

Автоматический экспорт HTML из облачных электронных таблиц

МАЛЫГИН Д. А., 24.М71

НАУЧНЫЙ РУКОВОДИТЕЛЬ: К.Ф.-М.Н. Д.В. ЛУЦИВ, ДОЦЕНТ КАФЕДРЫ
СИСТЕМНОГО ПРОГРАММИРОВАНИЯ

САНКТ-ПЕТЕРБУРГ

2025

- ▶ 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.
- ▶ Есть проблема - работа исключительно с загруженными на их сервер файлами.
- ▶ Решение – веб-сервис!

- ▶ Предлагаемый сервис должен:
 - Принимать публичную ссылку на хранимую в облаке таблицу
 - Конвертировать 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

Цель и задачи

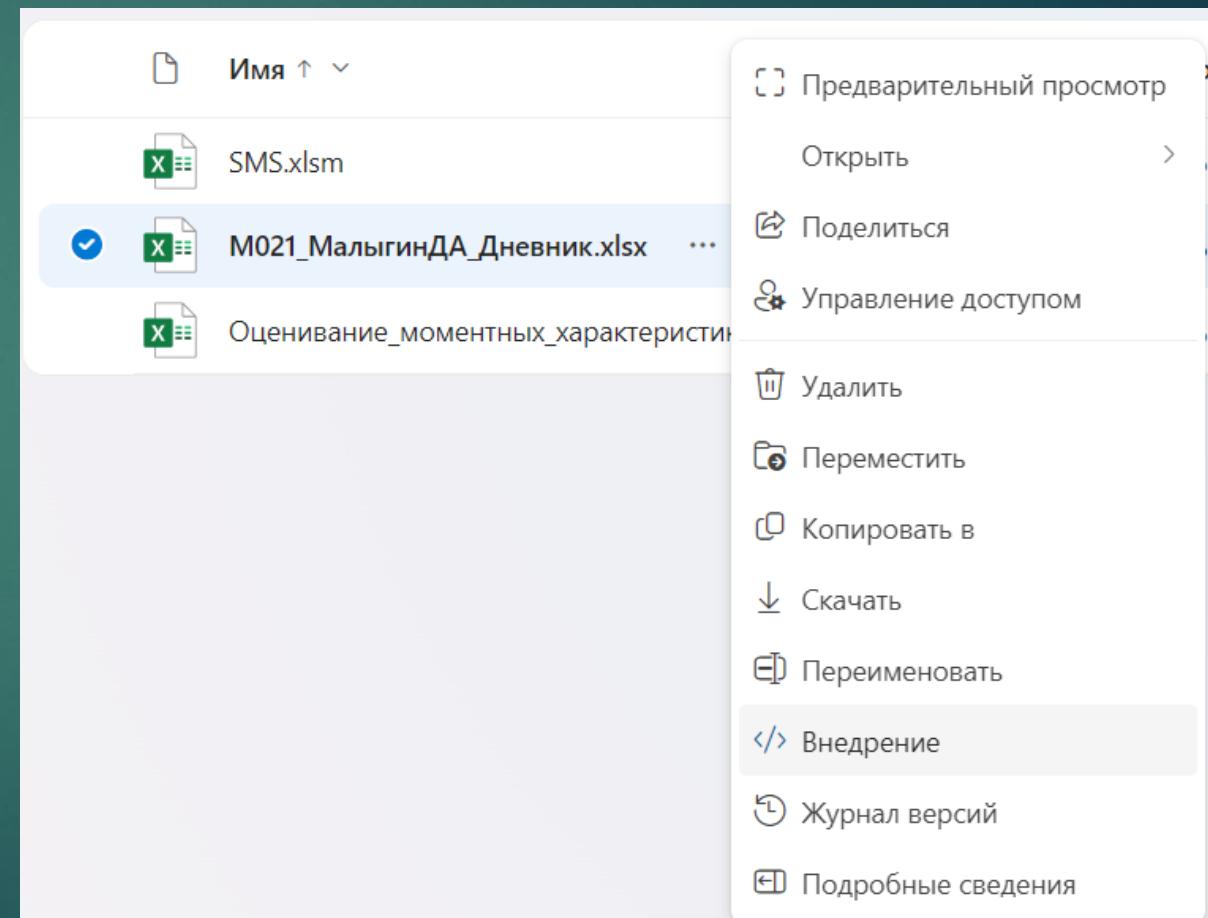
- ▶ Цель работы — проектирование и разработка веб-сервиса, обеспечивающего встраивание таблиц из облачных ресурсов в веб-сайты.
- ▶ В рамках данной работы необходимо выполнение следующих задач:
 - Выполнить обзор существующих решений и доступных технологий
 - Проанализировать требования к сервису
 - На основе требований к решению и выводов обзора выработать архитектуру сервиса
 - Реализовать ядро сервиса
 - Обеспечить возможность развёртывания сервиса
 - Провести тестирование и апробацию

