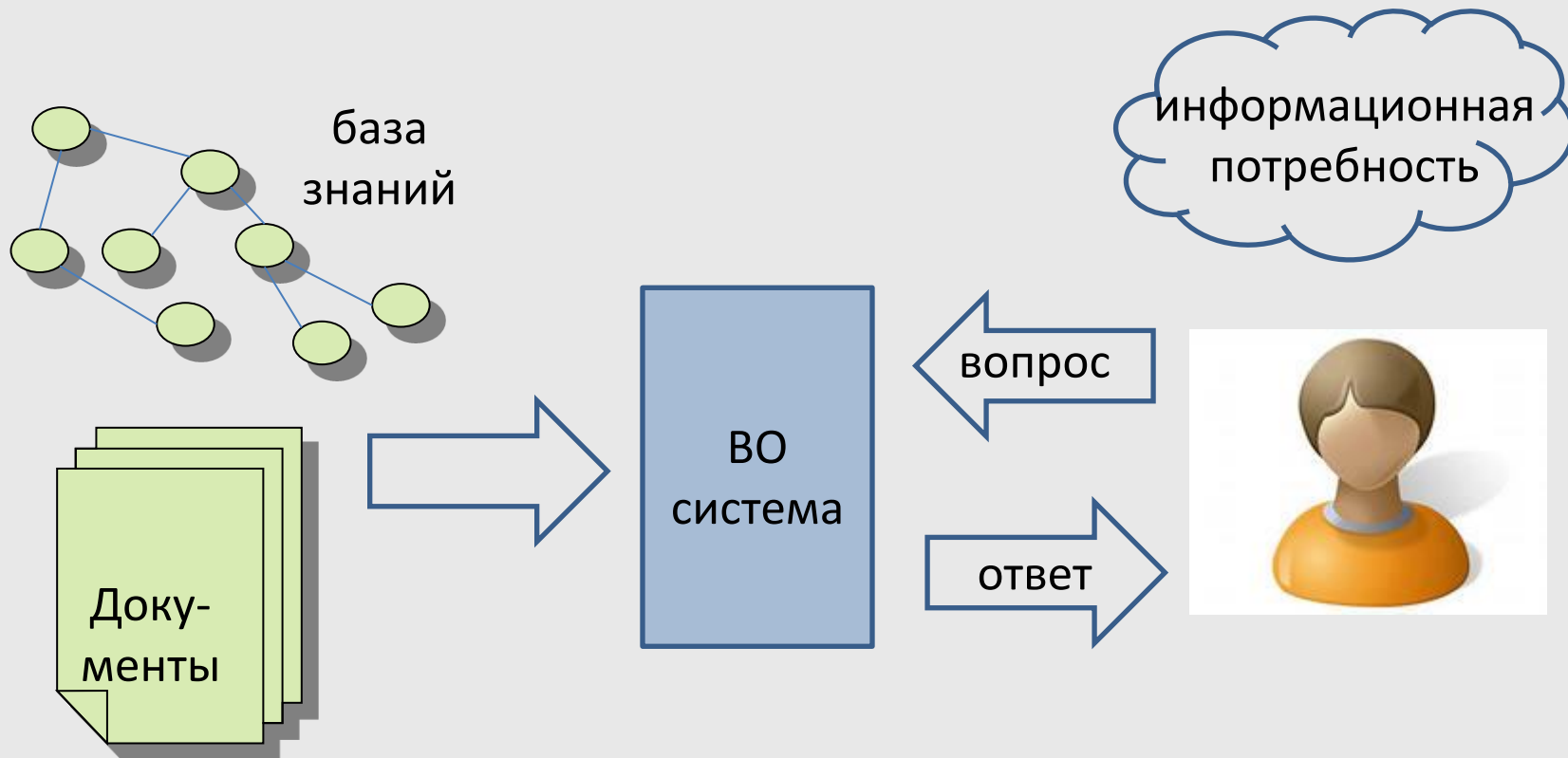


Вопросно-ответный поиск (question answering, QA)

Павел Браславский

ЧТО ЭТО ТАКОЕ?





usa

Search

[Advanced Search](#)

Web › [past 2 minutes](#)

[Hide options](#)

Results 1 - 1

› **All results**

[Videos](#)

[Forums](#)

[Reviews](#)

[Any time](#)

[Recent results](#)

[Past 24 hours](#)

[Past week](#)

[Past year](#)

[Specific date range](#)

› **past 2 minutes**

› **Sorted by relevance**

[Sorted by date](#)

› **Standard view**

[Related searches](#)

[Wonder wheel](#)

[Timeline](#)

› **Standard results**

[Images from the page](#)

[Prices from the page](#)

[More text](#)

[Reset options](#)

[G&L Legacy USA \(strat\)](#)

1 minute ago - 2004 G&L **USA** Legacy (stratocaster), trans red body (alder, nice grain), hard maple neck with thin nitro natural finish and no fret markers (standard 1 5/8 ... [denver.craigslist.org/msg/1372842961.html](#) - [Similar](#)

[Quotes from USA Today - News, photos, topics, and quotes](#)

1 minute ago - Somehow or other I think we will wind up seeing a range of estimates. [topics.treehugger.com/.../0fnpcSJ3wB7Hz?q...](#) - [Similar](#)

[USA Open: Womens Final - Page 2 - FC Barcelona Forum](#)

2 posts - 2 authors - Last post: 1 minute ago
USA Open: Womens Final Polls. ... Go Back, FC Barcelona Forum > General Chat > Polls · Reload this Page **USA** Open: Womens Final ... [www.barcaforum.com/showthread.php?p=140212](#) - [Similar](#)

[Sports News Headlines - 1stHeadlines](#)

55 seconds ago - **USA** Today-Sports. Del Potro routs Nadal, will take on five-time champ Federer in ... **USA** Today-Sports. Eli Manning, Osi Umenyiora lead Giants past Redskins ... [www.1stheadlines.com/sports.htm](#) - [Similar](#)

[How do you ship a dog in the USA? - Yahoo! Answers](#)

1 minute ago - We have puppies for sale online, and I have receiv... [answers.yahoo.com/question/index?qid...](#) - [Similar](#)

[What do you think about Miley Cyrus's song "party in the U.S.A ...](#)

1 minute ago - what do you think of the song. don't judge it by what you saw on TCA. like ... Bob Saget was on iTunes and over 14000 people gave it one star. LOL ... [answers.yahoo.com/question/index?qid=20090913163115AAcKiP8](#) - [Similar](#)

[Azcentral - Lidge struggles to close it out, as Phillies hang on ...](#)

1 minute ago - Most emailed stories from **USA** TODAY. Nadal finishes rain-interrupted win at Open to reach semis - USATODAY.com21h 19m ago; The sleepers: Five teams who ... [content.usatoday.net/diet/custom/acid/Page.aspx?cid...](#) - [Similar](#)



About 227,000,000 results (0.91 seconds)

Donald Trump / Age

70 years

June 14, 1946



Hillary Clinton
69 years



Melania Trump
Spouse
46 years



Barack Obama
55 years

Feedback

People also ask

Where was Trump born? ▾

Is Donald Trump's wife an American citizen? ▾

Why is Donald Trump so rich? ▾

How long has Donald Trump been married? ▾

[Donald Trump - Wikipedia](#)

https://en.wikipedia.org/wiki/Donald_Trump ▾

At **70 years old**, he will be the oldest person ever to assume the presidency. Trump was born and raised in New York City and received a bachelor's degree in economics from the Wharton School in 1968.

[Donald Trump Jr.](#) · [Donald Trump presidential](#) · [Las Vegas](#) · [Reform Party](#)

About 80,400,000 results (0.56 seconds)

The iPhone 6 measures 5.44 x 2.64 x 0.27 inches (138.1 x 67 x 6.9mm) and weighs **4.55 ounces (129g)** – a weight increase that is roughly proportional to its 16% volume increase compared to the iPhone 5S. Sep 9, 2014



[iPhone 6 vs iPhone 6 Plus: The Differences Between The ...](#)

www.forbes.com/.../iphone-6-vs-iphone-6-plus-what-is-the-differenc... Forbes

Feedback

[iPhone 6 - Technical Specifications - Apple](#)

www.apple.com > [iPhone](#) > [iPhone 6](#) Apple

Weight: 4.55 ounces (129 grams) 6.22 inches (158.1 mm) 3.06 inches (77.8 mm) 0.28 inch.

[iPhone - Compare Models - Apple](#)

www.apple.com/iphone/compare/ Apple

Available in silver and space gray, **iPhone 6 Plus** features an A8 chip, Touch ID, ultrafast LTE wireless, an 8MP iSight camera, and ... **Weight and Dimensions**² ...

[Apple iPhone 6 - Full phone specifications - GSMarena.com](#)

IBM Watson (2011)



Wolfram Alpha



WolframAlpha[™] computational...
knowledge engine

How old was Queen Elizabeth II in 1974? 

