

Wichtiger Hinweis!

Diese Software (einschließlich der Datenträger und gedruckten Dokumentation) wird Ihnen ausschließlich aufgrund der Bedingungen des Lizenzvertrags überlassen, aber nicht an Sie verkauft.

Lesen Sie den Lizenzvertrag aufmerksam durch, bevor Sie die Verpackung öffnen. Durch das Öffnen der Verkaufsverpackung erkennen Sie den Inhalt des Lizenzvertrags an.

Wenn Sie mit den Bedingungen des Lizenzvertrags nicht einverstanden sind, geben Sie innerhalb von 7 Tagen das ungeöffnete Paket (einschließlich der Datenträger und gedruckten Dokumentation) bei dem Händler zurück, von dem Sie es erworben haben.

Lizenzvertrag für Software

Dieser Lizenzvertrag ist die rechtliche Übereinkunft zwischen Ihnen (juristische oder natürliche Person) und ABBYY.

Das beiliegende Dokument »Nachweis des rechtmäßigen Erwerbs« ist ein wesentlicher Bestandteil dieses Lizenzvertrags.

Die Software, die ABBYY vertreibt, besteht aus Programmen, Datenbanken und Dokumentation, die urheberrechtlich geschützt sind. Das Wort »Dokumentation« bezeichnet alle gedruckten Materialien und Textdateien, die die Software beschreiben. Die Dokumentation ist integraler Bestandteil der Software.

1 Gegenstand des Vertrags

- 1.1 Der Gegenstand dieses Lizenzvertrags ist die Übertragung der nicht ausschließlichen Lizenz zur Nutzung der Software an den Endbenutzer durch ABBYY.
- 1.2 Alle nachfolgenden Punkte beziehen sich sowohl auf die Software als Ganzes als auch auf die einzelnen Komponenten der Software.

2 Urheberrecht

- 2.1 Das ausschließliche Urheberrecht an dieser Software hat ABBYY.
- 2.2 Sie erhalten die Lizenz zur Benutzung der Software, sofern Sie die folgenden Voraussetzungen und Bedingungen akzeptieren.
- 2.3 Diese Lizenz gilt ausschließlich für Sie. (Sofern keine gesonderte, schriftliche Erlaubnis von ABBYY vorliegt.)

3 Bedingungen zum Einsatz der Software

- 3.1 Sie dürfen die Software zu keinem Zeitpunkt auf mehr Computern benutzen, als auf dem »Nachweis des rechtmäßigen Erwerbs« unter »Anzahl der Lizenzen« angegeben ist. Sie dürfen nicht mehr Kopien des Programms besitzen (unabhängig davon ob diese Kopien installiert sind oder nicht), als Sie Lizenzen haben.
- 3.2 Sie können eine Sicherungskopie der Software erstellen, für den Fall des Verlustes oder der Beschädigung der Originale. Diese Sicherungskopie darf zu keinem anderen Zweck verwendet werden.
- 3.3 Sie können eigene Datenbanken für Programme erstellen, die Teil der Software sind, wenn diese Möglichkeit in der Dokumentation vorgesehen ist.
- 3.4 Wenn Sie die Software auf mehr als einem Medium erhalten (unabhängig von Größe und Typ des Mediums), dürfen Sie ausschließlich das Medium verwenden, das für Ihren Computer geeignet ist. Beide Medien enthalten dieselbe Kopie der Software.
- 3.5 Sie dürfen die Software nicht vertreiben. Vertreiben bedeutet Dritten Zugang zu Teilen der Software zu gewähren. Dabei ist es bedeutungslos, wer die Kopie erstellt und ob diese per Netzwerk oder auf anderem Wege verliehen, verkauft oder vermietet wird.
- 3.6 Es gelten die folgenden Einschränkungen:

3.6.1 Zu keiner Zeit darf die Software auf mehr als einem Computer installiert sein. Wird die Software im Netzwerk installiert, dann darf zu keinem Zeitpunkt mehr als eine Arbeitsstation die Software verwenden.

3.6.2 Sie dürfen Programme, Datenbanken und andere Komponenten der Software nicht dekompile und disassemblieren.

3.6.3 Sie dürfen den Objektcode der Programme und die Datenbanken nicht ändern. Ausgenommen hiervon sind die Stellen in der Software, wo dies explizit vorgesehen und der in der Dokumentation beschrieben ist.

3.6.4 Sie können Dritten keinerlei Rechte an der Software übertragen. Dies gilt auch für das Nutzungsrecht, das Ihnen dieser Vertrag gewährt.

3.6.5 Sie dürfen keine Handlungen vornehmen, die russische oder internationale Urheberrechts- und Softwareschutzgesetze verletzen.

4. Gültigkeitsdauer des Vertrags

4.1 Dieser Lizenzvertrag tritt mit der Installation auf einem Computer in Kraft und bleibt wirksam, solange die Software benutzt wird.

4.2 Wenn Sie die Voraussetzung und Bedingungen des Lizenzvertrags nicht erfüllen können, müssen Sie alle Kopien der Software (einschließlich der gedruckten Materialien, Datenträgern, Informationsdateien und Archivkopien der Software) vernichten. Damit wird der Lizenzvertrag aufgelöst.

5. Verantwortung

5.1 Unberechtigte Reproduktion, Verwendung oder nicht autorisierter Vertrieb der Software ist stellt eine Verletzung des Gesetzes der Russischen Föderation »Zum Schutz der Software« dar und wird gerichtlich verfolgt.

5.2 Im Falle der Verletzung des Lizenzvertrags entzieht ABBYY Ihnen mit sofortiger Wirkung das Recht zur Verwendung der Software. Gleichzeitig erlischt die Gewährleistungspflicht von ABBYY und der Anspruch auf technische Unterstützung und kostenlose Lieferung der Software.

6. Garantie

6.1 ABBYY garantiert die Qualität der Daten auf den Medien, die Arbeitsfähigkeit der enthaltenen Programme im Rahmen der in der Dokumentation beschriebenen Einsatzbedingungen, die Übereinstimmung der Komponenten der Software mit den Spezifikationen, und die Druckqualität der Dokumentation.

6.2 Was andere Komponenten betrifft, wird die Software geliefert, »wie sie ist«. ABBYY garantiert nicht, daß die Software fehlerfrei ist. Weiter übernimmt ABBYY keine Verantwortung für direkte oder indirekte Schäden, die durch den Fehler in der Software oder der Dokumentation entstehen können.

6.3 ABBYY übernimmt keinerlei Garantie für die Funktionsfähigkeit der Software, wenn sie zusammen mit anderer installierter Soft- oder Hardware betrieben wird, insbesondere wenn es sich dabei um ältere Soft- oder Hardware handelt.

6.4 Die Gewährleistung gilt 60 Tage ab dem Kauf der Software. In dieser Zeit akzeptiert ABBYY alle Forderungen, die aufgrund der Qualität der Software entstehen.

Mit FineReader haben Sie eine gute Wahl getroffen!

FineReader ist ein leistungsfähiges Werkzeug, das die Verarbeitung von Dokumenten beschleunigt. Wir haben viele Jahre Entwicklungsarbeit in dieses Programm investiert und freuen uns, wenn es Ihnen das Leben erleichtert!

Das Entwicklungsteam von ABBYY (BIT Software)

Programm zur optischen Texterkennung (OCR)

FineReader 4.0

Benutzerhandbuch

ABBYY (BIT Software), Moskau 1999

ABBYY (BIT Software)

FineReader 4.0

Benutzerhandbuch

Die Informationen in diesem können jederzeit und ohne Vorankündigung geändert werden. Sie stellen keine Verpflichtung seitens ABBYY (BIT Software) dar.

Die in diesem Dokument beschriebene Software wird Ihnen unter den Bedingungen des Lizenzvertrags übergeben. Es verstößt gegen das Gesetz zum Schutz von Software und Datenbanken der Russischen Föderation («On legal protection of software and databases») und gegen die internationalen Urheberrechtsgesetze, Software auf eine Medium zu kopieren, außer soweit dies explizit im Lizenzvertrag oder in der Vertraulichkeitsvereinbarung gestattet ist.

Kein Teil dieses Handbuchs darf, unabhängig vom Zweck, ohne vorherige schriftliche Genehmigung durch ABBYY in irgendeiner Form (elektronisch oder mechanisch) reproduziert, vervielfältigt oder verbreitet werden.

© ABBYY (BIT Software), 1993-98. Alle Rechte vorbehalten

ABBYY, BIT Software, FineReader, »fontain image transformation«, Lingvo, Scan&Read, Scan&Translate, »one button principle«, »Your computer reads by itself«, »Your computer reads and translates by itself« sind eingetragene Warenzeichen von ABBYY. Try&Buy und DOCFLOW sind Warenzeichen von ABBYY. Alle anderen Marken sind Warenzeichen oder eingetragene Warenzeichen ihrer Eigentümer.

105568 Moskau, box 19. ABBYY

Inhalt

Kapitel 1. Schnelleinsteig	11
Arbeitsschritte zum Einlesen von Text und Tabellen.....	11
Scannen und Erkennen eines gedruckten Dokuments	11
Erkennen einer bereits gespeicherten Bilddatei.....	14
Erkennen von Formularen (nur in FineReader Professional und Enterprise verfügbar)	15
Kapitel 2. Installation und Konfiguration	23
Software- und Hardwareanforderungen	23
Installation von FineReader	23
Installationsprobleme	24
FineReader starten	24
Arbeiten mit dem Scanner	24
Ratschläge zum Scannerkauf	26
Wenn Ihr Scanner nicht mit FineReader arbeitet	26
Wie geht es weiter?	27
Kapitel 3. FineReader 4.0 Überblick	29
FineReader 4.0 Standard Eigenschaften.....	29
FineReader 4.0 Professional Eigenschaften	31
FineReader 4.0 Enterprise Eigenschaften.....	31
Kapitel 4. Arbeiten mit FineReader 4.0.....	32
Das Hauptfenster.....	32
Die Scan&Read Werkzeugleiste	33
Die Standardwerkzeugeleiste.....	34

Die Struktur eines Stapels	34
Das Stapelfenster	35
Numerierung der Stapelseiten	37
Arbeiten mit Stapelseiten	37
Scannen	38
Gute und schlechte Texte	38
Scaneinstellungen für gute Ergebnisse	39
Mehrseitige Dokumente scannen	40
Bilder zu einem Stapel hinzufügen	42
Was geschieht mit Farb- und Graustufenbildern, wenn Sie in FineReader geöffnet werden?	42
Seitenanalyse (Blöcke markieren)	43
Tabellenstrukturen festlegen	44
Tabellenblöcke bearbeiten	45
Erkennung	46
Sprachen für die Erkennung	47
Texttyp	48
Speichern des Seitenlayouts und der Formatierung	48
Erkennung starten	49
Erkennung im Hintergrund	50
Prüfen und Bearbeiten der Ergebnisse	50
Speichern der Erkennungsergebnisse	52
Stapel zur späteren Bearbeitung speichern	53
Trainieren neuer Zeichen	54
Trainieren des Programms	54
Sprachen in FineReader	57
Benutzerdefinierte Sprachen und Sprachgruppen (für FineReader 4.0 Professional und Enterprise)	58
Stapelverarbeitung im Netzwerk (für FineReader 4.0 Professional und Enterprise)	58
Wie wird ein Stapel im Netzwerk verarbeitet?	59
Read&Wait Modus	59
Anhang	57
Verbessern der Erkennungsqualität: Tips und Tricks	57

Grundzüge.....	57
Das Wörterbuch erweitern.....	57
Training.....	57
Erzeugen neuer Sprachen	58
Beispiel zur Erkennung der altrussischen Sprache	58
Für FineReader Enterprise Administratoren	60
Von FineReader unterstützte Bildformate.....	60
Unterstützte Sprachen	60
Wörterbücher und Benutzeralphabete von FineReader 3.0 verwenden	61
Tastenkombinationen (Shortcuts) in FineReader	62

Wir alle müssen Text in unsere Computer eingeben. Der Text kann in Zeitungen, Zeitschriften, Verträgen, Briefen, Faxen, Preislisten, Fragebögen usw. stehen. Etwa 90% aller Dokumente erhalten wir in gedruckter Form auf Papier. Jahrelange gab es nur einen Weg, diese Daten in den Computer zu übernehmen: durch Eingabe über die Tastatur. Erinnern Sie sich noch an die Stunden, die Sie damit zugebracht haben, ein Dokument in den Computer einzutippen? Wie wundervoll wäre es doch, wenn der Computer den Text direkt von einem Blatt Papier einlesen könnte!

Manchmal werden Träume wahr – Das optische Texterkennungsprogramm **FineReader** (Optical Character Recognition, OCR) ermöglicht Ihrem Computer und Scanner, gedruckten Text einzulesen.

Kann der Scanner diese Aufgabe nicht alleine erledigen?

Nein. Der Scanner erzeugt nur ein Abbild des Texts, ähnlich einer Fotografie. Dieses Abbild besteht aus schwarzen und weißen Punkten oder in anderen Worten, es ist eine Bilddatei, die mit einem Texteditor, wie z.B. MS Word, WordPerfect, Word Pro usw. nicht bearbeitet werden kann.

Sie benötigen ein OCR-Programm, das in den schwarzen und weißen Punkten Symbole sucht, diese Symbole als Zeichen »erkennt« und so das Abbild in einen editierbaren Text umwandelt. Solch ein Programm erzeugt eine Datei, die mit einem Texteditor bearbeitet werden kann.

Also, kann ich meine Dokumente jetzt automatisch in meinen Computer einlesen?

Ja, mit FineReader können Sie Ihre Dokumente automatisch in Ihren Computer übernehmen, ohne sie per Tastatur einzugeben – Viel Vergnügen!

K a p i t e l 1

Schnelleinstieg

FineReader ist ein OCR-Programm, das entwickelt wurde, um Dokumente automatisch in den Computer zu übernehmen. Das Arbeitsprinzip ist einfach: Sie legen eine Textseite in den Scanner, der Scanner erzeugt ein Abbild des Texts, das FineReader erkennt und als Textdatei im Texteditor anzeigt. In nur 30–60 Sekunden haben Sie den gewünschten Text in Ihrem Texteditor, inklusive Seitenlayout, Farbbildern und Tabellen.

Arbeitsschritte zum Einlesen von Text und Tabellen

Sie haben FineReader auf Ihrem Computer installiert und einen Scanner angeschlossen. (Eine ausführliche Anleitung zur Installation finden Sie im Kapitel »Installation und Konfiguration«, Seite 23.)

Jetzt können Sie Dokumente in Ihren Computer einlesen. (Sofern FineReader bereits gestartet ist, fahren Sie bitte mit Schritt 4 fort.)

Scannen und Erkennen eines gedruckten Dokuments

1. Schalten Sie den Scanner ein.
2. Schalten Sie den Computer ein.
3. Starten Sie FineReader. Wählen Sie dazu im *Startmenü* den Eintrag *Programme>ABBYY FineReader>FineReader 4.0*.
Es erscheint das Hauptfenster von FineReader (siehe Abbildung 1) und der Dialog *Tips*. Dieser Dialog enthält hilfreiche Hinweise.
Anschließend erscheint der *Willkommen* Dialog.
4. Legen Sie das »Erkennungsbeispiel« (Sie finden das Beispieldokument am Ende dieses Handbuchs) in Ihren Scanner ein.
5. Klicken Sie auf die Schaltfläche *Scan&Read* in der linken oberen Ecke des Fensters.
6. Befolgen Sie die Anweisungen des Scan&Read Assistenten. In wenigen Minuten lernen Sie, ein gedrucktes Dokument in eine Datei umzuwandeln. Das Einlesen eines Dokuments ist einfach und besteht aus den folgenden drei Schritten: Scannen, Blöcke markieren und Erkennen.

