

Санкт-Петербургский государственный университет

Системное программирование

Группа 24.М71-мм

Автоматический экспорт HTML из облачных электронных таблиц

Малыгин Даниил Александрович

Отчёт по учебной практике в форме «Решение»

Научный руководитель:

доцент кафедры системного программирования, к.ф.-м.н. Д.В. Луцев

Санкт-Петербург

2025

Оглавление

Введение	3
Постановка задачи	4
1. Обзор	5
1.1. Microsoft OneDrive	5
1.2. Google Sheets	5
1.3. CloudConvert	6
2. Описание реализации	7
2.1. Архитектура решения	7
2.2. Требования к сервису	8
2.3. Предлагаемые инструменты	8
2.3.1. Django	9
2.4. Ядро проекта	10
2.5. Масштабируемость и возможность лёгкого развёртывания	13
Заключение	14
Список литературы	15

Введение

Таблицы были, есть и будут одним из удобных и структурированных способов как хранения, так и отображения данных. Хорошее применение они находят в том числе в веб-сайтах с отсутствующей серверной логикой от сервисов вроде Google Siter, WordPress, Tilda и тд. В таких сервисах эффективным является встраивание HTML-страниц, представляющих собой таблицу, из внешних ресурсов (например, Google Drive, OneDrive, Yandex Disk) для своевременного обновления содержимого страниц.

Казалось бы, хорошее, эффективное и простое решение! Но имеется главный и очевидный изъян: подобного рода сервисы работают исключительно с загруженными на их сервер файлами. Из-за это сужается набор используемых сервисов облачного хранилища.

Данная проблема подталкивает на создание решения, которое бы лишало пользователя зависимости от поставщиков. Таким решением является сервис, выполняющему следующий функционал:

- Принимал публичную ссылку на хранимую в облаке таблицу.
- Конвертировал эту самую таблицу в HTML-страницу с данными в форме таблицы.
- Загружал эту самую HTML-страницу в браузер.
- Автоматически выполнял обновления HTML-страницы.

Особенностью такого сервиса являются следующие свойства:

- Взаимодействие с популярными сервисами облачных документов: Google Drive, OneDrive, Yandex Disk.
- Доступность решения для использования неподготовленным пользователем.
- Удобный интерфейс.

Постановка задачи

Цель работы — проектирование и разработка веб-сервиса, обеспечивающего встраивание таблиц из облачных ресурсов в веб-сайты.

В рамках данной работы необходимо выполнение следующих задач:

1. Выполнить обзор существующих решений и доступных технологий.
2. Проанализировать требования к сервису.
3. На основе требований к решению и выводов обзора выработать архитектуру сервиса.
4. Реализовать ядро сервиса.
5. Обеспечить возможность развёртывания сервиса.
6. Провести тестирование и апробацию.

1. Обзор

В рамках данного обзора будут рассмотрены существующие реализации, способные импортировать электронные таблицы из облачного хранилища, а также экспортировать HTML.

Среди имеющихся реализаций можно выделить следующие:

- Microsoft OneDrive;
- Google Sheets;
- CloudConvert;

1.1. Microsoft OneDrive

OneDrive [1] – это онлайн-сервис от компании Microsoft, который обеспечивает возможность сохранения, синхронизации и обмена различными файлами через Интернет. Данный сервис полностью интегрирован с операционной системой Windows, что позволяет легко работать с файлами через файловый менеджер.

Проблема данного сервиса заключается в создании HTML-таблицы в виде кода `iframe`, используемого для встраивания в блог пользователя. Хотя метод и рабочий, но требует слишком много шагов для получения результата, плюс не совсем интуитивно очевидный. И это с учётом того, что работает это только с файлами, загруженными в сам сервис.

1.2. Google Sheets

Google Sheets [3] – аналог Excel в режиме онлайн от транснациональной корпорации Google. Это программа, которая помогает систематизировать информацию в табличном виде, считать простые и сложные примеры, решать математические, финансовые и статистические задачи.

Основная функция Google Sheets – создание таблиц и работа с ними совместно с разными пользователями. Главное его преимущество в данном случае – бесплатный доступ.

В случае данного сервиса, в сравнение с OneDrive, имеется возможность получить HTML-таблицу благодаря редактированию ссылки на исходную таблицу. Например, пусть оригинальная ссылка на изначальную таблицу выглядит следующим образом - https://docs.google.com/spreadsheets/d/1WYTl1wFRcuczEwWEtbcxlQPhHDpK_0tJT3Q9TATI-bU/edit?gid=0#gid=0, а чтобы получить HTML-страницу с таблицей внутри необходимо переделать ссылку на https://docs.google.com/spreadsheets/d/1WYTl1wFRcuczEwWEtbcxlQPhHDpK_0tJT3Q9TATI-bU/gviz/tq?tqx=out:html. При этом даже вне изменения в исходной таблице отображаются в измененной ссылке!

