



**Тенденции развития  
ПОИСКОВЫХ СИСТЕМ**

Александр Садовский

# Яндекс: что нового в 2005

## Робот

- быстрый (8-10 раз в сутки, 20 млн. URL)
- кластеризация основного (скорость в 1,5 раза больше)

## Поиск

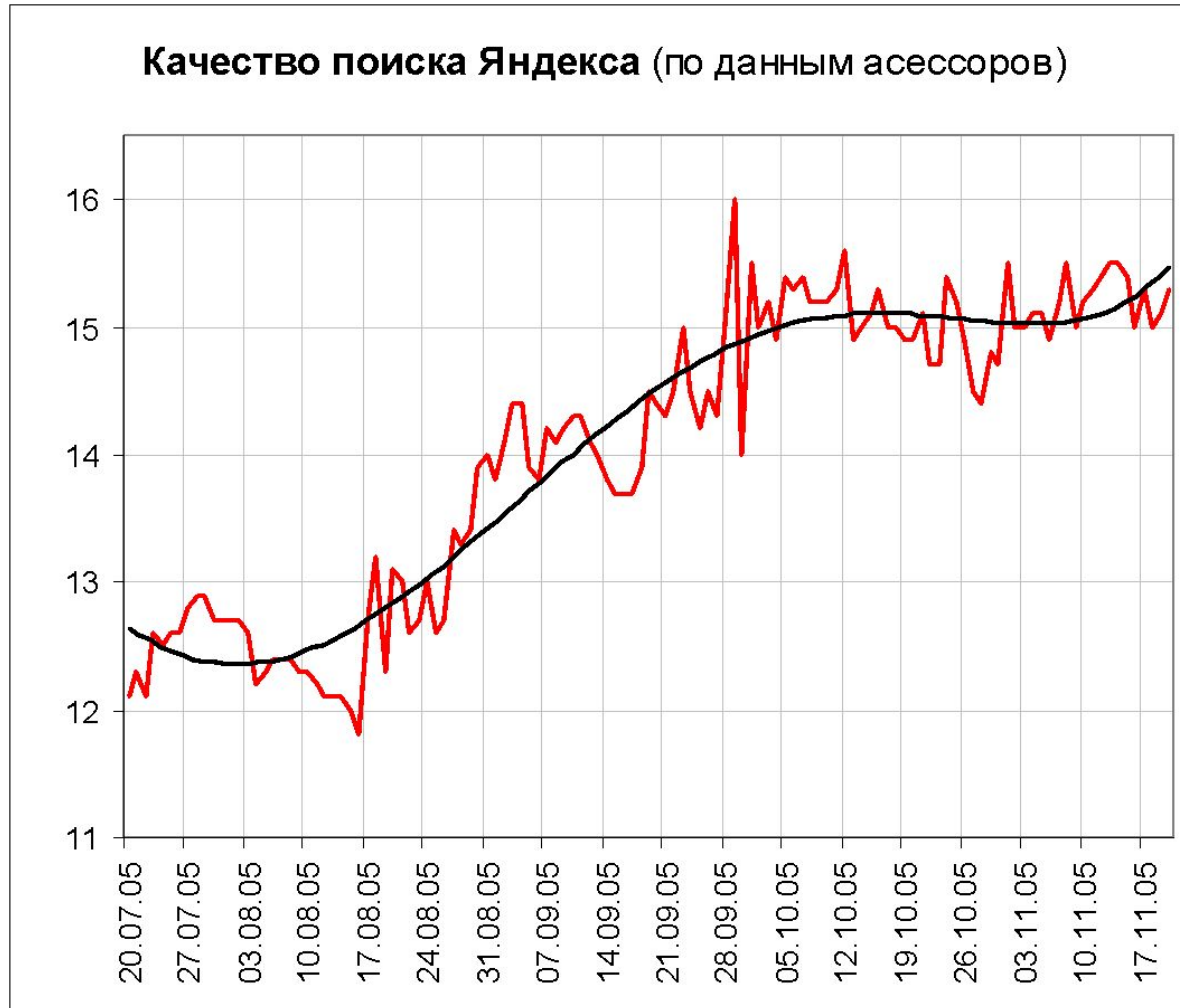
- ранжирование коротких запросов
- колдовство по опечаткам
- yandex.ua (морфология, регион сайта)