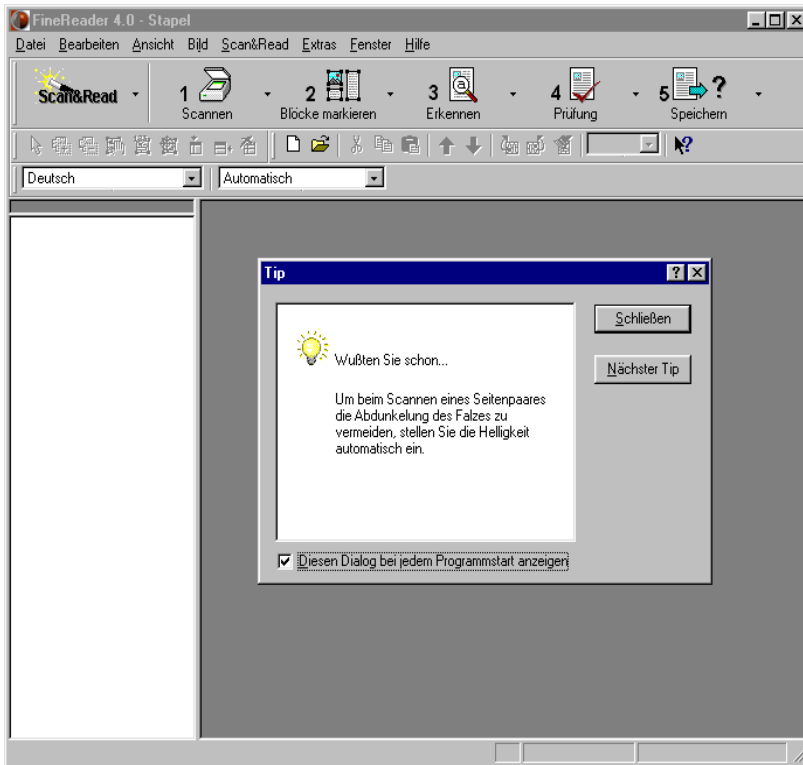


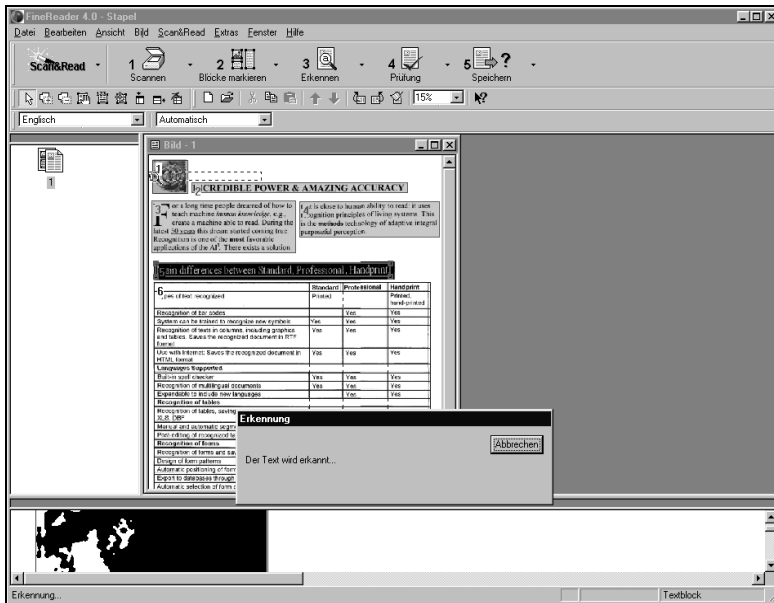
Abbildung 1. Das Hauptfenster von FineReader 4.0.

Bei **Flachbett-Scannern** hören Sie ein Geräusch und ein Licht bewegt sich unter der Scannerabdeckung.

Bei **Einzugscannern** wird die Vorlage wie bei einem Faxgerät eingezogen.

Bei **Handscannern** müssen Sie den Scanner über das Papier bewegen. Drücken Sie die Starttaste und bewegen Sie den Scanner über die Seite. Drücken Sie eine beliebige Taste der Tastatur, um den Scanvorgang zu beenden. (Diese Hinweise gelten für die meisten Scanner. Lesen Sie im Zweifelsfall die Dokumentation Ihres Scanners.)

*Falls der Dialog Scannerparameter erscheint, stellen Sie folgende Werte ein:
Auflösung – 300 dpi, Helligkeit – Normal oder Mittel, Modus - Schwarzweiß; und*



7. Nach dem Scannen erscheint das Fenster *Bild*, das ein Abbild der Seite enthält. Das Programm beginnt automatisch die Analyse des Abbilds. Dabei sucht es Text, Abbildung und Tabellen. Nach Ende der Analyse erkennt es den Text und Tabellen. Während der Erkennung markiert FineReader den gerade bearbeiteten Teil der Seite.

Das ist alles! Das Arbeitsergebnis von FineReader erscheint im Fenster »Text« (Abbildung 3, nächste Seite).

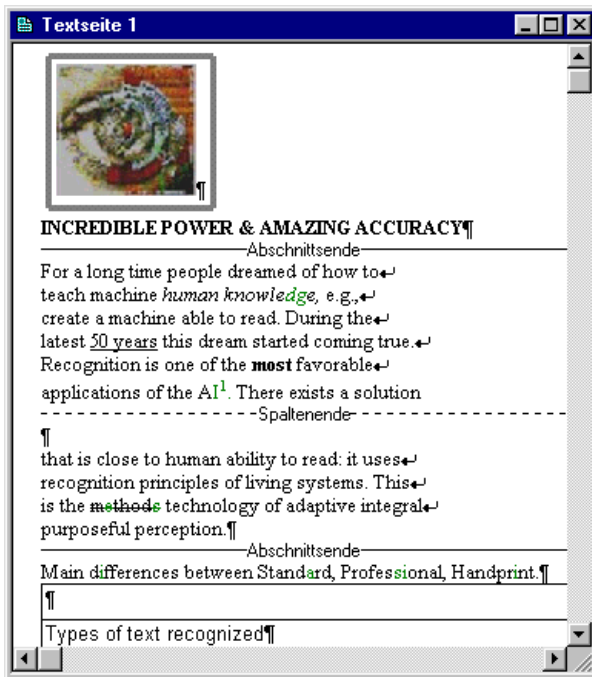


Abbildung 3. Das Textfenster.

Das Fenster *Text* ist der in FineReader integrierte Texteditor. In diesem Fenster können Sie den erkannten Text bearbeiten und das Erkennungsergebnis prüfen (Schaltfläche *Prüfung*).

8. Sie können den erkannten Text an MS Word, Excel usw. senden. Klicken Sie dazu auf den kleinen Abwärtspfeil rechts neben der Schaltfläche *Speichern* und wählen Sie im erscheinenden Menü den Eintrag *Senden an...* (MS Word, MS Excel usw.) Sie können das Erkennungsergebnis auch als Datei speichern. Klicken Sie dazu auf den Pfeil rechts neben der Schaltfläche *Speichern* und wählen Sie im erscheinenden Menü den Eintrag *Als Datei speichern*. Das Programm schlägt Ihnen vor, das Erkennungsergebnis im DOC Format zu speichern. Dieses Format erlaubt das Speichern von Text inklusive Seitenlayout, Abbildungen und Tabellen.

Erkennen einer bereits gespeicherten Bilddatei

Schalten Sie Ihren Computer ein.

1. Starten Sie FineReader. Wählen Sie dazu im Menü *Start* den Eintrag *Programme> ABBYY FineReader>FineReader 4.0*.
2. Klicken Sie anschließend auf den Pfeil rechts neben der Schaltfläche *Scan&Read* und wählen Sie im erscheinenden Menü den Eintrag *Open&Read*.
3. Wählen Sie das Laufwerk und den Ordner, der die gewünschte Bilddatei enthält im Dialog *Öffnen* und öffnen Sie diese Datei. Die Datei kann die Erweiterung TIF, PCX, BMP, DCX oder JPG haben (eine vollständige Liste der von FineReader unterstützten Dateiformate finden Sie im Anhang auf Seite 60).
4. Öffnen Sie zum Beispiel die Datei *demo.tif* (im Ordner \Programme\ABBYY FineReader\Demo). Im Fenster *Bild* erscheint ein Abbild dieser Seite. Das Programm beginnt mit der Erkennung und nach einigen Sekunden erscheint das Ergebnis im Fenster *Text*.

Jetzt können Sie das Ergebnis an ein anderes Programm senden oder als Datei speichern. Wie das geht, ist unter Punkt 8 im vorherigen Abschnitt beschrieben.

Erkennen von Formularen (nur in FineReader Professional und Enterprise verfügbar)

Was versteht man unter der Erkennung von Formularen?

Formularerkennung bedeutet die Erkennung von standardisierten Papierbögen, wie z.B. Fragebögen, Zahlungsbelege oder Kurzmittlungszettel. Formulare sind also Dokumente, die eine gleichbleibende Feldanordnung und Papiergröße haben.

Was soll hier erreicht werden? Sie haben 50, 100 oder 10.000 Formulare. Alle enthaltenen *Felder* mit Informationen, z.B. »Nachname«, »Vorname«, »Betrag in Worten«, »Betrag in Ziffern« und so weiter. Sie möchten diese Informationen in eine Datenbank übernehmen. Dazu muß jeder Wert eines *Feldes* aus jedem Formular in das entsprechende *Feld* der Datenbank eingelesen werden. Auf diese Weise werden alle Eintragungen von festgelegten Formularpositionen in gleiche Datenbankfelder übernommen, und die Reihenfolge der Formulare wird beibehalten.

Im folgenden wird erklärt, wie man Formulare einliest. Dazu werden die Beispielformulare »Job Application« (Bewerbung), die Sie am Ende dieses Handbuchs finden, verwendet.

Diese Formulare wurden mit Schreibmaschine ausgefüllt. In der Praxis werden Formulare auch von Hand ausgefüllt. Für die Erkennung von handschriftlich ausgefüllten Formularen benötigen Sie **FineReader Enterprise**.

Einen vorbereiteten Formularstapel finden Sie im Ordner \Programme\ABBYY FineReader\Forms\Demo. Die Formularschablone hat den Dateinamen *DEMO.FRM*. Eine Beschreibung zu dieser Datei finden Sie in diesem Kapitel.

Einlesen eines Formularstapels

Bereiten Sie ein Original des maschinenlesbaren Formulars vor (Sie finden drei Formulare am Ende dieses Handbuchs, eines davon ist nicht ausgefüllt). Das unausgefüllte Formular dient als Original.

1. Kopieren Sie dieses Original oder drucken Sie es mit einem Laserdrucker aus. Nehmen Sie keine Veränderungen an den Kopien vor und stellen Sie sicher, daß sich die Felder auf allen Kopien an exakt denselben Positionen wie auf dem Original befinden.
2. Füllen Sie die Formulare aus.
3. Starten Sie FineReader 4.0 Forms, indem Sie im Menü *Start* den Eintrag *Programme>ABBYY FineReader>FineReader 4.0 Forms* wählen.

Es erscheint der Dialog *Formularstapel*.



Abbildung 4. Der Dialog *Formularstapel*.

4. Wählen Sie *Einen neuen Formularstapel erstellen* und klicken Sie auf OK.
5. Geben Sie im Dialog *Name für neuen Stapel* (Abbildung 5) den gewünschten Namen und den Pfad zu den Stapeldateien an. Klicken Sie dann auf *Weiter*.



Abbildung 5. Der Dialog *Name für neuen Stapel*



Abbildung 6. Der Dialog *Formularschablone*

- Wählen Sie im Dialog *Formularschablone* (Abbildung 6) die Option *Eine neue Schablone erstellen*, klicken Sie dann auf *Weiter*. Es erscheint der Dialog *Schablonenvorlage*:



Abbildung 7. Geben Sie den Typ der Schablonenvorlage an.

7. Wählen Sie im Dialog *Schablonenvorlage* die Option *Scannen*. Jetzt benötigen Sie das nicht ausgefüllte Formular. Legen Sie diese Seite in den Scanner und klicken Sie auf *Weiter*.

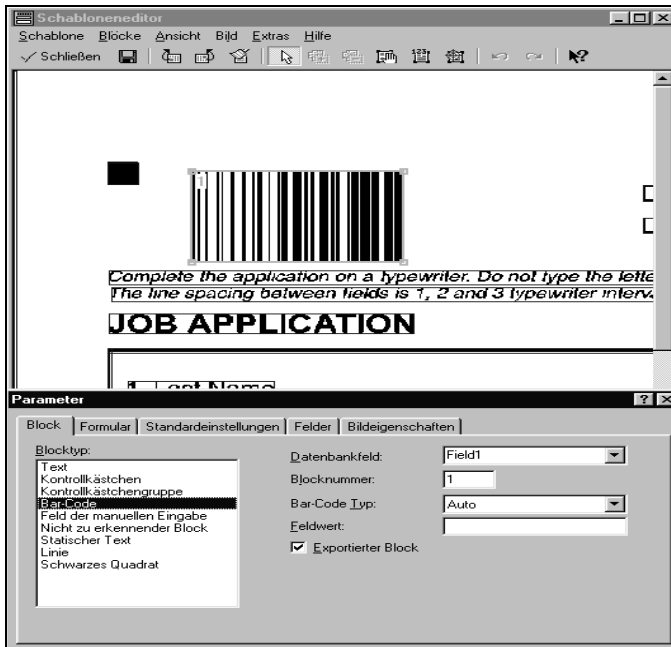


Abbildung 8. Eine Formelschablone im Dialog Schabloneneditor.

Nach dem Scanvorgang sehen Sie das gescannte Bild des Originalformulars im Dialog *Schabloneneditor*. Der Dialog *Parameter* wird ebenfalls angezeigt.

Erstellen der Formelschablone

1. Wählen Sie im Menü *Schablone* den Eintrag *Schablone erstellen*. Die schwarzen Rechtecke an den Ecken des Dokuments werden als Blocktyp *Schwarze Rechtecke* markiert. Alle übrigen Objekte werden als *Statischer Text* markiert.

Wichtiger Hinweis! Löschen Sie den statischen Text und die schwarzen Rechtecke nicht. Sie werden zur Auswahl und Ausrichtung der Formelschablone benötigt.

2. Nehmen Sie auf der Seite *Standardeinstellungen* des Dialogs *Parameter* (Abbildung 9) folgende Einstellungen vor: Sprache: Englisch, Texttyp: Schreibmaschine, Block säubern: aktiviert.



Abbildung 9. Seite Standardeinstellungen des Dialogs Parameter.

3. Markieren Sie jetzt die Textfelder, die erkannt werden sollen. Eines dieser Felder ist in diesem Beispiel das Feld »Last name«. Zeigen Sie mit der Maus auf die linke obere Ecke des Felds, drücken Sie die Maustaste und halten Sie sie gedrückt, während Sie die Maus zur rechten unteren Ecke des Felds bewegen. Lassen Sie dort die Maustaste los. Im Bild erscheint ein Block (ein grünes Rechteck), der das Textfeld markiert.

Hinweis: Alles innerhalb eines Blocks mit grünem Rahmen wird erkannt. Deshalb sollten Sie den Feldnamen und Begrenzungslinien nicht im Block einschließen. Im Beispiel sollte der Text »Last Name« nicht im Block enthalten sein. Der Block darf nur den Teil der Seite enthalten, der später Feldwerte (das, was beim Ausfüllen eingetragen wird) enthält.

Falls Sie die Ränder des Erkennungsbereichs genauer einstellen müssen, können Sie die Größe des Blocks ändern oder den Rahmen mit der Maus verschieben. Falls Sie beim Markieren der Blöcke einen Fehler machen, können Sie den aktiven Block durch drücken der Entf-Taste löschen.

4. Jetzt müssen Sie das Datenbankfeld angeben, in das das Erkennungsergebnis des aktuellen Felds exportiert werden soll. Geben Sie dazu »Nachname« in das Eingabefeld *Datenbankfeld* ein.

Hinweis: Sofern die Datenbank bereits existiert, wählen Sie das gewünschte Feld im Listenfeld Datenbankfeld aus. Gibt es die Datenbank noch nicht, dann berücksichtigen Sie eventuelle Einschränkungen bezüglich der Feldnamen, etwa Länge der Namen und verwendbare Zeichen. Sie können eventuelle Probleme vermeiden, wenn Sie nur lateinische Buchstaben und nicht mehr als 8 Zeichen verwenden.

