

# 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

- ▶ Воспользоваться HTML-таблицами можно везде, где необходимо структурировать информацию и отобразить ее в понятной пользователю форме
- ▶ Но в процессе работы могут возникать следующие задачи для получения подобной таблицы:
  - Необходимую HTML-таблицу требуется сгенерировать из существующей электронной таблицы
  - Извлечение в будущем используемой в процессе генерации электронной таблицы из облачного хранилища

# Цель и задачи

3/16

- ▶ Цель работы заключается в реализации независимого от поставщика сервиса публикации облачных электронных таблиц в вебе
- ▶ В рамках данной работы необходимо выполнение следующих задач:
  - Провести анализ функциональности существующих сервисов и возможностей доступных технологий
  - Выработать требования к сервису
  - Спроектировать архитектуру сервиса
  - Реализовать сервис
  - Провести тестирование и пробу сервиса

# Анализ существующих реализаций

4/16

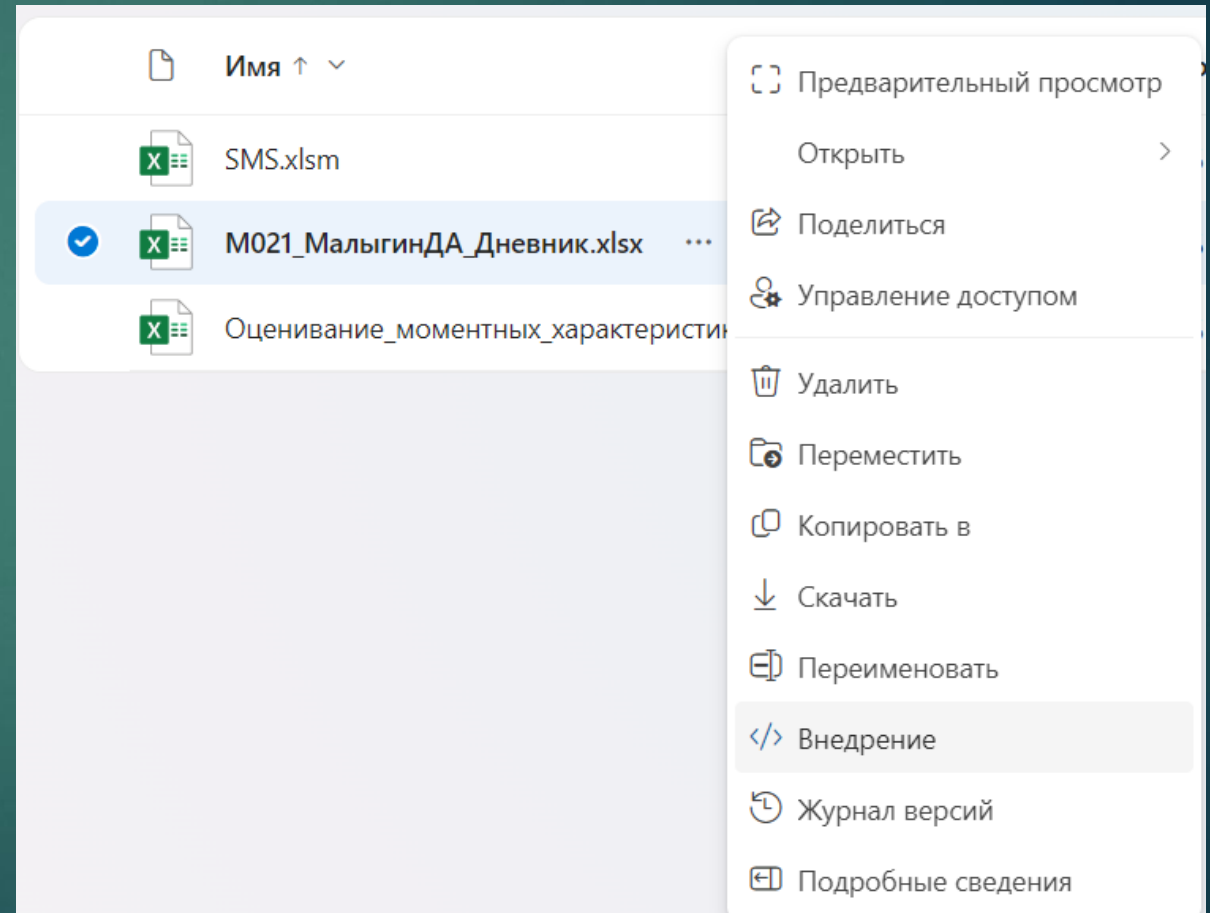
- ▶ В данном перечне реализаций, способных импортировать электронные таблицы из облачного хранилища , а также экспортировать HTML, можно выделить следующие:
  - Microsoft OneDrive
  - Google Sheets
  - CloudConvert
  - Сайт [ru.cloud-converter.com](http://ru.cloud-converter.com)
  - Dropbox
- ▶ Можно 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)