Анализ существующих реализаций

- ▶ В данном перечне реализаций, способных импортировать электронные таблицы из облачного хранилища , а также экспортировать HTML, можно выделить следующие:
 - Microsoft OneDrive
 - Google Sheets
 - CloudConvert
 - Сайт ru.cloud-converter.com
 - Dropbox
- ▶ Можно отметить, что последние две реализации хоть и предоставляет необходимый сервис, но требуют за него за плату, из-за чего не подходят

Анализ существующих реализаций (Microsoft OneDrive)

- OneDrive – онлайн-сервис от компании Microsoft, который обеспечивает возможность сохранения, синхронизации и обмена различными файлами через Интернет



Анализ существующих реализаций (Microsoft OneDrive)

- ▶ Как можно понять из описания, таблицу в формате HTML-страницы можно использовать лишь для встраивания её в блог
- ▶ Хоть 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

The screenshot shows a Microsoft Excel spreadsheet titled 'Лист1'. The table has column A and row numbers 1 through 11. The data includes:

A	
1	Показатели самоконтроля
2	Сон
3	Аппетит
4	Настроение
5	Самочувствие
6	Содержание занятия
7	Потоотделение (во время занятия)
8	Болевые ощущения
9	Перенес нагрузку
10	Пульс
11	Восстановимость пульса

Below the table, an 'iframe' tag is displayed in the formula bar:

```
<iframe src="https://1drv.ms/x/c/76544d8317217ed4/IQ7UfiEXg01UIIB2O3gAAAAAQW5FZoPNI42NnDqhXYauPk" width="402" height="346" frameborder="0" scrolling="no"></iframe>
```

Анализ существующих реализаций (Google Sheets)

- ▶ Google Sheets – аналог Excel в режиме онлайн от транснациональной корпорации Google
- ▶ В случае данного сервиса, в сравнение с OneDrive, имеется возможность получить HTML-таблицу благодаря редактированию ссылки на исходную таблицу

A	B	C
Показатели самоконтроля	11/30/2022	12/1/2022
Сон	-	-
Аппетит	Отличный	Отличный
Настроение	удовлетворительное	удовлетворительное
Самочувствие	хорошее	хорошее
Содержание занятия	Отдых	Отдых
Потоотделение (во время занятия)	-	-
Болевые ощущения	нет	нет
Перенес нагрузку	-	-
Пульс	-	-
Восстановимость пульса	-	-
Масса тела, кг	116	116
Нарушения режима	-	-
Другая физическая активность в течении дня	-	-

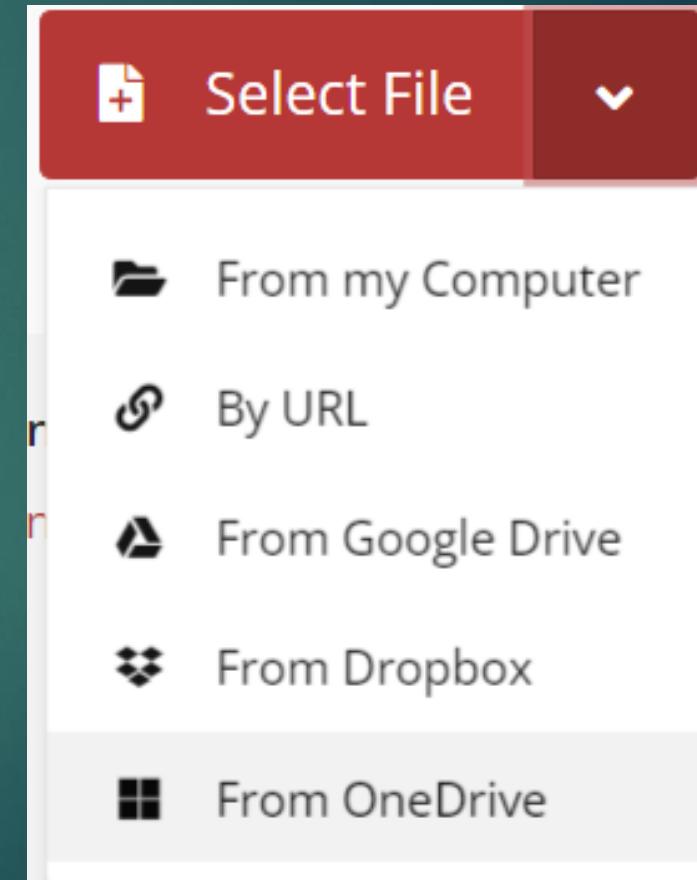
Анализ существующих реализаций (Google Sheets)

