



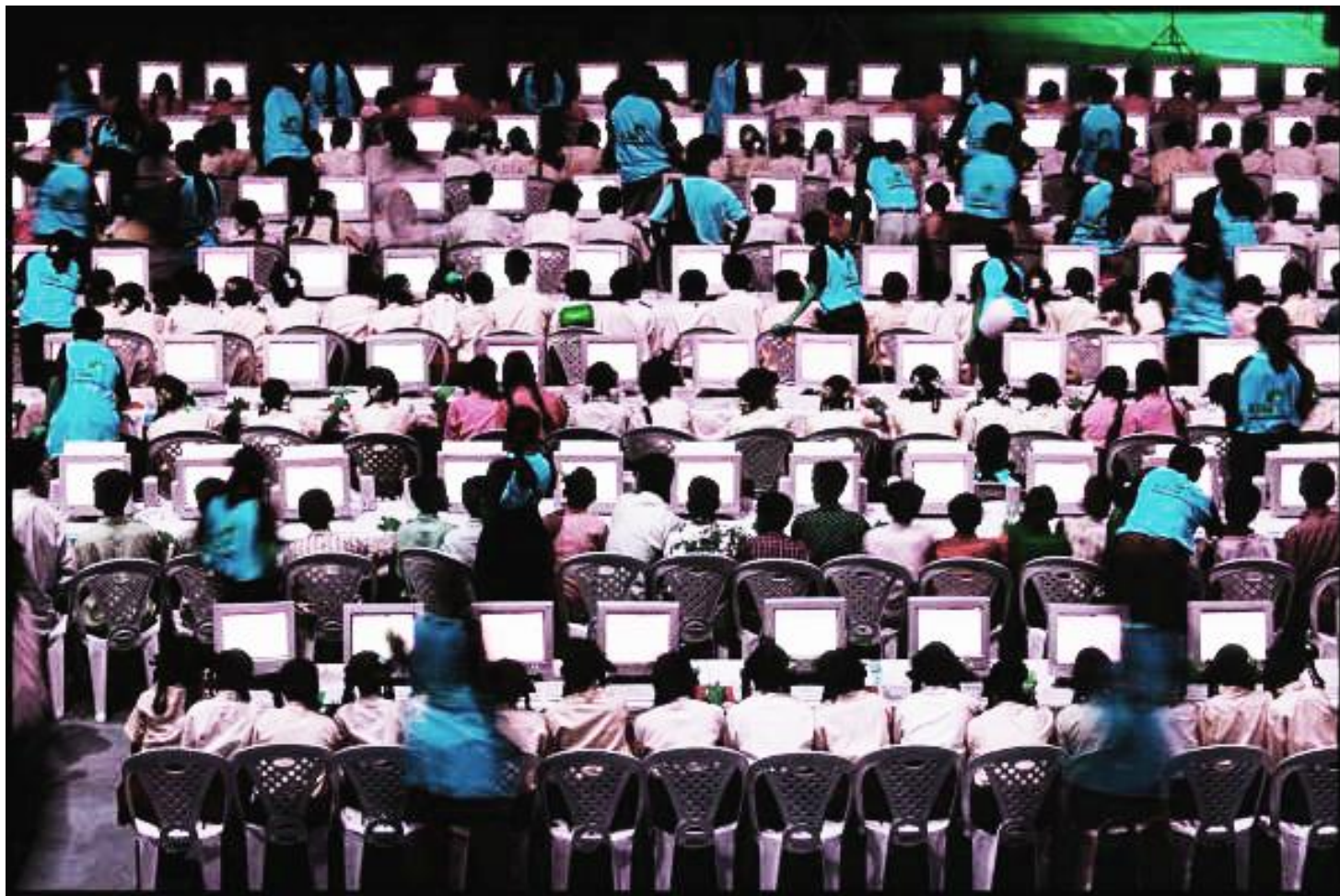
ТЕХНОСФЕРА

Введение в информационный поиск

Сергукова Юлия,
программист отдела инфраструктуры проекта
Поиск@Mail.Ru

Что такое Поиск?

Что такое Поиск?



Что такое Поиск?



Поиск@Mail.Ru





Найти



Интернет

Картинки

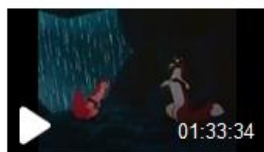
Видео

Новости

Ответы

Видео по запросу «**рики-тики-тави**»

3 тыс. результатов



01:33:34

Сборник Мультфильмов **Рикки-Тикки-Тави**
- **Рикки-Тикки-Тави**
Mail.ru

5 сентября 2013



00:19:35

Рикки-Тикки-Тави
Youtube
28 мая 2012



01:13:13

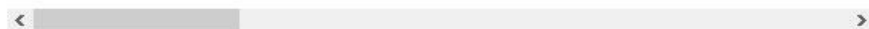
РИКИ. ТИКИ. ТАВИ. 1975 г. СССР
Ok.ru

7 сентября 2015



Рики - (1965)
Ok.ru

15 июля



📺 **Рикки-Тикки-Тави — Youtube.com** ✓

youtube.com/watch?v=ar-lv03EIXI

Мультки студии Союзмультфильм, 3 196 020 просмотров

Мультки : **Рикки тикки тави**. Советский мультфильм для детей по рассказу Р. Киплинга.

Нажми лайк, если любишь советские мультфильмы. Добрые советски...

📖 **Сказка Рикки Тикки Тави - Редьярд Киплинг** ✓

deti-online.com/skazki/skazki-...

Рикки Тикки Тави - сказка Киплинга. **Рикки Тикки Тави** читать онлайн или скачать в формате doc или pdf.



Рикки-Тикки-Тави

«Рі́ккі-Ті́ккі-Та́ви» — рассказ из «Книги джунглей» Редьярда Киплинга, а также имя главного героя этого рассказа. [Википедия](#)

Год первого издания: 1893

Авторы: [Редьярд Киплинг](#)

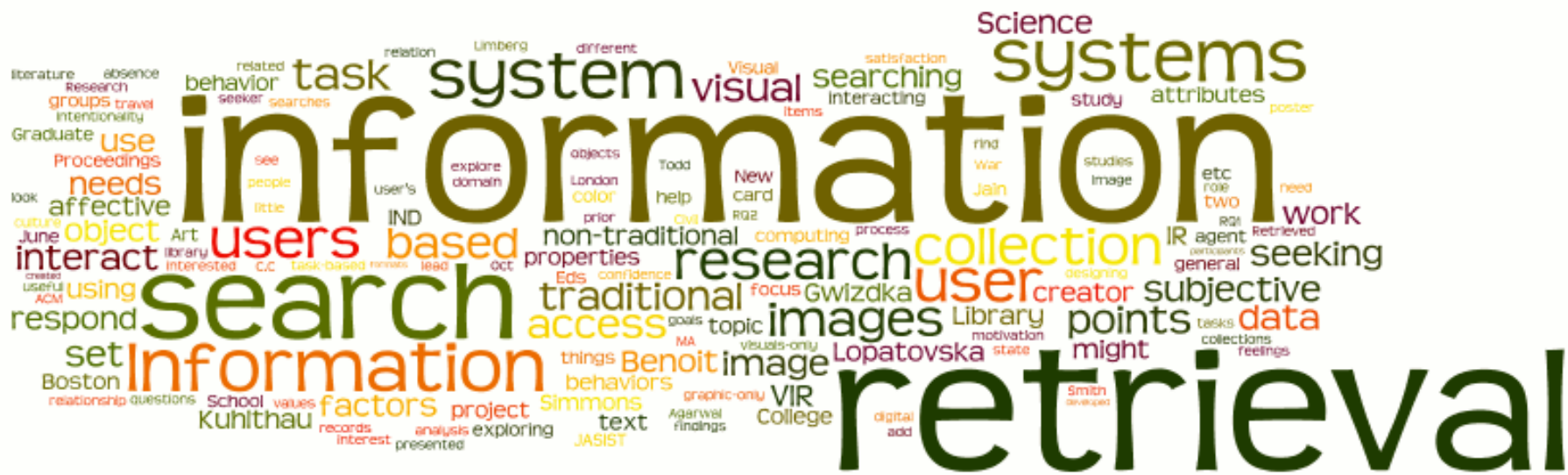
[Сообщить об ошибке](#) [Википедия](#) [Wikidata](#)

