

Background

package

# READ IRIS OCR

yananár

real

text

*Save Time, No More Retyping!*

typed

IRIS

Druckqualität

trainable

paper

učici modul

OCR

обраа

matrix

caractères accentués

omnifont

easy

recognition

fax

neue Symbole und Schriftar



IR

IMAGE RECOGNITION  
INTEGRATED  
SYSTEMS

**JUST READ IT!**

---

# **Readiris™** **Gebruikersgids**

**Intelligente karakterherkenningssoftware**

**voor IBM PCs, PS/2 en compatibele PCs**

**Image Recognition Integrated Systems**

Readiris is een merk van Image Recognition Integrated Systems.

De informatie in dit document is onderhevig aan veranderingen zonder voorafgaande mededeling en houdt geen verbintenis vanwege I.R.I.S. De software die in deze handleiding wordt geschreven wordt geleverd op basis van een gebruikerslicentie. De software kan enkel worden gebruikt overeenkomstig de voorwaarden van de gebruikerslicentie.

Proximity is een geregistreerd merk van Proximity Technology Inc. IBM is een geregistreerd merk van de International Business Machines Corp. DOS en Windows zijn (geregistreerde) merken van de Microsoft Corp. Alle bestandsformaten of scanners die in de de handleiding voorkomen zijn (geregistreerde) merken van hun respectievelijke eigenaars.

THE PROXIMITY/MERRIAM-WEBSTER LINGUIBASE

Copyright 1984 Alle rechten voorbehouden  
Proximity Technology Inc.

Copyright 1984 Alle rechten voorbehouden  
Merriam-Webster Inc.

THE PROXIMITY/MERRIAM-WEBSTER LINGUIBASE

Copyright 1984 Alle rechten voorbehouden  
Proximity Technology Inc.

Copyright 1984 Alle rechten voorbehouden  
William Collins Sons & Co. Ltd.

De spelling van dit product is gebaseerd op  
Proximity Linguistic Technology  
© Copyright Image Recognition Integrated Systems, 1995  
All rights reserved

Document nummer: READ-9501-DUT

# Inhoudstafel

## **Van start gaan**

---

|   |   |
|---|---|
| Systeemvereisten .....                    | 7 |
| Readiris installeren .....                | 7 |
| De Connect mogelijkheid installeren ..... | 8 |

## **Een eerste verkenning**

---

|  |    |
|--|----|
| Readiris opstarten .....                   | 9  |
| Startconfiguratie bepalen .....            | 10 |
| Een vooraf gescand beeld inlezen .....     | 10 |
| Readiris instellen .....                   | 10 |
| Een beeld openen .....                     | 11 |
| Inzoomen op een beeld .....                | 12 |
| Tekst- en grafische vensters tekenen ..... | 12 |
| Vensters wijzigen .....                    | 13 |
| Herkenning uitvoeren .....                 | 13 |
| De resultaten bekijken .....               | 14 |
| Een papieren document herkennen .....      | 15 |
| Scanner selecteren .....                   | 15 |
| Scannen met een vlakbedscanner .....       | 15 |
| Scannen met een handscanner .....          | 15 |
| Herkenning uitvoeren .....                 | 16 |

## **Meer over Readiris**

---

|                                     |    |
|-------------------------------------|----|
| Startconfiguratie bepalen .....     | 17 |
| Een vlakbedscanner selecteren ..... | 18 |
| Een handscanner selecteren .....    | 19 |
| Tekstformaten .....                 | 20 |
| Grafisch formaat .....              | 21 |
| Taal .....                          | 21 |
| Startconfiguratie bewaren .....     | 22 |
| Instellingen opslaan en laden ..... | 22 |
| Instelling opslaan .....            | 22 |
| Instellingen laden .....            | 23 |
| Alles over het controlepaneel ..... | 23 |

