

# Автоматическое реферирование

## Automatic Summarization

Павел Браславский

# ЗАДАЧА

# Задача

Сформировать сжатое представление (текстового) документа (связанных документов), передающее основное содержание

+ с учетом контекста/ особенностей пользователя

+ связное и удобочитаемое

# Методы

- Одна из самых ранних задач ОЕЯ
- Правила/знания vs. статистика

# Актуальность

- Большие объемы информации
- Быстрое принятие решений
- Мобильные устройства

NB: алгоритмы + интерфейс (!)

[FEATURE](#) | WSJ. MAGAZINE 2013 TECHNOLOGY INNOVATOR

# How Teen Nick D'Aloisio Has Changed the Way We Read

When a Hong Kong billionaire emailed a London tech startup to inquire about investing, he didn't realize its employee consisted of a single kid working in his bedroom. Meet the 18-year-old WSJ. Magazine Technology Innovator who became an overnight millionaire by inventing an app that may revolutionize how we read on the go



CME Group  
energy com  
manage the  
and plan fo

[LEARN MO](#)

*How the wo*



## Миссия для поиска жизни на Марсе успешно стартовала с Байконура



Аргументы и Факты 15:40

### Миссия для поиска жизни на Марсе успешно стартовала с Байконура

С космодрома «Байконур» успешно стартовала миссия для поиска жизни на Марсе «ЭкзоМарс-2016» (ExoMars-2016), передает РИА Новости. Сейчас разгонный блок «Бриз-М» отделился от ракеты-носителя.



Lenta.ru 14:34

### Стартовала ракета «Протон-М» с модулями проекта ExoMars

Ракета «Протон-М» успешно стартовала с модулями проекта ExoMars. Трансляция запуска проходит в прямом эфире на YouTube-канале Роскосмоса. Первая в истории совместная российско-европейская миссия ExoMars-2016 предполагает отправку к Марсу орбитального и демонстрационного спускаемого модулей.



TASS 12:56

### Французский ученый: проект «ЭкзоМарс» - важный вклад в космическую науку Европы и России статья

Цель - поиск метана в атмосфере Марса Ученый отметил, что важной частью программы "ЭкзоМарс" является поиск в его атмосфере следов метана марсианским научным орбитальным аппаратом Trace Gas Orbiter.

Свернуть все сообщения



Екатеринбург Он-лайн 16:46

### С Байконура стартовала миссия для поиска жизни на Марсе

Ракета с межпланетной станцией успешно стартовала с Байконура. Сегодня днём ракета "Протон-М" успешно стартовала с космодрома Байконур и вывела на околоземную орбиту космические аппараты российско-европейской миссии ExoMars-2016.



Московский Комсомолец 15:15

### Ракета-носитель «Протон-М» успешно стартовала к Марсу с

# Нобелевскую премию по экономике дали за теорию контрактов

Нобелевский комитет 14:49, 10 октября 2016 ✓ Надежный источник

Нобелевская премия по экономике — 2016 досталась Оливеру Харту и Бенгту Хольмстрему за их вклад в развитие теории контрактов. Эта теория отвечает на вопросы о том, почему мы заключаем разные контракты на выполнение разных типов работ и почему эти соглашения устроены именно так, а не иначе.



Share on Facebook

Share on your own Timeline

Pavel Braslavski Say something about this...

Say son

meduza | НОВОСТИ ✓ Надежный источник

## Нобелевскую премию по экономике дали за теорию контрактов

Нобелевскую премию по экономике дали за теорию контрактов — Meduza

Нобелевская премия по экономике — 2016 досталась Оливеру Харту и Бенгту Хольмстрему за их вклад в развитие теории контрактов. Эта теория отвечает...



# Не только новости

- Научные статьи
- Художественные произведения
- Отзывы о товарах
- Веб-страницы (сниппеты)
- ...



ELSEVIER

Information Processing and Management 40 (2004) 919–938

**INFORMATION  
PROCESSING  
&  
MANAGEMENT**

[www.elsevier.com/locate/infoproman](http://www.elsevier.com/locate/infoproman)

## Centroid-based summarization of multiple documents

Dragomir R. Radev<sup>a,\*</sup>, Hongyan Jing<sup>b</sup>, Małgorzata Styś<sup>b</sup>, Daniel Tam<sup>a</sup>

<sup>a</sup> *University of Michigan, Ann Arbor, MI 48109, USA*

<sup>b</sup> *IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA*

Received 5 January 2003; accepted 24 October 2003

Available online 13 December 2003

### Abstract

We present a multi-document summarizer, MEAD, which generates summaries using cluster centroids produced by a topic detection and tracking system. We describe two new techniques, a centroid-based summarizer, and an evaluation scheme based on sentence utility and subsumption. We have applied this evaluation to both single and multiple document summaries. Finally, we describe two user studies that test our models of multi-document summarization.

© 2003 Elsevier Ltd. All rights reserved.

*Keywords:* Multi-document summarization; Centroid-based summarization; Cluster-based relative utility; Cross-sentence informational subsumption

ГРНТИ 31.21

ISSN 0203-6088



*Российская Академия Наук*

ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ  
(ВИНИТИ)

# РЕФЕРАТИВНЫЙ ЖУРНАЛ

**19. ХИМИЯ**

19Ж. ОРГАНИЧЕСКАЯ ХИМИЯ

ВЫПУСК СВОДНОГО ТОМА

\*



1

МОСКВА

2012

ГРНТИ 25.73

ISSN 0203-6081



*Российская Академия Наук*

ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ  
(ВИНИТИ)

# РЕФЕРАТИВНЫЙ ЖУРНАЛ

**13. МАТЕМАТИКА**

СВОДНЫЙ ТОМ

\*



10

МОСКВА

2005

**06.02-13A.67** Обобщение теоремы Гаусса о группах классов вычетов mod  $m$ . Generalization of Gauss' theorem on the groups of residue classes mod  $m$ . *Yu Chen-yan, Zhou Xia. Zhejiang daxue xuebao. Lixue ban = J. Zhejiang Univ. Sci. Ed.* 2005. 32, № 3, с. 251–252. Кит.; рез. англ.

Упражнение по теории сравнений.

О. Фоменко

**05.10-13A.223** О многообразиях полугрупп, на свободных объектах которых почти все вполне инвариантные конгруэнции слабо перестановочны. *Верников В. М. Алгебра и логика.* 2004. 43, № 6, с. 635–649, 758. Рус.

Говорят, что многообразие полугрупп имеет индекс, не превосходящий 2, если все нильполугруппы из этого многообразия являются полугруппами с нулевым умножением. Дается описание многообразий полугрупп индекса, не превосходящего 2, на свободных объектах которых любые две вполне инвариантные конгруэнции, содержащиеся в наименьшей полурешеточной конгруэнции, слабо перестановочны, а также многообразий полугрупп индекса, не превосходящего 2, все подмногообразия которых обладают указанным свойством.

**06.02-13A.114** Строго транзитивные геометрические пространства: приложения к гипергруппам и теории полугрупп. Strongly transitive geometric spaces: Applications to hypergroups and semigroups theory. *Freni Domenico. Commun. Algebra.* 2004. 32, № 3, с. 969–988. Англ.

Гипергруппоидом называется множество  $H$  с отображением (умножением)  $\circ : H \times H \rightarrow \mathcal{P}^*(H)$ , где  $\mathcal{P}^*(H)$  — множество всех непустых подмножеств множества  $H$ . Полугипергруппа — полугипергруппоид, в котором  $(x \circ y) \circ z = x \circ (y \circ z)$  для всех  $x, y, z \in H$ . Гипергруппа — полугипергруппа, в которой  $x \circ H = H \circ x = H$  для всех  $x \in H$ . Геометрическое пространство — это пара  $(S, \mathcal{B})$ , где  $S$  — непустое множество (его элементы называются точками), а  $\mathcal{B}$  — непустое семейство подмножеств (называемых блоками). Геометрическое пространство называется строго транзитивным, если для любых  $A, B \in \mathcal{B}$  условие  $A \cap B \neq \emptyset$  влечёт, что  $\forall x \in B \exists C \in \mathcal{B}$  такое, что  $A \cup \{x\} \in C$ .

Для всякой гипергруппы  $H$  — можно построить геометрическое пространство, взяв в качестве точек элементы множества  $H$ , а блоков — всевозможные произведения  $z_1 \circ z_2 \circ \dots \circ z_n$ . Это пространство является строго транзитивным (Gutan, 1997). Если  $H$  — полугипергруппа, то можно в качестве блоков взять подмножества вида  $B(z_1, z_2, \dots, z_n) = \cup\{z_{\sigma(1)} \circ \dots \circ z_{\sigma(n)} \mid \sigma \in S_n\}$  (здесь  $S_n$  — симметрическая группа). В работе доказано, что построенное геометрическое пространство строго транзитивно. Кроме того, получены необходимые и достаточные условия, при



About 409,000 results (0.52 seconds)

## Борщ - рецепты - Povarenok.ru

[www.povarenok.ru/recipes/category/44/](http://www.povarenok.ru/recipes/category/44/) ▾ [Translate this page](#)

У каждой хозяйки, конечно, есть свой **рецепт** приготовления **борща**. Мне хотелось бы поделиться с вами одним из своих **рецептов**. Может быть вам ...

[Борщ](#) · [Самый красный борщ](#) · [Украинский борщ от тети Нюси](#) · [Мой борщец](#)

## Борщи, рецепты с фото на RussianFood.com: 652 рецепта как ...

[russianfood.com](http://russianfood.com) ▸ [Рецепты](#) ▸ [По типам блюд](#) ▾ [Translate this page](#)

Борщ – это первое блюдо, особая разновидность супа. Если вас интересуют вопросы: борщи, борщ **рецепт**, **борщ** рецепт с фото, как приготовить борщ, ...



## Как приготовить борщ по классическому рецепту - Лайфхакер

<https://lifehacker.ru/2014/12/27/classic-borshcht/> ▾ [Translate this page](#)

Дес 27, 2014 - В каждой семье свой **рецепт борща**. Кто-то добавляет грибы, кто-то фасоль, кто-то копчености. Мы расскажем, как приготовить борщ по ...



[lifehacker.ru](http://lifehacker.ru)

## Как приготовить борщ по классическому рецепту

*Настя Радужная*

В каждой семье свой рецепт борща. Кто-то добавляет грибы, кто-то фасоль, кто-то копчености. Мы научим вас готовить классический борщ. Незамужним девушкам читать обязательно!

