

MAE™ - Intelligent Document Processing

Автоматическое распознавание счетов и документов с OCR.

Извлекает: *Vendor, Invoice No, Internal ID & VAT ID*

Онлайн версия (без установки)

Попробуй прямо сейчас:

<https://mae-idp.onrender.com>

 Первый запуск может занять ~30 сек (сервер просыпается)

Быстрая установка (локальная версия)

Windows

1. **Скачай** → [Download ZIP](#)
2. **Распакуй** в любую папку
3. **Запусти** `install.bat` (первый раз)
4. **Запусти** `run.bat`
5. **Открой** <http://127.0.0.1:8766>

macOS

1. **Скачай** → [Download ZIP](#)
2. **Распакуй** в любую папку
3. **Открой Терминал** в этой папке
4. **Выполни:**

`chmod +x install.sh run.sh`

`./install.sh`

`./run.sh`

5. **Открой** <http://127.0.0.1:8766>

Требования

Windows

- **Windows 10/11**
- **Python 3.12+** — [скачать](#)
- **Tesseract OCR** — [скачать](#)
- **Poppler** — устанавливается автоматически через `install.bat`

macOS

- **macOS 10.15+ (Catalina или новее)**
- **Python 3.12+** — `brew install python@3.12`
- **Homebrew** — [установка](#)
- **Tesseract & Poppler** — устанавливаются автоматически через `install.sh`

Установка зависимостей

Windows — Tesseract

1. Скачай [tesseract-ocr-w64-setup-5.5.0.exe](#)
2. При установке выбери языки: **German, English**
3. Установи в `C:\Program Files\Tesseract-OCR`

macOS — Tesseract

`brew install tesseract tesseract-lang`

► Запуск

Windows

`run.bat`

Или напрямую:

`venv\Scripts\python app\mae.py`

macOS

`./run.sh`

Или напрямую:

`source venv/bin/activate`

`python3 app/mae.py`

После запуска откройте браузер: <http://127.0.0.1:8766>

Мобильный доступ (iOS / Android)

MAE IDP работает как Progressive Web App (PWA). Вы можете использовать его с мобильных устройств:

Шаг 1: Запустите сервер на компьютере

`./run.sh` # или `run.bat` на Windows

Шаг 2: Узнайте IP-адрес компьютера

```
# macOS  
ifconfig | grep "inet "
```

```
# Windows  
ipconfig
```

Шаг 3: Откройте в мобильном браузере

Откройте на телефоне: `http://ВАШИ_IP:8766`

Например: `http://192.168.1.100:8766`

Шаг 4: Установите как приложение (опционально)

iOS (Safari):

1. Нажмите кнопку "Поделиться" (квадрат со стрелкой)
2. Выберите "На экран «Домой»"

Android (Chrome):

1. Нажмите меню (три точки)
2. Выберите "Добавить на главный экран"

Примечание: Компьютер и телефон должны быть в одной Wi-Fi сети.

Возможности

-  **Drag & Drop** — перетащи файлы в окно
-  **Folder Watch** — автоматический мониторинг папки
-  **Cloud Support** — Google Drive Desktop, OneDrive, Dropbox
-  **Export** — выгрузка в CSV, Markdown, TXT
-  **Dark/Light Theme** — переключение темы

Поддерживаемые форматы

- PDF (первая страница)
- JPG / JPEG
- PNG
- TIFF / TIF

Извлекаемые данные

Поле	Описание	Confidence
Vendor	Автоматически распознаёт 20+ поставщиков (Amazon, DHL, IKEA и др.)	+30
Invoice Number	Номер счёта/накладной	+30
Internal Number	Внутренний входящий номер (из QR или угла)	+30
VAT ID	Идентификатор НДС (28 стран EU)	+10

Порог успеха: ≥90 баллов = **success**, иначе **review**

Известные поставщики

Amazon, DHL, UPS, FedEx, Deutsche Telekom, Vodafone, O2, IKEA, MediaMarkt, Saturn, Conrad, Reichelt, RS Components, Mouser, DigiKey, Farnell, Würth, Hoffmann, Grainger, Mercateo, Staples, Office Depot, Viking, Büroshop24.



Проблемы и решения

"Tesseract not found"

Windows: Убедись что Tesseract установлен в `C:\Program Files\Tesseract-OCR`

macOS: Запусти `brew install tesseract tesseract-lang`

"PDF conversion failed"

Windows: Poppler не установлен. Запусти `install.bat` заново.

macOS: Запусти `brew install poppler`

Окно сразу закрывается (Windows)

Запусти через CMD чтобы увидеть ошибку:

```
cmd /k run.bat
```

Низкий Confidence

- Проверь качество скана (минимум 300 DPI)
- Убедись что документ на немецком или английском
- Проверь что номера не обрезаны

Мобильный доступ не работает

- Убедись что телефон и компьютер в одной Wi-Fi сети
- Проверь что firewall не блокирует порт 8766
- Попробуй отключить VPN