

**Яндекс**

# Как работают Яндекс. Новости

Татьяна Исаева

# Задачи Яндекс.Новостей

**Миссия Яндекса** – отвечать на заданные и незаданные вопросы пользователей

Яндекс.Новости:

- **На главной странице** Яндекса удовлетворяют потребность пользователей в актуальной информации

Незаданный вопрос: *«Что сейчас происходит?»*

- **В Поиске** отвечают на явно сформулированные вопросы пользователей

Заданные вопросы: *«синагога в Твери», «Хамовнический суд», «танцующий мост в Волгограде»*

# Модель сервиса



## Издания

- Обеспечивают быструю трансляцию контента (~40 тыс. сообщений в сутки)
- Получают читателей (1,5 млн переходов в сутки)

## Читатели

- Узнают новости на Яндексе (15 млн человек в сутки)
- Читают подробности у источников (1 млн человек в сутки)
- Задают вопросы о новостях (500 тыс. запросов в сутки)

# Принципы Яндекс.Новостей

- Полностью автоматический сервис
- Партнерский сервис
- Единые требования для всех партнеров
- Зеркало русскоязычной медиа-среды

Цели:

отразить новостную картину дня и дать ссылки на источники, содержащие наиболее полную информацию о событиях

# Релевантный новостной ответ

Алгоритм, формирующий ответ на новостные запросы пользователей, ориентируется на следующие свойства сообщений:

- принадлежность первоисточнику
- оперативность
- цитируемость
- информативность

Источники, у которых доминируют такие сообщения, предлагаются пользователям в первую очередь.

# Требования к сообщению

- **Формат** (не подходят для трансляции реклама; не обработанные силами редакции пресс-релизы; новости сайта; аудио- и видеоматериалы, не сопровождаемые связным печатным текстом; таблицы и списки, не сопровождаемые связным печатным текстом)
- **Технология** (если текст сообщения на сайте был изменён, необходимо поместить его в экспортный файл в изменённом виде или прислать запрос на удаление неактуального сообщения)
- **Требования законодательства** (аккуратное цитирование; нормы литературного русского языка)

# Что такое \*дубликат

Дубликаты – документы, в большой степени совпадающие с оригинальным сообщением. Они автоматически определяются до формирования сюжета, **не представлены** в сюжетах и **не участвуют** в поиске по сервису.

Из каждой группы дубликатов остается по одному сообщению («мастеру»), которое может быть представлено на страницах сюжета и в результатах поиска. «Мастер» определяется:

- по времени публикации сообщения на сайте;
- по сравнительному анализу текстов;
- по цитированию источника (учитываются гиперссылки, текстовые упоминания).

# Статистика «копипейста»

- Доля дубликатов сообщений в российских интернет-СМИ - **20%**
- Доля изданий, которые ни разу не перепечатавали чужие материалы «вчистую» - тоже 20%
- Доля сообщений, содержащих ссылки на первоисточники (названия других СМИ или гиперссылки) – всего 8%

По данным аналитического отчета Яндекса «Медиа сфера Рунета», зима 2009

[http://download.yandex.ru/company/yandex\\_on\\_russian\\_internet\\_media\\_winter\\_2009.pdf](http://download.yandex.ru/company/yandex_on_russian_internet_media_winter_2009.pdf)

# Сюжет в Яндекс.Новостях

Фото (всего 30)



Вести.Ru



Голос России

Видео (всего 3)



Вести.Ru

ТРК Петербург - Пятый канал

Аудио (всего 4)

Эхо Москвы  
Радио Маяк  
РИА Новости

Общество / Лондон / Великобритания

## Евгений Чичваркин посетил вечер балета в Лондоне

Чичваркин в сопровождении троих спутников в антракте свободно прохаживался по Королевскому театру оперы и балета, где в этот вечер давали юбилейный концерт, посвященный 100-летию "Русских сезонов" [Сергея Дягилева](#). [Взгляд](#) 13:22

