Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Асманов Г.М. Разработка редактора шаблонов печати для медицинской информационной системы «Самсон» // Академия педагогических идей «Новация». – 2026. – №3 (октябрь). – АРТ 4-эл. – 0,7 п. л. – URL: http://akademnova.ru/page/875548

РУБРИКА: ПРОФЕССИОНАЛЬНОЕ ОБРАЗОВАНИЕ

УДК 004.9

Асманов Геннадий Матвеевич
студент 5 курса, группа ЗИС-51
кафедра «Вычислительная техника и программирование»
Технологический институт (филиал) ДГТУ в г. Азове
г. Азов, Российская Федерация
e-mail: Scarlett.mayer@yandex,ru

РАЗРАБОТКА РЕДАКТОРА ШАБЛОНОВ ПЕЧАТИ ДЛЯ МЕДИЦИНСКОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ «САМСОН»

Аннотация: В статье рассмотрены особенности проектирования и реализации программного модуля для автоматизации работы IT-отдела медицинского учреждения. Приведено описание функциональных возможностей редактора шаблонов печати, интегрированного с МИС «САМСОН», включая конвертацию документов, OCR-распознавание сканов и визуальное редактирование HTML-кода.

Ключевые слова: медицинская информационная система, шаблоны печати, OCR, Python, автоматизация документооборота.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Asmanov Gennady
5th year student, group ZIS-51
Department of Computer Engineering and Programming
Technological Institute (branch) of DSTU in Azov
Azov, Russian Federation

DEVELOPMENT OF A PRINT TEMPLATE EDITOR FOR THE MEDICAL INFORMATION SYSTEM «SAMSON»

Abstract: The article discusses the design and implementation of a software module for automating the work of the IT department of a medical institution. The functionality of the print template editor integrated with the MIS "SAMSON" is described, including document conversion, OCR recognition of scans, and visual editing of HTML code.

Keywords: medical information system, print templates, OCR, Python, document automation.

Введение

Современная система здравоохранения представляет собой сложный комплекс, эффективное функционирование которого напрямую зависит от качества информационного обеспечения и скорости обработки медицинских данных. В условиях активной цифровой трансформации отрасли, внедрения Единой государственной информационной системы в сфере здравоохранения (ЕГИСЗ) и повсеместного использования медицинских информационных систем (МИС), особое значение приобретает автоматизация рутинных процессов, выполняемых сотрудниками IT-подразделений. Медицинские учреждения ежедневно

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

сталкиваются с необходимостью обработки больших объемов документации: перевод в электронный вид бумажных архивов, распознавание отсканированных справок и выписок, а также формирование и поддержка актуальности шаблонов печати для врачей.

В ГБУЗ «Староминская ЦРБ» МЗ КК основной рабочей средой медицинского персонала является медицинская информационная система «САМСОН». Работа с печатными формами в данной системе требует создания HTML-шаблонов со специальными вставками, которые впоследствии заполняются данными из базы. На текущий момент значительная часть подготовки таких шаблонов выполняется IT-отделом вручную: специалистам приходится самостоятельно набирать текст со сканов или использовать разрозненные сторонние программы для конвертации документов. Это приводит к высоким временным затратам, появлению ошибок «человеческого фактора» и снижает оперативность внедрения новых печатных форм в лечебный процесс.

В связи с этим целью производственной технологической практики является закрепление и углубление теоретических знаний, изучение информационной среды медицинского учреждения, а также приобретение практических навыков разработки программного обеспечения путем создания специализированного модуля для автоматизации работы IT-отдела.

Для достижения поставленной цели в ходе практики решались следующие задачи:

1. Изучить организационную структуру ГБУЗ «Староминская ЦРБ», включая специфику работы отдела компьютерно-сетевого обеспечения (IT-отдела).
2. Охарактеризовать информационную среду предприятия и ее

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

назначение (взаимодействие с МИАЦ, работа в МИС «САМСОН» и защищенной сети ViPNet).

3. Разработать функциональную модель деятельности подразделения и провести анализ предметной области.

4. Разработать макет программного модуля для решения поставленной задачи — создания редактора шаблонов печати.

5. Обосновать выбор программных средств для реализации модуля (язык Python, фреймворк Flask, библиотеки PyQt, Tesseract и Poppler для OCR).

6. Реализовать программный модуль, провести его тестирование и сборку в виде автономного исполняемого файла.

7. Подготовить итоговый отчет по практике и освоить необходимые профессиональные компетенции.

Объектом практики является деятельность отдела компьютерно-сетевого обеспечения ГБУЗ «Староминская ЦРБ». Предметом практики является процесс разработки и редактирования шаблонов печатных документов для медицинской информационной системы «САМСОН».

В ходе выполнения работы были использованы следующие методы: анализ предметной области и технической документации, структурное и объектно-ориентированное программирование, методы оптического распознавания символов (OCR), а также методы тестирования программного обеспечения («черный ящик»).

Практическая значимость работы заключается в том, что разработанный программный модуль позволяет объединить процессы конвертации DOCX-файлов, распознавания текста из PDF-сканов и редактирования HTML-кода в едином окне. Внедрение данного

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

инструмента позволит сотрудникам IT-отдела существенно сократить время на подготовку шаблонов, исключить ошибки при ручном наборе и повысить общую эффективность информационного сопровождения лечебного процесса в больнице.

Отчет состоит из введения, трех глав, заключения и списка использованных информационных ресурсов. В первой главе представлен анализ предметной области, обзор существующих аналогов и обоснование выбора средств разработки. Во второй главе описана разработка структурной схемы, программная реализация, руководство пользователя и результаты тестирования. В заключении подведены итоги проделанной работы.

1. Аналитический раздел

1.1. Анализ предметной области

Предметная область автоматизации документооборота в медицинском учреждении представляет собой совокупность процессов и участников, взаимодействующих с целью обеспечения бесперебойной работы с медицинской документацией и шаблонами печати в информационной системе. Основными объектами обработки являются бумажные архивы, отсканированные PDF-документы, текстовые файлы форматов DOCX и HTML-шаблоны, используемые в МИС «САМСОН». Каждый из этих элементов имеет собственные требования к форматированию, структуре и корректности подстановки данных. Нарушения в работе с документацией могут приводить как к ошибкам в лечебном процессе, так и к снижению эффективности работы персонала и нарушению стандартов ведения медицинских карт.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Анализ процессов, существующих в IT-отделе учреждения, позволил выявить ключевые типы задач, возникающих в ежедневной практике: перевод устаревших бумажных документов и сканов в электронный вид, конвертация текстовых файлов в HTML-шаблоны, ручная вставка специальных переменных {...} для интеграции с базой данных МИС, а также проверка синтаксиса и корректности готовых форм.

Каждый из этих типов задач требует незамедлительной обработки, высокой точности распознавания текста и соблюдения строгой структуры HTML-кода. На текущий момент значительная часть работы выполняется вручную или с использованием разрозненных инструментов (сторонние онлайн-конвертеры, текстовые редакторы), что создаёт риски потери информации, появления ошибок при распознавании и невозможности быстрой адаптации шаблонов под меняющиеся требования МИС.