Поиск@Mail.Ru

- Текстовый поиск
- Картинки
- Видео
- Новости
- Ответы
- Карточка «быстрого ответа»

а еще:

- Приложения
- Обсуждения
- и многое другое



Что такое информационный поиск

Информацио́нный по́иск — процесс поиска неструктурированной документальной информации, удовлетворяющей информационные потребности пользователя

Примеры информационных поисков

- **grep**
- **find**
- **SQL**
- **Maps.Me / Google Maps**
- **Запросы в багтрекере**
- **Поиск по микроблогам**
- **Поиск по картинкам**
- **Вэб-поиск**

Информационный поиск

Наука

- математические модели
- вычислительная лингвистика
- статистика и теорвер

Практика

- огромный объем данных
- KISS («keep it simple, stupid»)
- хаки практической реализации

Цель

**«удовлетворение информационной потребности
пользователя»**

Цель

«удовлетворение информационной потребности пользователя»

У пользователя есть потребность в получении определенной информации

Цель – найти **релевантную** его запросу информацию

Входные данные

- **корпус документов** – набор документов (текстов, изображений, новостных блоков и т.д.)
- **ссылочная информация** – PageRank
- **обратная связь** – запросы пользователей и клики в выдачу

Пример

Задача

Поисковая система

Смерть («Плоский мир» Т.Пратчетта)



Пример



Пример



Пример



Пример



Что может пойти не так?

Что может пойти не так?

- Не поняли пользователя
 - Синонимы
 - Опечатки
 - Несодержательный запрос
- Нашли не то
 - Омонимы
 - Нет знаем нужный документ (непопулярные запросы)
 - Пользователь не согласен с результатами



Сложности обработки запросов от людей

- лингвистика: что написано в тексте
- запросы: что хотел пользователь
- статистика
- машинное обучение
- огромные объемы данных
- большие нагрузки
- пользовательский интерфейс
- форматы данных

Связанные задачи

- рекомендательные системы
- машинный перевод
- извлечение мнений
- распознавание и синтез речи
- семантический анализ текста
- диалог с пользователем
- и т.д.



Примеры поисковых систем


Южная Корея - Naver




NAVER



Google Scholar





Академия
Результатов: примерно 3 880 000 (0,05 сек.)

Статьи

Моя библиотека

За все время

С 2017

С 2016

С 2013

Выбрать даты

По релевантности

По дате

☒ включая патенты

☒ показать цитаты

☒ Создать оповещение

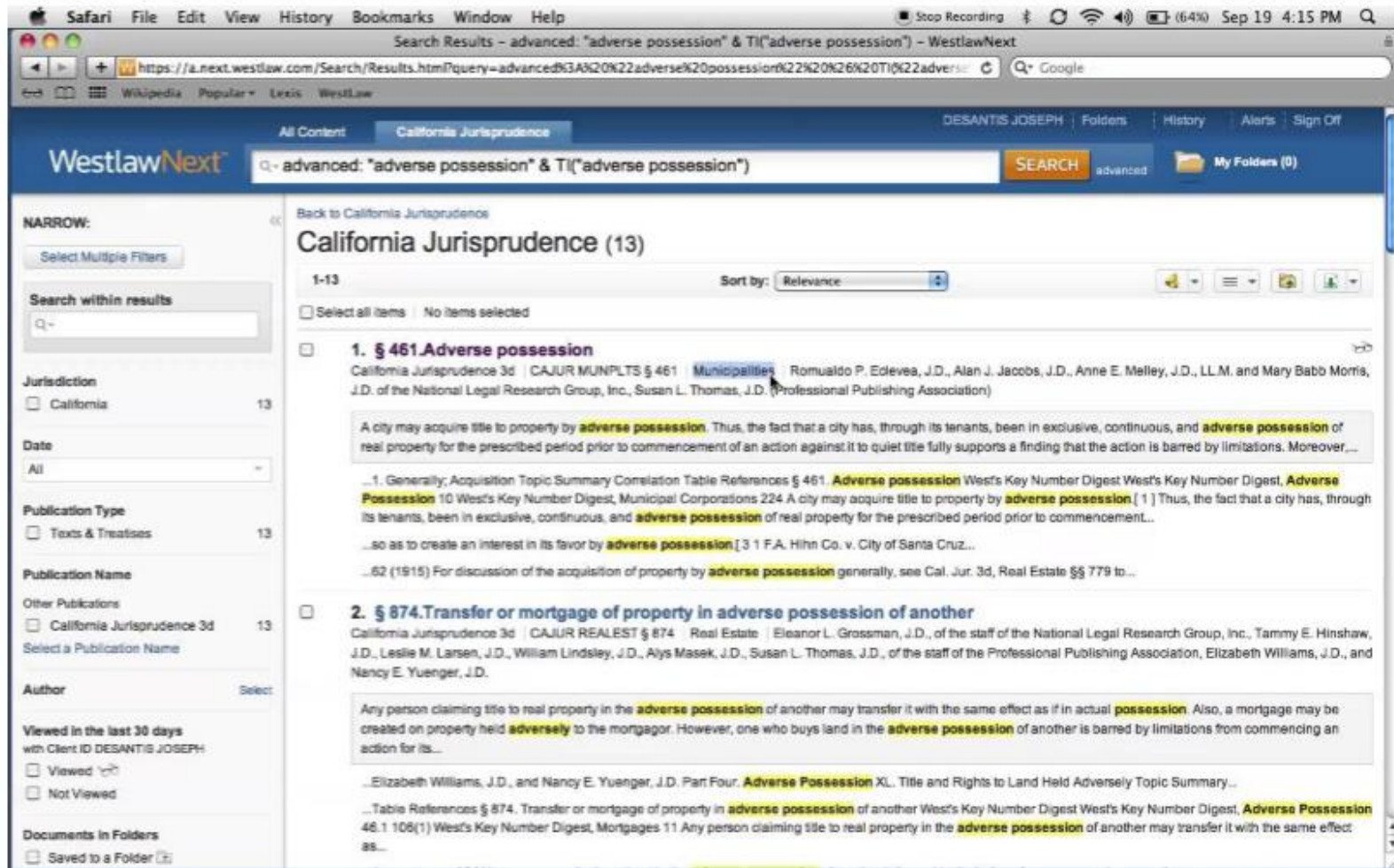
Information retrieval: data structures and algorithms
[WB Frakes](#), [R Baeza-Yates](#) - 1992 - citeulike.org
 Abstract Information retrieval is a sub-field of computer science that deals with the automated storage and retrieval of documents. Providing the latest information retrieval techniques, this guide discusses Information Retrieval data structures and algorithms,
 Цитируется: 2889 [Похожие статьи](#) [Все версии статьи \(4\)](#) [Цитировать](#) [Сохранить](#) [Ещё](#)

