

РОЧЕВ К. В., ВОЛЧИХИН Д. В.
РАЗРАБОТКА ИНФОРМАЦИОННОГО БОТА ДЛЯ ЧТЕНИЯ RSS-ЛЕНТ
 УДК 004.4:004.9, ВАК 2.3.1, ГРНТИ 50.41.25

Разработка информационного бота для чтения RSS-лент

Development of an information bot for reading RSS feeds

К. В. Рочев, Д. В. Волчихин

K. V. Rochev, D. V. Volchihin

Ухтинский государственный технический университет, г. Ухта

Ukhta State Technical University, Ukhta

Статья посвящена разработке бота для чтения RSS-лент. Поскольку интернет перегружен информацией на самые разные темы, целесообразно использовать такие способы структурирования и фильтрации этих данных, как RSS-канал, который содержит метаданные, представленные в виде одного или нескольких каналов, которые в свою очередь, разбиты на один или несколько пунктов. Очень часто, этот формат используется информационными порталами и блогами. В статье приведено описание проектирования и разработки бота, позволяющего обеспечить возможности, наиболее востребованные при работе с RSS-каналами.

The article describes the development of a bot for reading RSS feeds. Since the Internet is overloaded with information on a variety of topics, it is advisable to use such methods of structuring and filtering these data as an RSS feed, which contains metadata presented in the form of one or more channels, which in turn are divided into one or more items. Very often, this format is used by information portals and blogs. The article describes the design and development of a bot that allows you to provide the capabilities that are most in demand when working with RSS channels.

Ключевые слова: RSS, Технологии, Информационный поток, Новости, Бот, Веб-приложение, React

Keywords: RSS, Technologies, Information Flow, News, Bot, Web Application, React

Введение

Интернет перегружен информацией на самые разные темы, а главная проблема в том, что она повторяется, найти необходимую информацию становится сложнее.

С помощью приложений, которые могут считывать RSS файлы [1, 2], пользователь может подписаться на требуемые ему сайты. Подписавшись на RSS-канал, пользователь сможет отслеживать новости с сайта не посещая его. Приложение проверяет RSS файл, на наличие новых новостей с сайта, если они есть, приложение сразу оповестит об этом пользователя. В таких

приложениях, чаще встраивают функцию сохранения новостных статей, чтобы позже можно было прочитать их, в одном приложении [3, 4].

Цель этой работы – создание системы для автоматизации процесса сбора новостей из разных источников в одно приложение.

Для начала рассмотрим существующие приложения для чтения RSS-лент.

Приложения для чтения RSS-лент

1 Flipboard (flipboard.com)

Описание: Этот сервис подойдёт тому, кто чаще использует многофункциональные приложения. Он представляет из себя социальный журнал, который разработали в 2010 году. Он может составлять «журналы» с ссылками на понравившиеся материалы и приглашать в них соавторов, а также, делиться текстами с друзьями.

На этой платформе, есть много статей о дизайне, разработке, программировании, бизнесе, финансы и не только. Flipboard позволяет читать, как материалы СМИ, так же и мнения отдельно взятых людей. (Рисунок 1).

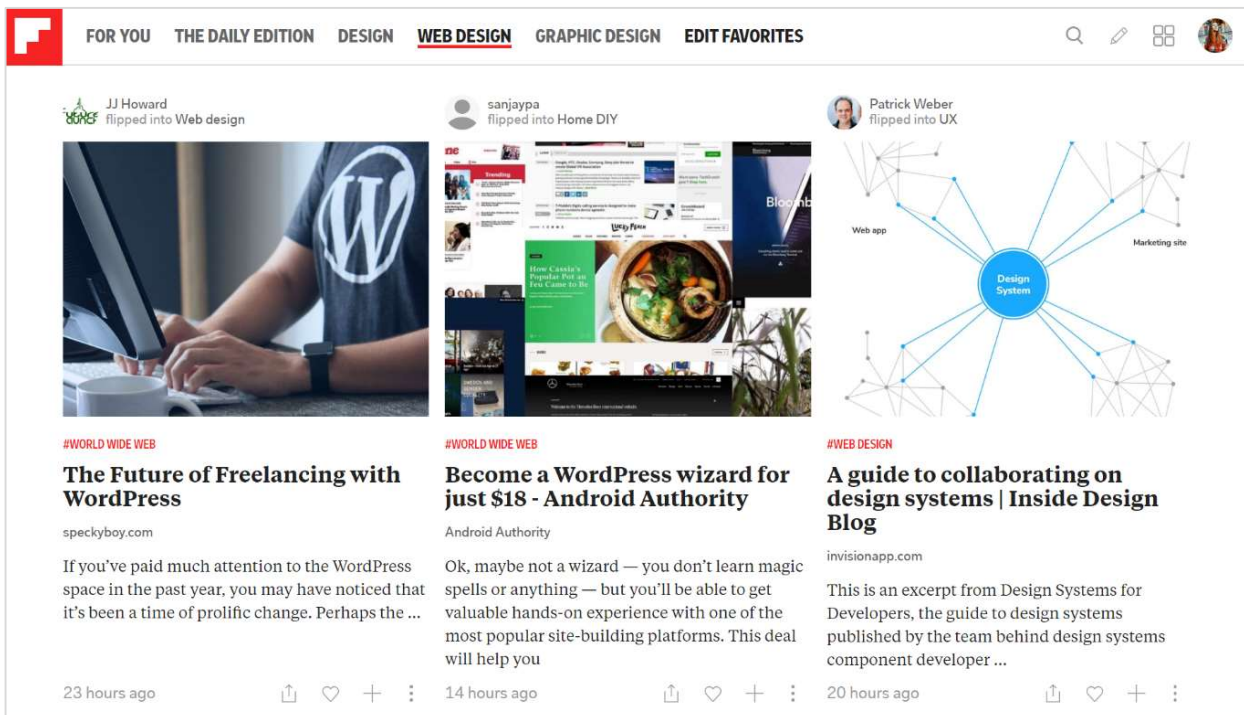


Рисунок 1. Flipboard

Цена: Бесплатно.

