

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
«Южно-Уральский государственный университет (национальный исследовательский университет)»  
Высшая школа электроники и компьютерных наук  
Кафедра системного программирования

# Разработка приложения для интеллектуального анализа отзывов пользователей магазина приложений Google Play

**Автор**

студент группы КЭ-401  
П.И. Шумилин

**Научный руководитель**

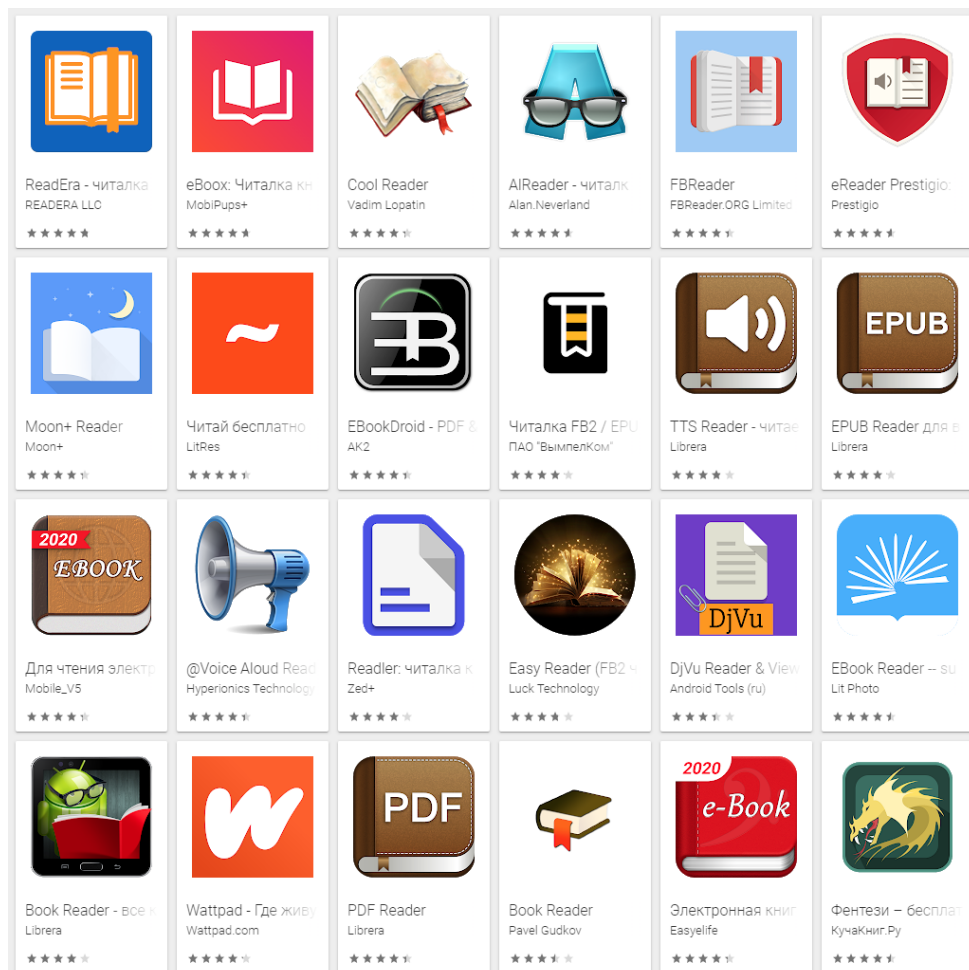
к.ф.-м.н., доцент  
М.Л. Цымблер

**Рецензент**

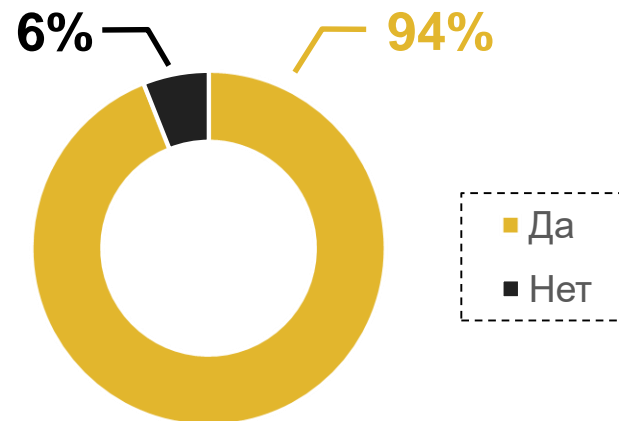
Зам. директора по инф. технологиям  
и инф. безопасности ГБУЗ «ЧОМИАЦ»  
А.С. Староверов

