Readiris™ Pro 12

Руководство пользователя

Содержание

Авторские права 1

Раздел 1 Введение в Readiris 3

Сэкономьте время — ручной ввод текста больше не нужен 3

Семейство Readiris 6

Раздел 2 Установка Readiris 11

Системные требования 11

Установка программного обеспечения 11

Удаление программного обеспечения 13

Регистрация программного обеспечения 13

Поддержка продуктов 14

Раздел 3 Начало работы 15

Запуск Readiris 15

Работа с мастером OCR 15

Пользовательский интерфейс 17

Изменение языка пользовательского интерфейса 20

Раздел 4 Readiris SmartTasks 21

Использование Readiris SmartTasks 21

Раздел 5 Сканирование документов 27

Выбор параметров 27

Открытие файлов изображений 28

Сканирование бумажных документов..... 30

Раздел 6 Настройка сканированных документов..... 37

Раздел 7 Сохранение документов в качестве файлов изображений 43

Раздел 8 Формирование окон для документов..... 45

Автоматическое формирование окон для документов 45

Формирование окон для документов вручную..... 47

Использование шаблонов формирования окон 52

Раздел 9 Распознавание документов..... 55

Введение 55

Выбор языка документа 56

Задание характеристик документа..... 59

Использование интерактивного обучения 61

Использование каталогов шрифтов..... 63

Раздел 10 Форматирование и сохранение документов ... 67

Форматирование документов..... 67

Форматирование текстовых документов 69

Форматирование табличных документов 73

Создание PDF-документов..... 77

Выбор параметров PDF 78

Сжатие iHQC для PDF-документов 80

Перепрофилирование PDF-документов 81

Создание XPS-документов 82

Выбор параметров XPS 84

Сжатие iHQC для XPS-документов 84

Выбор параметров графики 86

Раздел 11 Сохранение и загрузка параметров 89

Раздел 12 Распознавание многостраничных документов 91

Открытие и распознавание нескольких файлов изображений 91

Сканирование и распознавание многостраничных документов 93

Редактирование многостраничных документов 94

Раздел 13 Распознавание текста, написанного печатными буквами 97

Правила письма печатными буквами 99

Раздел 14 Распознавание штрихкодов 101

Указатель 103

Авторские права

ReadirisPro12-dgi-090211-02

©I.R.I.S., 1987-2009. Все права защищены.

Авторские права на программное обеспечение Readiris, интерактивную справочную систему и данную публикацию принадлежат компании I.R.I.S.

Информация, представленная в этом документе, является собственностью компании I.R.I.S. Содержимое документа может меняться без уведомления и не подразумевает обязательств со стороны компании I.R.I.S. Программное обеспечение, описанное в этом документе, предоставляется в соответствии с лицензионным соглашением, которое регулирует использование продукта. Программное обеспечение может использоваться и копироваться только в соответствии с условиями соглашения. Никакая часть данной публикации не может воспроизводиться, передаваться, сохраняться в файловой службе и переводиться на другой язык без письменного разрешения компании I.R.I.S.

В данном руководстве пользователя используются вымышленные имена для демонстрационных целей. Любые совпадения с реальными людьми, компаниями и организациями случайны.

Товарные знаки

Логотип Readiris и Readiris — товарные знаки Image Recognition Integrated Systems S.A.
OCR, ICR и технология распознавания штрихкодов принадлежат компании I.R.I.S.
AutoFormat и лингвистическая технология принадлежат компании I.R.I.S.
BCR и технология анализа полей принадлежат компании I.R.I.S.
Технология сжатия iHQC принадлежит компании I.R.I.S.
Средство синтаксического разбора XML разработано компанией Apache.
Данный продукт включает программное обеспечение, разработанное Apache Software Foundation.

Все прочие продукты, упомянутые в этом руководстве пользователя, являются товарными знаками или зарегистрированными товарными знаками соответствующих владельцев.

Раздел 1
Введение в Readiris

Сэкономьте время — ручной ввод текста больше не нужен

Поздравляем вас с покупкой Readiris. Это программный пакет, несомненно, станет серьезным подспорьем при считывании текста, таблиц, графики, штрихкодов и письма печатными буквами.

Как бы эффективны не были компьютеры, сначала в них нужно ввести информацию. Если вы когда-нибудь набирали 15-страничный отчет или большую таблицу с цифрами, вы знаете, каким длительным и трудоемким может быть этот процесс. Воспользуйтесь нашим передовым пакетом средств OCR для автоматического преобразования бумажных документов или сканированных изображений в редактируемые документы с возможностью поиска, которые можно архивировать и предоставлять для совместного доступа. Доступно два режима распознавания: один обеспечивает максимальную скорость, другой — оптимальную точность OCR.

Отсканируйте документ, напечатанный на принтере или печатной машинке, обозначьте области для распознавания с помощью Readiris — или позвольте системе сделать это за вас, — выполните распознавание текста и экспортируйте документы в текстовый процессор. Документы, включающие множество страниц, обрабатываются с начала до конца в рамках одной операции. Несколько щелчков мышью заменяют

многие часы работы, когда Readiris преобразует бумажные документы в редактируемые компьютерные файлы: в 40 быстрее, чем при использовании ручного ввода.

Мастер помогает пользователям ввести параметры, необходимые для работы Readiris, обеспечивая быстрое и простое получение результатов. Кроме того, можно воспользоваться задачами SmartTasks для еще большего ускорения процесса. Результаты чтения можно отправить непосредственно в текстовый процессор или редактор электронных таблиц, их можно архивировать в PDF- или XPS-файлы и др. Для распознавания факсов и преобразования PDF-документов перетащите файлы изображений из проводника Windows в окно приложения Readiris. Кроме того, можно мгновенно отправить изображение в Readiris с помощью контекстного меню.

Readiris распознает табличные данные и восстанавливает их в виде листов в редакторе электронных таблиц или в виде табличных объектов в текстовом процессоре. При этом цифровые данные будут немедленно готовы к дальнейшей обработке.

Readiris основывается на самых передовых методах распознавания. Распознавание текста, независимое от шрифтов, дополняется технологиями самообучения. Система может изучать новые символы и слова посредством контекстного и лингвистического анализа. Это значит, что точность OCR системы распознавания будет значительно улучшена по мере использования продукта.

Приложение Readiris способно распознавать тексты на 128 языках. Поддерживаются все американские и европейские языки, в том числе центрально-европейские, кириллические и балтийские, а также греческий и турецкий. Кроме того, Readiris считывает документы на арабском, фарси и иврите, также на четырех азиатских языках — японском, упрощенном китайском, традиционном китайском и корейском

(дополнительная функция). Readiris справляется даже со смешанными алфавитами: приложение обнаруживает ”западные” слова в кириллических документах, документах на греческом, арабском, иврите и азиатских языках, которые часто представляют собой непереводимые называния компаний и продуктов, написанные латиницей.

Readiris использует лингвистические технологии *во время* этапа распознавания, а не после него. В результате приложение Readiris распознает все типы документов с максимальной точностью, в том числе документы низкого качества, факсы и распечатки с матричного принтера. Оно отлично справляется с плохо отсканированными и скопированными документами, содержащими слишком темные или слишком светлые шрифты. Соединенные символы разделяются, а символы с прерывистыми контурами, например напечатанные на матричном принтере, восстанавливаются.

Кроме того, Readiris предлагает (дополнительную) функцию проверки пользователями. При активации функция проверки пользователем ("Интерактивное обучение") не только отмечает символы, в которых система распознавания не уверена, но и повышает ее точность. Все подтвержденные решения временно запоминаются во время распознавания, что повышает скорость системы и уверенность распознавания, делая ее более интеллектуальной по мере использования продукта. Это мощное средство обучение также позволяет "научить" Readiris работе со специальными знаками, такими как математические символы и графические метки, а также с искаженными шрифтами.

Для дальнейшего улучшения продуктивности Readiris предлагает форматирование текста в дополнение к их распознаванию. Поддерживаются различные уровни распознавания. При использовании ”автоформатирования” Readiris создает точную копию отсканированного документа: форматирование слов, параграфов и страниц сохраняется.

Используются те же типы шрифта. Размеры и стили, применяемые в исходном документе, сохраняются при распознавании. Расположение колонок, текстовых блоков и графики соответствует оригинальным документам. Более того, Readiris может включать фоновый фотоснимок сканированного изображения в распознанный документ. Readiris поддерживает цветное сканирование и сканирование в оттенках серого и позволяет считывать любые графические объекты — линейные рисунки, черно-белые фотографии и цветные иллюстрации. Если документ содержит таблицы, Readiris преобразует их в реальные ячейки и восстанавливает границы ячеек исходных таблиц.

Другими словами, Readiris позволяет создать точные копии документов в виде редактируемых компактных текстовых файлов, а не сканированных изображений.

Штрихкоды на сканированных страницах также могут быть считаны, как и текст, написанный ”печатными буквами” с выраженными пробелами.

ПО Readiris поддерживает широкий спектр Twain-совместимых планшетных сканеров и сканеров с подачей листов, универсальных устройств, MFP (многофункциональных устройств) и цифровых фотокамер. Если планшетный сканер не оборудован автоподатчиком (ADF), самым эффективным способом сканирования многостраничных документов является интервальное сканирование.

СЕМЕЙСТВО READIRIS

В таблице ниже представлен обзор расширенных версий:

Readiris Home 12 Ограниченные возможности 25 языков распознавания Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Создает файлы в формате PDF - изображение-текст, DOCX, ODT, WordML, SpreadsheetML, RTF, HTM, XML, TXT и TIFF

Readiris Pro 12 Основные возможности 128 языков распознавания Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Создает PDF-файлы 4-х типов, PDF-iHQC (уровень I), XPS-файлы 4-х типов, XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др.	**Readiris Corporate 12** Основные возможности 128 языков распознавания Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Создает PDF-файлы 4-х типов, PDF-iHQC (уровни I-III), PDF/A, XPS-файлы 4-х типов, XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др. Распознавание большого количества документов Автоматическая обработка Индексация документов Распознавание визитных карточек

Readiris Pro 12 Asian	**Readiris Corporate 12 Asian**
Основные возможности	Основные возможности
128 языков распознавания	128 языков распознавания
Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX	Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Создает PDF-файлы 4-х типов, PDF-iHQC (уровень I), XPS-файлы 4-х типов, XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др.	Создает PDF-файлы 4-х типов, PDF-iHQC (уровни I-III), PDF/A, XPS-файлы 4-х типов, XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др.
Распознавание текста на японском	Распознавание текста на японском
Распознавание текста на традиционном и упрощенном китайском	Распознавание текста на традиционном и упрощенном китайском
Распознавание текста на корейском языке	Распознавание текста на корейском языке
	Распознавание большого количества документов
	Автоматическая обработка
	Индексация документов
	Распознавание визитных карточек

Readiris Pro 12 Middle-East*	**Readiris Corporate 12 Middle-East***
Основные возможности	Основные возможности
128 языков распознавания	128 языков распознавания
Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX	Поддерживает изображения в формате PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Создает PDF-файлы 4-х типов, PDF-iHQC (уровень I), XPS-файлы 4-х типов, XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др.	Создает PDF-файлы 4-х типов, PDF-iHQC (уровни I-III), PDF/A, XPS-файлы 4-х типов,XPS-iHQC (уровень I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF и др.
Распознавание текста на фарси и арабском	Распознавание текста на фарси и арабском
распознавание иврита	распознавание иврита
	Распознавание большого количества документов
	Автоматическая обработка
	Индексация документов
	Распознавание визитных карточек

(**) Версия для Mac недоступна.

РАЗДЕЛ 2
УСТАНОВКА READIRIS

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Ниже приводится **минимальная конфигурация системы**, необходимая для использования Readiris:

- Персональный компьютер с процессором Intel 486 или совместимый. Рекомендуется компьютер с процессором Pentium.
- ОЗУ 256 Мбайт
- 120 Мбайт свободного места на диске.

  (105 Мбайт места на диске будет достаточно, если отказаться от установки примеров файлов)

- ОС Windows Vista, XP или 2000.

Обратите внимание, что драйверы некоторых сканеров могут не работать с последними версиями Windows. См. сведения о поддерживаемых платформах в документации по сканеру.

УСТАНОВКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Установка Readiris:

- Войдите в ОС Windows как администратор или убедитесь, что у вас есть необходимые права администратора.

- Подключите сканер к компьютеру и установите соответствующее ПО.
  Проверьте работоспособность сканера. При возникновении трудностей обратитесь к производителю сканера.

- Вставьте компакт-диск Readiris в дисковод для компакт-дисков и следуйте инструкциям по установке ПО.

- Выберите пункт **Readiris**, чтобы начать установку (будут предложены дополнительные продукты: Copernic Desktop Search Home Edition и Cardiris 4 LE).

- Выберите язык установки и нажмите кнопку **OK**.

- Примите условия лицензионного соглашения.

- Доступна **полная** и **пользовательская** установка. Выберите нужные параметры и нажимайте кнопку **Далее** для перехода к следующим экранам.
  Все словари, примеры изображений, интерактивная справка и электронное руководство пользователя будут установлены по умолчанию.

- Нажмите кнопку **Готово**, чтобы завершить установку.
  Подменю **Приложения I.R.I.S. - Readiris** в меню **программы** Windows создается программой установки автоматически.
  Кроме того, программа установки создает ярлык на приложение Readiris на рабочем столе Windows.

- Повторите процесс установки, чтобы установить дополнительные программы с компакт-диска.

УДАЛЕНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Существует только один корректный способ удаления Readiris: с помощью мастера установки и удаления программ Windows. *Не* рекомендуется удалять приложение Readiris или его программные модули вручную путем удаления файлов программы.

