Яндекс

От запроса до ответа

Тенденции в представлении страницы результатов поиска

Александр Садовский

11 ноября 2008

«Ты помнишь, с чего начиналось?»

- У кого *индекс* больше и свежее?
  1997-2005
- У кого *спама* меньше?
  1999-2000, 2004, 2006…
- Поиски *параллельные,* ортогональные, универсальные
  2001…
- Максимум *релевантности* в десятке
  всегда
- *Разнообразие* результатов, минимум дублей и «клонов»
  2002, 2008…
- Соответствие *прагматик* запроса и выдачи
  2007…

Что же дальше?

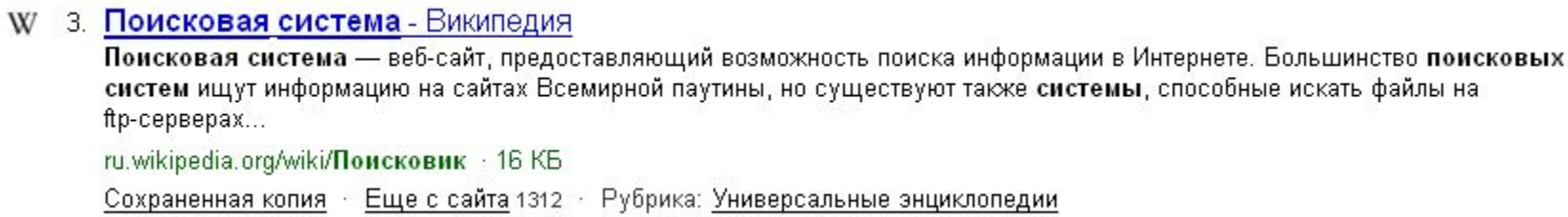
Я

Обогащение сниппетов

Я

Представление результатов: что было

- *До десяти результатов* поиска
- *Заголовок и сниппет* для каждого результата
- *«Сохраненная копия»* с подсветкой слов запроса, *«Еще с сайта»*
- Размер документа, формат и тема сайта (рубрика каталога)



Незаданные вопросы о сайте: новост

Стало

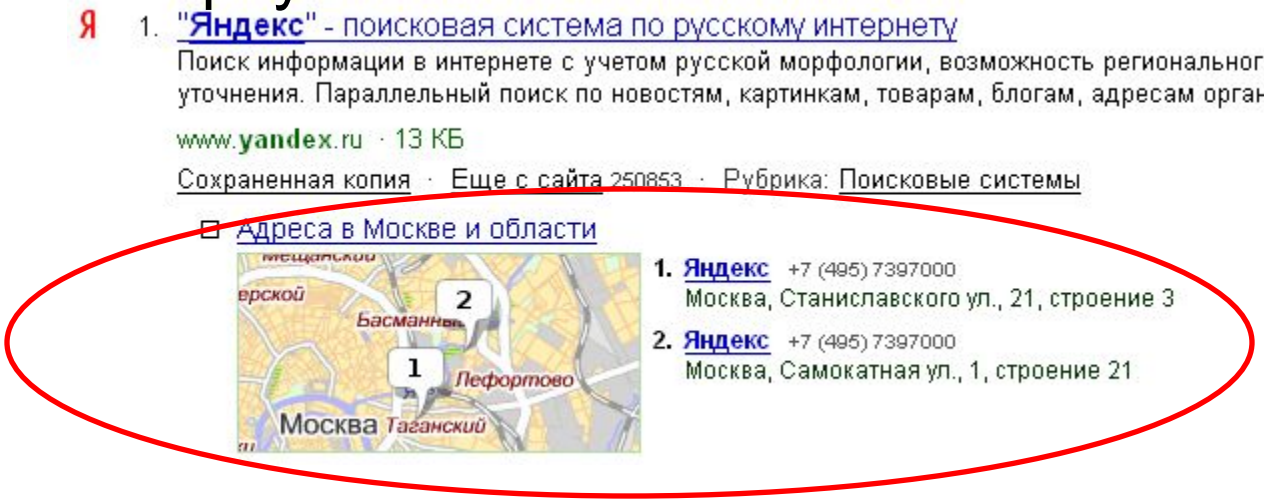
1. **ВГИК** - Главная
ВГИК не только учебный, но и научный центр. На кафедрах ведется научно-исследовательская работа. В аспирантуре **ВГИКа** готовятся специалисты для научной и педагогической деятельности.
www.vgik-edu.ru · 18 КБ
Сохраненная копия · Еще с сайта 640 · Рубрика: Высшее образование

Новости
12:04

В Москве открывается Международный фестиваль **ВГИК**
Сегодня в Москве открывается 28-ой Международный фестиваль **ВГИК**. В нем примут участие работы студентов из Аргентины, Великобритании, Германии, Южной Кореи, Швейцарии и других стран. Открытие пройдет в...
Все сообщения 10

Незаданные вопросы о сайте: адрес

Адрес развёрнут
свёрнут



Каждому случаю – свой сниппет

- Описание из

1. "Мобильные телесистемы (**МТС**)" - сотовый оператор
   Описание услуг в области сотовой связи. Тарифы. Схемы зон обслуживания (с детальными картами), сведения о роуминге. Технические данные мобильных телефонов. О компании.
   www.**mts**.ru · 24 КБ
   Сохраненная копия · Еще с сайта 32334 · Рубрика: Сотовые операторы
   ⊞ Адреса в Москве и области