Input interpretation:

age of Queen Elizabeth II (royalty) in 1974

Result for start of 1974: [Show details](#)

47 years

Computed by: [Wolfram Mathematica](#) [Source information »](#) [Download as: PDF](#) | [Live Mathematica](#)

Мотивация

- *ответ*, а не список документов
- быстрое принятие решений
- мобильные устройства
- голосовой ввод
- поиск сущностей
- поиск как диалог
- рост доли *вопросов* в логе

Вопросы в логе

- 2-4% всех запросов, доля растет [Pang and Kumar, 2011]
- ЕЯ и длинные запросы [Hearst 2011]
- до недавнего времени результаты хуже, чем по запросам из ключевых слов [Aula et al., 2010; Bendersky and Croft, 2009; Pang and Kumar, 2011]
- *как (how)* – самые частые вопросы (не факты)

Почему больше вопросов?

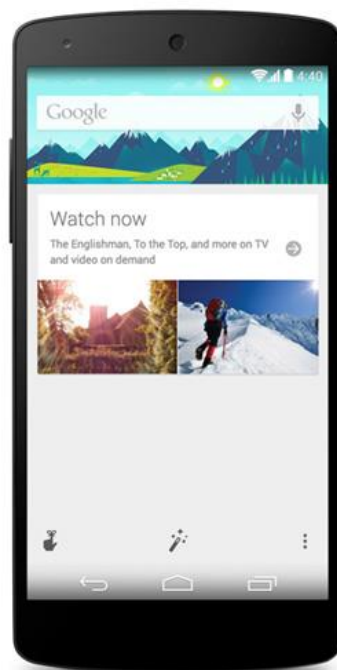
- Естественный интерфейс
 - голосовой поиск (!)
- Плохие навыки поиска/сложно сформулировать запрос
- Контент в форме «вопрос-ответ»
- Повышение качества поиска

Поиск – одна из функций

Apple Siri



Google Now



Windows Cortana



ПОДХОДЫ

Вопросы и ответы

- Вопросы:
 - поиск фактов /остальные (процедурные – «как»)
 - субъективные/объективные
 - информативные/риторические/коммуникативные
- Ответы:
 - тексты
 - базы знаний
 - гибридные

Поиск в текстовых коллекциях

- обработка запроса
 - тип вопроса, тип ответа, фокус, отношения
 - запрос к машине поиска
- поиск фрагментов
 - ранжированный список документов
 - разбивка на фрагменты, ранжирование
- обработка ответов
 - ответы-кандидаты (сущности)
 - ранжирование на основе отношений в тексте и внешних источников

Подход на основе баз знаний

- Семантические представление запроса
 - время, даты, места, сущности, числовые значения,...
- Запрос к структурированным базам
 - пространственные базы данных
 - онтологии (инфобоксы Wikipedia, DBPedia, WordNet, Yago)
 - Wolfram Alpha
 - ...

Например, Siri

Гибридный подход (IBM Watson)

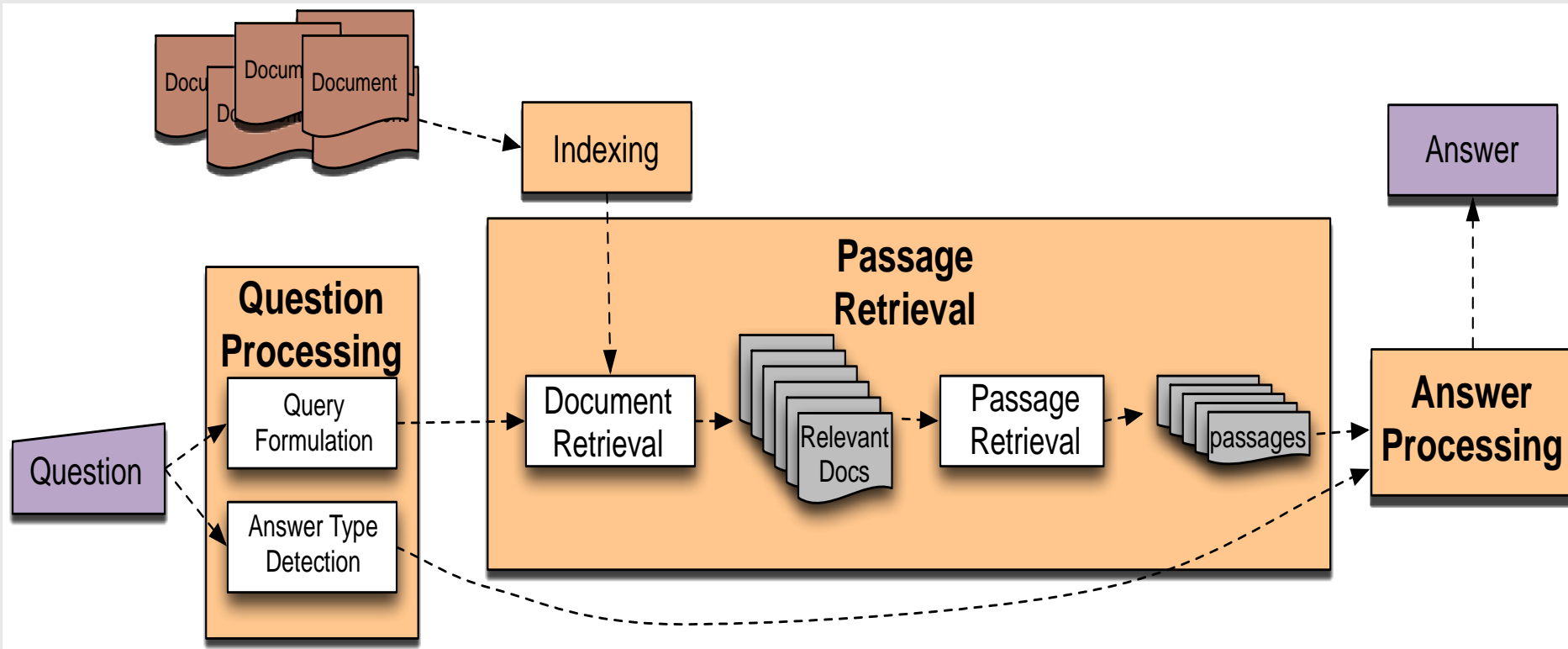
- поверхностное семантическое представление запроса
- генерация ответов-кандидатов на основе поиска
 - дополненная онтологиями и слабо структурированными данными
- оценка кандидатов на основе дополнительных источников /методов
 - пространственные базы данных
 - темпоральный вывод
 - таксономий

ПОИСК ФАКТОВ В КОЛЛЕКЦИЯХ ТЕКСТОВ

Основные этапы

- Анализ вопроса
 - Определение типа *ответа*
 - Определение типа *вопроса*
 - Вопрос → запрос к МП
 - Фокус вопроса
- Анализ ответа МП
 - сегментация на предложения/абзацы
 - взвешивание сегментов
 - формирование ответа

Поиск фактов в текстовой коллекции



[Dan Jurafsky]

Пример

Через сколько дней появляются корни у отростка традесканции?

- Тип ответа: **число (дней)**
- Тип вопроса: **факт**
- Запрос: **традесканция, корни, отросток**
- Фокус: **дни**
- Отношение: **появляются (отросток, корни) ?**



Ответы:

...достаточно отломить черенок любой длины и поставить его в воду или воткнуть во влажный грунт, и **через сутки** черенки уже выпускают корни.

Оборвите толстые кончики, поставьте их в воду, **дня через 3-4** на них уже будут корни.

Традесканции очень легко размножаются черенками в любое время года. Укореняются **в течение нескольких дней** в песчаной почве или даже в воде.

Ожидаемый тип ответа

Самая высокая гора?

→ географическое название

Какова высота Эвереста?

→ число (метры)

Кто первым покорил Эверест?

→ человек

Методы: правила и/или машинное обучение

Признаки: вопросительные слова, структура, NEs, POS, семантически близкие слова, ...

Типология вопросов UIUC

	Coarse	Fine	
6	ABBREVIATION	abbreviation, expansion	50
	ENTITY	definition, description, manner, reason	
	DESCRIPTION	animal, body, color, creation, currency, disease/medical, event, food, instrument, language, letter, other, plant, product, religion, sport, substance, symbol, technique, term, vehicle, word	
	HUMAN	description, group, individual, title	
	LOCATION	city, country, mountain, other, state	
	NUMERIC VALUE	code, count, date, distance, money, order, other, percent, period, speed, temperature, size, weight	

5,500 размеченных вопросов

<http://cogcomp.cs.illinois.edu/Data/QA/QC/>

Jeopardy/Своя игра

- Именно в этом году Суворов перешёл Альпы
- Этот писатель владел имением Ясная Поляна
- Именно это нужно сделать с нефтью, чтобы получить из неё керосин
- Однажды Леонид Кучма запнулся, позабыв, как по-украински „дрова“. Так и не вспомнил. А вы вспомните?