# Актуальность

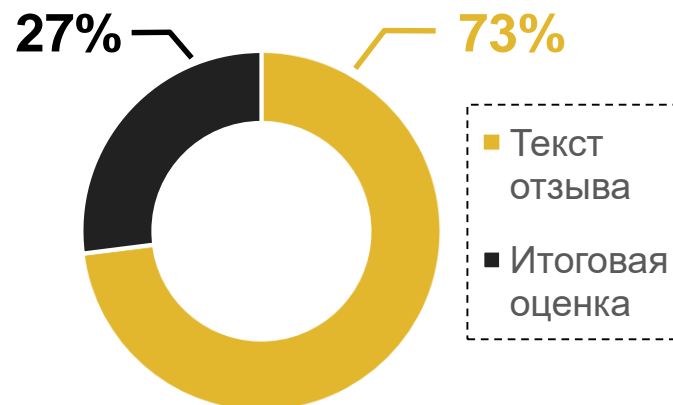
2.5 миллиона приложений













Вы обычно читаете  
текстовые отзывы?

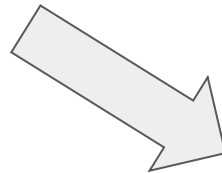
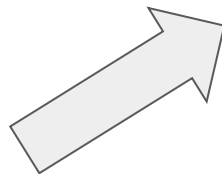


Какие элементы отзыва производят  
на вас наибольшее впечатление?








# Актуальность





-  **Александр Попов**  
★★★★★ 3 марта 2020 г.  
Самый лучший сервис видео связи. Вы молодцы.
-  **Яншин Вячеслав**  
★★★★★ 8 мая 2020 г.  
Виснет изображение и звук идут так себе.
-  **Федя Федя**  
★★★★★ 7 марта 2020 г.  
Что мне нужно есть и работает, я доволен.
-  **Денис Скрыпник**  
★★★★★ 6 мая 2019 г.  
стабильно работает. хорошо слышно
-  **Kim Mindcaster**  
★★★★★ 28 мая 2020 г.  
Кривой интерфейс, глюки звука, тупая система регистрации. Делалось для инопланетян
-  **Никита Кондратенко**  
★★★★★ 17 апреля 2020 г.  
Неконфиденциально, плохое качество связи, отвратительный UX-дизайн
-  **Илья Соломатин**  
★★★★★ 27 ноября 2019 г.  
Победитель номинации "Самый непонятный интерфейс"
-  **Вера Кузьмичёнок**  
★★★★★ 21 ноября 2019 г.  
Очень классно общаться.
-  **Dmitry Mazur**  
★★★★★ 16 февраля 2020 г.  
Звук есть, разберитесь сначала с настройками. И внимательно следите вышли ли вы из программы потому что она работает в фоновом режиме и можно попасть)
-  **Dmitry Z**  
★★★★★ 29 сентября 2018 г.  
Неудобный интерфейс, точнее значки путают где вкл микрофон и выключается звук



# Звук, качество связи

-  **Яншин Вячеслав**  
★★★★★ 8 мая 2020 г.  
Виснет изображение и звук идут так себе.
-  **Денис Скрыпник**  
★★★★★ 6 мая 2019 г.  
стабильно работает. хорошо слышно
-  **Kim Mindcaster**  
★★★★★ 28 мая 2020 г.  
Кривой интерфейс, глюки звука, тупая система регистрации. Делалось для инопланетян
-  **Никита Кондратенко**  
★★★★★ 17 апреля 2020 г.  
Неконфиденциально, плохое качество связи, отвратительный UX-дизайн
-  **Dmitry Mazur**  
★★★★★ 16 февраля 2020 г.  
Звук есть, разберитесь сначала с настройками. И внимательно следите вышли ли вы из программы потому что она работает в фоновом режиме и можно попасть)

# Интерфейс, UX-дизайн

-  **Kim Mindcaster**  
★★★★★ 28 мая 2020 г.  
Кривой интерфейс, глюки звука, тупая система регистрации. Делалось для инопланетян
-  **Никита Кондратенко**  
★★★★★ 17 апреля 2020 г.  
Неконфиденциально, плохое качество связи, отвратительный UX-дизайн
-  **Илья Соломатин**  
★★★★★ 27 ноября 2019 г.  
Победитель номинации "Самый непонятный интерфейс"
-  **Dmitry Z**  
★★★★★ 29 сентября 2018 г.  
Неудобный интерфейс, точнее значки путают где вкл микрофон и выключается звук

# Цель и задачи

## Цель

Разработка приложения для интеллектуального анализа отзывов пользователей магазина приложений Google Play

## Задачи

1. Провести обзор аналогов систем анализа отзывов
2. Разработать подход к анализу отзывов
3. Выполнить проектирование архитектуры системы
4. Реализовать систему
5. Провести эксперименты, исследующие эффективность предложенного подхода