- Начало текста

1. Lib.ru/Классика: **Толстой** Лев Николаевич. **Анна Каренина**
   Все счастливые семьи похожи друг на друга, каждая несчастливая семья несчастлива по-своему. Все смешалось в доме Облонских. Жена узнала, что муж был в связи с бывшею в их доме француженкою-гувернанткой, и объявила мужу, что не может жить с ним в одном доме.
   az.lib.ru/t/tolstoj_lew_nikolaewich/text_0080.shtml
   Сохраненная копия · Рубрика: Русская проза

- Найдено по ссылке

8. Welcome to Obama for America
   найден по ссылке: Сайт предвыборного штаба **Барака** Обамы... Предвыборная кампания **Барака** Обамы...
   www.barackobama.com · 1 КБ
   Сохраненная копия · Еще с сайта 4

Я

ЧПУ – человеко-понятные URLы

- Как было:

http://ru.wikipedia.org/wiki/%D0%9E%D0%B1%D1%84%D1%83%D1%81%D0%BA%D0%B0%D1%86%D0%B8%D1%8F

- Как стало:

1. **Обфускация** - Википедия
Обфускация (от лат. obfuscare, «затенять, затемнять») — запутывание кода программы, то есть приведение исходного текста или исполняемого кода к виду, сохраняющему функциональность программы, но затрудняющему анализ, понимание алгоритмов...
ru.wikipedia.org/wiki/**Обфускация**

Я

Где руда для обогащения сниппетов?

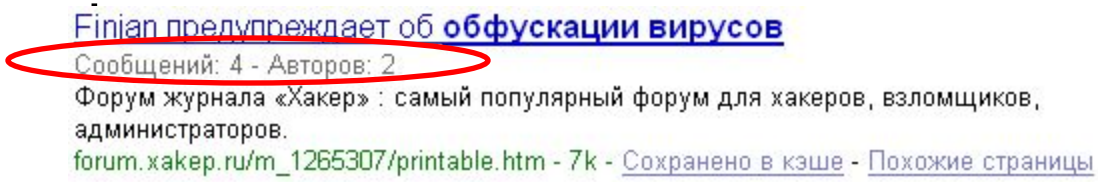
- ФИО, организации, должности
  уже выделяются в Новостях
- почтовые адреса, телефоны
  уже показываются за ссылкой адреса
- даты
- цены на товары
- типы ресурсов (магазин, энциклопедия и др.)
- ссылки на сюжеты новостей, блоги
- мультимедиа

и т.д.

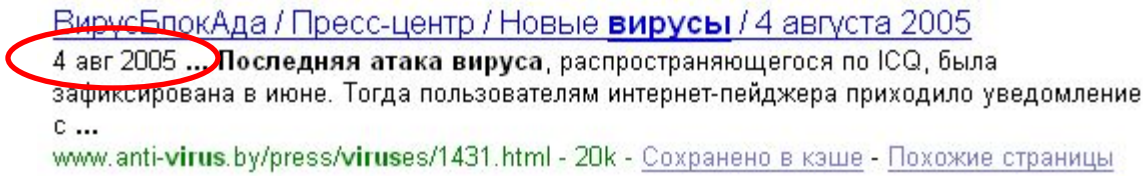
Я

Мы не одиноки во Вселенной!

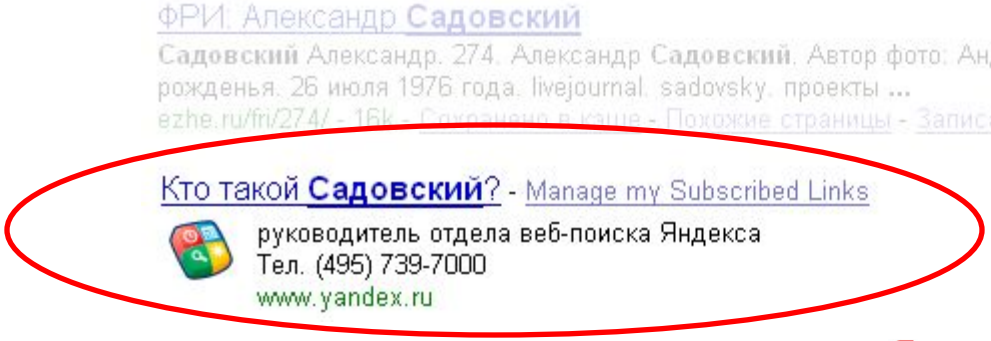
- Google: число сообщений в некоторых


Finjan предупреждает об **обфускации вирусов**
Сообщений: 4 - Авторов: 2
Форум журнала «Хакер» : самый популярный форум для хакеров, взломщиков, администраторов.
forum.xakep.ru/m_1265307/printable.htm - 7k - Сохранено в кэше - Похожие страницы

- …дата


ВирусБлокАда / Пресс-центр / Новые **вирусы** / 4 августа 2005
4 авг 2005 ... **Последняя атака вируса**, распространяющегося по ICQ, была зафиксирована в июне. Тогда пользователям интернет-пейджера приходило уведомление с ...
www.anti-**virus**.by/press/**virus**es/1431.html - 20k - Сохранено в кэше - Похожие страницы

- …информация третьих лиц (Subscribed Links)


ФРИ: Александр **Садовский**
Садовский Александр. 274. Александр **Садовский**. Автор фото: Андрей Себрант. день рожденья. 26 июля 1976 года. livejournal. sadovsky. проекты ...
ezhe.ru/fri/274/ - 16k - Сохранено в кэше - Похожие страницы - Записать

Кто такой **Садовский**? - Manage my Subscribed Links
руководитель отдела веб-поиска Яндекса
Тел. (495) 739-7000
www.yandex.ru

Я

Мы не одиноки во Вселенной-2!