Особое внимание в ходе исследования было уделено роли сотрудников IT-отдела. Они представляют собой ключевое звено процесса, поскольку непосредственно осуществляют конвертацию, распознавание и подготовку шаблонов для медицинского персонала. Однако отсутствие специализированных автоматизированных инструментов приводит к тому, что значительная часть времени тратится не на решение сложных технических задач, а на рутинные операции: ручной набор текста со сканов, исправление ошибок OCR (оптического распознавания символов) и поиск нужных переменных в документации. Это снижает эффективность работы, создаёт неравномерную загрузку и повышает риск появления «человеческого фактора».

Важным аспектом является взаимодействие IT-отдела с врачами и административным персоналом. Медицинские работники нуждаются в быстром и корректном обновлении печатных форм, а руководство - в

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

прозрачной системе контроля, позволяющей оценивать скорость внедрения новых шаблонов и качество подготавливаемой документации. Без автоматизации эти процессы осуществляются фрагментарно, что снижает точность анализа и ограничивает возможности совершенствования управленческих протоколов.

Также было выполнено изучение существующих цифровых решений, применяемых рынком. Обследование показало наличие отдельных инструментов для распознавания текста (например, ABBYY FineReader) и онлайн-конвертеров, но большинство из них не учитывают специфику работы с МИС «САМСОН», не поддерживают интеграцию со справочником переменных и не обеспечивают необходимого уровня безопасности при работе с персональными медицинскими данными. Это подтвердило необходимость разработки собственного программного модуля, учитывающего особенности предметной области, набор ролей пользователей, требования к безопасности данных и возможности масштабирования.

В результате анализа предметной области были выделены системные проблемы, объясняющие актуальность разработки комплексной автоматизации. Среди них - отсутствие единой среды для конвертации и распознавания, несогласованность форматов, разрозненность данных, низкая степень автоматизации вставки переменных МИС и невысокая скорость подготовки шаблонов. Все эти факторы определили методологическую основу для разработки проектной структуры системы и сформировали требования, учитываемые при построении её архитектуры, включая использование OCR-движка Tesseract и утилиты Poppler для работы с PDF-файлами.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

1.2. Анализ существующих аналогов

1.2.1. ABBYY FineReader

ABBYY FineReader - это признанный лидер на рынке программ для оптического распознавания текста (OCR) и конвертации документов. Программа позволяет сканировать бумажные документы, PDF-файлы и изображения, преобразуя их в редактируемые форматы, включая DOCX, PDF и HTML. Интерфейс программы представлен на рисунке 1.

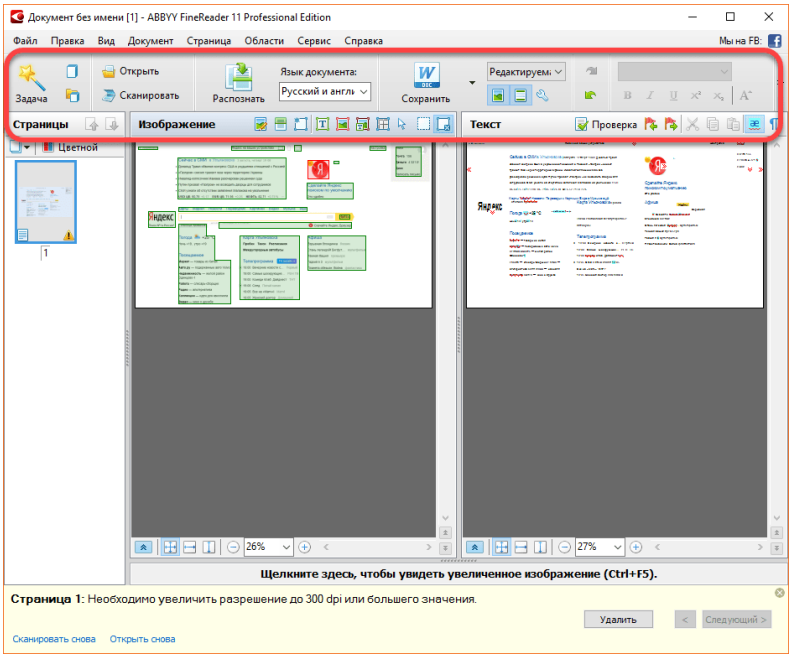


Рисунок 1 - Интерфейс программы ABBYY FineReader

Основные функции:

1. Высокоточное распознавание текста (OCR) на русском и иностранных языках.

2. Конвертация PDF и изображений в редактируемые форматы (DOCX, HTML).

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

3. Сохранение структуры документа (таблиц, списков) при конвертации.

4. Пакетная обработка большого количества файлов.

Преимущества:

1. Наивысшая точность распознавания текста на рынке.

2. Удобный визуальный интерфейс для проверки ошибок распознавания.

3. Поддержка огромного количества форматов.

Недостатки в контексте задачи:

Данный программный продукт является мощным, но избыточным и платным решением. Он не имеет встроенной палитры переменных МИС «САМСОН» и не позволяет напрямую проверять синтаксис специальных медицинских вставок {...}. Кроме того, он не интегрируется с базой данных шаблонов учреждения.

1.2.2. Pandoc

Pandoc - это универсальный конвертер документов с открытым исходным кодом, работающий через командную строку. Он широко используется IT-специалистами для преобразования файлов из одного формата в другой (например, из DOCX в HTML или Markdown). Интерфейс программы (командная строка) представлен на рисунке 2.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

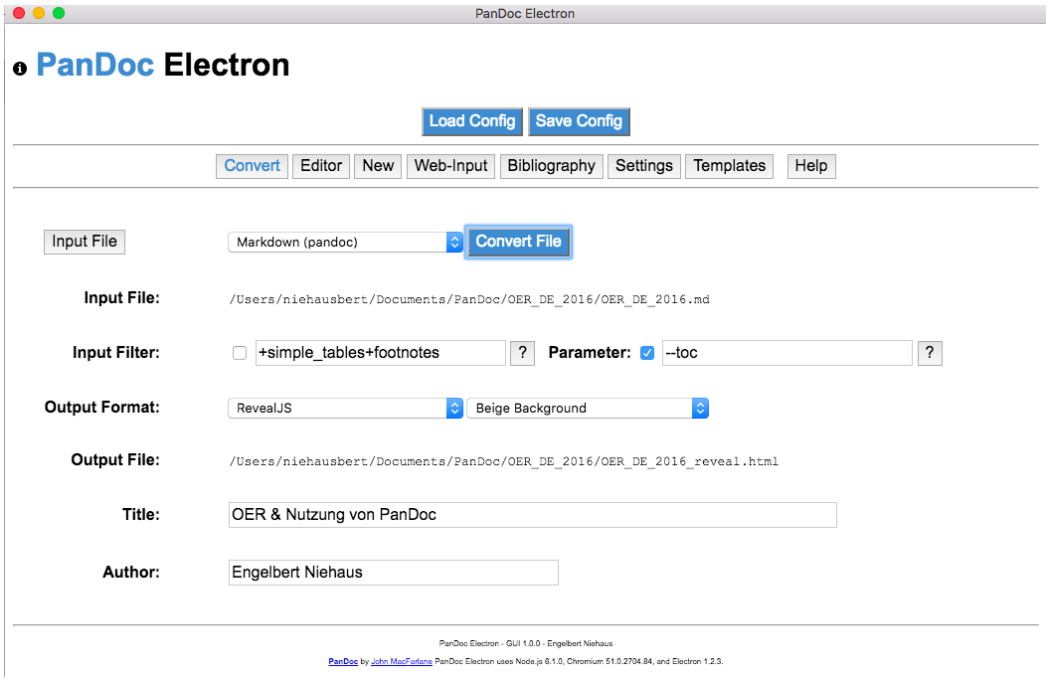


Рисунок 2 - Интерфейс программы Pandoc

Основные функции:

