

Эффективный поиск в интернете

Юрий Лифшиц
<http://yury.name>

Сентябрь 2007

Сегодня в программе

1 Поисковые системы изнутри

Сегодня в программе

- 1 Поисковые системы изнутри
- 2 Технические аспекты

Сегодня в программе

- 1 Поисковые системы изнутри
- 2 Технические аспекты
- 3 Техники эффективного поиска

Сегодня в программе

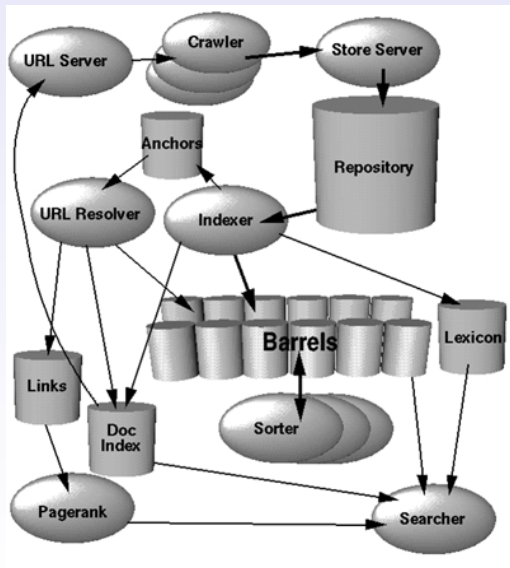
- 1 Поисковые системы изнутри
- 2 Технические аспекты
- 3 Техники эффективного поиска
- 4 Практика

Анатомия поисковой системы

Любая поисковая система содержит три базовые части:

- Робот (он же краулер, спайдер или индексатор)
- Базы данных
- Ядро (обработка запросов)

Схема из [Brin&Page, 1998]



Релевантность

Характеристики, влияющие на позицию в списке ответов?

Характеристики, влияющие на позицию в списке ответов?

- Наличие слов на сайте
- Частота слов
- Форматирование
- Близость слов друг к другу
- Количество ссылок с других страниц на данную
- Качество ссылок
- Соответствие тематик сайта и запроса

Технические аспекты: браузер

- Левая кнопка мыши → Open in a New Tab
- Хранение найденного: избранное, del.icio.us
- Ctrl + T — открыть еще один таб
- Не бояться одновременно работать со многими сайтами
- Поисковые формы — Яндекс Бар, Google Bar

Расширенный поиск

- Страницы, ссылающиеся на данную
- Страницы, “похожие на данную”
- Слова запроса должны “стоять подряд”
- Региональный поиск (например, только петербургские сайты)
- Поиск по каталогу
- Воскрешение “умерших” страниц
- Поиск по конкретному сайту
- AND, OR, NOT

Языковые аспекты

- На английском языке искать сложнее, но информация значительно полнее
- Используйте lingvo.yandex.ru для переводов слов, translate.google.com для перевода сайтов

Полезные сайты

- [Wikipedia.org](https://www.wikipedia.org)
- Каталоги и рейтинги: [Dmoz.org](https://www.dmoz.org), [Yahoo.com](https://www.yahoo.com), [List.ru](https://www.list.ru)
- Специализированные поисковики: [filesearch.ru](https://www.filesearch.ru), scholar.google.com, Яндекс-Маркет, [expedia.com](https://www.expedia.com)
- Мета-поисковики ([Nigma.Ru](https://www.Nigma.Ru))
- Личные коллекции ссылок

Для чего вы пользуетесь интернет-поиском?

Для чего вы пользуетесь интернет-поиском?

- Сбор информации
- Методы, советы, типичные ошибки и как их избегать
- Самообразование/творческий рост

Примеры:

- Обзор ноутбуков
- Как написать хороший пресс-релиз?
- Как проводить тренинги?