5. Führen Sie diese Schritte für alle *Textfelder* des Formulars aus: »First Name«, »Month of Birth«, »Town«, »Street«, »House«, »Apartment«, »Series«,

»Issued« und »Date«. Geben Sie für jedes Feld geeignete Datenbankfeldnamen an.

Führen Sie diese Schritte für alle *Zahlenfelder* des Formulars aus: »Index«, »Zip code«, »Phone number«, »Day of Birth«, »Year of Birth« und »Date«. Geben Sie für jedes Feld ein geeignetes Datenbankfeld an. Anders als bei den Textfeldern sollten Sie die Option *Standardeinstellungen* deaktivieren und als *Sprache* den Wert *Ziffern* einstellen.

6. Markieren Sie »Male (M)« und »Female (F)« als zwei *Kontrollkästchen* und erstellen Sie anschließend einen Block, der beide enthält und definieren Sie diesen als *Kontrollkästchengruppe* »Geschlecht«.
7. Im oberen Teil der Schablone finden Sie die Kontrollkästchen (»Rechtecke«) »Done« und »History«. Diese wurden automatisch als statischer Text markiert, Sie müssen den Typ in *Kontrollkästchen* ändern. Klicken Sie dazu mit gedrückter Umschalttaste nacheinander auf beide Blöcke und anschließend mit der rechten Maustaste auf einen der beiden Blöcke. Wählen Sie im erscheinenden Menü den Eintrag *Blocktyp* und dann *Kontrollkästchen*.
8. Damit ist die elektronische Schablone für dieses Formular fertiggestellt. Klicken Sie auf *Schließen*.
Das Fenster Stapel wird angezeigt.
9. Die Arbeit mit einem Stapel erfordert eine Reihe von aufeinanderfolgenden oder gleichzeitig ablaufenden Operationen: Scannen, Erkennung, Bearbeiten und Export des Erkennungsergebnisses.



Scannen

Bereiten Sie die aufgefüllten Formulare für das Scannen vor (im Beispiel sind das zwei ausgefüllte Formulare) und klicken Sie auf die Schaltfläche *Scannen* in der Werkzeugleiste.

Wichtiger Hinweis!

Alle Bilder eines Stapels müssen dieselbe Auflösung wie die Schablone haben.

Für jedes gescannte Formular erscheint ein neues Symbol im Stapelfenster und eine spezielle Markierung in der Bildspalte. Diese Bilder sind lediglich Abbilder der Formulare und können nicht bearbeitet werden.



Erkennung

Nachdem alle Formulare eingescannt sind, können Sie die Erkennung starten. Klicken Sie dazu auf den Pfeil rechts neben der Schaltfläche *Erkennung* und wählen Sie aus dem Kontextmenü den Eintrag *Alle Seiten erkennen*.

Das Programm beginnt mit der Erkennung. Ein spezielles Zeichen erscheint in der Textspalte für jede erkannte Seite.

Bearbeiten

Doppelklicken Sie auf das Symbol eines Formulars im Stapelfenster, um das Erkennungsergebnis zu sehen und (nötigenfalls) zu bearbeiten.

Rechts auf dem Bildschirm sehen Sie das Editorfenster mit dem erkannten Formular und unten das vergrößerte Abbild des Formulars.

Mit den Schaltflächen  und  können Sie zu anderen Stapelseiten wechseln

Export

Um das Erkennungsergebnis in einer Datei zu speichern, klicken Sie auf den Pfeil rechts neben der Schaltfläche *Speichern* und wählen Sie aus dem Menü *Als Datei speichern*. Wählen Sie Laufwerk, Ordner und Dateiname, unter dem Sie das Erkennungsergebnis speichern möchten und klicken Sie auf *Speichern*.

Das ist alles! Das Formular wird als TXT-Datei gespeichert.

Sie können neben TXT auch die Formate DBF, CSV oder XLS verwenden.

Mit FineReader Enterprise können Sie die Daten via ODBC auch in Datenbanken wie MS Access exportieren.

Verarbeitung im Netzwerk

Haben Sie mehrere Computer in einem Netzwerk, dann können Sie die Verarbeitung auf mehreren Computern gleichzeitig durchführen. Ein Beschreibung finden Sie im Abschnitt »Stapelverarbeitung im Netzwerk« auf Seite 58.

Installation und Konfiguration

Software- und Hardwareanforderungen

Sie benötigen folgende Dinge, um FineReader 4.0 zu installieren:

Computer mit Intel 486 Prozessor oder besser (Pentium 133 oder besser wird empfohlen), Windows 95, (Pan-European Version empfohlen), Windows 98 oder Windows NT 4.0 (Service Pack 3), 16 MB RAM für Windows 95, 98; 32 MB für Windows NT, 30 MB freien Speicherplatz auf der Festplatte für die minimale Installation und etwa 60 MB für die vollständige Installation, TWAIN-kompatibler Scanner oder gleichwertiges Gerät, Maus oder anderes Zeigegerät, CD-ROM-Laufwerk.

Installation von FineReader

FineReader führt Sie mit Hilfe von Anweisungen durch alle Installationsschritte. Für eine möglichst rasche Installation sollten während der Installation keine anderen Programme (z.B. MS Outlook, MS Word, etc.) laufen.

Wichtiger Hinweis! Wenn Sie Benutzeralphabeten und Wörterbücher von früheren Versionen weiterverwenden wollen, sollten Sie die alte Version vor der Installation nicht löschen. Sobald die neue Version installiert ist, können Sie die alten Dateien weiterverwenden. Weiter Hinweise hierzu finden Sie unter »Wörterbücher und Benutzeralphabeten von FineReader 3.0 verwenden« im Anhang.

1. Legen Sie die CD-ROM in das CD-ROM-Laufwerk.
2. Klicken Sie auf Start und wählen Sie *Einstellungen>Systemsteuerung*.
3. Öffnen Sie *Software* durch Doppelklick.
4. Wählen Sie die Seite *Installieren/Deinstallieren* und klicken Sie *Installieren...*
5. Befolgen Sie die Anweisungen auf dem Bildschirm.

*Hinweis: Schritte 2-4 sind gleichwertig mit dem Starten des Programms *install.exe* auf der «Start-up Diskette».*

Installationsprobleme

In seltenen Fällen können während der Installation Probleme auftreten, weil Inkompatibilitäten zwischen verschiedenen Softwarekomponenten vorliegen.

Falls das Programm eine Fehlermeldung anzeigt, lesen Sie bitte die Datei Readme.hlp (auf der »Start-up Diskette«) für Hinweise zur Fehlerbeseitigung.

FineReader starten

Wenn Sie vorhaben, Dokumente einzuscannen, dann stellen Sie sicher, daß Ihr Scanner an Ihrem Computer angeschlossen, eingeschaltet* und funktionsbereit ist, bevor Sie FineReader starten. Ist der Scanner nicht angeschlossen, dann befolgen Sie die Anweisungen der Scannerdokumentation zum Anschluß des Scanners.

Falls Sie keinen Scanner haben, dann können Sie die Funktion von FineReader anhand der Datei demo.tif im Ordner \Programme\ABBYY FineReader\FineReader\Demo sehen. Weitere Informationen zur Arbeit mit Bilddateien finden Sie unter »Erkennen einer bereits gespeicherten Bilddatei« auf Seite 14.

Um FineReader zu starten:

Klicken Sie *Start>Programme>ABBYY FineReader>FineReader 4.0.*

Arbeiten mit dem Scanner

Wichtiger Hinweis! Beachten Sie die Hinweise in Ihrer Scannerdokumentation zum Anschluß an Ihren Computer. Vergessen Sie nicht, die Scannersoftware (inklusive TWAIN-Treiber) zu installieren.



Um den Scanvorgang zu starten, bereiten Sie alles vor und klicken auf *Scannen*.

Ihr Scanner sollte sofort mit dem Scanvorgang beginnen. Nach etwa 20 Sekunden erscheint das Abbild der Seite auf dem Bildschirm.

Fahren Sie nun wie im Abschnitt »Scannen und Erkennen eines gedruckten Dokuments« auf Seite 11 beschrieben fort.

* Verfügt Ihr Scanner über eine eigene Stromversorgung, dann schalten Sie ihn **vor** dem Computer ein.

Beginnt Ihr Scanner nicht sofort mit dem Scanvorgang, dann erscheint der Dialog mit den Scannerparametern.

Hinweis: Die Oberflächen von verschiedenen Scannern können stark unterschiedlich sein, da der Scanmanagerdialog (die TWAIN-Schnittstelle des Scanners) vom Scannerhersteller entwickelt wird. Daher ist die Benutzerschnittstelle zwischen verschiedenen Scannern unterschiedlich und manchmal nicht benutzerfreundlich oder intuitiv.

Prüfen Sie die Scannereinstellungen. Wenn Sie bisher noch nicht mit einem Scanner gearbeitet haben, dann denken Sie daran, daß ein Scanner drei Haupteinstellungen hat, die die Qualität des Scanergebnisses beeinflussen: Auflösung, Helligkeit und Scanmodus (Farbeinstellung).

Korrekte Scannereinstellungen

- Auflösung: 300 dpi.

Hinweis: Manche TWAIN-Schnittstellen (z. B. HP PictureScan) lassen für OCR keine andere Auflösung als 300 dpi zu.

- Helligkeit: (manchmal Schwellwert genannt oder als Sonne oder schwarzweißer Kreis dargestellt) mittlerer Wert (50%).
- Scanmodus: Graustufen (256 Farben)*.

Nachdem Sie diese Parameter eingestellt haben, klicken Sie auf *Scannen* oder eine ähnlich benannte Schaltfläche, um den Scanvorgang zu starten.

Bei **Flachbett-Scannern** hören Sie ein Geräusch und eine Lichtquelle bewegt sich unter der Scannerabdeckung.

Bei **Einzugscannern** wird das Dokument wie bei einem Faxgerät eingezogen.

Bei **Handscannern** müssen Sie den Scanner über das Blatt bewegen. Klicken Sie den Startknopf und bewegen Sie den Scanner über die Seite. Drücken Sie eine beliebige Taste der Tastatur, um den Scanvorgang zu beenden.

Nach dem Scanvorgang erscheint ein Abbild der Seite im Bildfenster von FineReader.

Lesen Sie in Ihrem Scannerhandbuch nach, falls Sie noch Fragen zur Scannerbedienung haben.

Prüfen der Eigenschaften eines gescannten Bilds

Sie können die Eigenschaften eines gescannten Bilds prüfen. Klicken Sie dazu mit der rechten Maustaste in das Bildfenster und wählen Sie im Kontextmenü den Eintrag *Parameter*, anschließend die Seite *Bildeigenschaften*.

* Für alle Scanner außer Handscannern. Für Handscanner: schwarzweiß (Strichzeichnung, OCR, Text)

Wenn Sie sich an die Empfehlungen in diesem Handbuch gehalten haben, sollten Sie folgende Werte sehen: Bildtyp Graustufen*, Auflösung 300 dpi.

Diese Werte sind die gebräuchlichsten für Texterkennung.

Ratschläge zum Scannerkauf

Wenn Sie sich einen Scanner anschaffen wollen, dann wählen Sie einen TWAIN-kompatiblen. TWAIN ist ein Standardprotokoll, daß von fast allen Scannerherstellern unterstützt wird und sicherstellt, daß der Scanner mit FineReader zusammenarbeitet.

Hinweis: Unglücklicherweise erreicht man mit vielen Scannern nur dann gute Ergebnisse, wenn die Option TWAIN-Dialog anzeigen aktiviert ist. Ist diese Option deaktiviert, dann erlauben manche TWAIN-Treiber die Einstellung der Helligkeit und/oder Papiergröße nicht.

Sie finden eine Liste von FineReader-kompatiblen Scannern in der Datei *readme.hlp* (auf der Diskette »Start-up Diskette«) und auf der ABBYY Website (www.abbyy.ru).

Wenn Ihr Scanner nicht mit FineReader arbeitet

Arbeitet Ihr Scanner nicht mit FineReader, dann versuchen Sie folgendes:

1. Besorgen Sie sich eine neue Scannertreiberversion bei Ihrem Händler oder im Internet.

Hinweis: Falls Sie die Adresse Ihres Scannerherstellers nicht kennen, wenden Sie sich an die TWAIN-Organisation: www.twain.org. Außerdem gibt es unabhängige TWAIN-Treiber Hersteller für weitverbreitete Scanner (z. B. CFM www.cfm.de)

2. Besuchen Sie die Website von ABBYY: <http://www.abbyy.ru>. Möglicherweise finden Sie dort eine neue Konfigurationsdatei für Ihren Scanner.
3. Wenden Sie sich an den technischen Support von ABBYY: support@abbyy.ru.
4. Wenn Sie keinen neuen Treiber erhalten können, dann können Sie im folgenden »Notmodus« arbeiten. Scannen Sie die Dokumente mit der Scannersoftware und speichern Sie sie in einem von FineReader unterstützten Format (TIFF, BMP, JPEG, PCX, DCX*). Öffnen Sie diese Bilder in FineReader.

Hinweis: Sie können einen Stapel automatisch erstellen. Legen Sie dazu die gescannten Bilder in den Stapelordner. Die Dateinamen müssen dabei aus dem Namen des Stapels gefolgt von einer 4stelligen Nummer bestehen. Drücken Sie im Stapelfenster F5, um den Stapel zu aktualisieren.

* Eine Liste der unterstützten Formate finden Sie im Anhang.

Wie geht es weiter?

Nachdem Sie FineReader 4.0 installiert haben und Ihr Scanner mit FineReader zusammenarbeitet, fahren Sie im Kapitel »**Schnelleinstieg**« fort. Dort lernen Sie, wie Sie schnell die gewünschten Resultate erzielen.



Sie können die Funktion des Programms auch interaktiv mit dem *Scan&Read* Assistenten erlernen. Der Assistent zeigt Ihnen die Grundlagen des Scannens und der Erkennung. Lesen Sie einfach die Hilfetexte des Assistenten und folgen Sie dessen Anweisungen.

Viel Erfolg!

K a p i t e l 3

FineReader 4.0 Überblick

FineReader ist ein Omnifont OCR (Optical Character Recognition) Programm. Das bedeutet, daß es ohne vorheriges Training Texte in fast jeder Druckqualität und Schrift verarbeiten kann.

FineReader ist extrem unempfindlich gegen Druckfehler. Das ist ein Ergebnis der integrierten »Fontänenentransformations«-Technologie.

Dieses Kapitel gibt einen Überblick über die Funktionen von FineReader 4.0 in den Versionen Standard, Professional und Enterprise.

FineReader 4.0 Standard Eigenschaften

- FineReader ermöglicht das Einlesen von Text 5-10 mal schneller, als das eine Schreibkraft kann. Und mit wesentlich weniger Fehlern.
- Um ein Dokument in den Computer einzulesen, legen Sie es in den Scanner und klicken auf *Scan&Read*. 30 bis 90 Sekunden später erscheint der erkannte Text im Fenster *Editor*.
- Mit dem *Scan&Read* Assistenten erhalten Sie schnell Resultate, ohne sich in das Programm einarbeiten zu müssen.
- Mit FineReader können Sie auf einem Computer gleichzeitig Text erkennen lassen und erkannten Text bearbeiten.
- Unterstützt Drag&Drop.
- Unterstützt MMX.
- Arbeitet mit Lingvo 4.5 oder neuer zusammen.

Scannen und markieren von Blöcken

- Arbeitet mit allen Scannern, die das TWAIN-Protokoll unterstützen.
- Automatische Helligkeitseinstellung – liefert bestmögliche Ergebnisse bei hellen und dunklen Vorlagen.
- Automatische und manuelle Segmentierung von Tabellen.
- Scannen und Speichern von Farbbildern.

Arbeiten mit Bildern

- Automatische Erkennung der Vorlagenlage und Drehen in die korrekte Position.
- Säubern von Bildern.
- Drehung von Bildern um 90, 180 und 270 Grad.
- Automatische Farbumkehr von dunklen Blöcken mit heller Schrift.

