

Background

package

# READ IRIS OCR

yananár

real

text

*Save Time, No More Retyping!*

typed

IRIS  
paper

Druckqualität

trainable

učici modul

OCR

matrix

обработка

caractères accentués

omnifont

easy

recognition

fax

neue Symbole und Schriftarten

CR

IMAGE RECOGNITION  
INTEGRATED  
SYSTEMS

JUST READ IT!



---

# **Readiris™** **Benutzerhandbuch**

**System zur optischen Zeichenerkennung**

**für IBM Personal Computer, PS/2  
und kompatible Rechner**

**Image Recognition Integrated Systems**

Readiris ist ein Warenzeichen von Image Recognition Integrated Systems.

Geringfügige Abweichungen des Handbuches gegenüber der Software sind jederzeit ohne vorherige Ankündigung möglich. Image Recognition Integrated Systems kann nicht haftbar gemacht werden für den Inhalt dieses Handbuches. Die in diesem Handbuch beschriebene Software unterliegt einem Lizenzvertrag. Die Software darf nur im Rahmen der Bestimmungen des Vertrages benutzt oder kopiert werden.

Proximity ist eingetragenes Warenzeichen von Proximity Technology Inc. IBM ist eingetragenes Warenzeichen von International Business Machines Corp. Windows, PaintBrush und Microsoft sind eingetragene Warenzeichen von Microsoft Corp. Jede in einem Menü vorkommende Format- und Scannerbezeichnung ist Warenzeichen des jeweiligen Herstellers.

THE PROXIMITY/MERRIAM-WEBSTER LINGUIBASE

Copyright 1984, alle Rechte vorbehalten

Proximity Technology Inc.

Copyright 1984, alle Rechte vorbehalten

Merriam-Webster Inc.

THE PROXIMITY/MERRIAM-WEBSTER LINGUIBASE

Copyright 1984, alle Rechte vorbehalten

Proximity Technology Inc.

Copyright 1984, alle Rechte vorbehalten

William Collins Sons & Co. Ltd.

Die Orthographieprüfung dieses Produktes basiert auf Proximity  
Linguistic Technology

© Copyright, Image Recognition Integrated Systems, 1994

Alle Rechte vorbehalten

Dokumenten-Nr.: READIRIS-9502-GRM

# Anhalt

## Vor dem ersten Aufruf

---

Systemvoraussetzungen .....	7
Vorbereitung von Readiris .....	7
Die Installation von der Connect-Funktion .....	8

## Schneller Start

---

Aufruf von Readiris .....	9
Definieren der Standard-Konfiguration .....	10
Zeichenerkennung mit einem Bilddatei .....	10
Einstellungen .....	10
Grafikdatei öffnen .....	11
Vergrößern des Bildes .....	12
Definieren eines Textfensters und eines Grafikfensters .....	13
Veränderung eines Textfensters .....	13
Zeichenerkennung .....	14
Anzeige des Ergebnisses .....	15
OCR mit einem neu gescannten Bild .....	15
Einstellungen .....	16
Scannen mit einem Flachbett-Scanner .....	16
Scannen mit einem Handy Scanner .....	16
Texterkennung ausführen .....	17

## Mehr über Readiris

---

Definition der Standardeinstellungen .....	19
Konfiguration Ihres Flachbett-Scanners .....	20
Konfiguration Ihres Handy Scanners .....	21
Textformat .....	22
Grafikformat .....	23
Sprache .....	24
Abspeichern der Standardeinstellungen .....	24
Speichern und Neuladen von Parametern .....	24
Abspeichern der Einstellungen .....	24
Laden von Parametern .....	25
Alles über die Parametereinstellungen .....	26
Dokument .....	27

## II

Verarbeitung	27
Dateien	29
Scannen mit einem Flachbett Scanner	30
Scannen mit einem Handy Scanner	30
Vergrößern und Verkleinern des gescannten Bildes	31
Überprüfen der Qualität des gescannten Bildes	32
Fensterwerkzeuge	32
Erstellen eines rechteckigen Fensters	33
Erstellen eines polygonalen Fensters	33
Auswahl eines Fensters	34
Verändern eines Fensters	34
Verschieben eines Fensters	34
Löschen eines Fensters	35
Zeichenerkennung	35
Lernen von Schriftarten	36
OK	37
Weiter	38
Bestätigen	38
Rückgängig	38
Löschen	38
Beenden	39
Lernen mit Hilfe des Anwenderlexikons	39
On-Line Hilfe	40
Tabellen erkennen	40
Automatische Erkennung	41
Die Funktion Connect	42

### **Befehle von Readiris**

---

Datei	45
TIFF-Datei laden	45
Information	45
Quelle selektionieren	45
Erwerben	46
Layout speichern	46
Layout laden	46
Parameter speichern	46
Parameter laden	46
Standard Parameter speichern	46
Ende	46
Bearbeiten	47
Rückgängig	47
Löschen	47
Kopieren	47
Einfügen	47
Neu	47
Alle selektieren	48
Voreinstellungen	48
Scanner	48

Textformat .....	48
Grafikformat .....	48
Sprache .....	48
Funktionsleiste .....	49
Ansicht .....	49
Passend .....	49
50% Verkleinert .....	49
100 % Größe .....	49
200% Vergrößert .....	49
Alle Fenster .....	49
Selektierte Typ .....	50
Sort .....	50
Sort Ansicht .....	50
Sort Neu .....	50
Rotieren .....	50
Werkzeuge .....	50
Koordinaten .....	51
Verarbeitung .....	51
Parametereinstellungen .....	51
Scannen .....	51
OCR .....	51
Interaktives Lernen .....	51
Seitenanalyse .....	52
Anzeige nach Scan .....	52
Parametereinstellungen .....	52
Zeichenabstand .....	52
Größe .....	52
Schriftart .....	53
Seite(n) zu lesen .....	53
Seitenanalyse .....	53
Anzeige nach Scan .....	53
Interaktives Lernen .....	53
Starke Kontrolle .....	54
Eingabe .....	54
Ausgabe .....	54
Layout .....	54
Diktionär .....	54
Scan .....	54
Schließen .....	55
Hilfe .....	55
Inhalt .....	55
Über Readiris .....	55

## **Mehr über OCR**

---

Der OCR-Vorgang .....	57
Anwendungsgebiete für Readiris .....	59
Linguistischer Kontext .....	59
Isolierte Zeichen .....	60

# IV

Zeichengröße .....	60
Nicht lesbare Dokumente .....	60



# Einleitung

Readiris ist ein OCR-System (OCR oder Optical Character Recognition, optische Zeichenerkennung), welches einen PC in Verbindung mit einem Scanner in ein leistungsstarkes Lesesystem verwandelt.

OCR ist die schnellste Methode, um gedruckte und maschinengeschriebene Textvorlagen in ein Format zu bringen, das durch Textverarbeitungs- und Desktop Publishing-Programme erkannt wird.

Readiris verwendet die Vorteile der Windows-Umgebung: Befehlsmenüs, Dialogfenster, Sinnbilder usw. Sie können Readiris aufstarten in jeder Windows-Anwendung!





# Über Readiris

Installation und Einsatz von Readiris sind benutzerfreundlich gestaltet. Die Installation dauert nur zwei Minuten; mit Readiris arbeiten gelingt schon nach fünf Minuten! Meistens brauchen Sie nur eine Maus um Readiris zu bedienen.

Mit Readiris erhält der Anwender hochentwickelte OCR-Technik zu einem überaus günstigen Preis.

Readiris ist ein selbstlernendes System, es kann automatisch neue Zeichen lernen. Möglich wird dies durch eine Analyse des linguistischen Kontextes. Hierzu benutzt Readiris eine linguistische Datenbank, Lexika, um neue Lösungen bei schwierigen Fällen zu finden.

Ferner ist Readiris in der Lage, Bilder aus der gescannten Vorlage in Bilddateien für die Weiterbearbeitung mit einem Bildverarbeitungsprogramm zu erzeugen.



# Über dieses Handbuch

Das Handbuch beschreibt den Einsatz von Readiris. Das erste Kapitel, "Vor dem ersten Aufruf", befaßt sich mit den Hard- und Software-Voraussetzungen sowie mit der Installation.

Das zweite Kapitel, "Schneller Start", stellt die Arbeit mit Readiris im Überblick dar und enthält eine Einführung in das Grundkonzept der Software.

Um nun weiter in die einzelnen Leistungsmerkmale von Readiris einzusteigen, lesen Sie im dritten Kapitel, "Mehr über Readiris", weiter.

Alle Menüs und Befehle der Software werden im vierten Kapitel, "Befehle von Readiris", beschrieben.

Das letzte Kapitel, "Mehr über OCR", enthält eine Kurzbeschreibung von OCR-Verfahren und -Parametern. Hier geht es um den Einfluß der verschiedenen Parameter auf die Leistung eines OCR-Systems im allgemeinen und Readiris im besonderen.

Zusätzliche Informationen, die dieses Handbuch nicht enthält, befinden sich in der Datei README.TXT. Daher sollten Sie den Text dieser Datei vor Verwendung des Handbuches und der Software lesen.

Readiris unterstützt die meisten handelsüblichen Scanner. Informationen über Scanner befinden sich in der Datei READSCAN.TXT.



# 1 Vor dem ersten Aufruf

Dieses Kapitel enthält alles über die Installation von Readiris.

Vor dem ersten Aufruf von Readiris sollten Sie mit Ihrem Computer sowie mit DOS und Windows vertraut sein.

---

## Systemvoraussetzungen

Zur Installation und zum Einsatz von Readiris wird benötigt:

- 386 IBM PC, PS/2 oder 100%ig IBM-kompatibel. Multitasking wird unterstützt, sogar auf einen 386er!
- mindestens 4 MB RAM. Es ist möglich das Sie für die Connect-Funktion zusätzlichen Speicher brauchen.
- Das Betriebssystem MS-DOS, Version 3.3 oder höher oder höher. MS-DOS 6.x wird empfohlen.
- Windows, Version 3.1 oder höher. Windows darf aber nicht in "Standard-Modus" aufgestartet werden.

---

## Vorbereitung von Readiris

Fertigen Sie vorher eine Sicherheitskopie Ihrer Readiris Disketten an.

Installieren Sie Readiris wie folgt:

- In dem Program Manager van Windows wählen Sie "Ausführen" in dem "Datei" Menü. Tippen Sie ein **a:setup** und klicken Sie "OK".

- Die Installation fängt jetzt an. Wählen Sie die gewünschte Sprache und folgen Sie die Anweisungen auf dem Schirm.

Nach eine Weile müssen Sie die 2. und 3. Readiris-Diskette ins Laufwerk einlegen.

Die Programm-Gruppe und die Symbole werden von dem Installationsprogramm hergestellt.

---

## **Die Installation von der Connect-Funktion**

Readiris muß korrekt installiert sein bevor Sie die Connect Funktion installieren.

Starten Sie Readiris, wählen Sie Ihren Scanner und speichern Sie die Standardeinstellungen bevor Sie die Connenct-Funktion installieren. Wie Sie die Connect-Funktion installieren sollen, finden Sie in Abschnitt 3, "Mehr über Readiris".

In WordPerfect, Word und Excel müssen Sie ein Macro installieren um die Connect-Funktion zu gebrauchen. Das on-line Help-System wird Ihnen helfen bei der Installation. Wenn Sie Readiris installieren, bekommen Sie auch Informationen über die Installation von der Connect-Funktion.

Für jede andere Windows-Anwendung gibt es der Connect Task Server. Die Installation ist automatisch.

## 2 Schneller Start

Die Arbeitsweise von Readiris wird am besten durch einfaches Ausprobieren deutlich. Die folgenden Ausführungen richten sich an denjenigen, der sofort mit Readiris arbeiten will. Eine bereits gescannte Datei, GERMAN.TIF, befindet sich auf der Readiris Diskette, um zunächst einmal die Bedienung des OCR-Programmes ohne Scanner zu erlernen.

---

### Aufruf von Readiris

Vergewissen Sie sich davon daß Sie sich in den Program Manager von Windows befinden. Doppelklicken Sie das Readiris-Symbol.







Nach einige Augenblicken sehen Sie das Readiris-Schirm.

Das Readiris-Fenster enthält nicht nur Menüs, sondern auch einen **Funktionsleiste**. Für die meiste Befehlen brauchen Sie nur auf diese Symbole zu klicken. Um die Funktionsleiste zu aktivieren, klicken Sie auf "Funktionsleiste" unter das "Voreistellungen" Menü.

