

# Яндекс

Подготовка данных  
для Поиска по блогам

Антон Волнухин, Мария Игнатьева

- Как **открыть** Поиску по блогам данные?

С помощью открытых форматов для быстрой и полной индексации данных:

- RSS
- ATOM
- FOAF
- XML-RPC: [Weblogs.ping](#)

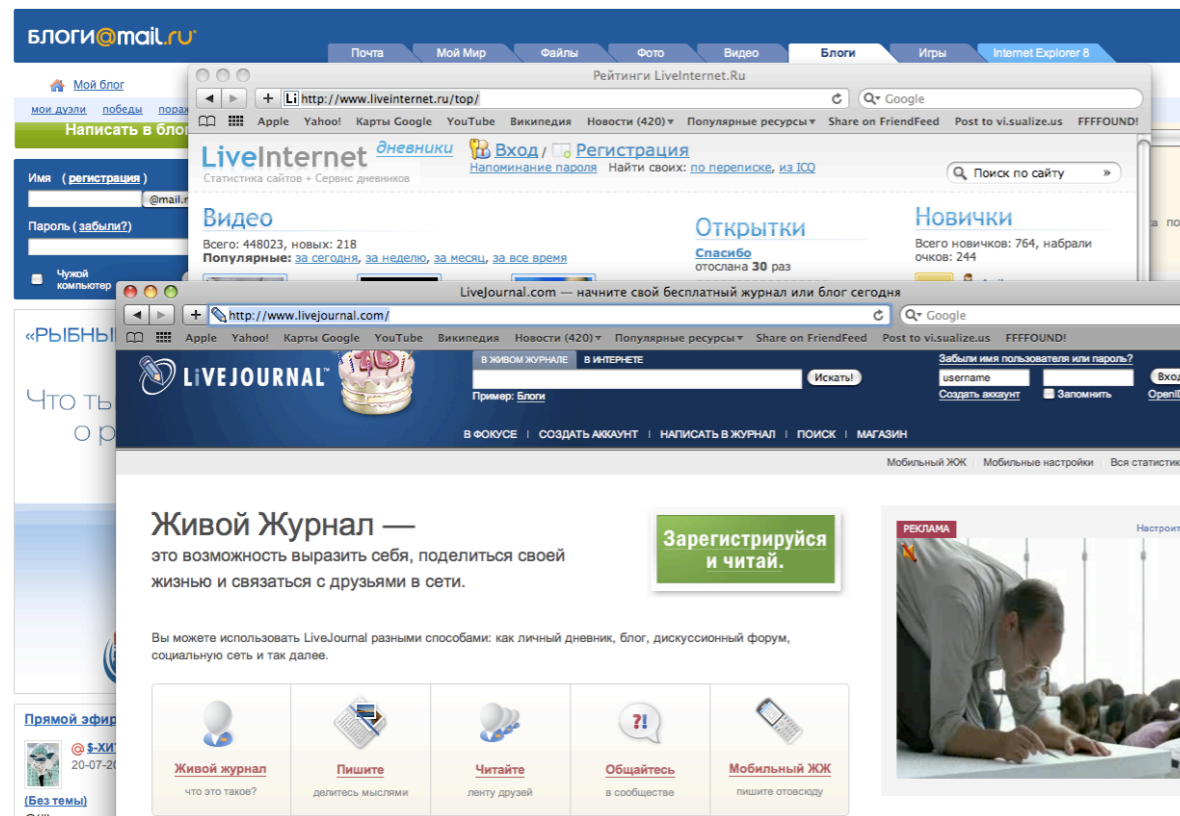
- Как **получить** данные от Поиска по блогам?

Результаты поиска, доступные всем в тех же открытых форматах:

- RSS (результаты поиска по записям и комментариям)
- ATOM+FOAF (результаты поиска по профилям)

Сервис основан на распространённых в интернете  
открытых форматах

Благодаря сотрудничеству с владельцами блог-хостингов  
эти форматы (RSS, FOAF, Weblogs.Ping) стали  
стандартом в российской блогосфере



