

САМОУЧИТЕЛЬ по работе с программой **ABBYY® FineReader 10**

Работать с «электронными» документами во многом удобнее и проще, чем с их «бумажными» аналогами. Электронный документ можно редактировать, использовать при создании собственных работ, его легко копировать и пересылать по электронной почте. Вместе с тем, многие материалы изначально доступны нам в нередактируемом виде (бумажные или отсканированные документы, цифровые фотографии). Программа ABBYY® FineReader – лучший инструмент для создания «электронных копий» любых печатных материалов: книг, справочников, журналов, договоров, бланков.

Книга включает описание приемов сканирования и распознавания разных оригиналов – от простых книжных страниц до сложно оформленных документов. А приведенные скриншоты программы позволят читателю быстро освоить интерфейс ABBYY® FineReader и получить практические навыки по работе с программой.

Изложение материала сопровождается практическими примерами. Читатели, которые еще не пробовали самостоятельно переводить печатные материалы в электронный вид, найдут в этой книге простое пошаговое руководство. Для тех же, кто хочет в совершенстве освоить работу с программой, книга откроет многочисленные тонкости настройки для эффективного использования ABBYY® FineReader.



*DVD содержит ознакомительные версии программ компании ABBYY®.



Internet-магазин: www.aliants-kniga.ru

Книга – почтой:

Россия, 123242, Москва, а/я 20

e-mail: book@aliants-kniga.ru

Оптовая продажа: «Альянс-книга»

Тел./факс: (495) 258-9195

e-mail: books@aliants-kniga.ru



978-5-94074-595-2



9 785940 745952

ABBYY® FineReader 10. Сканирование и распознавание текстов. Самоучитель

Жадаев А. Г.

Сканирование и распознавание текстов

Коллекция
программ ABBYY®
в подарок*



САМОУЧИТЕЛЬ
по работе с программой

ABBYY® FineReader 10



Жадаев А. Г.

Сканирование и распознавание текстов

Самоучитель по работе с ABBYY® FineReader 10



Москва, 2010

УДК 32.973.26-018.2
ББК 004.4
Ж15

Ж15 Жадаев А. Г.

Сканирование и распознавание текстов. Самоучитель по работе с ABBYY® FineReader 10. – М.: ДМК Пресс, 2010. – 248 с.: ил.

ISBN 978-5-94074-595-2

Работать с электронными документами во многом удобнее и проще, чем с их бумажными аналогами. Электронный документ можно редактировать, использовать при создании собственных работ, его легко копировать и пересылать по электронной почте. Вместе с тем, многие материалы изначально доступны нам в нередатируемом виде (бумажные или отсканированные документы, цифровые фотографии). Программа ABBYY® FineReader – лучший инструмент для создания электронных копий любых печатных материалов: книг, справочников, журналов, договоров, бланков.

Книга включает описание приемов сканирования и распознавания разных оригиналов – от простых книжных страниц до сложно оформленных документов. А приведенные скриншоты программы позволят читателю быстро освоить интерфейс ABBYY® FineReader и получить практические навыки по работе с программой.

Изложение материала сопровождается практическими примерами. Читатели, которые еще не пробовали самостоятельно переводить печатные материалы в электронный вид, найдут в этой книге простое пошаговое руководство. Для тех же, кто хочет в совершенстве освоить работу с программой, книга откроет многочисленные тонкости настройки для эффективного использования ABBYY® FineReader.

УДК 32.973.26-018.2
ББК 004.4

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

Материал, изложенный в данной книге, многократно проверен. Но поскольку вероятность технических ошибок все равно существует, издательство не может гарантировать абсолютную точность и правильность приводимых сведений. В связи с этим издательство не несет ответственности за возможные ошибки, связанные с использованием книги.