Борщ — заправочный овощной суп. Его любят не только в России и Украине, но и в Польше (баршч), Литве (баршчя), Румынии и Молдове (борш).

В Киевской Руси борщ готовили из съедобных листьев борщевика (отсюда название). Позднее стали варить со свёклой (отсюда цвет). С XIX века добавляют картофель.

Средняя стоимость борща в России — 220 рублей. Дешевле всего борщ с доставкой в Казани: 37 рублей за порцию. Самый дорогой борщ в столице: тарелка может обойтись в 700 и более рублей.

### Что взять

*Для бульона:*

- вода — 1,5–2 л.;
- свинина или говядина на кости — 400 г.

## 2000 книг в кратком изложении

### Школьная программа

/ 11 класс

10 класс

9 класс

8

7

6

5



#### Недоросль Фонвизин

Деревня помещиков Простаковых. Госпожа Простакова в гневе: крепостной портной Тришка, считает она, сшил слишком узкий кафтан её любимому сыну, шестнадцатилетнему недорослю Митрофанушке...



#### Бедная Лиза Карамзин



#### Гранатовый браслет Куприн



#### Гроза Островский



#### На дне Горький



#### Старуха Изергиль Горький



#### Олеся Куприн

Регион: [Екатеринбург](#)

(495) 733-91-74

 279041902

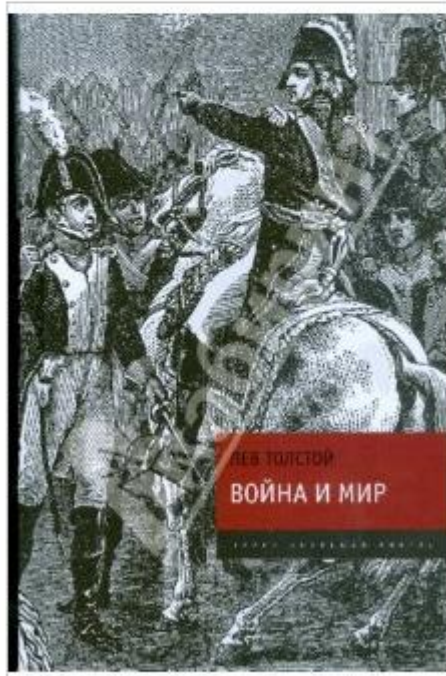


 [Корзина](#)

 [Доставка](#)

Жанр: [Книги](#) / [Художественная литература](#) / [Русская проза](#) / [Классика](#)

Война и мир: в 4 томах. Том 1-4



ID товара: 186469  
ISBN: 978-5-699-31874-2

Автор: [Толстой Лев](#)  
Издательство: [Эксмо](#)  
Серия: [Большая книга](#)  
Жанр: [Классика](#)  
Год выпуска: 2008

Страниц: 1296 ([Офсет](#))  
Тип обложки: 7бц - твердая, целлофанованная (или лакированная)  
Иллюстрации: [Без иллюстраций](#)

Масса: 1146 г  
Размеры: 218x147x56 мм

Наличие: **На складе (ограниченное количество)**

339 руб. [Купить](#)

Метки: [Добавить метку](#)


Регион: [Екатеринбург](#)

(495) 733-91-74

 279041902



 [Корзина](#)

 [Доставка](#)

Жанр: [Книги](#) / [Учебная, методическая литература и словари](#) / [Книги для школы](#) / [Литература](#) / [Краткое изложение произведений](#)

"Война и мир" Л. Н. Толстого. Краткое содержание. Особенности романа. Сочинения



ID товара: 72004  
ISBN: 5-8053-0442-2

Автор: [Шаповалова Оксана](#)  
Издательство: [Лицей](#)  
Серия: [По страницам литературной классики](#)  
Жанр: [Краткое изложение произведений](#)  
Год выпуска: 2005

Страниц: 304 ([Газетная](#))  
Тип обложки: обл - мягкий переплет (крепление скрепкой или клеем)  
Иллюстрации: [Черно-белые](#)

Масса: 276 г  
Размеры: 200x140x15 мм

Наличие: **На складе**

[Содержание](#)

92 руб. [Купить](#)

Метки: [Добавить метку](#)

## Рецензия «Афиши»



**Лев Данилкин**

839 рецензий · 779 оценок · 343 спасибо

★★★★★ 2

После оцетинившейся двенадцатью языками глоссологии «Сердце Пармы» «Географ» — легкая добыча: будто вместо раскаленных бронзовых скрижалей, смердящих паленой человечинной, подкинул вдруг бог прочесть классный журнал 10-го «В», заполненный шариковой ручкой. Еще в середине 90-х, до того как отлить цельнометаллическую эпопею про Пермское княжество в XV веке, историк искусств и краевед Алексей Иванов человеческим русским языком изложил сюжет про злоключения обычного учителя.

28-летний Служкин устраивается от безденежья в школу преподавать географию; в течение года он воюет с учениками, мается с женой, дочкой и любовницами, выпивает с приятелями и ведет девятиклассников в поход. Педагогико-авантюрно-эротический роман про географа написан умным и остроумным человеком — читать не бросите; но поскольку в нем нет ни туго скрученного сюжета, ни специальной лингвистической подоплеки, поскольку, наконец, это совсем-совсем реализм, диагностировать источники удовольствия от текста сложнее, чем в случае с «Пармой». В почти сплошь состоящем из анекдотов и диалогов «Географе» воспроизведен и реставрирован так называемый школьный дискурс — традиционные советские книги и фильмы про школу: от «Ералаша» и «Доживем до понедельника» до Драгунского, Носова, Виктора Голявкина и Анатолия Рыбакова. Такого рода литература и кино исчезли после развала империи, заменились сюжетами позффектней; в «Географе» бог весть с чего они вдруг снова проявились; оказалось, что и классный журнал может быть не менее интересным, чем вторая книга «Поэтики» Аристотеля.



# Не только текст

- Кино
- Музыка
- Видео (+ текст)
- Речь

MOVIECLIPS  
HOT NEW  
TRAILERS

THE GIRL  
ON THE  
TRAIN

OCTOBER 7, 2016

BUY TICKETS NOW!  
FANDANGO.COM

MOVIECLIPS

# TRAILERS

SEE OFFICIAL TRAILERS FIRST



## Movieclips Trailers

Subscribed 9,957,686

Home Videos Playlists Channels Discussion About



### Pirates of the Caribbean: Dead Men Tell No Tales Trailer - Teaser (2017) - Johnny Depp Movie

1,210,294 views 1 day ago

Starring: Johnny Depp, Kaya Scodelario, and Orlando Bloom

Pirates of the Caribbean: Dead Men Tell No Tales Trailer - Teaser (2017) - Johnny Depp Movie

Captain Jack Sparrow searches for the trident of Poseidon.

Subscribe to TRAILERS: <http://bit.ly/sxaw6h>  
Subscribe to COMING SOON: <http://bit.ly/H2vZUn> ...  
[Read more](#)

### Fandango MOVIECLIPS Channels

- Movieclips**
- Movieclips Coming Soon**
- Movieclips Indie Film Fest**
- Fandango All Access**
- Movieclips Extras**
- Movieclips Trailer Vault**
- Fandango Family**
- Movieclips Comedy**
- Movieclips Horror**
- Movieclips Action**

### HOT New Trailers & Movieclips Exclusives

Be the first to see the latest TRAILERS, clips, featurettes, & newest promo content for the most anticipated upcoming films.



Pirates of the Caribbean: Dead Men Tell No Tales Trailer - Teas...



Middle School: The Worst Years of My Life Movie CLIP - Rule 26...



Fantastic Beasts and Where to Find Them Official Trailer 2



Denial Featurette - The Making of Denial (2016) - Rachel Weisz

# ТИПЫ РЕФЕРАТОВ

# Типы рефератов

- Abstract  $\leftrightarrow$  Extract (*выжимка*)
- Индикативные  $\leftrightarrow$  информативные
- Один  $\leftrightarrow$  много документов
- Общие  $\leftrightarrow$  с учетом контекста/пользователя
- Критический

# Abstract/Extract

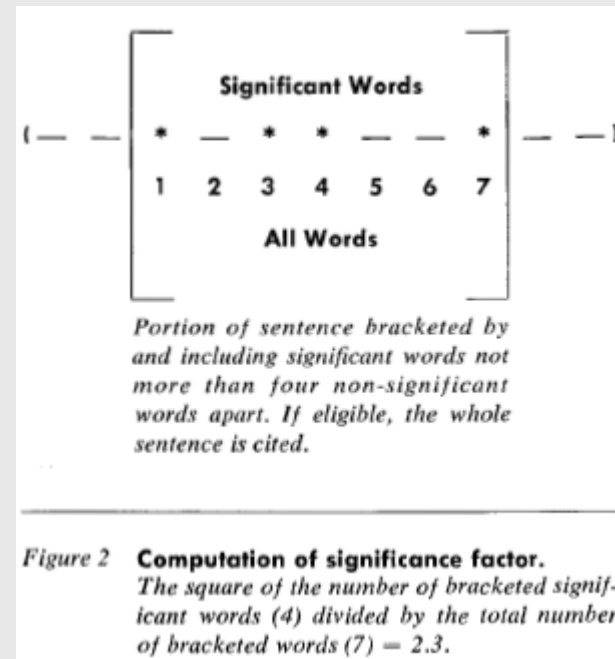
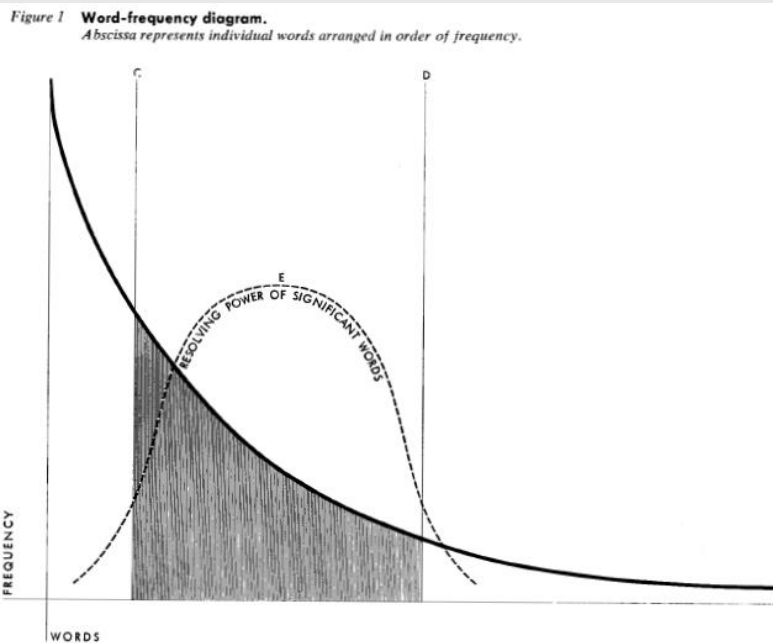
- Текст → «семантическое представление»  
→ генерация реферата  
(abstract generation)
- Выбор значимых фрагментов, обычно предложений  
(sentence extraction)