Erkennung

- Erkennung von Russisch, Englisch, Französisch, Deutsch, Spanisch, Italienisch, Schwedisch und anderen Sprachen, einschließlich gemischtsprachigen Texten (die Liste der unterstützten Sprachen finden Sie im Anhang auf Seite 60).
- Erkennung von mehrspaltigem Text mit Abbildungen und Tabellen. Das Seitenlayout wird mit dem Text gespeichert.
- Erkennung von Blöcken mit hellem Text vor dunklem Hintergrund.
- Erkennung von Farbbildern.
- Das Programm kann trainiert werden, neue Zeichen zu erkennen.

Bearbeiten und Export

- Der integrierte Texteditor erlaubt Ihnen, die Schrift, die Schriftgröße und die Formatierung von Text, Absätzen usw. zu ändern.
- Die eingebaute Rechtschreibprüfung erlaubt Ihnen, neue Wörter zum Wörterbuch hinzuzufügen (die Liste der unterstützten Sprachen finden Sie im Anhang auf Seite 60).
- Automatische Erstellung unterschiedlicher Stilvorlagen für Absätze mit unterschiedlicher Schriftart und Schriftgröße.
- Weiterverwendung im Internet: die Erkennungsergebnisse können im HTML-Format gespeichert werden.
- Die Erkennungsergebnisse können an andere Programme gesendet werden (z. B. MS Word 6.0, MS Word 95, MS Word 97, MS Excel 6.0, MS Excel 95, MS Excel 97, Word Pro 97, WordPerfect 7.0, WordPerfect 8.0, Stylus 3.0 und PROMT 98), ohne sie in vorher in einer Datei zu speichern.

FineReader 4.0 Professional Eigenschaften

- Alle Eigenschaften von FineReader 4.0 Standard.
- Verarbeiten von Dokumenten im Netzwerk zur deutlichen Beschleunigung der Verarbeitung von großen Dokumenten (vorausgesetzt, es wurden Lizenzen für mehrere Arbeitsplätze erworben).
- Erkennung von Formularen mit Hilfe von Schablonen zur Massenerfassung von Daten aus standardisierten Formularen.
- Ausrichten und Auswählen der Schablone anhand von Referenzpunkten.
- Schabloneneditor zur Erkennung verschiedenster Formulare.
- Erkennung von Text in nicht unterstützten Sprachen.
- Erkennung von Barcodes (Check Code 39, Check Interleaved 25, Code 128, Code 39, EAN 13, EAN 8 und Interleaved 25).
- Automatischer Ausgleich von Scannerverzerrungen durch Ausrichten der Schablone.

FineReader 4.0 Enterprise Eigenschaften

- Alle Eigenschaften von FineReader 4.0 Professional (siehe oben).
- Erkennung von handgeschriebenen Blockschrift-Buchstaben für Russisch, Englisch, Deutsch und Ukrainisch.
- Automatische Schablonenauswahl bei der Erkennung von unsortierten Formularstapeln.
- Export von Abbildungen in Datenbanken.
- Automatische Überprüfung des Erkennungsergebnisses durch reguläre Ausdrücke, Prüfregele und Vergleich mit der Datenbank.
- Datenbankexport via ODBC.

K a p i t e l 4

Arbeiten mit FineReader 4.0

In diesem Kapitel werden die wichtigsten Funktionen von FineReader 4.0 beschrieben. Zusätzliche Informationen und Schritt-für-Schritt-Anleitungen finden Sie in der Online-Hilfe.

Das Hauptfenster

Wenn das Programm gestartet wird, öffnet sich automatisch ein neues Stapelfenster, in dem die gescannten Seiten und die Erkennungsergebnisse erscheinen. Jedes eingescannte Dokument wird als einzelne Seite im Stapel angezeigt. Im Hauptfenster werden die Arbeitsfenster *Stapel*, *Bild*, *Text* und *Zoom* angezeigt (Abbildung 10).

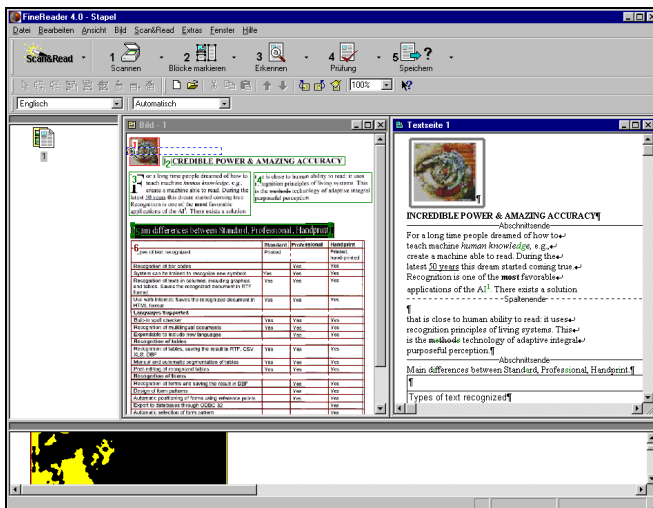


Abbildung 10. Das Hauptfenster von FineReader 4.0.

Im oberen Bereich des Fenster sehen Sie die Werkzeugleiste. Verwenden Sie das Menü *Ansicht>Werkzeugleiste* oder das Kontextmenü, um Werkzeugleiste ein- und auszublenden. Klicken Sie mit der rechten Maustaste auf eine Werkzeugleiste,

um das Kontextmenü (Abbildung 11) anzuzeigen. Die Werkzeugleisten, die am Bildschirm sichtbar sind, werden mit einem Häkchen versehen. Die Bearbeitungs-Werkzeugleiste ist standardmäßig ausgeschaltet. Wählen Sie die Werkzeugleiste, die Sie ein- oder ausblenden möchten.



Abbildung 11. Kontextmenü mit der Liste der Werkzeugleisten. Die markierten Werkzeugleisten werden am Bildschirm angezeigt.

Die Scan&Read Werkzeugleiste

Die *Scan&Read* Werkzeugleiste (Abbildung 12) enthält die Schaltflächen für die Hauptfunktionen des Programms. Die Zahlen in den Schaltflächen geben die Reihenfolge der Operationen an, um ein gedrucktes Dokument in ein elektronisches umzuwandeln: 1 - Scannen, 2 – Blöcke markieren, 3 - Erkennung, 4 – Ergebnis prüfen und 5 - Speichern.

Die Schaltfläche links außen *Scan/Open&Read* führt die Operationen Scannen, Blöcke markieren und Erkennung (im Modus *Scan&Read* für gedruckte Dokumente) oder Blöcke markieren und Erkennung (im *Open&Read* Modus für Bilddateien) aus. Diese Schaltfläche öffnet auch den *Scan&Read* Assistenten, der Sie in die Bedienung des OCR-Programms einführt.

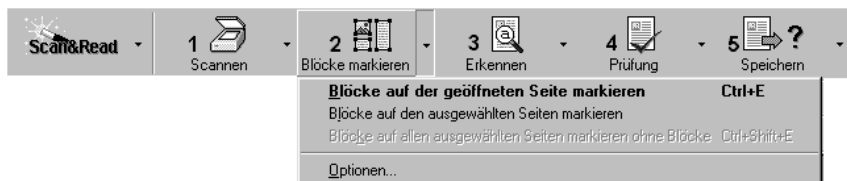


Abbildung 12. Die »Scan&Read«-Leiste.

Der Status der Schaltflächen *Blöcke markieren* und *Erkennung* hängt vom aktiven Fenster und davon ab, ob im Stapelfenster Seiten ausgewählt sind.

Ist das Bildfenster aktiv, dann wirken die Schaltflächen auf die jeweils geöffnete Seite.

Sind keine Bilder geöffnet und enthält der Stapel wenigstens eine markierte Seite, dann wirken die Schaltflächen auf die jeweils ausgewählten Seiten.

Hinweis: Der Modus der Schaltfläche kann manuell eingestellt werden. Klicken Sie auf den Pfeil rechts neben der Schaltfläche, um das Menü zu öffnen. Abbildung 12 zeigt das geöffnete Menü für die Schaltfläche 2 – Blöcke markieren. Einer der Menüeinträge ist fett. Dieser Eintrag wird aufgeführt, wenn Sie auf die Schaltfläche klicken. Wählen Sie einen anderen Menüeintrag, so wird dieser ausgeführt.

Die Standardwerkzeuggeste

Die Standardwerkzeuggeste (Abbildung 13) enthält folgende Schaltflächen: die fünf üblichen Schaltflächen (Neu, Öffnen, Ausschneiden, Kopieren, Einfügen), ferner Schaltflächen, um sich im Stapel vor und zurück zu bewegen (zur nächsten oder vorhergehenden Seite), Schaltflächen zum Rotieren und Säubern der Bilder und eine *Direkthilfe*-Schaltfläche.

Diese Schaltflächen wirken jeweils auf das geöffnete Bild oder die Bilder, die im Stapelfenster ausgewählt sind. Wenn Sie auf die Schaltfläche *Direkthilfe* geklickt haben, verändert der Mauszeiger seine Funktion: wenn Sie nun auf ein Objekt klicken (Schaltfläche oder Menüeintrag), erhalten Sie einen kurzen Hilfetext (die Form des Mauszeigers ist in diesem Modus ein Pfeil mit einem Fragezeichen).

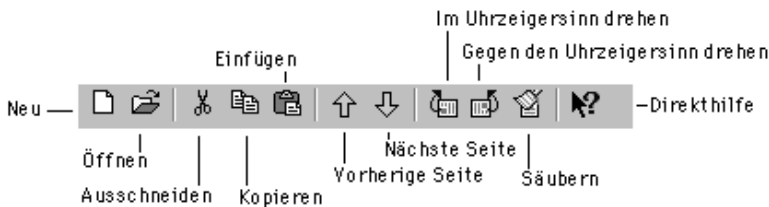


Abbildung 13. Die Standardwerkzeuggeste

Die Struktur eines Stapels

Beim Starten von FineReader wird automatisch ein neuer Stapel geöffnet. Ein Stapel verhält sich ähnlich wie ein Ordner, der gescannte und erkannte Seite enthält. Sie können Teile des Stapels oder den gesamten Stapel speichern. Im Stapelfenster wird eine Liste der Seiten im aktuellen Stapel angezeigt. Doppelklicken Sie das Symbol einer Seite oder die Seitennummer, um sie anzusehen. Sie sehen die Dateien, die zu dieser Seite gehören: den Text bzw. das Bild der Seite im *Textfenster* bzw. dem *Bildfenster*.

Die Seiten in einem Stapel werden auf die gleiche Weise ausgewählt, wie im Dateien im Explorer: mit einem Mausklick oder mit den Cursortasten. Genauere Informationen finden Sie im Abschnitt »Seiten markieren« auf Seite 37.

Das Stapelfenster

Standardmäßig ist das Stapelfenster am linken Rand des Hauptfensters verankert und Seiten werden mit großen Symbolen dargestellt.

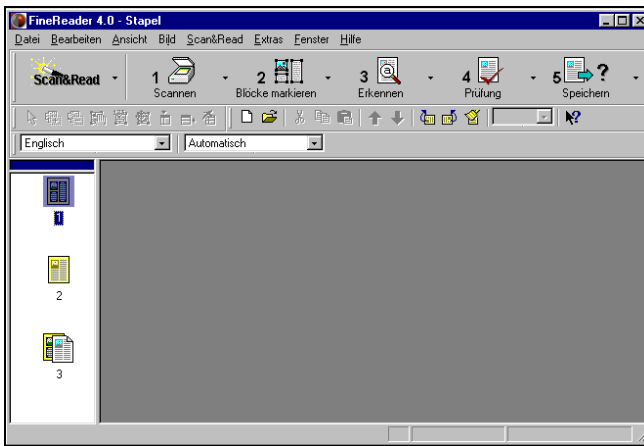


Abbildung 14. Das Stapelfenster im gedockten Zustand mit großen Symbolen für Seiten (die Standardeinstellung).

Wenn Sie das Stapelfenster bewegen möchten, können Sie es vom Rand lösen (klicken Sie mit der rechten Maustaste in das Fenster und deaktivieren Sie die Option *Verankerte Ansicht* im Kontextmenü.)

Das Stapelfenster kann außerdem Informationen über die Operationen, die für einzelne Seiten ausgeführt wurden, anzeigen: Ob sie erkannt, bearbeitet oder exportiert wurde, ob während der Erkennung ein Fehler aufgetreten ist usw. Damit diese Informationen angezeigt werden (Abbildung 15), müssen Sie im Kontextmenü die Option *Seitenparameter anzeigen* aktivieren.

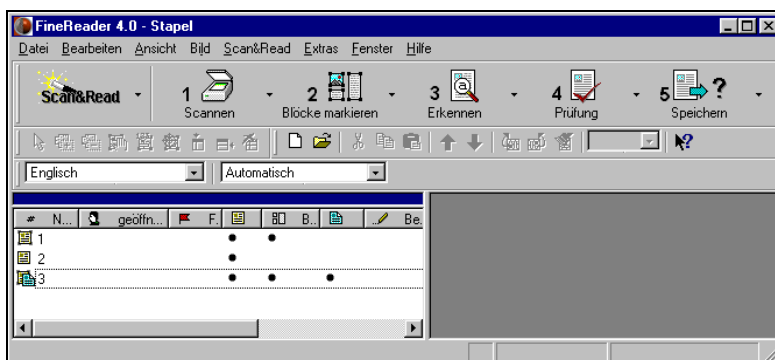


Abbildung 15. Das Stapelfenster mit eingeschalteten Seitenparametern. Die Komponenten einer Seite werden als Punkte markiert (Bilder, Blöcke, Text usw.)

Symbole für Seiten

Eine Stapelseite kann folgende Komponenten haben:



1 Eine *Bilddatei* (gescannt oder aus einer Datei). Solche Seiten werden durch ein gelbes Seitensymbol dargestellt (sofern die Seitenparameter ausgeschaltet sind).



2 Eine *Textdatei* (erkannter Text). Solche Seiten werden durch ein weißes (oder blaues) Seitensymbol dargestellt. Eine Seite kann nur dann als Textdatei dargestellt werden, wenn die Bilddatei nach der Erkennung gelöscht wurde (und die Seitenparameter ausgeschaltet sind).



3 *Bild und Textdatei*. Solche Seiten werden durch zwei überlappende gelbe und weiße Symbole dargestellt (sofern die Seitenparameter ausgeschaltet sind).

Numerierung der Stapelseiten

Ein Stapel kann bis zu 9999 Seiten aufnehmen. Sie sehen die Seitennummer unterhalb des Seitensymbols im Stapelfenster (vgl. Abbildungen oben) und in der Titelzeile des Fensters, in dem die Seite angezeigt wird. Etwa »Bild – 5« oder »Text – 7«.

Arbeiten mit Stapelseiten

Seiten markieren

Der Stapel unterstützt nicht zusammenhängende Markierungen. Das heißt, Sie können viele aufeinanderfolgende und nicht aufeinanderfolgende Seiten markieren.

- Um aufeinanderfolgende Seite zu markieren, klicken Sie mit der Maus auf die erste Seite und anschließend bei gedrückter Umschalttaste auf die letzte Seite der gewünschten Markierung.
- Um nicht aufeinanderfolgende Seiten zu markieren, klicken Sie bei gedrückter Strg-Taste auf die gewünschten Seiten.
- Um alle Seiten zu markieren, aktivieren Sie das Stapelfenster und wählen Sie den Befehl *Alles markieren* aus dem Menü *Bearbeiten*.

Hinweise:

1. Die Funktionsweise von *Alles markieren* hängt vom aktiven Fenster ab – *Stapel*, *Bild* oder *Text*: der Befehl markiert alle Seiten des Stapels, alle Blöcke des Bilds oder allen Text im Editor.
2. Die Markierung mehrerer Seiten ist im Stapelfenster am einfachsten, wenn die *Seitenparameter* eingeschaltet sind.

Arbeiten mit Bildern

Die Arbeitsweise der Schaltflächen *Blöcke markieren* und *Erkennung* sind im Abschnitt »Die Scan&Read Werkzeugleiste« auf Seite 33 beschrieben.