[ЦИТИРОВАНИЕ] Introduction to modern information retrieval
[G Salton](#), [MJ McGill](#) - 1986 - citeulike.org
 ... More... Brought to you by AQnowledge, precision products for scientists. x CiteULike uses cookies, some of which may already have been set. Read about how we use cookies. We will interpret your continued use of this site as your acceptance of our use of cookies. You may hide this
 Цитируется: 14231 [Похожие статьи](#) [Все версии статьи \(6\)](#) [Цитировать](#) [Сохранить](#) [Ещё](#)

Introduction
[DAC Manning](#) - Introduction to Industrial Minerals, 1995 - Springer
 Page 1. ~ ____ l_n_tr_o_d_u_c_t_i_o_n ____ ~1 ~ Human exploitation of minerals extends back for many thousands of years and, contrary to popular belief, mining may in fact be the 'oldest profession'. Early people used ...
 Цитируется: 12981 [Похожие статьи](#) [Все версии статьи \(16\)](#) [Цитировать](#) [Сохранить](#)

[книга] Modern information retrieval
[R Baeza-Yates](#), [B Ribeiro-Neto](#) - 1999 - mail.im.tku.edu.tw
Information retrieval (IR) has changed considerably in recent years with the expansion of the World Wide Web and the advent of modern and inexpensive graphical user interfaces and mass storage devices. As a result, traditional IR textbooks have become quite out of date
 Цитируется: 15957 [Похожие статьи](#) [Все версии статьи \(41\)](#) [Цитировать](#) [Сохранить](#) [Ещё](#)

Westlaw



The screenshot shows a Safari browser window displaying the WestlawNext search results page. The search query is "advanced: 'adverse possession' & T('adverse possession')". The results are filtered to "California Jurisprudence" and show 13 items. The first item is highlighted, titled "1. § 461. Adverse possession". The text of the item discusses the acquisition of title to property by adverse possession, citing California Jurisprudence 3d and CAJUR MUNPLTS § 461. The second item is titled "2. § 874. Transfer or mortgage of property in adverse possession of another". The text of this item discusses the transfer of property in adverse possession, citing California Jurisprudence 3d and CAJUR REALEST § 874. The left sidebar contains filters for Jurisdiction (California), Date (All), Publication Type (Texts & Treatises), Publication Name (California Jurisprudence 3d), and Author (Select). The bottom of the sidebar shows "Viewed in the last 30 days" and "Documents in Folders".

Search Results – advanced: "adverse possession" & T("adverse possession") – WestlawNext

Search Results – advanced: "adverse possession" & T("adverse possession") – WestlawNext

DESANTIS JOSEPH | Folders | History | Alerts | Sign Off

WestlawNext

Q - advanced: "adverse possession" & T("adverse possession")

SEARCH advanced

My Folders (0)

NARROW:

Select Multiple Filters

Search within results

Jurisdiction

California 13

Date

All

Publication Type

Texts & Treatises 13

Publication Name

Other Publications

California Jurisprudence 3d 13

Select a Publication Name

Author

Select

Viewed in the last 30 days with Client ID DESANTIS JOSEPH

Viewed

Not Viewed

Documents in Folders

Saved to a Folder

Back to California Jurisprudence

California Jurisprudence (13)

1-13

Sort by: Relevance

Select all items | No items selected

1. § 461. Adverse possession

California Jurisprudence 3d | CAJUR MUNPLTS § 461 | Municipalities | Romualdo P. Edevea, J.D., Alan J. Jacobs, J.D., Anne E. Melley, J.D., LL.M. and Mary Babb Morris, J.D. of the National Legal Research Group, Inc., Susan L. Thomas, J.D., of the Professional Publishing Association)

A city may acquire title to property by **adverse possession**. Thus, the fact that a city has, through its tenants, been in exclusive, continuous, and **adverse possession** of real property for the prescribed period prior to commencement of an action against it to quiet title fully supports a finding that the action is barred by limitations. Moreover,...

...1. Generally; Acquisition Topic Summary Correlation Table References § 461. **Adverse possession** West's Key Number Digest West's Key Number Digest, **Adverse Possession** 10 West's Key Number Digest, Municipal Corporations 224 A city may acquire title to property by **adverse possession** [1] Thus, the fact that a city has, through its tenants, been in exclusive, continuous, and **adverse possession** of real property for the prescribed period prior to commencement...

...so as to create an interest in its favor by **adverse possession** [3 1 F.A. Hihn Co. v. City of Santa Cruz...

...62 (1915) For discussion of the acquisition of property by **adverse possession** generally, see Cal. Jur. 3d, Real Estate §§ 779 to...