5/16

► Получения таблицы необходимо выполнить следующее:

- 1) Открыть веб-сайт OneDrive и перейти к рабочей книге
- 2) Выбрать опцию «Внедрение» для получения iframe-кода



# Анализ существующих реализаций (Microsoft OneDrive)

6/16

- 3) Скопировать источник iframe
  - 4) Вставить рабочую книгу путем переключения на редактирование блога в формате HTML и вставить код вставки
- ▶ Как можно понять из описания, HTML-таблицу можно использовать лишь для встраивания её в блог
  - ▶ Таким образом получение отдельной веб-страницы из данных электронной таблицы, расположенной в облачном хранилище, невозможно
  - ▶ Плюс реализация в виде копирования кода не является удобным с точки зрения пользователя вариантом

The screenshot displays a mobile interface of a Microsoft OneDrive document. At the top, a table with 11 rows and 2 columns is visible. The first column contains various health-related indicators, and the second column is empty. The row 'Болевые ощущения' (Pain) is highlighted in green. Below the table, the source code for an iframe is shown, which embeds the table into a web page.

	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

```
<iframe  
src="https://1drv.ms/x/c/76544d8317217  
ed4/IQTUfiEXg01UIIB2O3gAAAAAAQW5F  
ZoPNI42NnDqhXYauPk" width="402"  
height="346" frameborder="0"  
scrolling="no"> </iframe>
```