[Евгений Чичваркин](#), которого ищет Интерпол, появился в "Ковент-гардене" экс-владельца компании "Евросеть" [Евгения Чичваркина](#) видели в воскресенье на балетном гала-вечере в лондонском "Ковент-гардене". [Аргументы и факты](#) 11:02

После концерта он отправился на VIP-ужин, который помимо него посетили глава фонда "Русский мир" [Вячеслав Никонов](#), президент "Альфа-банка" [Петр Авен](#), внучка хореографа "Русских балетов" [Михаила Фокина](#) Изабель и прочие меценаты. [Vesti.kz](#) 09:44

Ранее глава национального центрального бюро Интерпола при МВД России Тимур Лахонин заявил, что когда заходит разговор об экстрадиции в Россию [Евгения Чичваркина](#), власти Великобритании "включают дурака".

Тимур Лахонин, генерал-майор [все цитаты сюжета](#) (6)

Еще на эту тему

[Названы лучшие хореографы и танцовщики 2009 года](#)  
05.06.2009

В сюжете: [Евгений Чичваркин](#), [Сергей Дягилев](#) • [Евросеть](#), [Королевский театр оперы и балета](#), [Альфа-банк](#), [Русский мир](#), [Пушкинского дома](#)

Все новости сюжета (всего 64) [Интервью](#) (всего 1)

[Евгений Чичваркин официально объявился в Лондоне](#) [Вести.Ru](#) 16:24

[Экс-владелец "Евросети" посетил вечер балета в Лондоне](#) [Голос России](#) 13:35

[Чичваркин побывал на спектакле в Лондоне](#) [Взгляд](#) 13:22

["Беглый" предприниматель Чичваркин объявился в Лондоне](#) [Право.ру](#) 12:32

[Беглый бизнесмен Чичваркин и не думает прятаться от Интерпола](#) [Белорусские новости](#) 12:31

[Разыскиваемый Интерполом Чичваркин был замечен в Ковент-Гардене](#) [Каспаров.Ru](#) 12:29

[Евгений Чичваркин посетил вечер балета в Лондоне](#) [Коммерсантъ-Online](#) 12:18

[Чичваркин появился на русском балете](#) [Утро.ру](#) 12:07

[Чичваркин: "Не обижайтесь, пожалуйста"](#) [Московский комсомолец](#) 11:16

[Чичваркин появился в королевском театре](#) [Дни.ру](#) 11:01

[Разыскиваемый Интерполом экс-владелец "Евросети" посетил лондонский театр](#) [NEWSru.com](#) 10:57

[Чичваркин посетил вечер балета в Лондоне](#) [Газета.Ru](#) 10:41

[Евгений Чичваркин вышел "в свет" в Лондоне](#) [РИА Новости](#) 10:12

# Схема создания сюжета

Создание поискового запроса из ключевых слов документа, построение матрицы близости документов



Аннотирование сюжета: выбор заголовка и фрагментов сообщений, содержащих основные факты сюжета



Выбор основных заголовков сюжета



Выбор статей и интервью

# Выбор заголовка сюжета

- Соответствие региону пользователя
- Лексическая и фактологическая «ядерность»
- Актуальность фактов
- Информативность и читаемость

Цель:

первый по времени заголовок, наиболее полно отражающий актуальную фактическую сторону сюжета, не содержащий нехарактерных для сюжета слов и фактов.

# Создание аннотации сюжета (дайджеста)

Из всех сообщений сюжета автоматически выделяются наиболее значимые объекты, имена людей, названия организаций, географические объекты, даты и числа. Они, наряду с ключевыми словами сюжета и новостными запросами, определяют выбор текстов для аннотации.