Платформы: Android, Windows, IOS, MacOS

2 Feedly (feedly.com)

Это приложение имеет простой и понятный интерфейс, а также, с большим количеством источников. Новостной агрегатор, который был разработан в 2008 году и изначально был предназначен для использования в качестве браузерного расширения. Данное приложение может комментировать и подчеркивать

прочитанное прямо в тексте, а важное подчеркнуть непосредственно на странице.

По категориям и тегам можно настраивать показ статей по ключевым словам. Dropbox, Evernot, OneNote и Pockets синхронизируются с Fedly. По состоянию на данный момент, сервис разрабатывает инструменты на основе искусственного интеллекта для улучшения качества персонализации.

Данный сервис предоставляет бесплатный план, по которому можно отслеживать до 100 каналов (Рисунок 2).

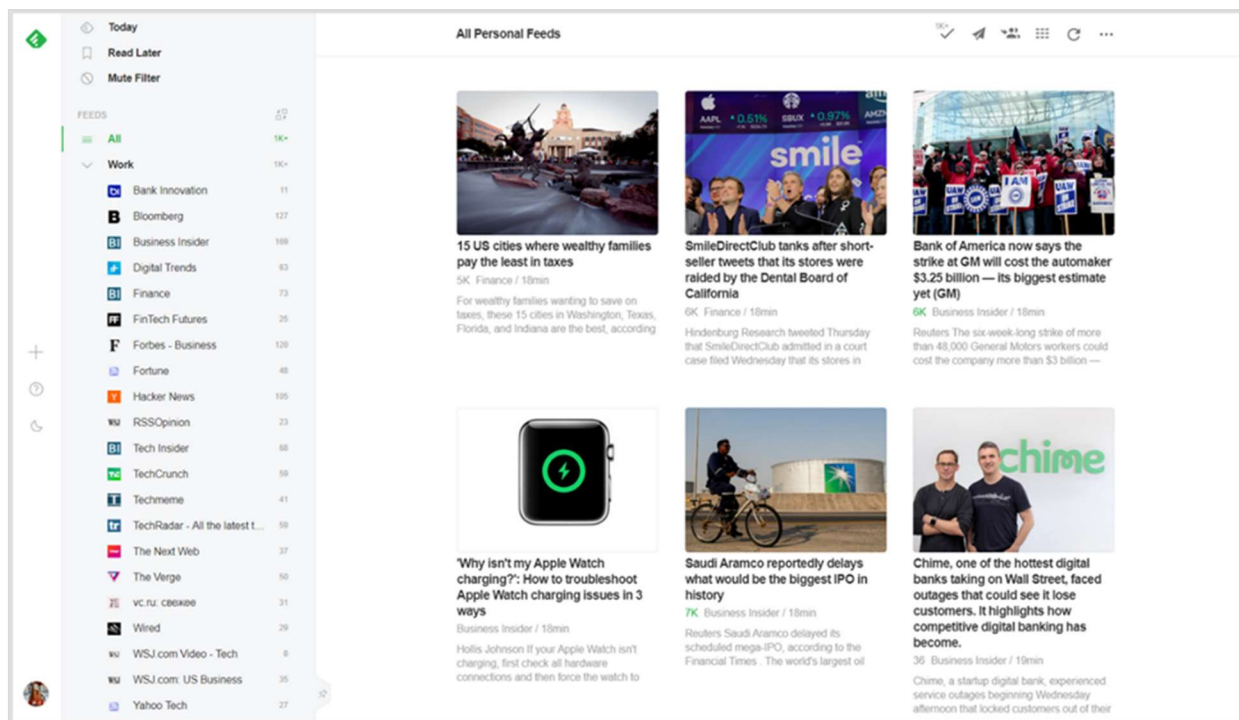


Рисунок 2. Feedly

3 Inoreader (inoreader.com)

С помощью этого приложения, вы сможете общаться с зарубежной прессой и подкастами.

В данном новостном агрегаторе, который придуман компанией Innologica в 2013 году, как замену для Google Reader. Кроме стандартных опций, это приложение поддерживает еще и подкасты. Pro-версия позволяет переводить материалы на другие языки. Пользователи могут также удалять фото к материалам: как в списке статей, так и на самих статьях. (Рисунок 3).

Цена: Базовая версия позволяет подписаться не больше, чем на 150 источников. Платная версия Supporter стоит 1280 рублей (\$19,99), а от бесплатной отличается блокировкой рекламы и большим количеством источников.

Pro-версии составит 320 рублей за месяц или 3200 рублей (\$49,99) за год. Она позволяет искать информацию во всех материалах по ключевым словам и поддерживает перевод статей.