1. Конвертация текстовых форматов (DOCX, Markdown, HTML, LaTeX и др.).
2. Сохранение базовой структуры документа (заголовки, списки, таблицы).
3. Работа без графического интерфейса (через терминал).
4. Кроссплатформенность (Windows, Linux, macOS).

Преимущества:

1. Полностью бесплатное решение с открытым исходным кодом.
2. Очень высокая скорость конвертации.
3. Легкость автоматизации и встраивания в скрипты.

Недостатки в контексте задачи:

Pandoc не умеет распознавать текст с отсканированных PDF

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

(отсутствует OCR). У него нет графического интерфейса, что делает его неудобным для обычных сотрудников IT-отдела без навыков программирования. Также он не поддерживает визуальное редактирование шаблонов и работу со специфическими переменными МИС.

1.2.3. Визуальные HTML-редакторы

Визуальные редакторы - это программные компоненты, встраиваемые в веб-приложения для редактирования HTML-кода в режиме реального времени. Подобные редакторы часто используются для создания шаблонов писем или веб-страниц. Интерфейс редактора представлен на рисунке 3.

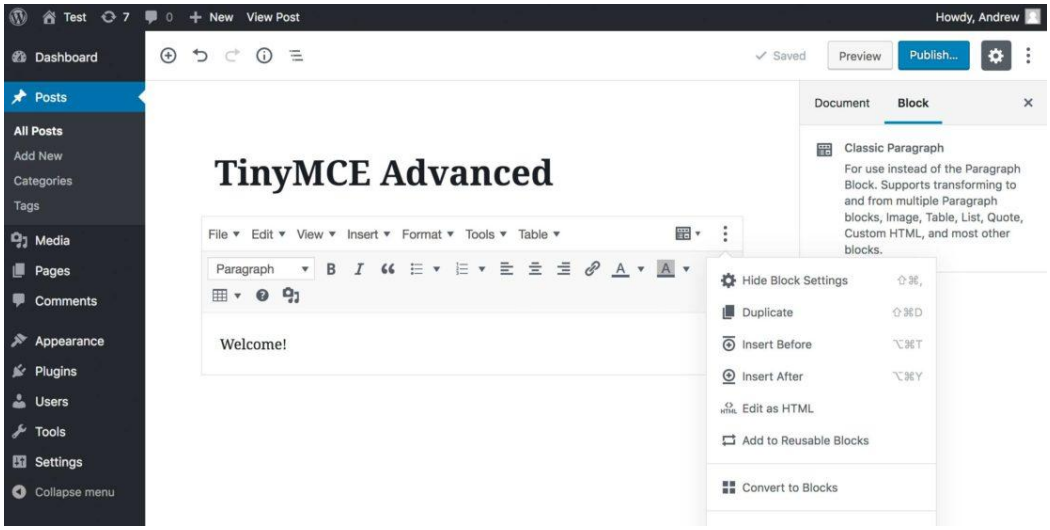


Рисунок 3 - Интерфейс визуального HTML-редактора

Основные функции:

1. Визуальное редактирование HTML-кода (как в текстовом процессоре).
2. Возможность просмотра и правки исходного кода.
3. Форматирование текста, вставка таблиц и изображений.
4. Расширяемость с помощью плагинов.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Преимущества:

1. Интуитивно понятный интерфейс, не требующий знания HTML.
2. Возможность встраивания в любые веб-системы.
3. Быстрота работы с текстом.

Недостатки в контексте задачи:

Редакторы сами по себе не умеют конвертировать PDF или DOCX в HTML, а также не распознают сканы. Кроме того, «из коробки» они не содержат специализированной палитры переменных для медицинской информационной системы и не проверяют баланс фигурных скобок в специфических вставках.

1.3. Анализ используемых средств разработки

При разработке автоматизированной информационной системы (редактора шаблонов печати) был выбран язык программирования Python, в качестве среды разработки использовался Visual Studio Code, а для хранения данных - встраиваемая СУБД SQLite в связке с ORM SQLAlchemy.

Python - это высокоуровневый язык программирования общего назначения, который является одним из самых популярных инструментов для создания программного обеспечения. Данный язык предлагает разработчикам широкие возможности для решения задач различной сложности и направленности.

Ключевым преимуществом Python является его обширная экосистема библиотек, позволяющая решать практически любые задачи: от веб-разработки (Flask) и работы с базами данных (SQLAlchemy) до создания графических интерфейсов (PyQt) и обработки документов (mammoth, pytesseract). Благодаря простому и читаемому синтаксису, Python обеспечивает высокую скорость разработки и упрощает сопровождение кода, что особенно важно при передаче проекта другим специалистам.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Мощная поддержка со стороны сообщества и наличие подробной документации позволяют быстро находить решения возникающих проблем. Кроссплатформенность Python гарантирует, что разработанное приложение будет работать как в среде Windows, так и в Linux (что критически важно для ГБУЗ «Староминская ЦРБ», где используются различные операционные системы). Интерпретируемость языка ускоряет цикл разработки, позволяя оперативно вносить изменения и тестировать новые функции без необходимости компиляции.

Visual Studio Code - это легкий, но мощный редактор кода от компании Microsoft, представляющий собой многофункциональный инструмент для разработки программного обеспечения. Данный редактор является одним из самых популярных среди разработчиков на Python.

Ключевым преимуществом Visual Studio Code является наличие официального расширения от Microsoft, обеспечивающего интеллектуальное автодополнение кода (IntelliSense), проверку синтаксиса, отладку и линтинг в реальном времени. Встроенная интеграция с системой контроля версий Git позволяет управлять историей изменений прямо из интерфейса редактора, не переключаясь в сторонние приложения. Гибкая система расширений (плагинов) позволяет настраивать среду разработки под индивидуальные потребности, добавляя поддержку HTML, CSS, JavaScript, SQL и других технологий. Постоянное обновление и развитие Visual Studio Code обеспечивает поддержку новейших технологий и фреймворков. Тесная интеграция с современными инструментами делает этот редактор оптимальным выбором для создания программных решений на Python.