- Yahoo: SearchMonkey – расширенные профили LinkedIn

Было:

Arkady Volozh - LinkedIn
View **Arkady Volozh's** professional profile on LinkedIn. **...** View **Arkady's** full profile: See who you and **Arkady Volozh** know in common **...**
www.**linkedin.com**/pub/0/10/818 - Cached

Стало:

Arkady Volozh
View **Arkady Volozh's** professional profile on LinkedIn. **...** View **Arkady's** full profile: See who you and **Arkady Volozh** know in common **...**
- Location: Russian Federation

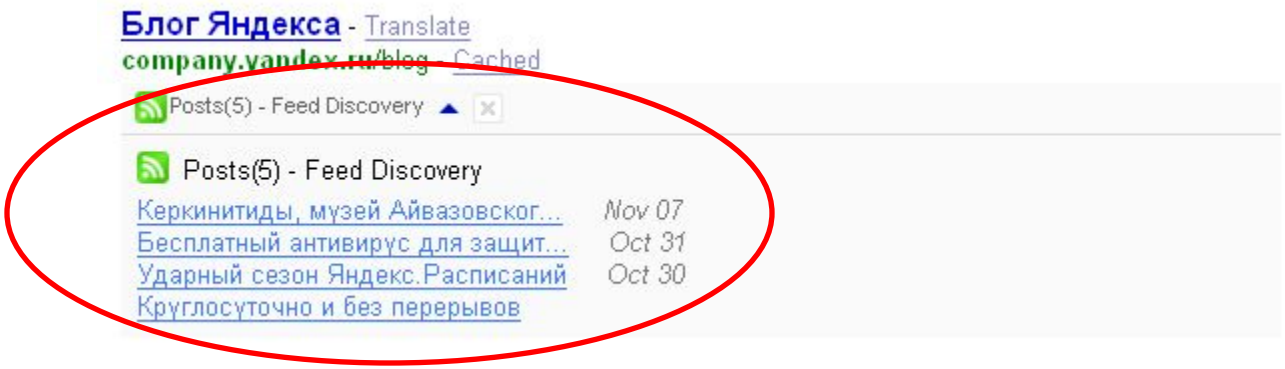
www.**linkedin.com**/pub/0/10/818 - Cached

Мы не одиноки во Вселенной-2.1!

- SearchMonkey – извлечение постов из фидов

Было:

Стало:



Перед тем, как двигаться дальше…

…вспомним, что было

Я

Качество поиска – это качество десяти

1. -------- --- ---------	Нерелевантный
2. ------ ------- --------- ------- ----- -- - -----	Спам
3. ---- -- ------ ------ ----- ----- ------ --------- ----------	Релевантный+
4. ------- -- -- --------	Нерелевантный
5. ----- --- ------ --- -- ----- --------- ----	Дубль

Плохой поиск

Я

1. -------- --- ---------	Полезный
2. ------ ------- --------- ------- ----- -- - -----	Релевантный+
3. ---- -- ------ ------ ----- ----- ------ --------- ----------	Полезный
4. ------- -- -- --------	Релевантный+
5. ----- --- ------ --- -- ----- --------- ----	Релевантный+

Хороший поиск

Парадигма перенаправления

- Поисковая система – умный *посредник*
  перенаправляет по запросу на подходящий URL
- Пользователь *наедине с собой*
  Сам выбирает запрос. Сам изучает выдачу, выбирая ответ. Сам ищет ответ на выбранной странице. «К пуговицам (релевантности) претензии есть?!»