## Антиспам

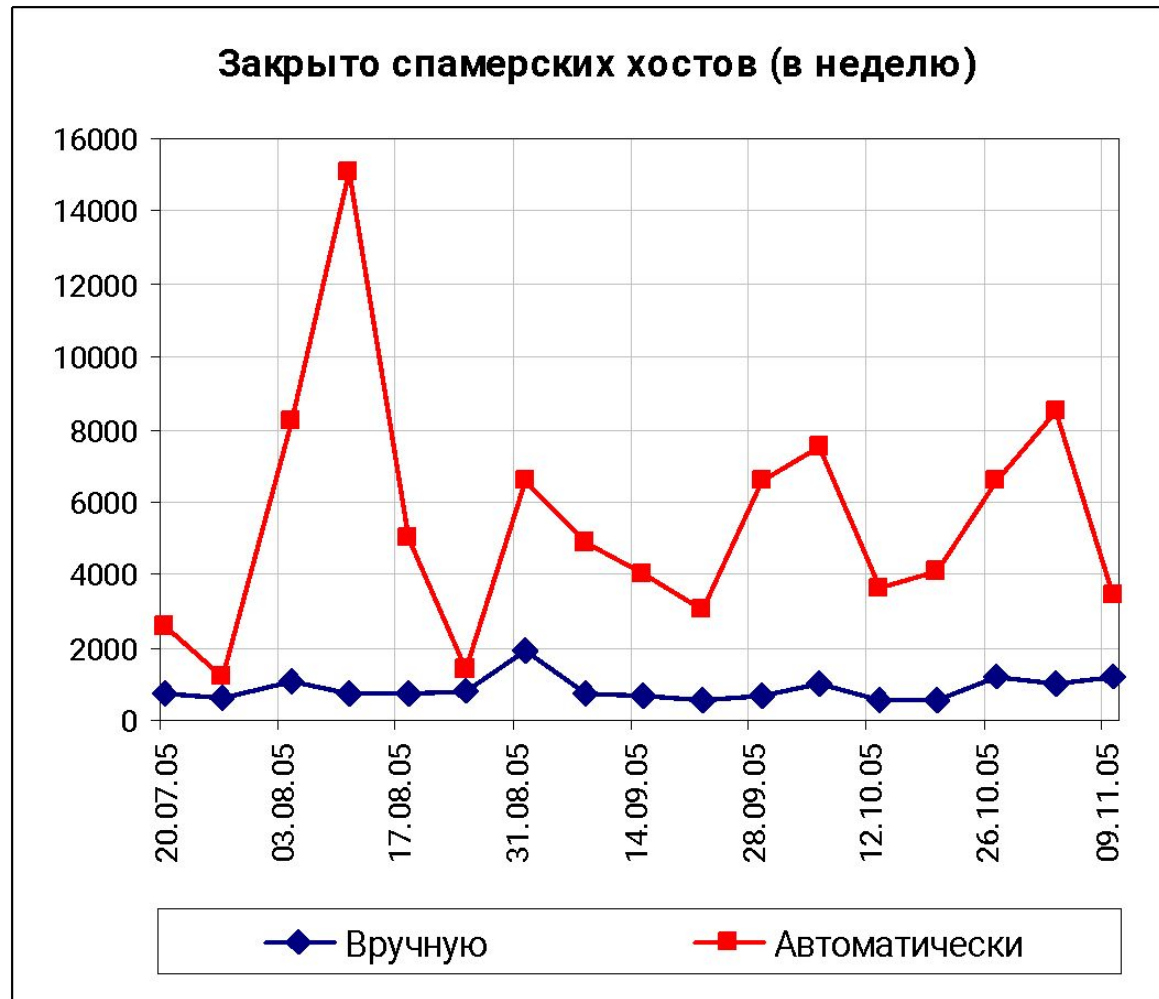
- внедрен автоматический отлов



# Яндекс: что нового в 2005



# Яндекс: что нового в 2005



# Поиск революции в поиске

## Принципы Web 2.0 и веб-поиск

- пользователи «улучшают» сервис:  
исправление опечаток на основе запросов;  
ИЦ; ссылки из блогов
- ценность уникальных данных: архивы  
блогов, словари, архивы запросов
- «длинный хвост»: поиск, а не только  
навигация; реклама по ключевым словам
- открытые API: Яндекс.XML, RSS  
поисковых результатов в Новостях, Блогах
- есть развитие, но нет версий: обратная  
связь от веб-мастеров и пользователей

# Поиск революции в поиске

## Борьба с мусором

- дубликаты и клоны: шинглы, лексические сигнатуры, сниппеты
- технический мусор: высокая изменчивость текстов и URL

## Антиспам

- статистический анализ текста
- анализ распределения ссылок
- обнаружение колец
- TrustRank

# Поиск революции в поиске

## Извлечение фактов: первый шаг

– «Новости в лицах»

The screenshot shows the Yandex search engine interface. At the top left is the Yandex logo with the tagline 'Найдётся всё'. A search bar is empty, and a 'Найти' button is to its right. Below the search bar is a navigation menu with tabs for 'Везде', 'Каталог', 'Новости', 'Маркет', 'Словари', 'Картинки', and 'Все службы...'. The 'Новости' tab is selected. Below the navigation menu, there is a status message: 'автоматически обработано 1024 источника, обновлено 22.11.2005 10:44 мск'. On the left side, there is a vertical menu with categories: 'Главные новости', 'Политика', 'В мире', 'Общество', 'Экономика', 'Спорт', and 'Происшествия'. The main content area is titled 'Культура' and includes