SQLite - это компактная, встраиваемая система управления базами данных (СУБД), разработанная для локального хранения данных. Она

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

предоставляет мощные средства для управления и обработки информации в настольных и веб-приложениях, от небольших до средних решений.

Важным преимуществом является бесплатность и кроссплатформенность - Visual Studio Code работает на Windows, Linux и macOS, что делает его универсальным инструментом для любых условий. Интерфейс этой среды представлен на рисунке 4.

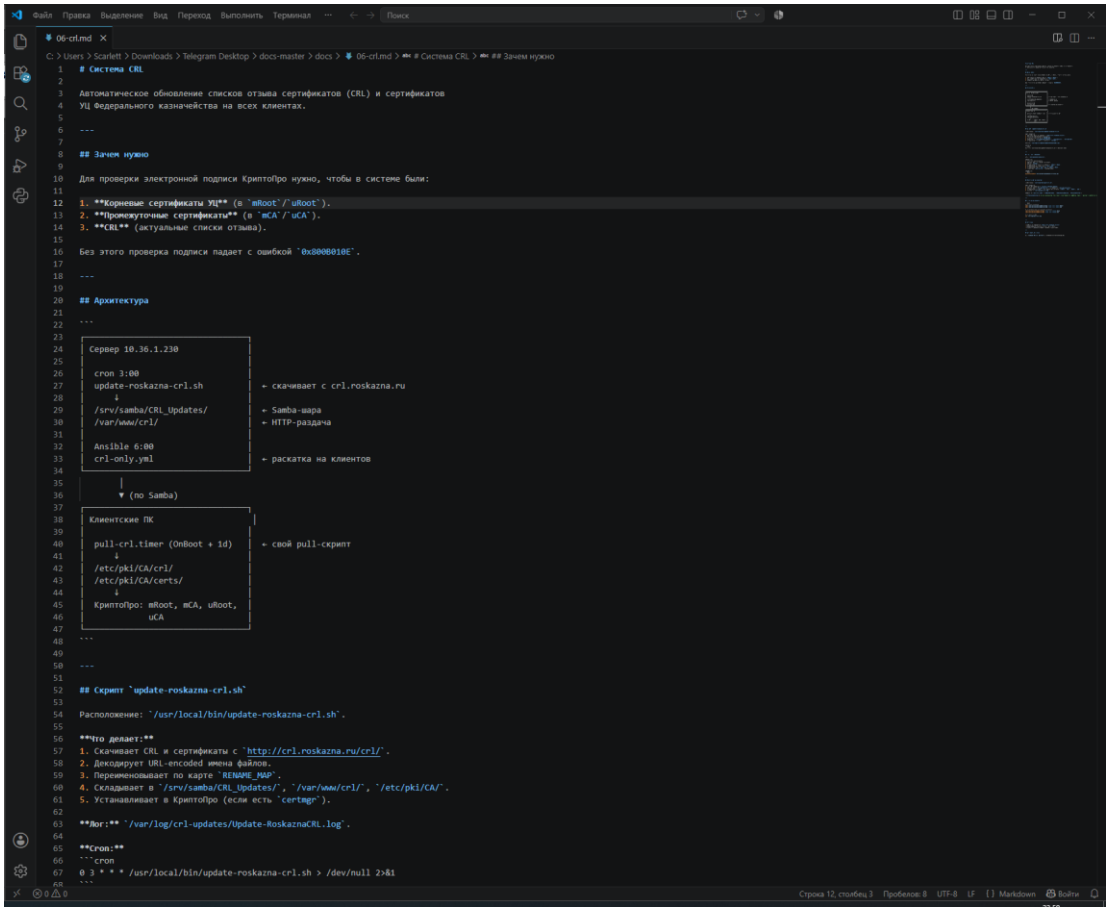


Рисунок 4 - Интерфейс программы Visual Studio Code

Одной из ключевых особенностей SQLite является отсутствие необходимости в отдельном сервере. Вся база данных хранится в одном файле (в нашем случае - templates.db), что делает её идеальной для десктопных приложений, где не предполагается многопользовательский доступ и высокие нагрузки. Это существенно упрощает развертывание

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

приложения на рабочих станциях IT-отдела и резервное копирование данных. SQLite поддерживает транзакции, что обеспечивает надежность и целостность данных.

Для удобства работы с базой данных в проекте используется SQLAlchemy - мощная библиотека ORM (Object-Relational Mapping) для Python. Она позволяет работать с таблицами базы данных как с обычными объектами Python, не отвлекаясь на написание сложных SQL-запросов. SQLAlchemy обеспечивает абстракцию от конкретной СУБД (при необходимости можно легко перейти на PostgreSQL). Интерфейс взаимодействия с базой данных представлен на рисунке 5.

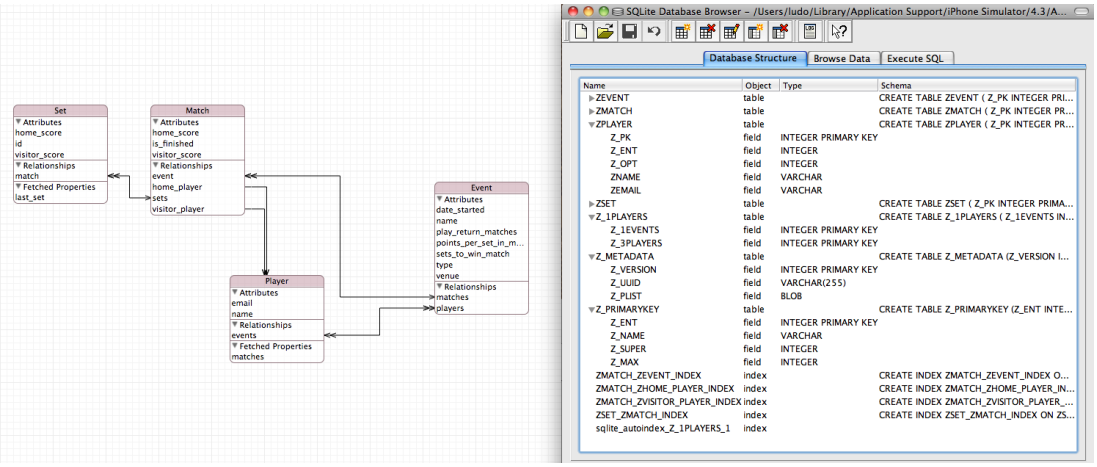


Рисунок 5 - Структура базы данных SQLite

SQLite и SQLAlchemy активно используются в разработке десктопных приложений, обеспечивая эффективное хранение и обработку информации. Система поддерживает интеграцию с различными библиотеками Python, что делает её универсальным инструментом для создания приложений, не требующих сложного администрирования.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Flask - это микрофреймворк для создания веб-приложений на Python. Он был выбран в качестве основы для серверной части приложения, так как не навязывает жесткую структуру проекта и позволяет гибко организовать код под специфику задачи.