ISBN 978-5-94074-595-2

© Жадаев А. Г., 2010

© Оформление, издание, ДМК Пресс, 2010

Содержание

Глава 1

Текст и графика в компьютере	7
Форматы файлов	9
Текстовые форматы	10
Графические файлы	11
Составные (сложные) документы	12
Оптическое распознавание символов (OCR)	14
От чего зависит качество распознавания?	16
Системные требования	17
Резюме	17

Глава 2

Быстрый старт	19
Сканирование в MS Word, MS Excel, PDF	23
Конвертирование изображений и PDF в документ Microsoft Word	28
Вызов сценариев из контекстного меню файла	30
Сканирование и сохранение изображений	31
Резюме	33

Глава 3

Работа в пошаговом режиме	34
Окно программы и настройка рабочего пространства	34
Рабочие окна	35
Окно Страницы	36
Окно Изображение	38
Окно Текст	39
Окно Крупный план	40
Изменение расположения рабочих окон	41
Настройка панелей инструментов	45
Диалоговое окно Опции	51
Документ FineReader	55
Резюме	57

Глава 4

Получение изображений	59
Работа со сканером	60
Параметры сканеров	60
Драйвер и настройки сканера	64
Разрешение	66
Режим сканирования	67
Яркость	67
Параметры страницы	68
Сканирование многостраничных документов	69
Работа с цифровой камерой	69
Параметры цифровых камер	69
Техника съемки	73
Расстояние	73
Освещение	73
Баланс белого	74
Повышение четкости изображения	74
Работа над ошибками	75
Частные случаи	77
Крупноформатные оригиналы	78
Книги	79
Резюме	81

Глава 5

Обработка и анализ изображений	83
Обработка изображения	84
Настройка автоматической обработки	85
Обработка в Редакторе изображений	86
Анализ изображений	92
Области изображения	92
Исправление разбивки на области	94
Свойства области	99
Использование шаблонов областей	101
Анализ таблиц	104
Резюме	107

Глава 6

Распознавание текстов	109
Применение пользовательского эталона	109
Общие правила работы с пользовательскими эталонами	109
Пример обучения и использования эталона	111
Редактирование пользовательских эталонов	117
Распознавание многоязычных документов	120
Пример распознавания двуязычного документа	120
Выбор языка для распознавания документа	123
Создание группы языков	124

Создание пользовательского языка	128
Пример распознавания текста с помощью регулярных выражений	131
Использование словарей	135
Общие правила работы со словарями	135
Пример редактирования и применения пользовательского словаря	139
Резюме	141

Глава 7

Проверка и корректировка распознанного документа 143

Проверка и корректировка документа в программе FineReader	143
Пример корректировки документа в окне Текст	143
Пример проверки и корректировки текста в диалоге Проверка	151
Использование стилей	158
Общие сведения о стилях в FineReader	159
Пример создания и применения пользовательского стиля	159
Окончательная обработка распознанного документа в программе Microsoft Word	162
Пример избавления импортированного в Word документа от стилей FineReader	162
Пример обработки распознанного документа в Word	168
Корректировка таблиц	173
Настройка панели инструментов для работы с таблицами	173
Пример корректировки таблицы	175
Резюме	177

Глава 8

Сохранение распознанного документа 178

Передача документа в приложение	180
Пример передачи распознанного документа в Microsoft Word	180
Пример передачи распознанного документа в Microsoft Excel	181
Пример передачи распознанного документа в Adobe Reader	182
Пример передачи распознанного документа в веб-браузер	183
Сохранение документа в файл	184
Настройка параметров сохранения	184
Примеры сохранения распознанного документа в формате Word	194
Пример сохранения распознанного документа в формате PDF	198
Пример сохранения распознанного документа в формате HTML	202
Пример сохранения распознанного документа в формате TXT	205
Пример сохранения распознанного документа в формате Excel	205
Как сохранить документ FineReader	208
Резюме	210

Глава 9

Сценарии 211

Создание пользовательского сценария	211
Пример: распознавание платежного поручения	212

Другие действия сценариев	222
Менеджер сценариев	225
Копирование сценария	225
Изменение сценария	226
Удаление сценария	226
Экспорт и импорт сценариев	226
Использование сценариев	228
Резюме	230

Глава 10

ABBYY Screenshot Reader	231
Интерфейс и настройки программы	231
Работа с программой	233
Захват изображения	234
Передача в буфер обмена	235
Передача в приложения Microsoft Office	237
Операция Изображение в ABBYY® FineReader	237
Сохранение в файл	238
Примеры использования программы	241
Копирование текста документов с экрана	241
Список файлов	242
Снимки интерфейса	245
Резюме	246