Платформы: iOS, Android

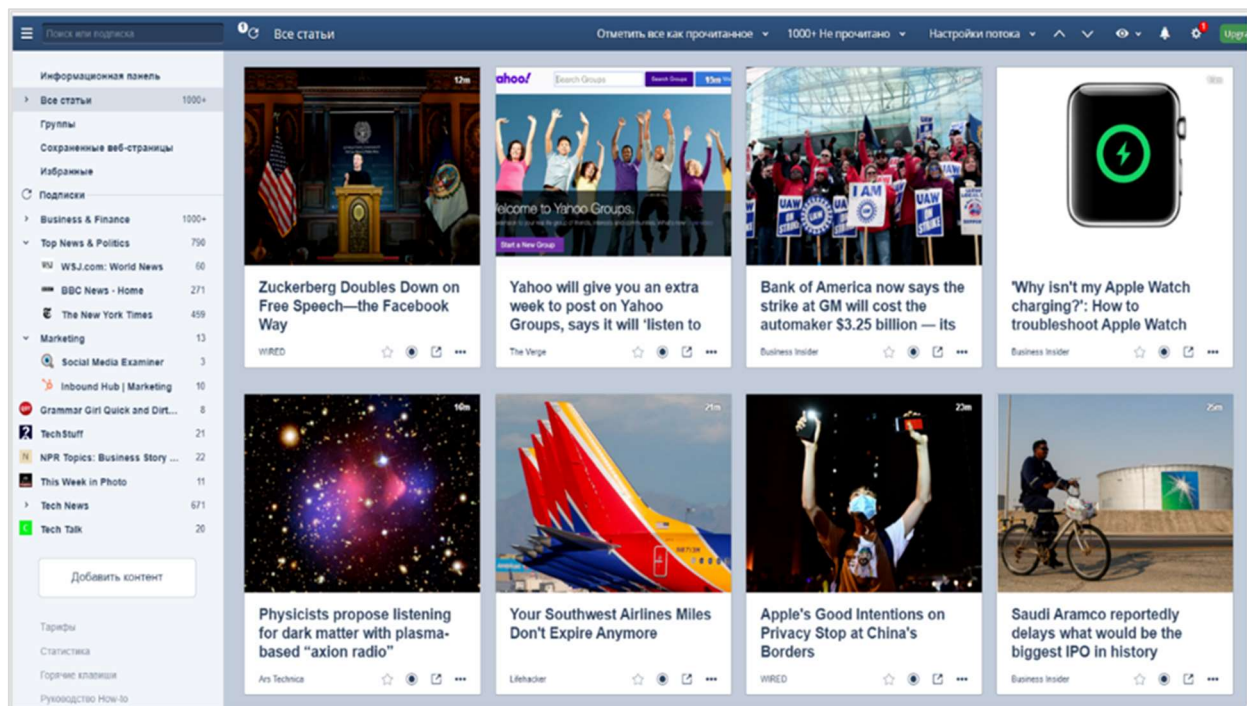


Рисунок 3. Inoreader

4 Instapaper (instapaper.com)

Это приложение создано для тех, кто не может сразу же прочитать все новости. Созданное Марком Арментом приложение, было разработано в 2008 году, спустя десять лет оно стало принадлежать Instant Paper.

В этом приложении есть возможность сохранять не только понравившиеся тексты, но и видео, формируя при этом личную ленту. (Рисунок 4).

Цена: Бесплатно. Расширенная версия стоит 190 рублей (2,99 доллара) в месяц, 1915 рублей (29,99 доллара) в год. Включает в себя неограниченное количество заметок, возможность прослушивания материалов и чтения их в Kindle.

Платформы: iOS, Android

5 Feeder (feeder.co)

Данное приложение подойдет тем пользователям, которые следят за трендами в индустрии дизайна.

Новая платформа, которую создали братья Ротофф из Швеции в 2010 году, для себя, позже стала доступна всем пользователям. Пользователь может изменять дизайн экрана, импортировать RSS-ленты и материалы с любых сайтов. Данный бесплатный вариант обновления обновляется раз в два часа, поэтому для оперативного мониторинга необходимо приобрести подписку. (Рисунок 1.5).

Цена: Pro-версия стоит 320 рублей (4,99 доллара) в месяц, Business-подпись – около 960 рублей (15 долларов) в месяц. По предварительной версии, в тестовом режиме она будет работать две недели.

