

# Яндекс

## Решение типовых задач с помощью сервиса Яндекс.Вебмастер

**М**ихаил Сенин

**А**лександр Шубин

Команда веб-поиска

Харьков, 14.10.2011

# Взаимоотношения

# Что нужно сайту от Яндекса? Только трафик?



# А что нужно Яндексу? Качественно отвечать на вопросы

- Нужно, чтобы сайт индексировался без ошибок
- Нужно, отличать важное от второстепенного
- Нужно понимать, что проиндексировали

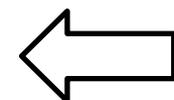
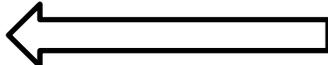
**Яндексу нужны понятные качественные сайты, которые легко индексируются!**

# Как это работает?

Сайт

База  
робота

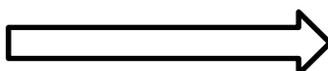
Загрузка  
роботом



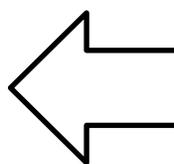
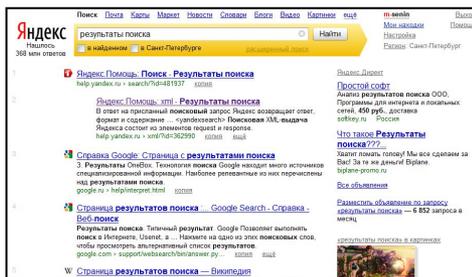
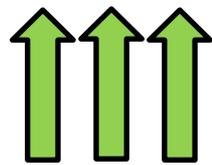
Построение  
индекса



Поисковый  
индекс



Ранжирование



Яндекс



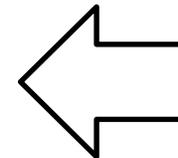
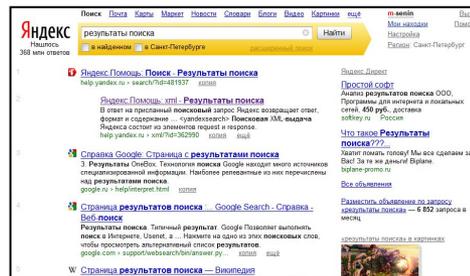
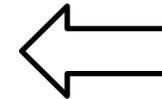
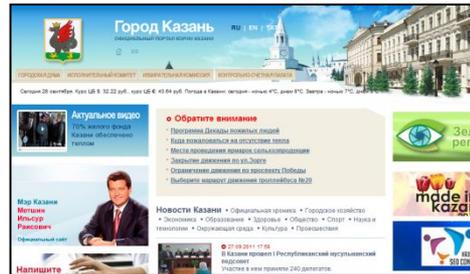
# Знакомимся

Я

# Новый сайт? Расскажите!

Сайт

База  
робота



Яндекс



# Добавление сайта Какой же тут подзаголовок? :)

## — Добавить сайт

<http://webmaster.yandex.ru/site/add.xml>

## — Проверка ответа сервера

<http://webmaster.yandex.ru/server-response.xml>

### Проверка ответа сервера

Проблемы с индексированием сайта могут быть связаны с тем, что ваш веб-сервер отвечает [роботам Яндекса](#) не так, как обычным пользователям. Чтобы увидеть, как веб-сервер отвечает роботу Яндекса, укажите URL и нажмите «Проверить».

- i** Поскольку проверка ответа сервера выполняется без учёта [директивы Crawl-Delay](#), инструмент доступен только для [подтверждённых сайтов](#).
- i** Роботы Яндекса делают запросы с различных ip-адресов. Данный инструмент может не выявить проблему, если ваш веб-сервер отвечает по-разному в зависимости от ip-адреса машины, выполняющей запрос.