# Экстракты

- + простота
- - СВЯЗНОСТЬ
  - дисбаланс (ср. разные точки зрения в статье)

# МЕТОДЫ

# Начало [Luhn, 1958]



+ [Edmundson, 1969], [Paice, 1981]



# Слова

- Вес предложения на основе весов входящих в него слов.
- Вес слова – на основе частоты.
- Близко расположенные важные слова усиливают друг друга (*keyword burst*)
- Сигнальные фразы и слова: *можно сделать вывод..., необходимо подчеркнуть...*

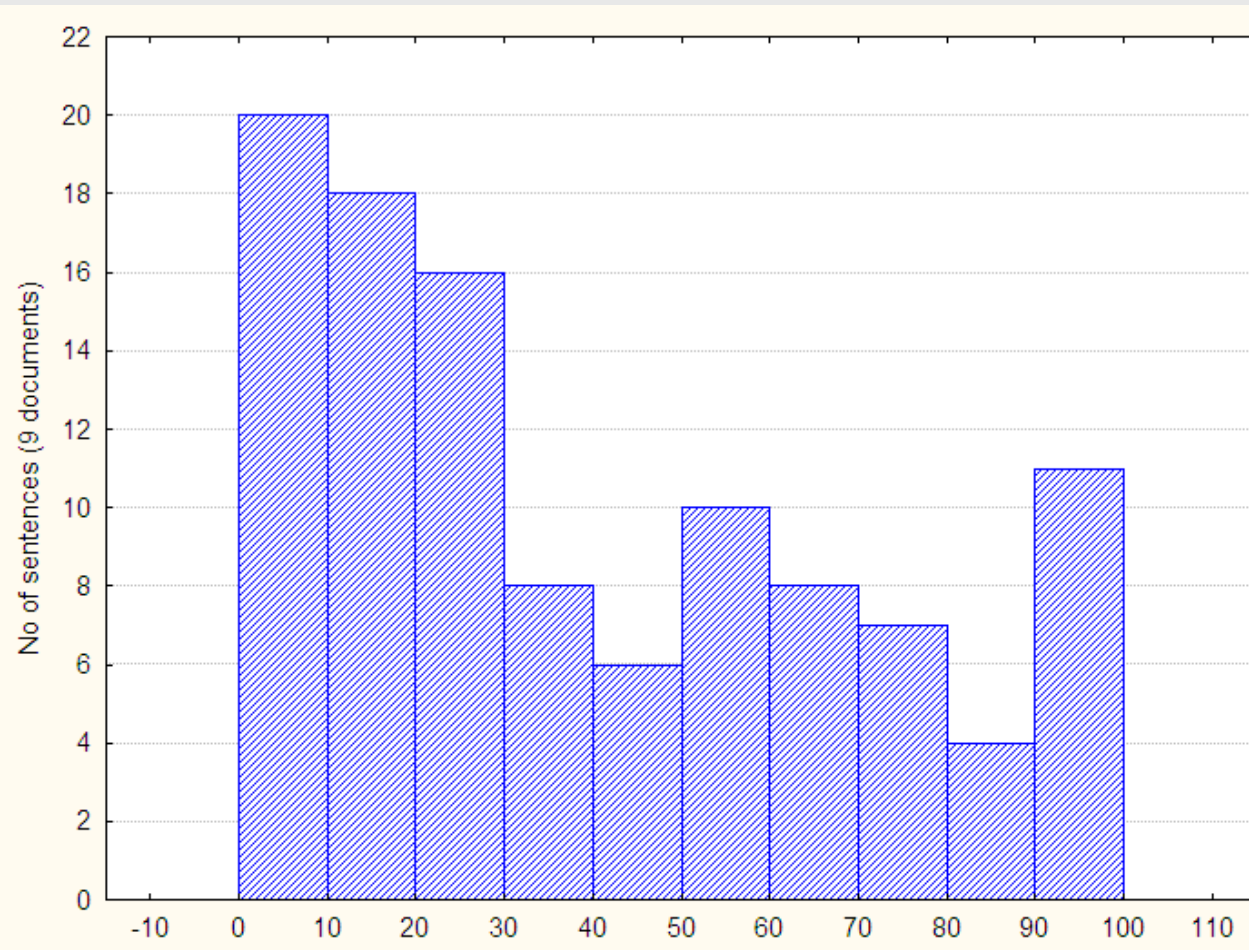
# Предложения

- Начало и конец документа/абзаца/раздела
- Похоже на заголовок
- Нормализация по длине предложения
- Вопросительные предложения скорее всего не подходят
- Предложения с анафорической отсылкой не подходят (или включаются в реферат вместе с предыдущим предложением)

*В отличие от классического токамака тороидальной формы, у NSTX-U, хотя и сохраняющего, строго говоря, такую же форму, крайне малая внутренняя полость. Это позволяет эффективно удерживать и разогревать плазму при менее сильных магнитных полях.*

<https://lenta.ru/articles/2016/10/10/fusion/>

# Положение предложений



Положение выделенных вручную «информативных» предложений (9 новостных текстов, каждый обрабатывался 3 ассессорами, длины текстов нормированы)

# Недостатки «ВЫЖИМОК»

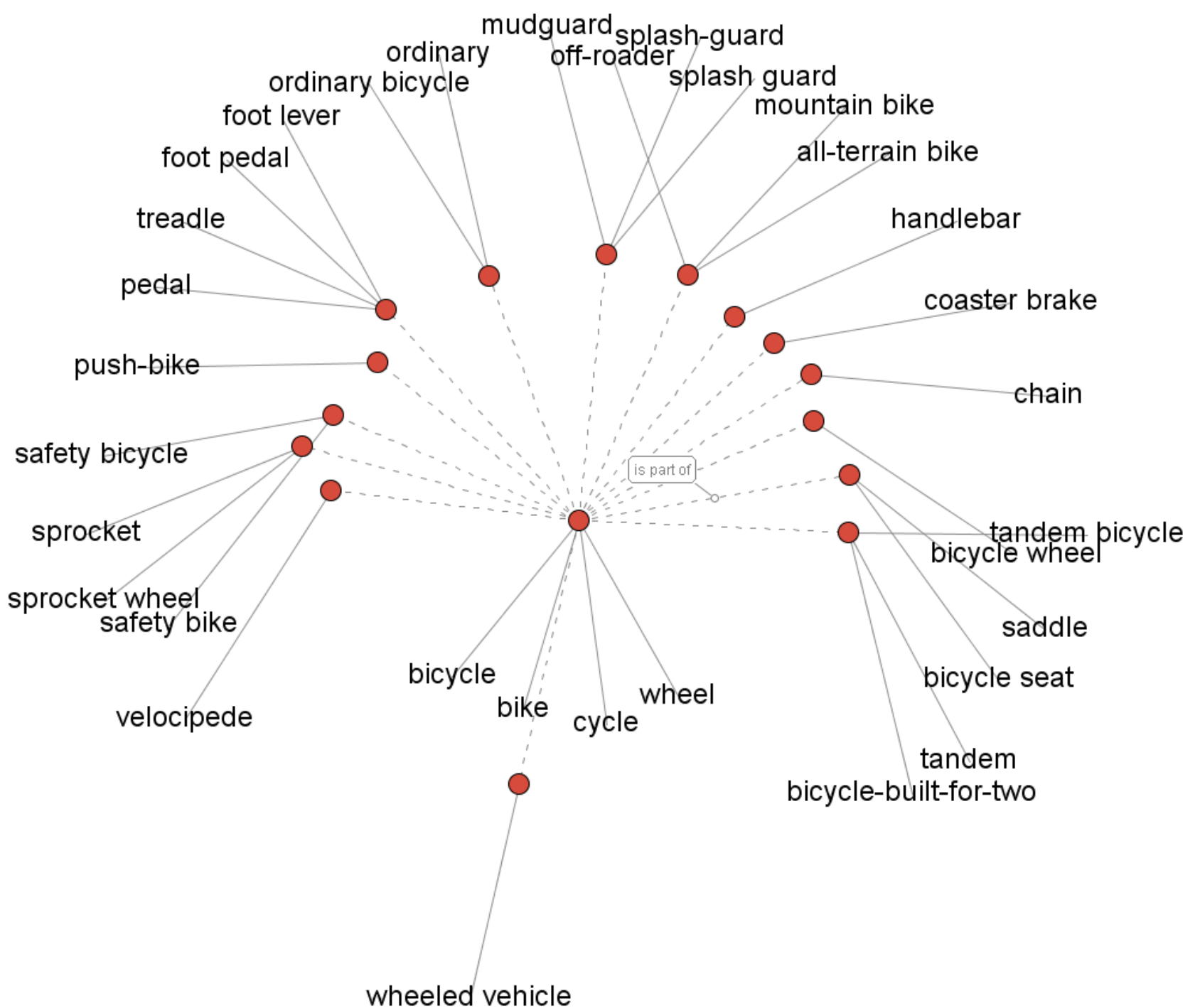
- Связность (анафорические отсылки)
- Баланс (например, разные мнения в статье)

# Методы на основе знаний

- Frump (deJong 1979, 1982)
  - Сценарии событий, которые заполняются на основе ключевых слов
- Summons  
(McKeown & Radev, 1996)

# Лексические цепочки

- Лексические цепочки (lexical chains) – последовательность семантически связанных слов:  
*{корова, овца, шерсть, шарф, шляпа}*
- Используются на нахождения связных (cohesive) структур в тексте, тематических областей и т.п.
  - Сегментация текста, **автоматическое реферирование**, вопросно-ответный поиск, разрешение лексической многозначности (word sense disambiguation)
- **Вычисляются на основе тезауруса**



# Лексические цепочки

**Червей** тутового шелкопряда кормили листьями шелковицы, опрысканными водными **растворами**, которые содержали 0,2 процента **углеродных одностенных нанотрубок**. Когда **гусеницы** окуклились и сформировали кокон, ученые собрали синтезированный насекомыми **шелк**. Такой способ является экологически чистым, в отличие от обработки **шелка графеном**, которая требует применения токсичных химических **растворителей**.