Чтобы удалить Readiris:

- Закройте приложение.
- В меню **Пуск** выберите пункт **Панель управления**.
- В окне **Программы** выберите **Удалить программу**.
- Выберите пункт **Readiris** и нажмите кнопку **Удалить**.
- Следуйте инструкциям на экране.

РЕГИСТРАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Обязательно зарегистрируйте лицензию Readiris. Это позволит вам:

- быть в курсе будущих разработок и новостей, связанных с другими продуктами I.R.I.S,
- получить право на бесплатную поддержку продуктов,
- получить право на специальные предложения I.R.I.S.

Для регистрации:

Воспользуйтесь **мастером регистрации** в меню **регистрация**. Следуйте инструкциям на экране.

Поддержка продуктов

После регистрации продукта вы получите право на поддержку от компании I.R.I.S. по всем основным возможностям ПО. Как связаться с I.R.I.S.:

Европа:
support@irislink.com
Тел.:+32 10 45 13 64

США:
support@irisusa.com
Тел.:+1 800 447 4744

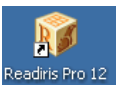
Азиатско-Тихоокеанский регион
support@irislink.com
Тел.: +852 22646133

РАЗДЕЛ 3
НАЧАЛО РАБОТЫ

ЗАПУСК READIRIS

Чтобы запустить Readiris:

- Запустите ПО **Readiris** из меню **Пуск** ОС Windows или дважды щелкните ярлык на рабочем столе.



- Щелкните в любой точке экрана запуска, чтобы запустить Readiris.

  Мастер OCR откроется автоматически.

РАБОТА С МАСТЕРОМ OCR

Мастер OCR позволяет задать все параметры, необходимые для эффективной работы с Readiris.

При запуске Readiris щелкните в любой точке экрана запуска, чтобы запустить мастер OCR.

Этап 1

Выберите **источник изображения**.
Можно оцифровать изображения с помощью **сканера** или открыть **файлы изображений**.
Выберите параметры **поворота** и **выравнивания**, которые необходимо использовать.

> Дополнительные сведения см. в разделе **Выбор параметров**.
>
> Чтобы привыкнуть к Readiris, воспользуйтесь примерами изображений, представленных в программном обеспечении.
>
> Их можно найти на компакт-диске Readiris и в подпапке **Samples** установочной папки Readiris.

Нажмите кнопку **Далее** для перехода к следующему этапу.

Этап 2

Если выбран сканер, нажмите кнопку **Изменить**, чтобы изменить параметры сканера.

> Дополнительные сведения см. в разделе **Сканирование бумажных документов**.

Для сохранения параметров нажмите кнопку **OK**.
Нажмите кнопку **Далее** для перехода к следующему этапу.

Этап 3

Нажмите кнопку **Изменить**, чтобы изменить язык документа.
Язык документа по умолчанию — американский английский.
Выберите нужный язык или сочетание языков, а также дополнительные языки в списке и нажмите кнопку **OK**.
Задайте нужный баланс **скорости и точности** с помощью бегунка.

> Дополнительные сведения см. в разделе **Выбор языка документа**.

Для сохранения параметров нажмите кнопку **OK**.
Нажмите кнопку **Далее** для перехода к следующему этапу.

Этап 4

Нажмите кнопку **Изменить**, чтобы изменить формат вывода или конечное приложение. Конечное приложение по умолчанию — Microsoft Word.
Выберите нужный формат вывода или приложение в раскрывающемся списке **Отправить** или **Внешний файл**.
Задайте нужные параметры на различных вкладках.
Параметры, недоступные для выбранного формата или приложения, будут затемнены.

> Дополнительные сведения см. в разделе Форматирование и сохранение документов.

Для сохранения параметров нажмите кнопку OK**.**
Нажмите кнопку Далее для перехода к следующему этапу.

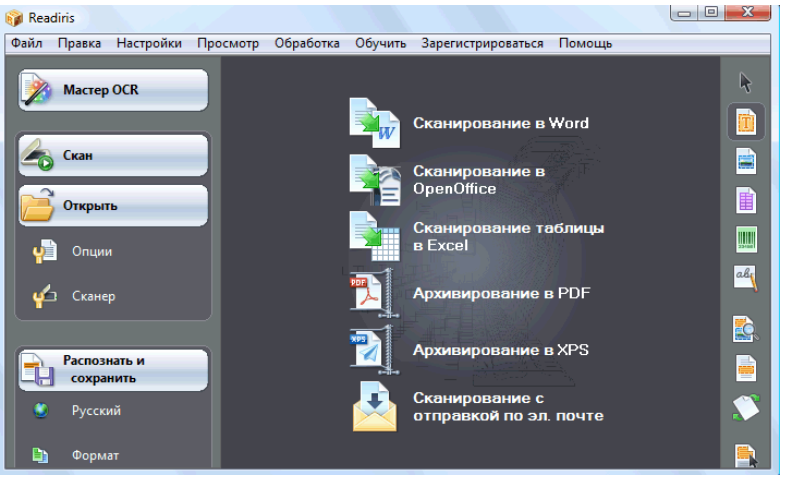
Этап 5

Нажмите кнопку **Вперед**, чтобы открыть или отсканировать и распознать документ.

ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС

Чтобы открыть пользовательский интерфейс Readiris, щелкните мышью в любой точке экрана запуска Readiris и нажмите кнопку Отмена, когда запустится мастер OCR.

Откроется пустой интерфейс Readiris.



Интерфейс Readiris состоит из следующих компонентов:

- **SmartTasks** (в середине)

  Задачи SmartTasks — это готовые команды, которые позволяют применять часто используемые функции Readiris одним нажатием кнопки.

  Задачи SmartTasks используются для сканирования, распознавания и отправки документов в конечное приложение или их преобразования в выбранный формат вывода.

  Для задач SmartTasks применяются параметры по умолчанию, но они могут быть настроены в соответствии с более конкретными потребностями. Для этого щелкните нужную задачу правой кнопкой мыши.

- **Главная панель инструментов** (левая панель)

  Команды главной панели инструментов используются для ручного сканирования и распознавания документов.

Рекомендуемый порядок выполнения этих команд дается в мастере OCR.

- **Панель изображения** (правая панель)

  Кнопки панели изображений используются для редактирования документов в интерфейсе Readiris. Наведите указатель мыши на кнопки, чтобы посмотреть подсказки по ним.

  Когда в приложении Readiris открыт или отсканирован документ, к интерфейсу добавляются три основные зоны:

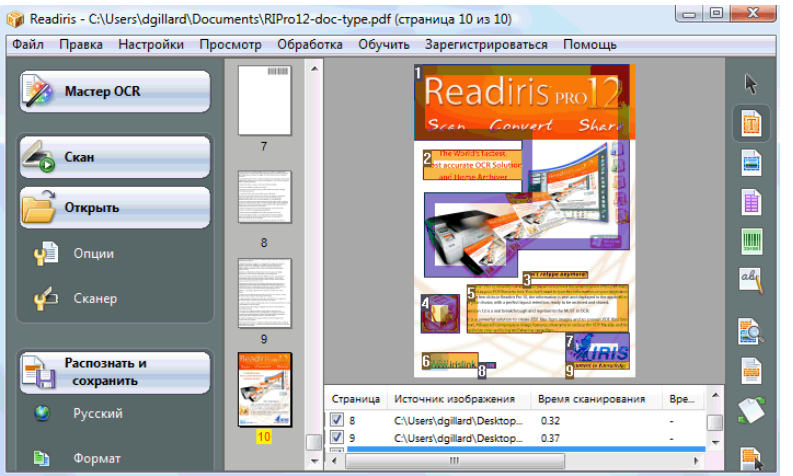
- **Панель страниц** (справа от главной панели)

  На панели страниц отображаются эскизы страниц. Если навести на них указатель мыши, появятся сведения о параметрах.

- **Окно изображения** (в середине)

- **Панель документов** (внизу)

  На панели документов отображаются статические сведения о документах, открытых в Readiris, такие как время сканирования и OCR, разрешение, ширина и высота документов и др.



Изменение языка пользовательского интерфейса

Пользовательский интерфейс Readiris доступен на широком спектре языков.

Изменение языка пользовательского интерфейса:

- В меню **Параметры** выберите **Язык пользовательского интерфейса**.
- В списке **Язык** выберите нужный язык, а затем нажмите кнопку **OK** для подтверждения.

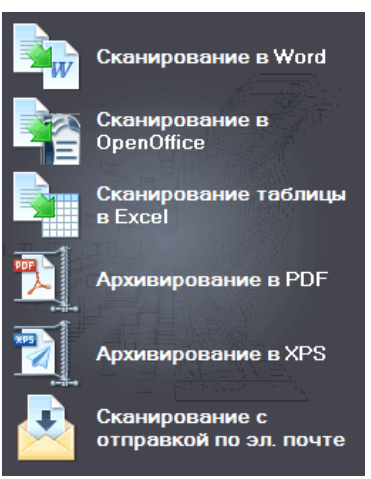
Примечание. Если был выбран неверный язык, нажмите Ctrl+U. Откроется диалоговое окно Язык, в котором можно выбрать другой язык из списка.

РАЗДЕЛ 4
READIRIS SMARTTASKS

ИСПОЛЬЗОВАНИЕ READIRIS SMARTTASKS

При запуске Readiris щелкните в любой точке экрана запуска, а затем нажмите кнопку **Отмена**, когда запустится мастер OCR. Откроются задачи Readiris SmartTasks.

Задачи SmartTasks — это готовые команды, которые позволяют применять часто используемые функции Readiris нажатием кнопки.



Кнопки SmartTask позволяют выполнить следующие действия:

1. Отсканировать и распознать документы, а затем отправить их непосредственно в **Word** для обработки текста.

Microsoft Word является конечным приложением по умолчанию. См. сведения о других доступных приложениях в разделе **Форматирование текстовых документов**.

2. Отсканировать и распознать документы, а затем отправить их непосредственно в **OpenOffice** для обработки текста.

OpenOffice.org Writer — это конечное приложение по умолчанию. См. сведения о других доступных приложениях в разделе **Форматирование текстовых документов**.

3. Отсканировать и распознать таблицы и отправить их непосредственно в **Excel** и другие редакторы электронных таблиц.

Microsoft Excel является конечным приложением по умолчанию. См. дополнительные сведения в разделе **Форматирование табличных документов**.

4. Отсканировать и распознать документы, а затем **заархивировать** их в **PDF**-файлы.

Формат вывода по умолчанию — Adobe Acrobat PDF - изображение-текст. См. сведения о других доступных форматах в разделе **Создание PDF-документов**.

5. Отсканировать и распознать документы, а затем **заархивировать** их в **XPS**-файлы.

Формат вывода по умолчанию — XPS - изображение-текст. См. сведения о других доступных форматах в разделе **Создание XPS-документов**.

6. Отсканировать и распознать документы, а затем отправить их по **электронной почте**.

Документы будут отправлены в формате "PDF - изображение-текст" посредством стандартного почтового приложения. См.

сведения о других доступных форматах в разделе **Форматирование документов**.

При первом запуске Readiris необходимо настроить функцию SmartTasks.

Для настройки задач SmartTasks:

- Щелкните нужную задачу SmartTask правой кнопкой мыши.
- Выберите **Сканер** или **Файлы изображений** в качестве источника изображений.
    - Если выбрать **Сканер**, Readiris начнет сканирование, как только пользователь щелкнет задачу SmartTask. Отсканированные документы будут открыты в интерфейсе, обработаны и сохранены.

      Чтобы задачи SmartTasks работали, сканер должен быть настроен верно.

      Для этого:

      - Нажмите кнопку **Сканер** на главной панели инструментов.
      - Нажмите кнопку **модель сканера** и выберите сканер в списке.

      Если нужного сканера нет в списке, выберите пункт "Twain - другие модели".

      - Нажмите кнопку **Настроить**, если применимо, чтобы выбрать источник Twain.
      - Затем нажмите кнопку **OK** для сохранения параметров.

Дополнительные сведения о параметрах сканера и сканировании бумажных документов см. в разделе **Сканирование бумажных документов**.

- Если выбран вариант **Файлы изображений**, при выборе задачи SmartTask откроется диалоговое окно **Ввод**, в котором можно выбрать файлы изображений для обработки.

  Дополнительные сведения об открытии файлов изображений см. в разделе **Открытие файлов изображений**.

- Нажмите кнопку **Настроить**, чтобы изменить **формат** вывода и его **параметры**.

  Обратите внимание, что доступные форматы вывода и параметры зависят от выбранной задачи SmartTask.

  См. сведения о других доступных форматах и параметрах в разделе **Форматирование документов**.

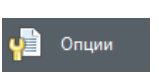
- Если выполняется **распознавание визитных карточек**, выберите **стиль карточки и формат вывода**.

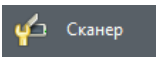
  Дополнительные сведения см. в разделе **Распознавание визитных карточек**.

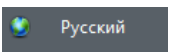
- Если выполняется **Разделение и индексация документов**, выберите **Обработка документов**, чтобы открыть параметры разделения и индексации.

  Дополнительные сведения см. в разделе **Разделение и индексация пакетов документов**.

- После настройки задач SmartTasks воспользуйтесь кнопками на главной панели инструментов, чтобы настроить параметры **языка** и улучшения изображений, а также **параметры сканера, если это необходимо**.







Дополнительные сведения об этих параметрах см. в разделе **Выбор параметров**, **Сканирование бумажных документов** и **Выбор языка документа**.

- И наконец, щелкните задачу, чтобы применить ее SmartTask.

