**Министерство науки и высшего образования Российской Федерации**

**Федеральное государственное автономное образовательное учреждение**

**высшего образования**

**«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И

ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

КАФЕДРА АНАЛИЗА ДАННЫХ И ТЕХНОЛОГИЙ ПРОГРАММИРОВАНИЯ

Направление: 09.03.03 – Прикладная информатика

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

**Экранная клавиатура для мобильных устройств с поддержкой лингвистических сервисов**

**Работа завершена:**

Студент 4 курса

группы 09-951

“\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_ 2023 г.    \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Фархетдинов Р.Р.

**Работа допущена к защите:**

Научный руководитель

ученая степень, ученое звание,

должность

“\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_ 2023 г.    \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Сафина Л.И.

Заведующий кафедрой

ученая степень, ученое звание

“\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_ 2023 г.    \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Вахитов Г.З.

Казань – 2023

Содержание

[Введение 3](#_Toc132051151)

[1. Формирование требований 5](#_Toc132051152)

[1.1. Исследование предметной области лингвистических сервисов 5](#_Toc132051153)

[1.2. Анализ существующих клавиатур с поддержкой лингвистических сервисов 8](#_Toc132051154)

[1.3. Техническое задание 10](#_Toc132051155)

[2. Проектирование сервиса-клавиатуры 13](#_Toc132051156)

[2.1. Проектирование баз данных 13](#_Toc132051157)

[2.2. Проектирование диаграмм сервиса-клавиатуры 13](#_Toc132051158)

# Введение

История всех экранных клавиатур начинается с 1992 года — года создания первой оптической виртуальной клавиатуры, изобретенной и запатентованной инженерами IBM. Хотя данный вид клавиатуры появился относительно недавно, история её развития достаточно обширна.

Первые виртуальные клавиатуры не обладали возможностью поддерживать какие-либо лингвистические сервисы. Чуть позже появились клавиатуры с данной возможностью. Но с покупкой нового устройства (смартфона, компьютера) все данные о стиле печати пользователя оставались на старых устройствах. Отличным решением было бы сохранять данные пользователя в «облаке», чтобы избавиться от зависимости от определенного устройства. В эпоху Интернета синхронизация данных в сети является неотъемлемой частью большинства людей.

Целью выпускной квалификационной работы является разработка экранной клавиатуры для мобильных устройств с поддержкой лингвистических сервисов. Такая клавиатура нацелена помогать любым пользователям набирать текст на мобильных устройствах максимально продуктивно: быстро, без ошибок, с индивидуальным «почерком» пользователя. Разработка экранной клавиатуры актуальна, потому что с появлением социальных сетей, мессенджеров и сенсорных телефонов люди во всем мире стали активно переписываться друг с другом целыми днями, тем самым клавиатура является необходимым элементом в жизни любого современного человека. Профессиональная терминология, новые слова в разных сферах, имена, фамилии, клички, названия брендов – все эти наименования не могут быть обобщены и заданы общими правилами. Персонализированная клавиатура позволяет хранить личный «словарный запас» в облаке и пользоваться им с любого устройства.

Для разработки данной системы выделяют следующие задачи:

1. Изучение предметной области и анализ существующих решений в разработке экранной клавиатуры;
2. Составление технического задания к функциям и интерфейсу программы;
3. Проектирование системы «экранная клавиатура»;
4. Выбор программных средств реализации;
5. Написание программного кода;
6. Тестирование приложения и его отладка.

# Формирование требований

## 1.1. Исследование предметной области лингвистических сервисов

В данной предметной области будут задействованы три основных лингвистических сервиса:

1. Орфокорректор — когда пользователь вводит слово, и оно неправильное, ему рекомендуются до трех правильных вариантов пока курсор находится в конце этого слова;
2. Предиктивный ввод — когда пользователь ввел правильное слово, ему рекомендуются три слова на выбор, каждое из которых может идти за этим словом;
3. Дополнение — когда пользователь набирает слово и, не дожидаясь его окончания, клавиатура начинает предлагать до трех вариантов слов на выбор, которые содержат в себе только что набираемое пользователем слово. Тем самым пользователю предоставляется возможность поскорей закончить слово, выбрав наиболее подходящее из списка.