Платформы: iOS, Android

Расширения: Google Chrome, Mozilla Firefox

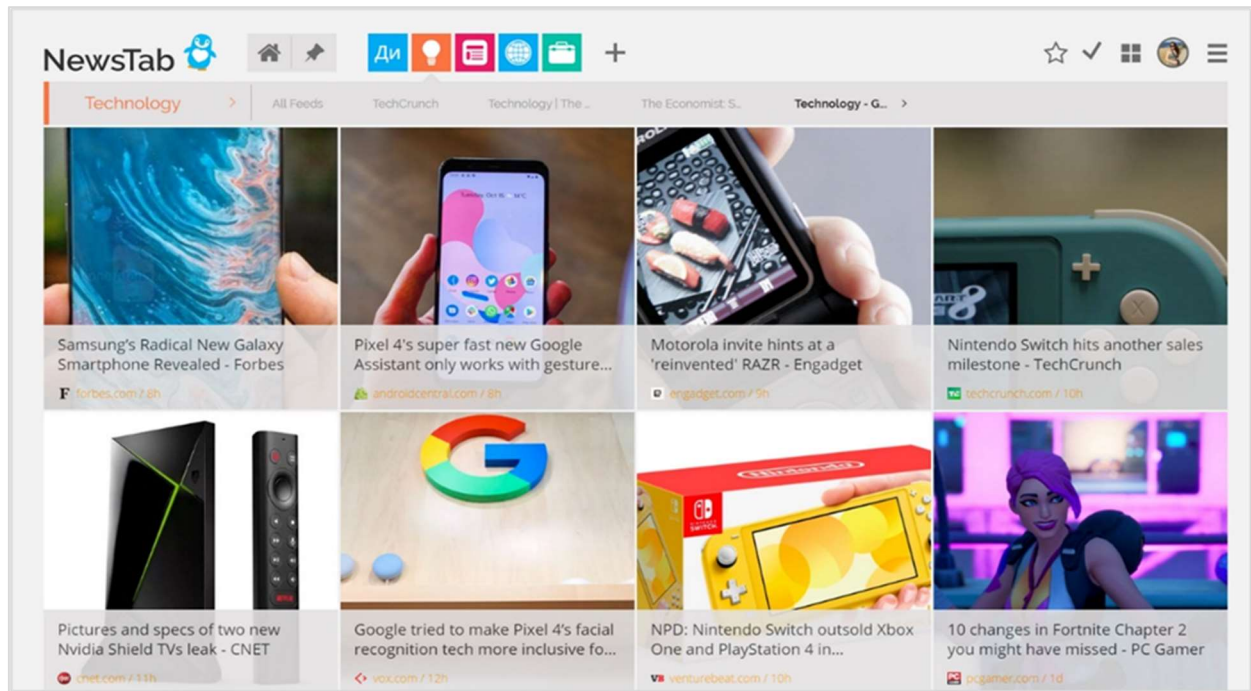


Рисунок 4. Instapaper

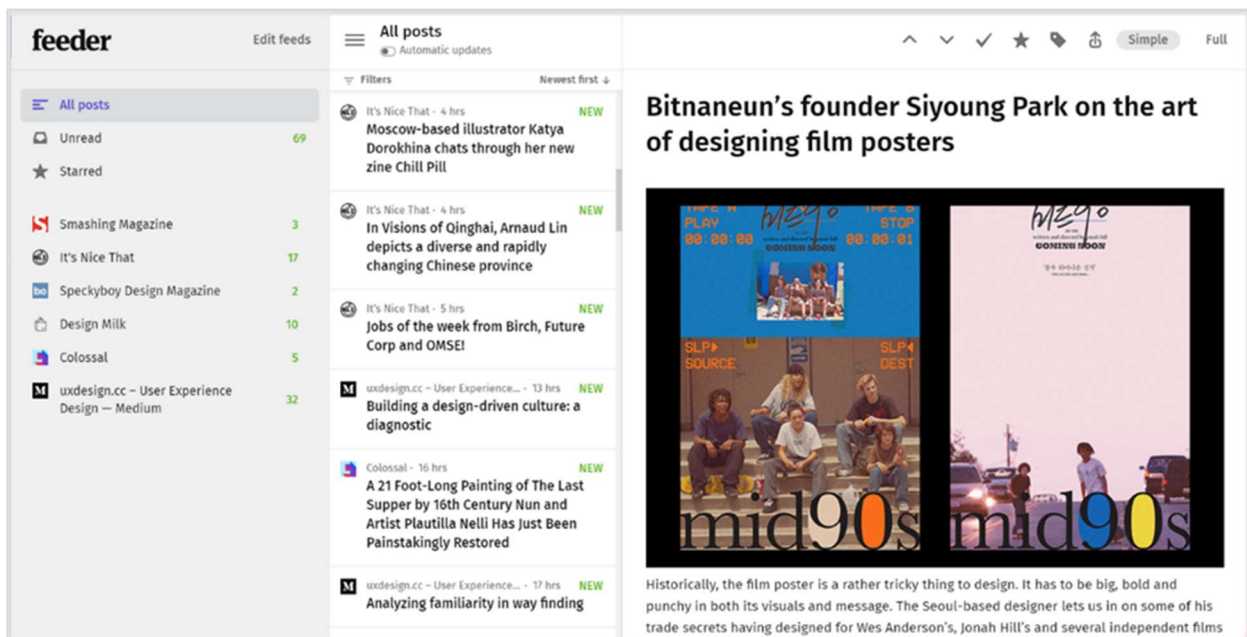


Рисунок 5. Feeder

6 Reeder 5 (reederapp.com)

Это приложение предназначено для владельцев iPhone и Mac, которым удобно работать с агрегатором с минималистичным интерфейсом. (Рисунок 6).

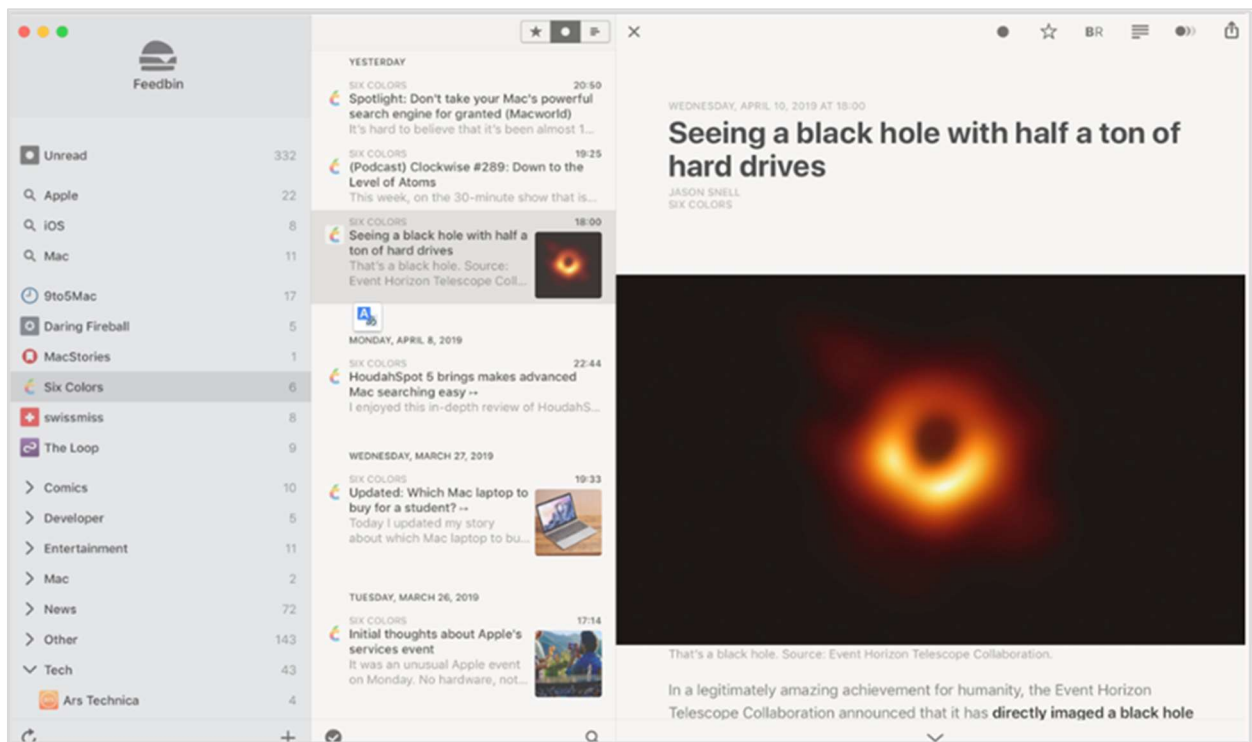


Рисунок 6. Reeder 5

6 Newsblur (newsblur.com)

Данное приложение создано для тех пользователей, которые увлекается программированием.

Этот новостной сервис, позволяет пользователю искать информацию по ключевым словам, во всех текстах в ленте. В платной версии пользователи могут добавлять теги. (Рисунок 7).