<input type="text" value="http://webmaster.yandex.ru/"/>	Основной робот Яндекса ▾	<input type="button" value="Проверить"/>
URL	User-Agent: Mozilla/5.0 (compatible; YandexBot/3.0; +http://yandex.com/bots)	
<input type="checkbox"/> добавить заголовок If-Modified-Since <input type="text"/>		

**Код статуса HTTP:** "200 ОК"  
**Время ответа сервера:** 72мс  
**IP сайта:** 87.250.250.243

#### Заголовки:

Date: Thu, 13 Oct 2011 22:07:05 GMT  
Transfer-Encoding: chunked  
Expires: Thu, 13 Oct 2011 22:07:05 GMT  
Set-Cookie: yandexuid=3538414791318543625; domain=.yandex.ru; path=/; expires=Tue, 19 Jan 2038 03:14:07 GMT  
Content-Type: text/html; charset=UTF-8  
Server: lighttpd/1.4.26  
Pragma: no-cache  
Cache-Control: max-age=0, proxy-revalidate

**Кодировка:** UTF-8

**Размер страницы:** 15.79Кб

# Подтверждение прав Сложно, но можно

С подтверждением справляются не все:

70% – успешно подтвердили

13% – не смогли подтвердить

17% – не пытались, добавили чужой сайт

Сайт [www.yandex.ru](http://www.yandex.ru) добавили 4 тыс. раз. Ну вот зачем?!

Новые способы подтверждения прав:

— Делегирование

— HTML-файл

— Email в записи WHOIS

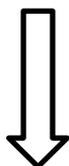
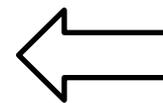
# Робот мешает?

База  
робота

Загрузка  
роботом

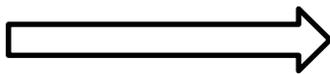


Сайт

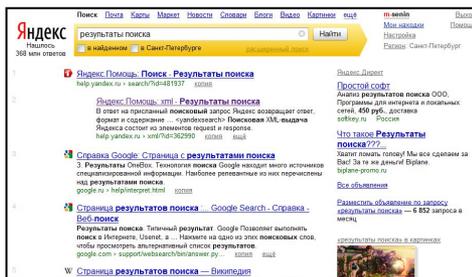


Построение  
индекса

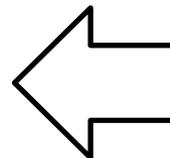
Поисковый  
индекс



Ранжирование



Яндекс



# Нагрузка со стороны робота

## Ситуация редкая, но всё же

### — Выставляем Crawl-delay

```
User-agent: Yandex  
Disallow: /catalog/  
Crawl-delay: 2,5 # таймаут в 2,5 сек
```

### — СМОТРИМ НА ПОРЯДОК ДИРЕКТИВ

```
User-agent: Yandex  
Disallow: *.doc  
Allow: /catalog/
```

/catalog/a.doc

Раньше был запрещён,  
будет разрешен!

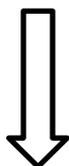
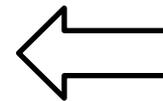
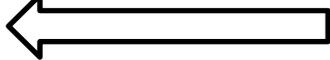
# Устраняем неполадки

# Уменьшился трафик

Сайт

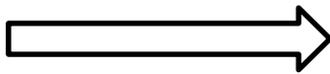
База  
робота

Загрузка  
роботом

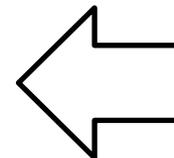
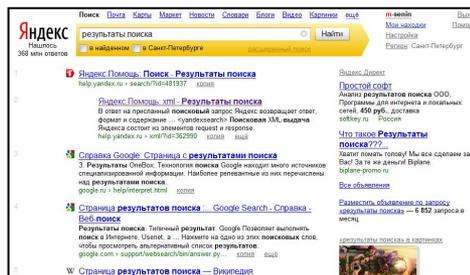


Построение  
индекса

Поисковый  
индекс



Ранжирование



Яндекс



# Пропал трафик Вирус?

# [Очень опасный сайт - трз скачать бесплатно без регистрации...](#)

[Сайт может угрожать безопасности вашего компьютера](#)

Посещение этого сайта может привести к заражению компьютера вредоносными программами, использованию его без вашего ведома, а также к порче или краже ваших данных. [Почему?](#)

✔ [Посмотреть сохранённую безопасную копию](#)

Это не угрожает вашему компьютеру и данным

[Всё равно перейти на эту страницу](#)