2. § 874. Transfer or mortgage of property in adverse possession of another

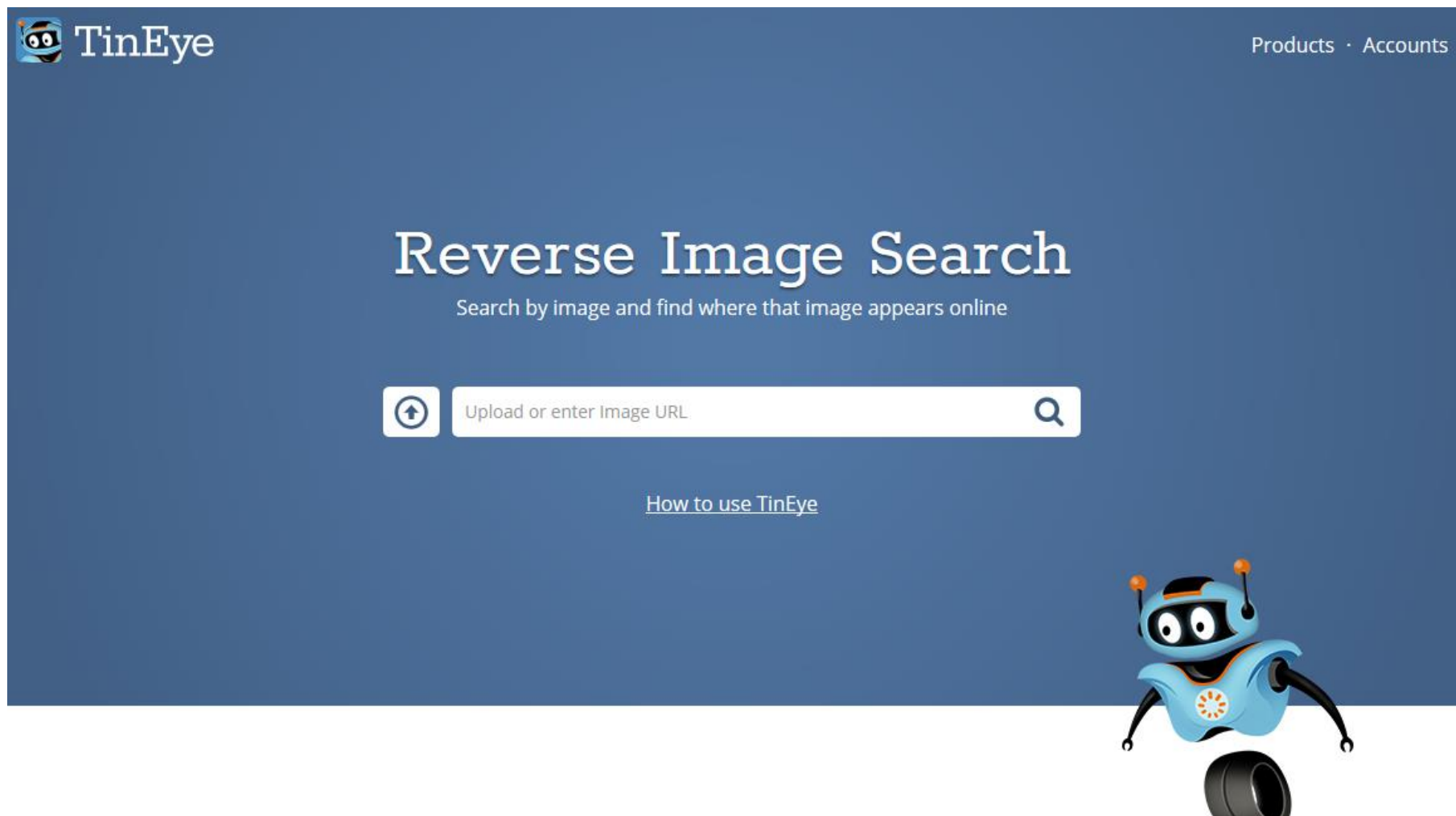
California Jurisprudence 3d | CAJUR REALEST § 874 | Real Estate | Eleanor L. Grossman, J.D., of the staff of the National Legal Research Group, Inc., Tammy E. Hinshaw, J.D., Leslie M. Larsen, J.D., William Lindsley, J.D., Alys Masek, J.D., Susan L. Thomas, J.D., of the staff of the Professional Publishing Association, Elizabeth Williams, J.D., and Nancy E. Yuenger, J.D.

Any person claiming title to real property in the **adverse possession** of another may transfer it with the same effect as if in actual **possession**. Also, a mortgage may be created on property held **adversely** to the mortgagor. However, one who buys land in the **adverse possession** of another is barred by limitations from commencing an action for its...

...Elizabeth Williams, J.D., and Nancy E. Yuenger, J.D. Part Four. **Adverse Possession** XL. Title and Rights to Land Held Adversely Topic Summary...

...Table References § 874. Transfer or mortgage of property in **adverse possession** of another West's Key Number Digest West's Key Number Digest, **Adverse Possession** 46.1 106(1) West's Key Number Digest, Mortgages 11 Any person claiming title to real property in the **adverse possession** of another may transfer it with the same effect 88...

Не-текстовый поиск. TinEye



Не-текстовый поиск.

Twenty-two settings

[DOWNLOAD ABC](#)

X: 1

T: Drowsy Maggie

R: reel

M: 4/4

L: 1/8

K: Edor

|:E2BE dEBE|E2BE AFDF|E2BE dEBE|BABc dAFD:|


d2fd c2ec|defg afge|d2fd c2ec|BABc dAFA|

d2fd c2ec|defg afge|afge fdec|BABc dAFD|

Added by *Jeremy* 17 years ago.

[DOWNLOAD](#)[PRINT](#)

Специализированный поиск.

 searchcode

TableMapReduceUtil.initTableMapperJob

search

AboutDevelopersUpdatessearchcode server

About 86 results: "TableMapReduceUtil.initTableMapperJob"

Page 1 of 5

◀ Previous

Next ▶

Filter Results

Remove

Apply

Sources

☐ Github

85

☐ Bitbucket

1

Languages

Filter Languages

☐ Java

79

☐ Perl

3

☐ HTML

3

☐ Scala

1

Filter Results

Remove

Apply

Try Search On

[GitHub Code](#)PreDownsampleInit.java in mon4h <https://github.com/bruceleexiaokan/mon4h.git> | 219 lines | Java20. import org.apache.hadoop.hbase.io.ImmutableBytesWritable;
21. import org.apache.hadoop.hbase.mapreduce.TableMapReduceUtil;
77.
78. TableMapReduceUtil.initTableMapperJob(tableRHD.ReadTable, scans[i],
81.
82. TableMapReduceUtil.initTableReducerJob(tableRHD.ReadTable,
102.
103. TableMapReduceUtil.initTableMapperJob(tableRHD.ReadTable, scans[i],
105. KeyValue.class, jobs[i]);
106. TableMapReduceUtil.initTableMapperJob(tableRHD.ReadTable, scan,
109.
110. TableMapReduceUtil.initTableReducerJob(tableRHD.ReadTable,
111. PreDownsampleReduceHourJob.class, jobs[i]);
112. TableMapReduceUtil.initTableReducerJob(tableRHD.ReadTable,
book.html in SecureBase <https://github.com/dishant/SecureBase.git> | 3804 lines | HTML1743.
1744. TableMapReduceUtil.initTableMapperJob(
1778.
1779. TableMapReduceUtil.initTableMapperJob(
1785. job);
1786. TableMapReduceUtil.initTableReducerJob(
1796. </pre><p>
An explanation is required of what `class="classname">TableMapReduceUtil</code> is doing, especially w`

Контентные рекомендательные системы

Когда Каламу повесили на шею медаль победителя, он решил разделить эту победу с дочкой и отдал награду ей. Малышка даже попробовала медаль на зуб. «Это действительно важный день для меня, и я хочу разделить эту победу с младшей дочерью. То, что нам это удалось, говорит о том, что быть папой и работающим человеком вовсе не значит, что нужно отказываться от своих мечтаний, надежд и хобби», – заявил Калам.

Он рассчитывает стать примером для других семей, показывая, что активный образ жизни на свежем воздухе приносит пользу и детям, и взрослым.

Также смотрите [15 трогательных фото об отцовской любви](#).


Материал подготовила Татьяна Щеглова

11 февраля 2016 [материалы](#) [Все темы](#)


Обновление в право собственности, прокатчика, выданных на и название On • Eke.

Все новости


Материалы по теме




Климова рассказала об интересной семейной традиции




Ученые предупредили об опасностях позднего отцовства



Какие подарки получили принцесса Шарлотта с братом




Анна Хильминович назвала дочь в честь себя и мужа




Константин Хабенский скоро снова станет папой

Директ

Системы Аквафор Морион от 5790 р.
[tek.shop.aquaphor.ru](#)



Рекуператор воздуха квартирный
[mailey-rus.ru](#)



Приточная вентиляция в квартире
свежий-воздух.рф

Приточная и приточно-вытяжная вентиляция в квартире. Монтаж. Распродажа без%

Адрес и телефон

Pandora's "Music Genome Project" explores the cold hard facts of how we interact with music

Twenty-five music analysts "grade" 10,000 songs a month. It's a mountainous job, writes Rob Pegoraro, but the results will be serious business.