Экранная клавиатура должна поддерживать вышеперечисленные лингвистические сервисы и «затачиваться» под определенного пользователя, который пользуется данной клавиатурой.

Разработка экранной клавиатуры в данной предметной области призвана решить проблему с набором текста на мобильных устройствах: медленный набор текста и утомительное исправление ошибок, а также индивидуальный «почерк» письма пользователя.

Постоянный набор слов требует определенных усилий со стороны пользователя, который, в свою очередь, хочет вводить слова быстро и корректно; в некоторых случаях предложения исправления слова как «неправильного», относящегося к сленгу пользователя, неуместно – всё это определяет практическую значимость данной клавиатуры.

Также можно затронуть одну интересную статью [-], где два студента-математика, учащиеся в Кембриджском университете, Фолькер Шлуе и Майкл Фестер разработали новый способ набора текста в мобильные устройства, а именно — ПО 8-pen, являющееся виртуальной клавиатурой. ПО 8pen имеет достаточно сильные расхождения от обычных клавиатур в плане организации кнопок — выглядит как «икс», вдоль линий которого расположены буквы от центра «икса» до края по убыванию частоты встречаемости, а в центре «икса» расположена большая точка черного цвета. По углам расположены такие кнопки, как «shift», «delete», «enter» и «?123». Принцип работы следующий: для того чтобы ввести символ, нужно нажать на черную точку в центре «икса» и произвести переход в один из секторов (всего их 4). Число секторов определяется порядком букв на «иксе», а направление движения — стороной, где размещены эти буквы.

Разумеется, в начале работы с данной клавиатурой пользователю может быть не комфортно, так как она в разрез отличается от классических клавиатур. Со временем пользователь начнет привыкать к данному ПО и будет набирать текст в разы быстрее, чем раньше. Также данная клавиатура занимает намного меньше площадь экрана, чем обычная, что добавляет свободу движения. Тем самым придумана новая возможность вводить текст быстро и без ошибок, не прибегая к лингвистическим сервисам.

В следующей обширной статье [-] рассказывается о способах ввода с помощью экранной клавиатуры и обсуждается ее дизайн. Разработка настольных компьютеров, особенно веб-разработка, приводит к вялому интерактивному дизайну: здесь делаются предположения, которые часто оказываются ложными за пределами этой узкой области. В эпоху массового распространения мобильных устройств такие предположения создают множество проблем. Одним из наиболее важных аспектов являются способы ввода информации.

iPhone по умолчанию использует виртуальную сенсорную панель для ввода информации. Этот взгляд очень интересен благодаря своей гибкости. Параметры ввода для определенных полей могут влиять на раскладку клавиатуры в целом и на функции отдельных кнопок в частности. Экранные клавиатуры могут использовать интересные способы взаимодействия с пользователем, начиная от набора текста непрерывным движением пальцев и заканчивая виртуальным колесом прокрутки, позволяющим выбирать значение, например, дату и время, из ограниченного набора элементов.

Клавиатуры используются для ввода слов, а панели используются для ввода цифр и символов. Хотя границы между различными функциями виртуальных клавиатур могут быть размыты, определение разницы между ними является довольно важной задачей.

Интересным моментом является изменение режима ввода. При переключении в новый режим могут измениться как названия, так и функции клавиш, а также расположение и размер кнопок. Это означает, что на самом деле количество режимов ввода бесконечно. Например, при вводе адреса электронной почты может отображаться клавиатура, включающая символ «@» и отдельную клавишу «.com», при этом в адресе могут отсутствовать недопустимые символы, например, запятая, в противном случае раскладка останется прежней.

Цифровой ввод — один из самых интересных режимов. Например, расположение кнопок для набора телефонного номера отличается от раскладки цифровой панели на клавиатуре. Иногда цифры расположены в верхнем ряду буквенно-цифровой панели, как на стандартной компьютерной клавиатуре.