Переход по ссылке может нанести вред вашему компьютеру

[оченьопасныйсайт.рф](#) [безопасная копия](#) [еще с сайта](#)

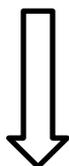
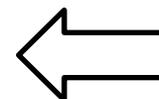
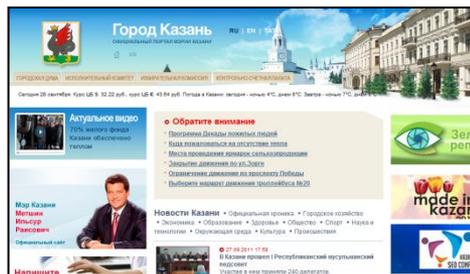
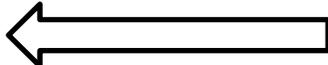
- Узнайте вовремя: Сообщения, Виджет, Яндекс. Бар
- Описание вердиктов в Я.Вебмастере
- Следите за обновлениями CMS
- Можно пожаловаться на вирусы и фишинг

# Уменьшился трафик

Сайт

База  
робота

Загрузка  
роботом

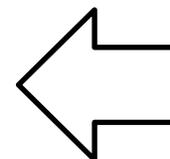
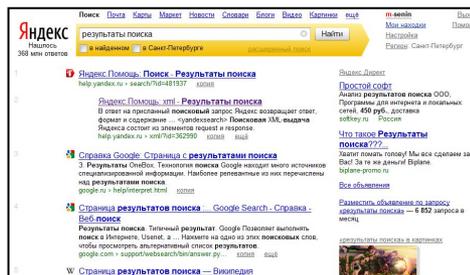


Построение  
индекса

Поисковый  
индекс



Ранжирование



Яндекс

# Уменьшился трафик

## Ранжирование

Конкуренция?

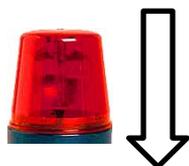
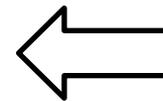
Доигрались?

- Накрутки
- Скрытый текст
- Переоптимизированные тексты
- Взрослый контент
- И То-О-Чём-Нельзя-Говорить

# Загрузили, но не то?

База  
робота

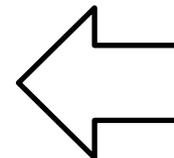
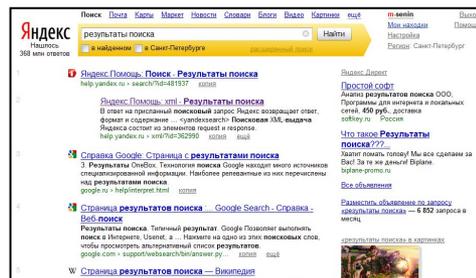
Загрузка  
роботом



Построение  
индекса

Поисковый  
индекс

Ранжирование



Яндекс

# Уменьшился трафик

## Построение индекса

- Смотрим на структуру сайта с точки зрения робота

пусть робот обходит то, что нужно

- Сколько url соответствуют одному документу?

у сайта есть квота, не тратьте её на дубликаты!

Я

- Насколько уникален сайт?

# Вообще не загрузили?

Сайт

База  
робота

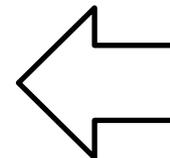
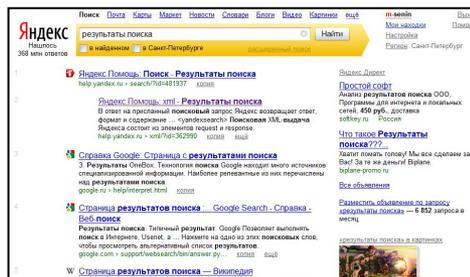
Загрузка  
роботом



Построение  
индекса

Поисковый  
индекс

Ранжирование



Яндекс



# Упал трафик Индексирование?

- Проблемы с хостингом
- Проблемы с DNS
- Неверный robots.txt или файл Sitemap
- Неверные ссылки

