

# Настоящее и будущее ПОИСКОВЫХ МАШИН

ИГОРЬ АШМАНОВ

\*



Ашманов  
и партнеры



# ПОИСКОВИКИ – СТЕРЖЕНЬ ИНТЕРНЕТ-РЕКЛАМЫ

- **Доходы от рекламы.** Поисковики зарабатывают больше половины всех доходов от интернет-рекламы (400 миллионов долларов в Рунете в 2009).
- **Посетители.** Поисковики дают до 70-80% всех заходов на сайты. Поисковики «излучают» в Рунет сотни миллионов переходов в день.
- **Производные рынки.** Поисковики кормят десятки тысяч оптимизаторов с годовым оборотом до 200 млн. долл.
- **Целевой посетитель с поиска** – самый качественный.



# ДОЛИ РЫНКА МЕНЯЮТСЯ БЫСТРО И ДРАМАТИЧЕСКИ

- Гугл быстро рос в России в 2006-2008, сейчас теряет долю.
- Яндекс терял долю, сейчас – быстро растёт.
- Рамблер с 2007 потерял долю с 22% до 3%.
- Яндекс потерял долю в Украине, Гугл набрал. Сейчас Яндекс начал рост.
- Гугл проиграл поисковый рынок Китая.

**Цена вопроса – миллиарды долларов.**



# СВЯЗЬ КАЧЕСТВА И ДОЛИ РЫНКА

- Есть ли прямая связь качества поиска и роста доли рынка?
- Маркетинг (раздача браузеров, работа с OEM, другие партнёрства, реклама в офлайне, реклама на ТВ) влияет очень сильно.
- Разработчики поиска абсолютно уверены, что качество является главным показателем.

**Качество – необходимое условие,  
маркетинг – достаточное.**



# ПРОБЛЕМЫ: ПОИСКОВИК КАК ПРОСТОЙ ТЕКСТОВЫЙ БРОКЕР

- Поисковики берут текстовую строку и возвращают её вхождения в тексты
- Поисковики не знают темы запроса и смысла запроса
- Поисковики не знают типа и темы возвращаемых документов
- Поисковики показывают не свои титулы и аннотации, а только то, что есть на сайте



# ПРОБЛЕМЫ: ПОЛЬЗОВАТЕЛЬ НЕДОВОЛЕН

- Высокое напряжение **борьбы** с вебмастерами, дорвейные войны
- Резкое **падение** качества результатов и замусоривание Интернета в целом
- **Падение полноты** (разнообразия) – даже при релевантной выдаче показ только одной коммерческой категории сайтов
- **Рост недовольства** пользователей, падение их лояльности одной поисковой машине



# ПРОБЛЕМА: СОВРЕМЕННЫЙ ИНТЕРФЕЙС ПОИСКОВИКА

Результат работы «яйцеголовых»:

- **Бесконечная лента:** аналог списка литературы в научной статье
- **Нечитаемость:** невразумительный заголовок, аннотация, ненужные дата, URL, размер,
- **Каша:** разные типы информации в одной ленте

Мучения простого пользователя:

- **Шарада:** о содержании сайта нужно догадываться по URL и нечитаемой аннотации
- **Метод проб и ошибок:** перебор ссылок вслепую
- **Программирование:** подбор слов и операторов



# ЧТО РЕАЛЬНО ВЛИЯЕТ

- Скорость отклика.
- Стабильная работа.
- Релевантность.
- Свобода от спама.
- Простой и удобный интерфейс.
- Бурный маркетинг.





# КАК РАСКРУТИТЬ НОВЫЙ ПОИСКОВИК

- Создать броскую поисковую «фишку»
- Создать легенду вокруг неё.
- Сделать инновационный интерфейс, выгодно показывающий эту фишку.
- Навалиться на продвижение именно этой фишки и легенды (ТВ, офлайновая реклама, статьи).
- Обеспечить скорость отклика и бесперебойную работу.



# КУДА ЖЕ ДЕЛИСЬ ПОИСКОВЫЕ СТАРТАПЫ?

- Два-три года назад каждый месяц появлялись поисковые стартапы.
- «Под поиск» охотно давали деньги.
- Сейчас про те поисковые стартапы почти ничего не слышно.
- Новые поисковые стартапы пока не появляются.
- Практически все инновации возникают внутри гигантов
- Венчурные деньги – в социальных сетях
- Более того, аутсайдеры тонут (Рамблер).



# ПРИЧИНЫ ПРОВАЛА ПОИСКОВЫХ СТАРТАПОВ

- Высокий порог входа:
  - в 1997 – 1 сервер, 2 разработчика
  - В 2000 -20 серверов, 15 разработчиков
  - В 2009 – сотни серверов, сотни разработчиков
  
- Кризис.
- Отсутствие кадров.
- Большая мобильность пользователей.
- Венчурные деньги – в социальных сетях.