Поиск по структурированной информации и неструктурированным запросам



☆☰

📄 📷 📊 🔊 🌐 Web Apps ☰ Examples 🔗 Random

Assuming "apple" is a food | Use the input as a [formula](#) instead

Assuming apple | Use [prepared apples](#) or [more](#) instead

Assuming any type of apple | Use [apple, with skin](#) or [apple, without skin](#) instead

Input interpretation:

apple	amount	1 apple	mass
-------	--------	---------	------

Result:

182 grams

[Show non-metric](#)

Unit conversions:

0.18 kg (kilograms)

0.4 lb (pounds)

6.4 oz (ounces)

Physical properties:

mass	182 grams
serving volume	233 mL (milliliters)

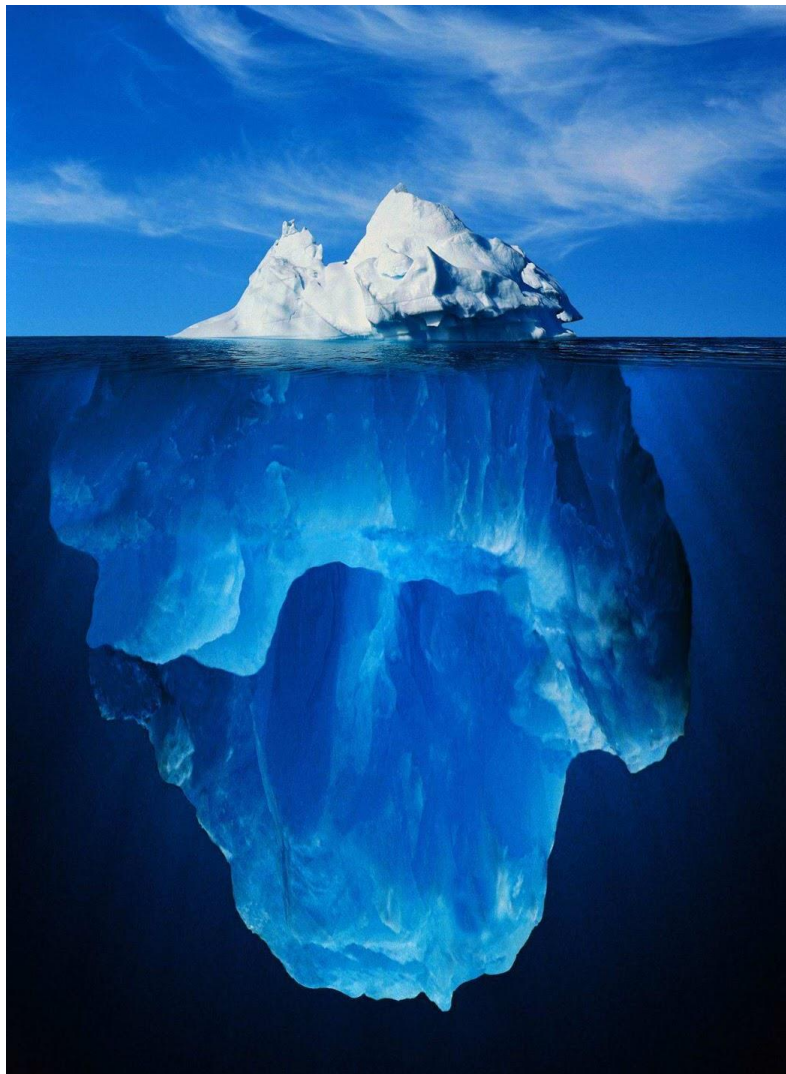
IBM Watson



Обзор курса



Обзор курса



Преподаватели



Юлия
Сергукова



Андрей
Мурашев



Евгений
Чернов



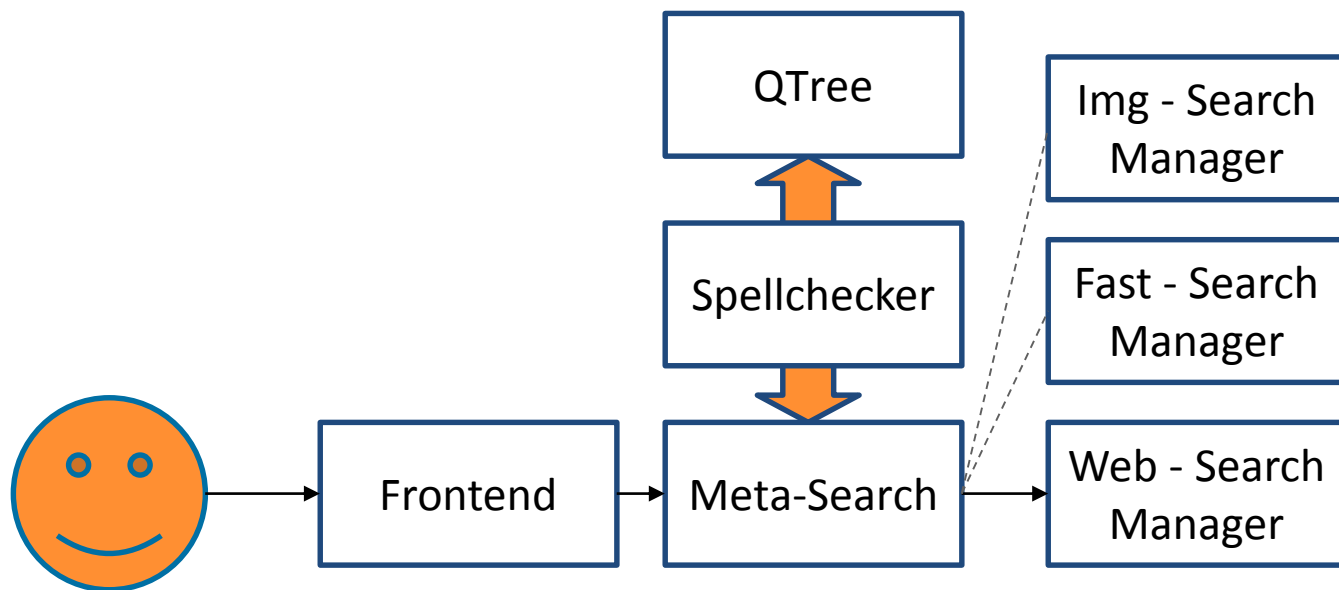
Владимир
Гулин

Как работает поиск

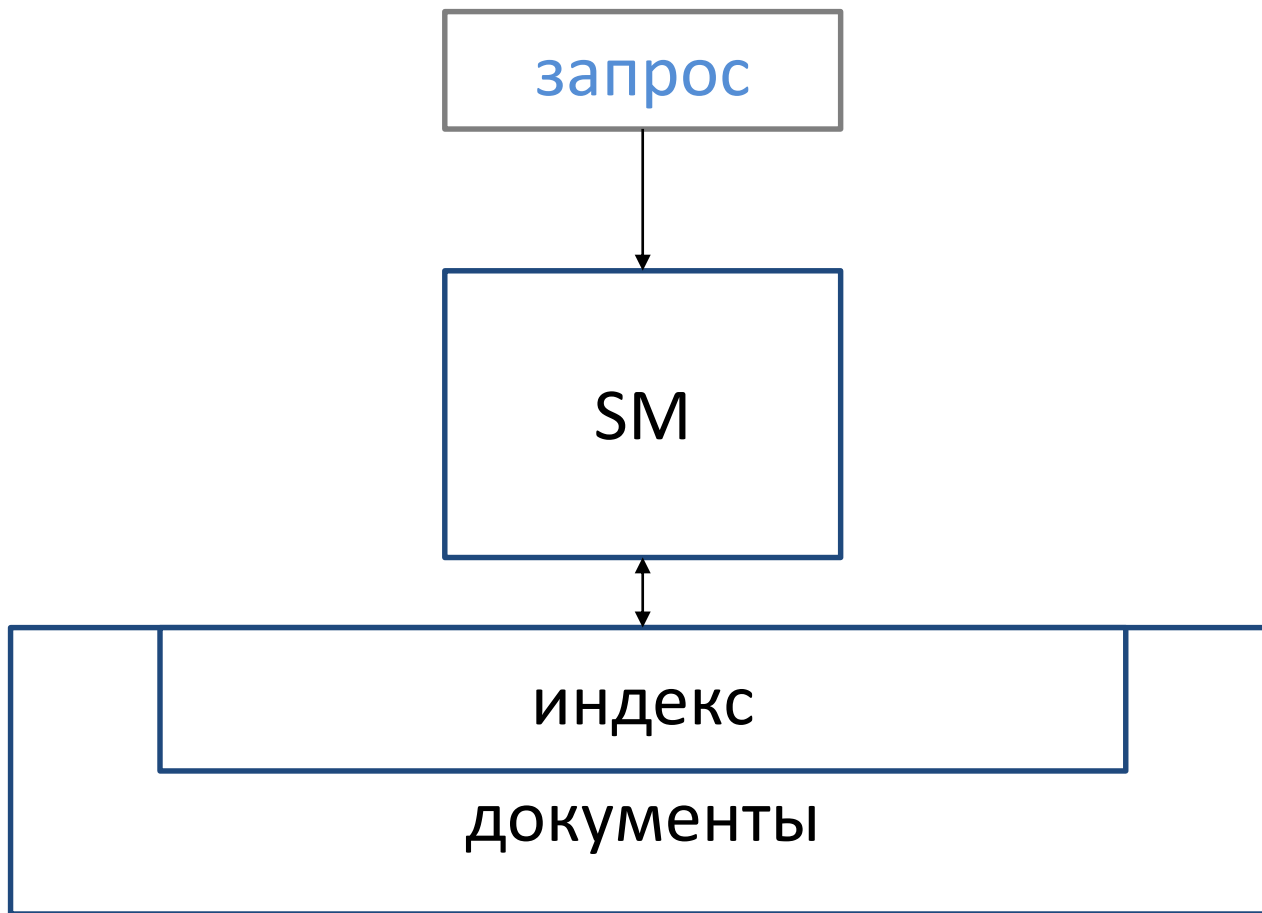
Как работает поиск



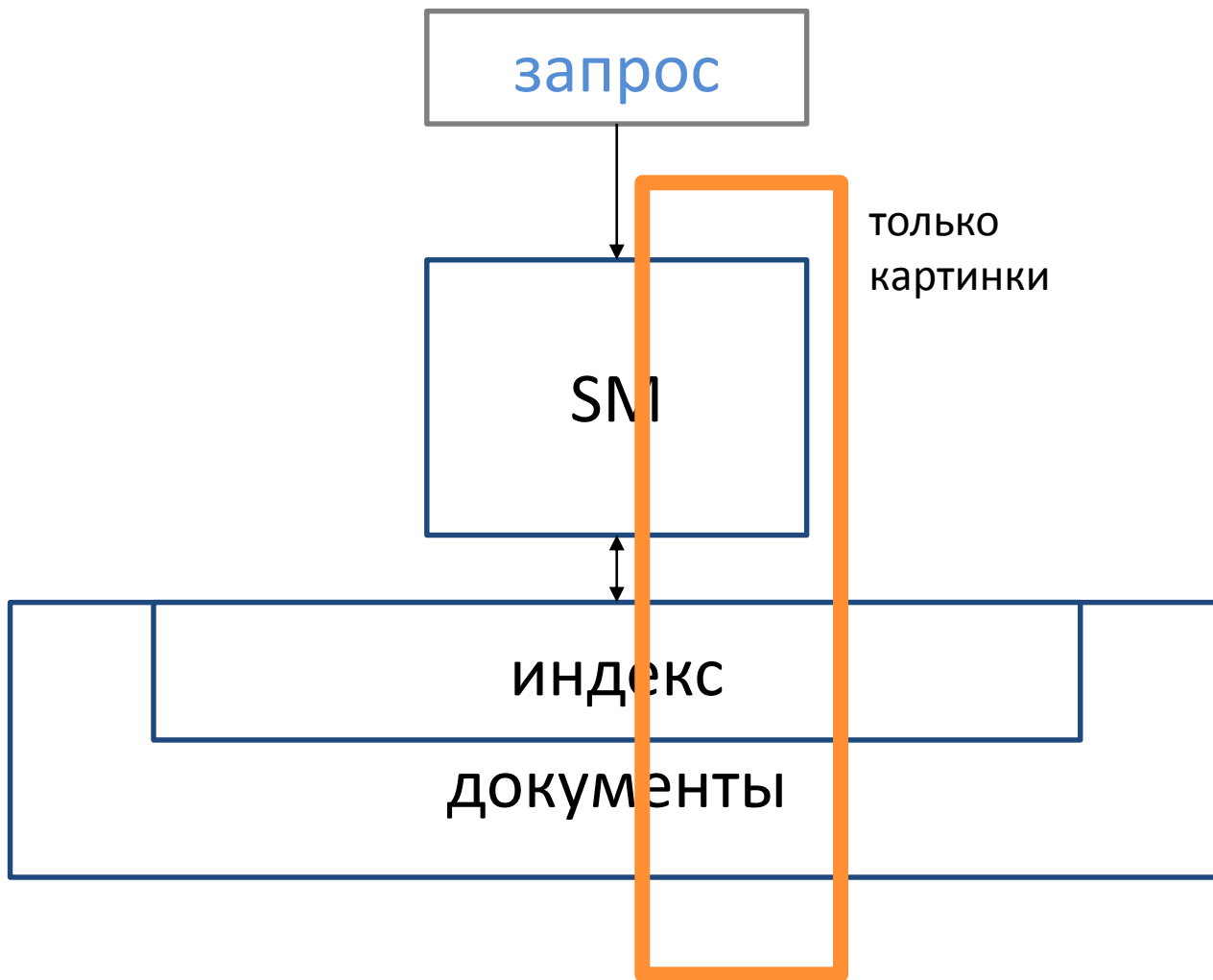
Как работает поиск



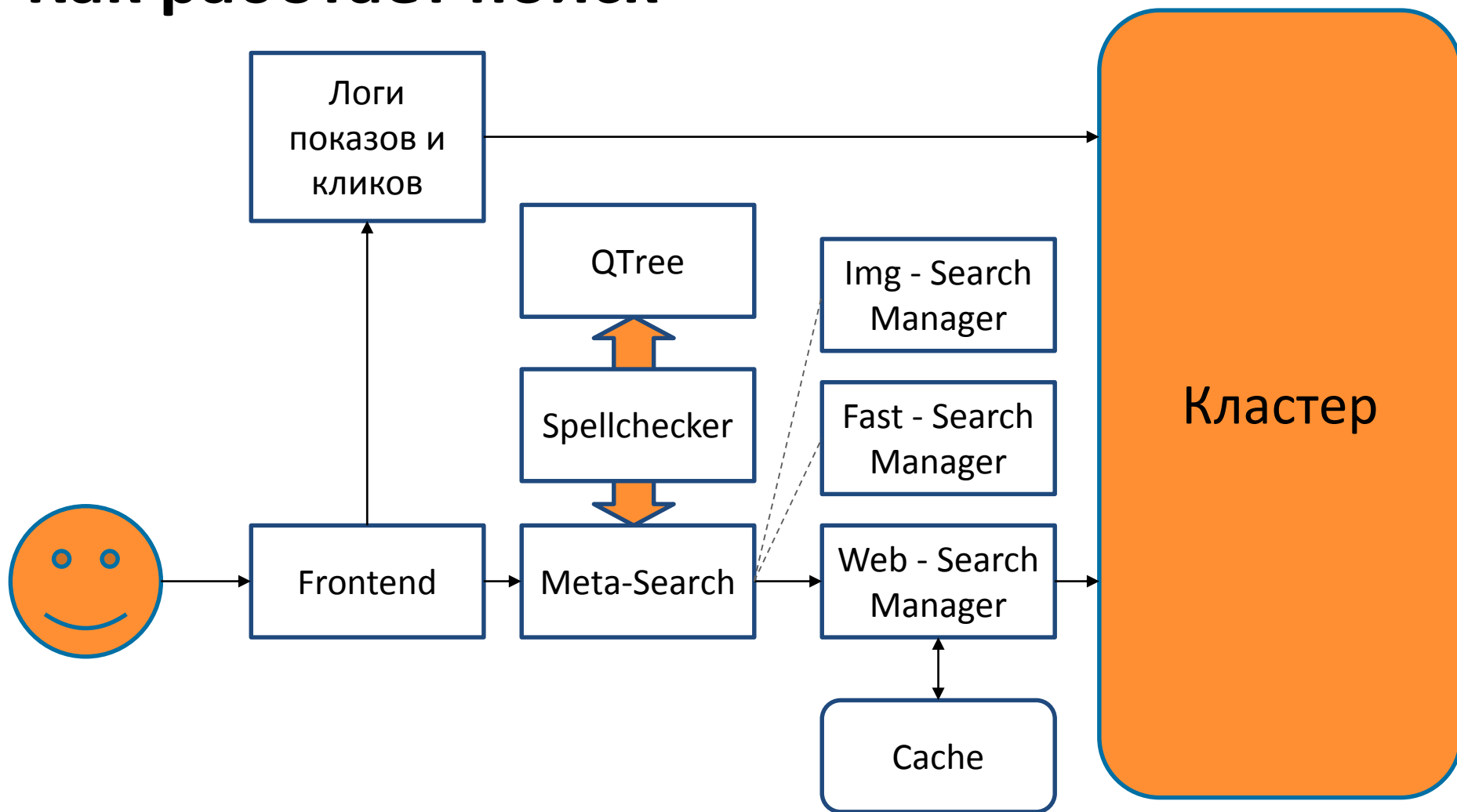
