

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)*

W 3. [Поисковая система](#) - Википедия

Поисковая система — веб-сайт, предоставляющий возможность поиска информации в Интернете. Большинство **поисковых систем** ищут информацию на сайтах Всемирной паутины, но существуют также **системы**, способные искать файлы на ftp-серверах...

ru.wikipedia.org/wiki/Поисковик · 16 КБ

[Сохраненная копия](#) · [Еще с сайта](#) 1312 · Рубрика: [Универсальные энциклопедии](#)

Незаданные вопросы о сайте: новости

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

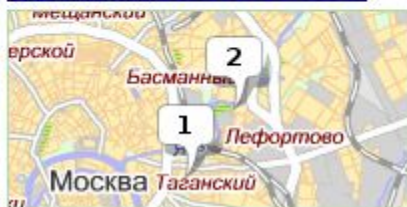
Я 1. ["Яндекс" - поисковая система по русскому интернету](#)

Поиск информации в интернете с учетом русской морфологии, возможность регионального уточнения. Параллельный поиск по новостям, картинкам, товарам, блогам, адресам организаций.

[www.yandex.ru](#) · 13 КБ

[Сохраненная копия](#) · [Еще с сайта 250853](#) · Рубрика: [Поисковые системы](#)

[Адреса в Москве и области](#)




1. [Яндекс](#) +7 (495) 7397000
Москва, Станиславского ул., 21, строение 3
2. [Яндекс](#) +7 (495) 7397000
Москва, Самокатная ул., 1, строение 21

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](#)

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:

W 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

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.

Мы не одиноки во Вселенной!

- Google: число сообщений в некоторых

Finjan предупреждает об обфускации вирусов

Сообщений: 4 - Авторы: 2

Форум журнала «Хакер» : самый популярный форум для хакеров, взломщиков, администраторов.

forum.hacker.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/viruses/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 рождения. 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:

[Блог Яндекса](#) - Translate
company.yandex.ru/blog - Cached

Стало:

[Блог Яндекса](#) - Translate
company.yandex.ru/blog - Cached

 Posts(5) - Feed Discovery ▲ ✕

 Posts(5) - Feed Discovery

[Керкинитиды, музей Айвазовског...](#) Nov 07
[Бесплатный антивирус для защит...](#) Oct 31
[Ударный сезон Яндекс.Расписаний](#) Oct 30
[Круглосуточно и без перерывов](#)

