

Описание функциональных характеристик программного обеспечения «ScanLex»

Версия ПО: 1.3.7. Дата документа: 02.09.2026.

1. Назначение

«ScanLex» - программа оптического распознавания документов для рабочего места. Предназначена для перевода сканов и фотографий бумажных документов, а также PDF-файлов в редактируемые форматы с сохранением структуры и внешнего вида документа.

Программа работает полностью автономно: обрабатываемые документы не покидают компьютер пользователя, подключение к сети Интернет и внешние (в том числе облачные) сервисы для распознавания не требуются.

2. Область применения

Оцифровка и подготовка документов в организациях и на рабочих местах: бухгалтерия, делопроизводство, кадровые и юридические службы, архивы. Ввод данных из первичных документов (счета, акты, накладные, справки, формы), перевод бумажных и отсканированных документов в редактируемый вид.

3. Функциональные возможности

3.1. Ввод документов

- Открытие файлов: PDF, а также изображений PNG, JPEG, TIFF, BMP, WEBP.
- Обработка как цифровых PDF (с текстовым слоем), так и сканов и фотографий.
- Захват изображения непосредственно с экрана (снимок выбранной области) - для случаев, когда документ открыт в другой программе или в браузере.
- Пакетная обработка многостраничных PDF, выбор диапазона страниц для распознавания.

3.2. Распознавание

- Распознавание печатного текста на русском и английском языках.
- Гибридный движок: нейросетевой детектор структуры страницы в сочетании с оптическим распознавателем символов (обеспечивает точность на кириллице).
- Автоматическое исправление перекоса скана и поворота страницы (90°/180°/270°) по ориентации текста.
- Увеличение изображений низкого разрешения (фотографии) перед распознаванием.

- Восстановление структуры документа: абзацы, порядок чтения, таблицы (в том числе без видимых линий), колонтитулы, заголовки и списки.
- Выделение графических элементов (печати, QR-коды, иллюстрации) в виде изображений.

3.3. Экспорт результата

- Microsoft Word (DOCX);
- OpenDocument Text (ODT, соответствует ГОСТ Р ИСО/МЭК 26300-2010) - для российских и свободных офисных пакетов (МойОфис, LibreOffice, Р7-Офис);
- Microsoft Excel (XLSX) - для распознанных таблиц;
- обычный текст (TXT);
- PDF с текстовым слоем (документ выглядит как оригинал, но текст доступен для поиска и копирования).

3.4. Редактирование поверх скана

- Встроенный редактор правки текста прямо на изображении документа: место исходной строки заливается цветом бумаги, поверх печатается новый текст; печати, подписи, линии и фон остаются неизменными.
- Точечная правка (изменяется только затронутое слово, остальное - пиксели оригинала), перетекание абзацев, перемещение, изменение размера и наклон блоков.
- Выбор гарнитуры (13 шрифтов), кегля, цвета (в том числе пипеткой по документу), начертания и выравнивания.
- Сохранение отредактированного документа в PDF с сохранением внешнего вида.

3.5. Просмотр

- Встроенный просмотрщик с масштабированием (25 - 400 %), постраничной навигацией.
- Выделение и копирование распознанного текста.

3.6. Управление нагрузкой и надёжность

- Распознавание выполняется в отдельном процессе под надзором основного приложения (изоляция сбоев, ограничение памяти и загрузки процессора).
- Режимы нагрузки на компьютер: «Щадящий», «Обычный», «Максимальный» - с автоматическим подбором под характеристики машины.
- Устойчивость к сбоям: постраничные контрольные точки, автоматическое возобновление, пропуск проблемной страницы.

4. Условия эксплуатации (системные требования)

- Процессор: x86-64 (AMD64), работает и на процессорах без AVX2.

- Оперативная память: не менее 2 ГБ (рекомендуется 4 ГБ и более).
- Свободное место на диске: не менее 2 ГБ.
- Наличие графического окружения (рабочего стола).
- Операционные системы:
- Astra Linux Special Edition 1.8.1 и 1.8.2 (в том числе с включённой замкнутой программной средой);
- ALT Linux (Рабочая станция 11);
- Windows 10 / 11 (64-разрядные).

5. Особенности реализации

- Собственная разработка правообладателя; алгоритмическое ядро поставляется в виде скомпилированного нативного кода.
- Дистрибутивы для отечественных ОС самодостаточны: не требуют доступа в Интернет и подключённых репозиториях (все компоненты, включая языковые модели, входят в состав пакета).
- Дистрибутив для Astra Linux подписан электронной подписью по ГОСТ Р 34.10-2012 и совместим с замкнутой программной средой.

6. Сетевое взаимодействие и автономность

Основная функция - распознавание документов, выполняется полностью автономно: обрабатываемые документы и их содержимое по сети не передаются и никогда не покидают компьютер пользователя. Программа запускается и работает без доступа в Интернет (пробный период включается автономно, активированная лицензия проверяется по сохранённому подписанному токenu).

Сетевые обращения программа совершает только к серверу правообладателя <https://scanlex.ru> (Российская Федерация); обращения к зарубежным серверам, внешним (облачным) сервисам и сторонним API отсутствуют. Перечень обращений:

- Лицензирование (проверка ключа и активация, включение пробного периода, периодическая перепроверка на предмет отзыва/продления). Передаётся: аппаратный отпечаток компьютера, лицензионный ключ. При отсутствии сети программа продолжает работать по ранее сохранённому токenu.
- Проверка обновлений программы и движка распознавания: передаётся текущая версия; в ответ - подписанный электронной подписью манифест. Автоматическую проверку можно отключить в настройках.
- Загрузка обновления - по инициативе пользователя и с его согласия.
- Статистика запусков (для правообладателя): передаётся отпечаток компьютера, состояние лицензии, версия и наименование ОС. Имя компьютера, IP-адрес и содержимое документов не передаются. Обращение необязательное: при

отсутствии сети молча игнорируется и на работу программы не влияет.
Скрытых или обязательных обращений к зарубежной инфраструктуре нет.

7. Ограничения

- Программа предназначена для рабочего места с графическим интерфейсом; работа в режиме сервера без графического окружения не предусмотрена.
- Для распознавания не используются и не требуются внешние (облачные) сервисы.