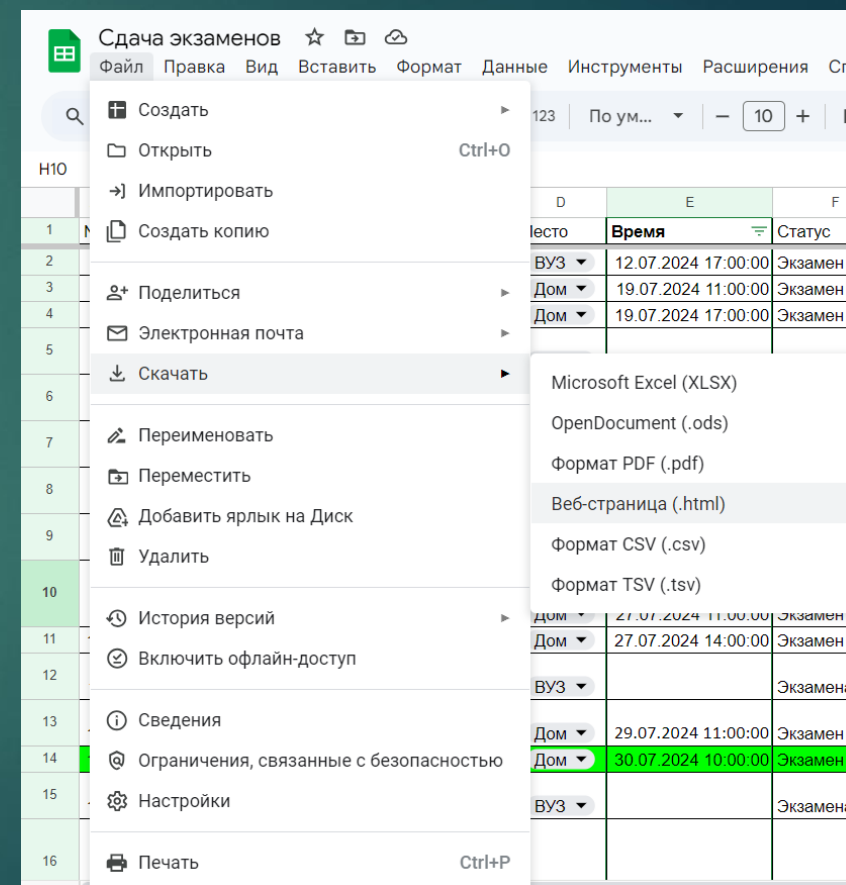


# Анализ существующих реализаций (Google Sheets)

7/16

► Чтобы получить HTML-таблицу требуется выполнить следующую последовательность действий:

- 1) Скачать электронную таблицу в формате «Веб-страница» (.html)
- 2) Открыть полученный архив и разархивировать его
- 3) Открыть HTML-документ для браузера



# Анализ существующих реализация (Google Sheets)

8/16

- ▶ В отличие от реализации в OneDrive, таблицу можно получить в виде отдельной веб-страницы
- ▶ Единственным минусом является получении результата через архив

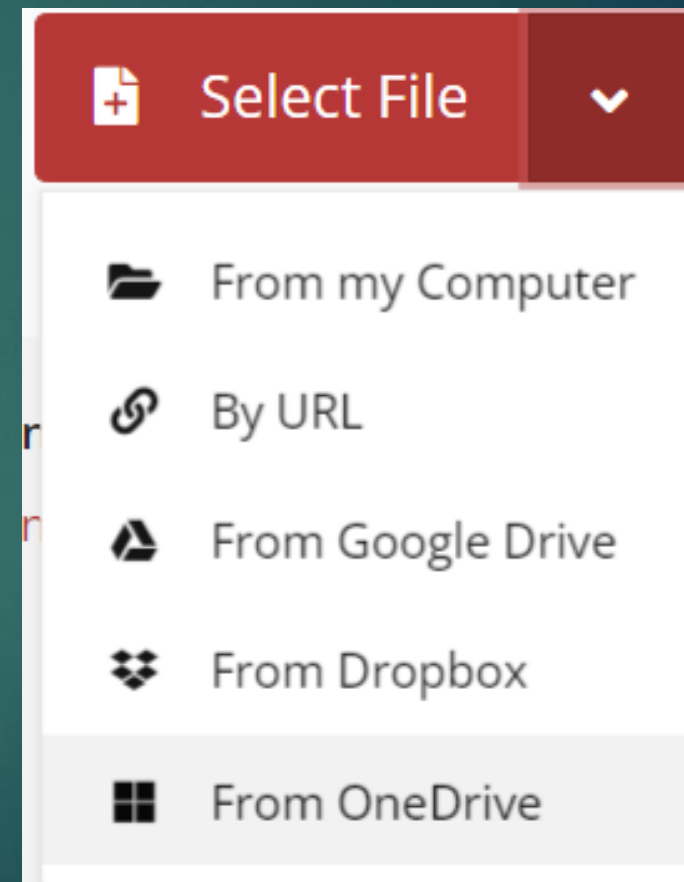
A	B	C	D	E	F	G	H
1 №	ВУЗ	Испытание	Место	Время	Статус	Оценка	Замечание
2 1	РАНХиГС	Бизнес-информатика	ВУЗ	12.07.2024 17:00:00	Экзамен был	74,0	
3 2	ВШЭ	Высшая математика	Дом	19.07.2024 11:00:00	Экзамен был	13,0	
4 3	ГУАП	Информационные науки	Дом	19.07.2024 17:00:00	Экзамен был	74,0	
5 4	ЛЭТИ	Информатика и вычислительная техника	Дом	23.07.2024 11:00:00	Экзамен был	50,0	
6 5	ЛЭТИ	Прикладная математика и информатика	Дом	23.07.2024 11:00:00	Экзамен был	48,0	