---

## Definieren der Standard-Konfiguration

- 1 Klicken Sie auf den Befehl "Voreinstellungen". Es erscheint ein Untermenü mit folgenden Befehlen: "Scanner", "Textformat", "Grafikformat" und "Sprache".
- 2 Klicken Sie jeden Befehl an und treffen Sie jeweils die gewünschte Auswahl. Dieser Schritt kann auch übergangen werden. Readiris verwendet dann seine Standardeinstellungen.
- 3 Bewegen Sie den Mauszeiger auf den Befehl "Datei" und klicken Sie mit der linken Maustaste. Klicken Sie auf "Standard Parameter Speichern". Klicken Sie auf "OK".



Readiris ist nun bereit. Wir empfehlen, zunächst das im folgenden noch beschriebene Beispiel durchzugehen, auch dann, wenn ein Scanner angeschlossen ist.

---

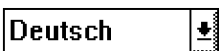
## Zeichenerkennung mit einem Bilddatei

Das Beispiel erfolgt anhand einer Seite einem gescannten Text über optische Zeichenerkennung.

### Einstellungen

Zuerst werden wir das System einstellen.

- 1 Vorher sollte als Sprache Deutsch ausgewählt sein aus der Drop-down Liste in die Funktionsleiste.



- 2 Wählen Sie im Menü "Voreinstellungen" mit Hilfe des Befehls "Textformat" das Dateiformat "ASCII" für die Textdatei und mit Hilfe des Befehls "Grafikformat" das PaintBrush-Format für die Bilddatei aus.
- 3 Wählen Sie "Parametereinstellungen" über das Menü "Verarbeitung" aus.
- 4 Wählen Sie als Größe normal und als Zeichenabstand proportional aus.
- 5 Wählen Sie als Diktionär-Datei READIRIS.DUS und als Modus "Zufügen zu Diktionär" aus.
- 6 Schließen Sie die Parametereinstellungen durch Klicken auf das Feld "schließen".



## Grafikdatei öffnen

- 1 Setzen Sie die Markierung "Anzeige nach Scan" im Menü "Verarbeitung".

Die Markierung ist gesetzt, wenn die Menüzeile mit einem Häkchen gekennzeichnet ist. Ist dies nicht der Fall, klicken Sie auf "Anzeige nach Scan", um das Häkchen zu setzen.

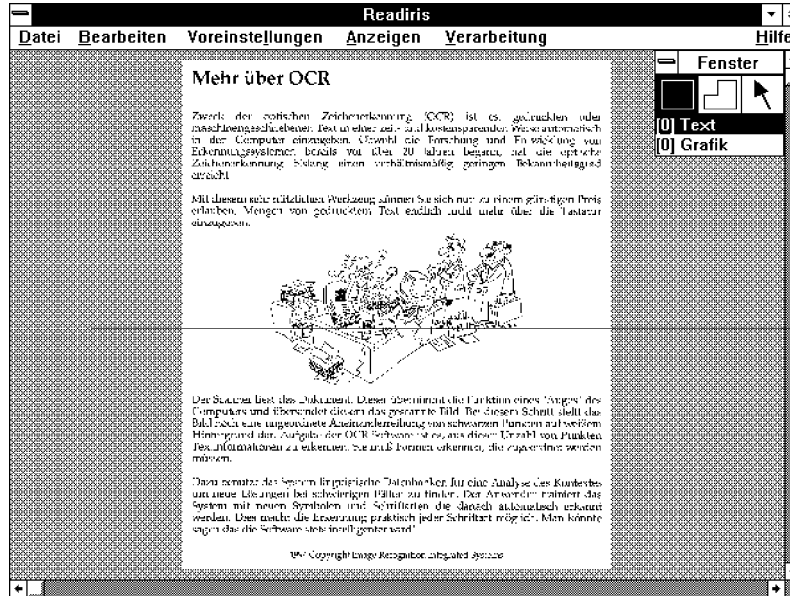
- 2 Klicken Sie auf das Diskette-Symbol um die richtige Eingabequelle zu wählen.



- 3 Klicken Sie auf das "Scan" Symbol in die Funktionsleiste. Jetzt müssen Sie die Grafikdatei wählen. Wählen Sie das vorgescannte Bild GERMAN.TIF in dem Readiris Verzeichnis.



- 4 Das Bild wird von der Festplatte gelesen und auf dem Bildschirm angezeigt. Desweiteren werden die Fenster-Werkzeuge rechts vom Bild eingeblendet.



## Vergrößern des Bildes

Einzelne Teile des Bildes können vergrößert werden.

1. Bewegen Sie den Mauszeiger auf den Bereich, den Sie vergrößern möchten.
2. Klicken Sie auf die *rechte* Maustaste. Klicken Sie ein weiteres Mal mit derselben Taste, erscheint das Bild wieder in Ganzansicht.

Sie können ebenso die Befehle aus dem "Anzeige" Menü und die entsprechenden Abkürzungstasten verwenden. Diese Tasten können sogar während der Arbeit mit der Maus benutzt werden; dies ist sehr hilfreich wenn Sie Blöcke in komplexen Layouts markieren.

## Definieren eines Textfensters und eines Grafikfensters

Jetzt müssen Sie angeben welche Teile der Seite sie verarbeiten möchten. Wir werden ein Textfenster um den zu erkennen Text setzen und ein Grafikfenster um die Bilder.

Setzen Sie nun ein Fenster um jeden zu lesenden Textbereich.

- 1 Klicken Sie zunächst auf das Rechteck-Symbol in den Fenster-Werkzeugen und dann auf "Text".
- 2 Der Mauszeiger nimmt nun die Gestalt eines Kreuzes an. Bewegen Sie ihn an die obere linke Ecke des Textes und klicken Sie.
- 3 Bewegen Sie den Mauszeiger an die untere rechte Ecke des Textes und klicken Sie.



Das Fenster ist nun definiert.

Alle oben beschriebenen Schritte gelten auch für Bildfenster, nur muß in den Fenster-Werkzeugen "Grafik" statt "Text" ausgewählt werden.

## Veränderung eines Textfensters

Größe und Position eines bereits gesetzten Textfensters können im nachhinein verändert werden.

- 1 Klicken Sie auf das Pfeil-Symbol in den Fenster-Werkzeugen.
- 2 Wählen Sie ein Fenster durch Klicken nahe dem Rand des gewünschten Fensters.
- 3 Um ein Fenster zu verschieben, klicken Sie zwischen zwei Fenster-Aktivpunkten. Halten Sie die linke Maustaste gedrückt und verschieben Sie das Fenster auf die gewünschte Position.
- 4 Um die Größe eines Fensters zu verändern, klicken Sie an den Fenster-Aktivpunkt. Halten Sie die linke Maustaste gedrückt und bringen Sie das Fenster in die gewünschte Größe.



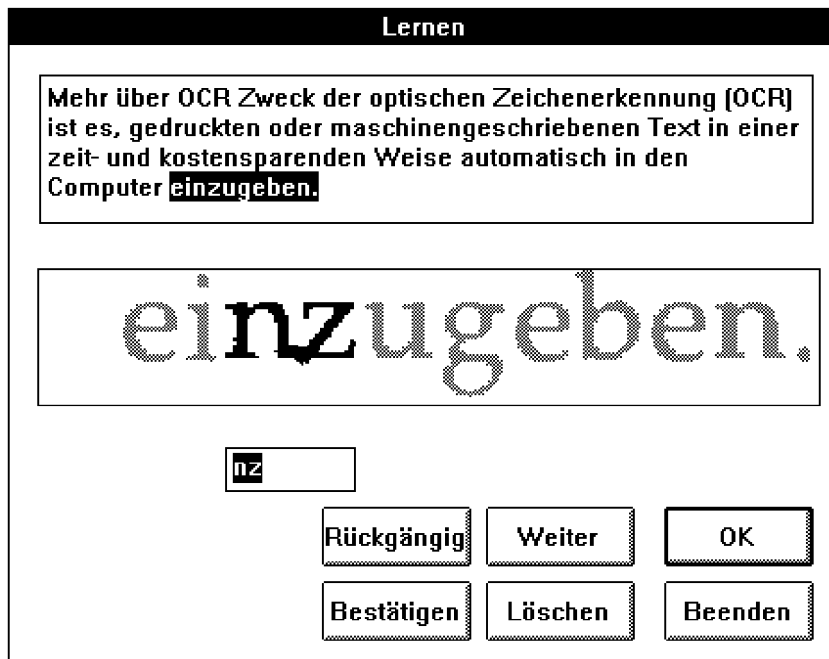
## Zeichenerkennung

Nach der Definition der Fenster kann die Zeichenerkennung aktiviert werden.



- 1 Klicken Sie das "OCR"-Symbol in der Funktionsleiste.

Nach ein paar automatischen Lernschritten schaltet das Programm in den Lernmodus. Der erkannte Text wird angezeigt. Das Programm hält bei zweifelhaften Zeichen an. Das Bild des Zeichens wird angezeigt und eine Lösung vorgeschlagen.



- 2 Ist die Lösung richtig, bestätigen Sie dies durch Klicken auf "OK" oder drücken ENTER. Ist die Lösung nicht richtig, geben Sie das richtige Zeichen über die Tastatur ein und drücken dann ENTER.

"Weiter" bewirkt die Korrektur eines Zeichens, ohne es jedoch zu lernen.

(Die genaue Funktion von den andere Optionen wird im nächsten Kapitell auseinandergesetzt.)

- 3 Verlassen Sie nun Readiris, aber bleiben noch in Windows.

Da das OCR-Symbol auch die Graphik-Datei speichert, ist die Graphik jetzt unsichtbar gespeichert.

## Anzeige des Ergebnisses

Es sollten nun eine Datei READIRIS.TXT und eine Datei READIRIS.PCX in Ihrem aktuellen Verzeichnis vorhanden sein.

- 1 Starten Sie den Notizblock und öffnen Sie die Datei READIRIS.TXT.

Der erkannte Text wird auf dem Bildschirm angezeigt. Was vorher eine Papiervorlage war, ist jetzt Text den Sie bearbeiten können!

- 2 Verlassen Sie die Notiz.

- 3 Starten Sie Paintbrush und öffnen Sie READIRIS.PCX.

Paintbrush stellt das Bild dar und Sie können es bearbeiten.

---

## OCR mit einem neu gescannten Bild

Wir setzen voraus, der Scanner ist korrekt installiert.

## Einstellungen



- 1 Klicken Sie auf das "Scanner"-Symbol in der Funktionsleiste. Die unterstützte Scanner werden angezeigt.
- 2 Wählen Sie den installierten Scanner und klicken sie "OK". Ein Dialogfenster gibt eine Beschreibung von Ihren Scanner-Einstellungen.
- 3 Setzen Sie die Auflösung auf 300 dpi oder auf 400 dpi, falls die Bildqualität niedrig ist.
- 4 Es ist möglich das Sie Helligkeit und Kontrast ändern müssen mit dem Rollbalken, falls vorhanden.

Bei Handy-Scannern sind Helligkeit und Kontrast durch entsprechende Schalter am Scanner vorgegeben.

- 5 Wählen Sie die korrekte Quelle durch auf das Scanner-Symbol in der Funktionsleiste zu klicken.

## Scannen mit einem Flachbett-Scanner

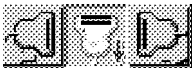


- 1 Legen Sie das zu lesen Dokument in den Scanner.
- 2 Wählen Sie die richtige Scan-Richtung in der Funktionsleiste.
- 3 Klicken Sie auf dem Scan-Symbol in der Funktionsleiste.

Die Seite wird gescannt und am Bildschirm dargestellt.

Wenn die gewählte Scanrichtung "Querformat" war, wird das Bild unmittelbar nach der Scavorgang rotiert.

## Scannen mit einem Handy Scanner



- 1 Wählen Sie die gewünschte Scan-Richtung in der Funktionsleiste.
- 2 Klicken Sie auf dem Scan-Symbol in der Funktionsleiste.
- 3 Der Cursor wird zu einem kleinen Handy-Scanner.
- 4 Legen Sie das zu lesende Dokument auf eine ebene Unterlage.
- 5 Positionieren Sie den Scanner auf dem zu bearbeitenden Teil des Dokuments.

- 6 Drücken Sie den Start-Knopf des Scanners und bewegen Sie den Scanner in die gewünschte Richtung.

Die Bewegung sollte so langsam und gleichmäßig wie möglich sein, ansonsten ist das gescannte Bild nicht für OCR zu verwenden. Das Bild wird entsprechend der Scannerbewegung am Bildschirm aufgebaut.

- 7 Wenn der Scan-Vorgang beendet ist, drücken Sie die Leertaste um Readiris die Kontrolle zu übergeben.