  Readiris автоматически выполнит весь процесс преобразования.

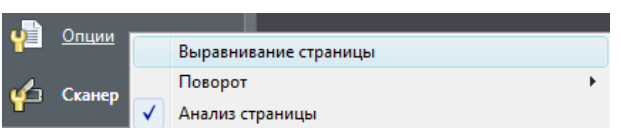
Раздел 5
Сканирование документов

Выбор параметров

Перед сканированием бумажных документов или открытием файлов изображений можно задать параметры улучшения изображений. Активированные параметры будут применены при открытии и сканировании документов.

Действия

- Нажмите кнопку **Параметры** на главной панели инструментов и выберите параметры улучшения изображения.



  - Выберите параметр **Выравнивание страницы** для выравнивания страниц, отсканированных под углом.
  - Наведите указатель мыши на поле **Поворот** определите, должно ли приложение Readiris поворачивать страницы **автоматически**, **на 90° влево**, **90° вправо** или **на 180°**.

    Обратите внимание, что эти два параметра замедляют процесс сканирования. Выбирайте их, только если это необходимо.

- **Анализ страницы** включен по умолчанию.

  Он обеспечивает автоматическое разделение открытых или отсканированных изображений в окнах.

  Инструменты формирования окон можно использовать для изменения результатов анализа страниц или ручного распределения документов по окнам. Дополнительные сведения см. в разделе **Распределение документов по окнам**.

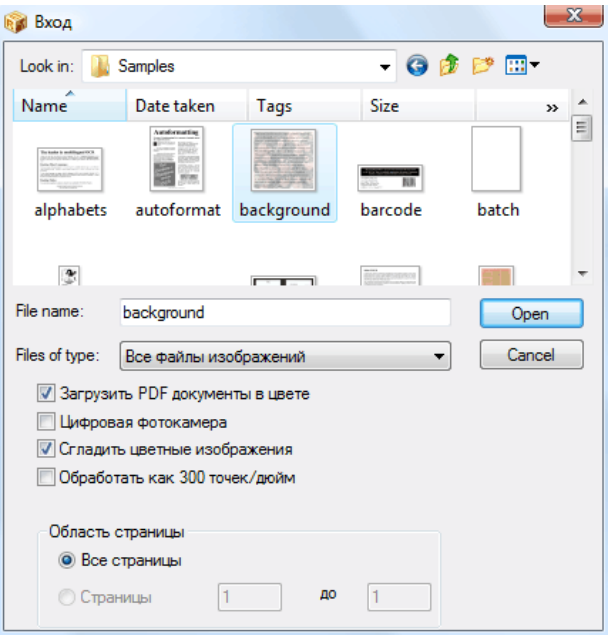
- После задания всех параметров (**Параметры сканера**, **Параметры**) нажмите кнопку **Сканировать**

ОТКРЫТИЕ ФАЙЛОВ ИЗОБРАЖЕНИЙ

Readiris позволяет обрабатывать сканируемые бумажные документы и существующие файлы изображений в различных форматах.

Чтобы существующие файлы изображений:

- Нажмите кнопку **Открыть**, чтобы найти файлы изображений.



Совет. Кроме того, можно **перетащить** файлы изображений в окно изображения Readiris, чтобы открыть их.

Совет. Щелкните правой кнопкой мыши файл изображения, который необходимо открыть, выберите пункт **Открыть с помощью**, а затем **Приложение IOCR**. ПО Readiris откроет и выведет изображение на экран.

Совет. При загрузке многостраничных файлов изображений (изображений в формате TIFF или факсов в формате DCX) и PDF-документов можно задать диапазон страниц (например, если необходим только один раздел документа). Для этого нажмите кнопку **Открыть** на главной панели инструментов. В разделе **Диапазон страниц** выберите поле **Страницы** и введите номера страниц, которые необходимо загрузить. См. также **Открытие и распознавание нескольких файлов изображений**.

Совет. Чтобы ускорить процесс, нажмите кнопку**Открыть** и снимите флажок **Загрузить PDF документы в цвете** при обработке PDF-документов.

- Readiris поддерживает следующие графические форматы:

Графические форматы	**Расширение файла**
Adobe Acrobat PDF	pdf
Факс DCX	dcx
Изображения DjVu	djv, djvu
Изображения JPEG	jpg, jpeg
Изображения JPEG 2000	j2c, j2k, jp2
Portable Network Graphics	png
Изображения TIFF (многостраничные и со сжатием) (без сжатия, со сжатием LZW, PackBits, Group 3, Group 4 и JPEG.	tif, tiff
Точечный рисунок Windows	bmp
Изображения ZSoft Paintbrush	pcx

- Выберите нужный файл изображения и нажмите кнопку **Открыть**.

Примечание. Параметры диалогового окна **Ввод** также применяются к сканированию документов и рассматриваются в разделе **Сканирование бумажных документов**.

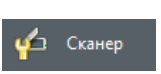
СКАНИРОВАНИЕ БУМАЖНЫХ ДОКУМЕНТОВ

Readiris позволяет обрабатывать сканируемые бумажные документы и существующие файлы изображений в различных форматах.

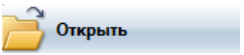
При обработке бумажных документов приложение Readiris запустит сканер, как только пользователь нажмет кнопку **Сканировать**, и отобразит отсканированные документы в интерфейсе.

Чтобы отсканировать документы:

- Нажмите кнопку **Сканер**, чтобы задать параметры сканера.



Обратите внимание, что ряд параметров диалогового окна **Сканер** также доступны в диалоговом окне **Открыть**.



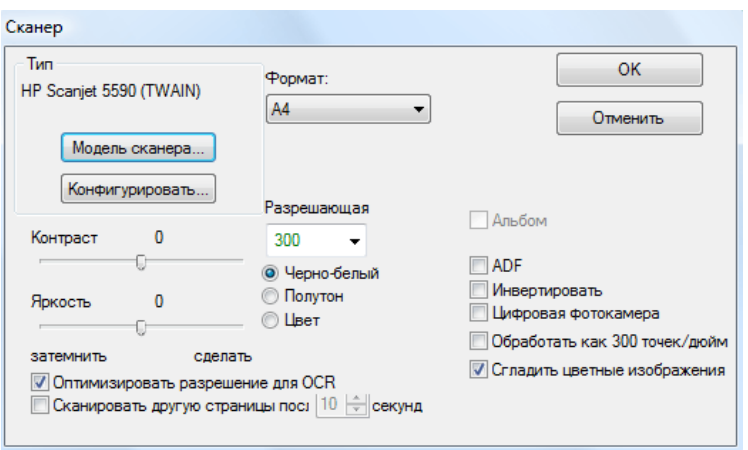
- Выберите нужную модель сканера.

  Если вашей модели нет в списке, выберите **Twain – другие модели** и нажмите кнопку **OK**.

- Readiris поддерживает почти все сканеры с подачей листов, планшетные сканеры, "универсальные" устройства ("MFP" — "многофункциональное периферийное устройство"), цифровые фотокамеры и стандарты сканеров. ПО Readiris соответствует стандарту Twain.

  Обратите внимание, что кнопка **Настроить** будет доступна, только если ее можно использовать в данной ситуации. Нажмите ее, чтобы выбрать источник Twain.

- Настройте параметры сканера.



Формат и разрешение

Readiris поддерживает широкий спектр форматов бумаги и разрешений.

Рекомендуемое разрешение — 300 точек/дюйм.

При распознавании , текста на азиатских языках и текста, напечатанного мелким шрифтом, используйте разрешение 400 точек/дюйм.

Режим цвета

Readiris может сканировать документы и открывать файлы изображений следующих видов: **цветные**, **черно-белые** и **оттенки серого**.

Контраст - яркость

Используйте бегунок, чтобы задать нужные параметры яркости и контраста для получения оптимальных результатов.

Оптимизация разрешения для OCR

Выберите пункт **Оптимизировать разрешение для OCR**, чтобы исправить разрешение изображений, отсканированных со слишком высокой детализацией (более 600 точек/дюйм). Readiris уменьшит разрешение.

> Обратите внимание, что эта функция не может увеличить разрешение изображений, отсканированных со слишком низкой детализацией.

Сканирование многостраничных документов

При сканировании многостраничных документов с помощью сканера с автоподатчиком, выберите параметр **автоподатчик**.

Поместите страницы, которые необходимо отсканировать, в автоподатчик и начните сканирование.

Или воспользуйтесь **интервальным сканированием**, если применяется планшетный сканер**:** выберите параметр **Сканировать другую страницу после** и укажите количество секунд, про прошествии которого приложение Readiris должно отсканировать следующую страницу.

> Дополнительные сведения см. в разделе **Сканирование и распознавание многостраничных документов**.
>
> **Важно**! Параметры, недоступные для выбранного сканера, будут затемнены.

Сканирование изображений с альбомной ориентацией

Выберите параметр **Альбомная** при сканировании изображений с альбомной ориентацией.

Автоэкспозиция

При использовании некоторых сканеров параметр **Автоэкспозиция** выбран по умолчанию. Он обеспечивает автоматическую регулировку параметров яркости и контраста.

Сканирование инвертированных изображений

Выберите параметр **Инвертировать** при сканировании инвертированных черно-белых изображений.

Использование цифровой фотокамеры

Выберите параметр **Цифровая фотокамера**, если в качестве источника сканирования используется цифровая фотокамера. Readiris использует специальные алгоритмы распознавания для обработки изображений, полученных с цифровой фотокамеры.

Советы по использованию цифровой фотокамеры в качестве источника сканирования:

- Проведите калибровку фотокамеры, сняв пустой документ белого цвета.
- Установите максимальное разрешение.
- При съемке на малом расстоянии переведите фотокамеру в режим макро.
- Используйте только оптическое масштабирование, но не цифровое.
- Держите фотокамеру прямо над документом. Избегайте съемки документов под углом.
- Не допускайте колебаний фотокамеры. При необходимости используйте штатив.
- Отключите вспышку при съемке документов, напечатанных на глянцевой бумаге.
- Не открывайте снимков, сжатых фотокамерой.
- Настройте параметры яркости и контраста Readiris (дневной свет, лампа накаливания, неоновая лампа).
- Используйте режимы цвета "цветной" или "оттенки серого"

Обработка документов как изображений с разрешением 300 точек/дюйм

При обработке изображений неверного или неизвестного разрешения выберите пункт **Обработать как 300 точек/дюйм**. Эти изображения будут обработаны так же, как изображения с разрешением 300 точек/дюйм.

Разрешение изображений, полученных с помощью цифровой фотокамеры, почти всегда неизвестно.

Сглаживание цветных изображений

Этот параметр выбирается по умолчанию, так как для успешного распознавания цветных и черно-белых изображений, а также изображений в оттенках серого, полученных на многих сканерах, необходимо сглаживание.

- После задания все параметров (**параметров сканера**, **других параметров**) нажмите кнопку **Сканировать**, чтобы отсканировать документы.

  Примечание. Обратите внимание на наклон строк. Наклон более 0,5° увеличит риск ошибок OCR.

РАЗДЕЛ 6
НАСТРОЙКА СКАНИРОВАННЫХ ДОКУМЕНТОВ

При открытии и сканировании очень светлых или очень темных цветных изображений или изображений в оттенках серого перед распознаванием может потребоваться их настройка. Это позволит получить удовлетворительные результаты OCR.

Для настройки изображений:

- Откройте или отсканируйте цветной документ или документ в оттенках серого.

  Убедитесь, что **параметры сканера** верны.

  Важно! **Параметры сканера** и **Параметры настройки** кажутся одинаковыми, но оба набора применяются на разных этапах процесса сканирования и распознавания.

- В меню **Обработка** выберите **Настроить изображение**.

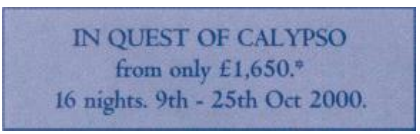
  В Readiris используются интеллектуальные алгоритмы **бинаризации**, которые преобразуют цветные изображения и изображения в оттенках серого в черно-белые изображения, которые применяются при OCR.

  - Выберите **Сгладить цветное изображение**, чтобы сгладить изображение.

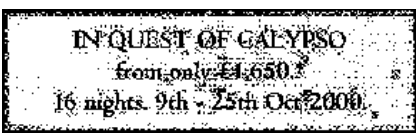
    Эта функция делает изображения в оттенках серого и цветные изображения более однородными за счет сглаживания различий

в насыщенности. В результате достигается более четкий контраст между объектами переднего плана (текстом) и фоном (иллюстрациями).

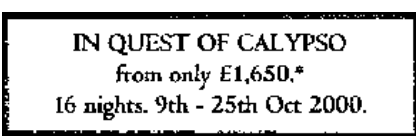
Примечание. Иногда сглаживание является единственным способом отделить текст от цветного фона.



(Оригинальное изображение)



(Бинаризированное черно-белое изображение)



(Сглаженное изображение)

- Используйте бегунок для увеличения или уменьшения **яркости**.

Параметры **яркости** регулирует общую яркость изображения. Используйте их для осветления или затемнения изображения, если текст неразборчив.

Пример 1. Осветление темного изображения для устранения фона.

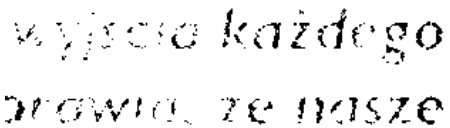
(Цветное изображение)

(Бинаризированное изображение. При использовании параметров бинаризации по умолчанию изображение станет черным.)