# Обзор аналогов

## Нативное решение



Google Play Console

Регистрация: 25\$

## Сторонние сервисы

APPFOLLOW

 **apptweak**

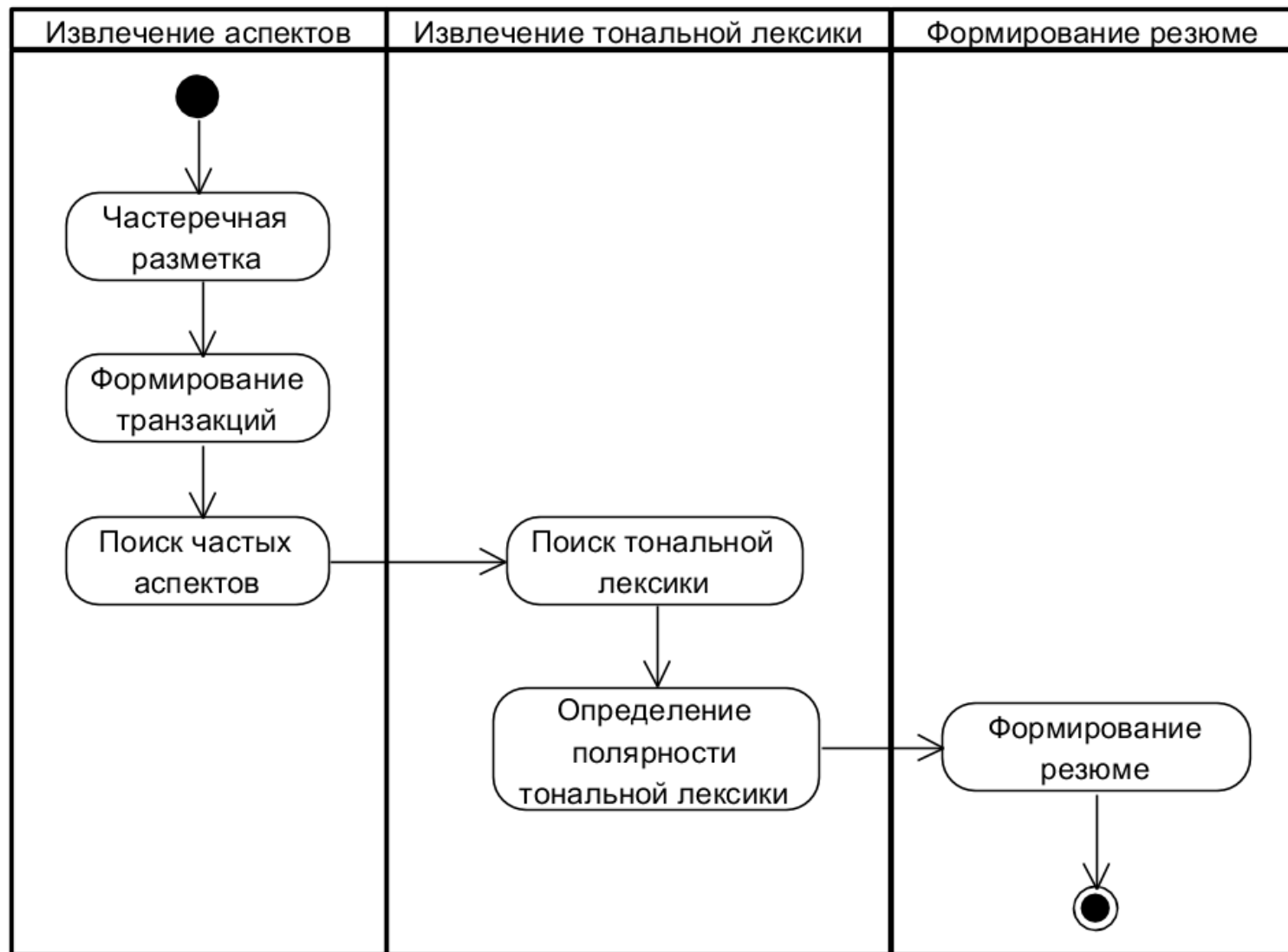
 **appbot**

Подписка: от 40\$ / месяц

## Недостатки:

- Не выполняют анализ отзывов на русском языке
- Не предназначены для конечного пользователя
- Платные

# Подход к анализу отзывов



# Частеречная разметка

Исходный  
отзыв:



Неплохое приложение! Приятный дизайн, удобно работать.  
Единственный замеченный недостаток это медленный поиск.  
Надеюсь, ситуация улучшится в следующих обновлениях.