Ввод жестов с виртуальной клавиатуры позволяет ускорить ввод текста (приложение Swype от Nuance studio). Пользователь вводит слово одним движением пальца, которое может остановиться или изменить направление на любом знаке. Дополнительный режим используется наряду со стандартным: пользователь всегда может вернуться к набору текста, нажимая клавиши поочередно — внешний вид клавиатуры не изменится.

Инструменты выбора даты и времени и другие механизмы ввода определенного значения позволяют пользователю произвольно выбирать способ взаимодействия, каждый из которых имеет свои особенности.

Некоторые из методов ввода, такие как колесо прокрутки даты и времени в Android, дополняют сенсорный язык и существуют наряду с ручным вводом. Один щелчок по значению открывает виртуальную клавиатуру или цифровую панель.

В статье [-] речь пойдет о типах элементов ввода с помощью экранной клавиатуры. Рассмотрим самые распространенные типы.

Date. Ни для кого не секрет, что календари на мобильных устройствах ужасно маленькие, что доставляет дискомфорт. Также есть требования на многих текстовых полях, которые принуждают пользователей при вводе даты соблюдать строгие правила проверки. Установив тип ввода date, мобильный браузер решит оставшиеся проблемы.

Email. Ввод адреса электронной почты на обычной клавиатуре требует усилий со стороны пользователя. Данный тип ввода добавляет необходимые клавиши на первый план, такие как «@» и «.com», тем самым ввод адреса значительно упрощается.

Text. Это, пожалуй, самый часто встречающийся тип элементов ввода, который задействован везде.

Tel. Если нужно получить от пользователя только число, то нужно использовать тип элемента ввода tel, так как данный тип предоставляет пользователю цифровую клавиатуру.

## Анализ существующих клавиатур с поддержкой лингвистических сервисов

В статье [-] рассмотрена одна из клавиатур с поддержкой лингвистических сервисов — Microsoft SwiftKey. Microsoft SwiftKey — это интеллектуальная клавиатура, изучающая стиль написания текста определенного пользователя, тем самым помогая ему быстрее и продуктивнее печатать текст. С помощью этой персонализированной клавиатуры есть возможность печатать сообщения, отправлять GIF-файлы и Emoji и так далее. Microsoft SwiftKey постоянно анализирует и подстраивается под уникальный стиль набора своего пользователя, сохраняя в памяти используемый им сленг и Emoji. Данная клавиатура предоставляет расширение сервиса орфокорректор в виде автоисправления. Настраиваемая клавиатура выдает полезные подсказки, помогающие печатать текст быстро и без ошибок. Основные функции этой клавиатуры:

1. Клавиатура Emoji подстраивается к стилю печати пользователя, обучается и предоставляет его любимые эмотиконы во время общения;
2. Проверка орфографии и автоматическая подстановка текста с прогнозированием на основе искусственного интеллекта;
3. Имеет поддержку около 400 языков;
4. Около 100 ярких тем.

В сравнение с Microsoft SwiftKey возьмем обычную встроенную экранную клавиатуру на более-менее современные Android-смартфоны от компании Google под названием Gboard. Сразу можно сказать, что Gboard также поддерживает все три лингвистических сервиса (орфокорректор, предиктивный ввод и дополнение), как и вышеописанная клавиатура.

Данная стандартная клавиатура также имеет возможность обучаться под определенного пользователя. Нужно просто активировать пункт Персонализация в ее настройках. Gboard будет некоторое время хранить на устройстве пользователя фрагменты напечатанного и продиктованного им текста.

Рассмотрим еще одну клавиатуру — экранная клавиатура Windows 10. Этот вид клавиатур поддерживает сервисы дополнение и предиктивный ввод, но не обучаем. Также есть возможность для более детальной ее настройки, например, использовать звуковой сигнал, если пользователь хочет слышать звук при нажатии клавиш.