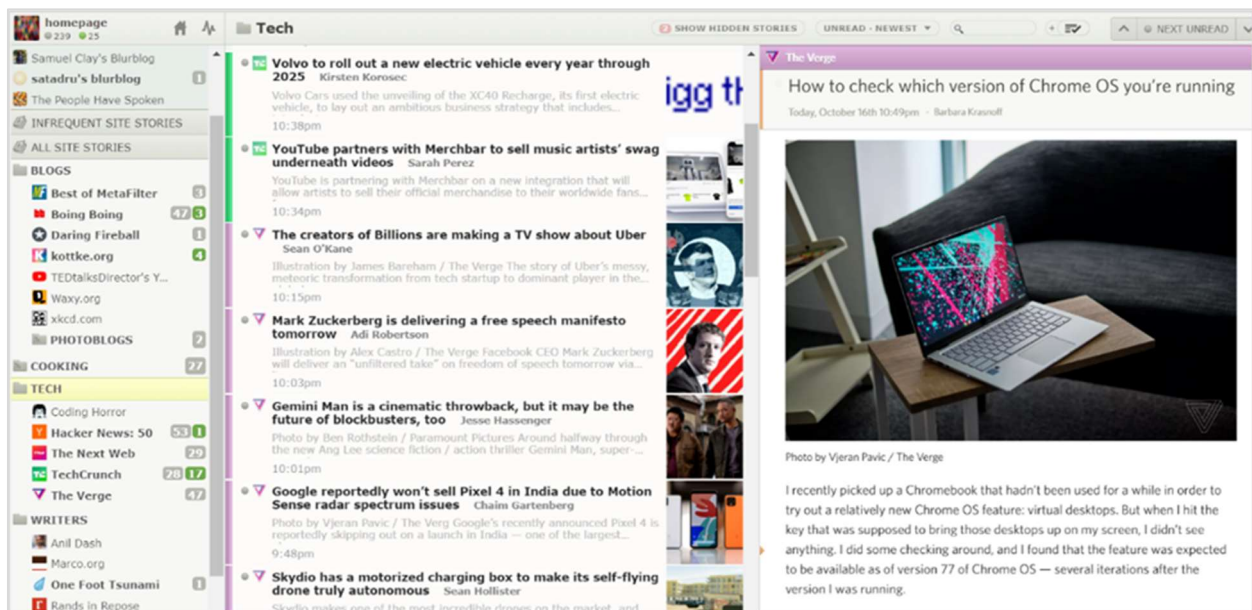


Рисунок 7. Newsblur

После сравнения всех аналогов, можно привести сравнительную таблицу основные и дополнительные функции (Рисунок 8, Рисунок 9).

№	1	2	3	4	5	6	7	8
Название	Inoreader	Feeder	Instapaper	Newsblur	Flipboard	Feedly	Reeder 5	RSS Bot
Возможность добавлять свои источники	+	+	+	+	+	+	\$	+
Возможность поиска по ключевым словам / авторам / тегам	\$	\$	\$	+	+	+	\$	+
Возможность удаления источника	+	+	+	+	+	+	\$	+
Всего функций	3	3	3	3	3	3	3	3

Рисунок 8. Сравнительная таблица основных функций

№	1	2	3	4	5	6	7	8
Название	Inoreader	Feeder	Instapaper	Newsblur	Flipboard	Feedly	Reeder 5	RSS Bot
Возможность переключение языка			+		+	+		+
Кроссплатформенность	+			+		+	\$	+
Возможность поиска RSS канала, по ключевым словам								+
Сохранение в избранное (прочитать потом)		+	+	+			\$	+
Возможность добавления реакции на новость					+			+
Возможность отключения уведомлений	+	+		+			\$	+
Поддержка синхронизации настроек на разных устройствах							+	+
Поддержка полной версии статьи								+
Всего функций	2	2	2	3	2	2	4	8

Рисунок 9. Сравнительная таблица дополнительных требований

Проектирование информационной системы

После изучения предметной области, сравнения аналогов и обоснования необходимости разработки системы, можно построить диаграмму потоков данных (Рисунок 10).

После чего, были выделены процессы, связанные с сущностями, указанными выше: а точнее, для сущности пользователь: задать вопрос, поделиться статьёй, получить полную версию статьи, добавление реакции на новость, добавить / удалить статью из избранного, добавить / удалить тег к подписке, подписаться / отписаться на RSS канал, поиск по подписанным RSS каналам, поиск RSS канала по ключевому слову.

Для сущности администратор были выявлены процессы: ответить на вопрос, заблокировать / разблокировать пользователя.

Для сущности суперпользователь выявлен процесс: установить роль.

И для последней сущности гость, выявлен один процесс - авторизация / регистрация.