## II

|   |    |
|---|----|
| Dokument  | 24 |
| Aktie   | 25 |
| Bestanden                                       | 26 |
| Scannen met een vlakbedscanner                  | 27 |
| Scannen met een handscanner                     | 28 |
| Gescande beelden van nabij bekijken             | 29 |
| De kwaliteit van het gescande beeld controleren | 29 |
| Bevenstering                                    | 30 |
| Rechthoekige vensters creëren                   | 30 |
| Veelhoekige vensters creëren                    | 30 |
| Vensters selecteren                             | 31 |
| Vensters wijzigen                               | 31 |
| Vensters verplaatsen                            | 32 |
| Vensters wissen                                 | 32 |
| Karakterherkenning                              | 32 |
| Lettertypes aanleren                            | 34 |
| OK  | 34 |
| Doorgaan  | 35 |
| Bevestigen                                      | 35 |
| Achteruit                                       | 35 |
| Wissen  | 35 |
| Einde   | 35 |
| De efficiëntie van woordenboeken                | 36 |
| Hulpinformatie opzoeken                         | 37 |
| Tabellen herkennen                              | 37 |
| Automatische herkenning                         | 38 |
| De Connect-mogelijkheid gebruiken               | 39 |

### Commando's

---

|                           |    |
|---------------------------|----|
| Bestand                   | 43 |
| Openen                    | 43 |
| Info                      | 43 |
| Bron Selecteren           | 43 |
| Verwerven                 | 44 |
| Layout opslaan            | 44 |
| Layout laden              | 44 |
| Instellingen opslaan      | 44 |
| Startconfiguratie opslaan | 44 |
| Verlaten                  | 44 |
| Bewerken                  | 45 |
| Ongedaan maken            | 45 |
| Knippen                   | 45 |
| Copiëren                  | 45 |
| Plakken                   | 45 |
| Wissen                    | 45 |
| Alles Selecteren          | 45 |
| Opties                    | 46 |
| Scanner                   | 46 |

|                         |    |
|-------------------------|----|
| Tekstformaat            | 46 |
| Grafisch Formaat        | 46 |
| Taal                    | 46 |
| Knoppenbalk             | 46 |
| Bekijken                | 47 |
| Hele Pagina             | 47 |
| 50% Verkleining         | 47 |
| Ware Grootte            | 47 |
| 200% Vergroting         | 47 |
| Alle Vensters           | 47 |
| Geselecteerd Type       | 47 |
| Sorteren                | 47 |
| Sorteren Bekijken       | 48 |
| Sorteren Reset          | 48 |
| Roteren                 | 48 |
| Funkties                | 48 |
| Coördinaten             | 48 |
| Aktie                   | 48 |
| Controlepaneel          | 49 |
| Scannen                 | 49 |
| OCR                     | 49 |
| Interactief Leren       | 49 |
| Pagina-analyse          | 49 |
| Bekijken na het Scannen | 49 |
| Controlepaneel          | 50 |
| Letterafstand           | 50 |
| Lettergrootte           | 50 |
| Lettertype              | 50 |
| Te Lezen Pagina('s)     | 50 |
| Pagina-analyse          | 50 |
| Bekijken na het Scannen | 50 |
| Interactief Leren       | 51 |
| Strengere Controle      | 51 |
| Invoer                  | 51 |
| Uitvoer                 | 51 |
| Layout                  | 51 |
| Dico                    | 51 |
| Scannen                 | 52 |
| Sluiten                 | 52 |
| Hulp                    | 52 |
| Inhoud                  | 52 |
| Aangaande Readiris      | 52 |

## Meer over OCR

---

|   |    |
|---|----|
| Het OCR-proces                          | 53 |
| Het bereik van Readiris                 | 54 |
| Aantal karakters en taalkundige context | 55 |
| Karakterisatie                          | 55 |

# IV

|                             |    |
|-----------------------------|----|
| Karaktergrootte .....       | 56 |
| Onleesbare documenten ..... | 56 |



# Welkom

Readiris is een krachtig intelligent OCR software-pakket (optische karakterherkenning) dat van uw PC en scanner een efficiënt invoersysteem maakt.

Het biedt u de beste en snelste manier om gedrukte en getikte informatie te converteren in editeerbare tekst die kan verwerkt worden met een tekstverwerker, rekenblad, DTP-pakket enz.

Readiris maakt gebruik van de vertrouwde conventies van de Windows omgeving. De gebruiksvriendelijkheid laat je toe Readiris op te starten vanuit je favoriete Windows-toepassingen!