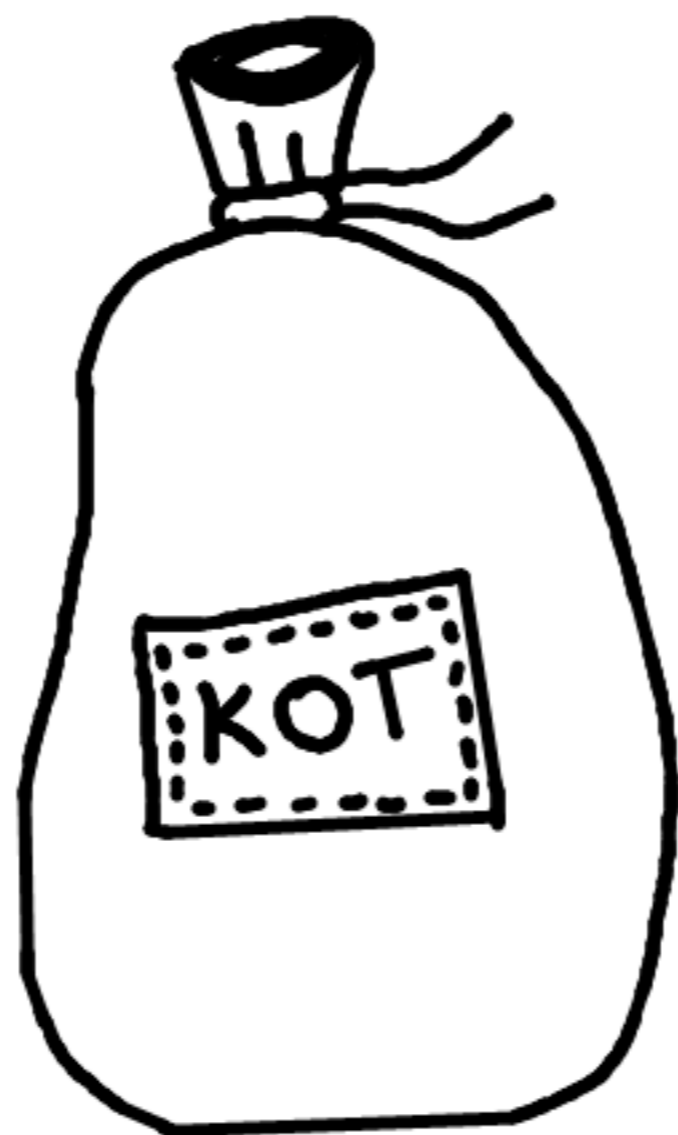
- Livejournal.com
- Blogs.mail.ru
- Liveinternet.ru

> 100 блог-ХОСТИНГОВ

# Что индексирует поиск по блогам

- блоги (RSS и ATOM)
- форумы (RSS и ATOM)
- профили (FOAF + Yandex FOAF extension)
- комментарии (RSS и ATOM)