7 6	СПбГУ	Технологии искусственного интеллекта и Big Data	Дом	26.07.2024 10:00:00	Экзамен был	49,0	
8 7	ВШЭ	Собеседование по программированию	Дом		Экзамена нет	0,0	
9 8	СПбГУ	Распределенные вычислительные технологии	Дом	26.07.2024 17:00:00	Экзамен был	40,0	
10 9	СПбГУ	Разработка и математическое обеспечение интеллектуальных информационных систем	Дом	27.07.2024 11:00:00	Экзамен был	53,0	
11 10	СПбГУ	Информационная безопасность	Дом	27.07.2024 14:00:00	Экзамен был	20,0	
12 11	СПбГУТ	Автоматизация технологических процессов и производств	ВУЗ		Экзамена нет	0,0	
13 12	ИТМО	Проектирование и разработка систем больших данных	Дом	29.07.2024 11:00:00	Экзамен был	1,0	
14 13	СПбГУ	Программная инженерия	Дом	30.07.2024 10:00:00	Экзамен был	35,0	
15 14	СПбГУТ	Информатика и вычислительная техника	ВУЗ		Экзамена нет	0,0	
16 15	СПбГУТ	Программная инженерия	ВУЗ	01.08.2024 11:00:00	Экзамен был	38,0	38 из 80. С учётом 10 баллов за индивидуальные достижения получается 48 баллов из 100 возможных
17 16	ИТМО	Веб-технологии	Дом	02.08.2024 11:00:00	Экзамен был	50,0	