Типы ответов Jeopardy

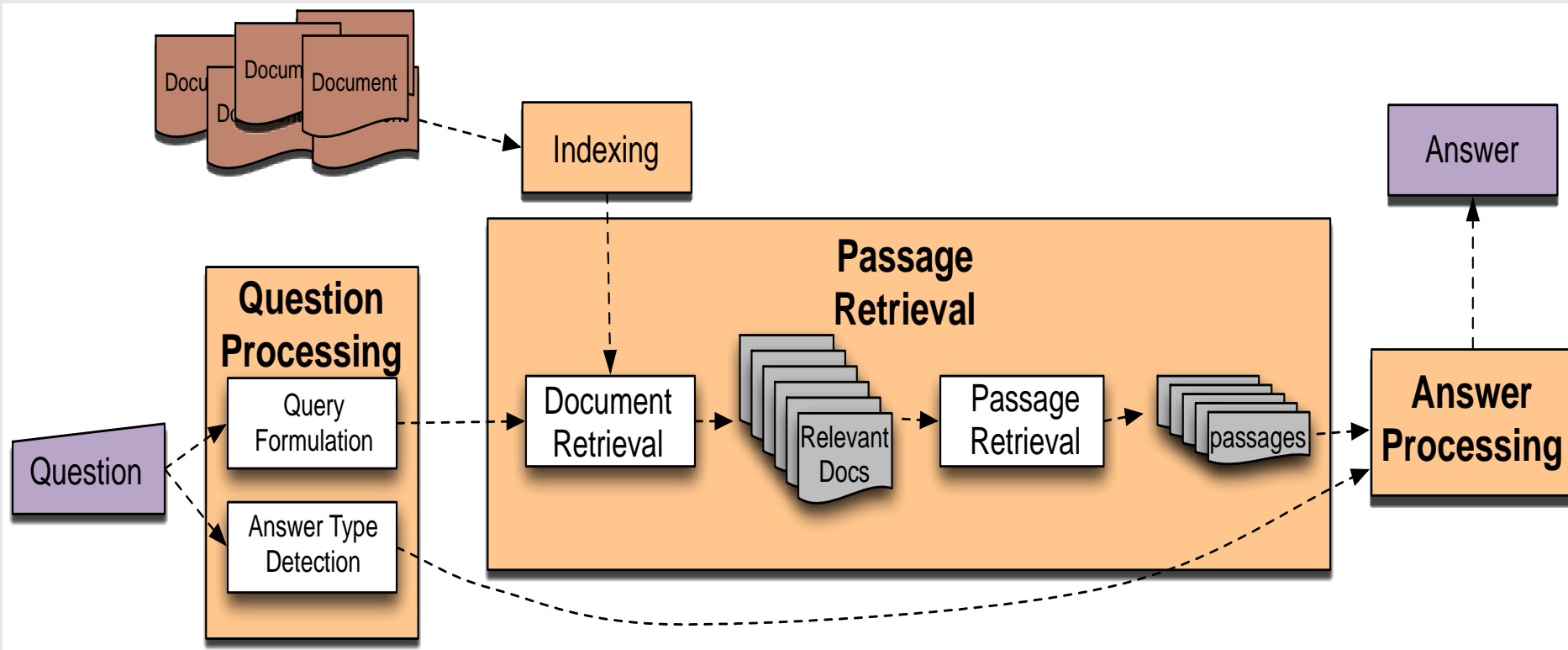
Ferrucci et al. Building Watson: An Overview of the DeepQA Project.
AI Magazine. Fall 2010.

- 2,500 типов в 20,000 вопросах
- Самые частотные 200 типов покрывают меньше половины вопросов
- Топ-40 типов ответов:

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

[Dan Jurafsky]

Поиск фактов в текстовой коллекции



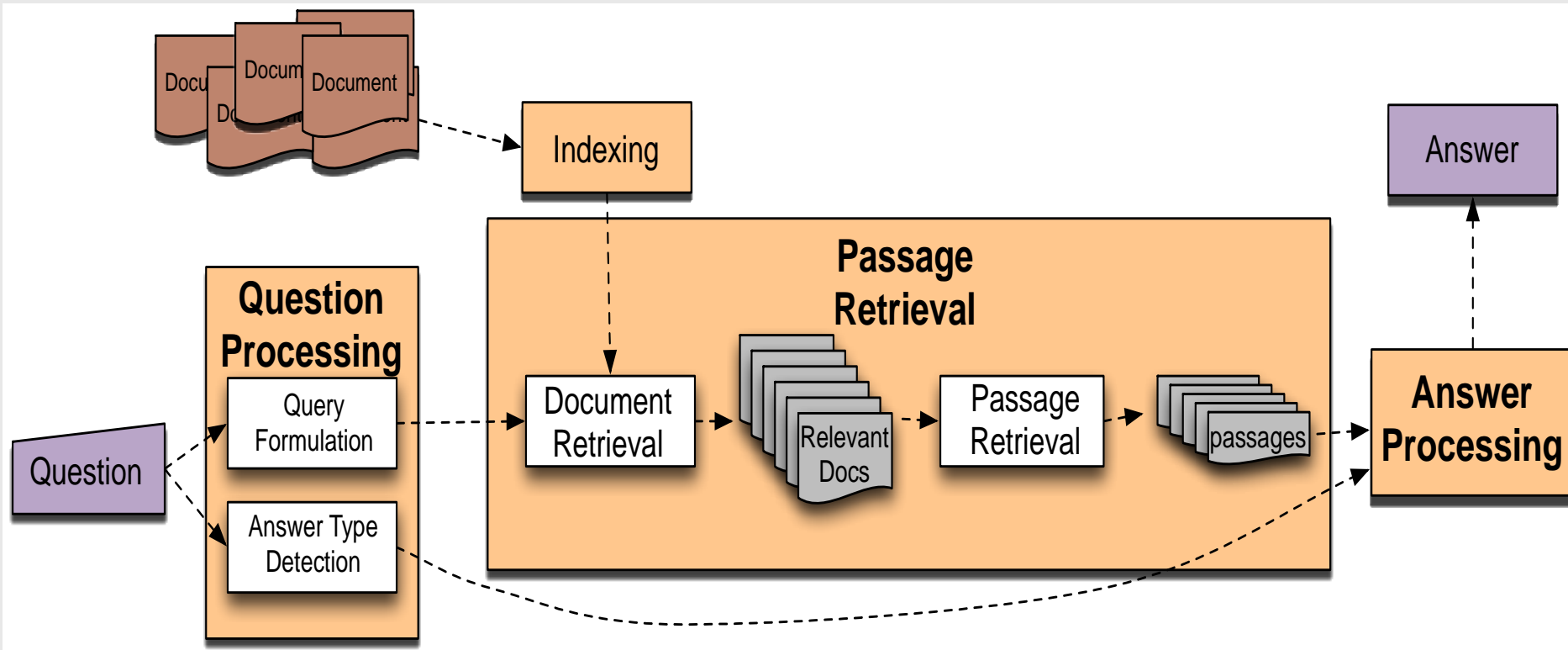
[Dan Jurafsky]

Выбор ключевых слов

Dan Moldovan, Sanda Harabagiu, Marius Poca, Rada Mihalcea, Richard Goodrum, Roxana Girju and Vasile Rus. 1999. TREC-8.

