

Расширение Chrome «Table OCR»

1. Общая идея

Расширение позволяет:

- выбрать область на текущей вкладке браузера и сразу получить **реальный скрин выбранной области**;
 - либо загрузить **один файл** (изображение или PDF);
 - показать миниатюру выбранной области / файла;
 - отправить изображение или файл на сервер для **OCR таблиц**.
-

2. Платформа и ограничения

- **Manifest:** Chrome Manifest V3
 - **Поддержка браузеров:** только Google Chrome
 - **Минимальная версия Chrome:** 109
-

3. Popur (всплывающее окно расширения)

- Расположение: стандартное popur рядом с иконкой расширения.
 - Размер окна: **фиксированный**, примерно **360x520 px**.
 - Элементы интерфейса:
 1. Заголовок: **Table OCR**
 2. Большая кнопка **Start**
 3. Область Drag&Drop с пунктирной рамкой
 - принимает **только изображения и PDF**
 - по клику открывает системный диалог выбора файла
 4. Область предпросмотра (миниатюра)
 - Состояния popur:
 - idle — ничего не выбрано
 - selecting — пользователь выбирает область экрана
 - processing — данные отправляются на сервер
 - Поведение:
 - кнопка **Start** активна только в состоянии idle
 - после закрытия popur **всё сбрасывается**, данные не сохраняются
 - в popur **только показ миниатюры**, без кнопок Copy / Download / Recognize
-

4. Drag & Drop / загрузка файлов

- Поддерживаемые форматы изображений:

- png, jpg, jpeg, webp, bmp, tiff
 - PDF:
 - только **1 файл**
 - многостраничность не анализируется
 - в preview показывается **иконка PDF**, не изображение
 - Ограничения:
 - размер файла: **до 10 МБ**
 - размер изображения: **не выше HD**
 - Поведение:
 - при загрузке нового файла предыдущий **заменяется без подтверждения**
 - файл хранится **только в памяти**, пока открыт в окне
-

5. Выбор области экрана

- Область выбора:
 - **только текущая вкладка**
 - через **DOM-оверлей**
- Поведение при нажатии **Start**:
 - страница затемняется
 - скролл страницы **блокируется**
 - курсор меняется на крестик
- Механика выбора:
 - выбор производится зажатой ЛКМ
 - формируется прямоугольник
 - выбранная область внутри прямоугольника освещается
 - выбор заканчивается при отпускании ЛКМ
- После завершения выбора:
 - сразу делается **реальный скрин области**
 - используется:
 - chrome.tabs.captureVisibleTab
 - crop изображения в canvas
 - координаты учитывают:
 - devicePixelRatio
 - масштаб браузера
 - прокрутку страницы

- в результатах отображается **миниатюра результата**
 - Управление:
 - Esc — отмена выбора
 - ресайз / перемещение области после выбора — **не требуется**
 - Повторный запуск Start:
 - предыдущая миниатюра **перезатирается**
-

6. OCR / Table OCR

- OCR выполняется **на сервере / через API**
 - Входные данные:
 - изображение (из выделенной области или загруженное)
 - либо PDF-файл
 - Результат OCR:
 - **CSV**
 - **Excel**
 - Тип распознавания:
 - **только таблицы**
 - с сохранением строк и колонок
 - Языки:
 - задаются списком (на стороне API)
 - Постобработка:
 - **не требуется**
 - без автосклейки ячеек
 - без очистки мусора
 - без автодетекции границ таблицы
-

7. Permissions и безопасность

- Разрешения:
 - activeTab
- Поддержка страниц:
 - chrome:///* — **да**
 - встроенный PDF Viewer Chrome — **да**
- Зависимости:
 - все библиотеки (pdf.js и др.) **упакованы локально**

- без CDN
 - CSP:
 - строгая, совместимая с MV3
-

8. Поведение при закрытии / повторном запуске

- Если пользователь нажал **Start** и закрыл попап:
 - оверлей немедленно отменяется
 - При повторном выборе:
 - предыдущий результат полностью сбрасывается
-

9. Структура проекта (MV3)

```
table-ocr-extension/
|
|--- manifest.json
|
|--- popup/
|   |--- popup.html      # UI popup
|   |--- popup.css       # стили popup
|   |--- popup.js        # логика popup, состояния, preview
|
|--- content/
|   |--- overlay.js     # затемнение страницы, выбор области
|   |--- overlay.css    # стили оверлея
|   |--- selection.js   # логика прямоугольника, координаты
|
|--- background/
|   |--- service_worker.js # координация popup ↔ content ↔ capture
|
|--- capture/
|   |--- crop.js        # crop изображения через canvas
|
|--- api/
|   |--- ocr_client.js  # отправка данных на OCR API
```

```
|  
|   |-- assets/  
|   |   |-- icons/  
|   |   |   |-- icon16.png  
|   |   |   |-- icon48.png  
|   |   |   \-- icon128.png  
|   |   \-- pdf_icon.svg  
|  
|  
\-- utils/  
    |-- file_validation.js  
    |-- constants.js  
    \-- helpers.js
```