PyQt5 и PyQtWebEngine - это набор библиотек для создания графических интерфейсов (GUI) на Python. В данном проекте они использовались для создания десктопной обёртки, которая запускает Flask-сервер в фоновом потоке и отображает веб-интерфейс во встроенном браузере на движке Chromium. Это позволяет пользователю работать с приложением как с обычной программой, без необходимости открывать внешний браузер.

PyInstaller - это инструмент для упаковки Python-приложений в автономные исполняемые файлы (.exe). Он использовался для сборки готового продукта, который можно запускать на рабочих станциях без предварительной установки Python и всех зависимостей.

Git - это распределённая система контроля версий, использовавшаяся для отслеживания изменений в коде, сохранения истории разработки и резервного копирования проекта в удалённом репозитории на GitHub. Интеграция с Visual Studio Code позволяет выполнять основные операции (commit, push, pull) прямо из редактора.

Кроме того, для решения специфических задач использовались специализированные библиотеки: mammoth (конвертация DOCX в HTML), pytesseract (обёртка для OCR-движка Tesseract), pdf2image (преобразование PDF-страниц в изображения) и BeautifulSoup (форматирование HTML-кода). Их использование позволило реализовать функции импорта и распознавания документов, что значительно упростило работу IT-отдела с шаблонами печати.

Таким образом, выбранный стек технологий обеспечивает высокую производительность, надежность и удобство сопровождения разработанного программного модуля, а также соответствует современным требованиям к разработке десктопных и веб-приложений.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

2. Специальный раздел

2.1. Разработка структурной схемы

Создание структурной схемы автоматизированной системы (редактора шаблонов печати для МИС «САМСОН») потребовало перехода от теоретического анализа предметной области к формализации процессов, определению функциональных модулей и взаимосвязей между ними. Структурная схема была разработана таким образом, чтобы обеспечить прозрачность направлений взаимодействия, единообразие обработки данных, возможность масштабирования и интеграцию с уже существующей инфраструктурой медицинского учреждения. В основе проектирования лежит принцип модульности, позволяющий независимо развивать отдельные компоненты системы, не нарушая целостности всей архитектуры.

Первым этапом разработки схемы стала формализация потока обработки документов. Была определена логика прохождения файла от момента его загрузки пользователем до полного сохранения в виде готового HTML-шаблона. Загрузка должна осуществляться в едином цифровом интерфейсе, доступном сотрудникам IT-отдела. После регистрации файла система автоматически определяет его тип: если загружен DOCX-документ, он направляется в модуль конвертации на базе библиотеки mammoth; если загружен PDF-файл (в том числе отсканированный), он передается в модуль оптического распознавания символов (OCR) на базе pdf2image, pytesseract и движка Tesseract. В схеме предусмотрены механизмы обработки ошибок (например, при сбое распознавания) и вывод уведомлений пользователю.

Таким образом, каждый документ проходит полный цикл преобразования внутри системы без необходимости использования сторонних онлайн-сервисов, что критически важно для соблюдения режима

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

конфиденциальности медицинских данных.

Следующим элементом структурной схемы стал модуль работы со справочником переменных. Он обеспечивает интеграцию с палитрой переменных МИС «САМСОН», позволяя пользователю вставлять в HTML-код специальные конструкции вида {ключ_переменной}. В схеме предусмотрена возможность проверки синтаксиса (баланс открывающих и закрывающих фигурных скобок), что позволяет выявлять ошибки в шаблонах до момента их загрузки в медицинскую информационную систему.

Важную роль в архитектуре играет модуль хранения данных и управления шаблонами. Его задача - обеспечение надежного сохранения готовых HTML-шаблонов в локальной базе данных SQLite. В основе модуля лежит ORM SQLAlchemy, которая управляет таблицами шаблонов, категорий и переменных. В схеме предусмотрены механизмы редактирования, удаления и предпросмотра сохраненных данных. Все действия фиксируются в базе данных с указанием даты и времени последнего обновления (created_at и updated_at), что позволяет отслеживать актуальность шаблонов.

Отдельный блок схемы посвящён взаимодействию десктопной обёртки и веб-сервера. В нём предусмотрена работа фреймворка Flask (серверная часть) и библиотек PyQt5/PyQtWebEngine (клиентская часть). Серверная часть запускается в отдельном фоновом потоке, а графический интерфейс отображает веб-страницу во встроенном браузере. Это позволяет пользователю работать с приложением как с обычной программой, без необходимости открывать внешний браузер и настраивать локальный сервер. Наличие этого блока обеспечивает кроссплатформенность и

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

удобство развертывания системы на рабочих станциях.

Завершающим этапом разработки структурной схемы стала формализация взаимодействия между пользователями системы. Были определены роли сотрудников - администратор (IT-специалист) и медицинский работник. Администратор обладает полномочиями по загрузке файлов, редактированию кода, управлению переменными и удалению шаблонов. Роль медицинского работника (на этапе использования готового шаблона) подразумевает только получение корректной печатной формы. В схеме также предусмотрено разграничение прав пользователей и защита данных за счет локального хранения (без передачи данных на сторонние серверы), что является обязательным требованием при разработке медицинских цифровых систем.

Созданная структурная схема представляет собой комплексную архитектурную модель, объединяющую процессы конвертации документов, OCR-распознавания, управления переменными, хранения шаблонов и взаимодействия пользователя с интерфейсом. Она служит основой для дальнейшего проектирования технической и программной реализации системы и демонстрирует принципы построения эффективного, безопасного и современного инструмента для IT-отдела медицинского учреждения.

2.2. Программная реализация

Функциональная схема или схема данных - это схема взаимодействия компонентов программного обеспечения с описанием информационных потоков, состава данных в потоках и указанием используемых файлов и устройств. Для изображения функциональных схем используют специальные обозначения, установленные стандартом.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Другими словами, функциональная схема - документ, разъясняющий процессы, протекающие в отдельных функциональных цепях изделия (установки) или изделия (установки) в целом. Функциональная схема является экспликацией (поясняющим материалом) отдельных видов процессов, протекающих в целостных функциональных блоках и цепях устройства.

Для сложного программного продукта разрабатывается несколько функциональных схем, поясняющих происходящие процессы при различных предусмотренных режимах работы. Количество функциональных схем, разрабатываемых на программном продукте, степень их детализации и объем помещаемых сведений определяется разработчиком с учетом особенностей изделия.

