

**Я**НДЕКС

ПОИСК



# Путь к Rich Content API. Контент «по-богатому»

Я.Субботник, Минск, 31 августа

Александр Гусак  
Бригадир

# Rich Content API

# Классическая ссылка

<http://clck.ru/8o9CX>

# Ссылка по-богатому





# Rich Content preview



Yandex Rich Content API: [Documentation](#)

## Yandex Rich Content API



### Я.Субботник в Минске — Яндекс.События

На Я.Субботнике вы узнаете: про разработку сервиса Sita и как с его помощью был сделан Rich Content API; как Яндекс разбил разработку фронтенда Островов на независимо-зависимые блоки; о технологии Cocaine, позволяющей построить своё собственное, приватное облако даже на обычном десктопе; как мы применили БЭМ к серверному коду на Node.js и получили общий модульный метод разработки веб-приложений; что скрыто за пуш-сообщениями и как они работают — на примере сервиса Яндекса...

# Rich Content API, пример ответа

<http://rca.yandex.com/?key=24LbkxMc&url=http://clck.ru/8o9CX>

```
{
  "url": "http://clck.ru/8o9CX",
  "title": "Я.Субботник в Минске – Яндекс.События",
  "finalurl": "http://events.yandex.ru/events/yasubbotnik/minsk-
aug-2013/",
  "img": [
    "http://yandex.st/events/_/b1HDrOVXFba-SipQXXKAPG9HaAI.png"
  ],
  "content": "На Я.Субботнике вы узнаете:\nпро разработку сервиса
Sita и как с его помощью был сделан Rich Content API;\nкак Яндекс
разбил разработку фронтенда Островов на независимо-зависимые
блоки;\nпо технологии Cosaine, позволяющей построить своё
собственное, приватное облако даже на обычном десктопе;\nкак мы
применили БЭМ к серверному коду на Node.js и получили общий
модульный метод разработки веб-приложений;\nчто скрыто за пуш-
сообщениями и как они работают – на примере сервиса Яндекса..."
}
```

# Rich Content API, где использовать

- форумы
- социальные сети
- блог-платформы
- мессенджеры
- браузеры
- сервисы укорачивания урлов
- ... и другие места, где есть урлы

# Sita

- сервисы контент-системы, Sita
- пример сценария
- HTTP API
- реализация Rich Content API



# Сервисы контент-системы

## Фетчер (качалка)

- скачивание документов из интернета
- соблюдение политик обхода
- кэширование
- предварительная обработка документов

# Сервисы контент-системы

## База данных

- хранилище документов и дополнительных данных
- платформа для вычислений
- «большой» кэш фетчера
- доступ по ключу (realtime)
- batch-режим для обхода всей базы

# Сервисы контент-системы

## Sita

- пользовательские сценарии поверх других сервисов
- комбинирование сценариев
- различные интерфейсы доступа к данным и сценариям
- кэширование