Falls das Bild von links nach rechts gescannt wurde, wird es durch drücken der Leertaste rotiert.

### **Texterkennung ausführen**

Jetzt können Sie Readiris wie im vorhergehenden Kapitel beschrieben benutzen: Zoomen um das Ergebnis des Scans zu kontrollieren, Fenster um die wichtigen Teile setzen, die Erkennung ausführen usw.





## 3 Mehr über Readiris

Die folgenden Seiten befassen sich mit den Leistungsmerkmalen von Readiris im einzelnen.

Es wird auf das Scannen einer Seite und die Steuerung der Intensität (Helligkeit) der gescannten Seite eingegangen. Ist die Seite korrekt gescannt, können Fenster um die Bereiche gesetzt werden, die Readiris lesen soll, oder Sie können Readiris der Layout automatisch analysieren lassen. Sodann beginnt Readiris mit der Zeichenerkennung und erstellt eine Ausgabedatei mit dem ausgewählten Text.

Wir werden es auch über die fortgeschrittene Merkmale von Readiris haben: automatische Erkennung und die Connect-Funktion. Wir werden auch lernen wie Sie Spalten einlesen und erkennen können.

---

### Definition der Standardeinstellungen

Zunächst definieren Sie die Standardeinstellungen:

- der verwendete Scanner und seine Parameter
- das Dateiformat für die von Readiris zu erstellenden Text- und Bilddateien
- die Sprache des zu erkennenden Textes.

Diese Einstellungen ändern sich im allgemeinen selten und können in einer Datei für Standardeinstellungen abgelegt werden.

## Konfiguration Ihres Flachbett-Scanners



- 1 Klicken Sie auf das Scanner-Symbol in der Funktionleiste. Die unterstützten Scanner werden angezeigt.

**Scanner**

Typ: Fujitsu M3094  
 Scanner ersetzen...

Kontrast: 127  
 Intensität: 127  
 Licht      Dokument      Dunkel

Format:  
 A3  
 A4  
 A5  
 B4  
 B5

Auflösung:  
 225  
 240  
 250  
 300  
 400

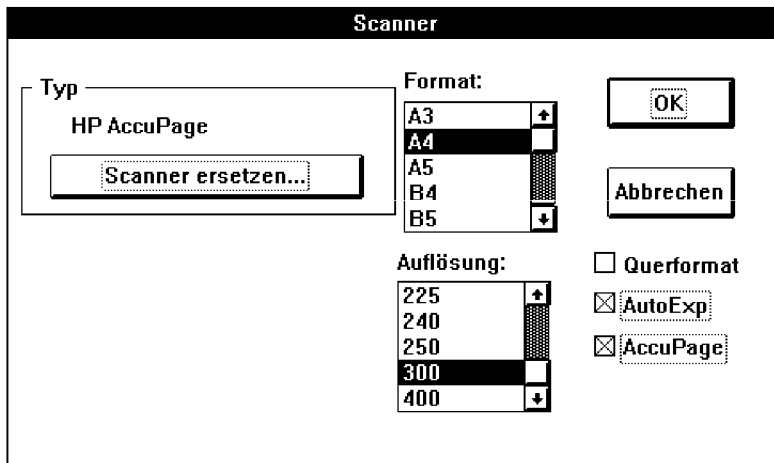
Querformat

OK  
 Cancel

- 2 Klicken Sie auf den Namen Ihres angeschlossenen Scanners. Die Standardauflösung, Seitenformat, Intensität und Kontrast werden angezeigt.
- 3 Klicken Sie auf das Format Ihres zu scannenden Dokumentes. Standardeinstellung: A4.
- 4 Klicken Sie auf die Auflösung, mit der das Dokument zu scannen ist. Standardeinstellung: 300 dpi.
- 5 Treffen Sie Ihre Einstellung im Intensität-Rollbalken gemäß der Helligkeit des Dokuments. Standardeinstellung: 127.
- 6 Treffen Sie Ihre Einstellung im Kontrast-Rollbalken, sofern bei Ihrem Scanner verfügbar.

Die Einstellung von Intensität und Kontrast ermöglicht es, zu helle und zu dunkle Dokumente auszugleichen. Mit HP-Scanner können Sie die AutoExp-Funktion benutzen. Diese Funktion stellt die Intensität automatisch ein.

Readiris unterstützt die **HP AccuPage** Technik. Diese Technik reguliert die Helligkeit und den Kontrast des Scanners automatisch



und verwendet neue Bildverarbeitungstechniken, um den Erkennungsprozeß des gescannten Bildes zu optimieren.

Diese Technik steht nur bei HP Scannern zur Verfügung. Installieren Sie die AccuPage Software, die Ihrem Scanner beiliegt, wählen Sie mit dem Befehl "Scanner" im Menü "Voreinstellungen" die Option "AccuPage" und klicken Sie "OK".

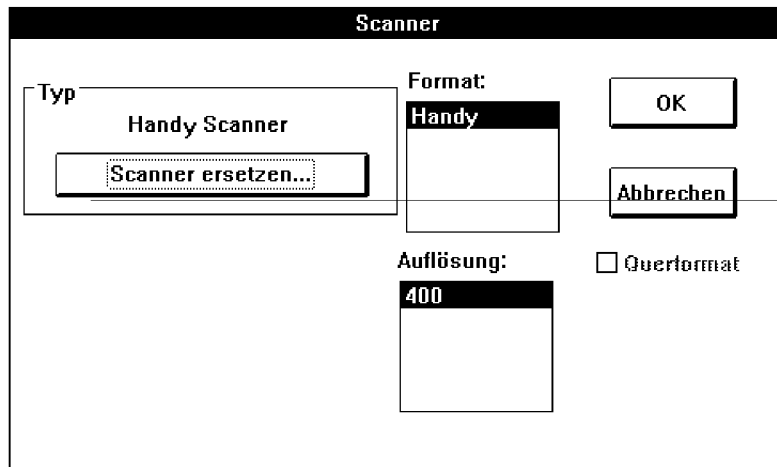
Das Benutzerhandbuch das mit Ihrem Scanner mitgeliefert ist, enthält weitere Informationen über die Bedienung und Installation vom Scanner. Die Datei READSCAN.TXT enthält auch Informationen über wie Sie ein Scanner einstellen müssen.

## Konfiguration Ihres Handy Scanners

- 1 Klicken Sie auf das Scanner-Symbol in der Funktionleiste. Die unterstützte Scanner werden angezeigt.
- 2 Klicken Sie auf den Namen Ihres angeschlossenen Scanners and klicken Sie "OK".

Die Scanner Einstellungen sind durch den Auflösungs-Schalter und den Helligkeits-Schalter an Ihrem Scanner vorgegeben.

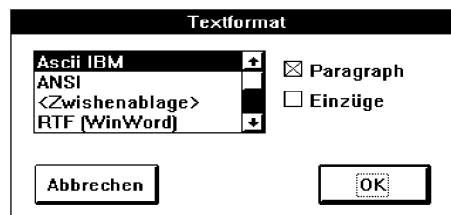
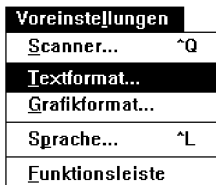
Das Benutzerhandbuch das mit Ihrem Scanner mitgeliefert ist, enthält weitere Informationen über die Bedienung und Installation



vom Scanner. Die Datei READSCAN.TXT enthält auch Informationen über wie Sie ein Scanner einstellen müssen.

## Textformat

- 1 Wählen Sie "Textformat" aus dem Menü "Voreinstellungen" aus.



- 2 Wählen Sie das gewünschte Textformat.

Readiris unterstützt viele **Textverarbeitungen und Textformate** . Wählen Sie ASCII oder ANSI für andere Textverarbeitungen.

Readiris ist auch in der Lage Spalten zu erkennen und das Spaltenformat in Spreadsheets einzubringen.

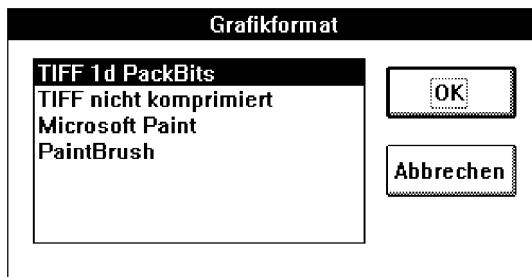
Das Zwischenablage Format erlaubt Ihnen die Ergebnisse direkt im Zwischenablage zu speichern und mit der "Einfügen" Option des "Bearbeiten" Menüs in Ihre Windows Anwendung einzufügen.

Readiris bietet nun zwei weitere Optionen zur Formatierung der Ausgabedatei an.

- **Paragraph:** aktiviert die automatische Absatzerkennung. Wird Paragraph nicht aktiviert, erhält jedes Zeilenende einen festen Zeilenumbruch. Die Option bewirkt, daß ein fester Zeilenumbruch nur am Ende eines Absatzes erscheint.
- **Einzüge:** aktiviert den Textrand-Code. Die Option sollte ausgewählt werden, wenn die Randeinstellung bzw. das Erscheinungsbild des Originals in die Ausgabedatei übernommen werden soll. Ist der Text zur Übernahme in ein DTP-Programm vorgesehen, sollte "Einzüge" nicht ausgewählt werden.

## Grafikformat

- 1 Wählen Sie "Grafikformat" aus dem Menü "Voreinstellungen" aus.



- 2 Wählen Sie das von Ihrem Bildverarbeitungsprogramm erkannte Grafikformat.

Readiris erstellt dann eine Bilddatei in diesem Format.

## Sprache

Es ist wichtig das Sie die richtige Sprache wählen wenn die **linguistische Datenbanken** die Erkennung erhöhen sollen.

Deutsch



- 1 Klicken Sie auf der "Drop-Down" Liste in der Funktionsleiste oder wählen Sie "Sprache" aus dem Menü "Voreinstellungen"
- 2 Wählen Sie die gewünschte Sprache.

Readiris hat viele linguistische Datenbanken die die Erkennung erhöhen.

Bei den Sprachen können Sie auch **numerisch** wählen um Ziffern mit sehr große Geschwindigkeit und Genauheit zu erkennen.

## Abspeichern der Standardeinstellungen

Nun können die Standardeinstellungen abgespeichert werden, um sie nicht jedesmal neu auswählen zu müssen.

Wählen Sie "Parameter speichern" aus dem Menü "Datei" aus. Readiris greift nun jedesmal beim Laden auf diese Angaben zu. Sie können diese Informationen sehen durch "Informationen" zu wählen aus dem Menü "Datei".

---

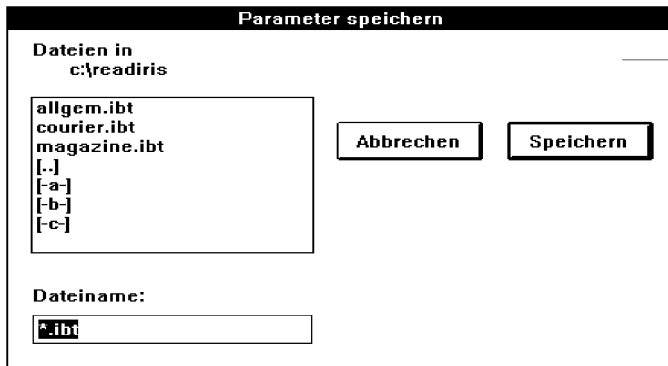
## Speichern und Neuladen von Parametern

Sie haben auch die Möglichkeit, für verschiedene Anwendungen unterschiedliche Einstellungen separat zu speichern. Das Abspeichern und Laden der Einstellungen geschieht mit Hilfe der Befehle "Parameter speichern" und "Parameter laden" aus dem Menü "Datei".

## Abspeichern der Einstellungen

- 1 Wählen Sie "Parameter speichern" aus dem Menü "Datei" aus.

Ein Dialogfenster zeigt eine Liste mit allen im aktuellen Verzeichnis vorhandenen Parameter-Dateien an.



- 2 Sie können nun einen neuen Dateinamen eingeben, um die Einstellungen in eine neue Datei abzulegen, oder eine vorhandene Voreinstellungs-Datei mit den neuen Einstellungen überschreiben.
- 3 Klicken Sie auf "Speichern"

Die komplette Readiris-Konfiguration wird nun in die gewählte Datei geschrieben. Die standardmäßige Dateinamenerweiterung lautet \*.ibt.

## Laden von Parametern

- 1 Wählen Sie "Parameter laden" aus dem Menü "Datei".  
Ein Dialogfenster zeigt nun eine Liste mit allen im aktuellen Verzeichnis vorhandenen Parameter-Dateien an.
- 2 Wählen Sie nun die Datei mit den gewünschten Einstellungen aus. Befindet sich die Datei in einem anderen Verzeichnis, kann sie auch mit der vollen Pfadbezeichnung eingegeben werden.
- 3 Klicken auf "Laden".