the text 'выпуск: Россия | Украина'. Below this, there is a news snippet about Maya Plisetskaya: 'Майя Плисецкая: Путин позвонил мне в машину, когда я ехала в театр! (109)'. The snippet text reads: 'Прославленная балерина подошла к оркестровой яме, к дирижеру: ну что, повторим еще раз? И с юбилеем, и с наградой - орденом "За заслуги перед Отечеством" I степени.' Below the snippet, it says: '"Гарри Поттер и Кубок огня" выйдет в российский прокат через месяц (55)'. On the right side, there is a yellow box with a red border containing a list of 'Культура в лицах' with counts: 'Майя Плисецкая 96', 'Гарри Поттер 68', 'Владимир Путин 56', 'Джоан Роулинг 29', and 'Андрей Малахов 26'. At the bottom center, there is a large red letter 'Я'.

Yandex  
Найдётся всё

Найти

Расширенный поиск +

Везде Каталог **Новости** Маркет Словари Картинки Все службы...

автоматически обработано 1024 источника, обновлено 22.11.2005 10:44 мск

Главные новости  
Политика  
В мире  
Общество  
Экономика  
Спорт  
Происшествия

Культура

выпуск: Россия | Украина

Майя Плисецкая: Путин позвонил мне в машину, когда я ехала в театр! (109)

Прославленная балерина подошла к оркестровой яме, к дирижеру: ну что, повторим еще раз?  
И с юбилеем, и с наградой - орденом "За заслуги перед Отечеством" I степени.

"Гарри Поттер и Кубок огня" выйдет в российский прокат через месяц (55)

**Культура в лицах**  
Майя Плисецкая 96  
Гарри Поттер 68  
Владимир Путин 56  
Джоан Роулинг 29  
Андрей Малахов 26

Я

# Яндекс: поиск революции в поиске

## Стимуляция исследований: научные стипендии

- 252 заявки, 34 стипендии
- сборник «Интернет-математика-2005»



**Яндекс®**

Найдется все.

Я