Обработанный  
отзыв:

неплохое приложение приятный дизайн удобно работать  
единственный замеченный недостаток это медленный поиск  
надеюсь ситуация улучшится в следующих обновлениях

Размеченный  
отзыв:

неплохое приложение приятный дизайн удобно работать  
единственный замеченный недостаток это медленный поиск  
надеюсь ситуация улучшится в следующих обновлениях

	- ПРИЛ		- СУЩ		- ПР
	- НАР		- ГЛ		
	- ПРИЧ		- МЕСТ		

# Формирование транзакций

Размеченный  
отзыв:

неплохое приложение приятный дизайн удобно работать  
единственный замеченный недостаток это медленный поиск  
надеюсь ситуация улучшится в следующих обновлениях



Транзакция:

приложение дизайн недостаток поиск ситуация обновление



- ПРИЛ



- СУЩ



- ПР



- НАР



- ГЛ



- ПРИЧ



- МЕСТ

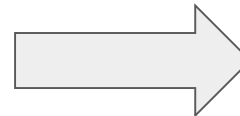


# Поиск частых аспектов

## Алгоритм Apriori

Минимальная поддержка: **5%**

№	Транзакция
1	<b>приложение дизайн</b> недостаток <b>поиск</b> ситуация обновление
2	молодец почтовик сервис <b>дизайн</b> тема <b>приложение</b> видео проект компания куча человек будущее леса технология удобство ряд
3	день <b>приложение</b> загрузка диалог написание письмо уведомление
4	реклама разработчик <b>приложение</b> комментарий очередь факт <b>дизайн</b> почта письмо список взгляд списание средство клик увеличение отказ космос
...	...



### Аспекты:

1. Приложение
2. Дизайн
3. Поиск
4. ...

# Поиск тональной лексики

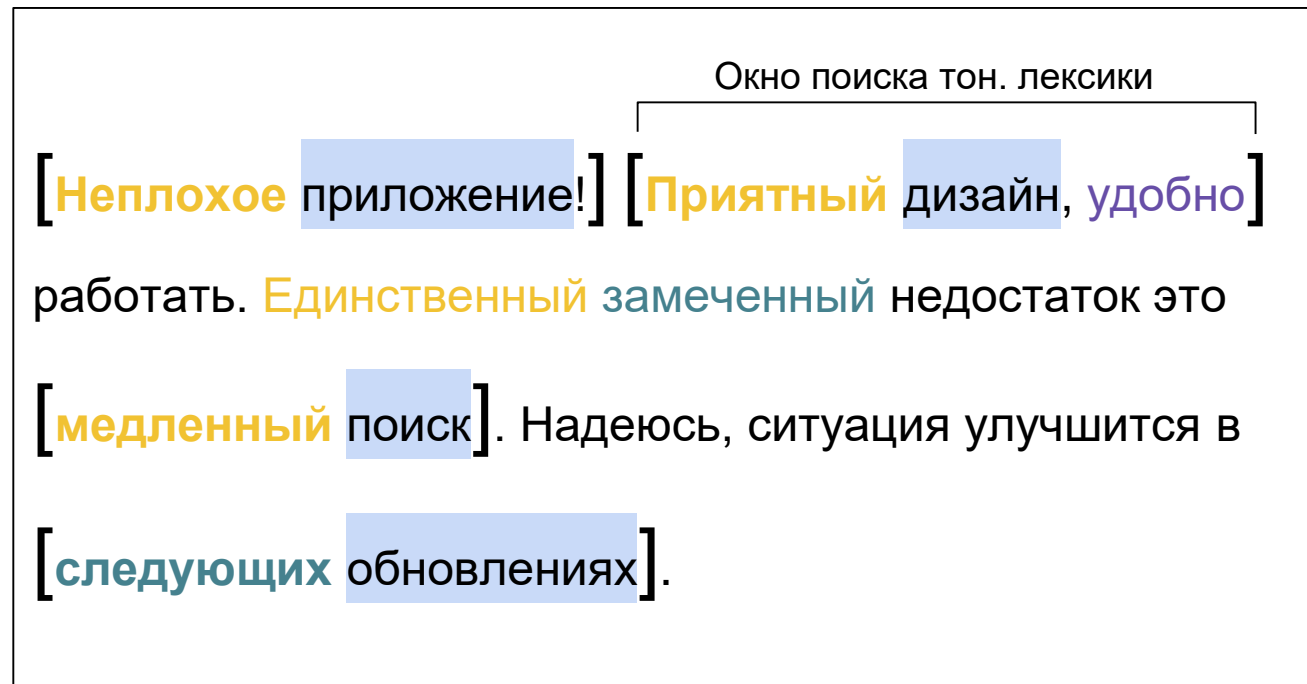
## Аспекты:

1. Приложение
2. Дизайн
3. Поиск
4. ...