<https://lenta.ru/news/2016/10/10/silcyworms/>



## [Barzilay & Elhadad, 1997]

1. существительные + именные фразы (*sea level*)
2. повторы слова < синсеты слов непосредственно связаны < есть путь длины 1+
3. построение всех возможных цепочек, снятие неоднозначности (выбирается наиболее связный граф, разные веса для разных типов связей)
4. взвешивание цепочек → «сильные цепочки»
5. Извлечение предложений:
  - первое появление звена каждой цепочки
  - первое появление самого частотного звена \*
  - предложения из «насыщенного» сегмента

# TextRank [Mihalcea & Tarau, 2004]

- Близость предложений на основе слов (~коэффициента Жаккара) → взвешенный граф предложений
- Упорядочение предложений на основе значения центральности (PageRank)

# Машинное обучение [Куріес'95]

- Корпус статей с рефератами для обучения
- Признаки:
  - длина, положение предложения
  - присутствие «тематических» слов, слов из заглавных букв, «ключевых» фраз
- Сопоставление предложений из реферата и основного текста
- Байесовский классификатор

# A Trainable Document Summarizer

**Julian Kupiec, Jan Pedersen and Francine Chen**

Xerox Palo Alto Research Center

3333 Coyote Hill Road, Palo Alto, CA 94304

{kupiec,pedersen,fchen}@parc.xerox.com

## Abstract

- To summarize is to reduce in complexity, and hence in length, while retaining some of the essential qualities of the original.
- This paper focusses on document extracts, a particular kind of computed document summary.
- Document extracts consisting of roughly 20% of the original can be as informative as the full text of a document, which suggests that even shorter extracts may be useful indicative summaries.
- The trends in our results are in agreement with those of Edmundson who used a subjectively weighted combination of features as opposed to training the feature weights using a corpus.
- We have developed a trainable summarization program that is grounded in a sound statistical framework.

**Keywords:** summary sentence, original documents, summary pairs, training corpus, document extracts

author-supplied indicative abstracts. In this paper, it is hoped that other, more easily generated summaries will serve.

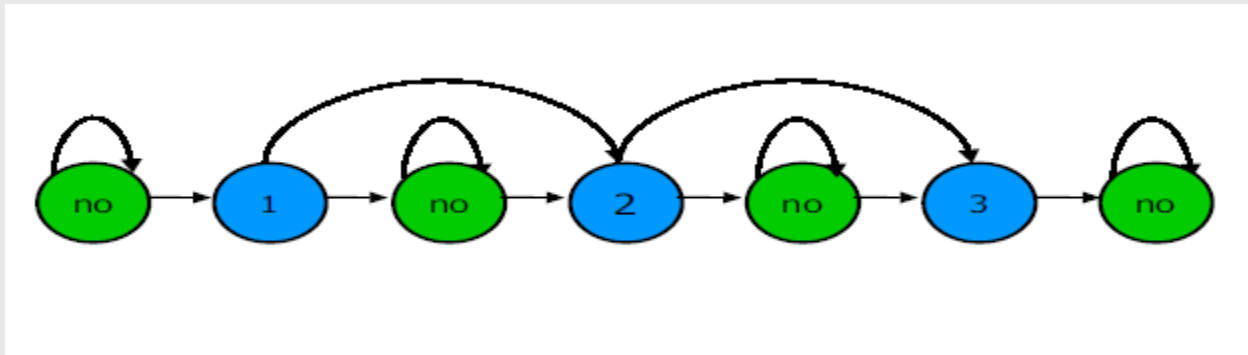
Numerous researchers have investigated document summarization (see [10] for an overview). Generating a coherent narrative summary from a document is considered too problematic since it requires a deep understanding of meaning, abstraction, and language. State-of-the-art edge intensive methods have been proposed [11, 5, 3, 13, 18]. For example, the Stanford sentence understanding system can generate a sentence summary. A simpler, more practical method is to extract a central sentence from the document. The central difficulties of natural language summarization are to be *summary by extraction* [7], i.e., to select a subset of the document that is indicative of the whole. The selection of sentences and presenting them in a coherent order of summaries are not guaranteed to be useful for rapid relevance assessment.

Document extracts consisting of roughly 20% of the original can be as informative as the full text of a document, which suggests that even shorter extracts may be useful indicative summaries. However, other studies [12, 2] suggest that such summaries are far from unique. Numerous heuristics have been proposed to

# Реферирование на основе НММ

[Conroy & O'Leary, 2001]

- $2s+1$  состояний
- признаки:
  - позиция (документ/абзац)
  - # терминов, статистика по коллекции



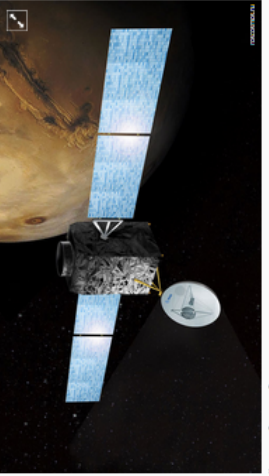
# Данные для обучения

- Машинное обучение
- «Дешево» получить данные
- Использовать разнообразный набор признаков

[Браславский, Густелев, 2007]

## Россия и Европа полетели на Марс

Россия и Европа вернулись к полету на Марс. Впервые в истории полетит аппарат «Экзомарс-2016».



Космический аппарат «Экзомарс»

На Марс отправился космический аппарат «Экзомарс» — совместный проект ESA и «Роскосмос», который исследует атмосферу Южного полюса. О том, как миссия поможет ответить на вопрос о жизни, рассказывает опенд наука «Светлы.РФ».

14 ноября на космодроме Байконур состоялся запуск ракеты-носителя «Союз-2.1В» с космическим аппаратом российской-европейской миссии ExoMars 2016 (сентябрь 2016).

«Экзомарс» — это совместный проект космических «Роскосмос» и Европейского космического агентства (ESA) по изучению Марса. Официальный договор о сотрудничестве был подписан 14 марта 2013 года в Париже в рамках «Евросоюза-Возрождения Полосы» и в штаб-квартире ESA в Брюсселе. До этого момента проект «Экзомарс» принадлежал только NASA, но европейцы на протяжении нескольких лет участвовали в разработке аппарата.

Проект состоит из двух этапов, первым из которых является запуск.

Поблизко южного автоматического станция, «Экзомарс-2016» состоит из двух блоков, орбитального и посадочного.

Этот орбитальный модуль TGO (Trace Gas Orbiter), который является автономной системой, который будет исследовать атмосферу Южного полюса и атмосферный состав, а также будет исследовать атмосферу Южного полюса. Он будет работать на высоте 400 км над поверхностью Марса. В состав аппарата входят также посадочный модуль и ровер.

Проект от ESA и России является самым сложным в истории космонавтики. Аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

После отделения модуля «Сциаларинде» орбитальный аппарат выйдет на эллиптическую орбиту. Он пролетит через верхнюю атмосферу Марса, а затем выровняет орбиту до круговой с высотой около 400 км.

Основной целью миссии является изучение атмосферы Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

Если эти данные не будут получены, то это будет означать, что атмосфера Марса является самой сложной в истории космонавтики. Аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

На Марсе тоже протекают биологические процессы. Однако газ может летать и химическое происхождение: он образуется в ходе различных геологических процессов.

Орбитальный аппарат сможет ответить на вопрос о происхождении жизни в атмосфере Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

Дополнительная миссия TGO — обеспечить связь между миссиями. Космический аппарат будет служить радиостанцией для аппарата «Экзомарс-2016», а также для полетевшего аппарата «Экзомарс-2018».

Главной задачей миссии «Сциаларинде» является изучение атмосферы Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

Научные задачи «Экзомарс» — изучение атмосферы Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

Второй этап миссии намечен на 2018 год: в ходе него на Марс будет отправлен космический аппарат «Экзомарс-2018». Он будет состоять из трех блоков: переносный модуль, адаптер с системой отделения и спускаемый аппарат, который разрабатывает НПО имени Лавочкина.

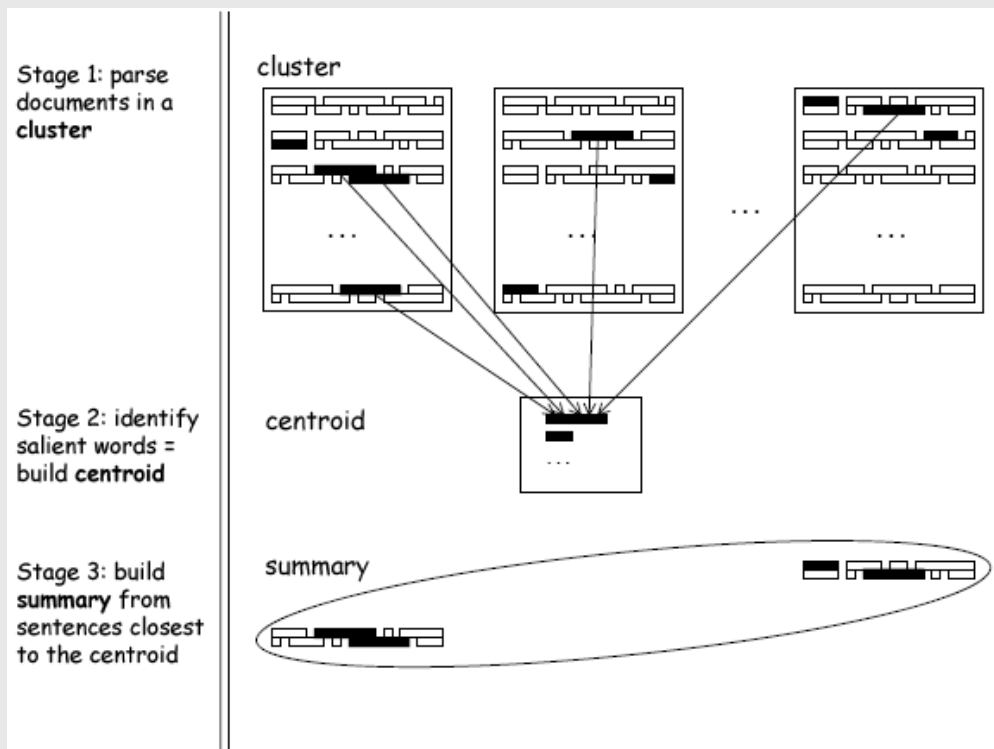
Аппарат «Экзомарс-2018» будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса. Это первый раз в истории, когда аппарат будет исследовать атмосферу Южного полюса Марса.