1. слова в кавычках
2. имена собственные в распознанных NEs
3. именные группы прилагательное + существительное
4. остальные именные словосочетания
5. все остальные существительные
6. глаголы
7. наречия
8. слова, отражающие фокус вопроса
9. остальные слова

Поиск фактов в текстовой коллекции



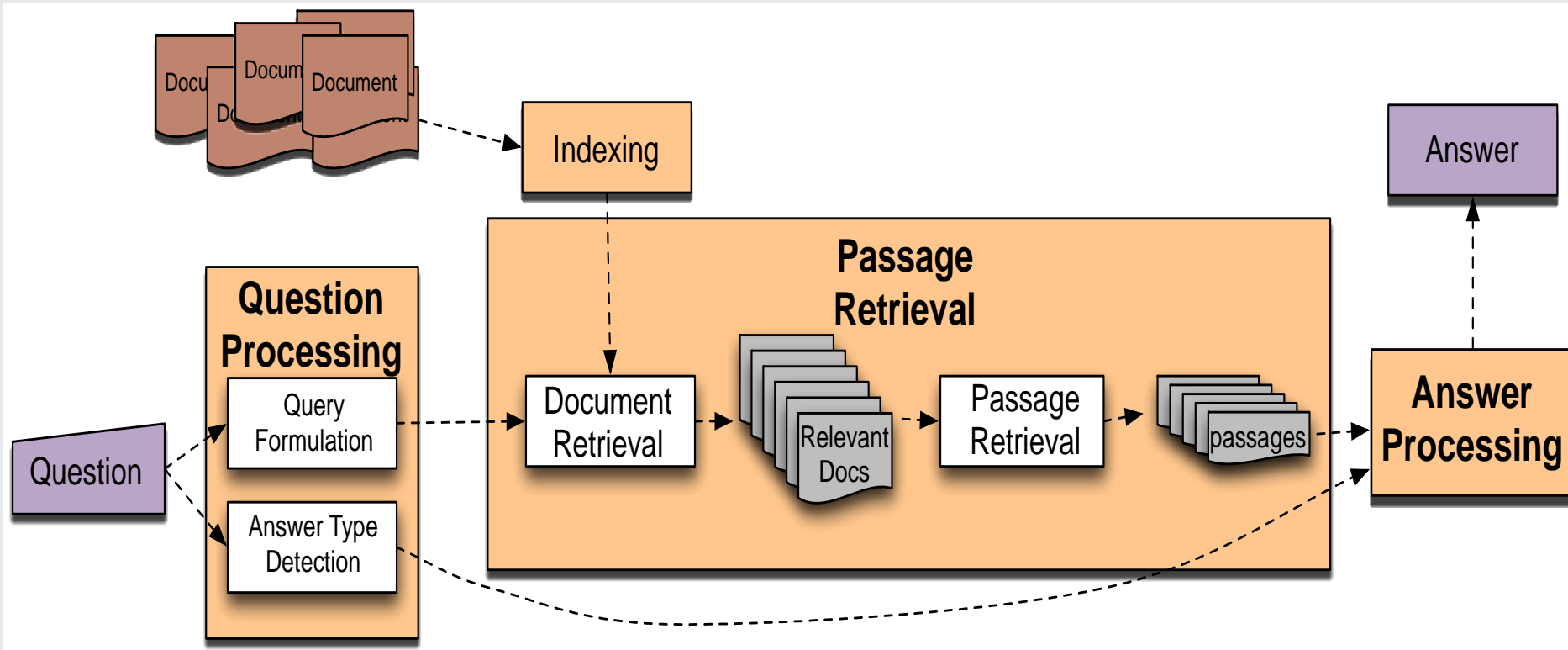
[Dan Jurafsky]

Поиск фрагментов: признаки

- Поиск именованных сущностей ожидаемого типа
- Количество слов, n-грамм из вопроса
- Близость слов запроса
- Последовательность слов вопроса наибольшей длины
- Позиция документа

[Dan Jurafsky]

Поиск фактов в текстовой коллекции



[Dan Jurafsky]

Признаки для ответов

- **Тип ответа**
- **Шаблон ответа** (*X дней*)
- **Ключевые слова из вопроса**
- **Расстояние между ответом и словами запроса**
- **Новизна** (слова не из вопроса)
- **Примыкание** (близость словам вопроса)
- **Пунктуация** (знак пунктуации непосредственно после ответа)
- **Длина последовательности слов из вопроса в ответе**

[Dan Jurafsky]

Не-фактографические вопросы

Pattern	Question	Answer
<AP> such as <QP>	What is autism?	“, <u>developmental disorders</u> such as autism”
<QP>, a <AP>	What is a caldera?	“the Long Valley caldera, a <u>volcanic crater</u> 19 miles long”

[Jurafsky & Martin, SLP3]

Ответы-кандидаты: пример

Вопрос:

Who was Queen Victoria's second son?

- Тип ответа: Person
- Фрагмент:

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

[Dan Jurafsky]

Взвешивание ответов в IBM Watson

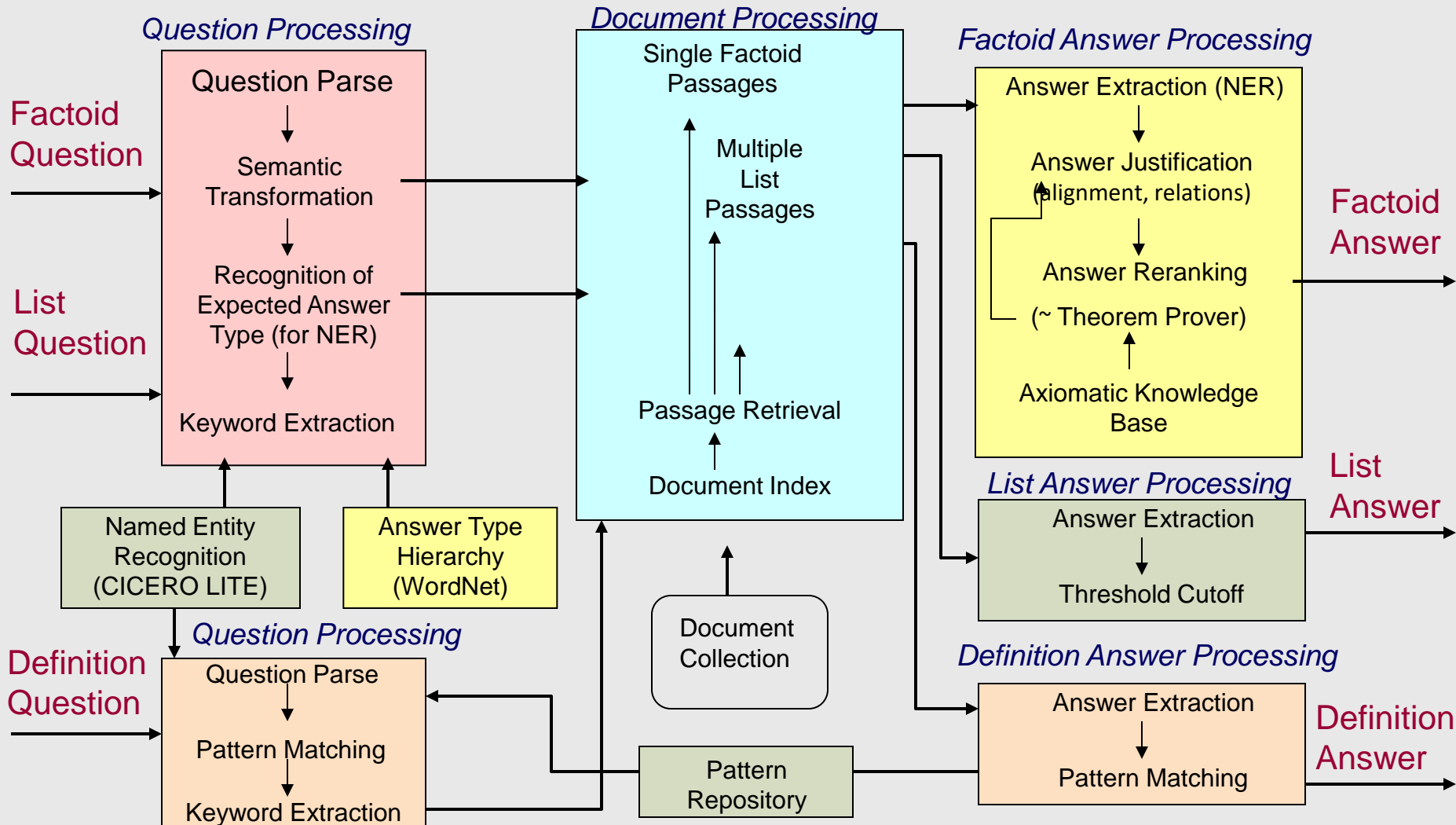
Каждый ответ-кандидат получает оценку от 50+ компонентов

- соответствие синтаксической структуры на основании разбора вопроса и ответа
- надежность источника
- географические и пространственные базы
- темпоральные отношения
- таксономические отношения

[Dan Jurafsky]

Full NLP QA: LCC (Harabagiu/Moldovan)

[below is the architecture of LCC's QA system circa 2003]



DrQA: Open-domain Question Answering

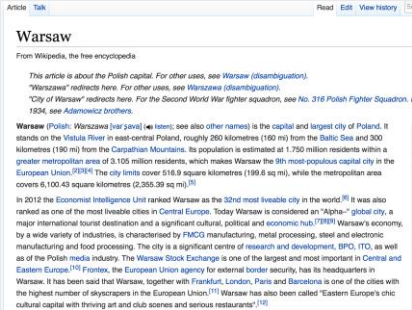
(Chen, et al. ACL 2017) <https://arxiv.org/abs/1704.00051>



Q: How many of Warsaw's inhabitants spoke Polish in 1933?