Немного о вертикалях



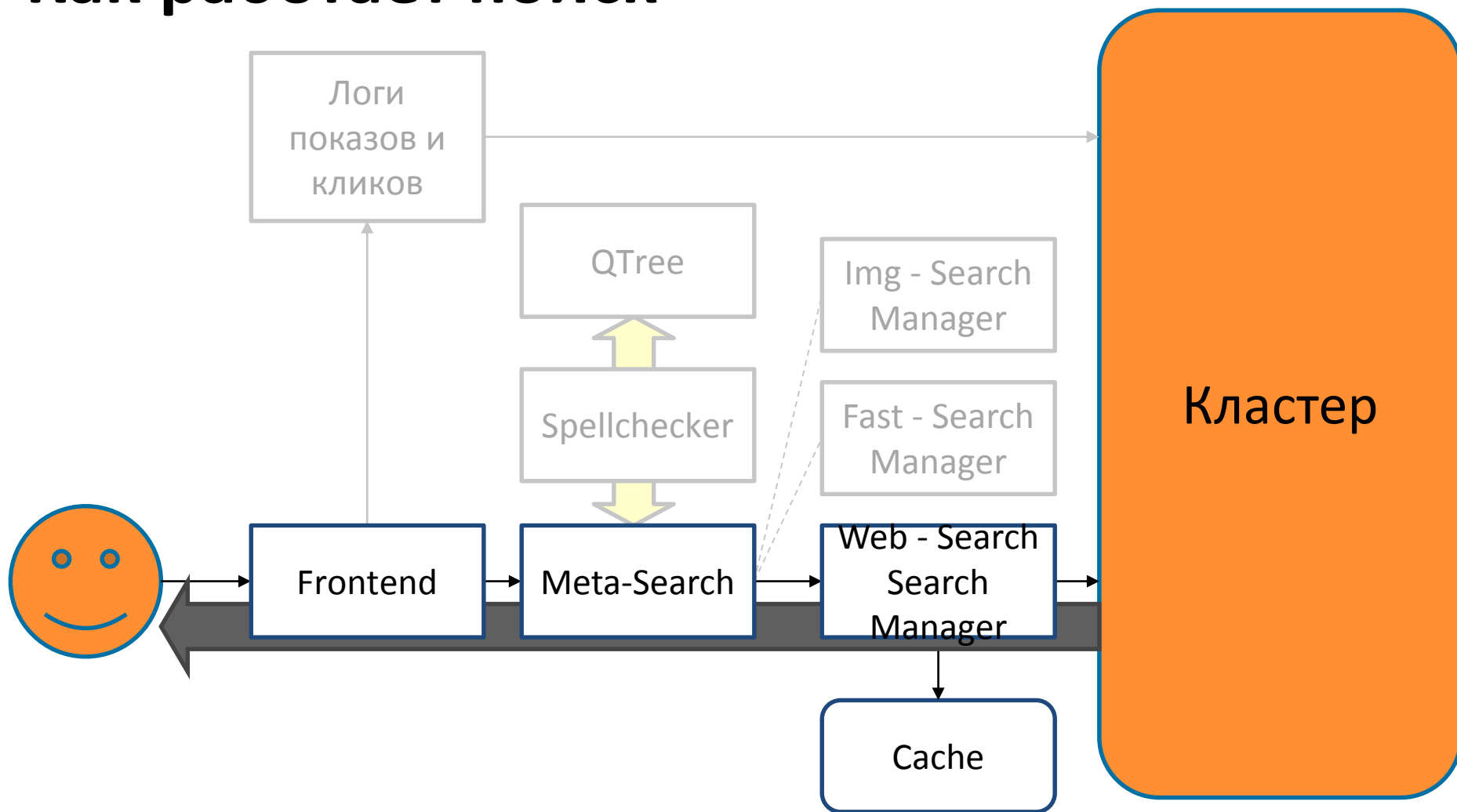
Немного о вертикалях



Как работает поиск



Как работает поиск



О чем мы будем говорить на лекциях

1. Поисковый спайдер (лекция 2)
2. Планировщик поискового спайдера (лекция 3)
3. Индексация (лекция 4) и оптимизация индекса (лекция 5)
4. Поиск дубликатов (лекции 6-7)
5. Антиспам и антипорн (лекции 8 и 9)
6. Микроразметка (лекция 10)
7. С니ппеты (лекция 11)
8. Исправление опечаток (лекция 12)
9. Саджесты и переформулировки (лекция 13)