2007-2008. Оказывается, **жизнь есть**

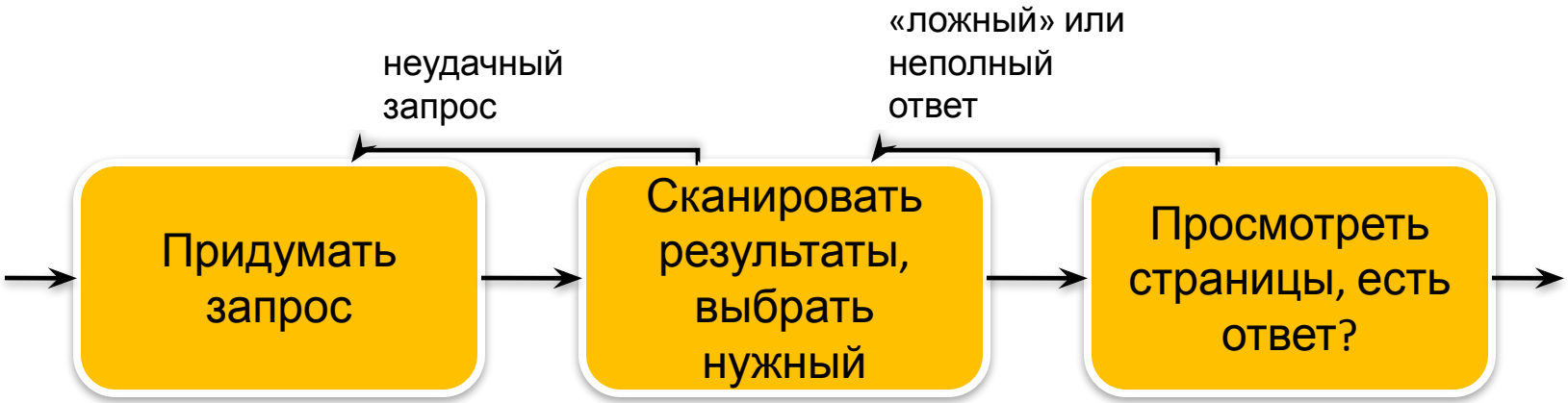
до запроса,

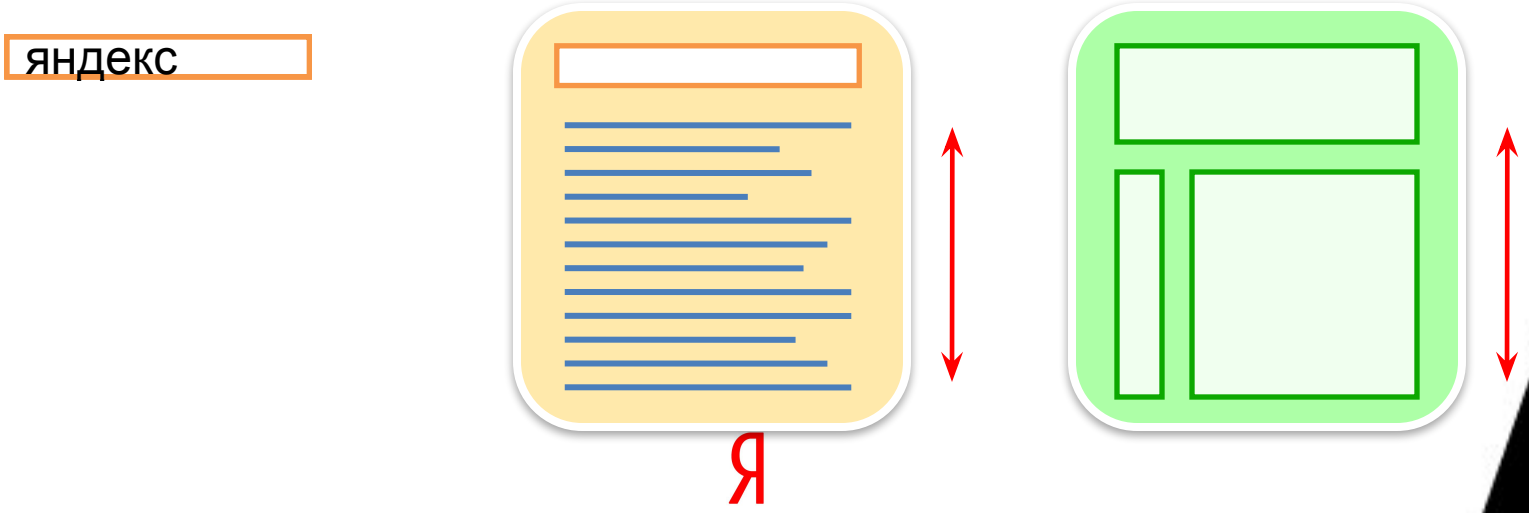
в момент выбора ответа на выдаче

и даже после клика.

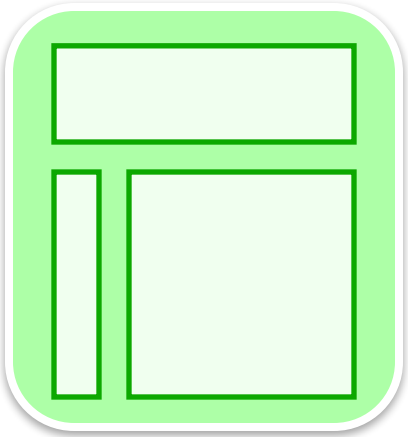
Я

На каждом шаге – потеря времени и си





Все может быть проще



Парадигма ответов на вопросы

- Давать ответы на *заданные и незаданные* вопросы
  перезадание запроса, адрес с картой и телефон по запросу с названием организации
- Структура выдачи позволяет уже сейчас показать многие типы ответов, *не меняя привычек* пользователей

Ответ – это не только факт, но и путь к нему:

- Помогать в *формулировании запроса*
- Показывать отличия между разными типами ответов
  например, блоги, новости и товары размечать по-разному
- Помогать *выбрать ссылку,* богаче аннотируя результаты
- Упрощать нахождение ответа *за пределами выдачи*
  предугадывать не полностью заданный вопрос, указывать ответ на странице

Еще раз:
заданные и
***незаданные* вопросы.**
И недозаданные.

Я

В чем разница?

Парадигма *перенаправления* концентрируется на «попадании» ответа в запрос.

Ощущения пользователя ее не волнуют.

Цель парадигмы *ответчика* – *снизить усилия пользователя* на пути к ответу.

Экономит клики (в том числе за счет повышения релевантности) и перемещения по странице, снижает требования к концентрации внимания.