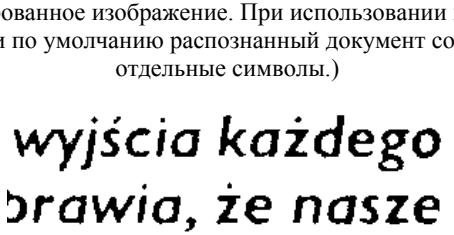
(Результаты распознавания осветленного изображения будут удовлетворительными.)

Пример 2. Затемнение изображения со слишком светлым текстом, который не отображается на бинаризированном изображении.

(Цветное изображение)



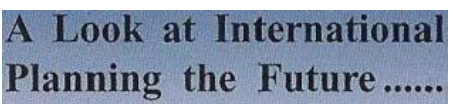
(Бинаризированное изображение. При использовании параметров яркости по умолчанию распознанный документ содержит отдельные символы.)



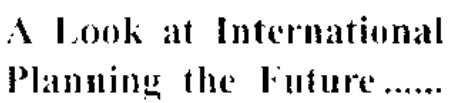
(Результаты распознавания затемненного изображения будут удовлетворительными.)

- Используйте бегунок для увеличения или уменьшения **контраста**.

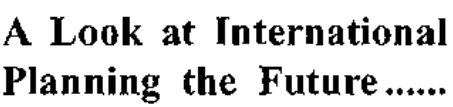
  Параметры **контраста** определяют контраст между темными и светлыми областями изображения. Используйте их, чтобы выделить контуры символов на цветном фоне.



(Цветное изображение)



(Использование параметров контраста по умолчанию приводит к повреждению символов.)



(Увеличение контраста обеспечивает удовлетворительные результаты распознавания.)

- Используйте бегунок для увеличения или уменьшения параметров **ретуши**.

  Ретушь удаляет небольшие пятна на черно-белых изображениях.

- Нажмите кнопку **Применить** для предварительного просмотра результатов.
- Если результаты удовлетворительные, нажмите кнопку **OK**. Если нет, измените параметры еще раз.

Нажмите **Распознать и сохранить**, чтобы распознать документ.

Раздел 7
Сохранение документов в качестве файлов изображений

Отсканированные документы не нужно распознавать немедленно. Из можно сохранить в виде файлов изображений.

Для этого:

- Отсканируйте документ.
- В меню **Файл** выберите команду **Сохранить всю страницу как изображение** или **Сохранить все и страницы как изображение...**.
- Затем откройте сохраненный файл, чтобы выполнить распознавание.

Сохранение только графики

Можно сохранить графические окна без текста документа.

Для этого:

- Отсканируйте или откройте документ.
- В меню **Файл** выберите **Сохранить графику**.
- Все графические элементы документа будут сохранены в одном файле.

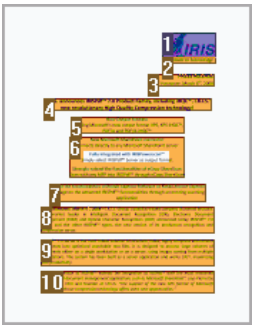
Раздел 8
Формирование окон для документов

Автоматическое формирование окон для документов

При сканировании и открытии документов Readiris автоматически применит функцию **Анализ страниц** для разделения документа на страницы.

По умолчанию используется функция **Анализ страниц** включена. Нажмите функцию **Параметры** и отключите функцию **Анализ страниц**, если автоматический анализ страниц нужно отключить.

Результаты автоматического анализа страниц можно изменить автоматически. Дополнительные сведения см. в разделе **Ручное формирование окон для документов**.



Типы окон

Существует пять типов окон: **текстовые блоки**, **графические области**, **таблицы**, **области штрихкодов** и **области письма печатными буквами**.

В ходе анализа страницы автоматически распознаются текстовые, графические и табличные области. Зоны штрихкодов и письма печатными буквами необходимо обозначить вручную.

> Дополнительные сведения см. в разделе **Ручное формирование окон для документов**.

Каждое окно имеет собственный цветовой код: Текстовым окнам соответствует оранжевый цвет, графическим – фиолетовый, а табличным – розовый. Области штрихкодов обозначаются зеленым цветом, а области письма печатными буквами – синим.

Окна сортируются сверху вниз, справа налево. Числа определяют порядок сортировки окон. Однако порядок сортировки и типы окон можно изменить. Дополнительные сведения см. в разделе **Ручное формирование окон для документов**.

Пропуск текстовых зон на границах страниц

Поскольку при сканировании вокруг изображения формируется темная граница, процесс анализа страницы может сформировать области, состоящие исключительно из шумов.

Чтобы избежать этого, выберите пункт **Не определять области на границах** в меню **Параметры** и отсканируйте документ повторно.

Игнорировать внешнюю область

Параметр **Игнорировать внешнюю область** можно использовать в качестве альтернативы автоматическому формированию окон. Эта функция будет полезна, только если необходимо распознать определенную область страниц.

Включите функцию **Игнорировать внешнюю область**, нажав соответствующую кнопку на панели изображений.

Нарисуйте рамку вокруг части страницы, которую приложение Readiris должно распознать.

> Выбранная часть страницы будет проанализирована автоматически.
>
> Система предложит исключить одну и ту же внешнюю область из процесса анализа всех страниц документа

ФОРМИРОВАНИЕ ОКОН ДЛЯ ДОКУМЕНТОВ ВРУЧНУЮ

Помимо автоматического формирования окон для документов с помощью анализа страниц, Readiris поддерживает ручное формирование окон.

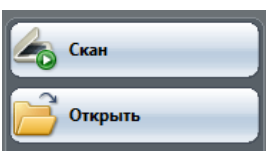
Ручное формирование окон будет полезно, если необходимо изменить результаты автоматического анализа страниц. Кроме того, его можно применить при создании шаблонов формирования окон.

> Дополнительные сведения о шаблонах формирования окон см. в разделе **Использование шаблонов формирования окон**.

Обратите внимание, что окна для зон с письмом печатными буквами и штрихкодами всегда формируются вручную.

Действия

- Для ручного формирования окон нажмите кнопку **Параметры** и снимите флажок **Анализ страниц**.
- Откройте или отсканируйте документ, нажав кнопку **Сканировать** или **Открыть**.



- Выберите правильный **тип окна**, нажав соответствующую кнопку типа окна на панели изображений.

Существует пять типов окон: **текстовые блоки**, **графические области**, **таблицы**, **области штрихкодов** и области **письма печатными буквами**.

- Нарисуйте рамку вокруг текстового блока, областей графики или таблиц, штрихкодов или письма печатными буквами, для которых необходимо сформировать окна.

Дополнительные сведения о распознавании штрихкодов и письма печатными буквами см. в разделах **Распознавание штрихкодов** и **Распознавание письма печатными буквами** соответственно.

- Закончив формирование окон для документа, нажмите кнопку **Распознать и сохранить**, чтобы выполнить OCR.

Сортировка окон

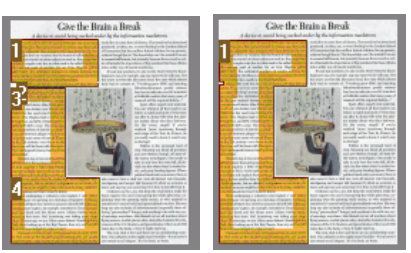
Чтобы изменить порядок сортировки окон, нажмите кнопку **Сортировка** на панели инструментов изображения и щелкните окна левой кнопкой мыши в нужном порядке.

Затем нажмите кнопку **Распознать и сохранить**, чтобы выполнить OCR.

> Окна, которые пользователь не щелкнет левой кнопкой мыши, будут исключены из распознавания.

Рисование многоугольников

Формирование окон вручную не ограничивается прямоугольниками. Многоугольные окна можно создавать путем наложения прямоугольных областей. При пересечении двух окон одного типа они автоматически преобразуются во многоугольник



Автоматический анализ страниц

Если текущая страница слишком сложна для ручного формирования окон, нажмите кнопку **Анализ страницы** на панели изображения, чтобы сформировать окна автоматически.

Обратите внимание, что окна для зон с письмом печатными буквами и штрихкодами всегда формируются вручную.

Изменение типа окна

Чтобы изменить тип окна, щелкните окно правой кнопкой мыши, наведите указатель на на **окно**, затем на **Тип** и щелкните нужный тип окна.

Кроме того, тип окна можно изменить для нескольких окон одновременно:

- Нажмите кнопку **указателя** на главной панели инструментов.

- Удерживайте клавишу **Shift** при выборе нескольких окон.
- Щелкните правой кнопкой любое из выбранных окон, наведите указатель на пункт **Окно**, затем на пункт **Тип** и щелкните нужный тип окна.

Изменение размера окна

Выберите окно для изменения.

Наведите указатель мыши на маркер (маркеры находятся на сторонах и по углам окна).

Щелкните и перетащите маркер для изменения размера окна.

Перемещение окон

Выберите окно для перемещения.

Щелкните внутри окна и перетащите указатель для изменения положения окна.

Распознавание отдельного окна

Щелкните правой кнопкой окно, которое необходимо распознать, и выберите **Копировать как текст**.

> Результаты будут переданы в буфер обмена как основной текст. Это можно сделать и для текста, написанным от руки печатными буквами.
>
> Графические окна и окна со штрихкодами также копируются в буфер обмена.

Удаление окон

Выберите окна, которые необходимо удалить, или щелкните команды **Выделить все** в меню **Правка**, чтоб выделить все окна.

Выберите команды **Вырезать** или **Очистить** в меню **Правка**, чтобы вырезать или удалить окна.

или

Щелкните выбранные окна правой кнопкой мыши, наведите указатель на пункт **Окно**, и выберите **Удалить**.

Удаление малых окон

Некоторые документы, такие как факсы, могут иметь "разрозненные" точки на страницах, которые приведут к созданию лишних окон без текста приложением Readiris.

Чтобы удалить все небольшие окна, выберите пункт **Удалить малые окна** в меню **Правка**.

Рассматриваемый параметр позволяет удалить все окна с размером стороны менее 1,27 см и провести повторную сортировку оставшихся областей.

ИСПОЛЬЗОВАНИЕ ШАБЛОНОВ ФОРМИРОВАНИЯ ОКОН

При распознавании многих документов с одинаковой раскладкой может быть полезно применить **шаблоны формирования окон** вместо автоматического анализа страниц. Таким образом, одинаковая структура окон будет применена ко всем отсканированным и открытым документам для ускорения процесса.

Действия

- Окно для первой страницы документа следует создать вручную с помощью кнопок панели изображения.

  Дополнительные сведения см. в разделе **Ручное формирование окон для документов**.

- В меню **Файл** выберите **Сохранить раскладку**.

- Откройте или сохраните другие страницы документы, нажав кнопку **Открыть** или **Сканировать** на главной панели инструментов.

- В меню **Файл** выберите **Загрузить раскладку**.
- Выберите сохраненный файл раскладки.
- Чтобы применить раскладку ко **всем** открытым или отсканированным страницам, **Применить раскладку ко всем страницам** в диалоговом окне **Файл раскладки**.
- Нажмите кнопку **Открыть**, чтобы загрузить файл раскладки.

Обратите внимание, что при добавлении документа в Readiris файл раскладки необходимо загрузить снова, так как функция анализа страниц включена по умолчанию.

Игнорировать внешнюю область

В качестве альтернативы шаблонам формированию окон можно использовать параметр **Игнорировать внешнюю область.** Таким образом, можно задать определенную область на странице, которую необходимо распознать. Все данные за пределами области OCR будут исключены из распознавания.

Действия

- Включите функцию **Игнорировать внешнюю область**, нажав соответствующую кнопку на панели изображений.

- Нарисуйте рамку вокруг части страницы, которую приложение Readiris должно распознать.

Выбранная часть страницы будет проанализирована автоматически.

Система спросит, нужно ли проигнорировать ту же внешнюю зону на всех страницах документа.

Чтобы отменить эту функцию, запустите процедуру **Анализ страницы**, нажав кнопку **Анализ страницы** на панели инструментов изображения.

- Нажмите **Распознать и сохранить**, чтобы выполнить распознавание.

Раздел 9
Распознавание документов

Введение

При распознавании документов Readiris применяет лингвистические технологии. Это позволяет распознавать текст, таблицы, графические объекты, штрихкоды и текст, написанный от руки печатными буквами, в документах всех типов. Более того, Readiris обрабатывает даже сложные документы с колонками, документы низкого качества, факсы, распечатки с матричного принтера, плохо отсканированные и скопированные документы со слишком темными или слишком светлыми цветами и др.

Readiris поддерживает 128 **языков**: все американские и европейские языки, в том числе центрально-европейские, кириллические и балтийские, а также греческий и турецкий. Кроме того, Readiris считывает документы на арабском, фарси и иврите, также на четырех азиатских языках — японском, упрощенном китайском, традиционном китайском и корейском (дополнительная функция). Readiris справляется даже со смешанными алфавитами: приложение обнаруживает ”западные” слова в кириллических документах, документах на греческом, арабском, иврите и азиатских языках, которые часто представляют собой непереводимые названия компаний и продуктов, написанные латиницей.

Readiris основывается на самых передовых методах распознавания. Распознавание текста, независимое от шрифтов, дополняется технологиями самообучения. Система может

изучать новые символы и слова посредством контекстного и лингвистического анализа. Это значит, что точность OCR системы распознавания будет значительно улучшена по мере использования продукта.

Кроме того, Readiris предлагает (дополнительную) функцию проверки пользователем. При активации функция проверки пользователем (**Интерактивное обучение**) не только отмечает символы, в которых система распознавания не уверена, но и повышает ее точность. Все подтвержденные решения временно запоминаются во время распознавания, что повышает скорость системы и уверенность распознавания, делая ее более интеллектуальной по мере использования продукта. Это мощное средство обучение также позволяет "научить" Readiris работе со специальными знаками, такими как математические символы и графические метки, а также с искаженными шрифтами.