С помощью пингов можно ускорить индексацию, сделав её почти мгновенной



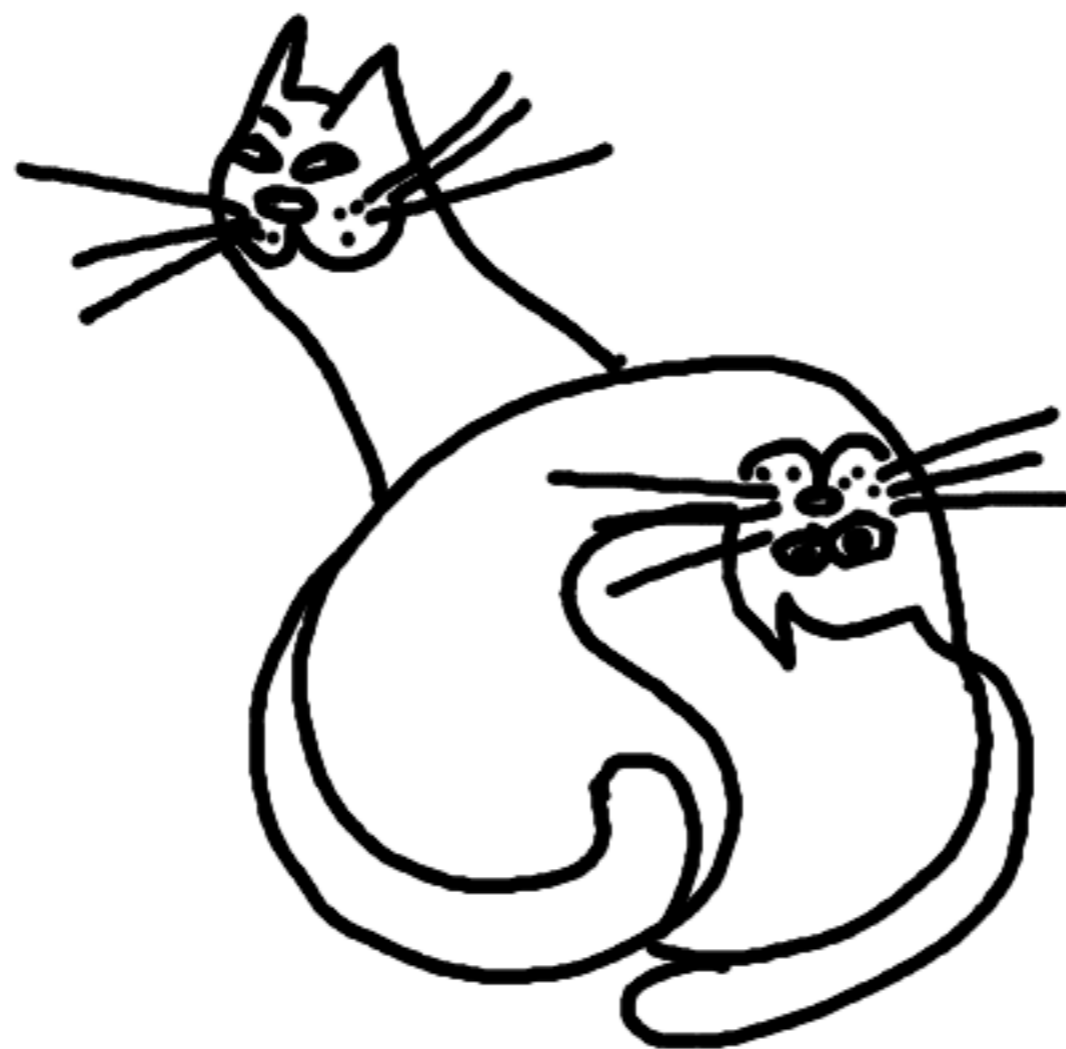
Я

# Преимущества индексации:

- Сервис лучше находится в Поиске по блогам
- Участие в рейтингах Поиска по блогам  
(в том числе, рейтинги сервисов, блогов и популярных записей)
- Улучшенная индексация веб-поиском  
(данные из Поиска по блогам используются в основном поиске)
- Возможность легко сделать поиск по своему сервису или блогу с помощью API ППБ



На что важно обратить **внимание**



**Я**

# Индексация записей

Индексируем стандартный RSS и ATOM

- уникальные адреса записей
- уникальные GUID
- валидный фид согласно [feedvalidator.org](http://feedvalidator.org)
- поддержка `<ya:more>` (для загрузки архивных записей)
- полные тексты записей
- тэг `<author>`, указанный в виде html-адреса на его журнал

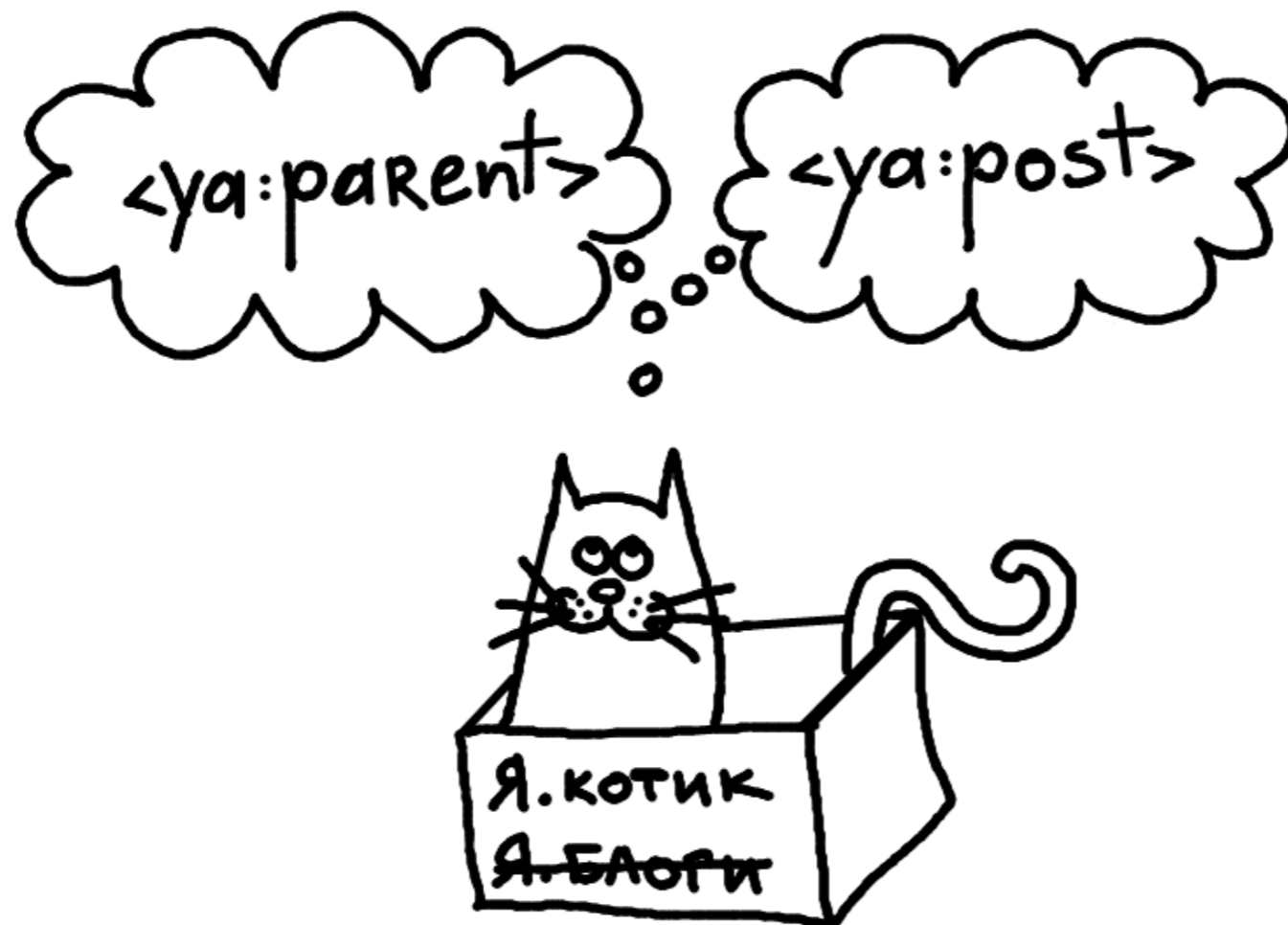
(хотя это противоречит стандартному описанию RSS)