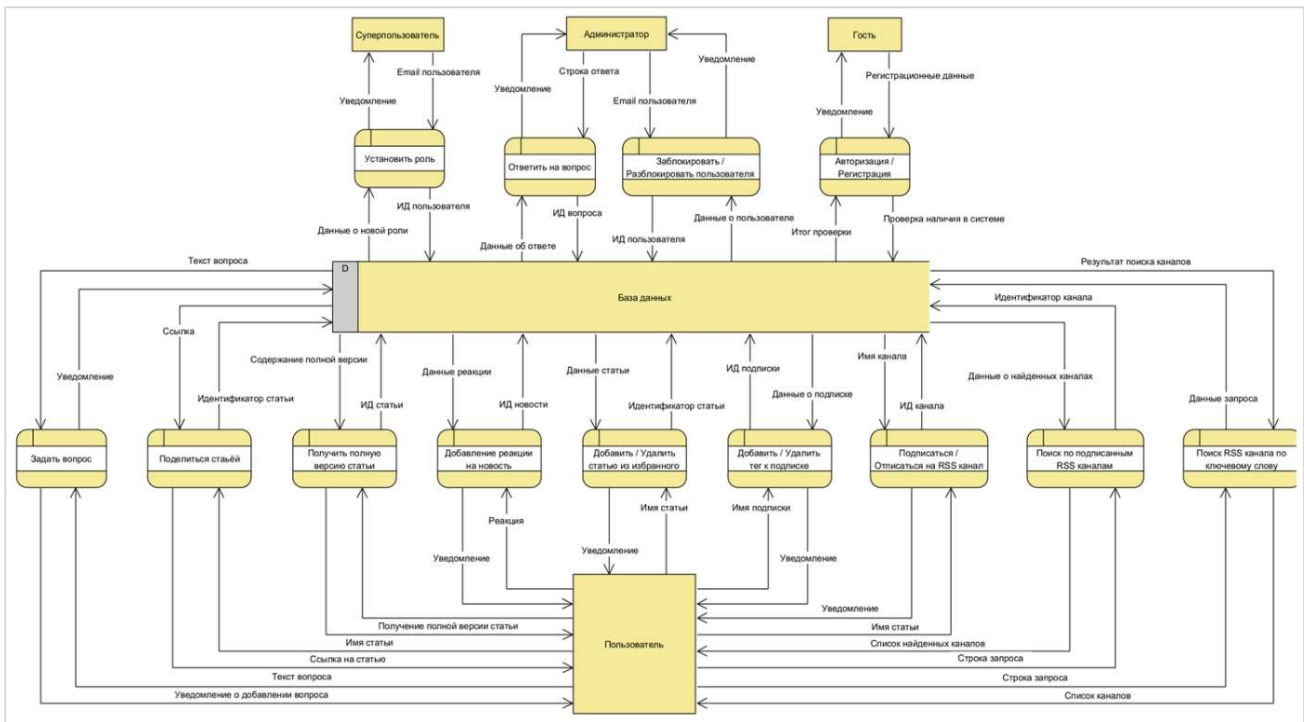


Рисунок 10. Диаграмма потоков данных

Так как в системе требуется хранение данных, была спроектирована и реализована база данных (Рисунок 11).

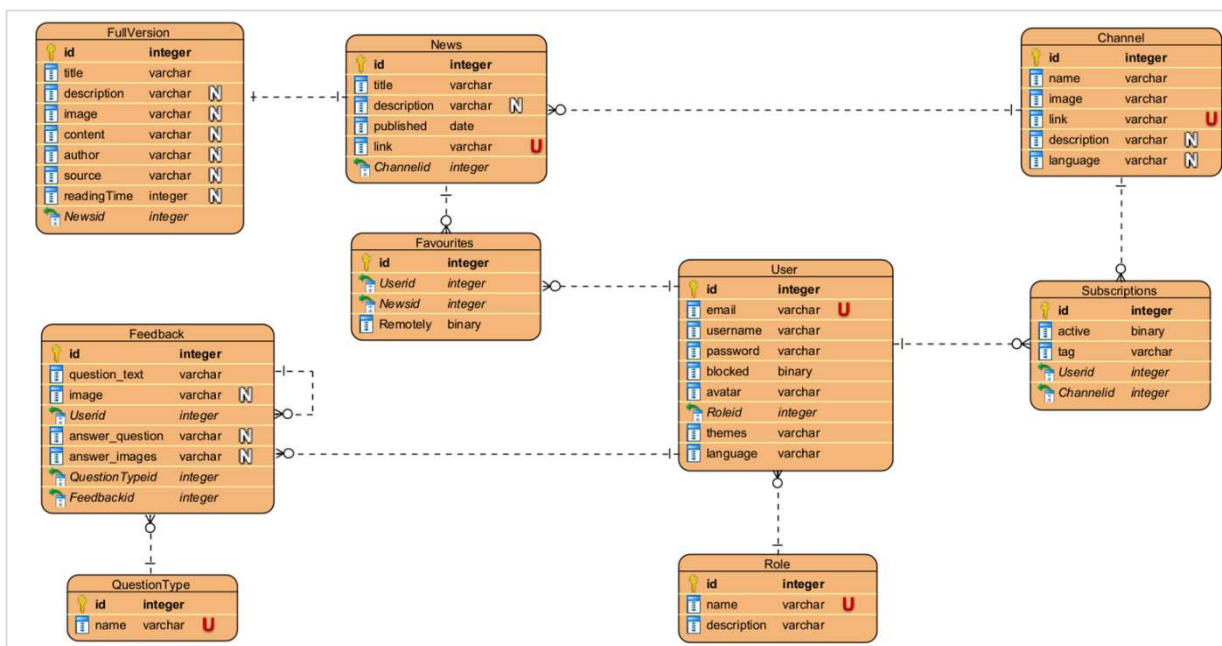


Рисунок 11. Логическая модель базы данных

Результаты разработки системы

В результате получилось реализовать приложение для чтения RSS-лент, на основе технологии podejs [5], которое имеет требуемые критерии:

- Кроссплатформенность;
- Возможность поиска RSS каналов по ключевым словам;

- Сохранение статей в избранное;
- Возможность добавление реакции на новость;
- Возможность отключения уведомлений;
- Поддержка синхронизации настроек на разных устройствах;
- Поддержка полной версии статьи.

Ниже приведены скриншоты основных форм программы.

Поиск каналов по ключевым словам (Рисунок 12)