Цель:

показать предложения из сообщений,  
содержащие основные факты события

# Аннотация сюжета.

## Информативность

[Общество](#) // [Лондон / Великобритания](#)

### Евгений Чичваркин посетил вечер балета в Лондоне

[Чичваркин](#) в сопровождении троих спутников в антракте свободно прохаживался по [Королевскому театру оперы и балета](#), где в этот вечер давали юбилейный концерт, посвященный 100-летию ["Русских сезонов"](#) [Сергея Дягилева](#). [Взгляд](#) 13:22

[Евгений Чичваркин](#), которого ищет Интерпол, появился в ["Ковент-гардене"](#) экс-владельца компании ["Евросеть"](#) [Евгения Чичваркина](#) видели в воскресенье на балетном гала-вечере в лондонском ["Ковент-гардене"](#). [Аргументы и факты](#) 11:02

После концерта он отправился на VIP-ужин, который помимо него посетили глава фонда ["Русский мир"](#) [Вячеслав Никонов](#), президент ["Альфа-банка"](#) [Петр Авен](#), внучка хореографа ["Русских балетов"](#) [Михаила Фокина](#) [Изабель](#) и прочие меценаты. [vesti.kz](#) 09:44

Ранее глава национального центрального бюро Интерпола при МВД России [Тимур Лахонин](#) заявил, что когда заходит разговор об экстрадиции в Россию [Евгения Чичваркина](#), власти [Великобритании](#) включают дурака".

Тимур Лахонин, генерал-майор [все цитаты сюжета](#) (6)

**В сюжете:** [Евгений Чичваркин](#), [Сергей Дягилев](#) • [Евросеть](#), [Королевский театр оперы и балета](#), [Альфа-банк](#), [Русский мир](#), [Пушкинского](#)

**Все новости сюжета** (всего 64) **Интервью** (всего 1)

[Евгений Чичваркин официально объявился в Лондоне](#) [Вести.Ru](#) 16:24

[Экс-владелец "Евросети" посетил вечер балета в Лондоне](#) [Голос России](#) 13:35

[Чичваркин побывал на спектакле в Лондоне](#) [Взгляд](#) 13:22

["Беглый" предприниматель Чичваркин объявился в Лондоне](#) [Право.ру](#) 12:32

[Беглый бизнесмен Чичваркин и не думает прятаться от Интерпола](#) [Белорусские новости](#) 12:31

[Разыскиваемый Интерполом Чичваркин был замечен в Ковент-Гардене](#) [Каспаров.Ru](#) 12:29

[Евгений Чичваркин посетил вечер балета в Лондоне](#) [Коммерсантъ-Online](#) 12:18

[Чичваркин появился на русском балете](#) [Утро.ru](#) 12:07

[Чичваркин: "Не обижайтесь, пожалуйста"](#) [Московский комсомолец](#) 11:16

[Чичваркин появился в королевском театре](#) [Дни.ру](#) 11:01

[Разыскиваемый Интерполом экс-владелец "Евросети" посетил лондонский театр](#) [NEWSru.com](#) 10:57

[Чичваркин посетил вечер балета в Лондоне](#) [Газета.Ru](#) 10:41

[Евгений Чичваркин вышел "в свет" в Лондоне](#) [РИА Новости](#) 10:12

# Заголовок сюжета. «Ядерность» лексики



вечер балета

только в этом сюжете

[расшир.](#)

[Главные новости](#)

[Политика](#)

[В мире](#)

[Общество](#)

[Экономика](#)

[Спорт](#)

[Происшествия](#)

[Культура](#)

[Наука](#)