# ML + внешние данные

- 1300+ документов, 3-4 выделенных предложения
- Парный метод ранжирования предложений
- Дополнительные признаки
  - на основе лога поиска по новостям
  - сущности Wikipedia

[Svore, Vanderwende, Burges, 2007]

# Multi-document summarization



MEAD

вес предложения:

+ слова из центроида

+ положение

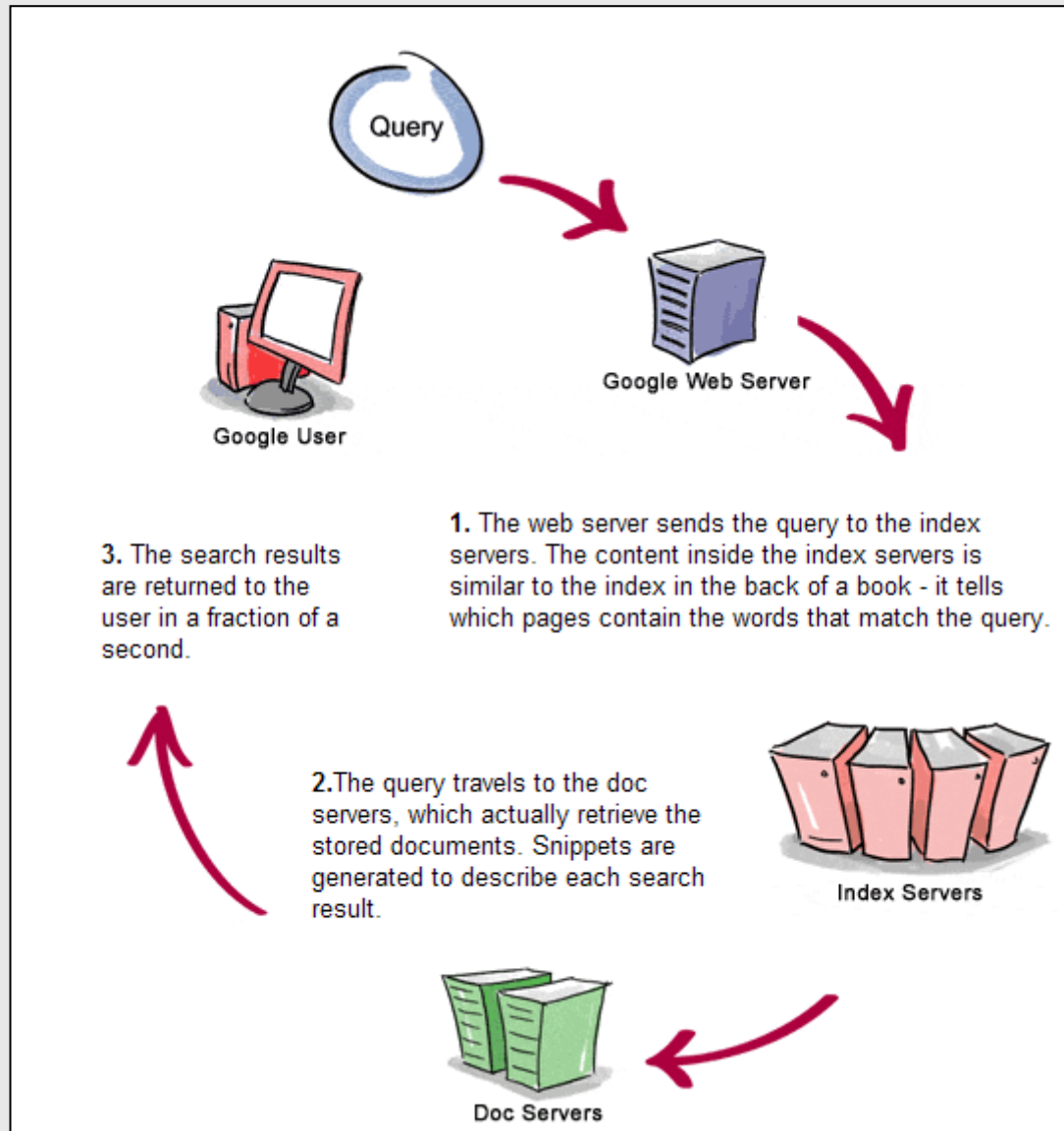
+ похожесть на первое предложение

-похожесть на другие

[Radev et. al, 2004]



# ВЕБ-СНИППЕТЫ



# Особенности веб-сниппетов

- Документы
  - ссылки, внешние описания
  - структура (title, meta, body, anchor)
  - очень разные (не все являются связным текстом)
- Сниппеты для разных запросов должны быть
  - индикативными
  - информативными
- UI
  - структура (заголовок + текст)
  - жесткое ограничение на длину
  - ранжированный список
- Спам/SEO
- Производительность



All Videos News Images Books More ▾ Search tools

About 409,000 results (0.52 seconds)

## Борщ - рецепты - Povarenok.ru

[www.povarenok.ru/recipes/category/44/](http://www.povarenok.ru/recipes/category/44/) ▾ Translate this page

У каждой хозяйки, конечно, есть свой **рецепт** приготовления **борща**. Мне хотелось бы поделиться с вами одним из своих **рецептов**. Может быть вам ...

[Борщ](#) · [Самый красный борщ](#) · [Украинский борщ от тети Нюси](#) · [Мой борщец](#)

## Борщи, рецепты с фото на RussianFood.com: 652 рецепта как ...

[russianfood.com](http://russianfood.com) ▸ [Рецепты](#) ▸ [По типам блюд](#) ▾ Translate this page

Борщ – это первое блюдо, особая разновидность супа. Если вас интересуют вопросы: борщи, борщ **рецепт**, борщ **рецепт** с фото, как приготовить борщ, ...



## Как приготовить борщ по классическому рецепту - Лайфхакер

<https://lifehacker.ru/2014/12/27/classic-borshcht/> ▾ Translate this page

Дес 27, 2014 - В каждой семье свой **рецепт борща**. Кто-то добавляет грибы, кто-то фасоль, кто-то копчености. Мы расскажем, как приготовить борщ по ...



```
<meta name="description" content="В каждой семье свой рецепт борща. Кто-то добавляет грибы, кто-то фасоль, кто-то копчености. Мы расскажем, как приготовить борщ по классическому рецепту." />
```

[lifehacker.ru](http://lifehacker.ru)

## Как приготовить борщ по классическому рецепту

*Настя Радужная*

В каждой семье свой рецепт борща. Кто-то добавляет грибы, кто-то фасоль, кто-то копчености. Мы научим вас готовить классический борщ. Незамужним девушкам читать обязательно!

Борщ — заправочный овощной суп. Его любят не только в России и Украине, но и в Польше (баршч), Литве (баршчя), Румынии и Молдове (борш).

В Киевской Руси борщ готовили из съедобных листьев борщевика (отсюда название). Позднее стали варить со свёклой (отсюда цвет). С XIX века добавляют картофель.

Средняя стоимость борща в России — 220 рублей. Дешевле всего борщ с доставкой в Казани: 37 рублей за порцию. Самый дорогой борщ в столице: тарелка может обойтись в 700 и более рублей.

### Что взять

*Для бульона:*

- вода — 1,5–2 л.;
- свинина или говядина на кости — 400 г.

анджелина джоли — 6 млн ответов

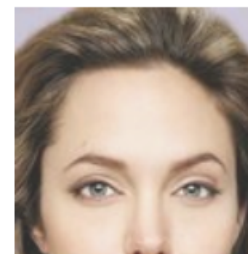


Найти

## Джоли, Анджелина — Википедия

[ru.wikipedia.org](https://ru.wikipedia.org) > [Джоли, Анджелина](#) ▾

**Анджелина́ Джоли́ Питт** (англ. **Angelina Jolie Pitt**, при рождении **Войт**; род. 4 июня 1975, Лос-Анджелес) — американская актриса, режиссёр и сценарист, фотомодель, посол доброй воли ООН.



# Джоли, Анджелина

Материал из Википедии — свободной энциклопедии

*В Википедии есть статьи о других людях с фамилией [Джоли](#).*

**Анджелина́ Джоли́ Питт** (англ. *Angelina Jolie Pitt*<sup>[3]</sup>, при рождении **Войт**; род. 4 июня 1975, Лос-Анджелес) — американская актриса, режиссёр и сценарист, [посол доброй воли ООН](#).

```
63 <meta property="og:title" content="&#x2605;&#x2605;&#x2605; Гостиница Варшава, Москва, Россия" />
64 <meta property="og:description" content="Трёхзвездочная гостиница «Варшава» находится в центре Москвы, в 3
65 <meta property="og:image" content="http://q-ec.bstatic.com/images/hotel/max300/584/5849012.jpg" />
66 <meta property="og:image:width" content="300" />
67 <meta property="og:image:height" content="200" />
```

# InCommonSense

- Выбран наиболее продуктивный шаблон ссылающегося абзаца
- Набор таких абзацев, ссылающихся на страницу, – кандидаты в сниппеты
- Проведена оценка полученных описаний в качестве сниппетов
- построен классификатор хороший/плохой сниппет
- признаки (15): длина, личные местоимения, аббревиатуры, оценочные прилагательные, слова, указывающие на содержание (about, information), повторы слов, оформление (заглавные буквы, пунктуация, глаголы)...

(The|This) [Anchor](#) --- ---  
-----  
-----

[Amitay&Paris, 2000]

Поиск

«**Bing**» — поисковая система от Microsoft

[bing.com](https://bing.com) ▼

Поиск по сайтам информации, изображений и видео.

Картинки

Видео

### Видео

**Bing** помогает принимать обоснованные решения...

Карты

### Сервисы

Discover new features in **Bing**

Маркет

Ещё

### Webmaster Tools

Войдите в службу "Средства веб-мастера – **Bing**" или...

### Учетная запись

Местоположение Укажите свой город и государство или почтовый код.