<u>Сменились объекты изучения:</u>

- с информации → на *восприятие ее пользователем*
- со страницы выдачи → на *процесс удовлетворения информационной потребности* (от запроса до ответа)

Я

Колдунщики и фишки

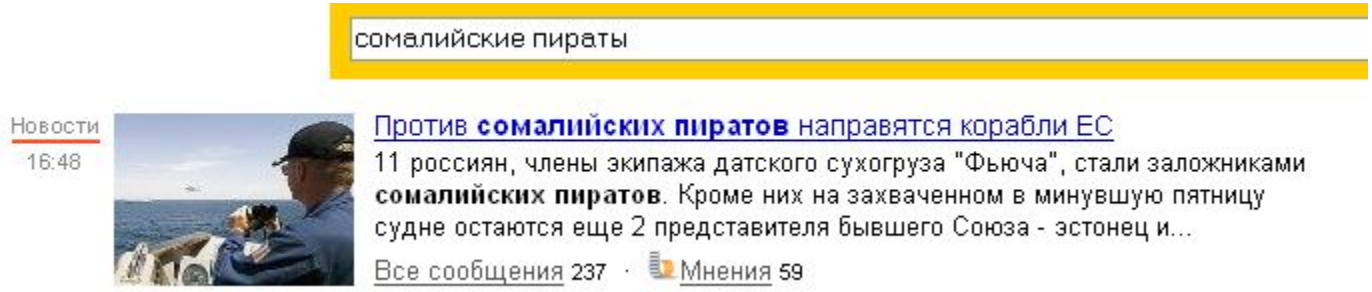
Я

Колдунщики: событийные

- Олимпиада-2008

Колдунщики: структурированные данные

- Новости



- Маркет

- Карты, Словари, Афиша и т.д.

Я

Колдунщики: ответы на вопросы

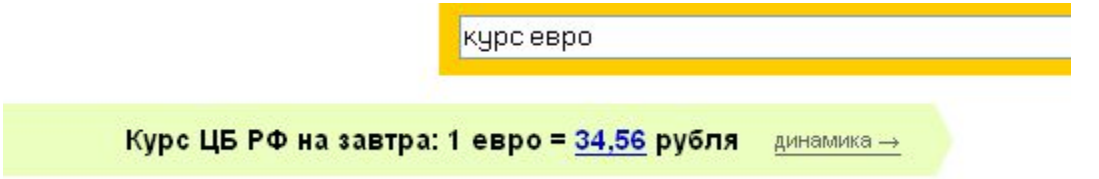
- Погода
  температура кодируется цветом

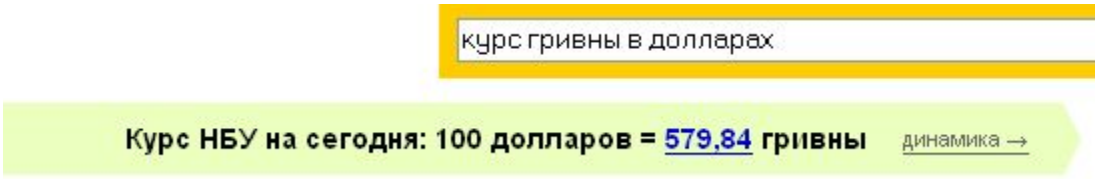
«Недозаданный» вопрос – регион определяем по IP

Я

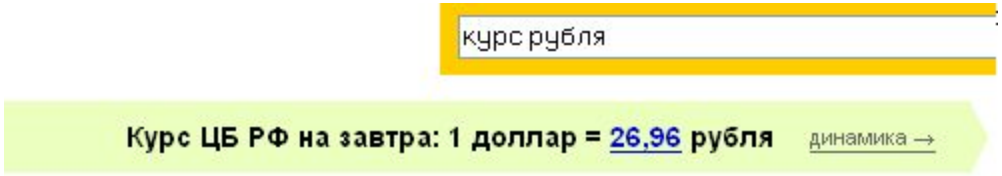
Колдунщики: ответы на вопросы

- Курсы валют





«Недозаданный» вопрос –
показываем курс к самой популярной валюте



Я

Фишки: не ждали?!

- Цвет
  а

Снова фишки

- Время

«Недозаданный» вопрос – смотрим регион по IP

- График отключения горячей воды, ваш IP, таблица Менделеева и др.

Я

Помогаем
в формулировании запроса

Бесплатно, 24x7

www.yandex.ru

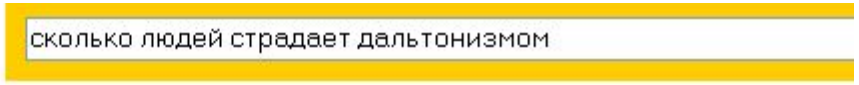
Я

Исправление ошибок

- Опечатки – 5% запросов предлагаем…

Быть может, вы искали: «чучхе»

…5% автоматически исправляем (незаданный вопрос)



По исходному запросу «сколько людей страдает вальтонизмом» ничего не нашлось.
Результаты поиска для «сколько людей страдает дальтонизмом».

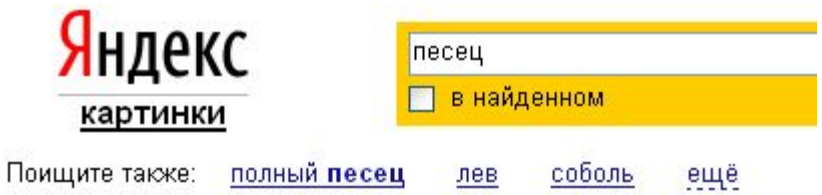
- Транслитерация, неправильная раскладка

Быть может, вы искали: «яндекс»

Я

Только где же это что-то?