Wie bei der Drehung, Farbumkehr und Säuberung werden alle Operationen auf das **aktive Fenster** angewandt. Ist das Stapelfenster aktiv, dann wirken die Befehle auf die markierten Seiten im Stapelfenster.

Arbeiten mit Text

Operationen mit Erkennungsergebnissen (Suchen, Rechtschreibprüfung usw.) werden auf der aktiven Seite des Stapels ausgeführt. Ist keine Seite geöffnet, dann wird die erste Seite der Markierung bearbeitet. Anschließend fragt das Programm, ob es fortfahren soll.

Eine ausführliche Anleitung finden Sie in der Online-Hilfe.

Scannen

Die Erkennungsqualität hängt sehr stark von der Qualität des gescannten Dokuments ab. In diesem Abschnitt wird erläutert, wie man »gute« von »schlechten« Texten unterscheidet und wie die Scannerparameter für das bestmögliche Ergebnis eingestellt werden.

Gute und schlechte Texte

»Gute« Texte

Text mit guter Qualität haben wenig ausgefranste, zusammenhängende, verschmutzte oder verzerrte Buchstaben. Ein Mensch muß sich nicht anstrengen, um den Text zu lesen.

Beispiele: dieses Handbuch; Ausdrücke von Laserdruckern, fast alle heutigen Magazine und fast alle farbigen Produktionen.

»Schlechte« Texte

Texte mit schlechter Qualität sind schwierig zu lesen, weil der Kontrast zu gering ist, d. h. der Unterschied zwischen hellen und dunklen Bereich ist klein. Solche Texte können zu dunkel oder zu hell sein.

Beispiele: Ausdruck von Nadeldruckern im Entwurfsmodus, mit einem verbrauchten Farbband, ein Schreibmaschinendurchschlag, Text von einer Schreibmaschine mit verbrauchtem Farbband, alle Texte, bei denen Buchstaben verzerrt und unvollständig sind oder ineinander laufen. (Abbildung 16).



Ineinanderlaufende Buchstaben



*Unvollständige
Zeichen*



*Verzernte, verschmutzte
Zeichen*

Abbildung 16. Beispiel für schlechte Textqualität. Um die Erkennungsqualität zu verbessern, können Sie die Helligkeitseinstellung anpassen. Erhöhen Sie die Helligkeit (machen Sie das Bild heller) bei ineinanderlaufenden, verzernten und verschmutzten Zeichen, und verringern Sie sie (machen Sie das Bild dunkler) bei unvollständigen Zeichen.

Scaneinstellungen für gute Ergebnisse

Wenn Sie FineReader zum erstenmal starten, werden die Standard-Scaneinstellungen verwendet: Graustufen (256 Farben*), mittlere Helligkeit und 300 dpi Auflösung.

Manchmal müssen Sie diese Einstellungen ändern, z. B. wenn Sie Farbbilder mit dem Text speichern wollen. Wählen Sie dann den Farbmodus. Für kleine Schriften (8 Punkt und darunter) benötigen Sie 400 bis 600 dpi. In diesen Fällen sind die Standardeinstellungen ungeeignet und Sie müssen sie ändern.

Ändern der Scannereinstellungen:

1. Wählen Sie *Scannerparameter* aus dem Menü *Extras*
2. Stellen Sie im Dialog *Scannerparameter* die gewünschten Optionen ein.


*Hinweis: Wenn Sie die Option zur Verwendung des TWAIN-Interface aktiviert haben (auf der Seite *Scanner des Dialogs Optionen*), können Sie den TWAIN-Dialog Ihres Scanners zur Einstellung der *Scannerparameter* verwenden. Die Bedienung dieses Dialogs ist in Ihrer Scannerdokumentation beschrieben. Bei manchen Modellen ist diese Option standardmäßig deaktiviert.*

* Eigentlich reichen 16 Graustufen aus, aber nicht alle Scannertreiber unterstützen das.

Wenn Sie den TWAIN-Dialog aktiviert haben, können Sie eine Vorschau des Dokuments sehen und entsprechend die Höhe und Breite, sowie die Helligkeit und den Kontrast einstellen. Haben Sie den TWAIN-Dialog deaktiviert, dann können Sie die Scannereinstellung in FineReader vornehmen. Die Scannerparameter können Sie in diesem Fall auch im Stapel speichern.

Bild säubern

Wenn Sie trotz korrekter Scanparameter Störungen und verschmutzte Hintergründe erhalten, können Sie das Bild säubern lassen.

Klicken Sie dazu auf die Schaltfläche . Verwenden Sie diese Funktion nicht zu intensiv, weil das Säubern des Bilds möglicherweise auch Punkte, Kommata und feine Elemente von Buchstaben entfernt. Das Resultat ist dann ein schlechtes Erkennungsergebnis.

Mehrseitige Dokumente scannen

Um eine große Anzahl von Seiten einzuscannen, bietet sich die Verwendung eines Vorlageneinzugs (ADF, Automatic Document Feeder) an. Natürlich ist das praktisch, wenn Einzelseiten vorliegen. Bei gebundenen Vorlagen muß erst die Bindung entfernt werden.

Damit der Scanner die Seiten aus dem Vorlageneinzug annimmt, müssen Sie im Dialog *Scannerparameter* (*Extras>Scannerparameter...*) die Verwendung des Vorlageneinzugs aktivieren. Falls Sie den TWAIN-Dialog verwenden, müssen Sie den Mehrseitenmodus einstellen.

Können Sie aus irgendeinem Grund den Vorlageneinzug nicht verwenden, dann gehen Sie wie folgt vor: nach dem Einscannen einer Seite macht der Scanner eine Pause (die Länge könne Sie in den Scannerparametern einstellen). In dieser Zeit können Sie die nächste Seite in den Scanner legen. Nach der Pause setzt der Scanner den Scanvorgang automatisch fort.

Sie müssen lediglich die Länge der Pause einstellen. Aktivieren Sie dazu die Option *Pause zwischen Seiten* im Dialog *Scannerparameter* (*Extras>Scannerparameter*). Geben Sie dann die gewünschte Pausenlänge in Sekunden ein. Das ist die Wartezeit zwischen der Verarbeitung von zwei Seiten.

Seitennumerierung

Beim Start eines Scanvorgangs fragt Sie das Programm, wie die Seiten numeriert werden sollen:

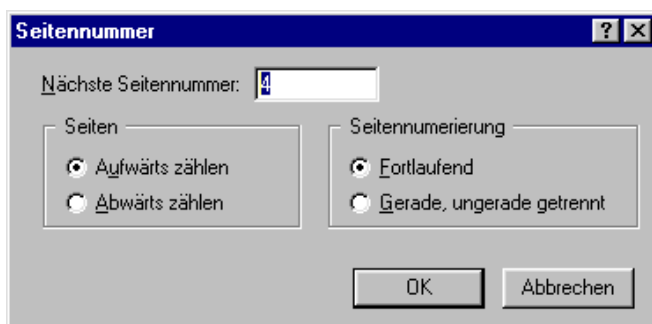



Abbildung 17. Seitennumerierung.

Geben Sie im Dialog *Seitennummer* die Nummer der ersten Seite ein. Die Seiten können aufsteigend oder absteigend numeriert werden. Sie können aufeinanderfolgend oder getrennt nach geraden und ungeraden Seiten gezählt werden. Wenn Sie zum Beispiel einen Stapel mit beidseitig bedruckten Seiten mit einem Vorlageneinzug scannen möchten, dann ist es ratsam, von allen Blättern erst die eine und anschließend von allen die andere Seite zu scannen. In diesem Fall sollten Sie gerade bzw. ungerade Seitennummern wählen.

Ob Sie aufsteigende oder absteigende Nummern wählen, hängt nur davon ab, wie Sie den Stapel in den Vorlageneinzug einlegen – hohe Seitennummern zuerst oder niedrige Seitennummern zuerst.

Bilder zu einem Stapel hinzufügen

Klicken Sie auf , um Bilddateien zu einem Stapel hinzuzufügen.

Wählen Sie im Dialog *Öffnen* ein oder mehrere Bilder, die Sie zum aktuellen Stapel hinzufügen möchten.

Hinweis: In diesem Dialog können Sie mehrere Dateien gleichzeitig auswählen. Drücken Sie die Umschalttaste für eine durchgehende Auswahl oder die Strg-Taste für eine nicht durchgehende Auswahl.

Sie können die Bilder in den Stapel bewegen, wenn Sie die entsprechende Option im Dialog *Öffnen* markieren.

Was geschieht mit Farb- und Graustufenbildern, wenn Sie in FineReader geöffnet werden?

Farb- und Graustufenbilder

Beim Öffnen von Farb- oder Graustufenbildern erzeugt FineReader eine Schwarzweißkopie für die Erkennung. Deshalb sehen Sie im Zoomfenster ein Schwarzweißbild des Farb- oder Graustufenbilds, wenn Sie den erkannten Text bearbeiten.

Schwarzweißbild statt Farb- oder Graustufenbilds speichern

Um Plattenplatz zu sparen, können Sie Farb- bzw. Graustufenbilder als Schwarzweißbild speichern. Aktivieren Sie dazu die Option *Nur Schwarzweißbilder im Stapel speichern* (*Extras>Optionen>Scanner*). Wenn Sie anschließend ein Farb- oder Graustufenbild laden, wird es von FineReader als Schwarzweißbild im Stapel gespeichert.

Fehlermeldung ungeeignete Auflösung

Manchmal, wenn Sie ein Bild mit einer Auflösung kleiner 100 dpi oder größer 800 dpi laden, wird Ihnen FineReader vorschlagen, die Auflösung anzupassen.

Es gibt folgende Gründe für solche Bilder:

- Es gibt Bildformate, deren Auflösung verschieden interpretiert werden kann.
- Ungebräuchliche Bildformate.
- Falsche Scannereinstellungen.
- Für den Großteil der Handscanner ist die wirkliche Scanauflösung für Farb- oder Graustufenbilder nicht 300 oder 400 dpi wie angegeben, sondern wesentlich weniger (75 – 100 dpi).

Abhängig vom Ursprung des Bilds sollten Sie die Auflösung anpassen.

Bedenken Sie, daß eine falsche Auflösung folgende Folgen haben kann:

- FineReader kann im Bild keine Blöcke markieren.
- Die Schriftgröße wird falsch erkannt.

Seitenanalyse (Blöcke markieren)

Vor der eigentlichen Erkennung muß das Programm wissen, welche Teile des Bilds erkannt werden sollen. Dazu analysiert das Programm die Seite und markiert Blöcke als Text, Abbildungen und Tabellen. Eine Seite kann manuell oder automatisch analysiert werden. FineReader kann komplexe Textlayouts erfolgreich analysieren. Klicken Sie dazu auf die Schaltfläche *Blöcke markieren* in der *Scan&Read* Werkzeugleiste. Weitere Informationen finden Sie im Abschnitt »Die Scan&Read Werkzeugleiste« auf Seite 33.

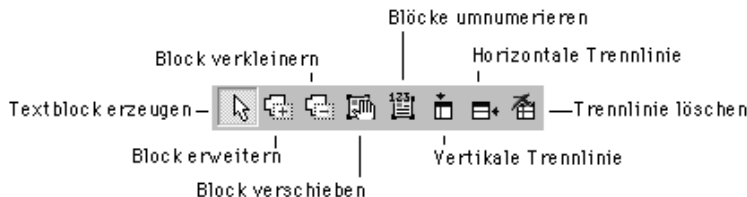



Abbildung 18. Die Ansichts-Werkzeugleiste. Verwenden Sie diese Werkzeuge, um das Ergebnis der automatischen Analyse zu bearbeiten.

Blöcke manuell markieren und bearbeiten

Um einen Block zu markieren, wählen Sie das  Werkzeug (die Voreinstellung) und klicken Sie in die linke obere Ecke des Blocks, den Sie markieren möchten. Ziehen Sie bei gedrückter Maustaste einen Rahmen auf, der den gewünschten Block umschließt. Lassen Sie die Maustaste los. Der ausgewählte Bereich wird von einem Rahmen eingeschlossen. Per Voreinstellung ist dies ein Textblock, d. h. er wird erkannt.

Abhängig von Inhalt kann ein Block folgenden Typ haben: Abbildung, Tabelle, Barcode*, Text oder nicht erkennen.

Um den Blocktyp zu ändern, klicken Sie mit der rechten Maustaste in den Block und wählen Sie aus dem erscheinenden Kontextmenü den gewünschten Blocktyp.

* Nur in FineReader 4.0 Professional und Enterprise.

Um die Blockgröße zu ändern, bewegen Sie den Mauszeiger über den Blockrahmen, bis er sich in einen Doppelpfeil verwandelt. Klicken Sie dann auf den Rahmen und ziehen Sie ihn auf die gewünschte Größe.

Blöcke in einem Teil der Seite automatisch markieren

Um Blöcke in einem Teil der Seite automatisch zu markieren, wählen Sie diesen Bereich mit gedrückter Strg-Taste aus (wie wenn Sie einen Block markieren möchten). Der Bereich wird automatisch analysiert.

Sie können den Bereich auch als Block markieren und dann im Kontextmenü *Block analysieren* auswählen.

Tabellenstrukturen festlegen

Wenn das Programm die Blöcke auf einer Seite markiert, findet es automatisch alle Tabellen und analysiert sie (d.h. markiert die Zellen). Um die Zuverlässigkeit der Erkennung bei einer hohen Zahl von Tabellen eines bestimmten Typs zu erhöhen, sollten Sie die Tabelleneigenschaften festlegen.

Folgende Eigenschaften können Sie angeben:

- Ob die Tabelle rechteckig ist.
- Ob die Tabelle einzeilig formatiert ist.
- Ob die Zellen durch schwarze Linien getrennt sind.

Tabelle B (siehe unten) ist rechteckig. Alle schwarzen Trennlinien durchziehen die gesamte Tabelle (im Gegensatz zu Tabelle A – sie ist nicht durchgängig rechteckig).

Tabelle A (unten) ist einzeilig formatiert, weil sie nur Zellen mit genau einer Textzeile enthält. Tabelle B ist nicht einzeilig, weil sie Zellen mit mehr als einer Textzeile enthält.

Temperatur	
Celsius	Kelvin
-273	0
100	373

Tabelle A. Einzeilig, nicht rechteckig

	Celsius
Siedepunkt des Wassers	100
Gefrierpunkt des Wassers	0

Tabelle B. Mehrzeilig, rechteckig.

Einzeilig formatierte Tabellen müssen keine Trennlinien enthalten, weder horizontal noch vertikal (Tabelle C).

Kilometer	Meilen
1	0,62
5	3,1
10	6,2

Tabelle C. Eine einzeilig formatierte Tabelle ohne Trennlinien.

Einstellen der Tabellenstruktur

1. Wählen Sie im Menü *Extras* den Eintrag *Optionen*.
2. Wählen Sie auf der Seite *Seitenanalyse* des Dialogs *Optionen* die gewünschten Einstellungen in der Kontrollgruppe *Tabelle*.

Wichtige Hinweise:

1. Die Tabellenstruktur sollte vor der Erkennung festgelegt werden.
 2. Falsche Angaben beeinträchtigen die Segmentierungsqualität.
 3. Wenn Sie eine große Zahl unterschiedlicher Tabellen einlesen wollen, dann wählen Sie keine der Einträge für Tabellen auf der Seite *Seitenanalyse* (Dialog *Optionen*).
-

Tabellenblöcke bearbeiten

Tabellen, die manuell analysiert werden müssen

Hat eine Tabelle keine horizontalen Trennlinien, mehr als eine Textzeile pro Zelle und gleiche Abstände zwischen den Zeilen (Abbildung 19), dann ist es besser, die Tabelle manuell zu analysieren.

Schließen Sie die Tabelle in einen Block ein und stellen Sie als Blocktyp *Tabelle* ein. Fügen Sie dann horizontale und vertikale Trennlinien mit Hilfe der Werkzeugleiste ein. (Abbildung 18, Seite 43).

1. Jim 1990.
Johnson
2. Mary 1996.
Lay
3. Boris 1997.
Namin

Abbildung 19. Eine Tabelle, die manuell analysiert werden muß. Die Tabelle ist nicht einzeilig formatiert und hat keine horizontalen Trennlinien.