О курсе

Основан на **Introduction to information retrieval (Stanford)**

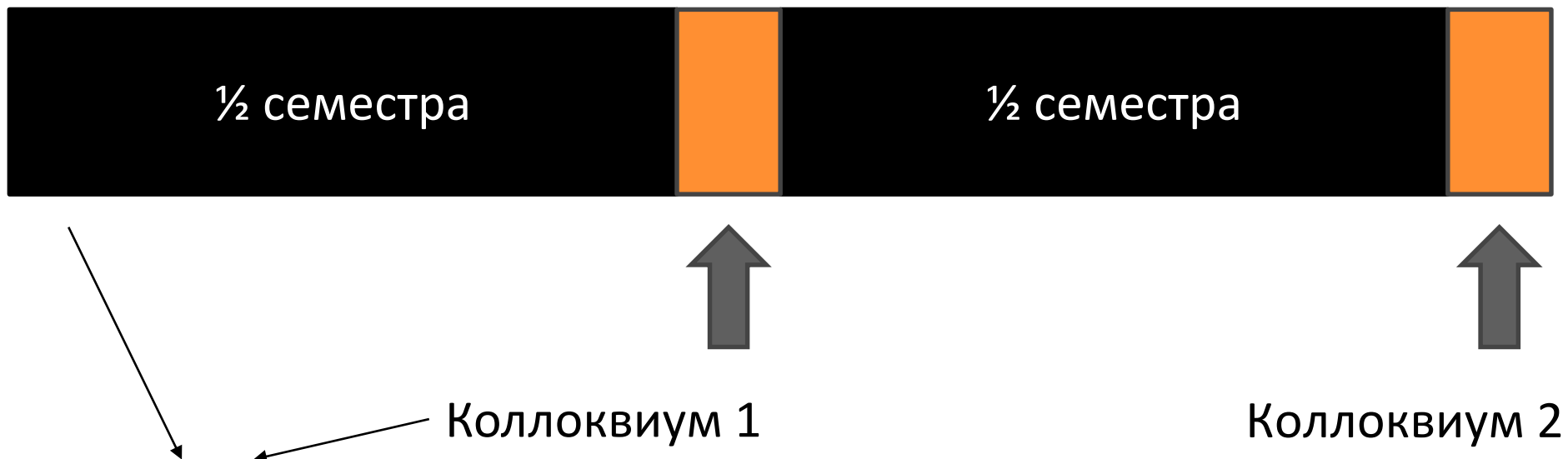
<https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>

Читаем более 10 лет

Основной упор – на практическое применение

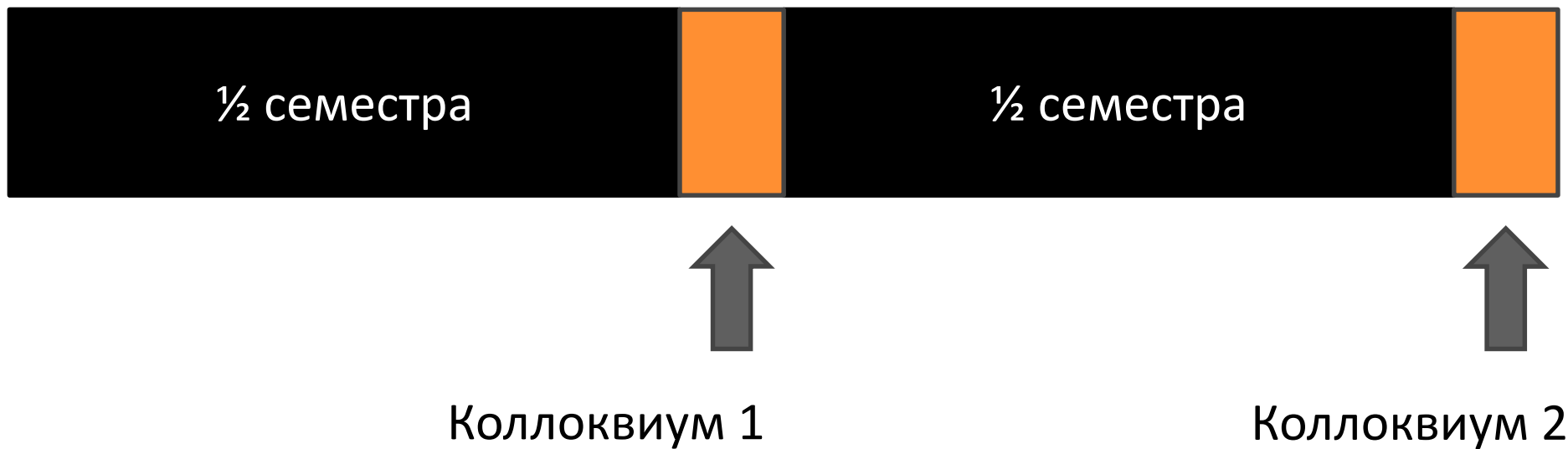
Делайте домашки!

Структура курса



3 домашние работы
коллоквиум (опционально, 0-10 баллов)
порог выживания = 35 баллов

Структура курса



Весь курс:

6 домашних работ

2 коллоквиума (опционально, 0-20 баллов)

порог выживания = 60 баллов

Делайте домашки!

Для допуска к первому коллоквиуму нужно набрать пороговое значение баллов - 35.

Не набрали – прощаемся с вами в середине семестра.

Оценки

За что можно получить баллы:

- домашние работы
- активная работа на лекция и семинарах
- сдача коллоквиума

Оценки:

60 – 79 == «3»

80 – 94 == «4»

95+ == «5»

Домашние работы

- проверяются преимущественно автоматически
- списывание и/или непроходимая глупость – 0 баллов
- дедлайн – обычно 2 недели

2 недели

до конца курса

N баллов

N/2 баллов

Если что-то непонятно:

в первую очередь спрашивайте преподавателя, который
который выдал домашку/читал лекции.

Блог – ваш друг

<https://sphere.mail.ru/blog/view/40/>

Зачем?

- Выкладываются лекции
- Перед семинаром даются указания
- Выкладываются домашки (и иногда результаты)
- Иногда бывают изменения в расписании

Лайки!



...а еще отзывы и посещаемость

Спасибо за внимание!

**Не забудьте отметить на сайте
и оставить отзыв**