Исходя из всего вышесказанного, можно сделать вывод, что на данный момент субъективно лучшей клавиатурой является Microsoft SwiftKey, так как она обладает наибольшим функционалом по сравнению со стандартными экранными клавиатурами на смартфонах и компьютерах. Наиболее отличительная особенность Microsoft SwiftKey — это показ статистики персонализированного ввода пользователю в наиболее удобном и развернутом виде, что отсутствует в других рассмотренных клавиатурах. Некоторые особенности всех трех вышеописанных клавиатур будут позаимствованы при выполнении выпускной квалификационной работы.

## 1.3. Техническое задание

Целью разработки экранной клавиатуры для мобильных устройств с поддержкой лингвистических сервисов является оптимизация, а также персонализация стиля ввода текста пользователем. Работа предлагает решение проблем с набором текста на мобильных устройствах: медленный набор текста и утомительное исправление ошибок. Функциональные требования к проектируемой системе:

1. Возможность набирать текст с помощью клавиш;
2. Рекомендации по исправлению слова до трех правильных вариантов;
3. Рекомендации до трех слов на выбор, каждое из которых может идти за текущим словом;
4. Рекомендации до трех слов на выбор, каждым из которых можно завершить написание текущего слова;
5. Обучение под конкретного пользователя, анализируя его набор текста и его использование рекомендаций, предложенных клавиатурой;
6. Возможность менять в настройках приложения цвет, шрифт и раскладку (поддержка русскоязычной и англоязычной раскладок) клавиатуры;
7. Должна иметься возможность авторизации и регистрации в приложении.

Данные о словах и необходимая информация о пользователях должны храниться в базах данных, причем пароли пользователей в целях безопасности должны храниться в виде хэша. Вся информация, используемая приложением, должна храниться в структурированном виде.

Далее обозначены требования к компонентам аппаратного и программного обеспечения, на которых развертывается программная реализация экранной клавиатуры:

1. ОС — Android;
2. Оперативная память — 2 GB;
3. Количество ядер — 2;
4. Тактовая частота процессора — 1000 МГц.

Данные от клиента на сервер, а с сервера в базу данных, и обратно должны передаваться в формате JSON. Протокол передачи данных, использующийся в приложении — http. Если возникнет сбой с сервером или с базами данных (на стороне сервера или локально), пользователь все равно сможет продолжить пользоваться клавиатурой, а также с подбором слов в качестве подсказок, кроме случая, если сбой произошел с локальной базой данных перед началом работы пользователя. Следует заметить, что обучаемость клавиатуры под пользователя при любых сбоях с локальной базой данных не предусмотрена.

Серверная часть (REST API) приложения будет разработана на языке программирования Java с использованием фрэймворка Spring.

Для взаимодействия с сервером выбрана СУБД PostgreSQL.

Клиентская часть будет также разработана на языке программирования Java и на локальном уровне будет обмениваться данными с СУБД SQLite. Чтобы взаимодействовать с сервером посредством отправки к нему запросов, а также для преобразования данных в формат JSON и обратно, клиентская часть будет использовать REST-клиент для Android и Java под названием Retrofit.

Интерфейс приложения должен быть удобен для конечного пользователя, который отвечает следующим требованиям:

1. Наличие удобного и не вызывающего напряженность глаз шрифта;
2. Использование русскоязычного интерфейса для авторизации и регистрации пользователя, а также для настройки клавиатуры;
3. Использование цветовой палитры спокойных оттенков для комфортного взаимодействия пользователя с программой.

# 2. Проектирование сервиса-клавиатуры

## 2.1. Проектирование баз данных

В приложении присутствуют две базы данных: на стороне сервера (серверная) и на стороне клиента (локальная). Разбирать диаграммы баз данных требуется для более глубокого понимания работы лингвистических сервисов. Сначала рассмотрим диаграмму серверной базы данных (рисунок 1).

Изображение выглядит как диаграмма

Автоматически созданное описание

Рисунок 1. Диаграмма серверной базы данных

Серверная база данных задействуется в случае, когда требуется выполнить авторизацию, регистрацию, синхронизацию данных или удаление аккаунта. Есть три таблицы: users, words, collocations.