## Евгений Чичваркин посетил вечер балета в Лондоне

Результат поиска — сообщений : 34

1. [В Лондоне отметили 100-летие дягилевских "Русских сезонов"](#) 16:10 Вести.Ru [посмотреть](#)  
Торжественный **вечер** собрал артистов **балета** со всего мира. Зрителям представили отрывки из знаменитых **балетов** постановки Михаила Фокина, который много лет был бессменным руководителем "Русского **балета**".
2. [Чичваркин, которого ищет Интерпол, сходил в Лондоне на балет](#) 15:32 Россия  
...троих спутников в антракте свободно прохаживался по Королевскому театру оперы и **балета**, где в этот **вечер** давали юбилейный концерт, посвященный 100-летию "Русских сезонов" Сергея Дягилева. От общения с журналистами Чичваркин отказался.
3. [Чичваркин после балета отправился на ужин для "випов"](#) 13:59 Вслух.Ру  
Объявленный в международный розыск экс-владелец компании "Евросеть" Евгений Чичваркин накануне посетил **вечер балета** в Лондоне, посвященный 100-летию открытию "дягилевского сезона".
4. [Экс-владелец "Евросети" посетил вечер балета в Лондоне](#) 13:35 Голос России  
...балетный гала-**вечер** в лондонском "Ковент-гардене". Бизнесмен в сопровождении троих спутников в антракте свободно прохаживался по Королевскому театру оперы и **балета**. Между тем, у присутствующих появление предпринимателя не вызвало удивления.
5. [Чичваркин побывал на спектакле в Лондоне](#) 13:22 Взгляд  
Чичваркин в сопровождении троих спутников в антракте свободно прохаживался по Королевскому театру оперы и **балета**, где в этот **вечер** давали юбилейный концерт, посвященный 100-летию "Русских сезонов" Сергея Дягилева.
6. [Евгений Чичваркин замечен на вечере балета в Лондоне](#) 13:07 Агентство бизнес новостей  
Он в сопровождении троих спутников в антракте свободно прохаживался по Королевскому театру оперы и **балета**, где в этот **вечер** давали юбилейный концерт, посвященный 100-летию "Русских сезонов" Сергея Дягилева.

# Выбор основных заголовков сюжета

Цитирование  
источника  
в сюжете

Дата публикации

Вес источника

Цель:

Должны быть представлены первоисточник(и), наиболее цитируемые источники и заголовки, отражающие актуальное состояние сюжета

В сюжете: [Евгений Чичваркин, Сергей Дагилев](#) • [Евросеть, Королевский театр](#)

Все новости сюжета ([всего 64](#)) [Интервью](#) ([всего 1](#))

[Евгений Чичваркин официально объявился в Лондоне](#) [Вести.Ru](#) 16:24

[Экс-владелец "Евросети" посетил вечер балета в Лондоне](#) [Голос России](#) 13:35

[Чичваркин побывал на спектакле в Лондоне](#) [Взгляд](#) 13:22

["Беглый" предприниматель Чичваркин объявился в Лондоне](#) [Право.ру](#) 12:32

# Выбор заголовков для первой страницы. Цитируемость в сюжете

Яндекс

НОВОСТИ

риа новости

только в этом сюжете

расшири

[Главные новости](#)

[Политика](#)

[В мире](#)

[Общество](#)

[Экономика](#)

[Спорт](#)

[Происшествия](#)

[Культура](#)

[Наука](#)

## Евгений Чичваркин посетил вечер балета в Лондоне

Результат поиска — сообщений : 37

### 1. Чичваркин, которого ищет Интерпол, сходил в Лондоне на балет 15:32 Россия

Чичваркин, как передает **РИА Новости**, в сопровождении троих спутников в антракте свободно прохаживался по Королевскому театру оперы и балета, где в этот вечер давали юбилейный концерт, посвященный 100-летию "Русских сезонов" Сергея Дягилева.

### 2. Объявленный в розыск Евгений Чичваркин "отметился" в лондонском театре 13:27 Общая газета.РУ

Как сообщает **РИА Новости**, в антракте экс-владелец "Евросети" в сопровождении троих спутников прохаживался по театру оперы и балета.

### 3. Чичваркин побывал на спектакле в Лондоне 13:22 Взгляд

Как передают корреспонденты **РИА "Новости"**, бизнесмен, известный своими экстравагантными пристрастиями в одежде, на чужбине изменил своему стилю и пришел на мероприятие в строгом темно-синем костюме с галстуком.

### 4. Разыскиваемый Интерполом Чичваркин "вышел в свет" в Лондоне 13:08 RB.ru - деловая сеть.

Не обижайтесь, пожалуйста", - сказал он в ответ на попытку корреспондента **РИА "Новости"** задать ему вопрос. После спектакля Чичваркин отправился на закрытый ужин, где присутствовали только VIP-гости.

### 5. Евгений Чичваркин замечен на вечере балета в Лондоне 13:07 Агентство бизнес новостей

После спектакля с участием российских и мировых звезд балета господин Чичваркин отправился на закрытый ужин для VIP-гостей в том же здании, сообщает "Коммерсант" со ссылкой на **РИА Новости**.