Sie können den Tabellenblock auch automatisch analysieren lassen (mit rechter Maustaste auf die Tabelle klicken, dann *Tabelle analysieren* aus dem Kontextmenü wählen), und anschließend das Ergebnis manuell korrigieren.

Tabelle manuell analysieren:

Aktivieren Sie das Bildfenster durch Anklicken.

Falls das Tabellenfenster zu klein ist, erhöhen Sie die Vergrößerung. Klicken Sie dazu mit der rechten Maustaste in das Bildfenster und wählen Sie im Kontextmenü die gewünschte Vergrößerung.

Falls die Tabelle noch nicht markiert ist, markieren Sie sie mit der Maus. (Siehe »Blöcke manuell markieren und bearbeiten«, Seite 43).

Verwenden Sie dann diese Werkzeuge:



Fügt eine vertikale Linie hinzu




Fügt eine horizontale Linie hinzu



Löscht eine Linie

Hinweise:

- 1. Um eine Linie zu löschen, verschieben Sie sie, bis sie auf dem Rahmen zu liegen kommt.*
- 2. Drücken Sie die Umschalttaste, um die Richtung der Linie, die Sie gerade hinzufügen, zu ändern (vertikal in horizontal und umgekehrt).*

Haben Sie das Hinzufügen oder Löschen der Linien abgeschlossen, dann klicken Sie  oder drücken Sie ESC, um den Mauszeiger in den Ausgangsmodus zurückzuverwandeln.

Sie können Zellen auch teilen oder zusammenfassen (Menü *Bearbeiten*>*Zellen zusammenfassen* bzw. *>Zellen teilen*)

Hinweis: Drücken Sie die Umschalttaste, um mehrere Zellen mit der Maus (oder den Pfeiltasten) auszuwählen.

Erkennung

Die Aufgabe der Erkennung ist es, das gescannte Bild in Text zu verwandeln. Anders ausgedrückt: jedes einzelne Zeichen durch den entsprechenden Computercode zu ersetzen.

Bevor Sie die Erkennung starten, stellen Sie bitte sicher, daß die korrekten Erkennungsparameter in der Erkennungs-Werkzeugleiste (Abbildung 20) eingestellt sind: die Sprache und der Texttyp.



Abbildung 20. Die Erkennungs-Werkzeugleiste

Sprachen für die Erkennung

- Um Englisch, Spanisch, Italienisch, Deutsch, Russisch, Englisch-Russisch, Ukrainisch, Französisch, Schwedisch, Dänisch, Holländisch, Polnisch, Finnisch, Norwegisch etc. und arabische Ziffern erkennen zu können, wählen Sie die gewünschte Sprache in der Erkennungs-Werkzeugleiste.
- Um andere Sprachen oder Sprachkombinationen zu erkennen, wählen Sie den Eintrag *Andere Sprache...* und fahren mit dem nächsten Punkt fort.

Es erscheint der Dialog *Benutzte Sprachen für die Erkennung* (Abbildung 21). In diesem Dialog können Sie die Sprachen auswählen, der Worte im Text vorkommen.

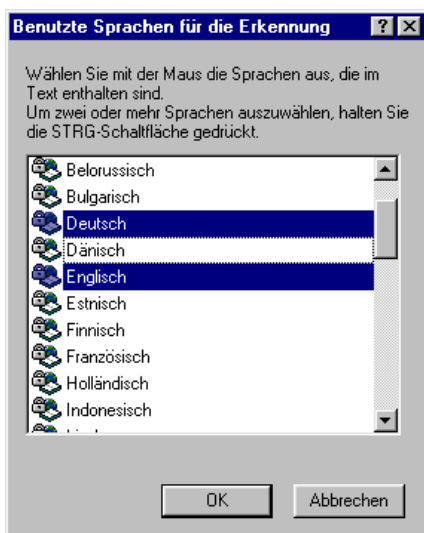


Abbildung 21. Dialog *Benutzte Sprachen für die Erkennung*. Auswahl der Sprachen für einen Englisch-Deutschen Text.

Nachdem Sie auf *OK* geklickt haben, erscheint die gewählte Sprachkombination in der Erkennungs-Werkzeuleiste:



Abbildung 22. Sprachen für die Erkennung. Kombination von Sprachen für Englisch-Deutsche Texte.

Texttyp

Texte, die auf einem Nadeldrucker im Entwurfsmodus oder auf einer Schreibmaschine gedruckt wurden, haben spezifische Eigenschaften (siehe unten) im Vergleich zu Standard-Schrifttypen. Symbole, die mit Nadeldruckern gedruckt wurden, bestehen aus einzelnen Punkten (Abbildung 24). Symbole von Schreibmaschinen (Abbildung 25) haben fast identische Breiten.

Bei den meisten Texten wird der Texttyp automatisch erkannt: Der Wert *Automatisch* ist im Listenfeld *Texttyp* ausgewählt (Abbildung 23). Für Texte von Nadeldruckern und Schreibmaschinen können Sie die Erkennungsqualität erhöhen, wenn Sie im Listenfeld *Texttyp* *Schreibmaschine* bzw. *Nadeldrucker* einstellen.



Abbildung 23. »Automatisch« ist eingestellt.

hardware

Abbildung 24. Ausschnitt einer Seite, die mit einem Nadeldrucker im Entwurfsmodus gedruckt wurde. Deutlich sind die einzelnen Punkte zu sehen.

direction

Abbildung 25. Ausschnitt aus einer Schreibmaschinenseite.

Speichern des Seitenlayouts und der Formatierung

Vor der Erkennung sollten Sie einstellen, wie die Erkennungsergebnisse zu speichern sind: je nachdem, ob Sie Schriften, Text und Bildlayout der Seite mit

gespeichert haben möchten. Sie können die zugehörigen Einstellungen auf der Seite *Formatierung* des Dialogs *Optionen (Extras>Optionen)* vornehmen.

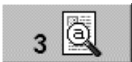
- Wenn Sie nur den Text ohne Formatierung benötigen (z.B. zum Einfügen in einen anderen Text), dann wählen Sie die Option *Formatierung entfernen*.
- Wenn Sie Zeilenschaltungen und Einzüge übernehmen wollen, dann wählen Sie die Option *Formatierung durch Leerzeichen darstellen*.
- Wenn Sie Einzüge nicht benötigen, aber das Schriftformat übernehmen wollen, dann wählen Sie die Option *Buchstabenform und Schriftart speichern*.
- Wenn Sie die gesamte Formatierung übernehmen möchten, dann wählen Sie die Option *Formatierung mit Rahmen speichern*. Diese Option erzeugt die bestmögliche Näherung zum Original.
- Rahmen erschweren aber die spätere Bearbeitung (siehe oben). Wenn Sie also planen, das Ergebnis mit Ihrem Texteditor zu bearbeiten, dann dürfte die Option *Buchstabenform, Schriftart und Spalten speichern* am besten geeignet sein. Rahmen werden hauptsächlich für Bilder benutzt.

Abbildungen im erkannten Text speichern

Um Abbildungen im erkannten Text zu speichern, aktivieren Sie die Option *Abbildungen speichern* auf der Seite *Formatierung* des Dialogs *Optionen (Extras>Optionen)*.

Hinweis: Sie können die Abbildungen auch in eigenen Dateien speichern. Markieren Sie dazu den Abbildungsblock und wählen Sie im Menü Datei den Befehl Bildkopie speichern. Wählen Sie im erscheinenden Dialog die Option Nur aktiven Block speichern.

Erkennung starten



Klicken Sie auf die Schaltfläche *Erkennen* in der Scan&Read Werkzeugleiste, um die Erkennung zu starten.

Ist ein Bild geöffnet, so wird dieses erkannt. Sind eine oder mehrere Seiten in Stapelfenster markiert, dann werden diese Seiten erkannt.


Sie können das Verhalten der Schaltfläche anpassen, indem Sie auf den Pfeil rechts neben der Schaltfläche klicken und die gewünschte Funktion im entsprechenden Menü auswählen.

Hinweis: Während der Erkennung zeigt das Programm einen Dialog an, der Sie warnt, wenn bestimmte Parameter nicht korrekt eingestellt sind. Die Meldungen haben die Form »Helligkeit zu hoch (oder zu gering)«, »Falscher Texttyp« usw. Diese Meldungen helfen Ihnen bei der Prüfung der eingestellten Parameter. Sie

können sie auch ignorieren oder ausschalten (Extras>Optionen > Erkennung>Tips während der Erkennung anzeigen).

Erkennung im Hintergrund

FineReader erlaubt Ihnen, die Erkennung im Hintergrund durchzuführen, während Sie gleichzeitig erkannte Seiten bearbeiten. Dies beschleunigt die Arbeit auf einem Computer erheblich.

Starten Sie die Erkennung im Hintergrund (*Scan&Read Menü, Erkennung im Hintergrund*). In der Statuszeile sehen Sie das Symbol . Sofern Sie die Seitenparameter im Stapelfenster eingeschaltet haben, (Kontextmenü, *Seiteneigenschaften anzeigen*), sehen Sie die Zeile *Erkennung im Hintergrund...* neben der Seite, die erkannt wird. Weiter sehen Sie die Punkte gegenüber den verarbeiteten Seiten erst bei Blöcken und dann in der Textspalte erscheinen.

Während die Erkennung läuft, können Sie bereits erkannte Seiten öffnen und bearbeiten.

Prüfen und Bearbeiten der Ergebnisse

Die Bearbeitungs-Werkzeugleiste (Abbildung 26) erlaubt Ihnen, den erkannten Text zu bearbeiten.

Hinweis: Die Bearbeitungs-Werkzeugleiste ist standardmäßig ausgeschaltet. Um sie anzuzeigen, klicken Sie mit der rechten Maustaste auf eine Werkzeugleiste und wählen Sie sie aus dem Kontextmenü.

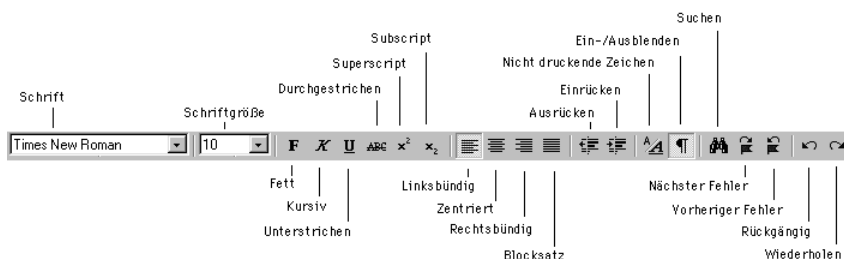


Abbildung 26. Die Bearbeitungs-Werkzeugleiste

Prüfen der Erkennungsergebnisse

Eine sehr nützliche Funktion des Editors von FineReader ist die eingebaute Rechtschreibprüfung für 40 Sprachen und Sprachkombinationen (eine vollständige Liste finden Sie im Anhang auf Seite 60).

Der Originaltext kann eventuell bereits Rechtschreib- und Druckfehler enthalten, die in den erkannten Text übernommen werden. Neben den Fehlern im Originaltext kann der erkannte Text unsicher oder nicht erkannte Worte enthalten (Worte, die unsicher oder nicht erkannte Zeichen enthalten). Damit die Rechtschreibprüfung an diesen Stellen stoppt, aktivieren Sie die entsprechende Option (Menü *Extras>Optionen>Rechtschreibprüfung*).

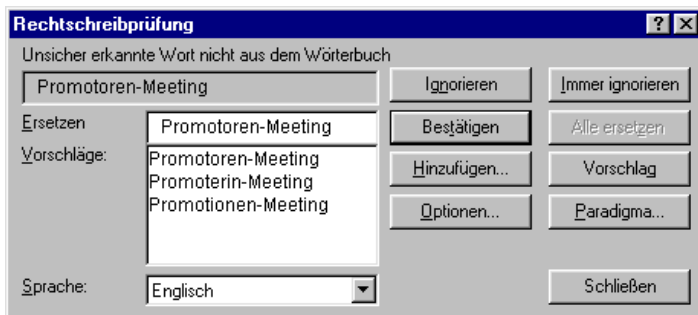


Abbildung 27. Der Dialog Rechtschreibprüfung.

Prüfen der Erkennungsergebnisse:





1. Klicken Sie auf *Prüfen*.
Der Dialog *Rechtschreibprüfung* (Abbildung 27) zeigt mögliche Fehler. Das Zoomfenster zeigt das Originalbild.
2. Wählen Sie eine der folgenden Optionen:
 - *Ignorieren* belässt das Wort, wie es ist.
 - *Immer Ignorieren* belässt alle Vorkommen des Worts.
 - *Hinzufügen* übernimmt das Wort in das Wörterbuch.
 - *Paradigma* zeigt das Paradigma (Gesamtsumme der Wortformen)¹ des Wortes in der Vorschlagsliste.

¹ Für Russisch, Englisch, Deutsch, Französisch, Ukrainisch, Spanisch und Italienisch.


- *Vorschlag...* zeigt Ersatzwörter vor.
- *Optionen...* stellt die Prüfoptionen ein².
- *Schließen* schließt den Dialog.

Zu Fehlern springen

Zum Prüfen und Bearbeiten von Fehlern können Sie die Schaltflächen  und  verwenden, um zum nächsten oder vorherigen Fehler zu springen. Sie können auch die Tastenkombinationen F4 und Umschalt+F4 verwenden.

Anzeige der Erkennungsergebnisse

Standardmäßig werden die Erkennungsergebnisse im Entwurfsmodus angezeigt. D. h. der gesamte Text wird in einer Schrift mit konstanter Größe dargestellt.

Durch Klicken auf  können Sie die Schriftgröße ändern (*Extras>Optionen>Allgemeine Einstellungen>Schriftgröße*).

Speichern der Erkennungsergebnisse

Sie können die Erkennungsergebnisse in einer Datei speichern, direkt an ein Programm senden (MS Word, MS Excel etc.) ohne sie auf der Festplatte zu speichern, in die Zwischenablage kopieren oder per E-Mail versenden. Sie können dabei nur ausgewählte oder alle Seiten exportieren.

Erkennungsergebnisse exportieren:

1. Klicken Sie auf den Pfeil rechts neben der Schaltfläche *5 - Speichern* und wählen Sie im erscheinenden Menü *Export Assistant*.
2. Wählen Sie im erscheinenden Dialog (Abbildung 28), wohin die Erkennungsergebnisse exportiert werden sollen. Sollen nur die im Stapelfenster markierten Seiten exportiert werden, dann aktivieren Sie die entsprechende Option.
3. Klicken Sie auf *Fertigstellen*.

² *Rechtschreibprüfungsoptionen stellen Sie unter Extras>Optionen>Rechtschreibprüfung ein*



Abbildung 28. Export-Assistent.

Hinweis: Wenn Sie an einem neuen (unbenannten) Stapel gearbeitet haben und das Programm abgebrochen wurde (z. B. Stromausfall), dann öffnet FineReader beim nächsten Start diesen Stapel. Nur die Seite, die Sie bearbeitet haben, geht verloren. Alle anderen Seiten werden wiederhergestellt.

Stapel zur späteren Bearbeitung speichern

Möchten Sie in der nächsten Sitzung mit dem gleichen Stapel weiterarbeiten, dann können Sie den Stapel speichern (*Dateimenü, Stapel speichern*). Es wird für jeden Stapel ein Ordner im Verzeichnis \Programme\ABBYY FineReader\Packets angelegt. Der Name des Ordners ist der Stapelname. Folgende Dateien werden im Stapelordner gespeichert:

- Die Stapeldatei (*.pac).
- Bild- (*.TIF) und/oder Textdateien (*.FRF) für jede Seite.
- Blockdateien (*.BOX) (nur für Seiten, die eine Bilddatei haben).
- Dateien mit Scaneinstellungen (*.SCN).
- Stapeleinstellungsdatei (*.OPT).
- Der Formularordner enthält ebenfalls eine Sprachdatenbank *textlang.dat* (für FineReader Professional und Enterprise).