- ▶ Пример:
 - оригинальная ссылка -
https://docs.google.com/spreadsheets/d/13IDxvFZ9Bh84xCL4dljQHWHp3k_BaZLq/edit?gid=42637984#gid=42637984
 - Ссылка на HTML-страницу таблицы -
https://docs.google.com/spreadsheets/d/1WYTI1wFRcuczEwWEtbcxIQPhHDpK_0tJT3Q9TATI-bU/gviz/tq?tqx=out:html
- ▶ В данном способе даже отображаются изменения в изначальной таблице

Показатели самоконтроля	11/30/2022	12/1/2022
Сон	-	-
Аппетит	Отличный	Отличный
Настроение	удовлетворительное	удовлетворительное
Самочувствие	хорошее	хорошее
Содержание занятия	Отдых	Отдых
Потоотделение (во время занятия)	-	-
Болевые ощущения	нет	нет
Перенес нагрузку	-	-
Пульс	-	-
Восстановимость пульса	-	-
Масса тела, кг	116	116
Нарушения режима	-	-
Другая физическая активность в течении дня	-	-

Анализ существующих реализаций (CloudConvert)

- ▶ CloudConvert 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



Анализ существующих реализаций (CloudConvert)

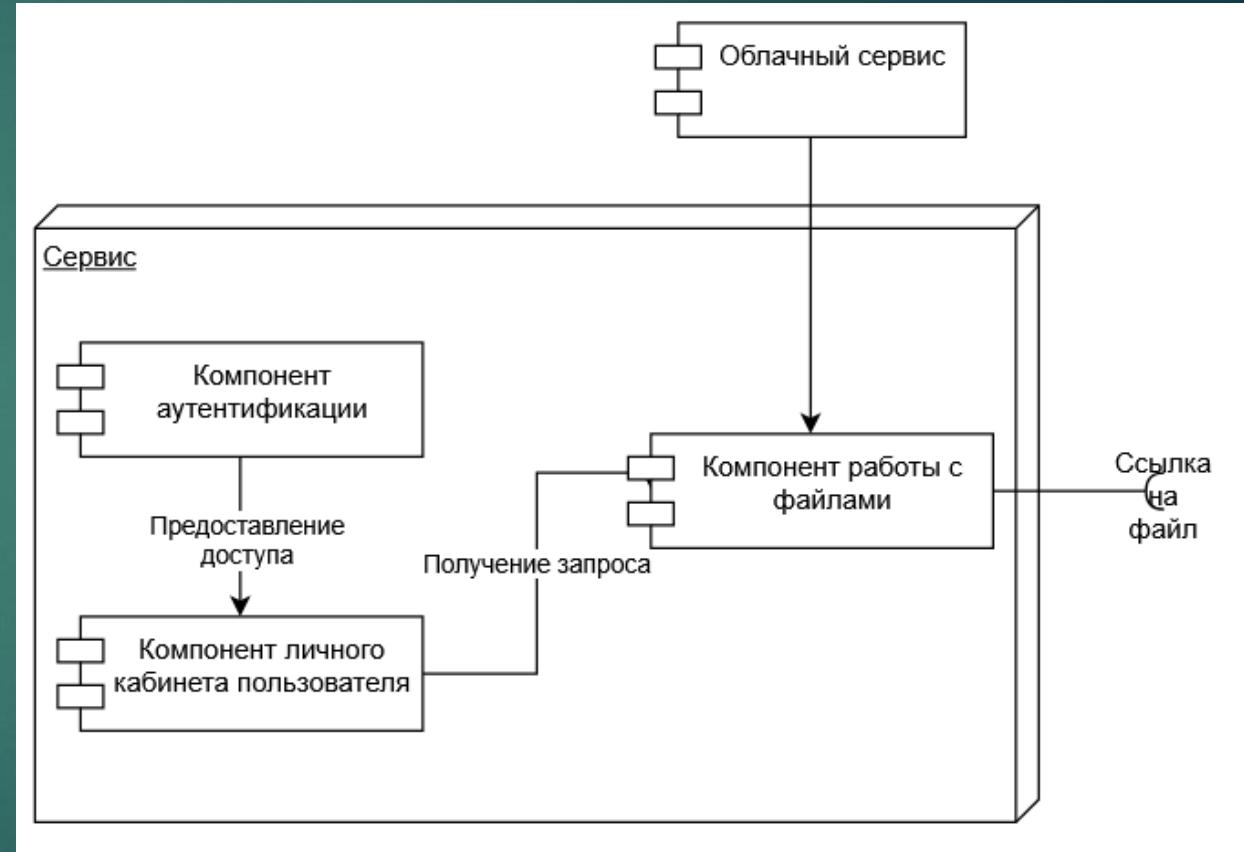
- ▶ Единственная проблема сервиса – он занимается просто конвертацией файла из одного формата в другой, результатом которого является сам файл
- ▶ Это означает, что полученная HTML-страница не будет получать изменения, выполненные на изначальной таблице, а также её нельзя будет встроить в создаваемые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