# Анализ существующих реализаций (CloudConvert)

9/16

- ▶ В рамках данного сервиса требуется следующая последовательность действий:
  - 1) С помощью функции выбора файла указать опцию необходимого облака для загрузки нужного файла с помощью выполнения процедуры авторизации
  - 2) Загрузить файл, выбрать выходной формат для каждого файла
  - 3) Начать процесс преобразования
  - 4) С завершением процесса можно выполнить загрузку на ПК или в облако





# Анализ существующих реализация

11/16

- ▶ Среди существующих реализаций можно выделить следующие проблемы:
  - Ограниченность рамками текущего облачного хранилища
  - Некоторый результат генерации HTML-таблицы из некоторых форматов первоначального файла

# Требования к сервису

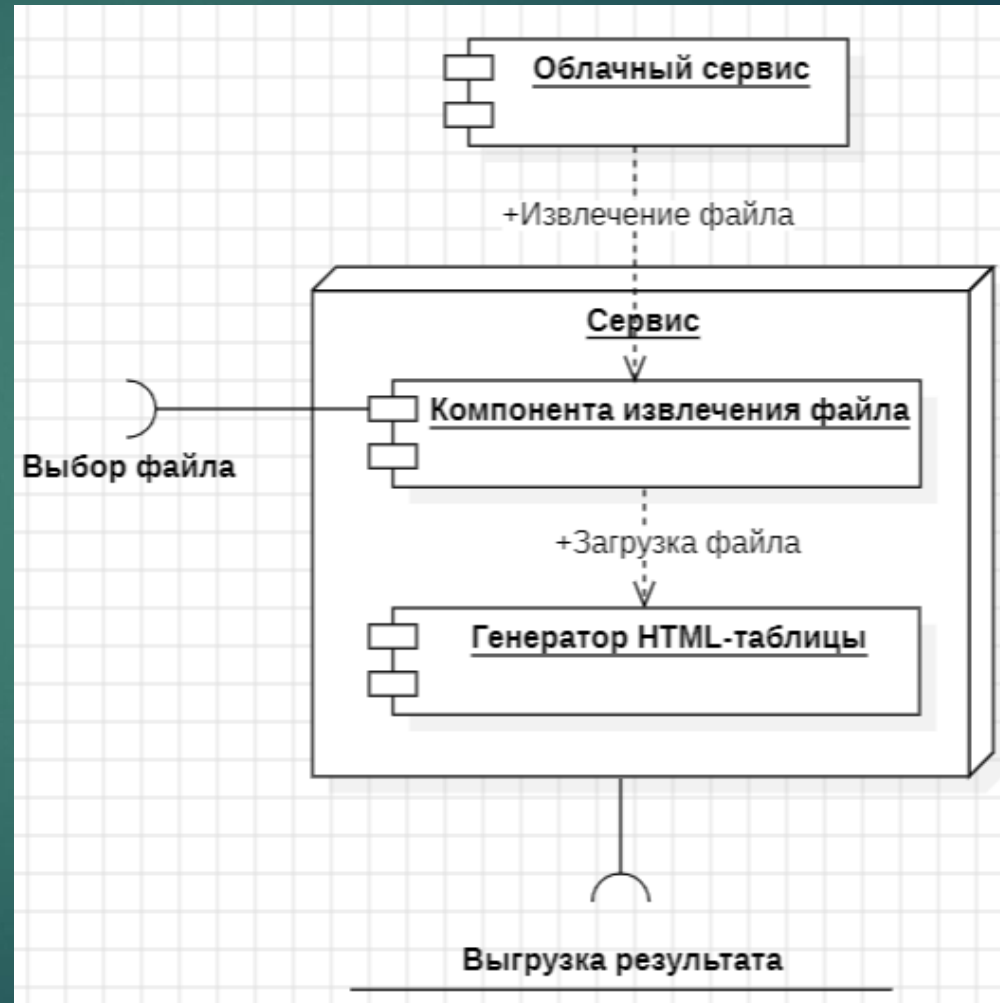
12/16

- ▶ Предлагаемый к реализации сервис должен включать следующие требования:
  - 1) Сервис должен поддерживать возможность извлечение файла из любого облачного хранилища
  - 2) Сервис должен использовать существующий инструмент генерации HTML-таблицы из файла формата .xlsx и .odb
  - 3) В процессе преобразования не должно быть использования архивирования
  - 4) Должна быть предусмотрена возможность сохранить результат преобразования как в облачном хранилище, так и в локальном
- ▶ Среди нефункциональных требований можно привести следующее:
  - 1) Итоговый результат должен быть представлен в виде готового HTML-документа

# Архитектура сервиса

13/16

- ▶ Предлагаемая архитектура сервиса состоит из двух компонентов:
  - Механизма, отвечающего за извлечение необходимой электронной таблицы из облачного хранилища
  - Генератора HTML-таблицы из загруженной электронной таблицы



# Реализация сервиса

14/16

- ▶ К сожалению, в данный момент у сервиса отсутствует реализация в какой-либо форме
- ▶ Но в планах имеется следующий перечень инструментов:
  - В качестве используемого языка программирования стоит выбор между Python и C, так как с ними легко работать, а также имеются необходимые среды разработки
  - Среды разработки PyCharm и VisualCode для работы с языком Python и C соответственно
  - Рассматриваются для работы Rclone и WebDAV, так как оба инструмента предоставляют возможность взаимодействия с локальным и облачным хранилищами
  - LibreOffice и MS Office для тестирования корректности работы сервиса с электронными таблицами
- ▶ При реализации и дальнейших доработках в архитектуру сервиса могут быть внесены изменения



# Тестирование и проба сервиса

15/16

- ▶ По причине отсутствия какой-либо реализации сервиса, каких-либо проверок работоспособности не было
- ▶ При дальнейшей реализации сервиса будет проводиться планирование проверки работоспособности основного функционала

# Заключение

16/16

- ▶ В текущий момент времени выполнена лишь проработка предлагаемой для реализации идеи, которая будет дополняться в дальнейшем
- ▶ Пока можно отметить, что выделены следующие задачи:
  - Проведён анализ функциональности существующих сервисов и возможностей доступных технологий
  - Выработаны требования к сервису
  - Спроектирована архитектура сервиса