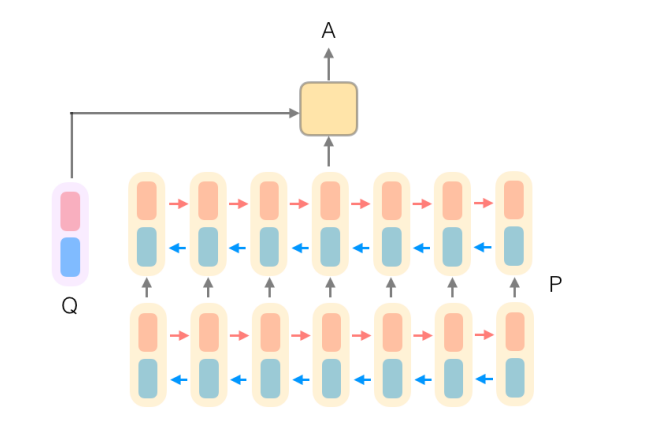
Document Retriever



Document Reader



833,500



[Manning & Nayak]


Open-domain Question Answering

SQuAD

Q: How many of Warsaw's inhabitants spoke Polish in 1933? **A:** 833,500

TREC

Q: What U.S. state's motto is "Live free or Die"? **A:** New Hampshire

WebQuestions (Berant et al, 2013) 

Q: What part of the atom did Chadwick discover? **A:** neutron

WikiMovies (Miller et al, 2016)

Q: Who wrote the film Gigli? **A:** Martin Brest

OMDb



Document Retriever

Dataset	Wiki Search	Doc. plain	Retriever +bigrams
SQuAD	62.7	76.1	77.8
CuratedTREC	81.0	85.2	86.0
WebQuestions	73.7	75.5	74.4
WikiMovies	61.7	54.4	70.3

Traditional
tf.idf
inverted
index +
efficient
bigram
hash

70-86% of questions we have that the answer segment appears in the top 5 articles

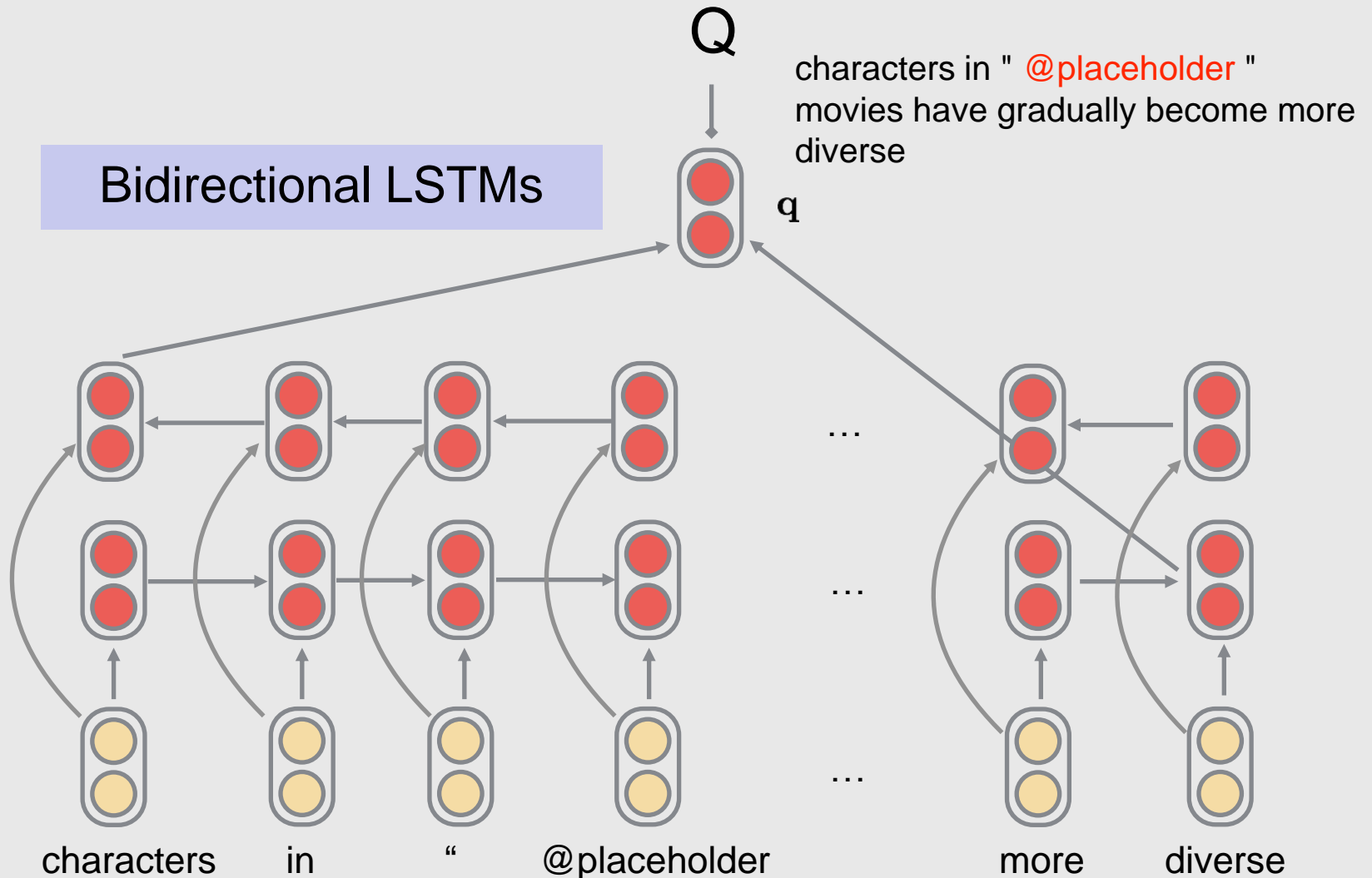
Document Reader

- Многослойный bLSTM
- Абзац (p), для каждого слова:
 - GloVe
 - точное совпадение со словом вопроса (0/1)
 - POS, NER, TF
 - взвешенная сумма близостей со словами вопроса
- Вопрос (q):
 - взвешенная сумма векторов слов
- Цель – вероятность
 - начального символа $p_i W_s q$
 - конечного символа $p_i W_e q$

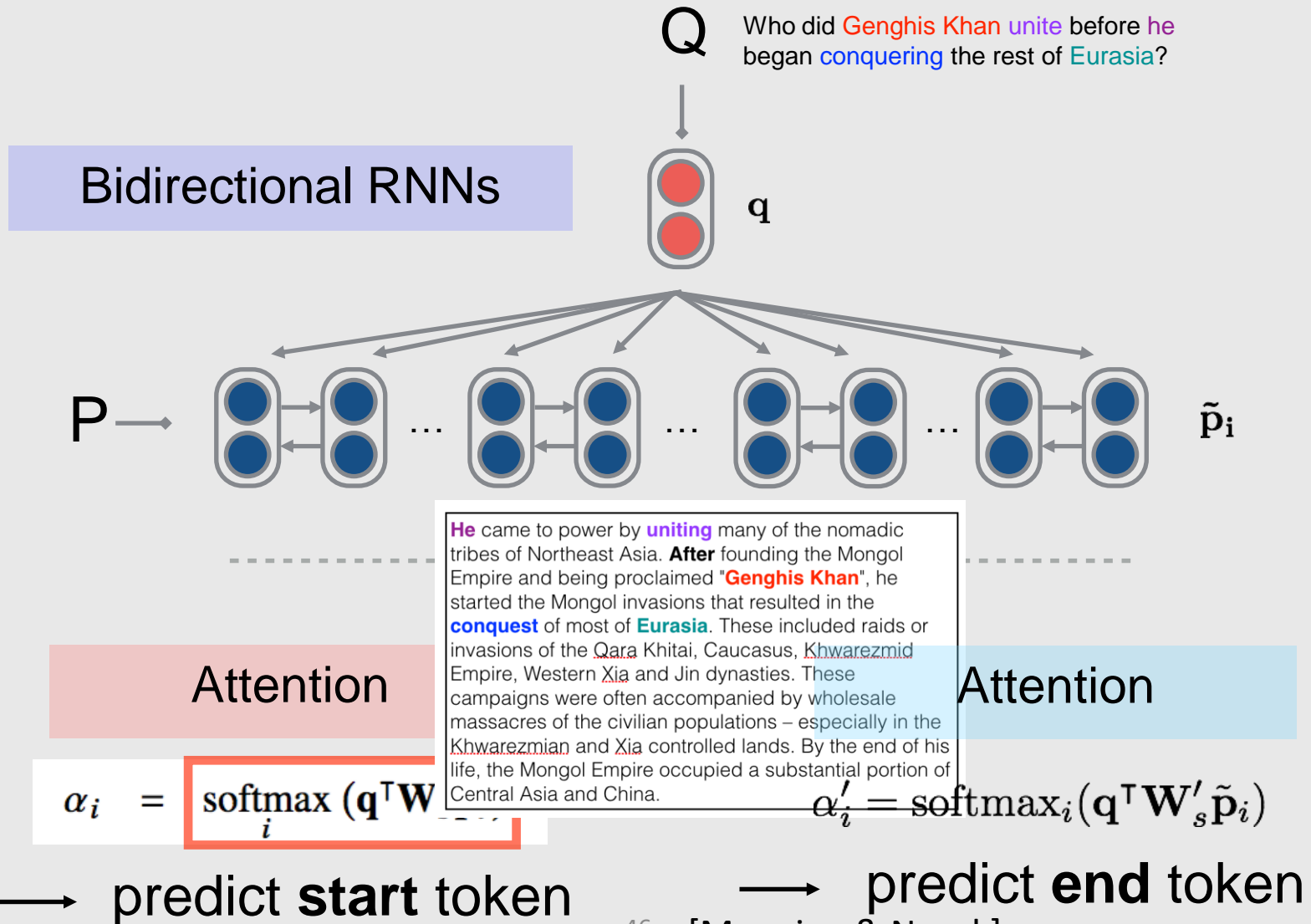
Данные

- Википедия (plain text, 5М статей)
- SQuAD
 - 87K + 10K: вопрос + абзац Википедии + фрагмент
- Другие наборы данных Q+A
 - CuratedTREC, WebQuestions, WikiMovies
 - distant supervision & тестирование