**Проверяем ответ сервера**

**Изучаем исключенные страницы**

**Улучшаем**

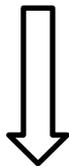
Я

# Где это можно сделать?

Сайт

База  
робота

Загрузка  
роботом

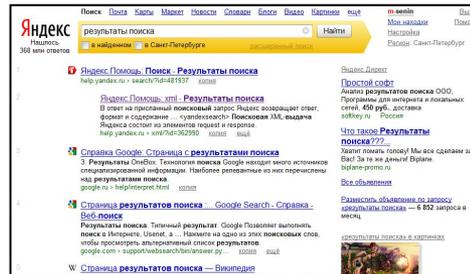
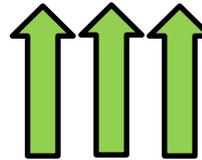
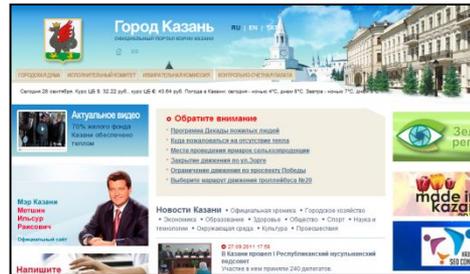


Построение  
индекса

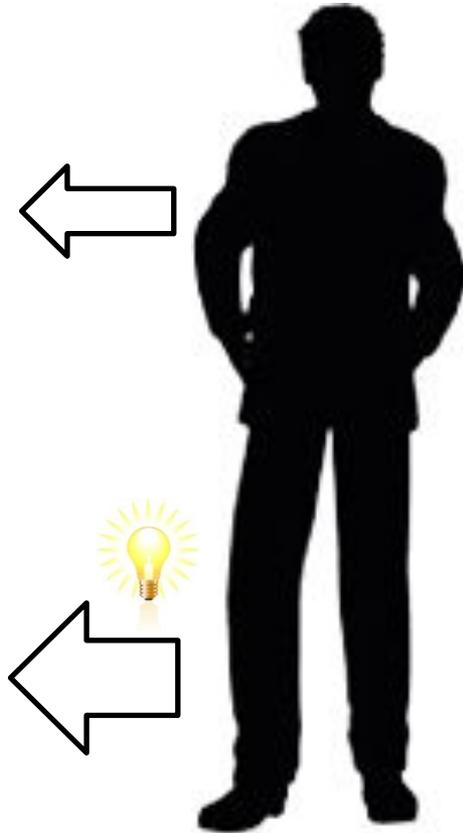
Поисковый  
индекс



Ранжирование



Яндекс



# ... ИНДЕКСАЦИЮ

## Robots.txt и Sitemap

### Robots.txt

- Clean-param: только нужные страницы
- Host: указываем главное зеркало
- Sitemap: урл файла Sitemap

### Sitemap

- Приоритет страницы

```
User-agent: Yandex  
Disallow:  
Clean-param: ref /get_book.pl  
Host: host.com.ua  
Sitemap: host.com.ua/sitemap.xml
```

# ... ранжирование в своём регионе И дополнительный источник трафика

- Инструмент “Регион сайта”  
указываем регион клиента
- Организация в Я.Справочнике  
([sprav.yandex.ru](http://sprav.yandex.ru))  
указываем регион организации

# ... нахождение на картах

## Фильтр по признакам

аптека доставка лекарств × Найти

круглосуточно  доставка лекарств  гомеопатическая  социальная ещё

ресторан круглосуточно × немецкая × Найти

круглосуточно кухня  летняя веранда  wi-fi ещё

# ... представление в Поиске И не только

- Микроформаты, schema.org, opengraph  
строим семантический веб вместе
- Партнёрские программы  
75 рубрик адресов, недвижимость, авто, ...



## [Отзывы владельцев Volvo \(Вольво\)](#)

11 мая 2011 81 **отзыв**

[Volvo XC90 2003 - отзыв владельца](#). Мне японские машины в обслуживании дороже обходились:  
Ниссан Скайлайн 35 кузов 01г.в- за год 6 расходомеров цена 4500 руб. с установкой...

[Volvo 850 1993 - отзыв владельца](#). Просто машина редко выставляется на продажу - когда купил  
понял почему: то двигатель не важно выглядит, то кузов коцанный, то салон убитый...