**Перед тем, как двигаться
дальше...**

Я

...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**

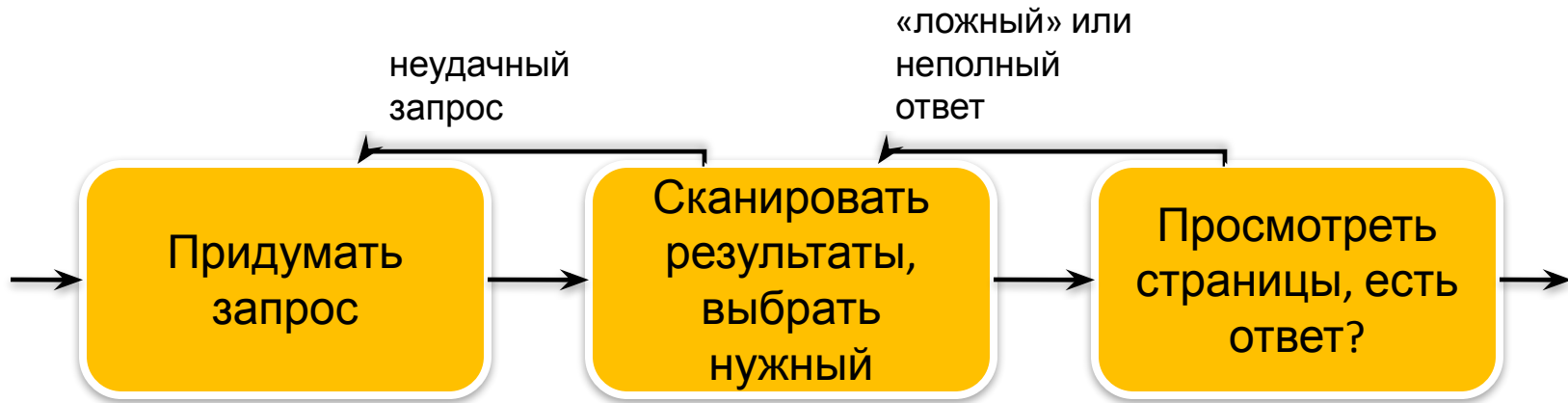
Парадигма **перенаправления**

- 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?!»
-

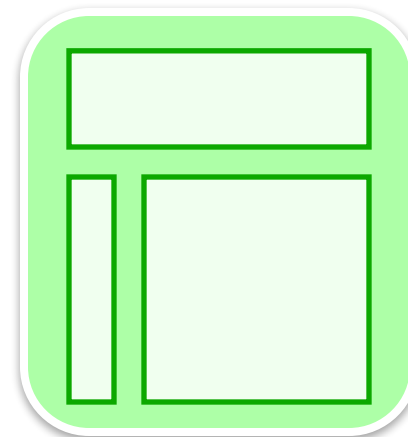
2007-2008. Оказывается, **жизнь есть**
до запроса,
в момент выбора ответа на выдаче
и даже после клика.

Я

На каждом шаге – потеря времени и сил

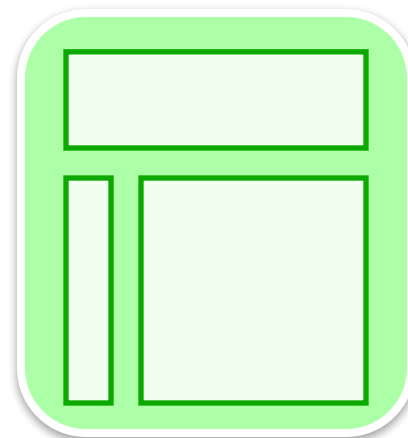
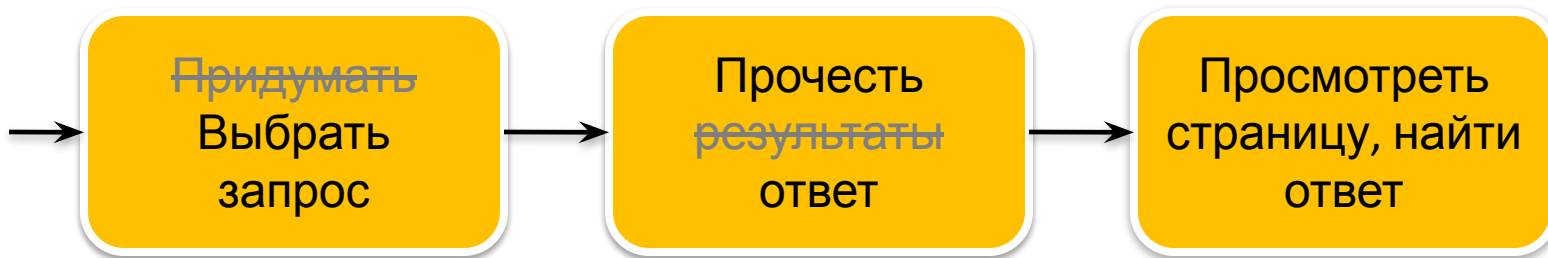


яндекс



Я

Все может быть проще



Я

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 **я** 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.