Кроме того, результаты интерактивного обучения можно сохранить в **каталогах шрифтов** для будущего использования.

ВЫБОР ЯЗЫКА ДОКУМЕНТА

Readiris обеспечивает оптическое распознавание символов на 128 языках. Поддерживаются все американские и европейские языки, в том числе центрально-европейские, кириллические и балтийские, а также греческий и турецкий.

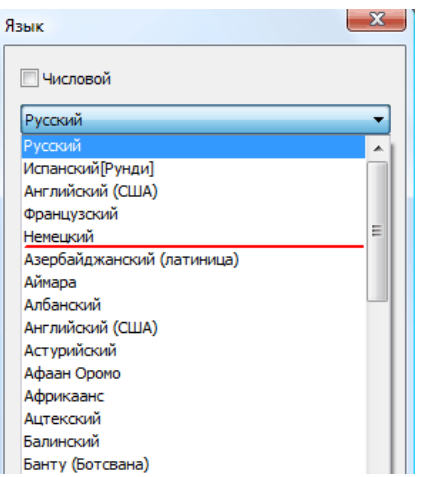
Версии Readiris Pro Asian и Readiris Corporate Asian также распознают документы на японском, упрощенном китайском, традиционном китайском и корейском.

Readiris Pro Middle-East и Readiris Corporate Middle-East также распознают документы на арабском, фарси и иврите.

Чтобы приложение Readiris могло распознать документ, необходимо указать его язык.

Для этого:

Нажмите кнопку **Глобус** на главной панели инструментов и выберите язык в раскрывающемся списке.



5 недавно выбранных языков будут находиться в верхней части списка языков.

Важно! При работе с документами на азиатских языках, иврите и арабском *перед* анализом страницы необходимо указать язык документа. Для обработки таких документов применяются специальные подпрограммы анализа страниц.

Кроме того, распознавание можно ограничить набором **цифровых** символов для оптимального распознавания таблиц и цифр. В этом случае Readiris будет распознавать только цифры 0-9 и следующие последовательности символов:

+ знак "плюс"	* звездочка	/ косая черта	% знак процентов	, запятая
. точка	(открывающая скобка	) закрывающая скобка	- дефис	= знак "равно"
$ знак доллара	£ знак фунта	€ знак евро	¥ знак йены	

Чтобы активировать цифровой режим, выберите параметр **Цифровой** в диалоговом окне **Язык**.

Распознавание документов, содержащих текст на разных языках

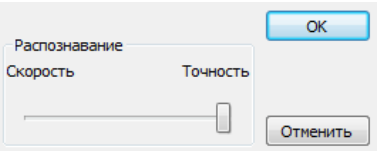
Readiris позволяет распознавать смешанные наборы символов. При этом Readiris автоматически переключается между языками в середине предложения и распознает английские слова English (имена продуктов и др.), которые могут встречаться в текстах на "экзотических" языках.

Нажмите кнопку **Глобус** на главной панели инструментов и выберите сочетание языков в раскрывающемся списке.

Примечание. При обработке документов на азиатских языках и иврите смешанные наборы символов применяются автоматически.

Баланс скорости и точности

Задайте соотношение между **скоростью OCR** и **точностью OCR**.



Это соотношение можно выбирать для латинских, кириллических и греческого алфавитов.

Совет. При обработке изображений низкого качества приоритет следует отдать точности.

ЗАДАНИЕ ХАРАКТЕРИСТИК ДОКУМЕНТА

Помимо языка документа другие характеристики, такие как **Тип шрифта** и **Расстояние между символами** играют важную роль в процессе распознавания.

Тип шрифта

Readiris различает "обычные" документы и документы, напечатанные на матричном принтере. Символы, напечатанные на 9-игольчатом матричном принтере, состоят из отдельных изолированных точек.

Для распознавания документов, напечатанных на матричном принтере, необходимы специальные методы сегментации и распознавания, которые нужно активировать.

Far out in the uncharted back

Чтобы выбрать тип шрифта:

- В меню **Параметры** наведите указатель на пункт **Тип шрифта**.
- Значение типа шрифта по умолчанию — **Автоматич.**.

Таким образом Readiris распознает документы, напечатанные на "25-игольчатом" матричном принтере ("NLQ") и другими "обычными" методами печати.

- Чтобы распознать только документы, напечатанные на матричном принтере, выберите **Матричная печать**.

  Readiris распознает так называемую "черновую" или "9-игольчатую" матричную печать.

Расстояние между символами

Расстояние между символами — это количество символов на дюйм в данном шрифте. Расстояние между символами может быть **фиксированным**, в этом случае все символы будут иметь одинаковую ширину или **пропорциональным**, в этом случае ширина символов будет различной.

Чтобы выбрать расстояние между символами:

- В меню **Параметры** наведите указатель на пункт **Расстояние между символами**.
- По умолчанию выбрано значение **Автоматич.**.
- Выберите **Фиксированный** если все символы шрифта имеют одинаковую ширину. Это часто встречается в старых документах, напечатанных на печатной машинке.
- Выберите **Пропорциональный**, если символы шрифта имеют разную длину. Практически все шрифты, которые применяются при печати газет, журналов и книг, относятся к категории пропорциональных.

Важно! Эти характеристики документов не относятся к документам на иврите, арабском и азиатских языках.

Использование интерактивного обучения

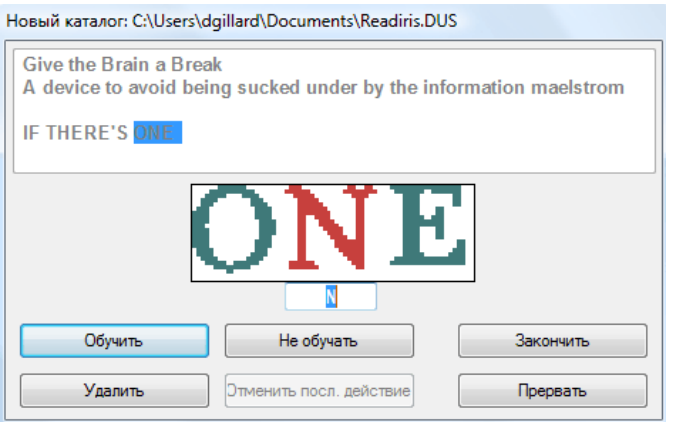
Readiris поддерживает функцию интерактивного обучения. С помощью **интерактивного обучения** можно научить систему работе с шрифтами и начертаниями символов и исправить результаты OCR при необходимости. Во время интерактивного обучения все символы, в которых система распознавания не уверена, отображаются в окне предварительного просмотра вместе с родительским словом и предложенным решением.

Интерактивное обучение может значительно повысить точность системы распознавания и будет особенно полезна при распознавании искаженных и трудночитаемых документов. Кроме того, эту функцию можно использовать для обучения системы Readiris работе со специальными символами, которые она не может распознавать изначально, например математическими и научными символами, а также графическими метками.

Чтобы включить интерактивное обучение:

- В меню **Обучение** выберите **Интерактивное обучение**.
- Задайте необходимые параметры и нажмите кнопку **Распознать и сохранить**, чтобы распознать документ.
- По окончании распознавания Readiris перейдет в режим интерактивного обучения.

  Символы, в которых система распознавания не уверена, будут выведены на экран.



Если результаты верны:

- Нажмите кнопку **Обучение**, чтобы сохранить результаты как верные.

  Результаты обучения временно сохраняются в памяти компьютера на время распознавания. Readiris не будет отображать изученные символы при распознавании оставшейся части документа.

  При распознавании нового документа результаты обучения будут удалены.

  Чтобы сохранить результаты распознавания, воспользуйтесь каталогом шрифтов. Дополнительные сведения см. в разделе **Использование каталогов шрифтов**.

- Нажмите кнопку **Готово**, чтобы сохранить решения, предложенные программным обеспечением.

Если результаты неверны:

- Введите верные символы и нажмите кнопку **Обучить**.

или

- Нажмите кнопку **Не обучать**, чтобы сохранить результаты как неточные.

  Используйте эту команду для поврежденных символов, которые могут быть приняты за другие символы после обучения. Например, цифра 1 и буква I, которые имеют одинаковое начертание во многих шрифтах.

- Нажмите кнопку **Удалить**, чтобы удалить символы из вывода.

  Используйте эту кнопку, чтобы предотвратить возникновение шума в файле вывода.

- Нажмите кнопку **Отменить**, чтобы исправить ошибки.

  В памяти Readiris хранятся сведения о последних 32 операциях.

- Нажмите кнопку **Прекратить**, чтобы прекратить интерактивное обучение.

Все результаты обучения будут удалены. В следующий раз, когда пользователь нажмет кнопку **Распознать и сохранить**, интерактивное обучение запустится снова.

ИСПОЛЬЗОВАНИЕ КАТАЛОГОВ ШРИФТОВ

При сканировании нескольких документов одного типа и с одинаковым качеством печати и шрифтов повторение процесса обучения для каждого из них может быть нежелательно. В этом случае можно воспользоваться каталогами шрифтов. Каталоги шрифтов содержат данные о шрифтах, полученные в ходе **интерактивного обучения** и могут значительно улучшить результаты распознавания.

Обратите внимание, что каталоги шрифтов могут содержать не более 500 начертаний. Рекомендуется создавать отдельные каталоги шрифтов для каждого приложения.

Для создания нового каталога шрифтов:

- В меню **Обучение** выберите **Новый каталог шрифтов**.

  По умолчанию Readiris открывает каталог **Readiris.dus**. Измените имя файла и нажмите кнопку **Сохранить**, чтобы сохранить его как новый каталог.

- Выберите пункт **Интерактивное обучение** в меню **Обучить**, чтобы активировать соответствующую функцию.
- Нажмите **Распознать и сохранить**, чтобы распознать документ.
- Readiris перейдет к этапу интерактивного распознавания. Используйте кнопки в диалоговом окне, чтобы сохранить символы в каталог шрифтов.

Чтобы применить существующий каталог шрифтов:

- В меню **Обучение** выберите **Каталог шрифтов**.
- Выберите нужный каталог и нажмите кнопку **Открыть**.
- В меню **Обучить** выберите **Присоединить каталог шрифтов** или **Читать каталог шрифтов**.

  Если выбран параметр **Присоединить каталог шрифтов**, обязательно включите **Интерактивное обучение**. Readiris распознает начертания шрифтов, сохраненные в словаре, и применит интерактивное обучение, чтобы сохранить новые данные в словаре.

  Если выбран параметр **Читать каталог шрифтов**, Readiris распознает начертания шрифтов, сохраненные в словаре, но не будет добавлять в него новое содержимое, даже если параметр **Интерактивное обучение** включен. Обратите внимание, что

функция **Интерактивное обучение** все еще может быть полезна для проверки и (при необходимости) исправления результатов распознавания, не сохраненных в каталоге шрифтов.

Внимание! Не выбирайте пункт **Каталог шрифтов** в меню **Обучить** и не открывайте существующий словарь, если включен режим каталога шрифтов **Новый каталог**. В противном случае содержимое существующего каталога шрифтов будет удалено.

- Нажмите **Распознать и сохранить**, чтобы распознать документ.

Раздел 10
Форматирование и сохранение документов

Форматирование документов

Документы, распознаваемые в Readiris, могут сохраняться в различных форматах. Readiris сохраняет результаты OCR в следующий форматах: файлы Adobe Acrobat PDF, текстовые файлы Microsoft XPS, Word, WordML, RTF, OpenDocument, HTML и XML, таблицы SpreadsheetML, а также текстовые файлы Ansi и Unicode.

Кроме того, Readiris может экспортировать результаты непосредственно в конечные приложения, такие как Microsoft Word и Excel, Adobe Reader, Microsoft XPS Viewer, основные веб-браузеры и почтовые клиенты.

> Примечание. RTF-файлы, созданные Readiris, также можно открыть в офисных веб-приложениях, таких как AjaxWrite, ThinkFree, Zoho Writer и Google Writely (поддерживающих RTF- и HTML-документы). При использовании одного из этих приложений, выберите параметр макета **Сохранить форматирование слов и абзацев**.

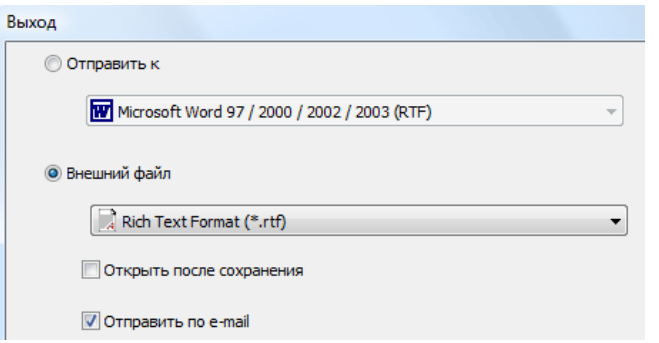
Действия

- Нажмите кнопку **Формат** на главной панели инструментов и выберите формат вывода.

  Readiris:

- отправит документы в приложение, которое откроется автоматически или
- сохранит документы во внешний файл.

Параметр **Отправить по эл. почте** помогает создать новое сообщение электронной почты и вставить в него распознанный документ в качестве вложения.



- Выберите соответствующие вкладки, чтобы задать нужные параметры.

Параметры, недоступные для выбранного формата вывода, будут затемнены.