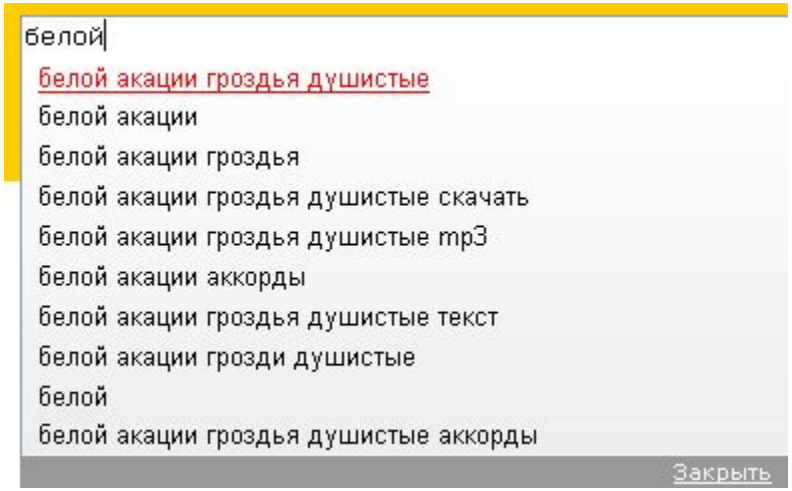
- запросы, связанные с вашим запросом



Я

Поисковые подсказки

- упрощают набор – раз

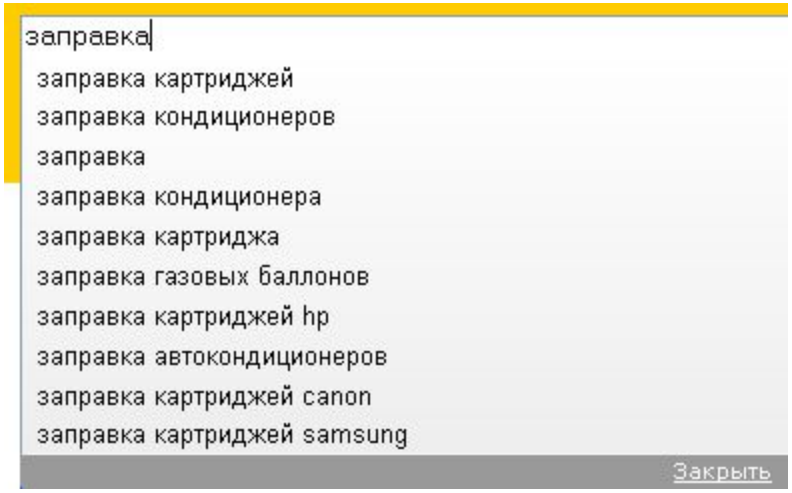


- исправляют ошибки – два

Я

Поисковые подсказки

- дают представление о спектре тем и прагматик – три

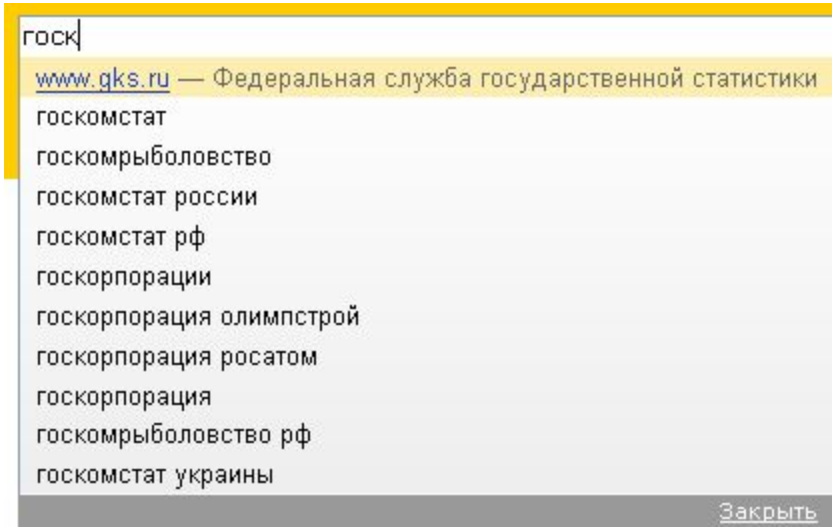


- «обучают» новичков, как задавать запрос – четыре

Я

Поиск до поиска!