Радиус окна поиска  
оценочной лексики: 1

## Тональная лексика:

1. Неплохой
2. Приятный
3. Медленный
4. ...

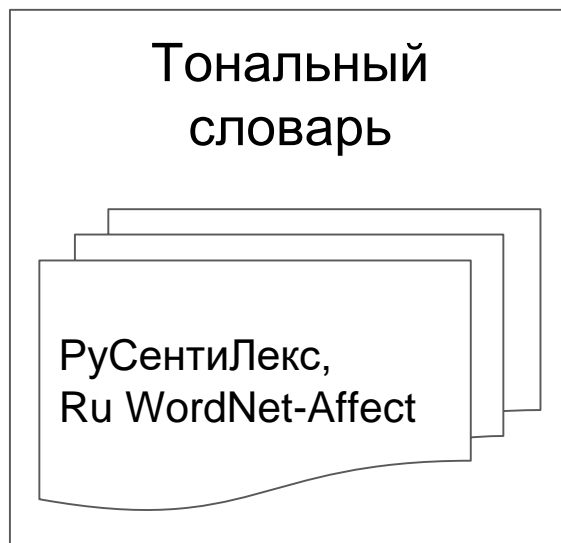
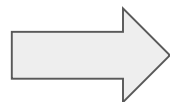


- ПРИЛ
- НАР
- ПРИЧ

# Определение полярности тональной лексики

## Тональная лексика:

1. Неплохой
2. Приятный
3. Медленный
4. ...



## Положительная полярность:

1. Неплохой
2. Приятный
3. Классный
4. ...

## Отрицательная полярность:

1. Медленный
2. Ужасный
3. Скучный
4. ...

## Не окрашенная лексика:

1. Следующий
2. Круглый
3. ...