Document Reader:



Stanford Attentive Reader++



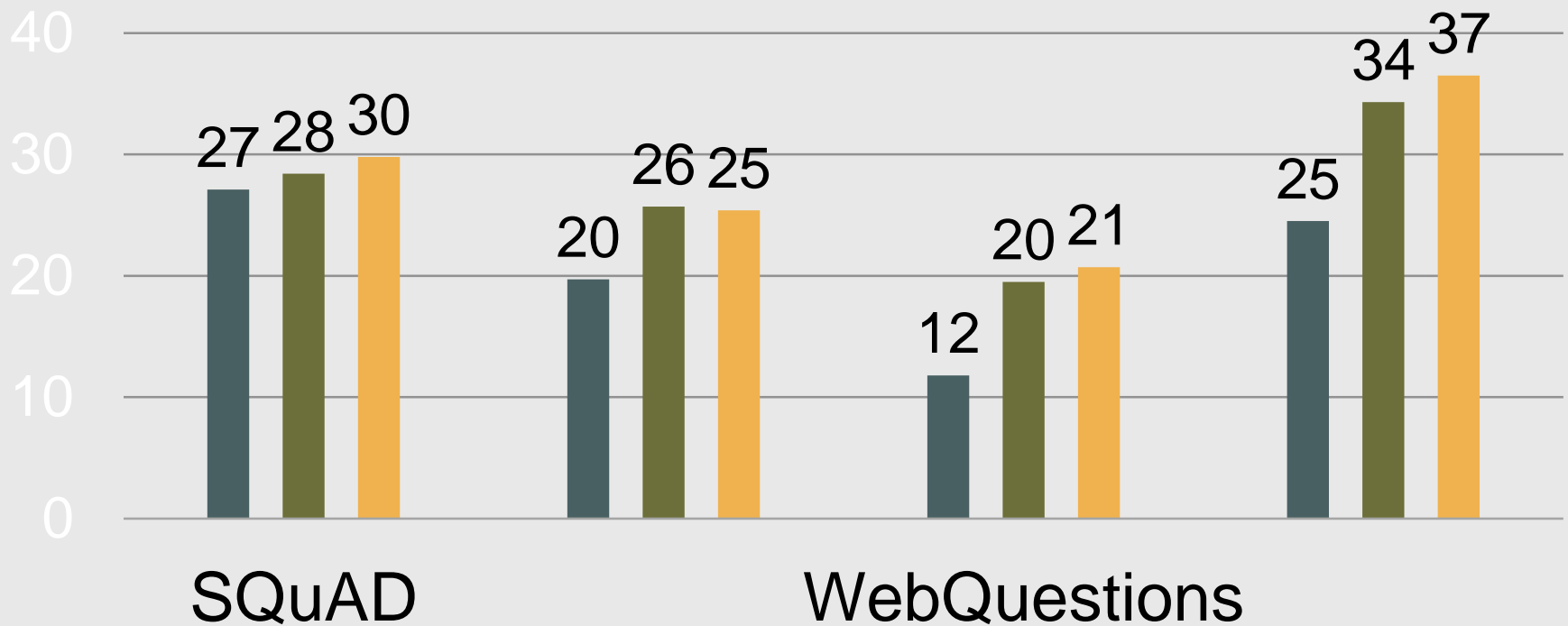
Results (single model)

	F1
Logistic regression	51.0
Fine-Grained Gating (Carnegie Mellon U)	73.3
Match-LSTM (Singapore Management U)	73.7
DCN (Salesforce)	75.9
BiDAF (UW & Allen Institute)	77.3
Multi-Perspective Matching (IBM)	78.7
ReasoNet (MSR Redmond)	79.4
DrQA (Chen et al. 2017)	79.4
r-net (MSR Asia)	80.8
Human performance	91.2

Results

- Pre-trained SQuAD
- Fine-tuning

Exact match
(top-1 prediction)

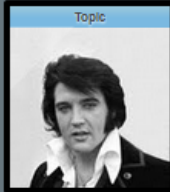


ПОИСК ПО БАЗЕ ЗНАНИЙ

Freebase

[Browse](#) [Query](#) [Help](#)[Sign In or Sign Up](#)[English](#) ▾

Important! Freebase is read-only and will be shut-down. [More](#).



Topic

Elvis Presley ^{en}

mid: /m/02jq1 notable type: /music/artist on the web: [Wikipedia.org](#)

Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply, "the King". Born in Tupelo, Mississippi, Presley and his family moved to Memphis, Tennessee, when he was 13 years old. His music career began there in 1954, when he started to work with Sam Phillips, the owner of Sun Records. Accompanied by guitarist Scotty Moore and bassist Bill Black, Presley was an early popularizer of rockabilly, an uptempo, backbeat-driven fusion of country music and rhythm and blues. RCA Victor acquired his contract in a deal arranged by Colonel Tom Parker, who was to manage the singer for more than two decades. Presley's first RCA single, "Heartbreak Hotel", released in January 1956, was a number-one hit in the US. He became the leading figure of rock and roll after a series of network television appearances and chart-topping records. [-]

Created by [mwcl_musicbrainz](#) on 9/27/2013

[Properties](#)[118n](#)[Keys](#)[Links](#)

Filter options: Show all domains and properties

Common /common

Freebase Commons

Topic /common/topic

X

Also known as /common/topic/alias

Also known as

- Elvis
- Elvis Aron Presley
- The King of Rock 'n' Roll
- Elvis Aaron Presley
- King of Rock and Roll
- Elvis, the pelvis
- The King
- The King of Rock and Roll
- "The Pelvis "

[44 values total](#)

Description /common/topic/description

Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply "the King".

Types:

Common
Topic

Film
Film music contributor
Film actor
Film subject
Person or entity appearing in film
Film song performer

Music
Musical Artist
Musician
Composer
Lyricist
Featured artist

Business
Product theme



About: Юрий Алексеевич Гагарин

[Goto](#) [Sponge](#) [NotDistinct](#) [Permalink](#)

An Entity of Type : <http://www.wikidata.org/ontology#Item>, within Data Space : dbpedia.org associated with source [document\(s\)](#)

Type: Command:

Attributes Values

[type](#) [Человек](#)
<http://www.wikidata.org/ontology#Item>

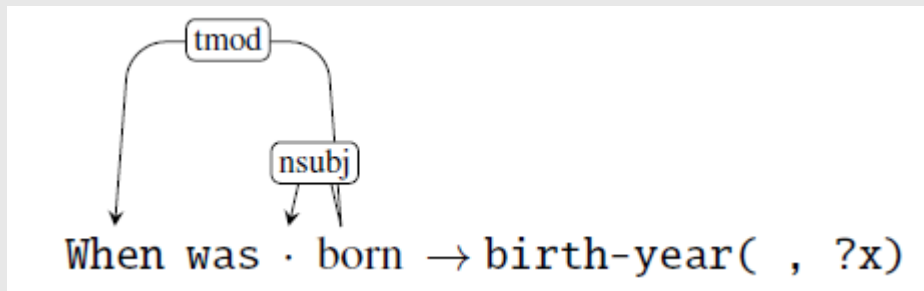
[wikidata:P1741c](#) 100616

[label](#) Yuri Gagarin
 Yuri Gagarin
 Yuri Gagarin

wikidata:P119c	Некрополь у Кремлёвской стены
wikidata:P227c	118537105
wikidata:P241c	Центр подготовки космонавтов имени Ю. А. Гагарина Военно-воздушные силы СССР
wikidata:P245c	500342574
wikidata:P26c	Гагарина, Валентина Ивановна
wikidata:P40c	Гагарина, Елена Юрьевна
wikidata:P935c	Юрий Алексеевич Гагарин
wikidata:P535s	wikidata:Q7327577B6D582-2641-4454-AA9A-A11CFE480837
wikidata:P1263c	666/000026588

Semantic parser

- Вопрос → формальное представление
 - логическое выражение, отношение с джокером, SQL, SPARQL, ...
- методы
 - Ручные правила для частых и «простых» отношений
 - Обучение на основе размеченных данных



Извлечение отношений

- Ответы

- Wikipedia, DBpedia, FreeBase, etc.