### Bing Homepage Gallery

Explore breathtaking images from the U.S. **Bing** homepage over the past...

### Bing Translator

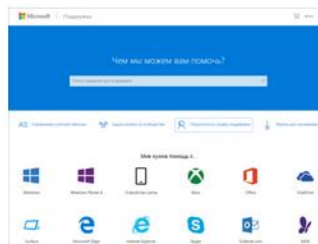
Бесплатная веб-служба перевода, предоставляющая мгновенный...

Нашлось **20 млн результатов**

77 тыс. показов в месяц

[Дать объявление](#)

Найдено сайтов: 83



### «Техническая поддержка Microsoft»

[support.microsoft.com](https://support.microsoft.com)

Подборка статей и информация об обновлениях продуктов **Microsoft**.

Рубрика: Windows

ТИЦ: 7400



### «Bing» — поисковая система от Microsoft

[bing.com](https://bing.com)

Поиск по сайтам информации, изображений и видео.

Рубрика: Поисковые системы

ТИЦ: 5200

# Производительность

- Сниметы генерируются «на лету» по оригинальному тексту документа
- Комбинация
  - схем сжатия (VB encoding),
  - кэширования,
  - предобработки документов

**может существенно повысить производительность.**

[Turpin A. et al., 2007]



# **ПОСТОБРАБОТКА РЕФЕРАТА**

# Борьба с избыточностью

- Maximal marginal relevance, MMR  
[Carbonell&Goldstein, 1998]
- кластеризация предложений  
[Nomoto&Matsumoto, 2001]
- Cross-sentence informational subsumption, CSIS  
[Radev et al., 2004]

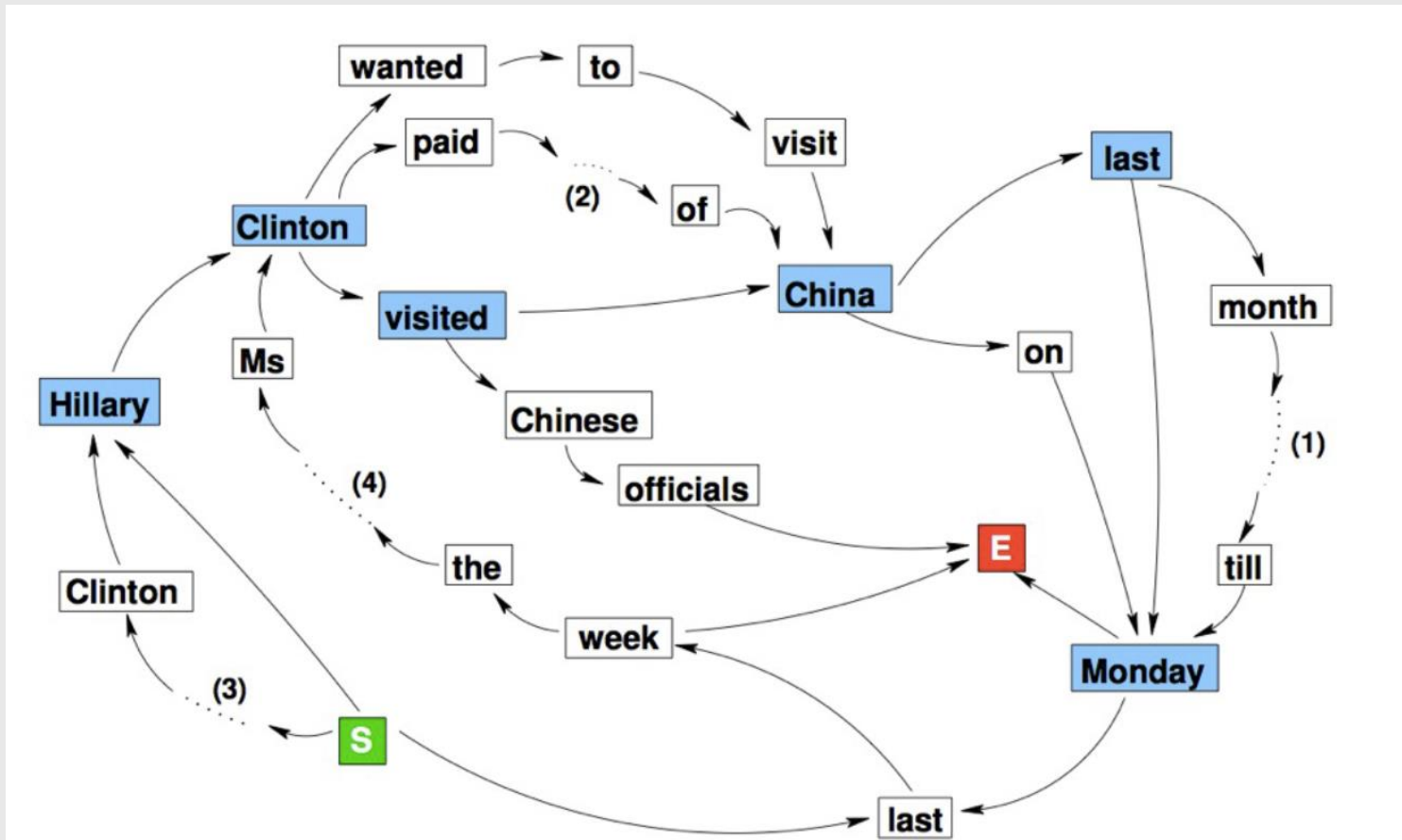
Но: избыточность может способствовать связности и удобочитаемости

# Постобработка

- Возможность сократить реферат
- Сделать его более читаемым
- «Середина» между выделением предложений и генерацией реферата

Методы:

- На основе правил (анализ синтаксического разбора)
- Обучение на основе корпуса переформулировок
- Подход на основе «канала с помехами»
- Синтез на основе графового представления (многодокументное реферирование)



[Filippova 2011]

# Промежуточные итоги

- простые универсальные методы позволяют получить приемлемый реферат «в среднем»
- повышение качества
  - более глубокий анализ документа
  - больше информации о документе (тема, структура...) → специализация
  - анализ и обработка «сырого» реферата
  - внешний семантические ресурсы

# **ОЦЕНКА РЕФЕРАТОВ**

# Оценка

- Сравнение с идеалом
- В рамках задачи (task-oriented)
- + связность, удобочитаемость...

# Согласие людей – 1

Проблема: низкое  
согласие людей

Пример:

16 текстов, 3 ассессора/текст,  
по 5 предложений/текст →  
182 уникальных  
предложения.

Согласие:

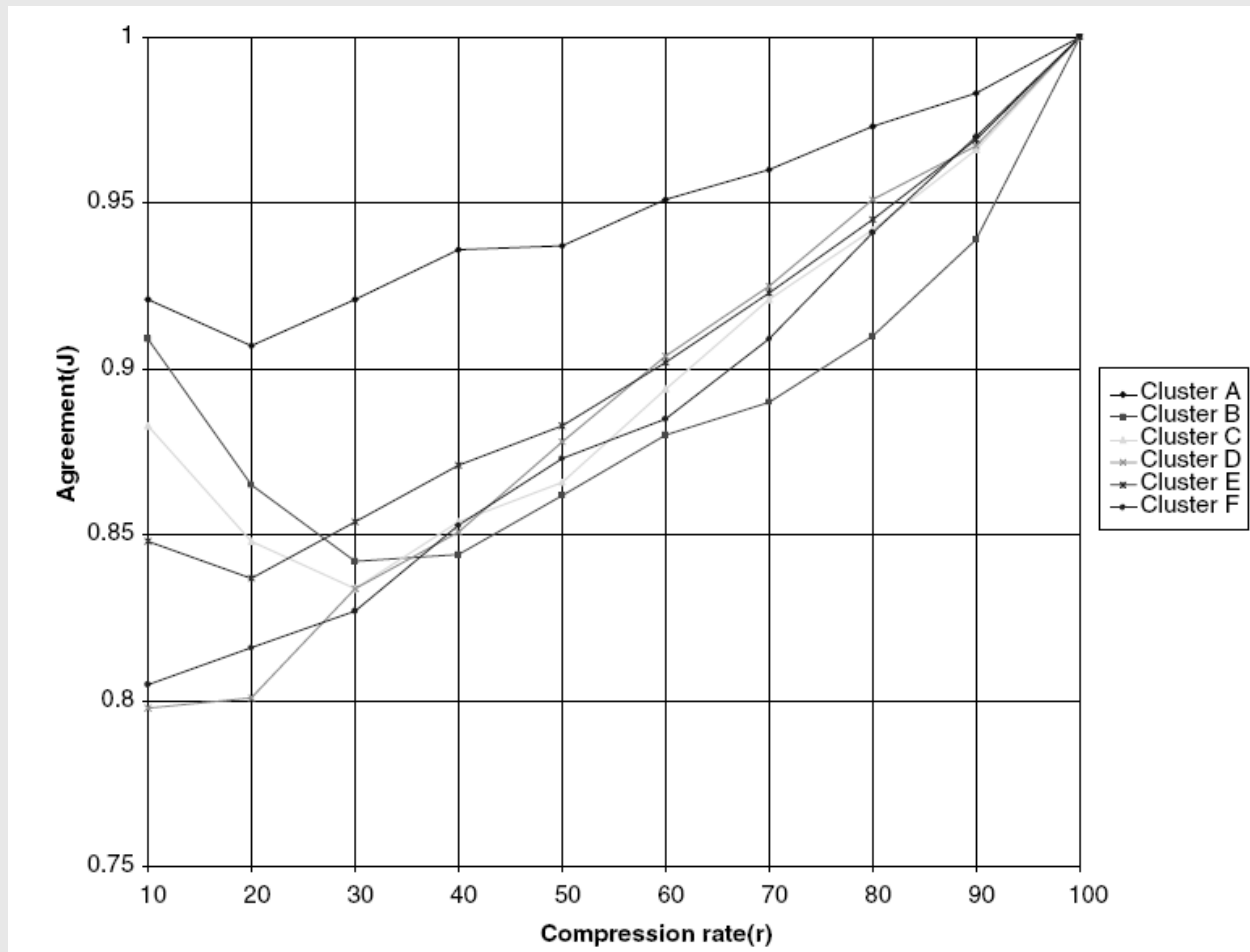
>1: 0,24;