Основные шаги

Основные шаги:

- 1 Принять решение о поиске

Основные шаги

Основные шаги:

- 1 Принять решение о поиске
- 2 Сформулировать тему. Определить тип ресурсов, которые надо найти

Основные шаги:

- 1 Принять решение о поиске
- 2 Сформулировать тему. Определить тип ресурсов, которые надо найти
- 3 Подобрать адекватные термины

Основные шаги:

- 1 Принять решение о поиске
- 2 Сформулировать тему. Определить тип ресурсов, которые надо найти
- 3 Подобрать адекватные термины
- 4 Найти сайты с информацией

Основные шаги:

- 1 Принять решение о поиске
- 2 Сформулировать тему. Определить тип ресурсов, которые надо найти
- 3 Подобрать адекватные термины
- 4 Найти сайты с информацией
- 5 Выбрать лучшие и изучить их

Основные трудности

Что мешает находить в интернете нужные материалы?

Основные трудности

Что мешает находить в интернете нужные материалы?

- Забываем про возможность поискать интернете
- Знаем проблему, но не знаем, что искать
- Теряем фокус (отвлекаемся на интересные, но не относящиеся к делу материалы)
- Делаем только поверхностный поиск
- Поисковый спам

Основные ошибки

- “Лишние” слова в строке запроса
- Неудачный выбор главного ключевого слова
- Плохой перевод (**Пример несоответствия:**
Фото архив — Photo sharing)
- “Уже пробовал”
- Недостаточная гибкость в смене ключевых слов и методов
- Разделять во времени поиск и чтение!
- “Поисковик ничего не нашел” \Rightarrow плохой запрос!

Что может быть объектом поиска?

Что может быть объектом поиска?

- Тематический ресурс (например Improvement.Ru)
- Сообщество
- Коллекция популярных статей
- Форум
- Каталог, коллекция ссылок
- Файл
- Энциклопедическая статья
- Электронная библиотека
- Контактная информация (E-mail)
- Сайты-сервисы

Краткий список служебных слов

- Скачать, download, free
- Pdf, ppt, doc, zip, txt, mp3
- Форум, каталог, ссылки, forum, directory, links
- Для начинающих, первые шаги, руководство, советы, правила, faq, for newbies, for beginners, guide, rules, checklist
- Конспект лекций, обзор, lecture notes, survey
- Как, где, почему, хорошо, правильно
- “ — это”

Ключевые слова

Развивайте внимание к ключевым словам. Они живут:

- В энциклопедических статьях
- В описаниях сайтов (в каталогах)
- В заголовках уже найденных сайтов

Мои лучшие находки

- Graduate studies advice
- Lecture notes
- Тайм менеджмент
- Поисковая оптимизация

Полномасштабный поиск

Шаги:

- Сформулировать цель
- Найти главные ресурсы области
- Найти специализированные каталоги
- Подобрать ключевые слова
- Найти личные коллекции ссылок
- Найти форумы и контакты экспертов
- Менять главное ключевое слово (3-4 раза)

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке
- Использовать каталоги и энциклопедии (Wikipedia.org)

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке
- Использовать каталоги и энциклопедии (Wikipedia.org)
- Искать не только информацию, но и ресурсы

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке
- Использовать каталоги и энциклопедии (Wikipedia.org)
- Искать не только информацию, но и ресурсы
- Главная техника — служебные слова

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке
- Использовать каталоги и энциклопедии (Wikipedia.org)
- Искать не только информацию, но и ресурсы
- Главная техника — служебные слова
- Развивайте внимание к ключевым словам

Повторим самое главное!

- Помнить про возможности интернета. Чаще принимать решение о поиске
- Использовать поиск на английском языке
- Использовать каталоги и энциклопедии (Wikipedia.org)
- Искать не только информацию, но и ресурсы
- Главная техника — служебные слова
- Развивайте внимание к ключевым словам
- Для важных вопросов делайте полномасштабный поиск

От теории к практике!

- Игра “Uno Google”: подобрать запрос из двух слов, в ответ на который Google выдаст ровно один результат
- Найти технику конькового хода на лыжах (с картинками)
- Найти список тиров Санкт-Петербурга
- День рождения Максима Яковлевича?
- Найти материалы на тему “как делать хорошие презентации”