Таблица users хранит логины и пароли пользователей, чтобы в дальнейшем их идентифицировать (необходимо для синхронизации данных определенного пользователя).

Таблица words хранит идентификаторы пользователей (userId), которым соответствуют определенные слова (word) и счетчик частоты появления для каждого слова (count). Данная таблица нужна для определения орфографически верного слова, а также для более быстрого завершения текущего слова. Благодаря этой сущности слово может определиться как «правильное» и «неправильное» для определенного пользователя. «Правильным» становится то слово, чей счетчик (count) больше или равен числу от 1 до 10 (пользователь может сам определить наиболее подходящее для себя, но по умолчанию число равно 3).

Таблица collocations хранит идентификатор на предыдущее слово (prevId), идентификатор на следующее слово (nextId) и счетчик данных пар (count). С помощью этой таблицы подбираются слова для продолжения, основываясь на частоту их использования (значение count) с определенным словом. В данной сущности могут находиться только «правильные» слова.

Теперь рассмотрим диаграмму локальной базы данных (рисунок 2).

Изображение выглядит как диаграмма, стол

Автоматически созданное описание

Рисунок 2. Диаграмма локальной базы данных

Локальная база данных взаимодействует с пользователем на постоянной основе, то есть всегда, независимо от того, вошел он в систему или нет, включил использование подсказок или отключил и так далее. Есть 4 таблицы: user, words, collocations и ime\_settings.

Таблица user, в целях большей безопасности, хранит только логин пользователя, чтобы в дальнейшем его идентифицировать и иметь возможность синхронизировать данные.

Таблица words идентична своему аналогу в «облаке», за исключением, разве что, идентификатора пользователя userId, так как клавиатура будет обучаться независимо от того, зарегистрирован пользователь или нет.

Таблица collocations полностью идентична своему аналогу в «облаке».

Таблица ime\_settings нужна для хранения следующих данных:

1. Использование звука;
2. Использование вибрации;
3. Использование подсказок;
4. Определение цвета фона подсказок;
5. Определение цвета текста сервиса дополнение;
6. Определение цвета текста сервиса орфокорректора;
7. Определение цвета текста сервиса предиктивного ввода;
8. Определение шрифта текста для подсказок;
9. Задание скорости обучения.

## 2.2. Проектирование диаграмм сервиса-клавиатуры

В системе, как уже было сказано, работают три лингвистических сервиса: дополнение, орфокорректор и предиктивный ввод. Работа данных сервисов спроектирована на одной общей диаграмме деятельности. Для лучшего понимания функциональности каждого из этих сервисов диаграмма будет разбита на три части, и каждая часть будет разобрана по отдельности. Начнем с сервиса дополнение.

Дополнение предоставляет возможность ускорить написание слова (быстро закончить его), подобрав наиболее подходящие для этого рекомендации. Модель данного лингвистического сервиса изображена в виде диаграммы деятельности (рисунок 3).

Изображение выглядит как диаграмма

Автоматически созданное описание

Рисунок 3. Модель дополнения

Рассмотрим алгоритм. Пользователь открывает клавиатуру и может начать писать слово, либо выбрать одно из трех слов в качестве подсказок, представленных сервисом дополнение.

Подсказки подбираются следующим образом: из таблицы words (эта таблица никогда не может быть пустой, так как при установке приложения она инициализируется значениями из встроенного словаря) берутся три слова, начальная часть которых совпадает с тем, что последнее написал пользователь на текущий момент, а также чьи счетчики count являются максимальными по значению (частота использования слова). При выборе пользователем подсказки клавиатура подставит выбранное слово на место предыдущего и счетчик count выбранного слова увеличится на 1.

Но, независимо от выбора пользователя, происходит следующая проверка: если символ в конце строки равен пробелу или знаку пунктуации и длина последнего слова не равна 0, то запускается сервис орфокорректор, иначе работа цикла продолжается и подбираются все новые слова для дополнения. Далее рассмотрим сервис орфокорректор.