# Индексация комментариев

Индексируем стандартный RSS и ATOM

Те же требования, что и к записям, плюс:

- тэг `<ya:parent>` (для древовидной структуры комментариев)
- тэг `<ya:post>` (определяет, к какой записи относится каждый комментарий из общего потока)
- в RSS с записями должна быть ссылка на RSS комментариев



# Индексирование форумов

Индексируем стандартный RSS и ATOM

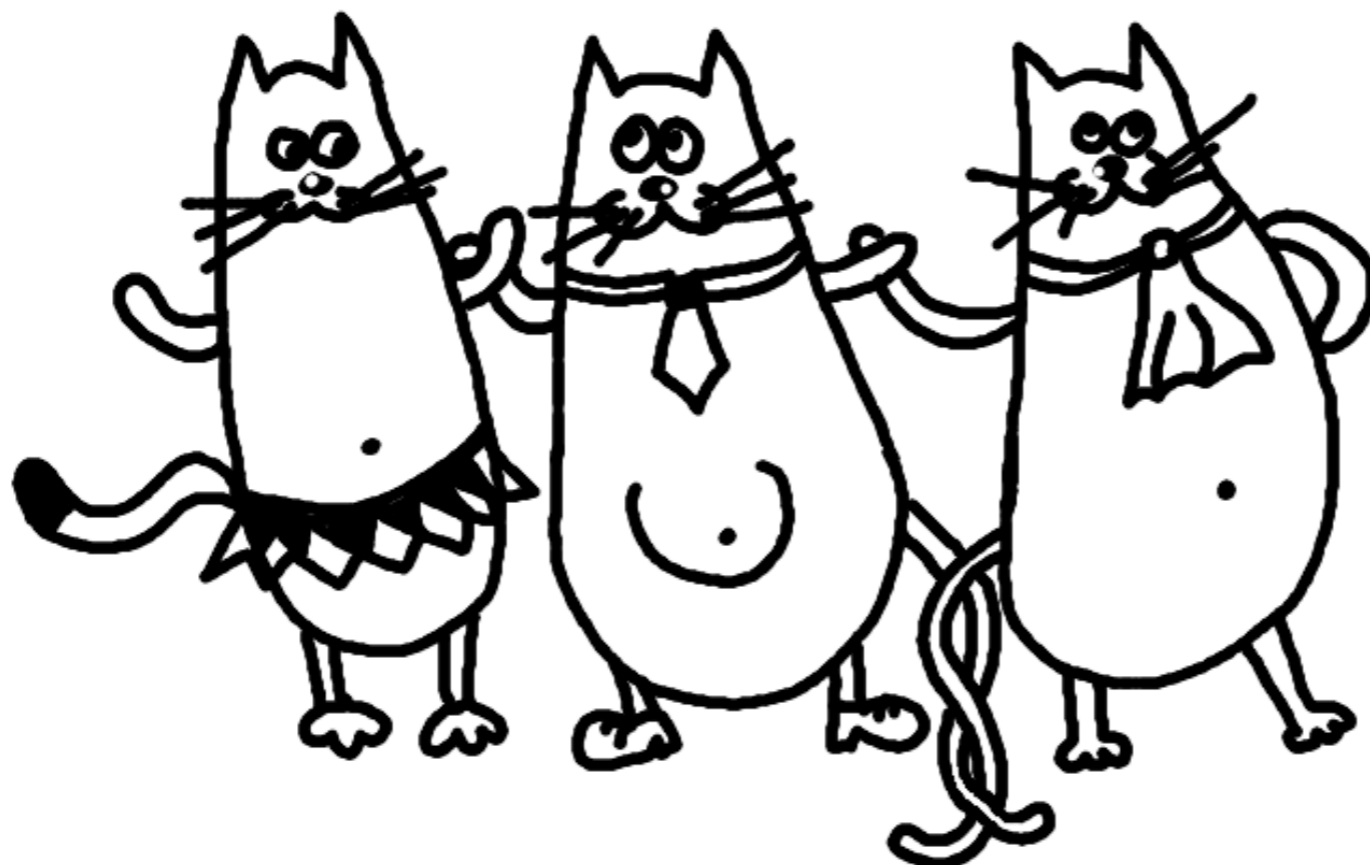
- стандартный тег RSS <[author](#)>
- тэг <ya:post>
- тэг <ya:parent>
- тэг <ya:more>
- **wtw:commentRss** (поток с записями может ссылаться на поток с комментариями)