Trainieren neuer Zeichen

FineReader kann Text in nahezu jeder Qualität und jeder Schrift erkennen. Texte mit guter und mittlerer Qualität und Standardschrift können ohne Training erkannt werden. Texte in schlechter Qualität und ungewöhnliche Schriften können nach einem Training erkannt werden.

Damit das Programm den gesamten Text verlässlich erkennen kann, muß es mit einigen Seiten trainiert werden. Ein Benutzeralphabet wird erzeugt, das das System verwendet, um den gesamten Text zu erkennen.

Das Training ist ein spezieller Modus der Erkennung. Das Ergebnis der Trainings ist eine Paarung im OCR-Speicher aus dem Bild des Symbols und seinem Namen. Die Gesamtsumme der Paare, die während dem Training erzeugt wird, wird Alphabet genannt. Dieses Alphabet dient der Erkennung des restlichen Texts.

Das Benutzeralphabet kann nur verwendet werden, um Text in derselben Schrift und Schriftgröße zu erkennen.

Trainieren des Programms

Gehen Sie wie folgt vor, um Texte in schlechter Qualität oder ungewöhnlichen Schriften zu erkennen:

1. Erzeugen Sie ein Alphabet (geben Sie ihm einen Namen).
2. Aktivieren Sie *Erkennung mit Training* (aktivieren Sie die entsprechende Option auf der Seite *Erkennung des Dialogs Optionen*).
3. Wenn Sie ungewöhnliche Schriften oder Sprachen trainieren, die von kyrillischen oder lateinischen Alphabet abweichen (z. B. Griechisch oder Irisch), dann schalten Sie Omnifont aus.
4. Lassen Sie 1 bis 2 Seiten im Trainingsmodus erkennen.

Hinweis: Ist der zu erkennende Text in einer Sprache, die FineReader nicht unterstützt, dann drucken Sie einen beliebigen Text mit der Schrift (nicht kursiv oder fett), die Sie trainieren möchten.

5. Bearbeiten Sie das Alphabet (*Extras>Alphabeteditor...*).
6. Beenden Sie die *Erkennung mit Training*.
7. Lassen Sie den restlichen Text unter Verwendung des Alphabets erkennen.

*Hinweis: Vergessen Sie nicht das Benutzeralphabet auszuschalten, wenn Sie anschließend Texte in anderen Schriften erkennen lassen (wählen Sie Kein als Benutzeralphabet auf der Seite *Erkennung des Dialogs Optionen*).*

Im folgenden werden diese Funktionen ausführlich beschrieben.

Erzeugen und Verwenden eines Alphabets

1. Wählen Sie *Optionen...* im Menü *Extras*.
2. Klicken Sie auf die Schaltfläche *Alphabet* in der Kontrollgruppe *Training* auf der Seite *Erkennung* des Dialogs *Optionen (Extras>Optionen)*.
3. Klicken Sie im Dialog *Alphabet* auf *Neu...*
4. Geben Sie einen Namen für das Alphabet ein.
5. Bevor Sie das Training starten, stellen Sie bitte sicher, daß Ihr Benutzeralphabet als *Aktuelles Benutzeralphabet* eingestellt ist und die Option *Erkennung mit Training* aktiviert ist.

Hinweis: Sie müssen Omnifont nur dann ausschalten, wenn Sie eine ungewöhnliche Schrift oder eine Sprache, die von Kyrillisch und Lateinisch abweicht (z. B. Griechisch oder Irisch), trainieren wollen.



Abbildung 29. Der Dialog *Alphabettraining*

Erkennung mit Training starten:

1. Stellen Sie sicher, daß die Option *Erkennung mit Training* aktiviert ist (*Extras>Optionen>Erkennung*).



2. Klicken Sie auf *Erkennen*. Das Programm beginnt mit der Erkennung und zeigt unsicher erkannte Zeichen im Dialog *Alphabettraining* (Abbildung 29) an.
3. Benennen Sie das Zeichen und klicken Sie in *Trainieren*.
Hinweis: Achten Sie beim Training darauf, daß Sie Großbuchstaben mit Großbuchstaben und Kleinbuchstaben mit Kleinbuchstaben benennen.
4. Das Rechteck im oberen Teil des Dialogs muß **exakt** ein Zeichen enthalten. Falls es nur einen Teil eines Zeichens oder mehr als ein Zeichen enthält, dann verschieben Sie es mit den Schaltflächen *nach links* / *nach rechts*, bis es nur noch ein Zeichen enthält.
5. Ist Ihnen während des Trainings ein Fehler unterlaufen, dann klicken Sie auf *Zurück*. Das Rechteck springt dann ein Zeichen zurück und das zuletzt trainierte Zeichen wird aus dem Alphabet gelöscht.
Hinweis: Zurück funktioniert nur innerhalb eines Worts.
6. Sie können das Benutzeralphabet auch nach dem Training noch bearbeiten. Siehe unten »Alphabet bearbeiten«

Hinweis: Bedenken sie die folgenden Einschränkungen, wenn Sie ein Alphabet trainieren: Das Programm unterscheidet manche Zeichen nicht und verbindet deren Bilder mit ein- und demselben Zeichen. So werden z.B. die drei einfachen Apostrophen glatt ('), links (') und rechts (') im Alphabet alle als glatter Apostroph gespeichert. Deshalb werden Sie in einem erkannten Text niemals linke und rechte Apostrophen sehen, obwohl Sie sie während des Trainings eingegeben haben.

Erkennung mit Training beenden

- Deaktivieren Sie die Option *Erkennung mit Training* auf der Seite *Erkennung* des Dialogs *Optionen* (*Extras>Optionen*), um die *Erkennung mit Training* auszuschalten.

Alphabet bearbeiten

Bevor Sie ein neues Alphabet für die Erkennung verwenden, sollten Sie es überprüfen und nötigenfalls korrigieren. Dieser Schritt minimiert die Anzahl der Erkennungsfehler, die aufgrund von Trainingsfehlern entstehen. Ihr Alphabet darf nur vollständige Zeichen oder Ligaturen (Kombinationen aus zwei oder mehr Zeichen, die nicht getrennt werden können und daher als ein Zeichen trainiert werden) enthalten.

Ein Alphabet bearbeiten:

1. Wählen Sie im Menü *Extras* den Befehl *Alphabeteditor*.
2. Markieren Sie das Alphabet in der Liste und klicken Sie auf *Bearbeiten*, um die Liste der Zeichen zu sehen (Abbildung 30).

3. Löschen Sie fehlerhafte Paare und schließen Sie den Dialog.

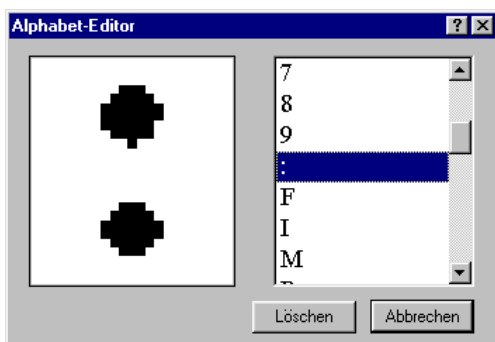


Abbildung 30. Der Alphabeteditor

Trainieren von Ligaturen

Manche Gruppen von zwei oder drei Buchstaben gehen aufgrund ihrer Form ineinander über. Wenn Sie während des Trainings das Rechteck nicht so verschieben können, daß nur exakt ein Zeichen enthalten ist, dann können Sie diese Zeichenkombination wie ein Einzelzeichen trainieren. Solche Kombinationen aus zwei oder drei Zeichen, die nicht zu trennen sind, werden Ligatur genannt. Beispiele sind: ff, ffi, ffl usw. Während des Trainings können Sie die Ligatur aus der Symbolliste auswählen. Ist die gewünschte Ligatur nicht vorhanden, dann können Sie eine neue erzeugen. Ein Alphabet kann über 100 neue Symbole aufnehmen, trotzdem sollten Sie nicht zu viele Ligaturen definieren, weil darunter die Erkennungsqualität leiden kann.

Hinweis: Trainieren Sie Zeichen, die nicht über die Tastatur eingegeben werden können, dann verwenden Sie Kombinationen aus zwei Zeichen (z. B. #A) oder die Zeichentabelle, um die Symbole zu benennen. Antworten Sie mit Ja, wenn das Programm Sie fragt, ob Sie eine neue Ligatur erstellen wollen.

Sprachen in FineReader

Um Texte in Russisch, Englisch, Deutsch, Französisch, Ukrainisch, Spanisch, Italienisch und beliebigen anderen Sprachen aus der Liste der unterstützten Sprachen erkennen zu können, wählen Sie die gewünschte Sprache in der Erkennungs-Werkzeugleiste. Falls Sie die gewünschte Sprache nicht finden, wählen Sie den Eintrag *Andere Sprache...* und erzeugen Sie eine neue Kombination (nähere Informationen finden Sie unter »Sprachen für die Erkennung« auf Seite 60).

Benutzerdefinierte Sprachen und Sprachgruppen (für FineReader 4.0 Professional und Enterprise)

FineReader unterstützt 40 Sprachen. Das reicht für den meisten Benutzer aus (die Liste der unterstützten Sprachen finden Sie im Anhang auf Seite 60). In manchen Fällen wird dennoch eine neue Sprache benötigt.

Wann wird eine neue Sprache benötigt:

- Um eine Textdatei oder einen regulären Ausdruck als Wörterbuch verwenden zu können. Hauptanwendungsfall ist die Erzeugung einer neuen Sprache zur Erkennung von Formularfeldern.
- Um Text in einer nicht unterstützten Sprache zu erkennen.
- Wenn häufig bestimmte Sprachkombinationen verwendet werden z.B. Englisch-Spanisch.

Sie können ein Wörterbuch für eine neue Sprache hinzufügen, es muß sich dabei um eine Textdatei in ANSI-Codierung handeln. In FineReader Enterprise können Sie auch eine Datenbankspalte verwenden.

Öffnen Sie den *Spracheditor*, um eine neue Sprache oder Sprachgruppe zu erzeugen (*Extras>Spracheditor*).

Ein Beispiel zur Erzeugung einer neuen Sprache finden Sie im Anhang auf Seite 58.

Stapelverarbeitung im Netzwerk (für FineReader 4.0 Professional und Enterprise)

Sie können einen Stapel auf mehreren Computern gleichzeitig bearbeiten. Das verringert die Bearbeitungszeiten für große Dokumente erheblich. Jeder beteiligte Computer benötigt dazu eine eigene Programmlizenz.

Das Einlesen des Texts in den Computer setzt sich aus mehreren Vorgängen zusammen: Scannen, Seitenanalyse, Erkennung, Bearbeiten und Export. Wird ein Stapel mit mehreren Computern gleichzeitig bearbeitet, dann können diese Aufgaben auf die Computer verteilt werden. Zum Beispiel könnte ein Computer mit Scanner und Vorlageneinzug das Scannen übernehmen, während ein zweiter Computer (oder mehrere) die Erkennung, Bearbeitung und den Export der Erkennungsergebnisse durchführt.

Wird die Erkennung des Stapels auf mehrere Computer verteilt, dann verteilen die Computer die Arbeit automatisch, solange es nicht erkannte Seiten im Stapel gibt.

Wie wird ein Stapel im Netzwerk verarbeitet?

1. Erzeugen Sie einen neuen Stapel (sofern nötig) und speichern Sie ihn an einem Ort, auf den alle beteiligten Computer Lese- und Schreibzugriff haben.
2. Schließen Sie den Stapel, um die Optionen zu speichern.
3. Starten Sie FineReader und öffnen Sie diesen Stapel auf allen beteiligten Computern.
4. Starten Sie den Scanvorgang auf dem Computer mit angeschlossenem Scanner und Vorlageneinzug.
5. Starten Sie die Erkennung auf allen Computern, die an der Erkennung mitarbeiten sollen, im Read&Wait Modus (*Scan&Read Menü, Alle nicht erkannten Seiten erkennen und warten*).
6. Sobald die ersten Seiten erkannt wurden, können Sie diese auf einem freien Computer (oder einem Computer, auf dem die Erkennung in Hintergrund läuft) bearbeiten, ohne auf die verbleibenden Seiten warten zu müssen.
7. Die erkannten (oder alle markierten) Seiten können in eine Datei exportiert oder an ein anderes Programm gesandt werden.

Hinweis: Damit Sie die Stapelverarbeitung besser überwachen können, klicken Sie mit der rechten Maustaste in das Stapelfenster und wählen Sie aus dem Kontextmenü Seitenparameter zeigen. Diese Ansicht des Stapelfensters erleichtert Ihnen die Arbeit mit dem Stapel im Netzwerk. (Abbildung 15, Seite 36).

Read&Wait Modus

Im *Read&Wait* Modus wird die Erkennung automatisch fortgesetzt, wenn neue, noch nicht erkannte Seiten im Stapel abgelegt werden.

Erkennung im Read&Wait Modus starten:

- Klicken Sie auf den Pfeil rechts neben der Schaltfläche *Scan&Read* und wählen Sie aus dem erscheinenden Menü *Alle nicht erkannten Seiten erkennen und warten*.

Anhang

Verbessern der Erkennungsqualität: Tips und Tricks

Grundzüge

Enthält ein Text viele nicht erkannte Zeichen, dann könnte eine schlechte Vorlage der Grund sein. Stellen Sie die korrekte Helligkeit, Texttyp und Auflösung ein und scannen Sie erneut. Ist die Vorlage verschmutzt, versuchen Sie den Befehl *Säubern*.

Eine schlechte Texterkennung kann auch durch falsche Erkennungsparameter verursacht werden, wenn z.B. die falsche oder zu viele Sprachen ausgewählt sind. Erkennungsfehler in jeder Zeile können entstehen, wenn ein englischer Text mit der Sprache Englisch-Russisch erkannt wird.

Das Wörterbuch erweitern

Das Programm verwendet das Wörterbuch während der Erkennung. Wenn Sie häufig vorkommende Worte und Abkürzungen in das Wörterbuch aufnehmen, verbessert sich die Erkennungsqualität. Sie können die Worte in das Rechtschreibwörterbuch einfügen (*Extras* Menü, *Wörterbücher anzeigen*) oder Sie können eine neue Sprache erzeugen und die Worte als Textdatei (in ANSI-Codierung) oder Datenbankspalte anfügen.

Weitere Informationen finden Sie im Abschnitt »Unterstützte Sprachen« auf Seite 60 und in der Online-Hilfe.

Training

Wenn Sie große Dokumente mit ungewöhnlichen Schriften oder in schlechter Qualität erkennen müssen, dann können Sie das Programm mit 2 bis 3 Seiten trainieren und anschließend den Text zuverlässig erkennen lassen. Das Training findet in einem besonderen Betriebsmodus während der Erkennung statt. Das Ergebnis ist ein Benutzeralphabet, daß während der Erkennung verwendet wird. Weitere Informationen finden Sie im Abschnitt »Trainieren neuer Zeichen« auf Seite 50.

Erzeugen neuer Sprachen

Um die Erkennung zu beschleunigen und die Genauigkeit zu erhöhen, können Sie neue Sprachen erzeugen und Wörterbücher hinzufügen. Sie können Textdateien in ANSI-Codierung als Wörterbuch verwenden.

Um zum Beispiel eine Tabellenspalte zu erkennen, die einen begrenzten Wortbereich hat, können Sie eine neue Sprache erzeugen und ein Wörterbuch anfügen, daß nur die Wörter enthält, die in der Spalte auftreten können.

Beispiel zur Erkennung der altrussischen Sprache

Gehen Sie wie folgt vor, um altrussische Texte zu erkennen:

1. Erzeugen Sie eine Kopie der russischen Sprache.
2. Stellen Sie die erzeugte Sprache als Erkennungssprache ein.
3. Erzeugen Sie ein neues, leeres Alphabet, verbinden Sie es und starten Sie die Erkennung mit Training.
4. Erkennen Sie 1 bis 2 Seiten altrussischen Text mit der *Erkennung mit Training*.
5. Prüfen und korrigieren Sie das Alphabet.
6. Schalten Sie die *Erkennung mit Training* aus (aber nicht das erzeugte Alphabet).
7. Führen Sie die Erkennung durch.
8. Ersetzen Sie im erkannten Text alle Ligaturen, die das Zeichen * enthalten, durch die korrekten Zeichen (siehe nächste Seite).