Часто используемые форматы вывода файлов, а также их параметры рассматриваются в разделах ниже.

- Нажмите **Распознать и сохранить**, чтобы распознать и сохранить документы.

Результаты OCR можно экспортировать несколько раз, не повторяя распознание. Нажмите кнопку **Формат** еще раз и измените формат текста и параметры форматирования. Нажмите кнопку **Распознать и сохранить** еще раз.

Приложение Readiris позволяет задавать **свойства документов** PDF, XPS, Word, RTF, WordML, SpreadsheetML и HTML для поиска и сортировки. Чтобы задать свойства документа, выберите

пункт **Свойства документа** в меню **Файл**. Обратите внимание, что свойства документа также доступны в диалоговом окне **Файл вывода**, которое открывается при нажатии кнопки **Распознать и сохранить**.

Обратите внимание, что при сохранении документа в качестве **внешнего файла** создается отдельный файл вывода для каждой страницы в Readiris или сохранить все страницы, принадлежащие к одному документу, в один файл. Просто выберите соответствующие параметры в окне **Файл вывода**: **Создать один файл на страницу** и **Создать один файл на документ**.

Однако обратите внимание, что параметры **Создать один файл на страницу** и **Создать один файл на документ** доступны только при сохранении документов во внешний файл, но не при открытии документов в конечном приложении.

ФОРМАТИРОВАНИЕ ТЕКСТОВЫХ ДОКУМЕНТОВ

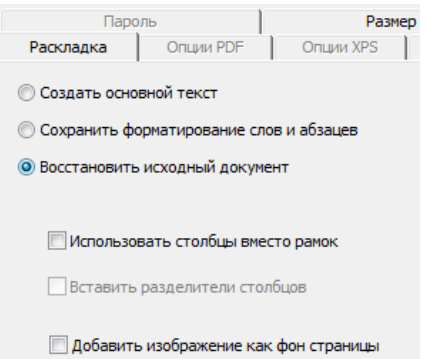
Readiris позволяет создавать несколько документы в нескольких форматах. Поддерживаются форматы Word, WordML, RTF, txt и текст OpenDocument.

Чтобы создать текстовые файлы вывода:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный формат вывода в раскрывающемся списке **Отправить** или **Внешний файл**.
- В зависимости от выбранного формата текста будут доступны следующие параметры.

  Параметры, недоступные для выбранного формата текста, будут затемнены.

Параметры макета



- Параметр **Создать основной текст** запрещает форматирование текста приложением Readiris.

  Если этот параметр установлен, Readiris возвращает сплошной текст без форматирования.

- Параметр **Сохранить форматирование слов и абзацев** занимает промежуточное положение между созданием основного текста и автоматическим форматированием.

  Тип, размер и стиль шрифта сохраняются при распознавании.

  Табуляции и выравнивание каждого блока восстанавливаются.

  В то же время, текстовые блоки и колонки не восстанавливаются, абзацы следуют друг за другом.

  Таблицы считываются верно.

- Параметр **Восстановить исходный документ** позволяет создать точную копию исходного документа.

  Readiris создает точную копию исходного документа, не сканированное изображение.

Кроме того, Readiris восстанавливает любые гиперссылки на адреса эл. почты и веб-сайты.

- Параметр **Использовать столбцы вместо рамок** позволяет создать документы с колонками.

  Редактировать текст в колонках проще, чем документы с несколькими рамками. Это связано с естественным "перетеканием" текста из одной колонки в другую.

  Примечание. Если системе не удается обнаружить колонки в исходном документе, то в качестве резервного варианта применяются рамки.

- Параметр **Вставить разрывы колонок** позволяет вставить жесткий разрыв после каждой колонки.

  Любой текст, который пользователь редактирует, добавляет или удаляет, остается в колонке; автоматический выход текста за пределы разрыва колонки невозможен.

  Совет. Если основной текст размещен в колонках, устанавливать этот параметр не следует. В таком случае текст будет свободно перетекать из одной колонки в другую.

- Параметр **Добавить изображение как фон страницы** позволяет использовать отсканированное изображение в качестве фона под распознанным текстом.

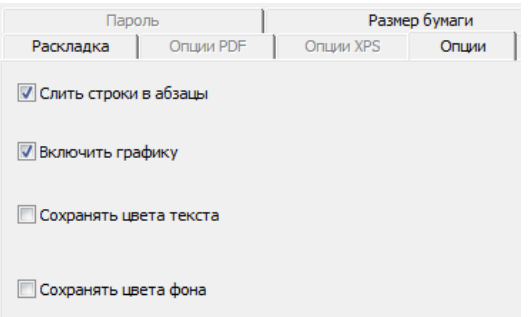
  Однако этот параметр значительно увеличивает размер файлов вывода.

  Формат **PDF - текст-изображение** позволяет добиться аналогичного результата при создании PDF-файлов.

  Параметр **Сохранять цвета фона** на вкладке **Параметры** предлагает менее радикальное и более компактное решение.

Общие параметры

Выберите вкладку **Параметры**, чтобы перейти к общим параметрам.



- Параметр **Слить строки в абзацы** делает возможным автоматическое обнаружение абзацев.

  Readiris выполняет перенос по словам, пока не начнется новый параграф, и "склеивает” слова с переносами в конце строки.

- Параметр **Включить графику** включает графику в файлы с автоматическим форматированием.

  Это необходимо при создании точной копии документа.

  Параметры графики на вкладке **Графика** определяют режим цвета и разрешение графических объектов, сохраняемых в файлах вывода.

- Параметр **Сохранять цвета текста** предназначен для сохранения исходных цветов текста в ходе распознавания.

- Параметр **Сохранять цвета фона** предназначен для сохранения исходных цветов фона в ходе распознавания.

  Для каждого параграфа файла вывода создается отдельный цветной фон.

Форматы бумаги

Можно указать предпочтительные форматы бумаги в зависимости от выбранного формата:

- Выберите вкладку **Формат бумаги** и воспользуйтесь кнопками со стрелками, чтобы применить и исключить форматы бумаги.
- Приложение Readiris перебирает активные форматы бумаги в указанном порядке и выбирает первый формат, достаточный для размещения сканированного документа.

ФОРМАТИРОВАНИЕ ТАБЛИЧНЫХ ДОКУМЕНТОВ

Readiris позволяет вывести табличные данные в электронные таблицы, текстовые процессоры и веб-браузеры: в редакторах электронных таблиц исходные таблицы воссоздаются ячейка за ячейкой, а в текстовых процессорах вставляются в качестве объектов таблиц.

Readiris распознает таблицы с сетками и без сеток.

Performance test optical media	
CD-ROM Digital Versatile Disk	Average time (n
CD-ROM 32x speed	60
DVD	58
Tested on 333 MHz Pentium II	

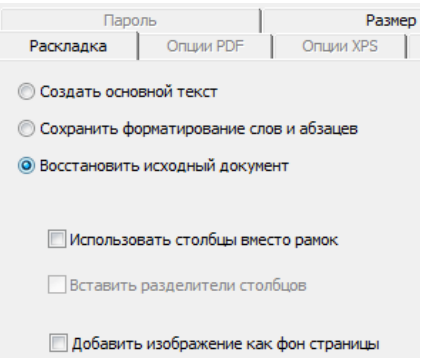
123.985	69	31
987.745	129	2
19	149	91
287.410	49	52

(с сеткой) (без сетки)

Для создания табличных документов:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный формат вывода в раскрывающемся списке **Отправить** или **Внешний файл**.

- Выберите параметры макета.

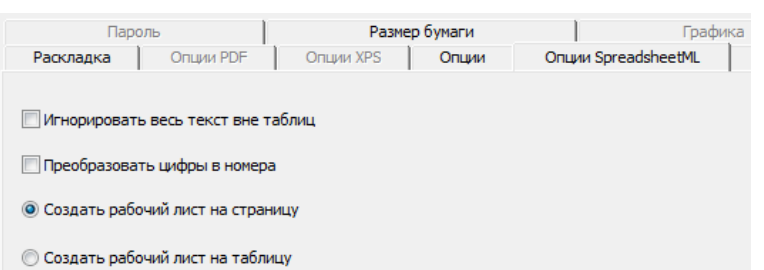


Дополнительные сведения о параметрах форматирования см. в разделе **Форматирование текстовых документов**.

Параметры SpreadsheetML

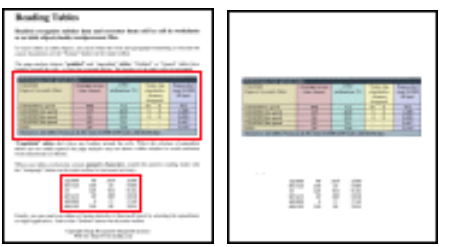
Если в качестве конечного приложения выбрано ПО **Microsoft Excel 2002/2003**, будут доступны параметры **SpreadsheetML**. Выберите вкладку **Параметры SpreadsheetML**, чтобы показать их:

Обратите внимание, что параметр **Восстановить исходный документ** будет недоступен, если выбран этот формат.



- Параметр **Игнорировать весь текст вне таблиц** позволяет сохранить таблицы и пропустить все прочие результаты распознавания.

  Все данные внутри таблиц будут считаны, данные вне таблиц — нет.



- Параметр **Преобразовать цифры в номера** распознает цифры как номера.

  Таким образом, этими ячейками можно производить арифметические операции. При этом текстовые ячейки (в любой таблице) остаются текстовыми.

  Обратите внимание, что только цифры внутри таблиц кодируются как номера.

  Excel выполняет математические операции только над данными, закодированными как числа.

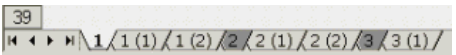
- Параметр **Создать рабочий лист на страницу** позволяет разместить в одной электронной таблице весь материал со сканированной страницы.

  Таким образом, даже если на странице содержатся таблицы и текст, все эти данные сохраняются в одной электронной таблице.

- Параметр **Создать рабочий лист на таблицу** позволяет разместить каждую таблицу на отдельном листе электронной таблицы, а весь распознанный текст,

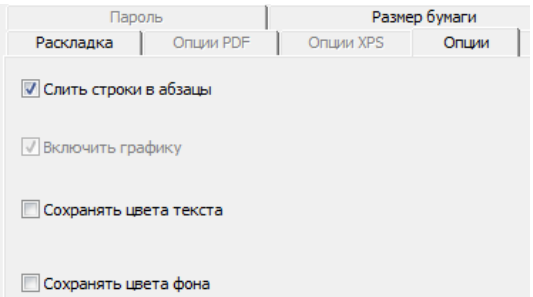
находящийся за пределами таблиц, сохранить на еще одном листе.

Если распознанный документ состоит из **нескольких страниц**, структура будет повторяться на каждой странице.



Общие параметры

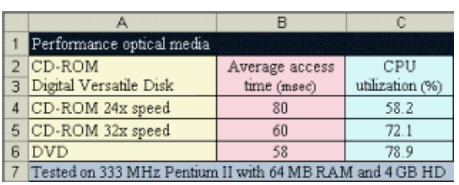
Выберите вкладку **Параметры**, чтобы перейти к общим параметрам.



- Параметр **Слить строки в абзацы** делает возможным автоматическое обнаружение абзацев.

  Readiris выполняет перенос по словам, пока не начнется новый параграф, и "склеивает” слова с переносами в конце строки.

- Параметр **Сохранять цвета фона** позволяет восстановить цвет фона каждой ячейки.



	A	B	C
1	Performance optical media		
2	CD-ROM	Average access	CPU
3	Digital Versatile Disk	time (msec)	utilization (%)
4	CD-ROM 24x speed	80	58.2
5	CD-ROM 32x speed	60	72.1
6	DVD	58	78.9
7	Tested on 333 MHz Pentium II with 64 MB RAM and 4 GB HD		

Форматы бумаги

Можно указать предпочтительные форматы бумаги в зависимости от выбранного формата:

- Выберите вкладку **Формат бумаги** и воспользуйтесь кнопками со стрелками, чтобы применить и исключить форматы бумаги.
- Приложение Readiris перебирает активные форматы бумаги в указанном порядке и выбирает первый формат, достаточный для размещения сканированного документа.

Создание PDF-документов

Readiris создает 4 типа PDF-документов: **текст**,**текст-изображение**, **изображение-текст** и **изображение**.

Для создания PDF-документа:

Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный тип PDF в раскрывающемся списке **Отправить** или **Внешний файл**:

PDF - изображение

Если выбран формат PDF - **Изображение**, Readiris создаст PDF-документы, включающие только изображения, и не будет выполнять OCR.

PDF - изображение-текст

Если выбран формат PDF **Изображение-текст**, Readiris распознает текст и создаст PDF-файлы с возможностью поиска, содержащие изображения страницы и распознанный текст.

Изображение страницы размещается над текстом.

В этом формате вы всегда будете видеть документ в исходной форме (как он отсканирован), но при этом сможете искать, копировать и вставлять распознанный текст, скрытый за изображением. Это значит, что данный формат можно использовать для **архивирования**.

PDF - текст

Если выбран формат PDF - **Текст**, Readiris распознает текст и создаст PDF-файлы с возможностью поиска.

Изображение страницы в таком однослойном PDF-файле отсутствует.