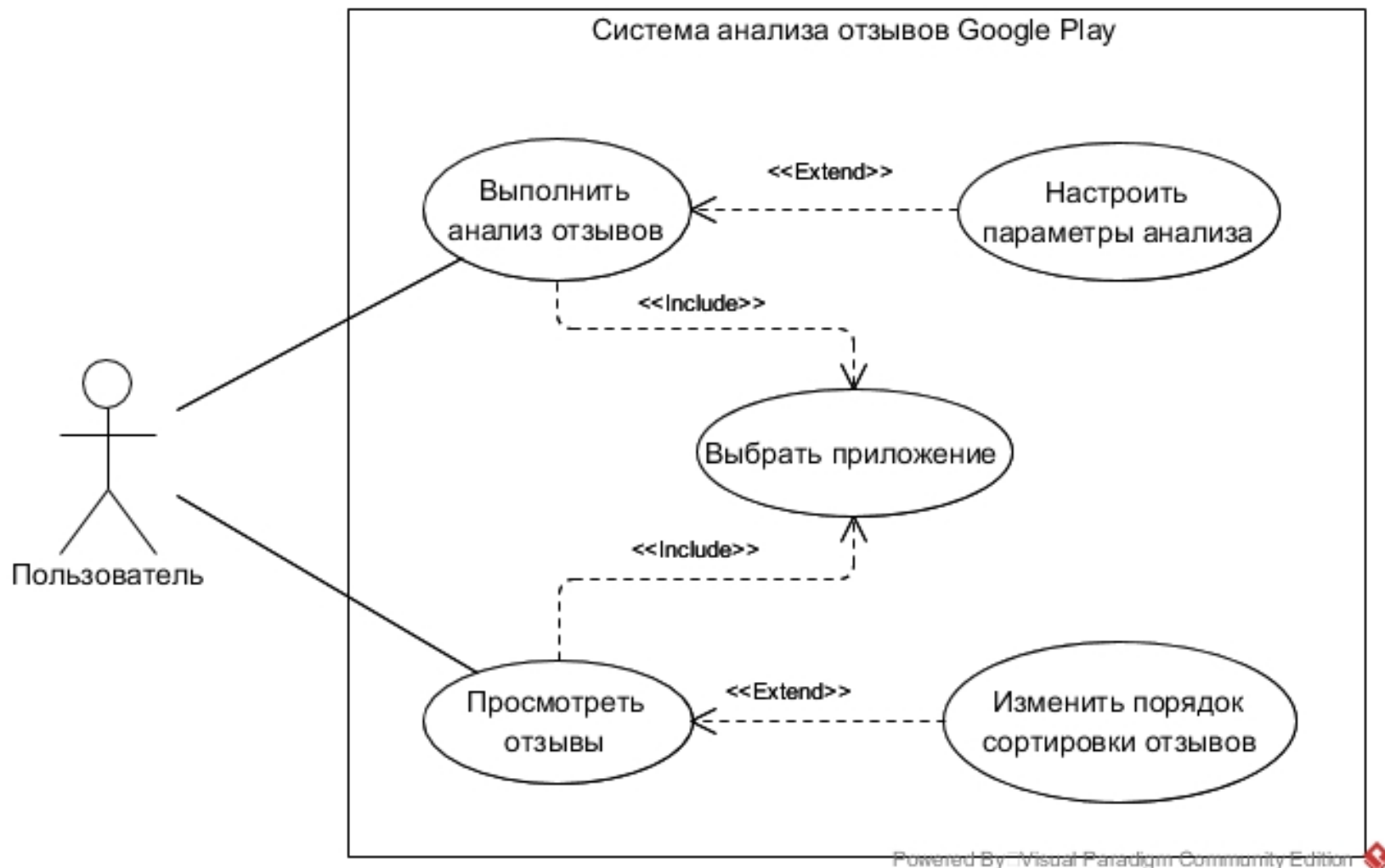
# Формирование резюме

ID	Текст отзыва
10	Неплохое приложение! Приятный дизайн, удобно работать. Единственный замеченный недостаток это медленный поиск. Надеюсь, ситуация улучшится в следующих обновлениях.
58	Ужасное приложение. Дизайн устаревший, пользоваться не хочется. Поиск очень медленный да и вообще работает приложение нестабильно, постоянные вылеты.
...	...
75	Классное обновление, наконец-то поправили раздражающую ошибку с синхронизацией и добавили возможность настроить размер шрифта.