Später erklären wir ausführlicher, wie eine Sprache erstellt und verbunden wird.

Erzeugen Sie eine neue Sprache:

1. Wählen Sie im *Extras* Menü den Eintrag *Spracheditor...*
2. Klicken Sie im angezeigten Dialog auf *Neu...*
3. Wählen Sie die Option *Eine vorhandene Sprache kopieren und bearbeiten*.
4. Wählen Sie *Russisch* in der Liste und klicken Sie auf *Weiter*.
5. Geben Sie den Namen für die Sprache ein: Altrussisch.
6. Sie sehen russische Zeichen im Dialog. Fügen Sie das Zeichen »*« (oder ein anderes Zeichen, das in der kyrillischen Codeseite vorhanden ist und per Tastatur eingegeben werden kann) ein. Dieses Zeichen wird verwendet, um die Ligaturen zu erzeugen. Damit das Zeichen während der Rechtschreibprüfung ignoriert wird, markieren Sie es als *Ignoriert* (klicken Sie auf die Schaltfläche rechts neben *Ignoriert* und wählen Sie es aus der Tabelle.)

Erzeugen Sie ein neues Alphabet:

1. Wählen Sie aus dem Menü *Extras* den Eintrag *Alphabeteditor*.
2. Klicken Sie im Dialog *Alphabet* auf *Neu...*
3. Benennen Sie das Alphabet, z.B. Altrussisch.

Verbinden Sie das Alphabet und schalten Sie »Erkennung mit Training« ein

1. Wählen Sie auf der Seite *Erkennung* des Dialogs *Optionen* (*Extras>Optionen*) in der Gruppe *Training* in der Liste *Aktuelles Benutzeralphabet* das Alphabet aus, das Sie eben erzeugt haben und jetzt trainieren möchten.
2. Aktivieren Sie *Erkennung mit Training* und *Omnifont*¹ verwenden.

Stellen Sie die Erkennungssprache ein:

1. Wählen Sie den Eintrag *Andere Sprache* im Listenfeld *Erkennungssprache* der Erkennungs-Werkzeugleiste.
2. Wählen Sie im erscheinenden Dialog Russisch und Altrussisch². Starten Sie die Erkennung.

Erkennung mit Training

Während der Erkennung mit Training zeigt das Programm den Dialog *Alphabettraining* an. Im oberen Bereich des Dialogs sehen Sie einen Ausschnitt der aktuellen Textzeile, in dem das unsicher erkannte Zeichen von einem Rechteck umschlossen ist (Abbildung 29, Seite 52).

Geben Sie das Zeichen mit der Tastatur ein und stellen Sie als Sprache Altrussisch ein. Wenn das Zeichen »i« angezeigt wird, geben Sie als Symbol »*й « an, wenn »v« angezeigt wird, die Ligatur »*и«, wenn »Θ« angezeigt wird, die Ligatur »*ф« und wenn » Ъ « angezeigt, wird die Ligatur »*ѣ«.

Wenn das Training abgeschlossen ist, wird das Alphabet erzeugt. Bevor Sie es zur Erkennung verwenden, sollten Sie es prüfen und nötigenfalls korrigieren (*Extras>Alphabeteditor*) (Abbildung 30, Seite 53).

Nachdem das Alphabet auf ein bis zwei Seiten trainiert wurde, können Sie den Rest des Texts erkennen und das Alphabet verbinden. Schalten Sie dann den Trainingsmodus aus und beginnen Sie die Erkennung mit der Sprachkombination Russisch-Altrussisch.

¹ *Omnifont wird nur in seltenen Fällen deaktiviert – wenn z. B. eine ungewöhnliche Schrift trainiert werden soll.*

² *Die Russische Sprache wird verbunden, weil das Wörterbuch die Erkennungsqualität verbessert und eine Rechtschreibprüfung ermöglicht.*

Ersetzen im Editor

Nachdem der Text erkannt wurde, müssen die Ligaturen durch die korrekten Zeichen ersetzt werden. Das heißt die Ligatur *ϕ wird durch Θ ersetzt, die Ligatur *н durch ъ und so weiter.

Für FineReader Enterprise Administratoren

Das Administrator Handbuch für FineReader Enterprise finden Sie als PDF-Datei in Ordner *Guide* auf der CD-ROM.

Von FineReader unterstützte Bildformate

BMP:	SW, Graustufen, Farbe
PCX, DCX:	SW, Graustufen, Farbe (16 und 256 Farben)
JPEG:	Graustufen, Farbe
TIFF:	SW– unpacked, CCITT3, CCITT3FAX, CCITT4, Packbits; Graustufen – unpacked, Packbits, JPEG; Farbe – unpacked, Packbits, JPEG

Unterstützte Sprachen

Afrikaans, Albanisch, Baskisch, Bretonisch, Bulgarisch, Byelorussisch, Dänisch, Deutsch, Englisch, Estnisch, Fidschianisch, Finnisch, Französisch, Griechisch, Hawaïisch, Indonesisch, Irisch, Isländisch, Italienisch, Kabardianisch, Katalanisch, Krim Tatar, Kroatisch, Lateinisch, Lettisch, Litauisch, Makedonisch, Maori, Moldavisches, Niederländisch (Standard, Belgien), Norwegisch (Bokmal, Nynorsk), Ossetisch, Polnisch, Portugiesisch (Standard, Brasilien), Rätoromanisch, Rumänisch, Russisch, Samoanisch, Schwedisch, Serbisch (Kyrillisch), Slowakisch, Slowenisch, Spanisch, Suaheli, Tagalog, Tartar, Tschechisch, Tschetschenisch, Türkisch, Ungarisch, Ukrainisch

Wörterbücher und Benutzeralphabete von FineReader 3.0 verwenden

Benutzerwörterbücher

Konvertieren Sie nach der Installation die alten Wörterbücher in das neue Format:

1. Wählen Sie *Extras>Optionen*
2. Klicken Sie auf der Seite *Rechtschreibprüfung* des Dialogs *Optionen* auf *Zum Wörterbuch hinzufügen...*
3. Geben Sie den Pfad zur alten FineReader Version an (den Ordner, der Fine32.exe enthält). Im oberen Bereich sehen Sie eine Liste der Sprachen mit Benutzerwörterbüchern.
4. Markieren Sie die gewünschten Sprachen und klicken Sie in *Hinzufügen*.

Benutzeralphabete

Kopieren Sie nach der Installation der neuen Version die Benutzeralphabete (*.ptn) aus dem Ordner der vorherigen Version in den Ordner der neuen Version
\\Programme\\ABBYY FineReader.

Tastenkombinationen (Shortcuts) in FineReader

	Funktion:	Tastenkombination:
Datei	Neuen Stapel erzeugen	Strg+N
	Stapel oder Bild öffnen, Bild zu Stapel hinzufügen	Strg+O
	Export Assistent starten	Strg+D
	In Datei speichern	Alt+Umschalt+S
	An MS Word senden	Alt+Umschalt+W
	An MS Excel senden	Alt+Umschalt+E
	An Word Pro senden	Alt+Umschalt+D
	An WordPerfect senden	Alt+Umschalt+P
	Per E-Mail senden	Alt+Umschalt+M
	An die Zwischenablage senden	Alt+Umschalt+C
Bearbeiten	Letzte Aktion widerrufen	Alt+Rück / Strg+Z
	Letzte Aktion wiederholen	Alt+Umschalt+Rück / Strg+Y
	Markierung ausschneiden und in Zwischenablage	Strg+X / Strg+Einf
	Markierung in die Zwischenablage kopieren	Strg+Einf / Strg+C
	Zwischenablage einsetzen	Strg+V / Umschalt+Einf
	Aktiven Block, Markierung, markierte Seiten löschen	Entf
	Seiten löschen	Alt+Entf
	Gesamten Text im Editor, alle Seiten im Stapel, alle Blöcke im aktuellen Bild auswählen	Strg+A
	Blocktyp auf «Text» setzen	Strg+1
	Blocktyp auf «Tabelle» setzen	Strg+2
	Blocktyp auf «Abbildung» setzen	Strg+3
	Zellen zusammenfassen	Strg+ENTER
	Zellen teilen	Strg+Umschalt+ENTER
	Suchtext im erkannten Text suchen	Alt+F3 / Strg+F
	Weitersuchen	F3
	Suchtext im erkannten Text suchen und ersetzen	Strg+H
	Dialog «Schrift» öffnen	Strg+Umschalt+F
	Dialog «Formatierung» öffnen	Alt+Umschalt+F
	Dialog zum Setzen der Sprache für die Markierung öffnen	Strg+L
	Markierung fett	Strg+B
Ansicht	Markierung kursiv	Strg+I
	Markierung unterstrichen	Strg+U
	Gesamtes Bild anzeigen	Strg+Umschalt+1
	Bildbreite	Strg+Umschalt+2
	Bildhöhe	Strg+Umschalt+3

	Funktion:	Tastenkombination:
	Vergrößern	Strg+Aufwärtspfeil
	Verkleinern	Strg+Abwärtspfeil
	Nächste Seite öffnen	Alt+Umschalt+N / Strg+ <Num>+
	Vorherige Seite öffnen	Alt+Umschalt+P / Strg+ <Num>-
	Seite mit Nummer öffnen...	Strg+G
	Seitenliste aktualisieren	F5
	Eigenschaften des aktiven Bilds (Blocks)	Alt+ENTER
Bild	Textblockwerkzeug	Alt+1
	Werkzeug zum Erweitern eines Blocks	Alt+2
	Werkzeug zum Verkleinern eines Blocks	Alt+3
	Werkzeug zum Bewegen eines Blocks	Alt+4
	Werkzeug zum Umnummerieren eines Blocks	Alt+5
	Werkzeug zum Löschen von Blöcken	Alt+6
	Werkzeug für vertikale Trennlinien (Tabellen)	Alt+Umschalt+1
	Werkzeug für horizontale Trennlinie (Tabellen)	Alt+Umschalt+2
	Werkzeug zum Löschen von Linien (Tabellen)	Alt+Umschalt+3
Scan&Read	Scannen und erkennen	F9
	Scannen und erkennen mehrerer Seiten	Umschalt+F9
	Bild öffnen und erkennen	Alt+Umschalt+O
	Scan&Read Assistent	Strg+J
	Scannen	Strg+K
	Scannen mehrerer Seiten	Strg+Umschalt+K
	Scannen stoppen	Alt+K
	Blöcke auf den ausgewählten Seiten markieren	Strg+E
	Blöcke auf allen Seiten ohne Blöcke markieren	Strg+Umschalt+E
	Ausgewählte Seiten erkennen	Strg+R
	Alle nicht erkannten Seiten erkennen	Strg+Umschalt+R
	Alle nicht erkannten Seiten im Hintergrund erkennen	Alt+R
	Rechtschreibprüfung	F7
	Nächstes unsicher erkanntes Wort	F4 / Strg+Umschalt+N
	Vorheriges unsicher erkanntes Wort	Umschalt+F4 / Strg+Umschalt+P
	Fehlermarke löschen	Strg+Q
	Ausgewähltes Wort oder Wortkombination übersetzen	Strg+Umschalt+T
	Dialog «Spracheditor» öffnen	Strg+Umschalt+L
	Dialog «Scannerparameter» öffnen	Strg+Umschalt+S
	Dialog «Optionen» öffnen	Strg+Umschalt+O
Fenster	Fenster anordnen	Strg+W
Hilfe	Direkthilfe für Objekte aktivieren (für Menü und Dialogbestandteile)	Umschalt+F1

Anhang

	Funktion:	Tastenkombination:
Andere	Ausgewählte Stapelseite öffnen	ENTER
Befehle	Stapelfenster aktivieren	Alt+0
	Bildfenster aktivieren	Alt+8
	Textfenster aktivieren	Alt+9
	Von einer Zelle zu einer anderen Zelle bewegen	Alt+ Linkspfeil, Rechtspfeil, Abwärtspfeil oder Aufwärtspfeil



INCREDIBLE POWER & AMAZING ACCURACY

For a long time people dreamed of how to teach machine *human knowledge*, e.g., create a machine able to read. During the latest 50 years this dream started coming true. Recognition is one of the **most** favorable applications of the AI¹. There exists a solution

that is close to human ability to read: it uses recognition principles of living systems. This is the **methods** technology of adaptive integral purposeful perception.

Main differences between Standard, Professional, Handprint.

	Standard	Professional	Handprint
Types of text recognized	Printed		Printed, hand-printed
Recognition of bar codes		Yes	Yes
System can be trained to recognize new symbols	Yes	Yes	Yes
Recognition of texts in columns, including graphics and tables. Saves the recognized document in RTF format	Yes	Yes	Yes
Use with Internet: Saves the recognized document in HTML format	Yes	Yes	Yes
Languages Supported			
Built-in spell checker	Yes	Yes	Yes
Recognition of multilingual documents	Yes	Yes	Yes
Expandable to include new languages		Yes	Yes
Recognition of tables			
Recognition of tables, saving the result in RTF, CSV, XLS, DBF	Yes	Yes	Yes
Manual and automatic segmentation of tables	Yes	Yes	Yes
Post-editing of recognized tables	Yes	Yes	Yes
Recognition of forms			
Recognition of forms and saving the result in DBF		Yes	Yes
Design of form patterns		Yes	Yes
Automatic positioning of forms using reference points		Yes	Yes
Export to databases through ODBC 32			Yes
Automatic selection of form pattern			Yes
Batch processing of documents and advanced functions			
Distributed processing of documents over a network		Yes	Yes
Integration with other applications through API, OLE-Automation, Active X.	Available in SDK ²		

1 - AI - artificial intelligence

2 - SDK - software developer kit



- ☐ Done
☐ History

Complete the application on a typewriter. Do not type the letters on the field names!
The line spacing between fields is 1, 2 and 3 typewriter intervals.

JOB APPLICATION

1. Last Name											
First Name											
2. Date of Birth	Day	Month (in Words)			Year						
							M	F			
3. Sex (tick the appropriate square with "x")											
4. The address where you live now											
Post or Zip Code	Town/City										
Street											
House Number	Apartment		Area Code		Phone						
			5. Telephone Number								
6. Passport	Series	Number		Day	Month (in Words)		Year				
		N		Issued							
7. Application Date	Day	Month (in Words)			Year						
Signature											

Place
photo
here

System for bulk input of documents - FINE READER 4.0 PROFESSIONAL



- ☐ Done
☐ History

Complete the application on a typewriter. Do not type the letters on the field names!
The line spacing between fields is 1, 2 and 3 typewriter intervals.

JOB APPLICATION

1. Last Name Galeev

First Name Valery

2. Date of Birth Day Month (in Words) Year
27 May 1960

3. Sex (tick the appropriate square with "x") M F
X

4. The address where you live now

Post or Zip Code Town/City

120931 Moscow

Street

Green

House Number Apartment:

120 31

Area Code Phone

7095 1258000

6. Passport

Series

Number

NSFII N 1231

Day

Month (in Words)

Year

Issued 12 March 1990

7. Application Date Day Month (in Words) Year
27 January 1996

Signature



System for bulk input of documents - FINE READER 4.0 PROFESSIONAL



☒ Done

☒ History

Complete the application on a typewriter. Do not type the letters on the field names!
The line spacing between fields is 1, 2 and 3 typewriter intervals.

JOB APPLICATION

1. Last Name Savinova

First Name Olga

2. Date of Birth Day Month (in Words) Year
4 April 1964

M F

3. Sex (tick the appropriate square with 'x')

X

4. The address where you live now

Post or Zip Code Town/City

141097 Moscow

Street

High

House Number Apartment

67

51

Area Code Phone

7095 267678

5. Telephone Number

6. Passport

Series

Number

BBCAI N 7098

Issued

Day

Month (in Words)

Year

11

October

1984

7. Application Date Day Month (in Words) Year
14 January 1995

Signature



System for bulk input of documents - FINE READER 4.0 PROFESSIONAL