- отец («Егор Гайдар», «Мария Гайдар»)
 - отец («Тимур Гайдар», «Егор Гайдар»)
 - отец («Аркадий Гайдар», «Тимур Гайдар»)

- Вопрос

Какие книги написал прадед Марии Гайдар?

прадед(?х, «Мария Гайдар»)

автор(?х, ?у)

по мотивам [Dan Jurafsky]

Пример

painkillers that don't thin blood



Search

painkillers that don't **thin blood**

painkillers that don't **cause constipation**

painkillers that don't **cause nausea**

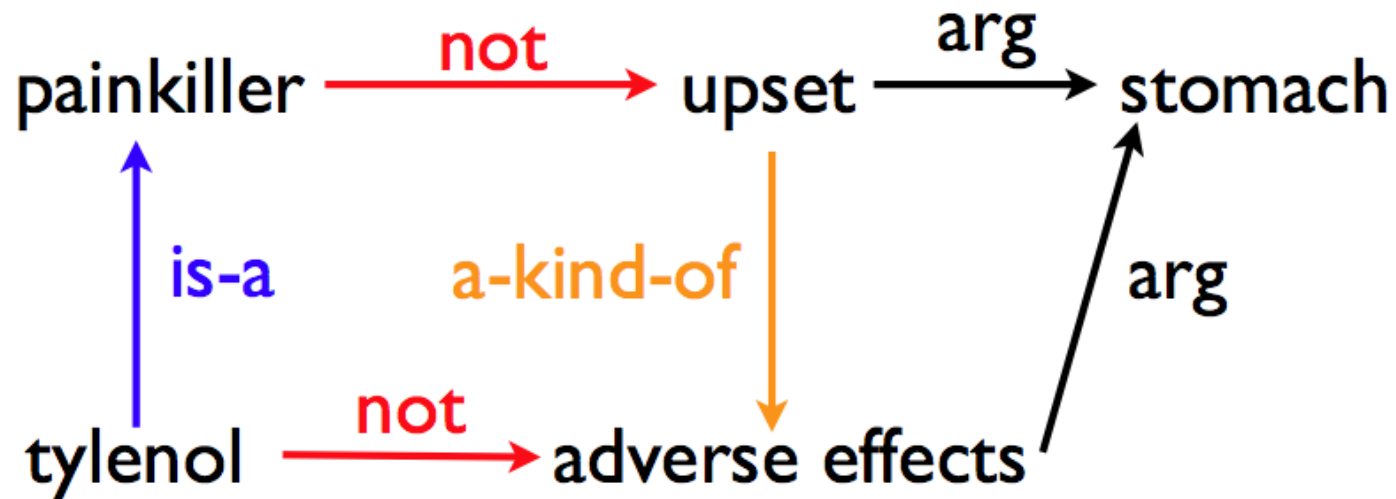
painkillers that don't **upset stomach** frequent query

painkillers that don't

About 89,300 results (0.14 seconds)

[Advanced search](#)

Пример



СОЦИАЛЬНЫЕ ВОПРОСНО- ОТВЕТНЫЕ СЕРВИСЫ (CQA)

Community Question Answering (CQA)



CQA – 2

List of question-and-answer websites

From Wikipedia, the free encyclopedia

The following is a list of websites that follow a question-and-answer format. The list contains only websites for which an article exists, dedicated either wholly or at least partly to the websites.

Website	Founded	Description/focus	Languages available	Copyrighting of user contributions	Registration required?
3form	1998	Many topics	English		
Answerbag	2003				
Answers.com	2000s				Yes
Ask.com	June 1996	Many topics			Yes
Ask.fm	2010	Many topics	38 languages		Yes
Ask Me Help Desk	2003				Yes
Ask MetaFilter	2003	Many topics	English	All posts are copyright to their original authors ^[1]	No to browse, yes to contribute
Askalo	August 3, 2010	Topics for answer by local residents	5 languages		No
Askbot	2009		English and a few others		Yes
Awo	2006	Legal and Medical	English		No
Baidu Knows		Many topics	Chinese		
BlikBook	2010	Academic	English		Yes
Blurtit	2006				No to browse, yes to contribute

~45, cm. http://en.wikipedia.org/wiki/List_of_question-and-answer_websites

Ответы@Mail.Ru

категории

профиль автора вопроса

вопрос



ОТВЕТЫ

профиль автора ответа

Категории



- [Вопросы друзей](#)
- [Авто, Мото](#)
- [Бизнес, Финансы](#)
- [Города и Страны](#)
- [Гороскопы, Магия, Гадания](#)
- [Досуг, Развлечения](#)
- [Еда, Кулинария](#)
- [Животные, Растения](#)
- [Знакомства, Любовь, Отношения](#)
- [Искусство и Культура](#)
- [Компьютеры, Интернет](#)
- [Красота и Здоровье](#)
- [Наука, Техника, Языки](#)
- [Образование](#)
 - [Домашние задания](#)
 - [ВУЗы, Колледжи](#)
 - [Детские сады](#)
 - [Школы](#)
 - [Дополнительное образование](#)
 - [Образование за рубежом](#)
 - [Прочее образование](#)
- [Общество, Политика, СМИ](#)
- [Путешествия, Туризм](#)
- [Работа, Карьера](#)
- [Семья, Дом, Дети](#)
- [Спорт](#)
- [Стиль, Moda, Звезды](#)
- [Темы для взрослых](#)
- [Товары и Услуги](#)
- [Философия, Непознанное](#)
- [Фотография, Видеосъемка](#)
- [Юридическая консультация](#)
- [Юмор](#)
- [О проектах Mail.Ru](#)




Главная » Образование » ВУЗы, Колледжи » Выбирается лучший ответ

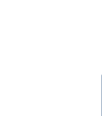
 [@Катя Фирсова](#) **Гуру** (3243)
На голосовании: 7 месяцев
 Подарить стикер! **NEW**



Что изучает и в чем особенности фундаментальной и прикладной лингвистики?

Оценка: 5 Рейтинг: 5 (1 голос)

 Ускорить выбор лучшего ответа  Нарушение!

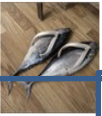
 Поделиться +1  Класс +1  Подписаться



 [@Владимир Савчинский](#) **Ученик** (205) 7 месяцев назад (ссылка)

 Нарушение!  Подарить стикер! **NEW**

Надо выпить чая))

Оценка: 0 Рейтинг: 0 [3 комментария](#)

 [@Некто в тапках](#) **Оракул** (68541) 7 месяцев назад (ссылка)

 Нарушение!  Подарить стикер! **NEW**

Кстати такая дифференциация встречается во многих науках. Например есть прикладная математика распространенные разделы, которые не так часто носят прикладной характер. Фундаментальная лингвистика нацелена на описание и изучения законов языка. Прикладная же лингвистика как инструмент в иных предметных областях. Например есть компьютерная лингвистика построение систем, распознавание речи, извлечение данных.