На рисунке 6 представлена структура программы, изображенная в виде блок-схемы.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

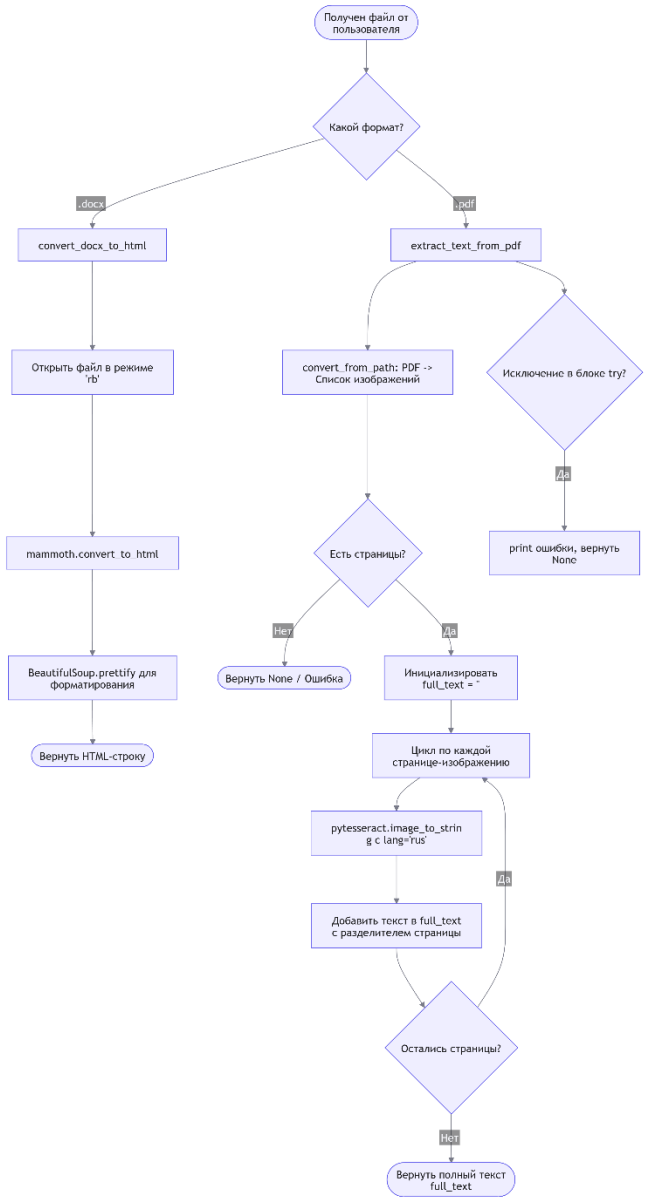


Рисунок 6 - Функциональная схема ключевых компонентов ПО

Функционал системы структурирован на четыре логических модуля. Модуль «Интерфейс» представляет собой десктопную обёртку на базе PyQt5, которая отображает веб-интерфейс для пользователя (сотрудника IT-отдела). Через этот модуль осуществляется загрузка исходных файлов (DOCX и PDF) и вызов диалоговых окон для работы с редактором.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Поступающие файлы попадают в модуль «Обработка файлов». Если загружен документ DOCX, он транслируется в библиотеку mammoth для конвертации в HTML. Если загружен PDF (в том числе скан), он передается в подсистему OCR, где с помощью pdf2image страницы превращаются в изображения, а затем pytesseract (Tesseract) распознает текст. Извлеченные данные конвертируются в предварительный HTML-код с помощью BeautifulSoup для форматирования.

После первичной обработки документ попадает в модуль «Редактирование и переменные». Здесь пользователь через визуальный редактор (CKEditor) корректирует HTML-код и вставляет специальные конструкции {...} из встроенной палитры переменных, которые в дальнейшем будут подставляться в медицинскую систему «САМСОН».

Завершающим этапом является модуль «База данных». Готовый HTML-шаблон сохраняется в локальную реляционную базу данных, где хранится полный набор шаблонов, категорий и переменных. После сохранения пользователь через модуль «Интерфейс» может просмотреть, отредактировать или удалить шаблон, а также скачать его в виде готового файла.

В рамках разработки информационной системы создана реляционная база данных для локального хранения шаблонов, а также реализовано десктопное приложение, автоматизирующее процессы конвертации и редактирования, с использованием СУБД SQLite и ORM SQLAlchemy.

2.3. Руководство пользователя

Разработанный программный модуль «Редактор шаблонов печати для МИС «САМСОН»» представляет собой кроссплатформенное десктопное

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

приложение, объединяющее в себе удобство веб-интерфейса и функциональность локальных программ. Руководство пользователя описывает пошаговый процесс работы с приложением, начиная с момента его запуска и заканчивая сохранением готового шаблона в базу данных. Приложение ориентировано на сотрудников IT-отдела медицинского учреждения, однако благодаря интуитивно понятному интерфейсу не требует глубоких навыков программирования для выполнения базовых операций.

При запуске приложения открывается главное окно, представленное на рисунке 7.

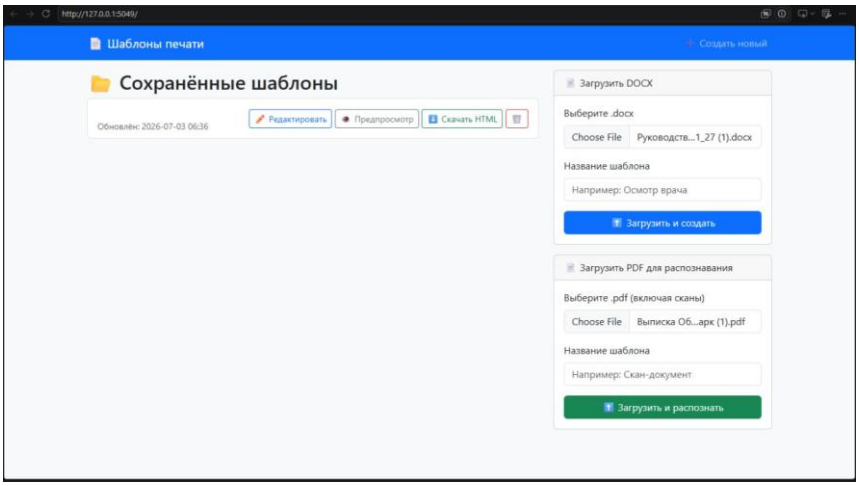


Рисунок 7 - Главная форма приложения

Главная форма представляет собой центральный интерфейс для управления существующими шаблонами печати и импорта новых документов. Окно программы условно разделено на две вертикальные области: основную (левую) часть и панель инструментов (правую).

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

В левой части окна располагается список всех ранее сохранённых в базе данных шаблонов. Каждая строка списка содержит название шаблона, а также точную дату и время его последнего обновления, что позволяет сотруднику IT-отдела всегда ориентироваться в актуальности документов. В крайней правой части каждой строки списка располагаются четыре кнопки управления. Кнопка «Редактировать» открывает окно редактора для внесения изменений в HTML-код. Кнопка «Предпросмотр» открывает страницу с визуализацией шаблона в том виде, в котором он будет отображаться при печати. Кнопка «Скачать HTML» позволяет выгрузить исходный код шаблона в виде отдельного файла на локальный компьютер. Завершает ряд кнопка с иконкой корзины «Удалить», предназначенная для безвозвратного удаления шаблона из локальной базы данных (при нажатии система запрашивает дополнительное подтверждение во избежание случайных потерь).