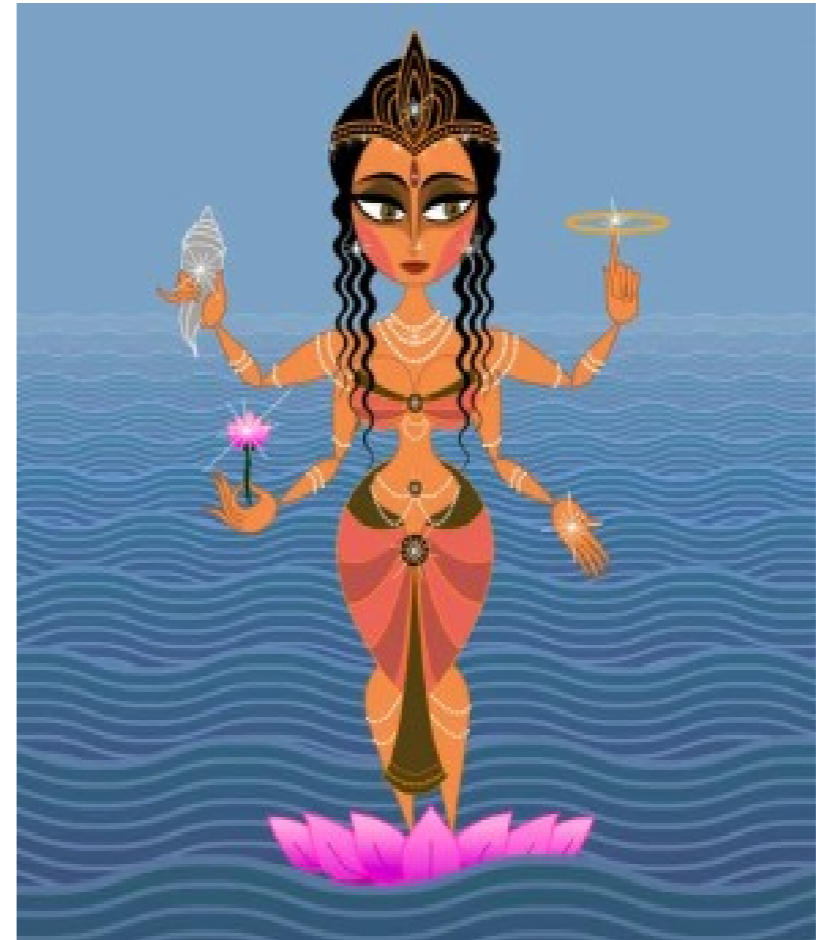
# Сервисы контент-системы

## Общее для всех сервисов

- отказоустойчивость
- масштабируемость
- передача данных в формате Google Protocol Buffers

# Sita

Сита — воплощение Лакшми



# Укорачивание урлов

<http://yandex.by>



<http://clck.ru>



<http://clck.ru/LiHj>



# Удлинение урлов в Sita

<http://clck.ru/LiHj>



<http://yandex.by>



<http://www.yandex.by>

# Что удлинение урлов делает в Sita?

- удлинить один урл можно с помощью `wget/curl`
- десяток тоже можно
- с миллионами могут быть проблемы

# Удлинение урлов в Sita

- ищем урл в базе данных («большой» кэш)
- если урла нет в базе, качаем
- котируем пользователей

Sita. HTTP API

**Интероперабельность — хорошо.**

# Sita. HTTP API

**Несогласованность интерфейсов —  
плохо.**

# Sita. HTTP API

Ключ к согласованности —  
в автоматической генерации одного  
интерфейса из другого.





# Sita. HTTP API

## JSON-запрос

```
{  
  "Url": "http://clck.ru/LiHj",  
  "Action": "UrlExtension"  
}
```



## protobuf-запрос

```
SitaRequest {  
  Url: "http://clck.ru/LiHj"  
  Action: UrlExtension  
}
```

# Sita. HTTP API

## protobuf-отвeт

```
SitaResponse {  
  Redirections {  
    Url: "http://clck.ru/LiHj"  
    HttpStatusCode: 302  
  }  
  Redirections {  
    Url: "http://yandex.by/"  
    HttpStatusCode: 302  
  }  
  Redirections {  
    Url: "http://www.yandex.by/"  
    HttpStatusCode: 200  
  }  
}
```



## JSON-отвeт

```
{  
  "Redirections": [  
    {  
      "Url": "http://clck.ru/LiHj",  
      "HttpCode": 302  
    },  
    {  
      "Url": "http://yandex.by/",  
      "HttpCode": 302  
    },  
    {  
      "Url": "http://www.yandex.by/",  
      "HttpCode": 200  
    }  
  ]  
}
```

# Rich Content API с помощью робота

- перевели алгоритмы построения превью на вычислительную платформу
- реализовали сценарий построения превью в Sita
- скомбинировали со сценарием удлинения урлов
- получили профит в виде отказоустойчивости, масштабируемости, скорости работы

Rich Content API. Где посмотреть?

<http://api.yandex.com/rca/>



**Я**ндекс

Александр Гусак

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

Спасибо