Я

# Определение жанра сообщения

- **Лента** – короткие новостные сообщения
- **Новость** – более развёрнутое сообщение
- **Интервью** – диалог, оформленный по правилам пунктуации русского языка
- **Статья** – большой по объёму текст (от 500 слов), представляющий собой анализ события, ситуации etc, в котором представлены разные точки зрения

# Выбор главных новостных сюжетов

Сегодня в новостях 19:10 все Москва

1. На шахте в Донецке произошла [авария, судьба 26 горняков неизвестна](#)
2. Американских журналисток [приговорили к 12 годам каторги в КНДР](#)
3. Каждый десятый выпускник [не сдал ЕГЭ по географии](#)
4. Генпрокуратура России реабилитировала [членов царской семьи Романовых](#)
5. Суд приговорил убийцу московской школьницы к [23 годам лишения свободы](#)

# Яндекс

Найдётся всё

Почта

логин

пароль

запомнить меня

Войти

[Забыли пароль?](#)

[Завести почтовый ящик](#)

 Фотки



Поиск [Карты](#) [Маркет](#) [Новости](#) [Словари](#) [Блоги](#) [Видео](#) [Картинки](#) [ещё](#) ▼

Например, день промыт, как стекло, только этого мало

Доступный мир  
полного привода

Специальное предложение  
на все модели до конца июня



Каталог сайтов

[Игры и развлечения](#)

[Спорт и отдых](#)

[Работа и учеба](#)

[Компьютеры](#)

[Бизнес](#)

[Дом и авто](#)

[Сайты Москвы](#)

 Маркет

[купить телевизор](#)

 Авто

[объявления о продаже](#)

 Мой Круг

[разместить резюме](#)

 Народ

[обмен файлами](#)

 Деньги

Москва. 8 июня,

Погода  +17

ночью +6 завтра +25

Котировки

06/06

USD ЦБ 30,6919 +0,38%

EUR ЦБ 43,6009 -0,27%

MMVB -2,33% 1117,66

Нефть -0,60% 68,09

Цель:  
отобрать самые освещаемые в СМИ,  
общезначимые, актуальные и вызывающие  
интерес пользователей события.

Я

# Определение веса сюжета

## Вес отдельного сообщения

- Время создания
- Вес источника
- Соответствие тематики сюжета специализации агентства

## Динамика сюжета

- Плотность потока сообщений
- Динамика кликабельности новости на главной странице
- Время жизни в топе

## Интерес пользователей

- Количество новостных запросов к Яндексу
- Записи в блогах
- Место новости на сайтах партнёров

# Вес источника

**Цитируемость** учитывает, насколько часто ссылаются на источник другие новостные ресурсы

**Оперативность** учитывает, насколько часто источник быстро реагирует на события

Вес источников автоматически пересчитывается каждую неделю

# Топ-10 самых цитируемых агентств (в алфавитном порядке)

«Ведомости»

«Интерфакс»

«ИТАР-ТАСС»

«Коммерсант»

«Комсомольская правда»

Lenta.ru

«Первый канал»

РБК

РИА «Новости»

«Российская газета»

По данным еженедельно обновляемого отчёта, формируемого роботом Яндекс.Новостей



# Робот и человек в Яндекс.Новостях

## Человек

1. Принимает решение о сотрудничестве
2. Удаляет ссылку на сообщение из базы данных Яндекс.Новостей в случае, если текст сообщения изменён или отсутствует на сайте источника

## Робот

1. Скачивает и индексирует новостные сообщения
2. Определяет и исключает из поиска по Яндекс.Новостям дубли
3. Определяет жанр сообщения
4. Выделяет ключевые слова и факты
5. Объединяет сообщения в сюжет
6. Определяет рубрику
7. Ранжирует сюжеты

**Спасибо за внимание!**  
**Вопросы?**



**info@news.yandex.ru**

Татьяна Исаева, руководитель службы по  
работе с партнерами Яндекс.Новостей