Верхняя панель предназначена для загрузки документов формата DOCX. Она содержит элемент выбора файла (кнопка «Выберите .docx»), однострочное текстовое поле для ввода названия будущего шаблона (по умолчанию система автоматически подставляет имя загружаемого файла без расширения) и кнопку «Загрузить и создать». При нажатии на кнопку программа сохраняет файл во временную директорию uploads, после чего на серверной стороне запускается процесс конвертации с использованием библиотеки mammoth. Данная библиотека преобразует DOCX в семантически чистый HTML-код. Дополнительно применяется библиотека BeautifulSoup для автоматического форматирования кода (расстановки отступов и переносов строк), чтобы полученный HTML был удобен для дальнейшего ручного редактирования. После успешной конвертации

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

пользователь автоматически перенаправляется в окно редактирования шаблона.

В правой части главной формы располагаются две специализированные панели для импорта документов, представленные на рисунке 8.

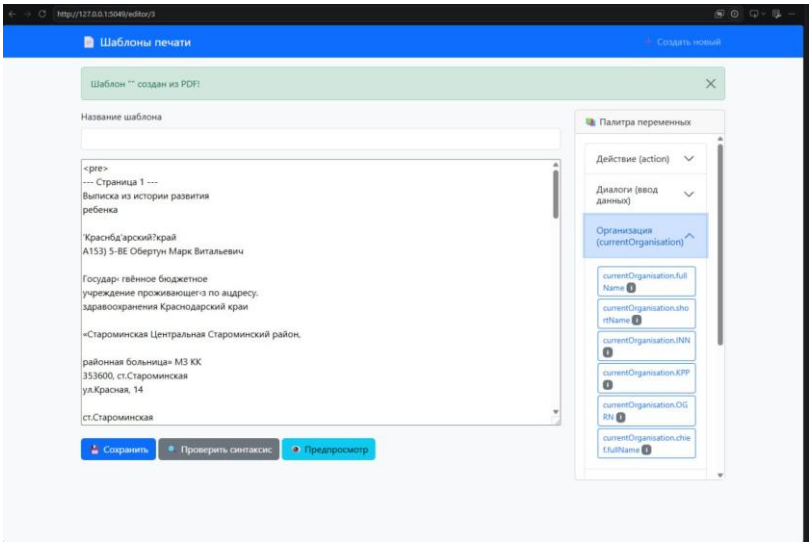


Рисунок 8 - Панели загрузки документов

Нижняя панель предназначена для загрузки PDF-файлов, в том числе отсканированных документов, не содержащих текстового слоя. Она содержит аналогичные поля ввода и кнопку «Загрузить и распознать». При нажатии на кнопку программа инициирует процесс оптического распознавания символов (OCR). На серверной стороне запускается скрипт, который с помощью библиотеки pdf2image и утилиты Poppler разбивает PDF-файл на отдельные изображения постранично. Затем каждое изображение передается в OCR-движок Tesseract (через обертку pytesseract), который распознает текст с поддержкой русского и английского языков.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Процесс распознавания больших файлов может занять некоторое время. Распознанный текст собирается в единую строку, разбитую по страницам, и оборачивается в тег <pre>, чтобы сохранить исходное форматирование и пробелы. По завершении процесса пользователь также переадресуется в редактор для дальнейшей работы.

В верхней части формы расположено поле для ввода или корректировки названия шаблона. Оно представляет собой однострочное текстовое поле, изменение значения в котором напрямую влияет на то, как шаблон будет отображаться в общем списке.

Центральную, самую объемную часть формы, занимает многострочный визуальный редактор HTML-кода (интегрированный в приложение CKEditor). Он позволяет пользователю работать с документом в режиме WYSIWYG (What You See Is What You Get - «что видишь, то и получишь»), не углубляясь в синтаксис языка разметки. Пользователь может добавлять таблицы, изменять шрифты, вставлять списки и выравнивать текст с помощью панели инструментов, расположенной сверху. Для опытных пользователей предусмотрена возможность переключения в режим исходного кода, где можно напрямую вносить правки в HTML-теги.

В правой части окна располагается панель «Палитра переменных», представляющая собой выпадающий список (аккордеон) с категориями переменных, используемых в МИС «САМСОН». Все переменные структурированы по логическим группам: «Действие», «Пациент», «Организация», «Диалоги» и др. При клике на любую переменную из списка её конструкция вида {переменная} автоматически вставляется в текст редактора в позицию установленного курсора. Это исключает

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

необходимость ручного набора длинных и сложных конструкций, исключает опечатки и значительно ускоряет работу.

В нижней части формы редактирования располагаются три кнопки управления. Кнопка «Проверить синтаксис» отправляет текущий HTML-код на сервер, где запускается скрипт проверки сбалансированности фигурных скобок { и }. Если количество открывающих и закрывающих скобок не совпадает, пользователю выводится всплывающее предупреждение с указанием точного количества найденных знаков, что позволяет быстро локализовать ошибку.

Кнопка «Предпросмотр» открывает новую вкладку с отображением того, как будет выглядеть готовый документ. Кнопка «Сохранить» инициирует запись внесенных изменений в базу данных SQLite. После успешного сохранения система обновляет метку времени updated_at и перенаправляет пользователя обратно на главную страницу, где измененный шаблон сразу становится доступен в общем списке.

2.4. Тестирование программного продукта

Тестирование разработанного программного модуля «Редактор шаблонов печати для МИС «САМСОН»» проводилось с целью проверки соответствия функциональных возможностей системы требованиям технического задания, а также для выявления и устранения ошибок на ранних этапах эксплуатации.

В связи с тем, что приложение включает в себя веб-сервер, десктопную обёртку и внешние библиотеки для распознавания текста, тестирование проводилось по нескольким ключевым направлениям: функциональное тестирование (проверка корректности работы основных

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

функций), интеграционное тестирование (проверка взаимодействия компонентов PyQt, Flask и SQLite), а также негативное тестирование (проверка поведения системы при некорректных действиях пользователя).

2.4.1. Методика и этапы проведения тестирования

Для проверки работоспособности системы использовался метод «черного ящика», при котором тестировщик (в роли которого выступал автор работы) взаимодействовал с графическим интерфейсом, не вникая в исходный код программы в момент проверки.

2.4.2. Негативное тестирование и обработка ошибок

Особое внимание уделялось проверке поведения системы при возникновении нештатных ситуаций. Программа должна корректно обрабатывать ошибки, не завершаясь аварийно и информируя пользователя о проблеме.

1. Загрузка файла неподдерживаемого формата. Если пользователь пытался загрузить файл .txt или .jpg через форму для DOCX, система выдавала всплывающее сообщение об ошибке с текстом «Разрешены только .docx файлы». Программа не зависала, база данных не повреждалась.

2. Ошибка распознавания PDF. При загрузке поврежденного PDF-файла или файла, не содержащего текста (например, фотография пейзажа), скрипт на Python перехватывал исключение (через блок try-except) и выводил сообщение: «Не удалось распознать текст из PDF. Проверьте файл и настройки». Пользователь оставался на главной странице.