---

## Alles über die Parametereinstellungen

Bevor Sie mit dem OCR-Prozess anfangen, müssen Sie die Eigenschaften des Dokumentes eingeben. Diese sind:

- Die Sprache
- Die Qualität
- Die Schriftart
- Der Zeichenabstand und die Größe

Sie können auch angeben ob es einen Diktionär gibt für die zu lesen Schriftart.

Im Dialogfenster "Parametereinstellungen" werden alle Parameter angezeigt, die eingestellt werden können. Die Parameter können hier je nach zu lesendem Dokument verändert werden.

Das Dialogfenster "Parametereinstellungen" ist in drei Bereiche aufgeteilt:

- Dokument
- Dateien
- Verarbeitung

**Parametereinstellungen**

**Dokument**

Zeichenabstand

fest

proportional

Größe

eng

normal

Schriftart

Normal

Nadeldruck

**Dateien**

Eingabe Scanner

Ausgabe Text: (N) READIRIS.TXT  
Grafik: READIRIS.TIF

Layout Standard

Diktionär  READIRIS.DUS

**Verarbeitung**

1 Seite(n) zu lesen

Seitenanalyse

Ansicht nach Scan

Interaktives Lernen

Starke Kontrolle

Scannen

Schliessen

## Dokument

Hier werden grundsätzliche Einstellungen der Schrift vorgenommen:

- Zeichenabstand: Proportionschrift oder Schrift mit festen Zeichenabständen
- Größe: normal oder eng
- Schriftart: normal oder Nadeldruck

**Feste Zeichenabstände** wählt man aus, wenn alle Zeichen des Dokumentes den gleichen Abstand zwischen den Zeichen aufweisen, z. B. bei Maschinenschrift.

**Proportionschrift** wählt man aus, wenn die Zeichen- und Wortabstände eine ungleiche Breite aufweisen. Beispiel hierfür: ein "m" ist breiter als ein "i". Dies trifft beispielsweise meist auf Zeitschriften zu.

Als Zeichengröße sollte **eng** angegeben werden, wenn die Buchstabengröße 10 Punkte unterschreitet oder der Zeilenabstand sehr klein ist. Wird der Parameter "eng" bei sehr geringen Zeilenabständen nicht gesetzt, könnte Readiris Schwierigkeiten bei der Trennung der einzelnen Zeilen haben.

Der Benutzer soll ebenso die Schriftart angeben: Sie können kein **Nadeldruck** erkennen ohne der richtige Parameter einzustellen. Nadeldruck-Buchstaben sind von sehr schlechter Qualität und oft aufgebaut aus einzelne Punkten, wie Sie es im Beispiel sehen können.

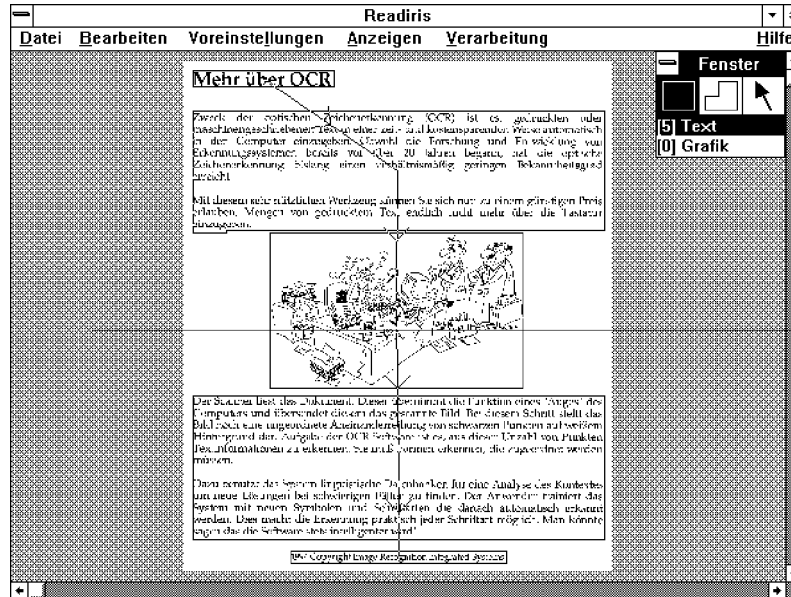
**ape--descended life**

## Verarbeitung

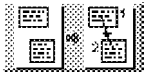
Hier können Sie die Art der Verarbeitung auswählen.

Es ist nicht nötig, die gescannten Seiten per Hand einzugrenzen! Aktivieren Sie die Option "Seitenanalyse" im Menü "Verarbeitung" oder über die Parametereinstellungen, die Readiris Software sucht und sortiert verschiedene Text- und Grafikblöcke auf der Seite automatisch. Dieses ist besonders nützlich, wenn Texte oder

Dokumente in Spaltenform mit einem komplizierten Layout über OCR eingelesen werden.



Die Seitenanalyse nimmt jeden erkannten Block als Textblock an. Greifen Sie nicht ein, wenn Grafiken auf der gedruckten Seite erscheinen: in einem zweiten Durchlauf erkennt die Software von allein, daß hier keine Textzeichen zu finden sind. (Diese Fenster werden auch nicht als Grafiken gespeichert.)



Die Textblöcke werden standardmäßig von oben nach unten und von links nach rechts sortiert. Sie können die Reihenfolge der Fenster ändern, indem Sie das "Sort"-Feld klicken und anschließend die Markierungspunkte der zu erkennenden Fenster. Nicht geklickte Fenster werden nicht erkannt.

Die Seitenanalyse ist schnell, präzise und toleriert schräge Einzüge.

Mit der Option "Anzeige nach Scan" können Sie angeben, ob das gescannte Bild vor seiner Verarbeitung **angezeigt** werden soll. (Für Handy-Scanner ist es am besten diese Option immer aktiv zu

lassen.) Ist diese Option nicht aktiv, können Sie mehrere Seiten automatisch hintereinander lesen lassen.

Der 3. Parameter ist "Interaktives Lernen". Hiermit können Sie die **Lernfunktion** aktivieren, womit Readiris Informationen von neuen Schriftarten lernt. Wählen Sie hier "OFF" für die Verarbeitung von mehreren Seiten von guter Qualität.

Sie können auch den Grad von der Interaktion beeinflussen. Wenn Sie "Starke Kontrolle" wählen, ist das Niveau der Kontrolle hoch: bei dem kleinsten Zweifel fragt Readiris einer Bestätigung statt einige Entscheidungen selber zu übernehmen. Benutzen Sie diese Option nur bei Dokumenten von zweifelhafter Qualität.

## Dateien

Hier treffen Sie die Auswahl aller Dateinamen, die bei Readiris eine Rolle spielen. Davon gibt es vier Arten.

**Eingabe:** hier wird der Scanner oder eine TIFF-Datei angegeben.

**Ausgabe:** hier wird außer der Ausgabedatei auch der Ausgabemodus angegeben - Überschreiben oder Anfügen. Wählen Sie entsprechend aus: "Neue Datei" oder "Zufügen Datei". Readiris kann zwei Ausgabedateien machen: eine Textdatei und eine Grafikdatei.

**Layout:** es kann der standardmäßige Name oder ein neuer Name gewählt werden. Die Layout-Datei enthält den Aufbau der vom Anwender definierten Fenster. Sie können Text- und Bilderfenster auf drei Arten definieren: mit der Hand, mit der Seitenanalyse und mit Layouts.

**Diktionär:** definiert das Anwenderlexikon und den Arbeitsmodus.

Diese Datei enthält den Namen des Schriften-Lexikons und enthält die Zeichen, die Sie dem System während der Lernphase beibringen.

Mit dem Lexikon kann auf dreierlei Art gearbeitet werden:

- Neuer Diktionär: Erstellung einer neuen Lexikondatei. Auf bereits bestehende Lexikondateien greift das System nicht zu.
- Zufügen zu Diktionär: Readiris öffnet eine bereits bestehende Lexikondatei und speichert die während des Lernvorgangs die gelernten Zeichen dazu.

- Diktionär verwenden wird in diesem Modus gearbeitet, verwendet Readiris eine bereits bestehende Lexikodatei, fügt jedoch keine Zeichen hinzu.

Der Auswahl des richtigen Modus kommt besondere Bedeutung zu, denn von ihr wird das Leseergebnis maßgeblich beeinflusst.

Soll nur eine Seite gelesen werden, wählen Sie den Modus "Neuer Diktionär" aus. Sollen mehrere Seiten mit etwa der gleichen Schriftart und Schriftqualität gelesen werden, wählen Sie "Neuer Diktionär" für die erste Seite, den Modus "Zufügen zu Diktionär" für die zweite und dritte Seite und den Modus "Diktionär verwenden" für den Rest der Seiten.

### Scannen mit einem Flachbett Scanner



- 1 Wählen Sie die richtige Scan-Richtung in der Funktionsleiste. (Sie können die "Querformat" Option auch aktivieren durch auf dem "Scanner"-Symbol in der Funktionsleiste zu klicken.)
- 2 Klicken Sie auf dem Scan-Symbol.

Das Dokument wird gescannt. Falls die "Anzeige nach Scan" Option aktiviert wurde, wird das Bild während des Scannens auf dem Bildschirm dargestellt.

Wenn die gewählte Scan-Richtung "Querformat" war, wird das Bild unmittelbar rotiert. Sie können das Bild auch nachher rotieren mit dem Befehl "Rotation" aus dem Menü "Anzeigen".

Achten Sie darauf das der **Zeilenversatz** nicht zu groß ist. Die Seitenanalyse und Erkennung von Readiris sind zwar Versatztolerant, aber wenn der Zeilenversatz zu groß ist, ist es schwer eine gute Texterkennung zu bekommen.

Klicken Sie auf dem "Scanner"-Symbol um die Einstellungen des Scanners, wie Auflösung usw., zu ändern.

### Scannen mit einem Handy Scanner

Wenn Sie die nötigen Optionen angegeben haben, können Sie mit dem Scannen des Dokuments beginnen.



- 1 Wählen Sie eine der drei Scan-Richtungen in der Funktionsleiste.

- 2 Klicken Sie auf das "Scan"-Symbol in der Funktionsleiste. Der Cursor wird zu einem kleinen Handy-Scanner.
- 3 Helligkeit und Kontrast sind durch entsprechende Schalter am Scanner vorgegeben. Setzen Sie die Auflösung auf 300 dpi oder auf 400 dpi, falls die Bildqualität niedrig ist.
- 4 Positionieren Sie den Scanner auf dem zu bearbeitenden Teil des Dokuments.
- 5 Drücken Sie den Start-Knopf des Scanners und bewegen Sie den Scanner in die gewünschte Richtung.

Die Bewegung sollte so langsam und gleichmäßig wie möglich sein, ansonsten ist das gescannte Bild nicht für OCR zu verwenden.

Das Bild wird entsprechend der Scannerbewegung am Bildschirm aufgebaut.

- 6 Wenn der Scan-Vorgang beendet ist, drücken Sie die Leertaste um Readiris die Kontrolle zu übergeben.

Falls das Bild von links nach rechts gescannt wurde, wird es durch drücken der Leertaste rotiert. Sie können das Bild auch nachher rotieren mit dem Befehl "Rotation" aus dem Menü "Anzeigen".

Achten Sie auch hier darauf das der **Zeilenversatz** nicht zu groß ist. Die Seitenanalyse und Erkennung van Readiris sind zwar Versatztolerant, aber wenn der Zeilenversatz zu groß ist, ist es schwer eine gute Texterkennung zu bekommen.

### Vergößern und Verkleinern des gescannten Bildes

Die Anzeige der gesamten Seite ist sinnvoll, um zur besseren Übersicht die genaue Positionierung von Bild und Text nebeneinander ansehen zu können. Die Option hierfür lautet "Passend" aus dem Menü "Ansicht".

Damit der Anwender beurteilen kann, ob er die Seite mit der richtigen Einstellung der Intensität (Helligkeit) gescannt hat, bietet Readiris folgende drei Optionen:

- 100 % größe: die Seite wird etwa in Originalgröße angezeigt.
- 50 % verkleinert: die Seite wird um die Hälfte verkleinert.