PDF - текст-изображение

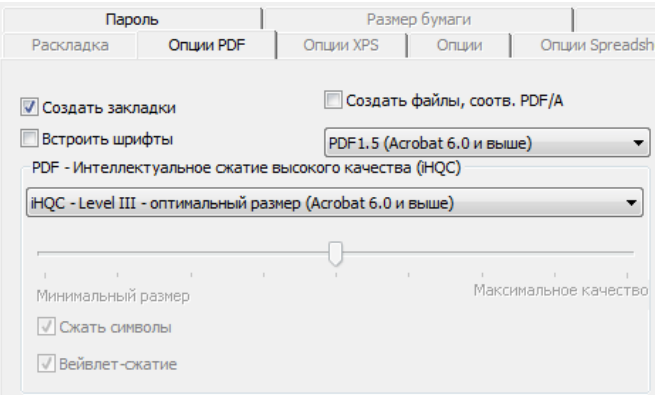
Если выбран формат PDF **Текст-изображение**, Readiris распознает текст и создаст PDF-файлы с возможностью поиска, содержащие изображения страницы и распознанный текст.

Изображение страницы будет находиться под текстом.

ВЫБОР ПАРАМЕТРОВ PDF

Чтобы выбрать параметры PDF:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный тип PDF в раскрывающемся списке **Отправить** или **Внешний файл**.
- В зависимости от выбранного **Типа PDF** будет доступно несколько параметров. Выберите вкладку **Параметры PDF**, чтобы получить доступ к ним:



Создать закладки

Параметр **Создать закладки** подразумевает создание закладок для каждого текстового блока, графического элемента и таблицы в файлах Adobe Acrobat PDF.

Встраивание шрифтов

Выберите параметр **Встроить шрифты**, чтобы встроить шрифты в файлы Adobe Acrobat PDF.

Это предотвратит замену шрифтов и позволяет пользователям просматривать текст с использованием оригинальных шрифтов независимо от конфигурации компьютера.

Встраивание шрифтов несколько увеличивает размер распознанных документов.

СЖАТИЕ IHQC ДЛЯ PDF-ДОКУМЕНТОВ

Помимо вывода в "обычные" PDF-файлы Readiris позволяет создавать PDF-документы со сжатием iHQC: Для PDF-документов типа **Изображение-текст** и **Изображение** можно обеспечить высокую степень сжатия по алгоритму iHQC без потери качества изображения.

Аббревиатура iHQC расшифровывается как **intelligent High-Quality Compression** (интеллектуальное сжатие высокого качества), и обозначает эффективную проприетарную технологию сжатия I.R.I.S. iHQC для изображений — это то же, что MP3 для музыки и DivX для фильмов.

Для создания PDF-документа со сжатием iHQC:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите один из двух форматов вывода.
- В раскрывающемся списке **Отправить** или **Внешний файл** выберите нужный тип PDF: **PDF - изображение-текст** или **PDF - изображение**.
- На вкладке **Параметры PDF** выберите нужный уровень сжатия.

  Readiris Pro поддерживает уровни сжатия **Level I - оптимальный размер** и **Level I - оптимальное качество**.

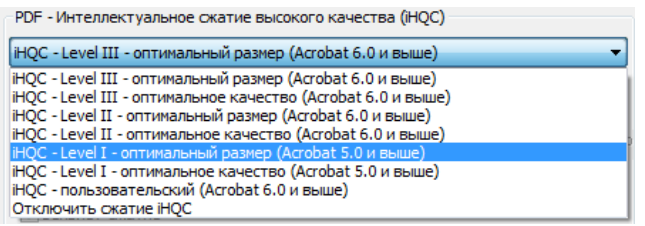
  Readiris Corporate также поддерживает **Level II и III - оптимальный размер** и **оптимальное качество**, а также **Пользовательский** уровень сжатия.

  При сжатии Level II параметр **Сжатие символов** включается автоматически для сжатия текста.

  При сжатии Level III параметр **Сжатие волновых пакетов**

включается автоматически для сжатия графики.

Если выбран **Пользовательский** уровень сжатия, эти параметры можно включить или отключить независимо друг от друга. Кроме того, можно использовать бегунок для задания соотношения **Оптимальный размер - Оптимальное качество**.



- Нажмите **Распознать и сохранить**, чтобы распознать документ и сохранить PDF-файл со сжатием iHQC.

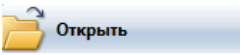
Перепрофилирование PDF-документов

Помимо создания PDF-файлов, Readiris поддерживает перепрофилирование PDF-файлов: Readiris преобразует **изображения** в формате PDF в **текст** PDF или любой поддерживаемый формат текста и **разблокирует** содержимое PDF только для чтения.

> **Предупреждение**. Приложение Readiris не может открывать документы PDF, защищенные паролем

Действия

- Нажмите кнопку **Открыть** на главной панели инструментов и выберите PDF-файл, который ПО Readiris должно перепрофилировать.



В разделе **Диапазон страниц** диалогового окна выберите поле **Страницы** и укажите страницы, которые приложение Readiris должно открыть.

- Нажмите кнопку **Открыть** в диалоговом окне, чтобы открыть выбранный PDF-файл.
- Нажмите кнопку **Формат** на главной панели инструментов и выберите тип PDF.

  Дополнительные сведения о параметрах форматирования см. в разделе **Создание PDF-документов**.

- Нажмите **Распознать и сохранить**, чтобы перепрофилировать документы.

СОЗДАНИЕ XPS-ДОКУМЕНТОВ

Readiris создает 4 типа XPS-файлов: **Текст**,**текст-изображение**, **изображение-текст** и **изображение**.

XPS расшифровывается как XML Paper Specification, это формат с фиксированной компоновкой, разработанный корпорацией Microsoft.

Для создания XPS-документа:

Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный тип XPS в раскрывающемся списке **Отправить** или **Внешний файл**:

XPS - изображение

Если выбран формат XPS - **Изображение**, Readiris создаст PDF-документы, включающие только изображения и не будет выполнять OCR.

XPS - изображение-текст

Если выбран формат XPS **Изображение-текст**, Readiris распознает текст и создаст XPS-файлы с возможностью поиска, содержащие изображения страницы и распознанный текст.

Изображение страницы размещается над текстом.

В этом формате вы всегда будете видеть документ в исходной форме (как он отсканирован), но при этом сможете искать, копировать и вставлять распознанный текст, скрытый за изображением. Это значит, что данный формат можно использовать для **архивирования**.

XPS - текст

Если выбран формат XPS - **Текст**, Readiris распознает текст и создаст XPS-файлы с возможностью поиска.

Изображение страницы в таком однослойном XPS-файле отсутствует.

XPS - текст-изображение

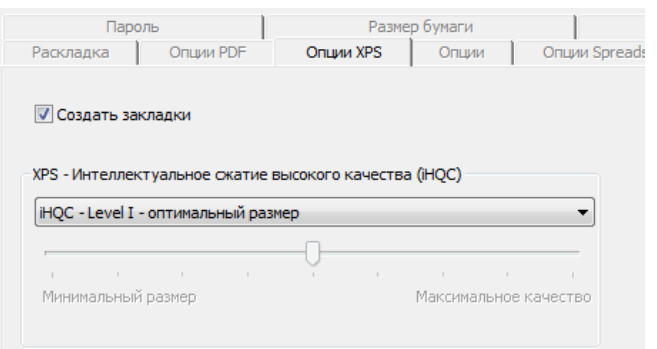
Если выбран формат XPS **- Текст-изображение**, Readiris распознает текст и создаст XPS-файлы с возможностью поиска, содержащие изображения страницы и распознанный текст.

Изображение страницы будет находиться под текстом.

Выбор параметров XPS

Чтобы выбрать параметры XPS:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный тип XPS в раскрывающемся списке **Отправить** или **Внешний файл**.
- В зависимости от выбранного **типа XPS** будет доступно несколько параметров. Выберите вкладку **Параметры XPS**, чтобы получить доступ к ним:



Создать закладки

Параметр **Создать закладки** подразумевает создание закладок для каждого текстового блока, графического объекта и таблицы в файлах Microsoft XPS.

Сжатие iHQC для XPS-документов

Помимо вывода в "обычные" XPS-файлы Readiris позволяет создавать XPS-документы со сжатием iHQC: Для XPS-

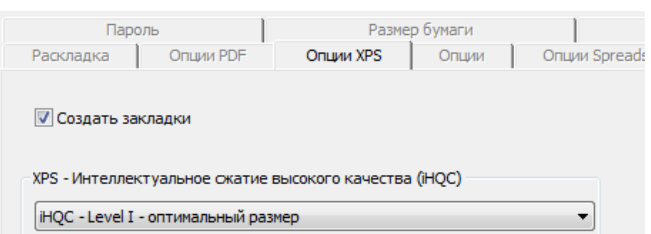
документов типа **Изображение-текст** и **Изображение** можно обеспечить высокую степень сжатия по алгоритму iHQC без потери качества изображения.

Аббревиатура iHQC расшифровывается как **intelligent High-Quality Compression** (интеллектуальное сжатие высокого качества), и обозначает эффективную проприетарную технологию сжатия I.R.I.S. iHQC для изображений — это то же, что MP3 для музыки и DivX для фильмов.

Для создания XPS-документа со сжатием iHQC:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите один из двух форматов вывода.
- В раскрывающемся списке **Отправить** или **Внешний файл** выберите нужный тип XPS: **XPS - изображение-текст** или **XPS - изображение**.
- На вкладке **Параметры XPS** выберите нужный уровень сжатия.

  Readiris Pro поддерживает уровни сжатия **Level I - оптимальный размер** и **Level I - оптимальное качество**.

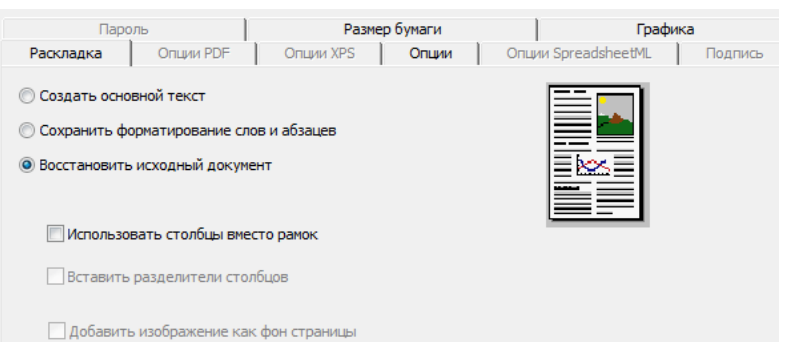


Выбор параметров графики

В зависимости от выбранного формата вывода и конечного приложения могут быть доступны расширенные параметры графики. Параметры графики можно использовать для изменения качества и разрешения изображения.

Чтобы открыть параметры графики:

- Нажмите кнопку **Формат** на главной панели инструментов и выберите нужный формат вывода в раскрывающемся списке **Отправить** или **Внешний файл**.
- Выберите вкладку **Графика**, чтобы открыть параметры.



Режим цвета

По умолчанию Readiris сохраняет графические объекты в **цвете**.

Выберите **Черно-белый** режим, чтобы сохранить графические объекты в черно-белом формате.

Разрешение

По умолчанию Readiris **сохраняет разрешение сканирования**.

Кроме того, **разрешение можно уменьшить**.

Совет. При сохранении документов в виде HTML-файлов для публикации в Интернете, уменьшите разрешение до 70 точек/дюйм (разрешение экрана).

Качество JPEG

Графические объекты в документах формата PDF, XPS, Word и RTF сохраняются в формате JPEG.

Для регулировки качества JPEG используется бегунок.

Сжатие JPEG 2000

При сохранении файлов в форматах PDF и XPS ПО Readiris может применит сжатие JPEG 2000 к цветным изображениям и изображениям в оттенках серого, сохраненных в этих файлах.

> JPEG 2000 — это новая, более компактная версия стандарта JPEG.

Выберите параметр **Сжатие JPEG 2000**, чтобы применить его.

Раздел 11
Сохранение и загрузка параметров

Все параметры, указанные в приложении Readiris, автоматически сохраняются для использования после закрытия приложения. Чтобы восстановить заводские параметры, выберите команду **Восстановить заводские параметры** в меню **Файл**.

При сканировании различных групп документов, которые требуют различных параметров, будет полезно сохранить отдельный **файл параметров** для каждой группы.

Действия

- Выберите параметры, которые следует использовать для той или иной группы документов.
- В меню **Файл** выберите **Сохранить параметры**.
- Во время последующего сканирования или открытия документа из этой группы выберите команду **Загрузить параметры** в меню **Файл**.
- Выберите нужный файл параметров и нажмите кнопку **Открыть**.
- Нажмите **Распознать и сохранить**, чтобы распознать документы с использованием верных параметров.

Примечание. Команда **Информ** в меню **Файл** предоставляет обзор наиболее важных из выбранных параметров.

Раздел 12 Распознавание многостраничных документов

Открытие и распознавание нескольких файлов изображений

Приложение Readiris разработано для одновременной обработки нескольких файлов изображений.

Чтобы открыть несколько файлов изображений:

- Нажмите кнопку **Открыть** на главной панели инструментов.

- Выберите файлы изображения, которые необходимо открыть:
    - Выберите первый файл изображения и держите клавишу **Ctrl** нажатой при добавлении дополнительных изображений или;
    - выделите непрерывный диапазон изображений, для этого щелкните первое из них, а затем выберите последнее, удерживая клавишу **Shift**.

Примечание. При открытии одного файла, включающего несколько страниц, например многостраничного TIFF-файла или PDF-документа, можно указать диапазон страниц. В разделе **Диапазон страниц** диалогового окна выберите поле **Страницы** и укажите страницы, которые необходимо распознать.

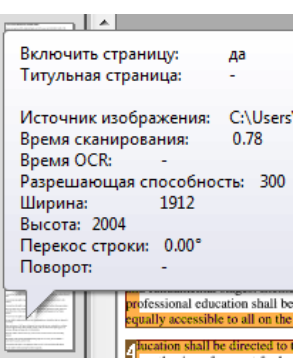
- Нажмите кнопку **Открыть**, чтобы открыть изображения.