Но главная проблема в том, что данная возможность неизвестна пользователю, плюс сервис не даёт автоматизированное предоставление подобной ссылки. Из-за всё редактирование ссылки возлагается на пользователя. И это не упоминания возможность делать только с файлами, загруженными на сервис от Google.

1.3. CloudConvert

CloudConvert [4] позволяет конвертировать более 140 различных форматов в браузере без загрузки приложений на ПК. Удобной возможностью этого приложения является то, что можно выбрать файлы с аккаунта в облаке Google Drive или Dropbox и конвертировать их без сохранения файла на ПК.

В сравнение с предыдущими реализациями, она реализует работу с файлами, расположенных в разных облачных хранилищах.

Единственная проблема сервиса – он занимается просто конвертацией файла из одного формата в другой, результатом которого является сам файл. Это означает, что полученная HTML-страница не будет получать изменения, выполненные на изначальной таблице, а также её нельзя будет встроить в создаваемые веб-сайты.

2. Описание реализации

2.1. Архитектура решения

Предлагаемая архитектура сервиса состоит из двух основных компонентов:

1. Компонент аутентификации, который включает процедуры регистрации, входа в личный кабинет и выхода из него.
2. Компонент личного кабинета, в котором пользователь может взаимодействовать с файлами: создавать с помощью отправки ссылки на свою таблицу, удалять созданную HTML-страницу;
3. Компонент работы с файлом, который и отвечает за основные манипуляции с ними: получение из облака исходного файла с таблицей, генерация из него HTML-страницы, обновление, удаление и отображение самой HTML-страницы.

Диаграмма компонентов предлагаемого решения представлена на рис. 1.

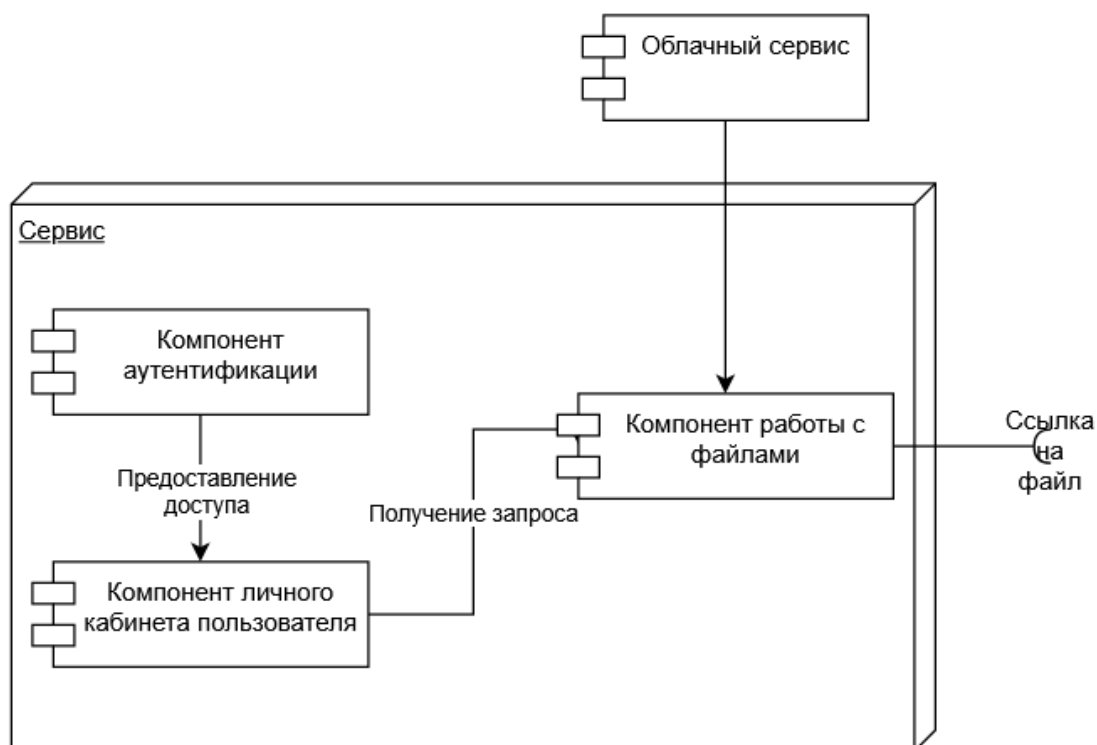


Рис. 1: Диаграмма компонентов

2.2. Требования к сервису

Предлагаемый к реализации сервис должен включать следующие требования:

1. Добавление ссылок на таблицу в облаке для генерации HTML-страницы, отображение и удаление сгенерированных страниц.
2. Извлечение файла из любого облачного хранилища.
3. Отображение по ссылке готовой HTML-страницы.
4. Автоматизированное обновление HTML-страницы в случае обновления изначальной таблицы в облаке.