<b>Anzeigen</b>	
✓ <b>Passend</b>	^W
50% Verkleinert	^5
100% Größe	^1
200% Vergrößert	^2
✓ <b>Alle Fenster</b>	
Selektierte Typ	
<b>Sortieren</b>	
✓ <b>Sort Ansicht</b>	
<b>Sort Neu</b>	
<b>Rotation...</b>	^R
<b>Werkzeuge...</b>	
<b>Koordinaten...</b>	^C

- 200 % vergrößert: die Seite, oder ein Ausschnitt davon, wird in doppelter Größe angezeigt.

Sie können auch zwischen "100% Größe" und "Passend" durch Klicken mit der *rechten* Maustaste beliebig hin- und herschalten.

## Überprüfen der Qualität des gescannten Bildes

Die Leistung der optischen Zeichenerkennung hängt in hohem Maße von der Qualität des gescannten Bildes und vom Scanner ab. Bei näherer Betrachtung des gescannten Bildes läßt sich abhängig von den verwendeten Einstellungen die Qualität feststellen.

Bei guten Dokumenten liefert Readiris selbstverständlich auch bessere Resultate. Als gut zu bezeichnen sind Dokumente mit gut ausgeformten Zeichen und deutlicher Trennung zwischen den Zeichen.

Bei einem sehr hellen Dokument sind Zeichen mitunter zwei Teile zerrissen. In einem solchen Fall könnte beispielsweise ein "O" als "()" gelesen werden. Bei einem sehr dunklen Dokument hingegen entstehen häufig Zeichen mit überaus dicken Linien, und offene Formen laufen mitunter zu. Beispielsweise könnte ein "c" als "o" erkannt werden. Bei dunklen Dokumenten laufen oft auch Zeichen zusammen (Ligaturen) und tragen somit nicht gerade zu einem guten Leseergebnis bei.

Der Anwender kann die Qualität durch Veränderung der Scan-Helligkeit beeinflussen. Versuchen Sie verschiedene Einstellungen des Intensität-Knopfes, um die Unterschiede auf dem Bildschirm zu überprüfen.

---

## Fensterwerkzeuge

Wenn sich das Bild auf dem Bildschirm befindet können Sie die Fenster definieren.

Es gibt verschiedene Weisen für Fensterdefinition: die Seitenanalyse sortiert und gibt die Text- und Grafikfenster automatisch an. Sie können die Sortierreihenfolge ändern durch auf dem Sortiersymbol zu klicken, oder Sie können die Fenster mit



der Hand definieren. Drittens können Sie "Layouts" speichern um sie nachher wieder zu verwenden.

Verschiedene Fenster angeben oder erstellen ist erforderlich wenn es sich um Text oder Ziffern in Spaltenformat handelt. Ziehen Sie ein Fenster um jede Spalte, um die Ausgabe in den entsprechenden Zellen präzise zu gestalten.

Hier ist eine Beschreibung von den Readiris Fensterwerkzeuge.

### Erstellen eines rechteckigen Fensters

- 1 Wählen Sie das Rechteck-Werkzeug durch Klicken auf das entsprechende Feld in den "Fenster-Werkzeugen".
- 2 Wählen Sie den Fenstertyp: Textformat oder Grafikformat.
- 3 Bewegen Sie den Mauszeiger an eine Ecke des zu definierenden Fensters und klicken Sie.
- 4 Ziehen Sie das Fenster durch Bewegen des Mauszeigers und klicken Sie.



Es erscheint nun ein Rahmen um das definierte Fenster.

Seiten von weniger als 1mm sind nicht erlaubt!

### Erstellen eines polygonalen Fensters

- 1 Wählen Sie das Polygon-Werkzeug aus durch Klicken auf das entsprechende Feld in den Fenster-Werkzeugen.
- 2 Bewegen Sie den Mauszeiger zu einer Ecke des zu definierenden polygonalen Fenster und klicken Sie.
- 3 Durch Bewegen der Maus wird eine Seite des Polygons gezeichnet, welche sich mit der Maus bewegt.
- 4 Durch Klicken wird die angezeigte Seite des Polygons bestätigt.
- 5 Bewegen Sie den Mauszeiger und klicken Sie überall dort, wo sich eine Seite des Polygons befinden soll.
- 6 Sie können keine bestehenden Seiten kreuzen.
- 7 Klicken Sie mit der *rechten* Maus-Taste um das Polygon zu schließen.





Eine Seite die Sie schon gezeichnet haben können Sie nicht kreuzen.

Sie können auch polygonalen Fenster erstellen durch rechteckigen Fenster zusammenzufügen. Wenn zwei Fenster von derselben Art einander berühren, macht Readiris ein Fenster daraus.

### Auswahl eines Fensters



- 1 Wählen Sie das Pfeil-Werkzeug aus.
- 2 Klicken Sie den Fensterrand an. An jeder Ecke und in der Mitte des Fensters erscheint nun ein Aktivpunkt.
- 3 Um ein zweites Fenster auszuwählen, ohne das erste zu deaktivieren, halten Sie die Shift-Taste gedrückt und klicken dabei das zweite Fenster an.
- 4 Um ein Fenster einschließlich der sich wiederum in diesem Fenster befindlichen Fenster zu aktivieren, halten Sie die Control-Taste gedrückt und klicken dabei das Hauptfenster an.
- 5 Um Fenster zu deaktivieren, klicken Sie mit der linken Maustaste in einen Bereich außerhalb der Fenster.

Zur Beachtung: Den Typ und die Koordinaten eines Fensters können Sie sich anzeigen lassen durch Auswahl eines Fensters bei gedrückter Alt-Taste.

### Verändern eines Fensters

- 1 Wählen Sie das Fenster aus.
- 2 Bewegen Sie den Mauszeiger auf einen der Aktivpunkte.
- 3 Klicken Sie die linke Maustaste und halten Sie sie gedrückt.
- 4 Bewegen Sie die Maus. Die Fenstergröße verändert sich entsprechend der Mausbewegung.
- 5 Lassen Sie die Maustaste los, wenn das Fenster die gewünschte Größe erreicht hat.

### Verschieben eines Fensters

- 1 Wählen Sie das Fenster aus.

- 2] Bringen Sie den Mauszeiger an eine Seite des Fensters, etwa in die Mitte zwischen den beiden Aktivpunkten.
- 3] Klicken mit der linken Maustaste und halten Sie sie gedrückt.
- 4] Bewegen Sie die Maus. Das Fenster bewegt sich nun entsprechend der Mausbewegung.
- 5] Lassen Sie die Taste los, wenn das Fenster die gewünschte Position erreicht hat.

### Löschen eines Fensters

- 1] Wählen Sie das Fenster aus.
- 2] Wählen Sie entweder den Befehl "Löschen" oder "Neu" aus dem Menü "Bearbeiten", abhängig davon, ob Sie den Ausschnitt mit den Befehlen "Einfügen" oder "Kopieren" später wieder verwenden möchten.

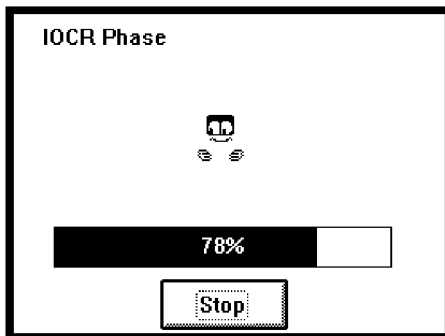
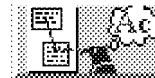
<b>Bearbeiten</b>	
Rückgängig	Alt+BkSp
Löschen	Sh Del
Kopieren	*Ins
Einfügen	Sh Ins
<b>Neu</b>	<b>Del</b>
Alle selektieren *A	

---

### Zeichenerkennung

Klicken Sie auf dem "OCR"-Symbol in der Funktionsleiste um mit der Texterkennung anzufangen, auch wenn Sie nur eine Grafik-Datei speichern möchten.

Sie können die gerade aktive Phase jederzeit abbrechen durch Klicken auf "Abbrechen".

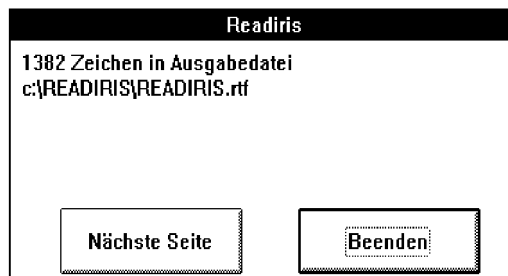


Da Readiris **Multitasking** unterstützt, kann die Texterkennung auch auf dem Hintergrund laufen. Beschränken Sie das

Readiris-Fenster auf ein Minimum, und erfüllen Sie eine andere Aufgabe auf dem Rechner; das interaktive Lernfenster warnt Sie wenn die Software fertig ist.

Während der interaktive Entscheidungsphase wird ein Dialogfenster eingeblendet, in welches der erkannte Text einfließt. Ist sich Readiris bei einer gefundenen Lösung nicht sicher, wird das Bild des Wortes angezeigt: Das betreffende Zeichen wird dabei hervorgehoben dargestellt. Readiris wartet nun auf Ihre Eingabe. Im nächsten Abschnitt wird näher auf dieses Dialogfenster eingegangen.

Wenn die Entscheidungsphase beendet ist, wird der erkannte Text in der Ausgabedatei gespeichert, und Readiris fragt, ob eine weitere Seite eingelesen werden soll.



Im Falle eines mehrseitigen Dokuments, ändert der "Nächste Seite" Befehl das Zeichensatz-Lexikon und die Ausgabedatei in den "Zufügen" Modus. Die nächste Seite wird dann direkt von Ihrem Scanner gelesen. Mit einem Flachbett Scanner wird die nächste Seite gescannt, mit einem Handy Scanner können Sie die nächste Seite scannen.

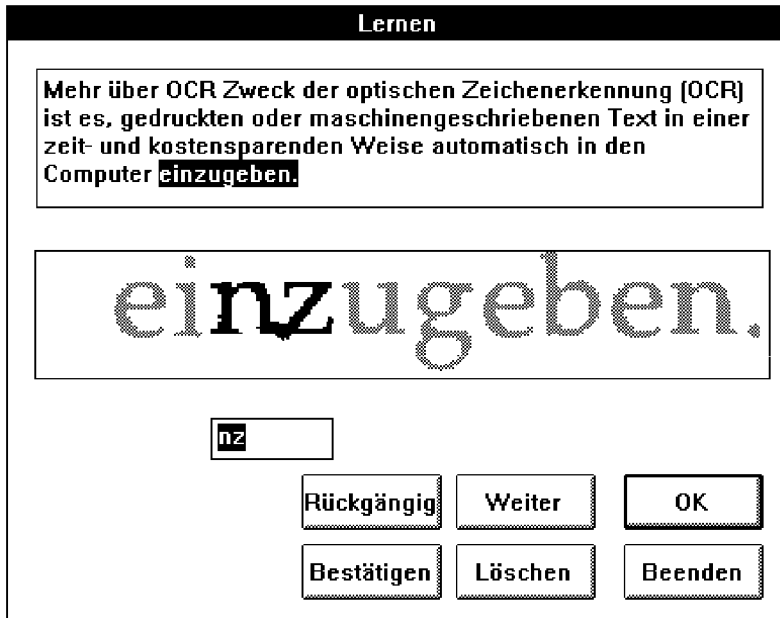
Sie werden feststellen das die Überprüfungs- und die Korrekturphase für die folgenden Seiten weniger Eingriffe erfordert wenn ein ähnlicher Schrifttyp verwendet wird.

---

## Lernen von Schriftarten

Der letzte Schritt bei der Zeichenerkennung ist die **Entscheidungsphase**. (Der Grad von der Interaktion hängt von der Option "Starke Kontrolle" in der "Parametereinstellung" ab.)

Ist sich Readiris bei einem Zeichen nicht sicher, so wird dieses mitsamt dem betreffenden Wort angezeigt. Das Zeichen - oft auch eine Zeichengruppe, die zusammenhängt - wird hervorgehoben dargestellt, und eine Lösung wird vorgeschlagen. Zeichen die Readiris nicht erkennt werden dargestellt durch "~".



Wenn erforderlich, geben Sie über die Tastatur das richtige Zeichen - ggf. die Zeichengruppe - ein und wählen Sie einen der folgenden Befehle.

## OK

Dieser Befehl erlaubt es, die vorgeschlagene Lösung entweder zu akzeptieren oder zu korrigieren. Readiris speichert das Zeichen dann mit dem Merkmal "sicher" im Lexikon ab. Readiris wird dieses Zeichen nicht wieder anzeigen, wenn es nochmals vorkommt.

Alternative: drücken Sie die Eingabe-Taste.

## **Weiter**

Dieser Befehl erlaubt es, die vorgeschlagene Lösung entweder zu akzeptieren oder zu korrigieren. Der Unterschied besteht darin, daß Readiris dem Zeichen im Lexikon das Merkmal "unsicher" gibt.