# Индексирование профилей

Индексируем FOAF + Yandex FOAF extension

**FOAF** (акроним от Friend of a Friend - "друг друга") является машиночитаемым языком для описания людей, групп и отношений между ними

- создать FOAF-представление для профильных данных каждого пользователя
- указать ссылку на пользователя на странице, указанной в теге [author](#) в RSS-потоках записей и комментариев



# Пинги: мгновенная индексация



При появлении новой записи на блогахостинге можно посылать Яндексу специальное сообщение по протоколу [Weblogs.Ping](http://weblogs.ping.yandex.ru/).

**Адрес:**

<http://ping.blogs.yandex.ru/RPC2>

**Описание и инструкция:**

<http://blogs.yandex.ru/pings/info/>



# Подробная документация

Яндекс

API Поиска по блогам

[О проекте](#) [Документация](#)

поиск

## Документация

В этом разделе собрана техническая документация ко всем компонентам, составляющим программный интерфейс (API) службы Яндекс.Поиск по блогам. Чтобы лучше понимать их, рекомендуем ознакомиться с [гlossарием](#), который содержит краткие объяснения основных терминов, используемых в наших текстах. Вашему вниманию предлагается полная документация на русском языке, а также наиболее важная её часть в переводе на английский.

## Руководство по индексированию блогов, форумов и профилей

В этом разделе содержится описание особенностей индексирования блогов, форумов и профилей, которое поможет вам сделать поиск по вашему блогу или соц сети наиболее эффективным.

- [О руководстве](#)
- [Индексирование RSS — общие сведения](#)
- [Индексирование профилей \(FOAF\) — общие сведения](#)
- [Примеры подготовленных для индексирования документов](#)
- [RDF-схема расширений Яндекса для FOAF](#)

## Справочник по API

Здесь можно найти информацию о том, как правильно формировать запросы к Поиску по блогам, а также о форматах, в которых возвращаются результаты поиска

- [О справочнике](#)
- [Поиск по записям, комментариям и профилям — общие сведения](#)
- [Формирование запросов в адресной строке — обзор GET-параметров](#)
- [Формирование запросов при помощи языка запросов Яндекс.Поиска по блогам](#)
- [Поиск по профилям — представление результатов](#)
- [Поиск по записям и комментариям — представление результатов](#)

## Yandex extension of FOAF (in English)

Yandex extension of FOAF provides the main classes and properties required to describe users' profiles and their blog activities. It represents an attempt to make Semantic even more convenient for automated search.

- [Detailed description of Yandex extension of FOAF](#)



© 2008—2009 «Яндекс»  
О компании · Обратная связь  
Работайте в Яндексе

<http://api.yandex.ru/blogs/doc/>

Я

# Яндекс

**Антон Волнухин**

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