Орфокорректор позволяет предлагать слова для исправления текущего неправильно написанного слова. Также, если несколько раз не исправлять определенное слово, то орфокорректор «запомнит» его и не будет в дальнейшем предлагать его исправление — еще одна особенность данного лингвистического сервиса. Модель этого сервиса изображена в виде диаграммы деятельности (рисунок 4).

Изображение выглядит как диаграмма

Автоматически созданное описание

Рисунок 4. Модель орфокорректора

Рассмотрим алгоритм. Если слово определилось как «правильное», то счетчик count данного слова увеличивается на 1, орфокорректор заканчивает свою работу и следом запускается сервис предиктивный ввод. Если слово определилось как «неправильное», то у пользователя появляется два варианта действий: либо нажать на клавишу, либо выбрать слово из только что подобранных рекомендаций для исправления в порядке слева направо по убыванию вероятности наиболее подходящего слова.

Подсказки подбираются следующим образом: начинается поиск в таблице words последнего написанного пользователем слова. Если данное слово отсутствует в таблице или же его счетчик count меньше определенного значения (по умолчанию это значение равно 3, но у пользователя есть возможность поменять это значение от 1 до 10), то слово считается «неправильным» и предлагаются наиболее подходящие слова для его исправления.

Если пользователь воспользовался подсказкой от клавиатуры, то система подставит выбранное слово на место предыдущего, счетчик count выбранной подсказки увеличится на 1, орфокорректор закончит свою работу и начнет работу предиктивный ввод. Если пользователь проигнорирует подсказку и начнет дальше писать текст, то «неправильное» слово запишется в таблицу words (если оно там не находится) или увеличит его счетчик count на 1 (если оно там есть), орфокорректор закончит свою работу и начнет работу дополнение. Если пользователь нажал на символ удаления и после этого символ в конце строки не равен символу алфавита, то ничего не произойдет в плане подбора подсказок и фиксации новых слов. Иначе если при удалении курсор дошел непосредственно до самого слова (символ в конце строки равен символу алфавита), то орфокорректор закончит свою работу и начнет работу дополнение. Далее рассмотрим последний сервис — предиктивный ввод.

Предиктивный ввод позволяет предугадать следующие наиболее вероятные слова, которые могут идти за текущим словом. Модель данного лингвистического сервиса изображена в виде диаграммы деятельности (рисунок 5).

Изображение выглядит как диаграмма, схематичный

Автоматически созданное описание

Рисунок 5. Модель предиктивного ввода

Рассмотрим алгоритм. Сначала сервис проверяет на равенство предыдущее слово (слово, которое находится на данный момент перед текущим словом) с сохраненным ранее словом. Если равенство соблюдается, то данная пара слов добавляется в таблицу collocations (если ее там нет) или ее счетчик count увеличится на 1 (если она там есть). Иначе ничего не происходит. В любом случае у пользователя будут два варианта действий: либо нажать на клавишу, либо выбрать слово для продолжения из только что подобранных наиболее вероятных слов слева направо по убыванию вероятности.

Подсказки подбираются следующим образом: система обращается к данным таблицы collocations и по значению prevId текущего слова находит все пары, связанные с этим словом. С помощью счетчика count производится сортировка по убыванию. В полях подсказок оказываются самые верхние слова, чьи индексы соответствуют значениям nextId.

Если пользователь воспользовался подсказкой от клавиатуры, то на место курсора клавиатура подставит выбранную подсказку, данная пара слов добавляется в таблицу collocations (если ее там нет) или ее счетчик count увеличится на 1 (если она там есть), текущее слово сохранится в оперативную память, предиктивный ввод закончит свою работу и начнет работу дополнение. Если пользователь нажал на клавишу и данный символ не равен символу алфавита, то идет возврат в начало цикла. Иначе текущее слово сохранится в оперативную память, предиктивный ввод закончит свою работу и начнет работу дополнение.