Dies ist z.B. dann von Vorteil, wenn das angezeigte Zeichen zerrissen ist und dadurch die Gefahr von Verwechslungen beim Lesen der noch folgenden Zeichen nach sich ziehen würde. Readiris wird dieses Zeichen wieder anzeigen, wenn es nochmals vorkommt, und die von Ihnen dann eingegebene Lösung abspeichern.

## **Bestätigen**

Bei vielen Schriftarten haben die Zeichen "1" und "l" die gleiche Gestalt. Um sie auseinanderzuhalten, führt Readiris eine Kontextanalyse durch. Genügt der Kontext nicht für eine eindeutigen Entscheidung, fragt Readiris das Zeichen ab.

Unterscheiden sich die Zeichen "1" und "l" im gescannten Dokument in ausreichender Weise, klicken Sie bei beiden Zeichen auf "Bestätigen".

## **Rückgängig**

Haben Sie versehentlich ein falsches Zeichen eingegeben, so können Sie die Eingabe mit diesem Befehl widerrufen. Sie können bis zu neun Zeichen widerrufen.

## **Löschen**

Wenn Sie Readiris diesen Befehl geben, speichert es das betreffende Zeichen nicht ab. Zweck dieses Befehls ist es zu vermeiden, daß Flecken auf dem Papier versehentlich als falsch erkannte Zeichen abgespeichert werden.

## Beenden

Dieser Befehl bricht den interaktiven Lernvorgang ab. Die Entscheidungsphase wird jedoch nicht unterbrochen.

---

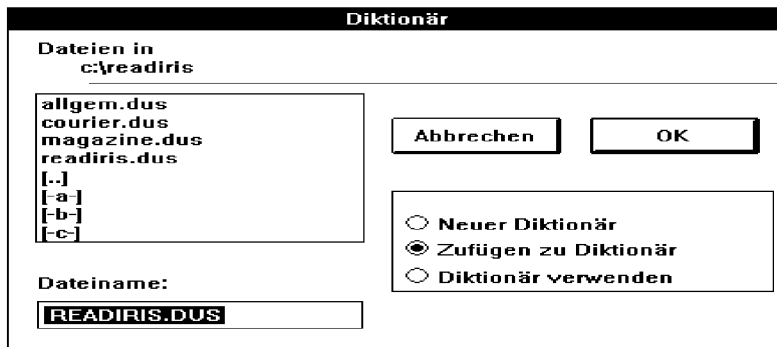
## Lernen mit Hilfe des Anwenderlexikons

Wenn Sie eine höhere Anzahl von Vorlagen mit derselben Schriftart lesen, möchten Sie Readiris wahrscheinlich nicht jedesmal dieselben Zeichen beibringen müssen! Aus diesem Grunde können Sie die gelernten Zeichen in ein Anwenderlexikon abspeichern und dieses für gleiche Vorlagen immer wieder benutzen.

Um ein Anwenderlexikon zu erstellen oder aufzurufen, gelten folgende Schritte:

- 1 Wählen Sie "Parametereinstellungen" aus dem Menü "Verarbeitung" aus.
- 2 Klicken Sie auf "Diktionär".

Das Dialogfenster des Lexikons zeigt eine Liste aller im aktuellen Verzeichnis vorhandenen Anwenderlexika an. Die standardmäßige Dateinamenerweiterung lautet \*.dus.



- 3 Geben Sie einen Dateinamen in das Textfenster ein, um ein neues Anwenderlexikon zu erstellen, oder wählen Sie ein vorhandenes durch Klicken auf dessen Namen aus.

- 4] Wählen Sie einen der verfügbaren Arbeitsmodi aus: "Neuer Diktionär", um ein neues Anwenderlexikon zu erstellen, "Zufügen zu Diktionär", um ein bestehendes Anwenderlexikon zu erweitern und "Diktionär verwenden", wenn Readiris das Anwenderlexikon nur verwenden, jedoch nicht erweitern soll.
- 5] Klicken Sie auf "OK".

Der Dateiname des Anwenderlexikons wird neben dem Feld "Diktionär" im Dialogfenster "Parametereinstellungen" angezeigt. Der ausgewählte Modus steht in Klammern.

---

## On-Line Hilfe

Dies beschließt unsere allgemeine Übersicht von Readiris. Ein **on-line Hilfe-System**, das ausführliche Informationen enthält, ist beigefügt. Wählen Sie das Menü "Hilfe", um auf die Hilfe-Informationen zuzugreifen. Die on-line Hilfe kann auch noch einige neue Informationen enthalten.

Hilfe
-------

Inhalt
--------

Über Readiris
---------------

"Inhalt" gestattet Ihnen den Zugriff auf on-line Hilfe-Informationen. Alternative: drücken sie die F1 Taste. "Über Readiris" zeigt den ursprünglichen Bildschirm mit Ihren Registrierdaten an.

Readiris hat einige wichtige Merkmale die noch weiter erklärt werden: Tabellenerkennung, automatische Erkennung und die Connect-Funktion.

---

## Tabellen erkennen

Sie können erkannte Tabellen direkt in Ihr Tabellenformulare importieren.

Ziehen Sie ein Fenster um jede Spalte, um die Ausgabe in den entsprechenden Zellen präzise zu gestalten. Siehe Beispiel unten.

MB	FIRST Plattform	SECOND Plattform	THIRD Plattform	FOURTH Plattform
0.0	E2,857	E2,143	E2,143	E2,143
0.1	E2,921	E2,207	E2,207	E2,207
0.5	E3,178	E2,454	E2,454	E2,454
1.0	E3,500	E2,766	E2,766	E2,766
2.0	E4,153	E3,429	E3,286	E3,143
3.0	E4,798	E3,929	E3,843	E3,700
4.0	E5,439	E4,256	E3,987	E3,857
5.0	E6,071	E4,643	E3,930	E3,999
6.0	E6,671	E4,857	E4,288	E3,857
7.0	E7,071	E5,071	E4,536	E3,893
8.0	E7,571	E5,286	E4,871	E4,143
9.0	E8,071	E5,500	E4,807	E4,390
10.0	E8,571	E5,714	E4,643	E4,643
11.0	E9,029	E5,857	E4,293	E4,607
12.0	E9,296	E6,000	E5,143	E4,571
13.0	E9,843	E6,143	E5,289	E3,536
14.0	E10,000	E6,289	E5,536	E4,643
15.0	E10,257	E6,429	E5,536	E4,821
16.0	E10,714	E6,571	E5,571	E5,400
17.0	E11,071	E6,714	E5,607	E5,178
18.0	E11,429	E6,857	E5,643	E5,257
19.0	E11,786	E7,000	E5,879	E5,536
20.0	E12,143	E7,143	E5,714	E5,714
25.0	E13,908	E7,679	E6,607	E6,607
30.0	E14,649	E8,214	E7,500	E7,600
35.0	E15,000	E8,750	E8,389	E8,333
40.0	E17,143	E9,286	E9,200	E9,886
45.0	E18,000	E10,179	E10,179	E10,179
50.0	E18,250	E11,071	E11,871	E11,871
55.0	E19,921	E11,964	E11,964	E11,964
60.0	E20,714	E12,857	E12,857	E12,857

Zwei neue Textformate wurden hinzugefügt: **Excel ASCII** erstellt eine formatierte Text-Tabelle, die direkt in Ihr Spreadsheet transferiert werden kann. **Excel Zwischenablage** bietet ebenfalls diese Möglichkeit, die Ausgabedatei erscheint hier in der Windows Zwischenablage. Verwenden Sie den Befehl "Einfügen".

Beachten Sie, daß die Suche nach Absätzen und Randeinstellungen nicht aktiviert ist, wenn Sie das Excel Textformat auswählen.

Wenn es keinen Text sondern nur Ziffern gibt, können Sie die Erkennung verbessern und beschleunigen durch "Numerisch" wählen in der Drop Down Liste in der Funktionsleiste.

## Automatische Erkennung

Readiris ist in der Lage, Ihre Seiten automatisch zu erkennen. Die Intensität der Automatik hängt von Ihrem Scanner ab.

**Flachbett Scanner** gestatten vollautomatische Erkennung einer großen Anzahl von Seiten. Deaktivieren Sie die Optionen "Interaktives Lernen" und "Anzeige nach Scan" und aktivieren Sie die Option "Seitenanalyse" im Menü "Verarbeitung" oder in den

Verarbeitung	
Parametereinstellung...	*P
Scannen	*S
OCR	*O
Interaktives Lernen	
<input checked="" type="checkbox"/> Seitenanalyse	
<input type="checkbox"/> Anzeige nach Scan	



Parametereinstellungen. (Sie können auch ein Fenster-Layout anwenden um die wichtige Bereiche anzugeben.) Geben Sie die Anzahl der Seiten an, die Sie automatisch erkannt haben möchten.

Legen Sie einen Stapel Seiten in den Feeder Ihres Scanners, starten Sie den Scan-Vorgang und lassen Sie Ihren PC den Rest erledigen. Die Dokumente werden gescannt, eingegrenzt und erkannt, ohne daß ein Eingreifen des Anwenders notwendig ist! Falls der Dokument-Feeder leer ist vor Beendigung der Erkennung, erscheint eine Fehler-Nachricht.

Bei **Handy Scannern** liegen die Dinge anders. Da das Scannen hier stets einen gewissen physischen Aufwand seitens des Anwenders erfordert, kann die Erkennung niemals völlig automatisch geschehen. Dennoch können Sie die Phasen des Eingrenzens in Fenster und der Erkennung automatisch durchführen lassen.

Da die automatische Erkennung die Deaktivierung der Interaktiven Lernphase verlangt, ist es zu empfehlen, Readiris zunächst mit einigen Seiten zu trainieren, um die Erkennungsrate der Ausgabe zu optimieren.

---

## **Die Funktion Connect**

Die Connect-Funktion ermöglicht es Ihnen, Readiris über Ihr Windows-Programm zu aktivieren. Jetzt können Sie Texte über Ihre Windows Applikation OCR-einlesen!

### **WordPerfect, Word und Excel**

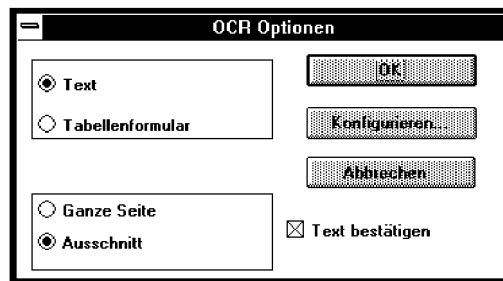
Vergewissen Sie sich, daß Readiris nicht im Hintergrund läuft. Klicken Sie den neuen Befehl in Ihrem Applikationsmenü an, um Readiris über Ihr Windows Programm laufen zu lassen. Für WordPerfect 5.x erscheint dieser Befehl im Menü "Macro", für WordPerfect 6.0 im Menü "Tools" und für Word und Excel im Menü "Datei".

Die Connect-Schnittstelle wird Ihnen einige Fragen stellen, bevor Readiris beginnen kann. Gehen Sie vor wie gewöhnlich, und verlassen Sie die Connect-Schnittstelle nach Beendigung. Die Ergebnisse der Erkennung werden an der gegenwärtigen Cursorposition eingefügt.

## Andere Windows-Applikationen

Klicken Sie das Connect Symbol zweimal an, um den Connect Task Server im Hintergrund laufen zu lassen, bevor Sie Ihre Windows Applikation beginnen.

Klicken Sie den neuen Befehl "OCR" im Menü "Datei" Ihrer Applikation an. Ein Dialogfenster erscheint.



Geben Sie an, ob Sie einen Text ("Text") oder eine Tabelle ("Tabellenformular") einlesen möchten. Wählen Sie "Ganze Seite", um die Seitenanalyse zu aktivieren oder "Ausschnitt", um ein zu scannendes Dokument manuell einzugrenzen. Aktivieren Sie die Option "Text Bestätigen" um die interaktive Lernphase in den Vorgang einzuschließen. Die Validation durch den Anwender ist standardmäßig deaktiviert.



Das Feld "Konfigurieren" erlaubt den Zugriff auf höhere Parameter. Wählen Sie "Tasten Simulieren", um den erkannten Text so einzufügen, als ob er über die Tastatur eingegeben worden wäre oder wählen Sie "Zwischenablage Einfügen", um den Text unsichtbar von der Zwischenablage einzufügen.