**Сейчас большинство инноваций будет возникать в уже существующих поисковиках.**



# ЭТАПЫ РАБОТЫ ПОИСКОВИКА

1. Выбор сайтов для обхода
2. Скачивание и индексация
3. Получение запроса от пользователя
4. Разбор запроса
5. Вычисление запроса (собственно поиск)
6. Показ результатов поиска

**Улучшать поиск можно на каждом из ЭТИХ ЭТАПОВ**



## ВЫБОР САЙТОВ/КОНТЕНТА ДЛЯ ИНДЕКСАЦИИ

- Выбор тематической **сферы** – специализированные поисковики (товары, автомобили, пр.).
- Выбор **типа сайтов** (поиск по блогам, картинкам, видео, энциклопедиям, словарям, файлам, пр.).
- **Новые виды контента:** социальные сети, мобильный контент, быстрый контент

**Специализированные поисковики  
могут быть прибыльны, но не  
заменяют универсальные поисковики**



## ВЫКАЧКА И ИНДЕКСАЦИЯ

- Распознавание типа данных на этапе выкачки (форумы, блоги, товарные предложения, статьи, новости, описания товаров).
- Распознавание темы страницы (семантическое индексирование)
- **Синтаксический и семантический разбор текстов, выделение объектов и фактов**
- **Разные индексы для разных типов сайтов.**
- **Индексация в реальном времени (Твиттер, блоги)**

**Здесь рулят большие универсальные  
ПОИСКОВИКИ**



## ПОЛУЧЕНИЕ ЗАПРОСА ОТ ПОЛЬЗОВАТЕЛЯ

- Подсказка и склейка вариантов запросов
- Последовательное уточнение запросов
- Персонализация, запоминание истории запросов пользователя
- Запрос на естественном языке

**Ведутся активные работы в больших поисковиках, в некоторых стартапах**





## ВЫЧИСЛЕНИЕ ЗАПРОСА ПО ИНДЕКСУ

- Повышение релевантности (тематическое разнообразие, навигационный поиск и прочее)
- Учёт прошлых интересов пользователя
- Учёт поведения пользователей на поисковике
- Использование фактов и объектов
- Использование структуры сайта
- Вычисление ответа как логический вывод

**Активно ведутся работы в больших поисковиках, стартапам труднее – нет базы текстов и логов запросов**



## ПОКАЗ РЕЗУЛЬТАТОВ ПОЛЬЗОВАТЕЛЮ

- Подсказка и склейка вариантов запросов
- Последовательное уточнение запросов
- Персонализация, запоминание истории запросов пользователя
- Запрос на **естественном языке**



# ТЕХНОЛОГИИ БЛИЖАЙШЕГО БУДУЩЕГО

- Автоматическое обучение по множеству факторов
- Автоматическое подавление обвязки и неважных элементов
- Анализ сайта в целом
- Учёт кликабельности ссылок и другого пользовательского поведения.
- Понижение веса ссылок, повышение веса контента.
- Обнаружение новых спам-приёмов: синонимизаторов, генерации контента.
- Индексация в реальном времени
- Извлечение фактов со страницы (цены, адреса и пр.).
- Голосовые интерфейсы



# ПОИСКОВАЯ ВЫДАЧА БЛИЖАЙШЕГО БУДУЩЕГО

- **Структурированная выдача:** по теме и типам документов, воспроизведение структуры сайта
- **Энциклопедическая выдача:** названия, аннотации, дополнительная информация, тэги, картинки, карты
- **Понимание запроса:** запросы на ЕЯ, распознавание темы запроса, выявление скрытых прагматик (интенций), уточнение запроса
- **Подсказка:** исправление запроса, предложение вариантов, склейка вариантов написания, история запросов
- **Дайджесты** и сложные структуры вместо плоского списка



# ПОИСКОВЫЕ ТЕХНОЛОГИИ БОЛЕЕ ОТДАЛЁННОГО БУДУЩЕГО

- **Вопросо-ответные машины:** не список сайтов, а исчерпывающий ответ на вопрос пользователя.
- **Распределённый поиск:** поисковая машина в формате торрентов
- **Поиск картинок и видео по смыслу:** семантическая индексация изображений и звуков.
- **Модель пользователя:** истинная персонализация, социальная кластеризация, психотипы пользователей.



## У ЧЕГО НЕТ БУДУЩЕГО

- «Библиографическая лента»
- Дальнейшие усилия по повышению традиционной «релевантности»
- Отношение к вебмастерам, как мёртвому материалу
- Ссылочное ранжирование в его текущем виде
- Автоматическая кластеризация результатов по темам
- Персонализация как настройка пользователем



# БУДУЩЕЕ ПОИСКОВОЙ ОПТИМИЗАЦИИ (предположения)

- **Риски:** непрерывные апдейты, рост цен на продвижение, непредсказуемость бизнеса
- **Консолидация:** разорение маленьких оптимизаторских контор, превращение крупняка в агентства полного цикла цифрового продвижения (digital agencies)
- **Смена парадигмы:** закрытие бирж ссылок через 2-3 года
- **Автоматизация:** системы автоматического продвижения и управления продвижением
- **Полный цикл:** консалтинг, управление контекстной рекламой и копирайтинг как основное содержание услуги SEO
- Симбиоз с поисковиками



## БУДУЩЕЕ КОНТЕКСТНОЙ РЕКЛАМЫ (предположения)

- **Высокотехнологичные платформы** для управления КР (автоподбор слов, управление ставками, длинным хвостом запросов, анализ эффективности)
- **Интеграция:** сращивание с системами управления предприятием (бухгалтерия, склад, CRM)
- **Широкий охват каналов:** включение покупки трафика, сетей аффилиатов (партнёрок) и SMO в общий пакет продвижения, симбиоз с оптимизацией
- **Смена ключевого показателя:** поставка лидов, регистраций, звонков (CPA)
- **Консолидация:** разорение маленьких контор, превращение крупняка в агентства полного цикла



# СПАСИБО!

**Игорь Ашманов**

Генеральный директор

[igor@ashmanov.com](mailto:igor@ashmanov.com)

Информация о компании,  
услугах и технологиях

[www.ashmanov.com](http://www.ashmanov.com)

Анализаторы

ПОИСКОВИКОВ

[www.analyzethis.ru](http://www.analyzethis.ru)

\*



Ашманов  
и партнеры