Из нефункциональных требований можно выделить следующее:

1. Интуитивно понятный дизайн веб-приложения для взаимодействия с ним пользователя.
2. Визуальный дизайн, который был бы приятным глазам пользователя.

2.3. Предлагаемые инструменты

Проект было решено реализовать в виде веб-приложения. Это связано с тем, что необходимо от пользователя получить ссылку на изначальный файл, а также отображать на браузере готовую HTML-страницу с необходимой таблицей.

Благодаря этому можно выделить следующие инструменты:

- Основной язык программирования Python версии выше 3. Выбор обусловлен легкостью работы с ним и доступностью необходимой среды разработки.
- Интегрированная среда разработки (IDE) PyCharm из-за возможности создания проектов на языке Python.
- Веб-фреймворк Django.
- Библиотеки для скачивания файлов из облака в лице OneDriveLoader[5], который используется для загрузки файлов из OneDrive, и

Requests[6], который используется в проекте для отправки запроса на скачивание файлов из Yandex Disk с помощью ссылки конкретной формы.

2.3.1. Django

Django — это фреймворк, с помощью которого можно реализовывать на Python сайты и приложения, которые работают в браузере [7].

Он предоставляет следующие основные возможности:

- Настроенный веб-сервер, который будет обрабатывать запросы от пользователей к веб-сервису.
- Готовые механизмы для авторизации пользователей.
- Простые шаблоны веб-страниц.
- Административный интерфейс для управления контентом сервиса — наполнения, изменения, обновления используемых данных.
- Система кэширования для увеличения скорости загрузки и открытия страниц через браузеры, внешние клиенты или приложения.
- Интерфейсы и адаптеры для подключения к различным типам баз данных.

С помощью Django можно очень быстро, как из конструктора, настроить и запустить работающий веб-сервис — а потом программировать только специфичные функции и бизнес-логику.

Django работает по модели MVT — Model-View-Template, или «Модель–Представление–Шаблон». Она разделяет внутреннюю логику работы сайта, внешний вид страниц для пользователя и реакции веб-сервиса на внешние воздействия:

- Модель обеспечивает внутреннюю работу сайта: подключение к базе данных, формат хранения информации и другие функции.
- Представление определяет данные, которые нужно показать пользователю, и отправляет их туда, где они должны быть показаны. Также

оно принимает запросы пользователей и реагирует на них — например, обновляет страницу после отправки данных формы.

- Шаблон — это каркас страницы на Django, которую увидит пользователь.

Выбор данного фреймворка обусловлен следующими факторами [8]:

- Работа с языком Python;
- Является открытым программным обеспечением;
- Просто для изучения;
- Имеет замечательное сообщество.

2.4. Ядро проекта

В проекте реализованы следующие приложения проекта:

- authentication, который требует ограничивает доступ к личному кабинету.
- account - личный кабинет с необходимым функционалом работы с файлами для пользователя.
- periodic – компонент выполнения основных операций с файлами, в том числе обновление HTML-страницы, но при получении запроса от пользователя, то есть без реализации автоматического механизма.

Специально в приложении periodic для работы с HTML-страницами создана база данных на СУБД SQLite3. Также используется автоматически созданная для проекта таблица для аутентификации пользователей.

В данный момент проект развёртывается благодаря сервису CloudPub, отвечающего за создание безопасного туннеля.

Реализован также набор HTML-файлов для отображения страниц сайта (рис. 2-8).

Добро пожаловать!

Данный сервис создан для конвертации ваших хранимых в облаке электронных таблиц в HTML-файл.

Для конвертации вашего текущего файла вам необходимо в свой существующий личный кабинет или создать новый.

Рис. 2: Главная страница личного кабинета аутентификации

Логин:

Эл. почта:

Пароль:

Подтвердите пароль:

Рис. 3: Вариант страницы входа с формой регистрации

Логин:

Пароль:

Рис. 4: Вариант страницы входа с формой логина

Рады вас видеть, daniil!

У вас отсутствуют конвертированные файлы!

Добавить таблицу

Выйти

Удалить аккаунт

Рис. 5: Главная страница личного кабинета

Рады вас видеть, daniil!