3: 0,05

	# of sent	2 or 3	3	ave	ave 2
T_Imps	12	3	0	0,25	0,00
T_Used	9	5	0	0,56	0,00
T_Weakness	14	1	0	0,07	0,00
T_Awkward	12	2	0	0,17	0,00
T_Use	12	3	0	0,25	0,00
NY_Frantic	12	2	1	0,17	0,08
H_Helm	12	2	0	0,17	0,00
H_Plotting	11	4	1	0,36	0,09
E_Schools	13	0	1	0,00	0,08
H_Tamiflu	12	2	1	0,17	0,08
H_Straight	12	3	0	0,25	0,00
E_Summer	9	4	2	0,44	0,22
E_Decline	11	4	0	0,36	0,00
E_Little	9	3	2	0,33	0,22
H_Treatment	9	4	2	0,44	0,22

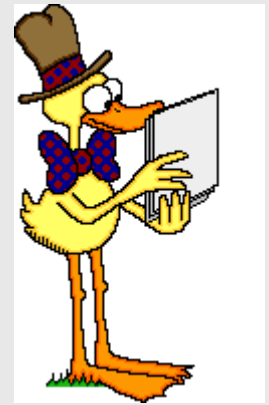


# Согласие людей – 2



[Radev et al., 2004]

# Document Understanding Conference (DUC)



- Мероприятие по оценке автоматического реферирования
- Задачи
  - Реферирование новостного документа (2001, 2002) – 100 слов
  - Реферирование кластера новостей (рефераты разной длины, в т.ч. «заголовка»)
  - Рефераты с учетом контекста/типа события/вопроса/профиля пользователя...
- С 2008 – Text Analysis Conference (TAC)

# Оценка @ DUC

- Сравнение автоматического реферата (peer) с рефератами-образцами (models):
  - Разбивка на предложения, сопоставление с последующим ответом на вопрос «Насколько информация, содержащаяся в образце, передана в оцениваемом реферате?» (0/20/40/60/80/100%)
- Недостатки:
  - Сильно зависит от автора образца
  - Низкое согласие ассессоров (в т.ч. согласие ассессора с самим собой)
- Лингвистическая оценка (1..5)
  - Оформление (grammaticality)
  - Лаконичность (non-redundancy)
  - Корректные отсылки (referential clarity)
  - Фокус (focus)
  - Структура и связность (structure and coherence)

# ROUGE (Lin, 2004)

- Следует BLEU (bilingual evaluation understudy) – автоматическая метрика качества машинного перевода
- Recall-Oriented Understudy for Gisting Evaluation
- сравнение реферата с набором «идеальных»
- подсчет общих элементов
- варианты: +стемминг –стоп-слова
- отбор метрик, по которым ранжирование систем коррелирует с ранжированием на основе экспертных оценок (разные для разных задач)

# ROUGE-N

- ROUGE-N: N-gram co-occurrence statistics is a recall oriented metric

$$\text{ROUGE - N} = \frac{\sum_{S \in \{\text{Refs}\}} \sum_{n\text{-gram} \in S} \text{count}_{\text{match}}(n\text{-gram})}{\sum_{S \in \{\text{Refs}\}} \sum_{n\text{-gram} \in S} \text{count}(n\text{-gram})}$$

# Другие метрики ROUGE

- ROUGE-L: longest common subsequence
- ROUGE-W: weighted longest common subsequence, favours consecutive matches
- ROUGE-S: Skip-bigram recall metric
- ROUGE-SU adds unigrams to ROUGE-S

# Пример

Модель1: красный синий желтый зеленый

Модель2: красный белый синий желтый

$\text{count}(1) = 8, \text{count}(2) = 6$

Реферат1: красный синий голубой

$R-1 = 4/8 \quad R-2 = 1/6$

Реферат2: красный белый красный синий

$R-1 = 7/8 \quad R-2 = 2/6$

# Люди vs. ROUGE

Method	DUC 2001 100 WORDS SINGLE DOC						DUC 2002 100 WORDS SINGLE DOC					
	1 REF			3 REFS			1 REF			2 REFS		
	CASE	STEM	STOP	CASE	STEM	STOP	CASE	STEM	STOP	CASE	STEM	STOP
R-1	0.76	0.76	0.84	0.80	0.78	0.84	0.98	0.98	<b>0.99</b>	0.98	0.98	<b>0.99</b>
R-2	0.84	0.84	0.83	<b>0.87</b>	0.87	0.86	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
R-3	0.82	0.83	0.80	0.86	0.86	0.85	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
R-4	0.81	0.81	0.77	0.84	0.84	0.83	<b>0.99</b>	<b>0.99</b>	0.98	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
R-5	0.79	0.79	0.75	0.83	0.83	0.81	<b>0.99</b>	<b>0.99</b>	0.98	<b>0.99</b>	<b>0.99</b>	0.98
R-6	0.76	0.77	0.71	0.81	0.81	0.79	0.98	<b>0.99</b>	0.97	<b>0.99</b>	<b>0.99</b>	0.98

Method	DUC 2003 10 WORDS SINGLE DOC							
	1 REF		4 REFS		1 REF		4 REFS	
	CASE	STEM	STOP	CASE	STEM	STOP	CASE	STEM
R-1	0.96	0.95	0.95	0.95	0.90	0.90		
R-2	0.75	0.76	0.75	0.75	0.76	0.77		
R-3	0.71	0.70	0.70	0.68	0.73	0.70		
R-4	0.64	0.65	0.62	0.63	0.69	0.66		
R-5	0.62	0.64	0.60	0.63	0.63	0.60		
R-6	0.57	0.62	0.55	0.61	0.46	0.54		
R-7	0.56	0.56	0.58	0.60	0.46	0.44		
R-8	0.55	0.53	0.54	0.55	0.00	0.24		
R-9	0.51	0.47	0.51	0.49	0.00	0.14		
R-L	<b>0.97</b>	<b>0.96</b>	<b>0.97</b>	<b>0.96</b>	0.97	0.96		
R-S*	0.89	0.87	0.88	0.85	0.95	0.92		
R-S4	0.88	0.89	0.88	0.88	0.95	0.96		
R-S9	0.92	0.92	0.92	0.91	0.97	0.95		
R-SU*	0.93	0.90	0.91	0.89	0.96	0.94		
R-SU4	<b>0.97</b>	<b>0.96</b>	0.96	0.95	<b>0.98</b>	<b>0.97</b>		
R-SU9	<b>0.97</b>	0.95	0.96	0.94	0.97	0.95		
R-W-1.2	0.96	<b>0.96</b>	0.96	<b>0.96</b>	0.96	0.96		

35	0.79	0.80	0.76	0.98	0.98	0.97	<b>0.99</b>	<b>0.99</b>	0.97
31	0.78	0.78	0.72	0.98	0.98	0.96	<b>0.99</b>	<b>0.99</b>	0.97
59	0.76	0.76	0.69	0.97	0.97	0.95	0.98	0.98	0.96
33	0.86	0.86	0.86	<b>0.99</b>	<b>0.99</b>	0.99	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
30	0.78	0.77	0.82	0.98	0.98	0.98	0.98	0.97	0.98
34	<b>0.87</b>	<b>0.88</b>	<b>0.87</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
34	<b>0.87</b>	<b>0.88</b>	<b>0.87</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
31	0.78	0.77	0.83	0.98	0.98	0.98	0.98	0.98	0.98
35	<b>0.87</b>	0.87	<b>0.87</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
35	<b>0.87</b>	0.87	<b>0.87</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
35	<b>0.87</b>	0.87	<b>0.87</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>



# **ОЦЕНКА ВЕБ-СНИППЕТОВ**

# Нет документов – есть сниппеты

- Первичная оценка пользователями качества поиска – не на основании документов, а на основании сниппетов
- Сниппеты плохо отражают содержание страницы – потерянный результат
- вводящие в заблуждение сниппеты – лишние клики

# [Tombros&Sanderson, 1998]

- простой метод реферирования (title, location, term occurrence, query terms)
- рефераты по запросу vs. начало документа
- сравнение:
  - оценки релевантности на рефератах/полных текстах (данные TREC);
  - скорость;
  - обращения к полным документам

# [White et al., 2003]

- простой метод выделения предложений: WebDocSum (title, location, query-biased, text-formatting)
- единый интерфейс
- 4 системы: Google +/- WebDocSum, AltaVista +/- WebDocSum
- task-oriented evaluation:
  - поиск факта
  - поиск нескольких объектов
  - поиск для принятия решения
  - поиск информации по теме
- Анкеты, мысли вслух, логгирование поведения

# РОМИП

- релевантность на сниппетах/полных документах (2005-2007)
  - Ассессор оценивает релевантность документа запросу по аннотации этого документа. Сам документ ассессор не увидит. При оценке ассессор будет видеть название (первые 100 символов) и аннотацию (300 символов).
- 2008
  - информативность (плохо/хорошо/отлично)  
(«Достаточно ли информации для принятия решения о релевантности?»)
  - читабельность (плохо/хорошо/отлично)
  - оценка релевантности документа

# Опыт оценки сниппетов

- Полевые эксперименты (user studies)
- Массовая оценка
- Автоматические метрики
- Поведение пользователей (клики)