Prüfen Sie, ob die korrekten Abkürzungsschlüssel für den Befehl "Einfügen" in Ihrer Windows Applikation spezifiziert sind. Einige Tasten representieren ein spezielles Zeichen:

TASTE	VORSTELLUNG
Shift	+
Ctrl	^
Alt	%
Insert	{INSERT}

Einige Beispiele: wenn der Abkürzungsschlüssel für den Befehl "Einfügen" Ctrl+V lautet, sollten Sie **^V** eingeben. Wenn der Abkürzungsschlüssel Shift+Insert lautet, geben Sie **+{INSERT}** ein usw.

# 4 Befehle von Readiris

Dieses Kapitel enthält eine komplette Übersicht über alle Menüs von Readiris. Die Befehle aller Menüs werden hier beschrieben. Die meisten davon kennen Sie bereits, falls Sie das Kapitel "Schneller Start" durchgearbeitet haben.

---

## Datei

Dieses Menü gibt Ihnen die Möglichkeit, Voreinstellungen und Fenster abzuspeichern und wieder aufzurufen. Über dieses Menü können Sie auch Bilddateien auswählen und sich anzeigen lassen.

### TIFF-Datei laden

Öffnet Bilddateien im Format TIFF und bringt das Bild auf den Bildschirm. Dieser Befehl steht nur Vorschau-Modus zur Verfügung.

### Information

Zeigt Ihre Readiris-Konfiguration an.

### Quelle selktionieren

Readiris ist kompatibel mit dem neuen Scanner Standard Twain für Handy und Flachbett Scanner. "Quelle Selektieren" gestattet das Speichern von Parametern.



**Erwerben**

Beginnt den Scan-Vorgang und entspricht damit dem Befehl "Scannen" im Menü "Verarbeiten" sowie dem "Scannen"-Feld in den Parametereinstellungen.

**Layout speichern**

Speichert die vom Anwender im angezeigten Bild definierten Fenster ab. Die standardmäßige Dateinamenerweiterung lautet .wdw. Dieser Befehl steht nur im Vorschau-Modus zur Verfügung.

**Layout laden**

Mit diesem Befehl können Sie die zuvor in einer Datei mit der Erweiterung .wdw abgespeicherten Fenster aufrufen und aktivieren. Der Befehl ist nur im Vorschau-Modus verfügbar.

**Parameter speichern**

Speichert die Readiris-Konfiguration in eine Datei mit der standardmäßigen Erweiterung .ibt ab.

**Parameter laden**

Mit diesem Befehl können Sie die zuvor in eine Konfigurationsdatei abgespeicherten Parameter laden und aktivieren.

**Standard Parameter speichern**

Dieser Befehl bewirkt, daß Readiris beim nächsten Aufruf mit den aktuellen Einstellungen arbeitet.

**Ende**

Beendet Readiris.

## Bearbeiten

Die sechs Befehle dieses Menüs sind nur während der Anzeige eines Dokumentes verfügbar. Sie betreffen die Fenster, die Sie bei der Vorschau des Dokumentes definieren.

Um ein Fenster auszuwählen, klicken Sie an eine Seite des Fensters. Möchten Sie gleichzeitig mehrere Fenster auswählen, halten Sie die Shift-Taste gedrückt und klicken dabei jedes dieser Fenster an.

Bearbeiten	
Rückgängig	Alt+BkSp
Löschen	Sh Del
Kopieren	^Ins
Einfügen	Sh Ins
Neu	Del
Alles Selektieren *A	

### Rückgängig

Widerruft den zuvor erteilten Befehl.

### Löschen

Löscht die ausgewählten Fenster und speichert sie in einen internen Puffer.

### Kopieren

Kopiert die ausgewählten Fenster in einen internen Puffer.

### Einfügen

Fügt den Inhalt des internen Puffers ein. Die eingefügten Fenster erscheinen jeweils an der Position, an der sie zuvor gelöscht oder kopiert wurden. Um sie zu verschieben, klicken Sie auf die Seite eines der Fenster und bewegen Sie die Maus. Die Fenster verschieben sich entsprechend der Mausbewegung.

### Neu

Löscht alle ausgewählten Fenster.

## Alle selektieren

Entfernt alle Fenster durch einen Mausklick.

---

## Voreinstellungen



Mit dem Menü "Voreinstellungen" können Sie die Voreinstellungen setzen, die standardmäßig gelten sollen: Text- und Grafikformat und die Sprache des Dokumentes. In diesem Menü können Sie die Funktionsleiste ein- oder ausschalten.

### Scanner

Hier werden die Scanner und die Scanner-Einstellung angezeigt: Modell, Auflösung usw.

### Textformat

Bestimmt das Textformat der Ausgabedatei, Formate verschiedener Textverarbeitungsprogramme stehen zur Auswahl.

Aktivierung von Paragraph oder Absatzerkennung und Randeinstellung in der Ausgabedatei.

Wenn Sie Spalten erkennen wollen, müssen Sie das Excel-Format wählen.

### Grafikformat

Auswahl eines Grafikformats der Ausgabedatei, welches Ihr Bildverarbeitungsprogramm unterstützt.

### Sprache

Auswahl einer Sprache. Readiris hat verschiedene linguistische Datenbanken. Bei den Sprachen können Sie auch "numerisch" wählen um Ziffern mit sehr große Geschwindigkeit und Genauheit zu erkennen.

## Funktionsleiste

Hiermit können Sie die Funktionsleiste ein- oder ausschalten. Die Funktionsleiste wird an der linken Seite gezeigt. Der Gebrauch der Funktionsleiste wird empfohlen, da man mit den Symbolen fast alle Befehle ausführen kann.

---

## Ansicht

Die Befehle dieses Menüs lassen sich nur während der Vorschau des Dokuments aufrufen.

### Passend

Anpassung der Vergrößerungs- oder Zoom-Stufe, so daß das gescannte Bild exakt die Größe Ihres Bildschirms annimmt.

### 50% Verkleinert

Das Bild wird um 50 % verkleinert (gezoomt).

### 100 % Größe

Das Dokument wird auf seine tatsächliche Größe gezoomt: ein Pixel des Bildschirms stimmt überein mit ein Pixel des Bildes. Durch Klicken auf die *rechte* Maustaste können Sie beliebig zwischen "100 % Größe" und "Passend" umschalten.

### 200% Vergrößert

Vergrößern des Bildes um 200 %.

### Alle Fenster

Zeigt alle im Layout des Dokuments definierten Fenster an.

<b>Anzeigen</b>	
<input checked="" type="checkbox"/> <b>Passend</b>	<b>^W</b>
<input type="checkbox"/> <b>50% Verkleinert</b>	<b>^5</b>
<input type="checkbox"/> <b>100% Größe</b>	<b>^1</b>
<input type="checkbox"/> <b>200% Vergrößert</b>	<b>^2</b>
<b>Alle Fenster</b>	
<input type="checkbox"/> <b>Selektierte Typ</b>	
<b>Sortieren</b>	
<input checked="" type="checkbox"/> <b>Sort Ansicht</b>	
<input type="checkbox"/> <b>Sort Neu</b>	
<input type="checkbox"/> <b>Rotation...</b>	<b>^R</b>
<b>Werkzeuge...</b>	
<input type="checkbox"/> <b>Koordinaten...</b>	<b>^C</b>



## **Selektierte Typ**

Zeigt alle Fenster des ausgewählten Typs an.

## **Sort**

Sortiert die Textfenster und teilt Readiris dadurch die Reihenfolge mit, in der es den Text in den Fenstern lesen und in der Ausgabedatei abspeichern soll. Dies ist vor allem bei mehrspaltigen Dokumenten nützlich. Klicken Sie auf die Ecken jedes Fensters um die Ausgabereihenfolge festzulegen.

## **Sort Ansicht**

Zeigt die mit dem Befehl "Sort" definierte Reihenfolge der Fenster an. Dies geschieht durch Pfeile, die die einzelnen Textfenster verbinden.

## **Sort Neu**

Setzt die Sortierung der Textfenster zurück.

Alternative: Klicken Sie auf dem "Sortieren"-Symbol in der Funktionsleiste

## **Rotieren**

Rotiert das gescannte Bild. Das Befehl zeigt ein "Rotieren"-Fenster, in dem man die Rotation des Bildes im oder gegen den Uhrzeigersinn wählt, an. Wir empfehlen das Sie beim scannen die richtige Scanrichtung wählen.

## **Werkzeuge**

Hier können der Typ der Fenster (Text- oder Grafikenster) sowie die Fensterwerkzeuge (Rechteck, Polygon und Pfeil) ausgewählt werden.

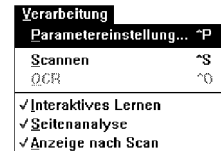
## Koordinaten

Zeigt die Koordinaten des Mauszeigers an. Die Zahlen beziehen sich auf die Pixel des Dokuments.

---

## Verarbeitung

Dieses Menü enthält alle Befehle mit Bezug auf dem OCR-Prozess: Scannen, OCR und alle Parameter dafür.



## Parametereinstellungen

Zeigt das Dialogfenster für die "Parametereinstellungen" an.

## Scannen

Abhängig von der Auswahl der Eingabe im Dialogfenster "Parametereinstellungen" beginnt entweder der Scanner mit dem Abtasten des Dokuments, oder Readiris lädt die ausgewählte Bilddatei.

Alternative: Klicken Sie auf dem "Scannen"-Symbol in der Funktionsleiste.

## OCR

Startet den Vorgang der Zeichenerkennung. Dieser Befehl entspricht dem Feld "Scannen" bei den Parametereinstellungen.

## Interaktives Lernen

Aktiviert das interaktive Lernen. Diese Option bestimmt ob Sie zweifelhafte Zeichen kontrollieren wollen, oder das Sie die Entscheidungen von Readiris akzeptieren. (Die Option "Starke Kontrolle" bei den Parametereinstellungen bestimmt den Grad der Kontrolle.)

Das interaktive Lernen kann auch in der Funktionsleiste aktiviert werden.

Sie können auch der Grad von der Interaktion beeinflussen. Wenn Sie "Starke Kontrolle" wählen, ist das Niveau der Kontrolle hoch: bei der kleinste Zweifel fragt Readiris eine Bestätigung statt einige Entscheidungen selber zu übernehmen. Benutzen Sie diese Option nur bei Dokumenten von einer zweifelhaften Qualität.

### **Seitenanalyse**

Aktiviert die Seitenanalyse. Abhängig davon wird die gescannte Seite automatisch analysiert.

Diese Option können Sie auch aktivieren durch klicken auf dem "Seitenanalyse"-Symbol in der Funktionsleiste.

### **Anzeige nach Scan**

Aktiviert den Vorschau-Modus. Das gescannte Dokument wird auf dem Bildschirm angezeigt. Danach können die einzelnen zu lesenden und als Bilddatei abzuspeichernden Bereiche definiert werden.

---

## **Parametereinstellungen**

Alle Auswahlmöglichkeiten in den Standardeinstellungen sind in Kapitel 3, "Mehr über Readiris", beschrieben.

### **Zeichenabstand**

Bestimmt den Zeichenabstand: Proportionalschrift oder Schrift mit festen Zeichenabständen.

### **Größe**

Hier können Sie die Größe der Zeichen im Dokument auswählen: eng oder normal. Sie sollten "eng" auswählen, wenn die Buchstabengröße 10 Punkte unterschreitet oder wenn die

Software auf Schwierigkeiten bei der Trennung aufeinanderfolgender Zeilen stößt.

## **Schriftart**

Gibt die Druckqualität von dem Text an: Normal oder Nadeldruck.

## **Seite(n) zu lesen**

Hier geben Sie die Anzahl der zu scannenden Seiten ein. Diese Option ist nicht vorhanden bei Handy-Scanner

## **Seitenanalyse**

Aktiviert die Seitenanalyse. Abhängig davon wird die gescannte Seite automatisch analysiert.

Diese Option können Sie auch aktivieren durch klicken auf dem "Seitenanalyse"-Symbol in der Funktionsleiste.

## **Anzeige nach Scan**

Aktiviert den Vorschau-Modus. Die Wirkung dieses Befehls entspricht der von der Option "Anzeige nach Scan" im Menü "Verarbeitung".

## **Interaktives Lernen**

Aktiviert das interaktive Lernen. Diese Option bestimmt ob Sie zweifelhafte Zeichen kontrollieren wollen, oder das Sie die Entscheidungen von Readiris akzeptieren.

Das interaktive Lernen kann auch in der Funktionsleiste aktiviert werden. Schalten Sie diese Option aus für automatische Texterkennung. Sie können auch den Grad der Interaktion beeinflussen mit der nächsten Option.