Сменились объекты изучения:

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

Колдунщики и фишки

Я

Колдунщики: событийные

- Олимпиада-2008

олимпиада 2008 волейбол



Олимпиада 2008: Волейбол



Ближайшие соревнования, 10 августа



Женщины. Олимпийский турнир  Польша —  Куба начало в 10:30

Мужчины. Олимпийский турнир  Италия —  Япония начало в 16:00

Последние результаты

Женщины. Олимпийский турнир Групповой этап. Группа В

 Италия 2:3  Россия

 Сербия 3:1  Казахстан

Колдунщики: структурированные данные

- Новости

сомалийские пираты

Новости
16:48



Против **сомалийских пиратов** направятся корабли ЕС

11 россиян, члены экипажа датского сухогруза "Фьюча", стали заложниками **сомалийских пиратов**. Кроме них на захваченном в минувшую пятницу судне остаются еще 2 представителя бывшего Союза - эстонец и...

Все сообщения 237 ·  Мнения 59

- Маркет

htc touch pro

Маркет



КПК

HTC Touch Pro — от 21465 р. до 39925 р. · Все предложения (63)

MS Windows Mobile 6.1, 480x640, Bluetooth 2.0, Wi-Fi, фотокамера, GSM/GPRS, GPS, microSD, microSDHC, Qualcomm MSM 7201A 528 МГц, 288 Мб RAM, 512 Мб ROM, 165 г, 51x102x18 мм.

- Карты, Словари, Афиша и т.д.

Я

Колдунщики: ответы на вопросы


- **Погода**

температура кодируется цветом

погода в новом уренгое

Погода: [Новый Уренгой](#) -27 °С  [подробный прогноз →](#)

погода в дакаре

Погода: [Дакар](#) +30 °С  [подробный прогноз →](#)

«Недозаданный» вопрос – регион определяем по IP

погода

Погода: [Москва](#) +5 °С  [подробный прогноз →](#)

Я

Колдунщики: ответы на вопросы

- Курсы валют

курс евро

Курс ЦБ РФ на завтра: 1 евро = **34,56** рубля [динамика →](#)

курс гривны в долларах

Курс НБУ на сегодня: 100 долларов = **579,84** гривны [динамика →](#)

«Недозаданный» вопрос –
показываем курс к самой популярной валюте

курс рубля

Курс ЦБ РФ на завтра: 1 доллар = **26,96** рубля [динамика →](#)

Я

Фишки: не ждали?!

- Цвет
а

цвет лягушки в обмороке

Аквамариновый
Панг
Лягушки в обмороке
Маренго
Серой спаржи

R	G	B	H°	S%	V%
123	145	123	120	15	57

7b917b

Detailed description: This block shows a color palette for the color 'лягушки в обмороке'. On the left is a vertical bar with a gradient from light green to dark grey, with labels: 'Аквамариновый', 'Панг', 'Лягушки в обмороке' (highlighted), 'Маренго', and 'Серой спаржи'. To the right is a circular color wheel with a square color selector. Further right is a table with color data: R=123, G=145, B=123, H°=120, S%=15, V%=57. Below the table is a text box containing the hex code '# 7b917b'.

ультрамариновый

Полуночно-синий
Темно-синий
Ультрамариновый
Темно-лазурный
Синей пыли

R	G	B	H°	S%	V%
18	10	143	244	93	56

120a8f

Detailed description: This block shows a color palette for the color 'ультрамариновый'. On the left is a vertical bar with a gradient from dark blue to light blue, with labels: 'Полуночно-синий', 'Темно-синий', 'Ультрамариновый' (highlighted), 'Темно-лазурный', and 'Синей пыли'. To the right is a circular color wheel with a square color selector. Further right is a table with color data: R=18, G=10, B=143, H°=244, S%=93, V%=56. Below the table is a text box containing the hex code '# 120a8f'.

Снова фишки

- Время

сколько времени в шемышейке

⌚ Точное [время](#): Шемышейка 19 50

«Недозаданный» вопрос – смотрим регион по IP

который час

⌚ Точное [время](#): Москва 19 51

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

Я

Помогаем в формулировании запроса

Бесплатно, 24x7

www.yandex.ru



Исправление ошибок

- Опечатки – 5% запросов предлагаем...