Кошмар на улице Вязов

- 10 мая 2010 17:11 Поиск: как бы с собой первый раз смотрели «Кошмар на улице Вязов» и как это было? Вспомнил, как в детстве смотрели этот фильм — мы втроём подкакивали и шугались в комнате страшной морозной ночью. Лучший кошмар современности — фильм «Кошмар на улице Вязов» [фото](#) [видео](#)
- 2 Кошмар на улице Вязов  
Отзывы зрителей о фильме «Кошмар на улице Вязов: Но мне он понравился! С удовольствием посмотрел фильм и рекомендую своим коллегам! Обратите пожалуйста внимание на постер к новому фильму «Кошмар: Преподание во имя Фредди» [фото](#) [видео](#)
- 3 Кошмар на улице Вязов  
«Кошмар на улице Вязов» отзыв о фильме. Слабый фильм! Раньше Крюгер убивал психически больных людей (то есть он убивал невинных людей) — разочарован. Фредди это самоубийца, потому что он убивает невинных людей [фото](#) [видео](#)
- 4 Кошмар на улице Вязов  
А как новый «Кошмар на улице Вязов» понравится? Многие могут не согласиться, что ремейк лучше оригинала, но все же все бездарные сценарии новый «Кошмар на улице Вязов» перелопачивает [фото](#) [видео](#)
- 5 Фильм «Кошмар на улице Вязов» отзывы и рецензии  
Снова на экраны Кошмар на улице Вязов это классика мирового кино в жанре ужасов. Можно посетовать на некоторые проблемы с логикой — но главная проблема все-таки в нехватке смелости, а главной неприятностью начинают гадать, чтобы не разочароваться [фото](#) [видео](#)
- 6 Кошмар на улице Вязов (Nightmare on Elm Street, A, 2010)  
Елена Тихонова Через двадцать шесть лет после классического «Кошмара на улице Вязов» на экраны выходит его ремейк. Действительно, смотреть можно. Однако, при этом не вспоминать, что все это самое происхоило великого друга [фото](#) [видео](#)
- 7 скачать фильм Кошмар на улице вязов бесплатно  
В 80 «Кошмар на улице вязов» был по настоящему фильмом ужасов... Но теперь просто скачать и часто перематывать за это 4 из 10! - Конечно всего и сразу - а получаешь ничего и постепенно [фото](#) [видео](#)
- 8 Первые отзывы на КОШМАР НА УЛИЦЕ ВЯЗОВ 2010 \* Зона Ужасов  
Но Хейли проделал великолепную работу. ... Я наслаждался Кошмаром на улице Вязов от начала и до конца Фредди показан жестоким и мрачным. Возможно, он мрачнее, чем мы когда-либо видели его раньше. [фото](#) [видео](#)
- 9 Скачать фильм Кошмар на улице Вязов 7. Последний кошмар  
Отзывов: 0 Название: Кошмар на улице Вязов 7. Последний кошмар. Оригинальное название: Wes Craven's New Nightmare. ... И кошмары на улице Вязов начинаются снова, преследующие теперь сына нашей героини. В этом фильме Фредди существует на самом деле [фото](#) [видео](#)
- 10 YouTube: Кошмар на улице Вязов 2010 (трейлер)  
Трейлер к фильму Кошмар на улице Вязов 2010 год. Просмотров: 7435KinoEframov [фото](#) [видео](#)

Страницы — [сп](#) [предыдущая](#) [следующая](#) [сп](#) —

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) ...

Отсортировано [по релевантности](#) [по дате](#)

«кошмар на улице вязов отзывы» Запросов за месяц кошмар — 355 764, улице — 2 369 261, вязов — 630 105.  
в других поисковых системах: [Google](#) [Bing](#) [Yahoo!](#) [Rambler](#) [Mail.ru](#) [Яндекс.Каталог](#)

Яндекс Директ

«Кошмар на улице Вязов»  
Все о фильме «Кошмар на улице Вязов»: рецензии, новости, фотографии, афиша [www.kinoafisha.msk.ru](#) Москва

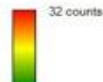
[Все объявления 2](#)

[Дать объявление](#)

«кошмар на улице ...» в картинках



[Все картинки](#)



32 counts

# Группировка сниппетов

sql import from

query

1 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

MS **sql import** includes a wizard, which allows you to set all the **import** options for different. Tags: **Import Data, Import Data To Sql Server, Import From Excel, Import Table. EMS Data Import ...**

7 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

EMS data **import** for **sql** server is a powerful tool to **import** your data quickly from MS Access. Tags: **Import Data, Import Data To Sql Server, Import From Excel, Import Table. EMS Data Import ...**

2 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

Tags: **Import Data, Import Data To Sql Server, Import From Excel, Import Table. EMS Data Import ...**EMS data **import** for **sql** server is a powerful tool to **import** your data quickly from MS Access, MS...

3 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

... **Import** 2007 for **SQL** Server 3.0 download by EMS Database Management Solutions, Inc. EMS Data **Import** 2005 for DB2 3.0 download by EMS Database Management Solutions, Inc.

4 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

...Excel In **Sql** Server. ...Run **sql** command to see **importing** results. 3 Support Unicode.

5 **Import Data To Sql Server Downloads - EMS MS SQL Data...**

EMS data **import** for MySQL is a powerful tool to **import** your data quickly from MS Excel, MS Access.....see **importing** results. 3 Support Unicode.

group 1

group 2



# Автоматические метрики

- Подсветка
- Пустые сниппеты
- Оформление
- Уникальные слова

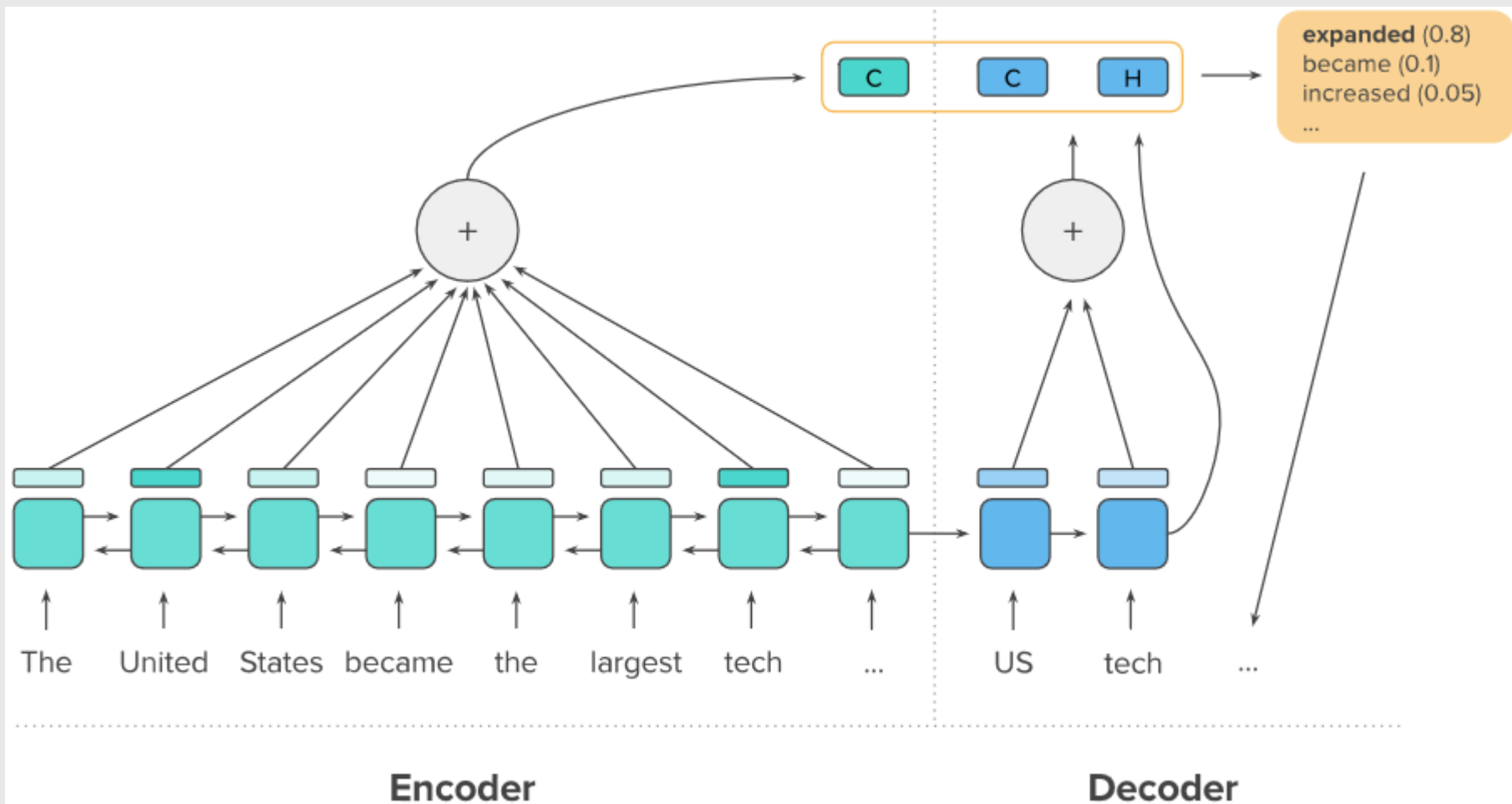
# Поведение пользователей

- Время первого клика
- Инверсии кликов
- Время между кликами
- Отсутствие кликов

# **СОВРЕМЕННОЕ СОСТОЯНИЕ**

# DNN для реферирования

- двунаправленная рекуррентная нейронная сеть
- механизмы внимания
  - внимание на разных частях статьи → разнообразие/покрытие реферата
  - внимание на уже сгенерированных последовательностях -- против повторов
- целевая функция
  - обучение с подкреплением (reinforcement) на уровне всего реферата
  - обучение с учителем (supervised) на уровне слов
- данные
  - DeepMind Q&A Dataset
  - NY Times



Paulus, Xiong, and Socher, 2017.  
 A Deep Reinforced Model for Abstractive Summarization

## Detectives fear Russian spy was sprayed with poison in street

172-218 минут

- **Sergei Skripal and his daughter Yulia collapsed near shopping centre on Sunday**
- **Two policemen who helped them were among up to 10 others later treated**
- **His status as a former MI6 spy led to claims the incident is a Kremlin 'hit'**
- **Counter-terror police have now assumed control of the probe, Scotland Yard say**
- **Police have widened a Salisbury cordon as investigations at the scene continue**
- **Were you at The Maltings centre on Sunday? Contact [news@mailonline.co.uk](mailto:news@mailonline.co.uk)**

By [Chris Greenwood for the Daily Mail](#) and [Paul Thompson In Salisbury](#) and [Khaleda Rahman](#) and [Richard Spillett](#) and [Tim Stickings For Mailonline](#)

Published: 22:38 GMT, 6 March 2018 | Updated: 10:05 GMT, 7 March 2018

Former Russian double agent Sergei Skripal may have been ambushed by attackers who sprayed him with poison in the street.

Investigators suspect the Russian army colonel collapsed so quickly because he inhaled the deadly chemical.

One line of inquiry is that his daughter Yulia is 'collateral damage', coming into contact with the substance as she attempted to help him.

Another line of inquiry investigators are looking into is the possibility Mr Skripal's drink was spiked at a pub he and his daughter visited shortly before they collapsed.

CNN/Daily Mail  
287,113 train  
13,368 validation  
11,490 test

статья 632 слова  
реферат 53 слова

NYT  
600,000+ рефератов

Model	R-1	R-2
First sentences	28.6	17.3
First $k$ words	35.7	21.6
Full (Durrett et al., 2016)	42.2	24.9
ML+RL, with intra-attn	<b>42.94</b>	<b>26.02</b>