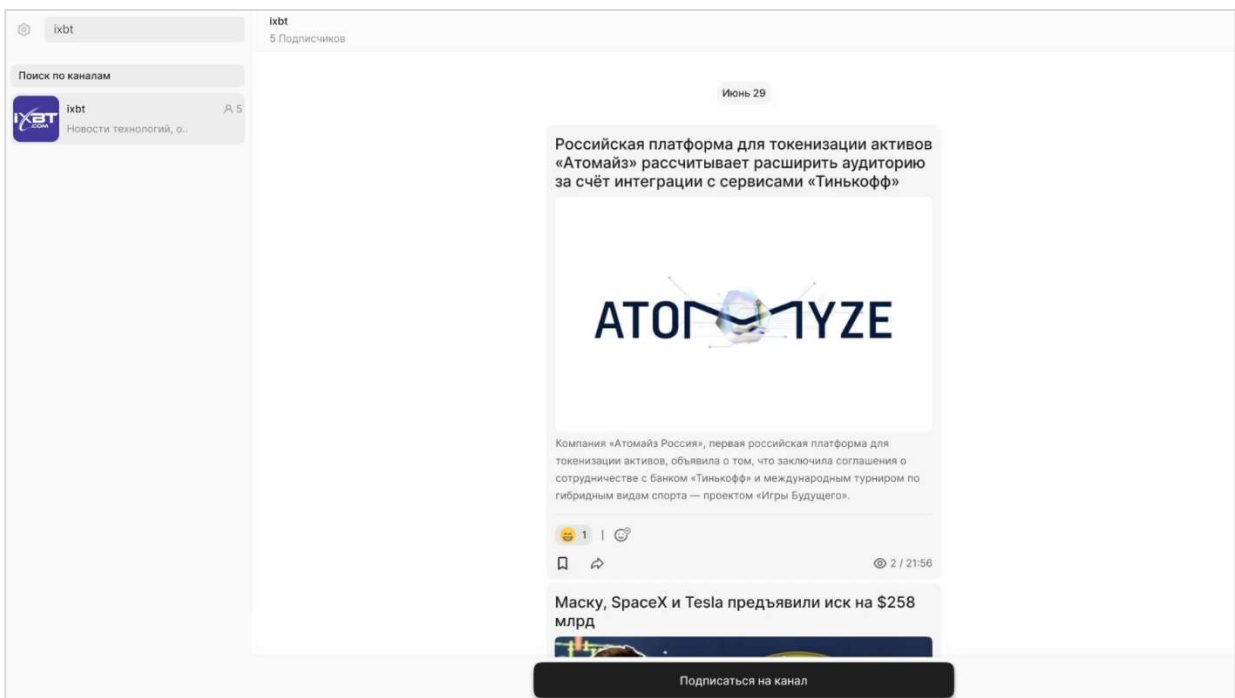


Рисунок 12. Поиск каналов по ключевым словам

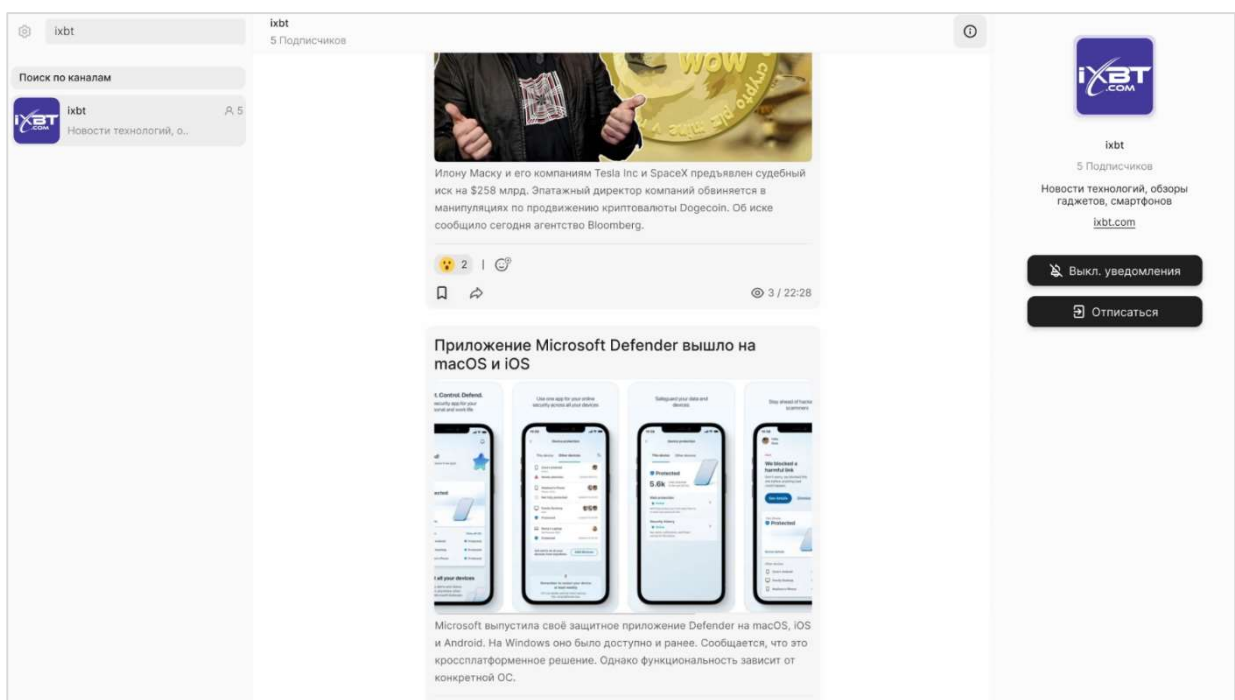


Рисунок 13. О канале

Информация о канале, где можно отписаться от него, а также отключить уведомления (Рисунок 13).

Меню приложения, с помощью которого, можно перейти в настройки или выйти из аккаунта, сменить тему (Рисунок 14).