## **Starke Kontrolle**

Wenn Sie "Starke Kontrolle" wählen, ist das Niveau der Kontrolle hoch: bei der kleinste Zweifel fragt Readiris eine Bestätigung statt einiger Entscheidungen selber zu übernehmen. Benutzen Sie diese Option nur bei Dokumenten von einer zweifelhaften Qualität.

## **Eingabe**

Hier wählen Sie aus, ob die Eingabe über den Scanner oder eine zuvor auf der Festplatte abgespeicherte Datei erfolgen soll.

Diese Option können Sie auch aktivieren mit dem "Eingabe"-Symbol in der Funktionsleiste.

## **Ausgabe**

Mit diesem Befehl legen Sie den Namen der Ausgabedatei fest. Existiert bereits eine Datei mit diesem Namen, müssen Sie entscheiden, der Text angehängt oder die Datei überschrieben werden soll. Readiris vergibt automatisch die Dateinamenweiterung, je nachdem, welches Ausgabeformat Sie zuvor ausgewählt haben.

## **Layout**

Auswahl der Layout-Datei mit den Fensterdefinitionen für einen Dokumententyp. Die Standardweiterung lautet .wdw.

## **Diktionär**

Hier wählen Sie den Namen des Anwenderlexikons und den Arbeitsmodus aus: Neuer Diktionär, Zufügen zu Diktionär, Diktionär verwenden. Standardmäßig lautet die Dateinamenerweiterung \*.dus.

## **Scan**

Startet den Scanvorgang.

Alternative: klicken Sie auf dem "Scan"-Symbol in der Funktionsleiste.

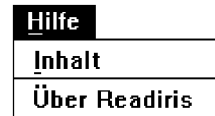
## Schließen

Schließt das Dialogfenster "Parametereinstellungen".

---

## Hilfe

In diesem Menü gibt es das on-line Hilfe-System und das "Über Readiris" Fenster.



## Inhalt

"Inhalt" gestattet Ihnen den Zugriff auf dem on-line Hilfe-System das ausführliche Informationen enthält über jeden Aspekt des Softwares. Die on-line Hilfe kann auch noch einige neue Informationen enthalten.

Alternative: drücken Sie die F1 Taste.

Diese Informationen können nicht nur auf dem Bildschirm angezeigt werden, sondern können auch abgedruckt werden. Wie, finden Sie in Ihrem Windows Handbuch.

## Über Readiris

Zeigt den ursprünglichen Bildschirm mit Ihren Registrierdaten an.



## 5 Mehr über OCR

Zweck der optischen Zeichenerkennung - OCR oder Optical Character Recognition - ist es, gedruckten oder maschinengeschriebenen Text in einer zeit- und kostensparenden Weise automatisch in den Computer einzugeben. Obwohl die Forschung und Entwicklung von OCR-Systemen bereits vor über 20 Jahren begann, hat die optische Zeichenerkennung bislang einen verhältnismäßig geringen Bekanntheitsgrad erreicht.

Mit diesem sehr nützlichen Werkzeug können Sie sich nun zu einem günstigen Preis erlauben, Mengen von gedrucktem Text endlich nicht mehr über die Tastatur einzugeben.

OCR-Systeme können für viele Dokumentenarten eingesetzt werden, wenngleich nicht für alle... Im folgenden fassen wir die einzelnen Schritte zur Verarbeitung eines Dokuments zusammen. In diesem Zusammenhang geben wir auch einen Überblick über Readiris.

---

### Der OCR-Vorgang

Man legt das zu lesende Dokument in den Scanner. Dieser übernimmt die Funktion eines "Auges" des Computers und übersendet diesem das gescannte Bild. Bei diesem Schritt stellt das Bild noch eine scheinbar ungeordnete Aneinanderreihung von schwarzen Punkten auf weißem Hintergrund dar. Aufgabe der OCR-Software ist es, aus dieser Unzahl von Punkten Textinformationen zu erkennen. Sie muß Formen erkennen, die zugeordnet werden müssen.

Es gibt zwei Arten von OCR-Systemen: automatisch lesende Systeme und lernende Systeme. Die erstgenannten können



automatisch eine ganze Reihe von Schrifttypen erkennen, aber auch nicht alle. Beim zweiten Typ bringt der Anwender dem System bei jeder neuen Dokumentenart neue Zeichen bei. Der Anwender ist hier mehr in den Vorgang eingespannt. Dies hat jedoch den Vorteil, daß im Endeffekt mehr Schriftarten erkannt werden können.

Beide Typen von OCR-Systemen verwenden letztlich die gleiche Methode zur Texterkennung. Hier die wichtigsten Schritte:

- Trennung der Zeilen
- Trennung der Zeichen voneinander
- Erkennung der Zeichenbedeutung
- Ausgabe von Text

Die **Trennung der Zeilen** voneinander besteht daraus, einzelne Blöcke auf der Textseite zu isolieren. Dieser Schritt umfaßt auch die Erkennung des Zeilenabstandes und des Grades der Zeilenausrichtung, d.h. ob die Zeilen auf der Vorlage gerade oder schräg positioniert sind. Am schwierigsten gestaltet sich das Lesen von Vorlagen mit gleichzeitig sehr geringen Zeilenabständen und einer relativ schrägen Ausrichtung.

Sodann werden die Zeilen zunächst in **Wörter zertrennt**, die wiederum in **Buchstaben** aufzuteilen sind. Weisen alle Buchstaben die gleiche Breite auf, ist dieser Schritt relativ einfach auszuführen. Dies ist bei maschinengeschriebenen Vorlagen meist der Fall.

Schwieriger wird es schon, wenn die Breite je nach Form des Buchstaben variiert, nämlich bei Proportionalchrift. Noch größer wird die Herausforderung, wenn sich außerdem noch Zeichen berühren. Nadeldruck-Zeichen sind aus einzelne Punkte aufgebaut, und erfordern eine besondere Trennungstechnik.

Die **Zeichenerkennung** untersucht die Merkmale aller isolierten Zeichen und gibt ihnen einen Namen.

Die **Textausgabe** besteht aus dem Abspeichern des aus den erkannten Mustern erstellten Textes in eine Datei, die Sie in Ihr Textverarbeitungs- oder Desktop Publishing-Programm und Spreadsheets übernehmen können.

---

## Anwendungsgebiete für Readiris

Readiris ist ein leistungsstarkes und zugleich einfach zu bedienendes Anwenderpaket. Seine Funktionen wurden aus allen von I.R.I.S. entwickelten OCR-Technologien ausgewählt.

Dieses OCR-Programm umfaßt die folgenden Merkmale:

- Es kann im **automatischen Lernmodus** arbeiten. Dadurch wird es zum "offenen" System. Readiris ist ein lernendes System und arbeitet nur teilweise automatisch. Einen großen Teil des Dokumentes kann Readiris selbständig erkennen.
- Readiris enthält leistungsstarke Funktionen zur **Verwendung des linguistischen Kontextes** für die Zeichenerkennung. Es kann jedoch auch ohne linguistischen Kontext arbeiten, z.B. bei numerischen Tabellen.
- Die Software führt den Anwender durch einen **einzigen Vorgang**, der sich aus fünf Schritten zusammensetzt: Scannen, Vorschau, Fensterfunktionen, topologische Onmifont-Erkennung und Auswertung der automatisch erlernten Zeichen.

Wie jedes andere OCR-System kann auch dieses Programm nicht alle Arten von Dokumenten fehlerfrei verarbeiten, jedoch: was diese Software liest, das liest sie gut! Zuverlässigkeit und Effektivität sind dabei großgeschrieben. Die Software liest sowohl Original-Dokumente als auch Kopien guter Qualität.

Aus dem Leistungsmerkmal des automatischen Lernens kann der Anwender am größten Nutzen ziehen, wenn die folgenden Voraussetzungen erfüllt sind.

### Linguistischer Kontext

Readiris benötigt bei jeder Schriftart eine gewisse Anzahl von Zeichen, um daraus einen Kontext zu bilden. Erst mit dessen Hilfe kann es den automatischen Lernvorgang so durchführen, daß die Texterkennung nahezu automatisch erfolgt.

Ist wenig Kontext vorhanden, arbeitet das System zwar auch, fragt jedoch den Anwender wesentlich häufiger nach Zeichen.

## **Isolierte Zeichen**

Zwecks Einsparung von Speicherplatz und Lesezeit verfügt Readiris über weniger Funktionen zur Trennung von Zeichen.

Der erste Schritt bei der Zeichenisolierung kann bei deutlich voneinander getrennten Zeichen durchgeführt werden. Im zweiten Schritt hingegen verwendet Readiris einen leistungsstarken Algorithmus zur Isolierung von unterschrittenen Zeichen und Ligaturen.

Infolgedessen muß Readiris bei dem betreffenden Dokument nur in der Lage sein, wenigstens ein Minimum an elementaren Mustern zu erkennen, um die zusammenhängenden Zeichen korrekt trennen zu können.

Dies bedeutet, daß die Erkennung jeder Schriftart möglich ist, wenn es dem System gelingt, eine ausreichende Anzahl von elementaren Mustern zu isolieren.

## **Zeichengröße**

Die Zeichengröße sollte zwischen 8 und 20 Punkten liegen. Der Textkorpus fast eines jeden Dokuments liegt in diesem Bereich. Kleinere Schriften werden ebenfalls gelesen, vorausgesetzt, Sie wählen eine Auflösung von 400 dpi.

Es ist nicht ratsam, große Buchstaben in den wenigen Wörtern einer Überschrift mit dem übrigen Textkorpus gleichzeitig zu verarbeiten. Ohne dies wird sie ein Desktop Publishing-Anwender bei der Weiterverarbeitung zumeist per Hand eingeben.

## **Nicht lesbare Dokumente**

Das System kann keinen allzu dichten Text lesen, d. h. wenn fast alle Zeichen zusammenhängen. Aus diesem Grunde ist das Lesen von Dokumenten mit sehr schlechter Qualität nicht möglich. Dies gilt besonders für dunkle Photokopien.

# Index

## !

100 %	49
200 %	49
50 %	49

## A

Abbrechen	35
Absatz	
Absatzerkennung	48
Alle Fenster	49
Alle selektieren	48
Anfügen	54
Ansicht	
100 %	49
200 %	49
50%	49
Alle Fenster	49
Menü	49
Passend	49
Selektierte Fenster	50
Anwenderlexikon	
Anlegen	39
Aufrufen	39
Anzeige nach Scan	52 - 53
ASCII	11
Auflösung	
400 dpi	60
Ausgabedatei	54

## B

Bearbeiten	
Einfügen	35
Kopieren	35
Löschen	35
Menü	47
Neu	35

Beenden	39
Befehl	
100 %	49
200 %	49
50%	49
Alle Fenster	49
Alle selektieren	48
Anzeige nach Scan	52
Einfügen	47
Fensterwerkzeuge	50
Grafikformat	48
Information	45
Koordinaten	51
Kopieren	47
Layout laden	46
Layout speichern	46
Löschen	47
Nächste Seite	36
Neu	47
OCR	51
Parameter laden	46
Parameter speichern	46
Parametereinstellungen	51
Passend	49
Rückgängig	47
Scannen	30, 51
Scanner	48
Selektierte Fenster	50
Sort	50
Sort Ansicht	50
Sort Neu	50
Sprache	48
Standard Parameter speichern	46
Textformat	48
TIFF-Datei laden	45
Bestätigen	38
Bild	
Bilddatei	11
Bilddatei öffnen	45
Ganzansicht	12
Zoomen	12
Bilddateien	3





Seite(n) zu lesen	53
Selektierte Fenster	50
Sort 50	
Sort Ansicht	50
Sort Neu	50
Spalte	50
Sprache	10, 48
Systemvoraussetzungen	7

## T

Textformat	10, 48
ASCII	11
Clipboard	23
Textverarbeitungsprogramm	48, 58

## V

Verarbeitung	
Menü	51
Vergrößern	12
Voreinstellung	
Aufrufen	45
Menü	48
Speichern	45
Standard Parameter speichern	46
Voreinstellungen	10
Vorschau	53

## W

Weiter	38
Windows	
NOTIZ	15
Version	7

## Z

Zeichen	
Größe	52
Größenbereich	60
Unterscheidung	60
Zeichenabstand	52
Zeichenerkennung	14
zusammenhängend	60
Zeichenabstand	52
Auswahl	27
fest	58
fester Zeichenabstand	27
Maschinenschrift	27
Proportionalschrift	27, 58
Zeichenerkennung	35
Zeichengröße	27
Zeichentrennung	58, 60
Zeile	
Zeilenabstand	58
Zeilenabstand	27
Zeilentrennung	58
Zoomen	12