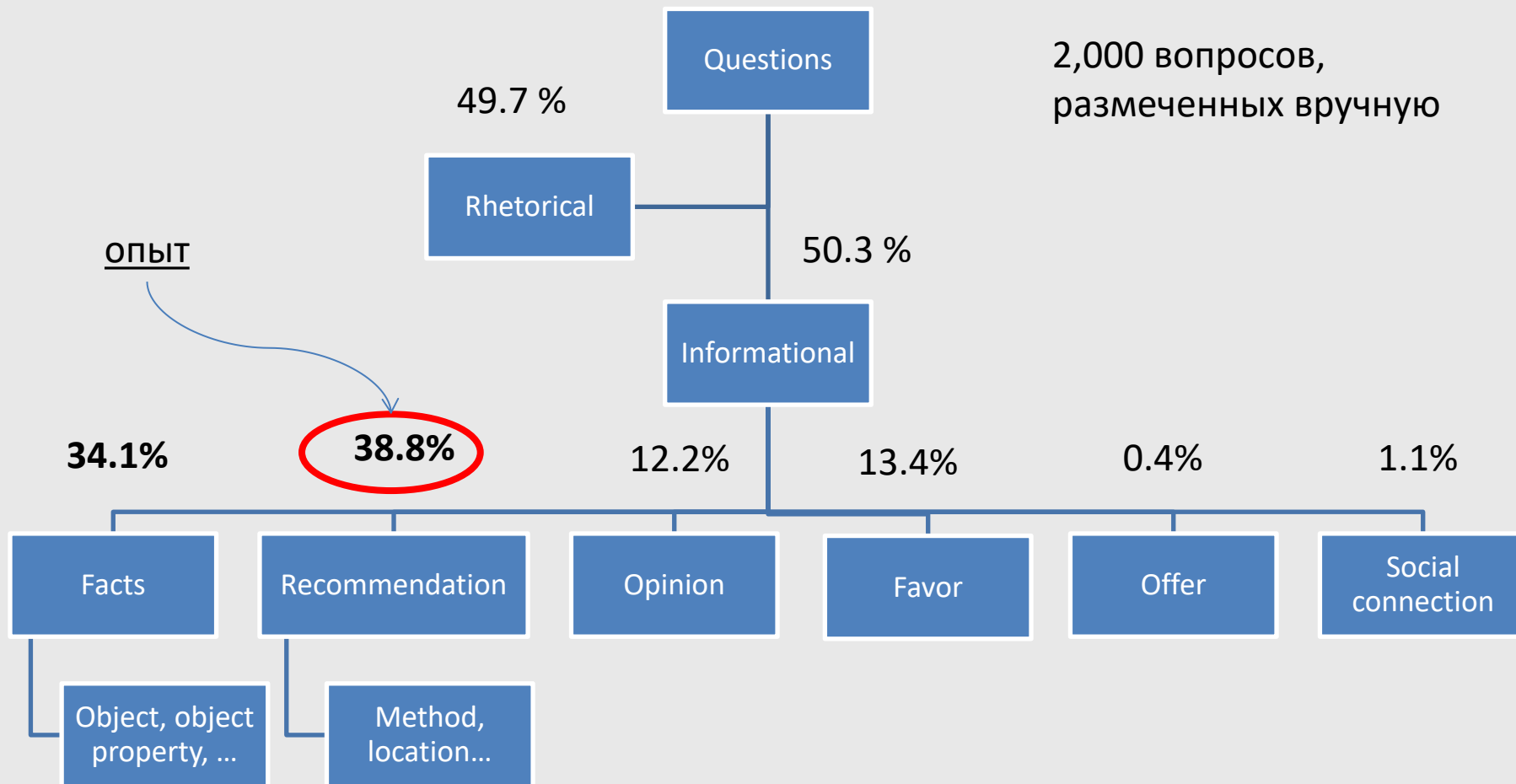
Источник: Не проф. лингвист. Просто мнение.

Оценка: 5 Рейтинг: 5 (1 голос)

Цифры

- Yahoo!Answers:
 - 1 миллиард ответов (Май 2010)
- Otvety@Mail.Ru
 - 80 млн вопросов, 400 млн ответов (август 2012)
- Stackoverflow
 - 13 млн вопросов, 20 млн ответов
 - 92% с ответами, медиана времени первого ответа – 11 минут

Типы вопросов Ответы@Mail.Ru



<http://kansas.ru/cqa/data/>

QA как поиск по архивам CQA

- Модели поиска
 - модели перевода Вопрос → Ответ
- Определение качественного контента
- Определение риторических/
субъективных/ироничных/
коммуникативных вопросов
- Рейтинг/специализация пользователей

Поиск по вопросам «как»

- 140,000 пар «вопрос-ответ» Yahoo!Answers
- лингвистические признаки++
- машинное обучение ранжированию (LETOR)

Table 6

Contribution of natural language structures in each feature group. Scores are MRR changes of the Perceptron on the development set over the baseline model (FG1 with W), for $N = 15$. The best scores for each feature group (i.e., column in the table), are marked in **bold**.

	FG1	FG2	FG3
W	-	+4.18	-6.80
N	-13.97	+2.49	-13.63
N_g	-18.65	+3.63	-15.57
D	-15.15	+1.48	-15.39
D_g	-19.31	+3.41	-18.18
R	-27.61	+0.33	-27.82
R_g	-28.29	+3.46	-26.74
W + N	+1.46	+5.20	-4.36
W + N + N_g	+1.51	+5.33	-4.31
W + N + N_g + D	+1.56	+5.78	-4.31
W + N + N_g + D + D_g	+1.56	+5.85	-4.21
W + N + N_g + D + D_g + R	+1.58	+6.12	-4.28
W + N + N_g + D + D_g + R + R_g	+1.65	+6.29	-4.28

W – слова,
N – n-граммы,
D – зависимости,
R – семантические роли.

FG1 – близость,
FG2 – перевод,
FG3 – плотность и частота.

ОЦЕНКА

TREC

- 1999-2007
- Фактографические вопросы
 - факты, списки, определения
 - коллекция новостей
 - тест: несколько сотен вопросов
 - системы возвращают ранжированный список фрагментов (50/250 символов)/единственный ответ и «уверенность»
- поиск в коллекции новостей (~1М документов, 3Гб)
- оценка для ранжированных списков:
Mean Reciprocal Rank (MRR)

<http://trec.nist.gov/data/qamain.html>

Вопросы TREC QA

- Who is the author of the book, "The Iron Lady: A Biography of Margaret Thatcher"?
- What was the monetary value of the Nobel Peace Prize in 1989?
- What does the Peugeot company manufacture?
- How much did Mercury spend on advertising in 1993?
- What is the name of the managing director of Apricot Computer?
- Why did David Koresh ask the FBI for a word processor?
- What debts did Qintex group leave?
- What is the name of the rare neurological disease with symptoms such as: involuntary movements (tics), swearing, and incoherent vocalizations (grunts, shouts, etc.)?

РОМИП

- Коллекция ВУ.WEB
- 10К запросов-вопросов из лога поисковой машины
 - gta san andreas как сделать машину призрак?
 - монгольские полевки как ухаживать?
 - берут ли с экземой в армию?
 - перелёт до екатеринбурга от москвы сколько по времени?
 - через сколько дней появляются корни у отростка традесканции?
 - всем ли девушкам важны деньги?
 - как заполучить парня своей мечты?
 - где пройдет финал кубка уефа 2009?
- До 5 ответов системы: docID, краткий ответ, фрагмент (до 300 символов)
- мало точных ответов в коллекции
<http://romip.ru/ru/2010/tracks/qa.html>

TREC LiveQA (с 2015)

Вопросы сервиса Yahoo!Answers: очень разные, сложные, субъективные

В реальном режиме времени: системы должны отвечать в течение 1 минуты

Нет ограничения на источник ответа

~1,000 вопросов из популярных категорий

<https://sites.google.com/site/trecliveqa2016/>

[Agichtein et al.]

CNN/Daily Mail

- ~200K документов
- ~1M «запросов»

Context

The BBC producer allegedly struck by Jeremy Clarkson will not press charges against the “Top Gear” host, his lawyer said Friday. Clarkson, who hosted one of the most-watched television shows in the world, was dropped by the BBC Wednesday after an internal investigation by the British broadcaster found he had subjected producer Oisin Tymon “to an unprovoked physical and verbal attack.” ...

Query

Producer X will not press charges against Jeremy Clarkson, his lawyer says.

Answer

Oisin Tymon

SQuAD

The Stanford Question Answering Dataset

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a new reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage. With 100,000+ question-answer pairs on 500+ articles, SQuAD is significantly larger than previous reading comprehension datasets.

[Explore SQuAD and model predictions](#)

[Read the paper \(Rajpurkar et al. '16\)](#)

Getting Started

We've built a few resources to help you get started with the dataset.

Download a copy of the dataset (distributed under the [CC BY-SA 4.0](#) license):

[Training Set v1.1 \(30 MB\)](#)

[Dev Set v1.1 \(5 MB\)](#)

Leaderboard

Since the release of our dataset, the community has made rapid progress! Here are the ExactMatch (EM) and F1 scores of the best models evaluated on the test set of v1.1. Will your model outperform humans on the QA task?

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> <i>(Rajpurkar et al. '16)</i>	82.304	91.221
1 <small>Jan 22, 2018</small>	Hybrid AoA Reader (ensemble) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	82.482	89.281
1 <small>Mar 06, 2018</small>	QANet (ensemble) <i>Google Brain & CMU</i>	82.744	89.045
1 <small>Feb 19, 2018</small>	Reinforced Mnemonic Reader + A2D (ensemble model) <i>Microsoft Research Asia & NUDT</i>	82.849	88.764
2 <small>Jan 05, 2018</small>	SLQA+ (ensemble) <i>Alibaba iDST NLP</i>	82.440	88.607
3 <small>Feb 02, 2018</small>	Reinforced Mnemonic Reader (ensemble model) <i>NUDT and Fudan University</i>	82.283	88.533

~500 страниц → `23К абзацев → ~100К вопросов + ответов-фрагментов