3. Проверка синтаксиса с ошибкой. Если в шаблоне была открыта скобка {, но не было закрывающей }, система при проверке выдавала

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

красное предупреждение с указанием точного количества найденных знаков. Это позволило избежать загрузки «сломанного» шаблона в МИС.

4. Отсутствие соединения с сервером (PyQt + Flask). При закрытии приложения в диспетчере задач не оставалось «зависших» фоновых процессов Python. Это подтверждает корректную работу многопоточности (запуск Flask в отдельном потоке с флагом daemon=True).

2.4.3. Интеграционное и нагрузочное тестирование

В ходе проверки интеграции компонентов было подтверждено, что десктопная обёртка (PyQt5) и веб-сервер (Flask) работают слаженно. Данные, введенные в веб-интерфейсе, корректно сохраняются в локальную базу данных SQLite через ORM SQLAlchemy. Перезапуск приложения не приводил к потере ранее сохраненных шаблонов.

В рамках нагрузочного тестирования проверялась обработка многостраничных PDF-документов. Было установлено, что время распознавания отсканированного документа объемом 10 страниц занимает в среднем от 15 до 30 секунд. На время работы OCR-движка интерфейс программы оставался отзывчивым, так как обработка файлов происходит на стороне сервера, а клиентская часть ожидает ответа. Для ускорения работы пользователей с большими архивами в будущем возможна реализация пакетной обработки файлов.

2.4.4. Результаты тестирования и сборка приложения

По итогам проведенного тестирования критических ошибок, приводящих к потере данных или краху программы, выявлено не было. Все заявленные функции выполняются в полном объеме. Выявленные в процессе мелкие недочеты (например, некорректное отображение длинных

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

имен файлов в списке) были устранены в ходе отладки.

Завершающим этапом стала проверка собранной версии приложения. С помощью утилиты PyInstaller проект был скомпилирован в единый исполняемый файл main.exe. Тестирование показало, что собранный файл корректно запускается на компьютере без установленного Python и сторонних библиотек, успешно разворачивает локальный сервер и открывает интерфейс программы.

Заключение

В ходе прохождения производственной технологической практики в ГБУЗ «Староминская ЦРБ» МЗ КК был спроектирован и реализован программный модуль «Редактор шаблонов печати» для медицинской информационной системы «САМСОН», а также разработано приложение для автоматизации деятельности IT-отдела учреждения.

Тестовый запуск программы и анализ её работы на основе контрольных примеров (загрузка реальных документов и сканов) подтвердили её работоспособность. Процесс тестирования и отладки, включая проверку обработки нештатных ситуаций (загрузка файлов неподдерживаемого формата, ошибки распознавания или синтаксиса), позволил выявить и устранить недостатки, что привело к улучшению качества и надежности приложения.

В созданной программе был реализован следующий функционал:

1. Работа с локальной базой данных (SQLite) для хранения шаблонов и справочника переменных;
2. Конвертация текстовых документов формата DOCX в HTML-код с сохранением форматирования;
3. Распознавание текста из PDF-файлов (включая

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

отсканированные страницы) с использованием технологий OCR (Tesseract и Poppler);

4. Интерактивный визуальный редактор с палитрой переменных МИС «САМСОН» и встроенной проверкой синтаксиса;

5. Возможность редактирования, предпросмотра, сохранения и удаления данных.

Разработка приложения была осуществлена с использованием языка программирования Python, веб-фреймворка Flask и библиотеки PyQt5 для создания десктопной обёртки, а в качестве среды разработки применялся Visual Studio Code. Использование данного стека технологий обеспечило удобство в написании кода и отладке, а также кроссплатформенность решения. Это позволило сосредоточиться на реализации функционала и улучшении пользовательского интерфейса, что в конечном итоге положительно сказалось на качестве конечного продукта, готового к внедрению в работу отдела компьютерно-сетевого обеспечения больницы.

Всероссийское СМИ
«Академия педагогических идей «НОВАЦИЯ»
Свидетельство о регистрации ЭЛ №ФС 77-62011 от 05.06.2015 г.
(выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций)

Сайт: akademnova.ru
e-mail: akademnova@mail.ru

Список использованной литературы:

1. Гвоздева, В. А. Информатика, автоматизированные информационные технологии и системы: учебник / В. А. Гвоздева. - Москва: ФОРУМ: ИНФРА-М, 2022. - 542 с.
2. Гринберг, М. Разработка веб-приложений с использованием Flask на языке Python: руководство / М. Гринберг. - Москва: ДМК Пресс, 2023. - 274 с.
3. Егошина, И. Л. Информационные системы и технологии в здравоохранении: учебное пособие / И. Л. Егошина. - Йошкар-Ола: Поволжский государственный технологический университет, 2021. - 90 с.
4. Омельченко, В. П. Информатика, медицинская информатика, статистика: учебник / В. П. Омельченко, А. А. Демидова. - Москва: ГЭОТАР-Медиа, 2021. - 608 с.
5. Прохоренок, Н. А. Python 3 и PyQt 5. Разработка приложений / Н. А. Прохоренок, В. А. Дронов. - Санкт-Петербург: БХВ-Петербург, 2021. - 832 с.
6. Лукин, В. В. Технология разработки программного обеспечения: учебное пособие / В. В. Лукин. - Москва: Гостехиздат, 2021. - 208 с.
7. Мезенцев, К. Н. Автоматизированные информационные системы: учебник / К. Н. Мезенцев. - Москва: Академия, 2021. - 176 с.
8. Рудаков, А. В. Технология разработки программных продуктов: учебник / А. В. Рудаков. - Москва: Академия, 2022. - 208 с.
9. Парфилова, Н. И. Программирование. Структурирование программ и данных: учебник / Н. И. Парфилова, А. Н. Пылькин, Б. Г. Трусов. - Москва: Академия, 2022. - 256 с.
10. Хетагуров, Я. А. Проектирование автоматизированных систем обработки информации и управления (АСОИУ): учебник / Я. А. Хетагуров. - Москва: Бином. Лаборатория знаний, 2022. - 327 с.
11. Канцедал, С. А. Алгоритмизация и программирование: учебное пособие / С. А. Канцедал. - Москва: ФОРУМ: ИНФРА-М, 2024. - 352 с.
12. Соколова, В. В. Вычислительная техника и информационные технологии: учебное пособие / В. В. Соколова. - Москва: Юрайт, 2023. - 253 с.
13. Шапиро, Л. Компьютерное зрение: учебное пособие / Л. Шапиро, Дж. Стокман. - Москва: Лаборатория знаний, 2023. - 763 с.
14. Дейт, К. Дж. Введение в системы баз данных / К. Дж. Дейт. - Москва: Вильямс, 2021. - 1328 с.
15. Тормасов, А. Г. Параллельное программирование многопоточных систем с разделяемой памятью / А. Г. Тормасов. - Москва: Физматкнига, 2021. - 288 с.

Дата поступления в редакцию: 19.10.2026 г.
Опубликовано: 19.10.2026 г.

© Академия педагогических идей «Новация», электронный журнал, 2026
© Асманов Г.М., 2026