чучхэ

Быть может, вы искали: [«чучхе»](#)

...5% автоматически исправляем (незаданный вопрос)

сколько людей страдает дальтонизмом

По [исходному запросу](#) «сколько людей страдает вальтонизмом» ничего не нашлось. Результаты поиска для «сколько людей страдает дальтонизмом».

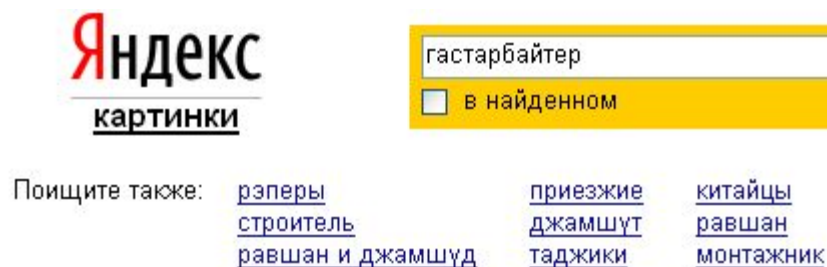
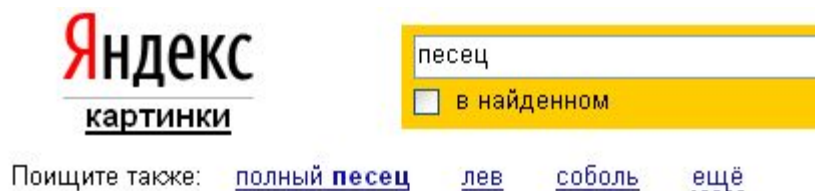
- Транслитерация, неправильная раскладка

zyltrc

Быть может, вы искали: [«яндекс»](#)

Только где же это что-то?