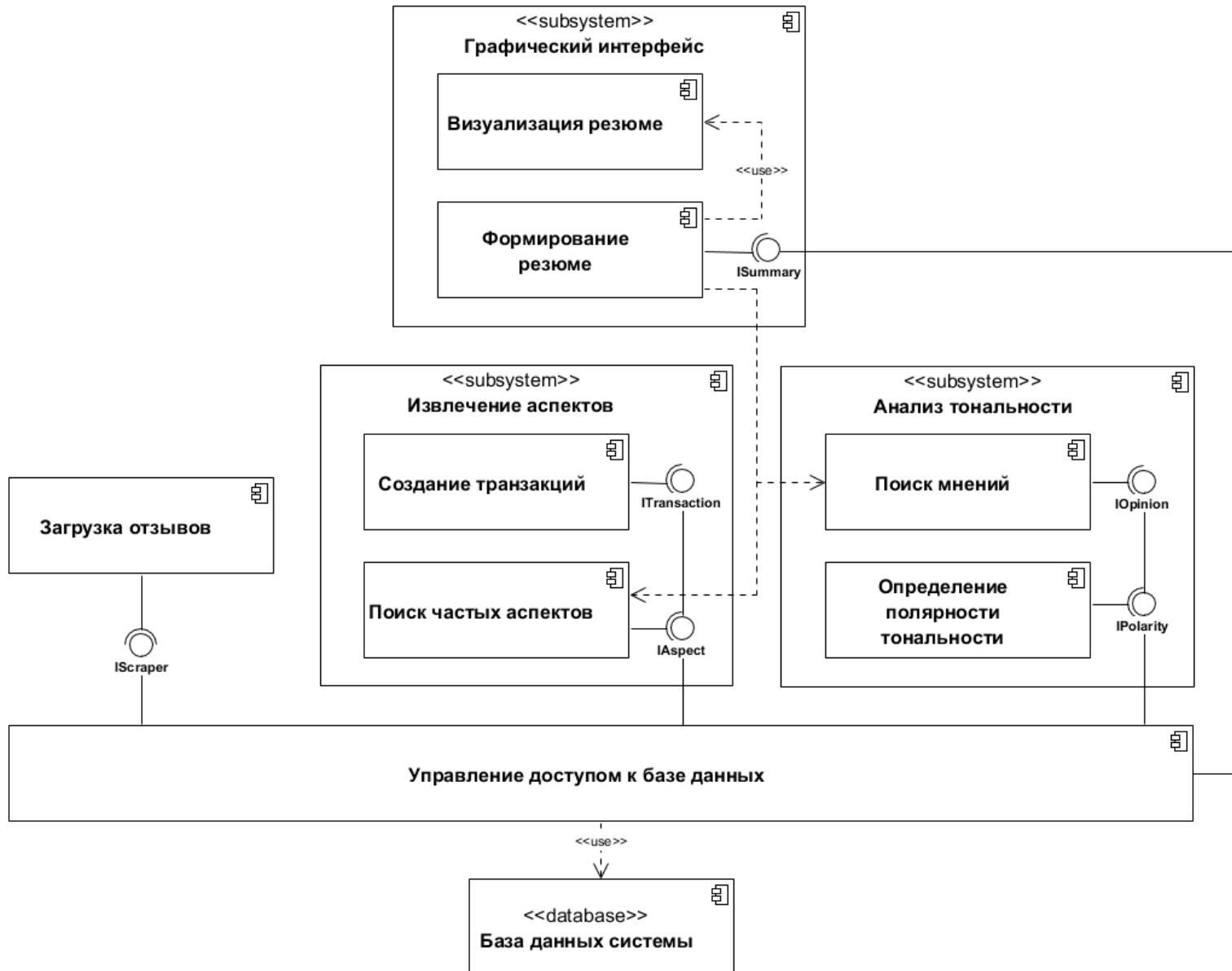
## Резюме

Аспект	Положительно	Отрицательно
Приложение	ID: 10, ...	ID: 58, ...
Дизайн	ID: 10, ...	ID: 58, ...
Поиск	...	ID: 10, 58, ...
Обновление	ID: 75, ...	...
...	...	...

# Варианты использования



# Архитектура системы



# Реализация

**Язык программирования:** Python

## **Библиотеки**

- Загрузка отзывов: Selenium
- Обработка естественного языка: nltk, rymorphy2
- Поиск частых наборов: mlxtend.frequent\_patterns.apriori, pandas
- База данных: pymongo
- Графический интерфейс: PyQt5

**СУБД:** MongoDB

Исходные коды доступны по адресу <https://github.com/ShumilinPavel/diploma>

# Графический интерфейс

Просмотр отзывов

Приложение: Яндекс.Почта  
Аспект: интерфейс (35) Параметры...

Положительные  Отрицательные  Нейтральные

	Автор	Оценка	Дата	Полезность	Текст
20	Виталий Евгеньевич	3	2019-04-24	0	из плюсов - подтягивает тему оформления, довольно удобный интерфейс, доступ по пину или отпечаткам, из минусов: нет возможности просмотра всех сообщений по всем ящикам, не умеет архивировать сообщения в gmail ящиках (аутлук умеет) и перестал открывать сообщения из нотиса - уведомляет - открываем и бесконечно грузит это сообщение, хотелось бы еще фильтровать почту клиентом, мечты... пока удаляю
21	Сергей Шаров	5	2019-04-09	1	Удобное приложение с приятным интерфейсом и главное - выключатель рекламы, хочешь поддержать Яндекс? оставь рекламу, отвлекает или мешает реклама? включи. И никакого доната для этого не надо, очень понравился такой подход
22	Nikita Lavrishev	5	2019-04-02	1	Всегда пользовался gmail ом, яндекс ничем не уступает, интерфейс удобный и интуитивный, очень доволен качеством работы.
23	Vlada Pak	4	2019-03-16	0	Пока что всем довольна, вроде никаких багов нет, но вот разочаровало то, что при темной теме не все становится темным, часть остаётся белым, ну ладно, в дальнейшем может быть с этим что-то сделают. Интерфейс красивый и удобный, спасибо. Удачи в дальнейших разработках 🍀
24	Сергей Горбунов	5	2019-02-28	3	Хорошее приложение, понятный интерфейс, оперативная техподдержка. Хорошая мобильная версия браузерного почтового клиента. Удобно работать с метками. Почему-то не выходятся уведомления о почте на экран блокировки (возможно, особенности моего телефона). Спасибо разработчикам.
25	Артём Матвеев	5	2019-02-18	1	Удобная почта, есть синхрон со всеми сервисами Яндекса и даже с другими почтами. Интерфейс милый, всё сделано удобно и красиво. И... там есть кнопочка бесплатного отключения рекламы! В чем подвох, Яндекс?))
26	Марина Полякова	5	2019-02-13	1	Очень удобное приложение! всегда выручает, так как письма нужно отслеживать 24 часа в сутки! Прекрасный интерфейс, легкость в использовании, минусов нет или не нашла, а если не нашла, значит, их и нет! И в этот раз Яндекс большие молодцы!
27	Ivan Sharapenkov	4	2019-01-31	0	Интерфейс удобный, возможность связать аккаунты разных почтовых сервисов. НО: приложение автоматически не обновляет список писем и не приходит уведомления. В gmail и vk, например, все приходит автоматически. В чем может быть причина? Устройство OnePlus 6t
28	Серафим Иванов	5	2019-01-15	1	Яндекс, как всегда, не полкачал. Приятный интерфейс, богатый функционал, возможность собрать почту из других ящиков и защитить приложение пин кодом. Супер, благодарю!
	Максим				Отличный клиент, продуманный интерфейс, качественная работа. Очень не хватает общей виртуальной папки. Выполнив гл. почта с разных аккаунтов показывалась бы

Окно просмотра отзывов

Параметры а...