Список файлов

№	Название	Ссылка на отображение	Опции
20	Yandex Disk	Файл	Изменить Обновить файл Удалить

Добавить таблицу

Выйти

Удалить аккаунт

Рис. 6: Главная страница личного кабинета с имеющимися файлами

Имя таблицы:

Ссылка на источник:

Подтвердить

Назад

Рис. 7: Вариант страницы внесения изменений файла с формой добавления

Имя таблицы:

Подтвердить

Назад

Рис. 8: Вариант страницы внесения изменений файла с формой редактирования

Ниже представлен пример отображения созданной HTML-страницы (рис. 9).

Показатели самоконтроля	2022-11-30 00:00:00	2022-12-01 00:00:00	2022-12-02 00:00:00	2022-12-03 00:00:00	2022-12-04 00:00:00	2022-12-05 00:00:00	2022-12-06 00:00:00
Сон	-	-	-	-	-	-	-
Аппетит	Отличный	Отличный	Отличный	Отличный	Отличный	Отличный	Отличный
Настроение	удовлетворительное	удовлетворительное	удовлетворительное	удовлетворительное	удовлетворительное	удовлетворительное	удовлетворительное
Самочувствие	хорошее	хорошее	хорошее	хорошее	хорошее	хорошее	хорошее
Содержание занятия	Отдых	Отдых	Отдых	отдых	отдых	Отдых	Отдых
Потоотделение (во время занятия)	-	-	-	-	-	-	-
Болевые ощущения	нет	нет	нет	нет	нет	нет	нет
Перенес нагрузку	-	-	-	-	-	-	-
Пульс	-	-	-	-	-	-	-
Восстановимость пульса	-	-	-	-	-	-	-
Масса тела, кг	123	116	116	116	116	116	116
Нарушения режима	-	-	-	-	-	-	-
Другая физическая активность в течении дня	-	-	-	-	-	-	-

Рис. 9: Пример отображения сгенерированной HTML-страницы

Текущий код проекта опубликован на ресурсе GitHub [9].

2.5. Масштабируемость и возможность лёгкого развёртывания

В план масштабирования проекта планируется:

- Автоматизировать обновление HTML-страниц.
- Внедрение CCS-кода для создания необходимых требований касательно внешнего вида.
- Увеличить безопасность, связанной с доступом к личным данным пользователей.
- Развернуть веб-приложение благодаря nginx.

Заключение

В текущий момент времени проработана предлагаемой для реализации идеи и частично спроектировано ядро сервиса.

Из обозначенных изначально задач получилось выполнить следующие:

1. Проведён анализ функциональности существующих сервисов и возможностей доступных технологий;
2. Выработаны требования к сервису;
3. Спроектирована архитектура сервиса;
4. Частично реализовано ядро сервиса.

Список литературы

- [1] Умнова А. One Drive — что такое. — 2024. — URL: <https://skyeng.ru/magazine/wiki/it-industriya/chto-takoe-one-drive/> (дата обращения: 2025-01-10).
- [2] Share it: Embed an Excel workbook on your web page or blog from OneDrive. — URL: <https://support.microsoft.com/en-us/office/share-it-embed-an-excel-workbook-on-your-web-page-or-blog-from-onedrive-804e1845-5662-487e-9b38-f96307144081> (дата обращения: 2025-01-09).
- [3] Кирякова А. Гайд по Google Sheets: универсальный инструмент планирования. — 2022. — URL: <https://4brain.ru/blog/gajd-po-google-sheets-universalnyj-instrument-planirovaniya/> (дата обращения: 2025-01-10).
- [4] Leonardo Конвертируем в "облаке" любые файлы в любые другие форматы при помощи CloudConvert. — 2013. — URL: <https://webtun.com/webservices/5443-convert-any-files-to-any-other-format-in-cloud-with-cloudconvert.html> (дата обращения: 2025-01-09).
- [5] OneDriveLoader. — URL: <https://webtun.com/webservices/5443-convert-any-files-to-any-other-format-in-cloud-with-cloudconvert.html> (дата обращения: 2025-09-24).
- [6] Requests: HTTP for Humans™. — URL: <https://docs.python-requests.org/en/latest/index.html#requests-http-for-humans> (дата обращения: 2025-09-24).
- [7] Шпрингер Л., Павлова А., Федосеев Н. Django «освобождающий»: как фреймворк помогает быстро разрабатывать веб-сервисы на Python. — 2023. — URL: <https://practicum.yandex.ru/blog/framework-django/> (дата обращения: 2025-06-17).
- [8] When to Use Django? Comparison with other Development Stacks. — URL: <https://www.geeksforgeeks.org/python/when-to-use-django-comparison-with-other-development-stacks/> (дата обращения: 2025-06-17).

[9] Программа проекта. – URL:
<https://github.com/MalyginDaniil2002/HTML-Converter.git> (дата обращения:
2025-09-24).