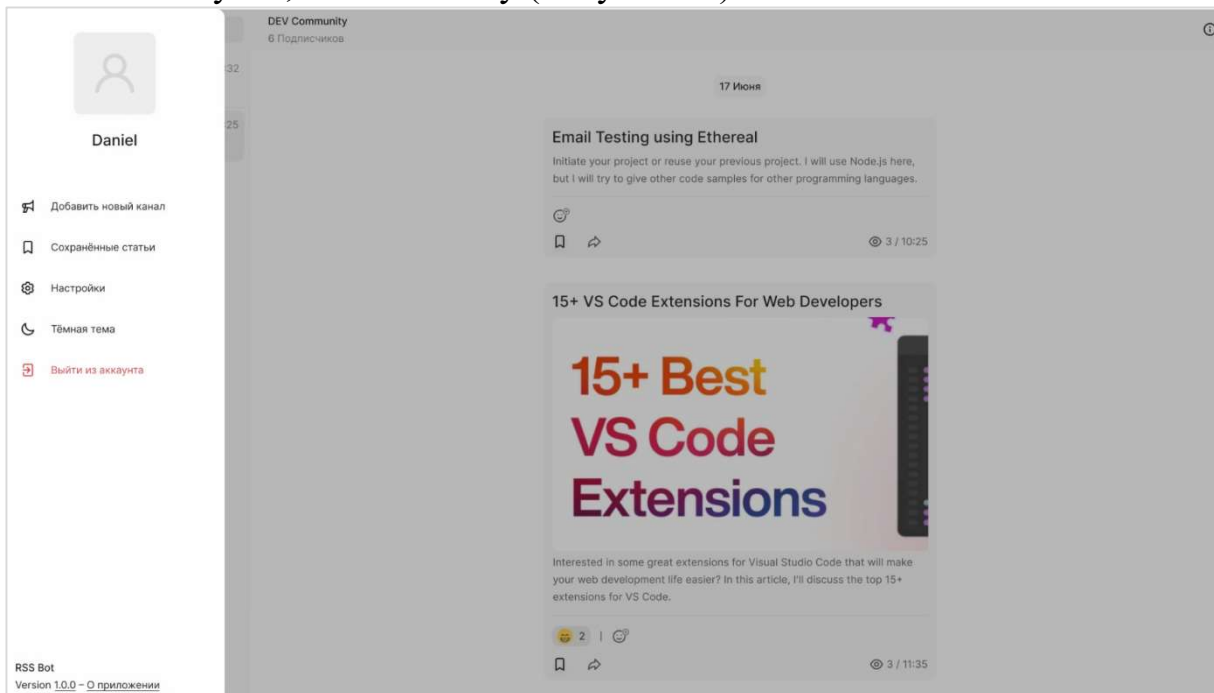


Рисунок 14. Меню приложения

Настройки приложения, с помощью которых, можно редактировать информацию о профиле, сменить язык приложения, сменить тему приложения (Рисунок 15).

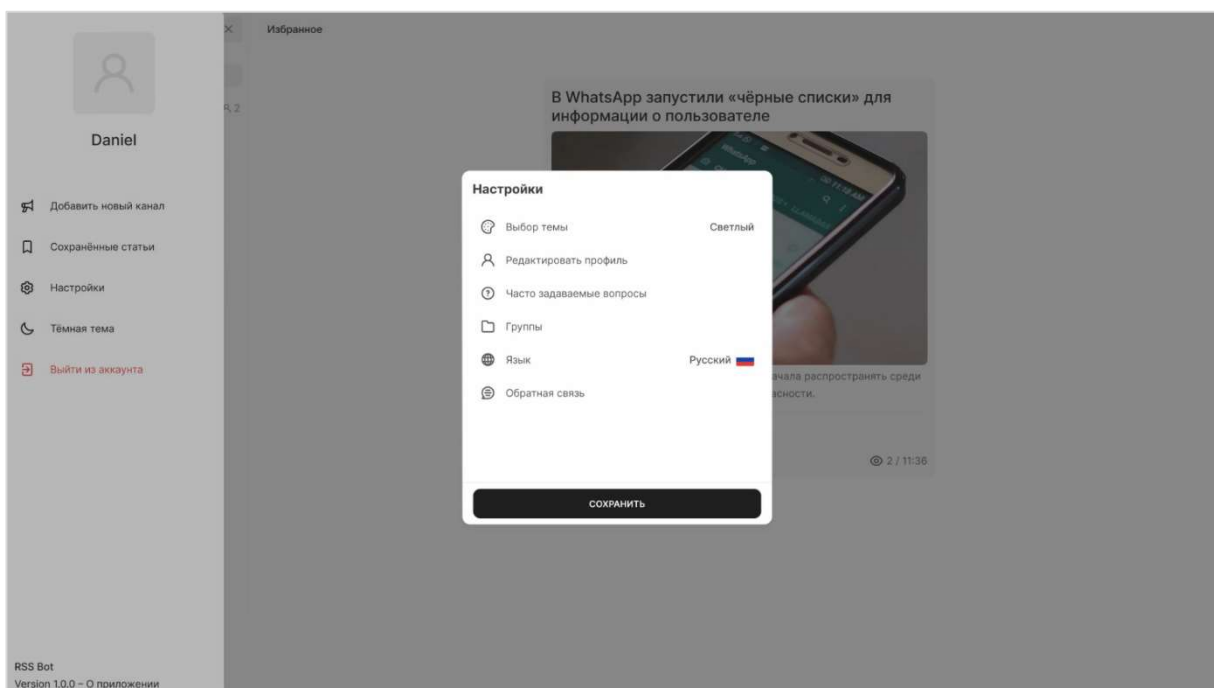


Рисунок 15. Настройки

Заключение

В ходе работы проведено проектирование и реализация информационной системы – бота, позволяющего обеспечить возможности, наиболее востребованные при работе с RSS-каналами, такие как: поиск каналов по ключевым словам, добавление статей в избранное, реакции на новость, включение-отключение уведомлений, синхронизации настроек на разных устройствах, сохранение полной версии статьи.

В перспективе, возможно расширение приложения дополнительными функциями и улучшение уже существующих, с целью повышения эффективности использования данного приложения.

Список использованных источников и литературы

1. Земсков А.И. Что такое RSS? // Научные и технические библиотеки. – 2007. № 6. – С. 4.
2. Герасевич В. Блоги и RSS. Интернет-технологии нового поколения. Санкт-Петербург, – 2006. – 256 с.
3. Храмова Н.Н. Спецификация генерации новостей через RSS на примере работы агрегатора яндекс.новости // Знак: проблемное поле медиаобразования. – 2015. № 3 (17). – С. 36-41.
4. Всё о RSS каналах [Электронный ресурс]. – Режим доступа: <https://www.wpbloging.com/vedenie/chto-takoe-rss.html#hrss> (дата обращения: 18.02.2022).
5. Программная платформа, основанная на движке V8, превращающей JavaScript из узкоспециализированного языка в язык общего назначения - [Электронный ресурс]. – Режим доступа: <https://nodejs.org/en/docs/> (дата обращения: 18.02.2022).

List of references

1. Zemskov A.I. What is RSS? // Scientific and Technical libraries. 2007. No. 6. p. 4.
2. Gerasevich V. Blogs and RSS. New generation Internet technologies. St. Petersburg, 2006. 256 p.
3. Khramova N.N. Specification of news generation via RSS on the example of the Yandex aggregator.news // Sign: the problematic field of media education. 2015. No. 3 (17). pp. 36-41.
4. All about RSS feeds - [Electronic resource]. Mode of treatment: <https://www.wpbloging.com/vedenie/chto-takoe-rss.html#hrss> (accessed: 02/18/2022).
5. A software platform based on the V8 engine that turns JavaScript from a highly specialized language into a general-purpose language - [Electronic resource]. Mode of treatment: <https://nodejs.org/en/docs/> (Accessed: 02/18/2022).