Параметры для поиска аспектов

Минимальная поддержка, %

Параметры для поиска мнений

Радиус окна поиска тон. лексики

Выполнить анализ Отмена

Окно параметров



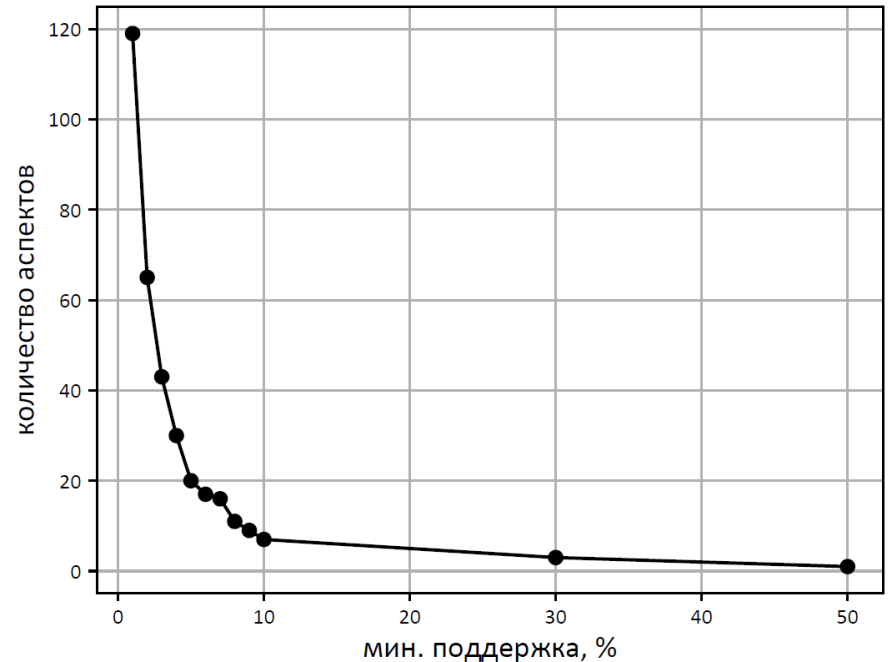
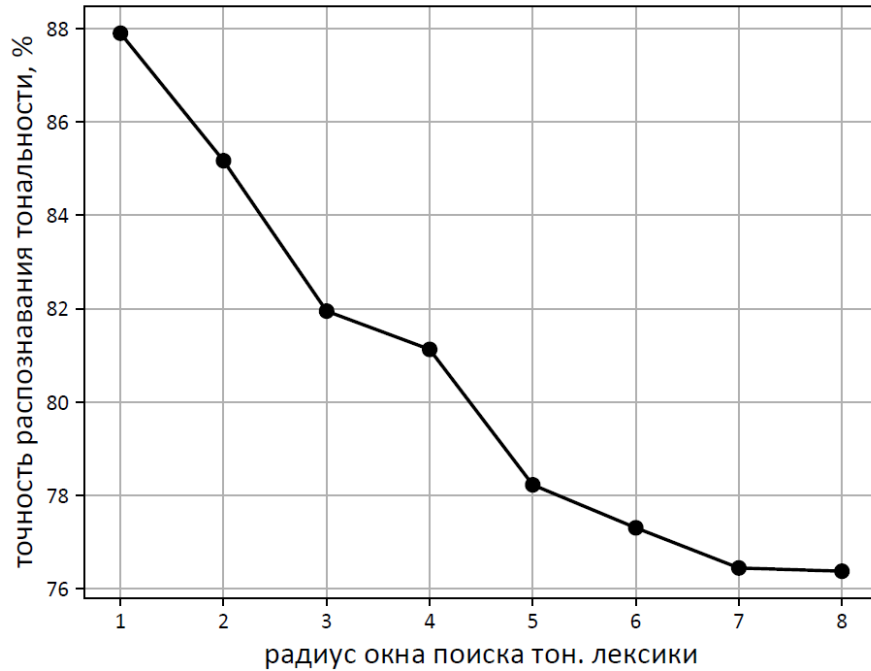
# Эксперименты

Минимальная поддержка	Полнота (recall)	Точность (precision)
30%	50.69%	76.04%
10%	49.30%	81.60%
3%	44.50%	80.20%

Набор данных: 500 отзывов

$$recall = \frac{correctRecognised}{total}$$

$$precision = \frac{correctRecognised}{totalRecognised}$$



# Основные результаты

1. Разработан подход к анализу отзывов, основанный на поиске частых аспектов, тональной лексики по отношению к аспектам и определении ее полярности
2. Выполнено проектирование архитектуры системы
3. Выполнена реализация системы
4. Проведены эксперименты, исследующие эффективность предложенного подхода к анализу отзывов