- запросы, связанные с вашим запросом



Поисковые подсказки

- упрощают набор – раз

белой|

белой акации гроздя душистые

белой акации

белой акации гроздя

белой акации гроздя душистые скачать

белой акации гроздя душистые mp3

белой акации аккорды

белой акации гроздя душистые текст

белой акации грозди душистые

белой

белой акации гроздя душистые аккорды

Закреть

- исправляют ошибки – два

атист|

аттестат

атистур

аттестация

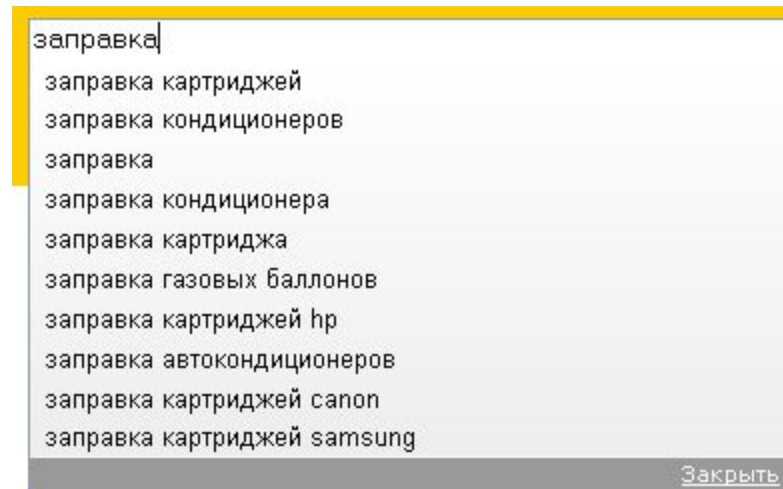
fnbcn

Закреть

Я

Поисковые подсказки

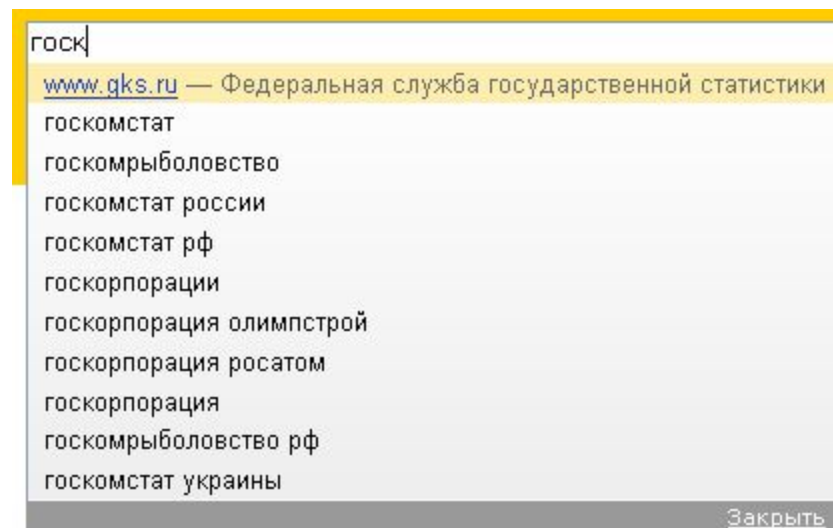
- дают представление о спектре тем и прагматик – три



- «обучают» новичков, как задавать запрос – четыре

Поиск до поиска!

- дают ответы с URL нужных сайтов – пять



За пределами выдачи

Я

Подсветка слов, всегда доступный URL

<http://www.litera.ru/stixiya/authors/pushkin/ya-pomnyu-chudnoe.html/>

Слова: **Я ПОМНЮ ЧУДНОЕ МГНОВЕНИЕ**

[к первому](#) | [к последнему](#)

! Яндекс никак не связан с авторами и содержимым страницы

Александр Пушкин

К ***

Я ▶ ◀ **помню** ▶ ◀ **чудное** ▶ ◀ **мгновенье** ▶ :

Передо **мной** явилась ты,
Как мимолетное виденье,
Как гений чистой красоты.

В томленьях грусти безнадежной
В тревогах шумной суеты,
Звучал мне долго голос нежный
И снились милые черты.

Шли годы. Бурь порыв мятежный
Рассеял прежние мечты,
И я забыл твой голос нежный,
Твой небесные черты.

[не видно кириллицу](#)

[А. Пушкин](#)

[Страница автора:](#)
[стихи, статьи.](#)

СТИХИЯ:
[крупнейший архив](#)
[русской поэзии](#)

[отправьте это стихотворение другу](#)

Я

Изучая сайт, советоваться с поиском

- отзывы о странице в Баре; ТИЦ как уровень доверия

Yandex search results for <http://leo.aha.ru/>. The browser address bar shows the URL. The search bar contains the text "Я помню чудное". The search results show a profile picture of Leonid Kaganov and the text "Леонид Каганов моя домашняя страница". A sidebar on the right displays "Отзывы о странице (последние 10)".

Отзывы о странице (последние 10)

Username	Text	Date
iefian	Мегамонстр http://leo.aha.ru	10.11.2008 13:12
pbl6a	опаньке Каганова крякнули	10.11.2008 03:09
bocharik	очень неприятно хакнули один из самых интересных и читаемых дневников в сети - стандалон блог Леонида Каганова.	10.11.2008 02:54
fuksito	интересно правда, или это очередной развод от ЛLeo	10.11.2008 01:36
Покусай	вот это чувство юмора... или разгильдяйство. У мной любимого, почитаемого (почитываю я его, да) и уважаемого писателя Леонида Каганова вместо дневника сейчас висит вот это.	09.11.2008 23:44
ichthuss	П8:#146;ять копійок щодо генномодифікованих продуктів	09.11.2008 20:07

С чего начать осмотр этого...

Мой дневник

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](#)

[Развлечения](#)

[Русская клавиатура](#)

[Новости](#)

[Игры](#)

[Фотки](#)

[Спорт](#)

[Картинки](#)

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»

**«...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

The logo for Yandex, featuring the word "Яндекс" in a bold, sans-serif font. The letter "Я" is red, and the remaining letters "ндекс" are black.

**Александр
Садовский**

руководитель отдела веб-
поиска

эл. почта: funsad@yandex-team.ru