[reviews.drom.ru](http://reviews.drom.ru) > **Volvo** [копия](#) [ещё](#)

# ... представление в Поиске Немного статистики по **schema.org**

- За три месяца размечено 50 млн страниц
- Рунет на 2м месте, Уанет – на 21м
- Популярные форматы: товары, люди, адреса, отзывы
- Следите за анонсами...

# ... ранжирование авторских Как это работает ТЕКСТОВ

1. Сообщаем Яндексу свой текст
2. Яндекс проверяет наличие текста в интернете
3. Публикуем текст на сайте

...

4. Авторство за текстом закреплено!

# ... ранжирование авторских Немного статистики ТЕКСТОВ

- Всего ~57 000 текстов (63% оригинальные)
- В день загружается ~1 000 текстов
- На пользователя приходится от 1 до 40  
ТЕКСТОВ

# ... ранжирование авторских Некоторые темы ТЕКСТОВ

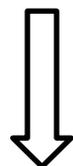
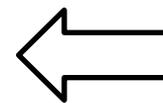
- Пластиковые окна, газогенераторы, тракторы, строительство срубов, прокладка труб
- Гадания, предсказания, гороскопы
- Порно-тексты
- Теория мирового заговора, масоны, конец света

# Хорошо сработано!

Сайт

База  
робота

Загрузка  
роботом

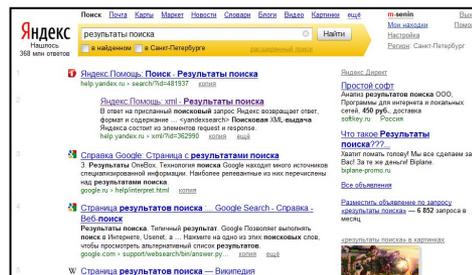
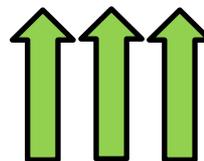


Построение  
индекса

Поисковый  
индекс



Ранжирование



Яндекс



# ... свою информированность :) Подписываемся на сообщения

- Про атаки и вирусы

XSS атака, вирусы на сайте, скрытый текст

- Про партнёрские программы

результаты обработки данных, сообщения от модераторов

- И многое другое...

# ... СВОЙ БИЗНЕС Все инструменты в одном

Яндекс

[Вебмастер](#) [Метрика](#) [Виджеты](#) [Рекламная сеть](#) [Директ](#) [Поиск для сайта](#) [API](#) [ещё](#)



[Войти](#)



[Понятная схема проезда](#)

Помогите клиентам не потеряться по пути к вам

Яндекс для бизнеса — это готовые решения, которые помогут развивать ваше дело и зарабатывать с помощью наших технологий.

Если у вас возникли какие-либо вопросы, пожалуйста, задавайте

**Привлекайте аудиторию**

Среди 20 миллионов пользователей Яндекса найдутся ваши клиенты.

**Улучшайте сайт**

Сайт станет привлекательнее без лишних расходов.

**Следите за результатом**

Узнайте, насколько эффективны ваши действия в интернете.

[Партнерские программы](#)

[Семинары и вебинары](#)

[Подписка на анонсы мероприятий](#)

Персональные наборы



[Компаниям без сайта](#)

Привлекайте аудиторию из интернета, даже если у вас нет сайта.



[Сайтам-визиткам](#)

Продолжайте развивать свой бизнес в интернете с помощью Яндекса.



[Интернет-магазинам](#)

Повышайте продажи и привлекайте новых покупателей.



[Тематическим и городским порталам](#)

Улучшайте сервисы и создавайте новые, увеличивайте посещаемость.

© 2011 ООО «Яндекс»

Я

<http://business.yandex.ru/>



**М**ихаил Сенин

**А**лександр Шубин

Команда веб-поиска

119021, Россия, Москва,  
ул. Льва Толстого, 16

+7 (495) 739-00-00

+7 (495) 739-70-70 — факс

[senin@yandex-team.ru](mailto:senin@yandex-team.ru)

[ajax@yandex-team.ru](mailto:ajax@yandex-team.ru)