Обратите внимание, что файлы изображений можно перетащить из проводника Windows в окно изображения Readiris, чтобы открыть их.

- На панели страницы будут отображаться открытые файлы изображений.

Совет. Удерживайте указатель мыши на эскизах изображений, чтобы вывести сведения о параметрах каждой страницы.

Эскизы на панели страниц можно использовать для редактирования многостраничных документов. Дополнительные сведения см. в разделе **Редактирование многостраничных документов**.



- Задайте параметры распознавания и нажмите кнопку **Распознать и сохранить**, чтобы начать распознавание.

- Чтобы открыть или отсканировать дополнительные изображения для добавления в текущий документ, нажмите

кнопку **Сканировать** или **Открыть** на главной панели инструментов.

Появится запрос о том, следует ли удалять текущий документ. Нажмите кнопку **Да**, чтобы удалить текущий документ, или кнопку **Нет**, чтобы добавить дополнительные отсканированные изображения в текущий документ.

СКАНИРОВАНИЕ И РАСПОЗНАВАНИЕ МНОГОСТРАНИЧНЫХ ДОКУМЕНТОВ

ПО Readiris поддерживает обработку документов, включающих несколько страниц.

Readiris Pro обрабатывает документы до 50 страниц.

ПО Readiris Corporate обрабатывает документы с неограниченным числом страниц.

Для сканирования многостраничных документов в Readiris можно использовать **автоподатчик** (сканер с подачей листов) или **интервальное сканирование** (планшетный сканер).

Сканирование многостраничных документов через автоподатчик (сканер с подачей листов)

- Нажмите кнопку **Сканер** на главной панели инструментов и выберите пункт **Автоподатчик**.
- Поместите страницы в податчик и нажмите кнопку **Сканировать**, чтобы начать сканирование.
- Нажмите **Распознать и сохранить**, чтобы распознать документы.

Сканирование многостраничных документов с помощью интервального сканирования (планшетный сканер)

- Нажмите кнопку **Сканер** на главной панели инструментов.
- Выберите пункт **Сканировать другую страницу после** и задайте временной интервал с помощью кнопок со стрелками.

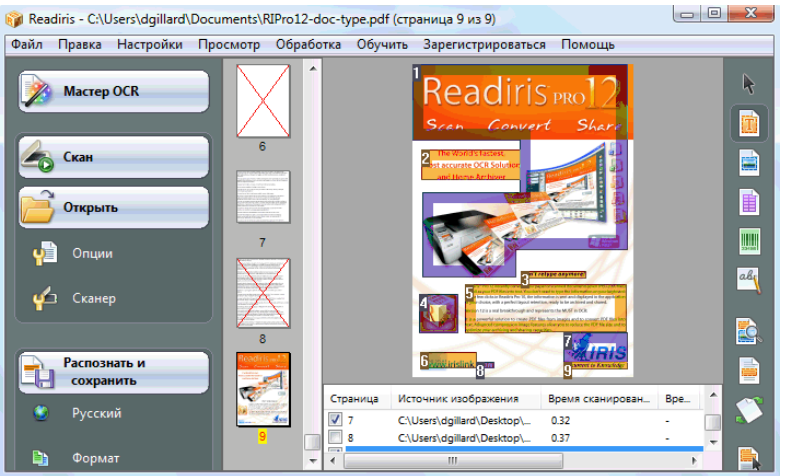
Сканер автоматически отсканирует следующую страницу по истечении указанного числа секунд. При этом пользователю не придется нажимать кнопку **Сканировать** для сканирования каждой страницы.

Чтобы прекратить автоматическое сканирование, нажмите кнопку **Прервать** в диалоговом окне интервального сканирования или клавишу **ESC** на клавиатуре.

Чтобы приостановить автоматическое сканирование, нажмите кнопку **Пауза** в диалоговом окне интервального сканирования или клавишу ПРОБЕЛ на клавиатуре. Для возобновления сканирования нажмите кнопку **Продолжить**.

РЕДАКТИРОВАНИЕ МНОГОСТРАНИЧНЫХ ДОКУМЕНТОВ

При открытии или сканировании нескольких документов в Readiris на панели страниц будут отображаться их эскизы. Эскизы на панели страниц можно использовать для редактирования многостраничных документов.



Перемещение страницы внутри документа:

- Щелкните правой кнопкой страницу, которую необходимо перенести, и нажмите **Выбрать страницу**.
- Перетащите страницу в нужное положение.
- Или щелкните страницу правой кнопкой и нажмите **Переместить страницу вверх** или **вниз**.

Удаление страницы:

- Щелкните правой кнопкой страницу, которую необходимо перенести, и нажмите **Удалить страницу**.
- Или выберите страницу и нажмите кнопку **Delete** на клавиатуре.

Исключение страницы из процесса распознавания:

- Щелкните правой кнопкой страницу, которую необходимо исключить, и нажмите **Исключить страницу**.

- Или очистите поле номера страницы на панели документа.

Изображения исключенных страниц будут перечеркнуты на панели страниц.

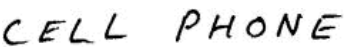
Исключенные страницы игнорируются при печати и сохранении сканированных изображений в многостраничный файл изображения.

Страница	Источник изображения	Время сканирования
☑ 3	C:\Users\dgillard\Desktop\Samples\...	0.92
☐ 4	C:\Users\dgillard\Desktop\Samples\...	0.78

Совет. Команды **Включить все страницы** и **Исключить все страницы** в **меню "Правка"** позволяют включить или исключить все страницы одновременно.

Раздел 13
Распознавание текста, написанного печатными буквами

Помимо печатного текста, таблиц, графики и штрихкодов, Readiris поддерживает распознавание текста, **написанного печатными буквами**. Такой текст состоит из отдельных печатных букв.



Распознавание письма печатными буквами требует крайне специализированного программного обеспечения ICR (интеллектуального распознавания символов).

Для распознавание письма печатными буквами:

- Нажмите кнопку письма печатными буквами на главной панели инструментов.

- Нарисуйте рамку вокруг текста, написанного печатными буквами.
- Нажмите кнопку **Распознать и сохранить** на главной панели инструментов.

Содержимое документа будет распознано полностью, включая текст, написанный печатными буквами.

Примечание. Щелкните область с текстом, написанным печатными буквами, правой кнопкой мыши и выберите пункт **Копировать как текст**, чтобы распознать только эту область и отправить результат в буфер обмена.

Распознаваемые символы

Распознавание письма печатными буквами доступно только для латиницы и поддерживает цифры (0–9), буквы в верхнем регистре (A–Z), а также знаки препинания — запятую, знак "плюс" и дефис.

Диакритические знаки, умляуты и специальные символы не поддерживаются.

Примечание.

- Readiris поддерживает **письмо печатными буквами**, а не рукописное **курсивное письмо**.

  Дополнительные сведения см. в разделе **Правила письма печатными буквами**.

- При распознавании заглавные буквы заменяются строчными за исключением случаев, когда они находятся в начале предложения.
- Такие характеристики документа, как язык, тип шрифта и расстояние между символами, неприменимы к письму печатными буквами.
- Интерактивное обучение также невозможно при распознавании письма печатными буквами. Наша комплексная технология интеллектуального распознавания символов учитывает свыше миллиона образцов письма.

Правила письма печатными буквами

При распознавании письма печатными буквами в Readiris следует учесть несколько правил:

- Символы должны разделяться пробелами и быть единообразными.

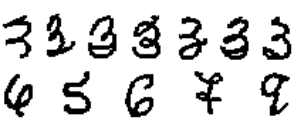
Обратите особое внимание на способ написания букв A, G и Q.

- Используйте шариковую ручку достаточной толщины.

Черные ручки дают лучшие результаты по сравнению с синими.

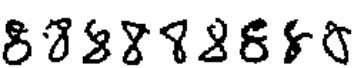
Не пользуйтесь карандашами.

- Постарайтесь отказаться от декоративных элементов



Символы с декоративными элементами увеличивают риск ошибок OCR.

- Не разрывайте петли, которые должны быть замкнуты, и наоборот.



- Не разрывайте символы.

- Не обводите написанные буквы.

Повторная обводка ранее написанных символов снижает качество и четкость письма.

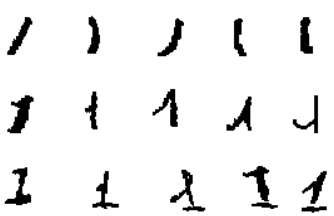
Полностью перечеркнутые символы не будут распознаны.

- Пишите без ошибок.

Цифру "1" можно писать как в "английском”, так и в "европейском” стиле.

Подчеркивание единиц допускается.

При этом горизонтальная черта под цифрой может не соприкасаться с остальными ее элементами.



Совет. Если полученные результаты неоптимальны, воспользуйтесь **трафаретом** I.R.I.S., чтобы адаптировать почерк. Трафарет I.R.I.S. может служить в качестве полностраничного шаблона для письма печатными буквами с единообразными расстоянием между символами и размером. Трафарет можно найти на компакт-диске Readiris и в установочной папке Readiris.

Раздел 14
Распознавание штрихкодов

Введение в считывание штрихкодов

Помимо оптического распознавания символов на 128 языках, Readiris поддерживает считывание штрихкодов.

Поддерживаются все распространенные символики штрихкодов: Codabar, Code 128, Code 39, Code 39 extended, Code 39 HIBC, Code 93, Datalogic 2 of 5, Discrete 2 of 5, EAN-13, EAN-8, Interleaved 2 of 5, MSI Pharmaceutical, MSI-Plessey, патчкод Kodak, PDF-417, PostNet, UCC-128, UPC-A и UPC-E.

Обратите внимание, что для прохождения процедуры OCR в приложении Readiris штрихкод должен быть напечатан на лазерном или струйном принтере. Штрихкоды, напечатанные на матричном принтере, не поддерживаются, так как имеют недостаточную контрастность и их разрешение, как правило, не превышает 60 точек/дюйм.

Ручное считывание штрихкодов

- Определите, какие штрихкоды должны распознаваться приложением Readiris:

- В меню **Параметры** выберите **Штрихкоды**.
- Выберите символики, которые должно распознавать приложение Readiris.
- Определите, должно ли приложение Readiris проверять или удалять контрольные разряды.

- Щелкните штрихкод на панели изображений и нарисуйте рамку вокруг областей штрихкодов в документе.

- Нажмите кнопку **Распознать и сохранить** на главной панели инструментов.

Содержимое документа будет распознано полностью, включая штрихкод.

Примечание. Щелкните область штрихкода правой кнопкой и нажмите кнопку **Копировать как данные**, чтобы скопировать ее содержимое в буфер обмена.

Указатель

A

Asian Edition 5, 8, 56

M

Middle-East edition.....5, 8, 56

S

SmartTasks 18, 21

А

автоматическое формирование окон 45

автоподатчик 93

Азиатские документы .5, 8, 56

анализ страницы..... 28

Арабские документы ..5, 8, 56

В

восстановление заводских параметров..... 89

восстановление исходного документа 70

Вывод в Excel 67

Вывод в HTML..... 67

Вывод в OpenDocument.....67

Вывод в PDF 67, 77

Вывод в PDF iHQC.....80

Вывод в RTF 67

Вывод в SpreadsheetML.....67

Вывод в Unicode 67

Вывод в Word.....67

Вывод в WordML.....67

Вывод в XML.....67

Вывод в XPS 67, 82

Вывод в XPS iHQC.....84

выравнивание.....27

выравнивание страницы 27

Г

главная панель инструментов 18

Д

Документы на иврите . 5, 8, 56

З

заводские параметры 89

загрузка параметров..... 89

И

изменение языка пользовательского интерфейса 20

изображение в оттенках серого 32

инвертированные изображения 33

интерактивное обучение ... 61

интервальное сканирование 94

К

каталоги шрифтов 63

контраст 32, 40

Л

листы электронных таблиц 73

М

Мастер..... 15

матричный принтер 60

многостраничные документы 91, 93

Н

наклон строки 35

настройка сканированных документов.....37

О

отправка документов по электронной почте.....68

П

панель документа 19

панель изображений.....19

параметры 27

параметры PDF 78

Параметры XPS 84

параметры графики 86

параметры макета 70

параметры сканера 30

перепрофилирование PDF-документов.....81

письмо печатными буквами 97, 99

поворот 27

поддерживаемые форматы изображений 30

поддержка продукта.....14

пользовательский интерфейс 17

Р

разрешение 32

расстояние между символами 60

регистрация 13

редактирование многостраничных документов 94

ретушь..... 41

С

свойства документа..... 67

сглаживание цветных изображений..... 35, 37

Сжатие JPEG 2000 87

соотношение скорости и точности..... 58

сохранение параметров 89

страницы..... 94

выбор..... 94

исключение..... 94

перемещение..... 94

удаление..... 94

Т

таблицы 73

текстовые документы.....69

тип шрифта 59

У

удаление Readiris 13

установка.....11

Ф

файл параметров.....89

файлы макета 52

фоновый цвет.....71

фоновый цвет ячеек таблицы 75

форматы бумаги.....73

форматы вывода 67

формирование окон вручную 47

Х

характеристики документа 59

Ц

цветное изображение ...32, 37

цветовой режим 32

цифровая фотокамера.....33

цифровой 57

Ч

черно-белое изображение.. 32

Ш

шаблоны окон..... 52

штрихкоды..... 101

Э

электронные таблицы.....73

Я

язык.....56

язык пользовательского интерфейса 20

яркость.....32, 38