Показатели самоконтроля	11/30/2022	12/1/2022	12/2/2022
Сон	-	-	-
Аппетит	Отличный	Отличный	Отличный
Настроение	удовлетворительное	удовлетворительное	удовлетворительное
Самочувствие	хорошее	хорошее	хорошее
Содержание занятия	Отдых	Отдых	Отдых
Потоотделение (во время занятия)	-	-	-
Болевые ощущения	нет	нет	нет
Перенес нагрузку	-	-	-
Пульс	-	-	-
Восстановимость пульса	-	-	-
Масса тела, кг	116	116	116
Нарушения режима	-	-	-
Другая физическая активность в течении дня	-	-	◆

Архитектура сервиса

12/21

- ▶ Предлагаемая архитектура сервиса состоит из двух компонентов:
 - Компонент аутентификации – процедуры регистрации, входа в личный кабинет и выхода из него
 - Компонент личного кабинета – инструмент взаимодействия пользователей с файлами
 - Компонент работы с файлом – взаимодействие с файлами



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.

Предлагаемые инструменты

14/21

- ▶ 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, который используется для загрузки файлов из OneDrive, и Requests, который используется в проекте для отправки запроса на скачивание файлов из Yandex Disk с помощью ссылки конкретной формы.

Предлагаемые инструменты (Django)

- ▶ Django — фреймворк, с помощью которого можно реализовывать на Python сайты и приложения, которые работают в браузере
- ▶ Он предоставляет следующие основные возможности:
 - Настроенный веб-сервер для обработки запросов пользователей к веб-сервису
 - Готовые механизмы для авторизации пользователей
 - Простые шаблоны веб-страниц

Предлагаемые инструменты (Django)

- Административный интерфейс для управления контентом сервиса
- 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

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
- ▶ Специально в приложении periodic для работы с HTML-страницами создана база данных на СУБД SQLite3, а также используется автоматически созданная для проекта таблица для аутентификации пользователей
- ▶ В данный момент проект развёртывается благодаря сервису CloudPub, отвечающего за создание безопасного туннеля

Масштабируемость и возможность лёгкого развёртывания

- ▶ В план масштабирования проекта планируется:
 - Автоматизировать обновление HTML-страниц.
 - Внедрение CSS-кода для создания необходимых требований касательно внешнего вида.
 - Увеличить безопасность, связанной с доступом к личным данным пользователей.
 - Развернуть веб-приложение благодаря nginx.

Ссылка на веб-приложение

19/21



Ссылка на GitHub проекта

20/21



- ▶ В текущий момент времени проработана предлагаемой для реализации идеи и частично спроектировано ядро сервиса
- ▶ Пока можно отметить, что выделены следующие задачи:
 - Проведён анализ функциональности существующих сервисов и возможностей доступных технологий
 - Выработаны требования к сервису
 - Спроектирована архитектура сервиса
 - Частично реализовано ядро сервиса