- дают ответы с URL нужных сайтов – пять

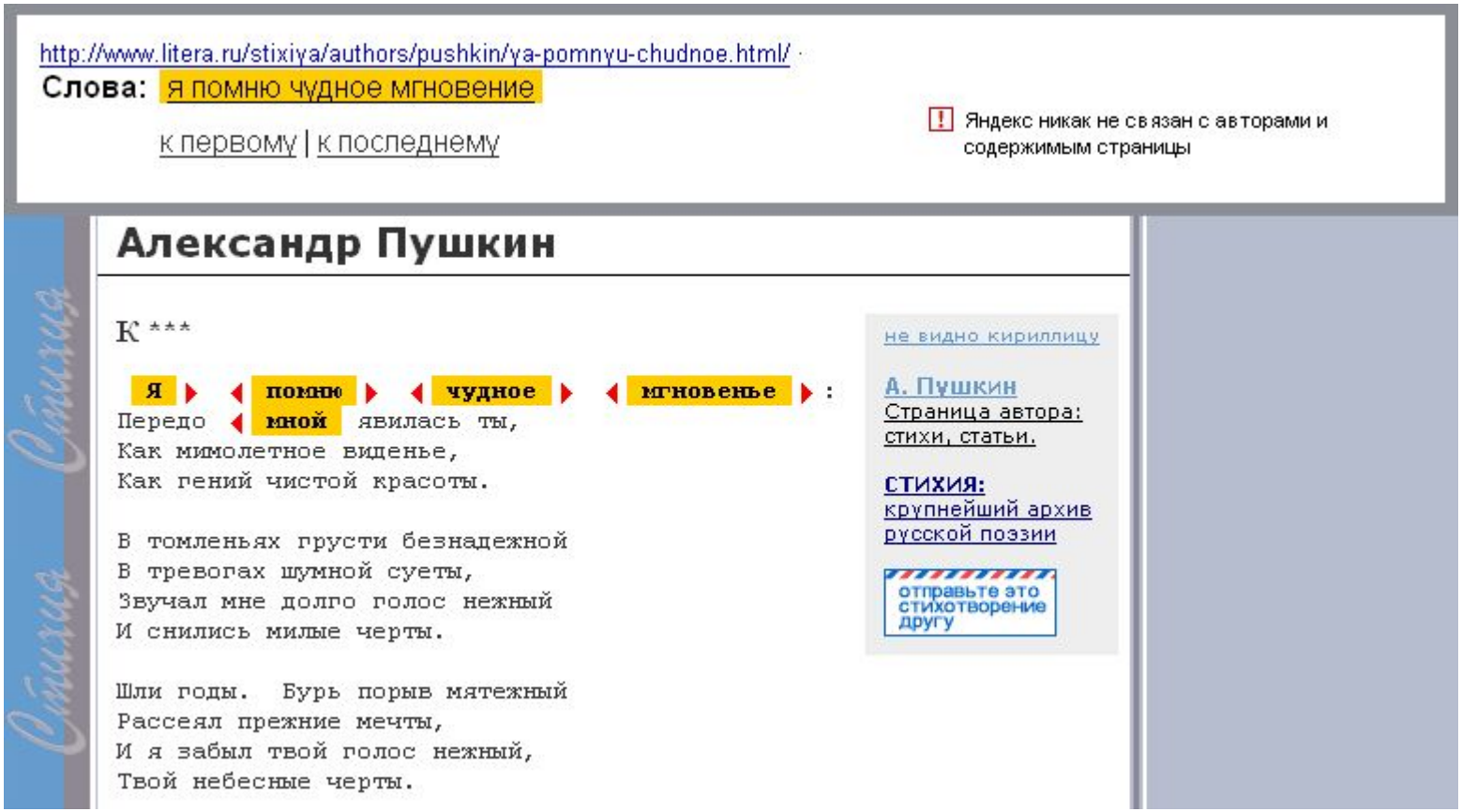


Я

За пределами выдачи

Я

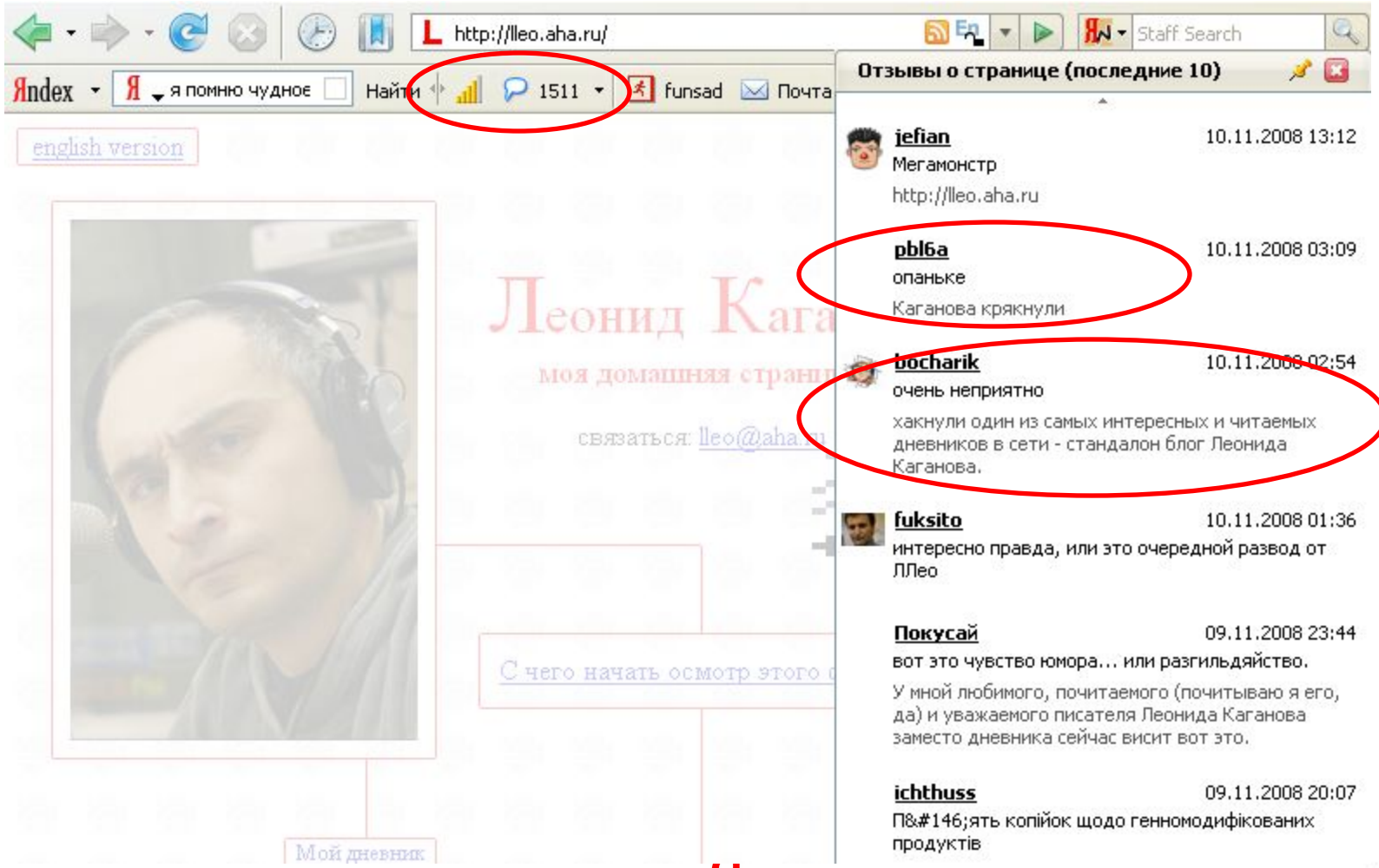
Подсветка слов, всегда доступный UR



Я

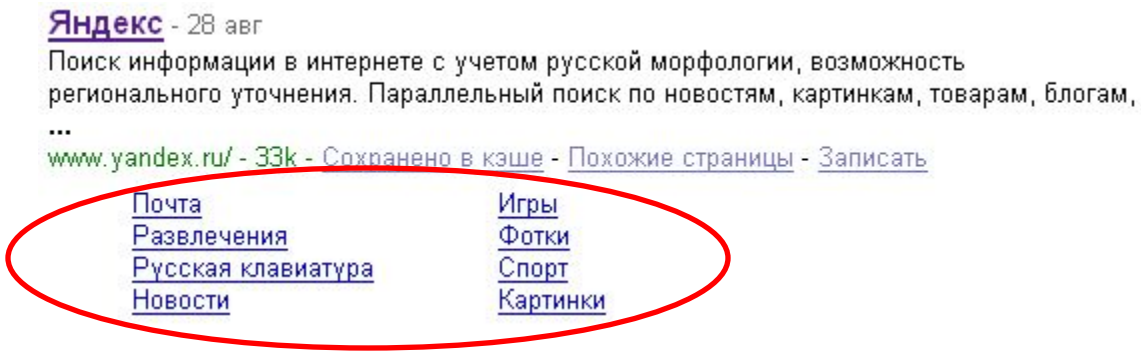
Изучая сайт, советоваться с поиском

- отзывы о странице в Баре; тИЦ как уровень доверия



Есть ли где еще разумная жизнь? :)

- Google: популярные разделы сайта



Яндекс - 28 авг
Поиск информации в интернете с учетом русской морфологии, возможность регионального уточнения. Параллельный поиск по новостям, картинкам, товарам, блогам, ...
www.yandex.ru/ - 33k - Сохранено в кэше - Похожие страницы - Записать

Почта	Игры
Развлечения	Фотки
Русская клавиатура	Спорт
Новости	Картинки

Я

Как все это создается?

Я

Некоторые аспекты измерения качества

- Оценка *страницы в целом* (привлекательность выдачи)
  колдунщик – такой же ответ, должен быть релевантным в соответствии с позицией
- *Распределение трафика* по странице
  кликабельность всех результатов, в т.ч. смежных с колдунщиками
- Эксперименты *на части аудитории*
- Оценка сниппетов включает:
  - наличие *ответа* на запрос
  - *помощь в выборе* релевантного
  - *впечатление* о выдаче
- *Usability*-исследования
  проблема «баннерной слепоты»

Я

«…А сейчас со всей этой байдой мы попытаемся взлететь…»

Я

Проблемы будущего

- *Консистентность* vs. *Перегруженность* выдачи
- *Унификация формы ответов* на разные вопросы

  Где и в какой форме пользователь должен их ждать? «Слепота» от привычки. Скорость восприятия.
- Переключение *«абстрактное ↔ конкретное»* требует усилий

  перейти от поиска слов запроса к чтению описания сайта или рубрики непросто;

  факты могут классифицировать ответ не менее информативно, чем рубрика
- *Ветвление усилий* разработчиков поиска

  Рост сегментации запросов и ответов
  - конкуренция данных (решение – data mining, партнерство)
  - покрытие редких типов ответов (решение – социальный путь)
- *Следование за толпой*

  сокращение разнообразия запросов и ответов («богатые богатеют»). Одно из решений – персонализация

Я

Яндекс

Александр Садовский
руководитель отдела веб-поиска

эл. почта: funsad@yandex-team.ru