# Over Readiris

Het software-pakket Readiris is zeer eenvoudig te installeren en te gebruiken. Readiris installeren duurt twee minuten, van start gaan vijf minuten. Het gebruik beperkt zich in de meeste gevallen tot klikken in de knoppenbalk.

Readiris is een betaalbaar produkt dat u niettemin toegang verleent tot de meest geavanceerde AI-technieken op het gebied van de optische karakterherkenning (OCR).

Readiris is een zelflerend systeem. Anders gesteld kan het autonoom nieuwe karakters aanleren. Een uitgebreid aantal linguïstische databanken, lexicons, wordt gebruikt om de linguïstische context te analyseren. Op die manier beschikt Readiris over een krachtig hulpmiddel om moeilijke gevallen op te lossen.

Readiris biedt je ten slotte ook de mogelijkheid om beelden van gedrukte documenten via uw scanner in te lezen en met uw grafische software te bewerken.



# Over deze handleiding

Deze handleiding beschrijft het gebruik van Readiris. Het eerste hoofdstuk, "Van start gaan", geeft informatie over de systeemvereisten en de installatieprocedure.

Het tweede hoofdstuk, "Een eerste verkenning", beschrijft kort de werking van Readiris en wijdt u in de basisbegrippen in.

Daarna wenst u meer details over de verschillende kenmerken en functies van Readiris. Daar zorgt het derde hoofdstuk, "Meer over Readiris", voor.

Een kort overzicht van de menus en hun commando's vindt u in het vierde hoofdstuk, "Readiris commando's".

Het laatste hoofdstuk, "Meer over OCR", beschrijft kort enkele OCR-concepten en parameters. Tegelijk wordt de impact van deze parameters op OCR-systemen en op Readiris verduidelijkt.

Alle bijkomende informatie die niet in deze handleiding opgenomen is, staat in het bestand README.TXT. Raadpleeg dit bestand indien je problemen hebt met de installatie of gebruik van Readiris. Voor meer informatie over scanners kunt u terecht in het bestand READSCAN.TXT.



# 1 Van start gaan

Dit hoofdstuk legt uit hoe je Readiris moet installeren.

Wij veronderstellen hierbij dat u vertrouwd bent met uw computer, het DOS besturingssysteem en de Windows omgeving.

---

## Systeme vereisten

Dit zijn de minimale systeemvereisten om Readiris te kunnen gebruiken:

- een IBM PC, PS/2 of compatibele computer met een 386 processor. Readiris ondersteunt multitasking, zelfs op een 386 PC!
- 4 MB RAM geheugen. Om van de Connect mogelijkheid gebruik te maken kan je wel eens meer geheugen nodig hebben.
- het MS-DOS besturingssysteem, 3.3 of later. MS-DOS 6.x wordt aanbevolen
- Windows 3.1. Gebruik Windows 3.1 niet in de standaard modus!

---

## Readiris installeren

Als veiligheidsmaatregel maakt u best een kopie van uw Readiris diskettes; bewaar deze op een veilige plaats.

Voer de volgende stappen uit om Readiris te installeren:

- Ga na of u in het Programmabeheer van Windows zit. Selecteer "Starten" onder het menu "Bestand", voer **a:setup** in en klik op "OK".
- Het installatieprogramma start. Geef de gewenste gebruikerstaal van Readiris aan en voer de instructies uit die op uw scherm verschijnen.

Op een bepaald ogenblik wordt je gevraagd de tweede diskette in te voeren.

De programmagroep en de applicatie worden door het installatieprogramma gecreëerd.

---

## **De Connect mogelijkheid installeren**

Readiris moet correct geïnstalleerd zijn eer je de Connect mogelijkheid kan installeren.

Start Readiris op, selecteer je scanner en sla de startconfiguratie op eer je de Connect mogelijkheid installeert. Hoofdstuk 3, "Meer over Readiris", vertelt je hoe dat moet.

Voor WordPerfect, Word en Excel moet je wel nog een macro installeren. Het on-line hulpssysteem legt uit hoe dat gaat - de betreffende hulpbladzijde wordt overigens vanzelf op het scherm gebracht wanneer het Readiris installatieprogramma loopt.

Voor alle andere Windows-toepassingen wordt er door het Readiris installatieprogramma een Connect task server geïnstalleerd; de gebruiker is hier niet bij betrokken.

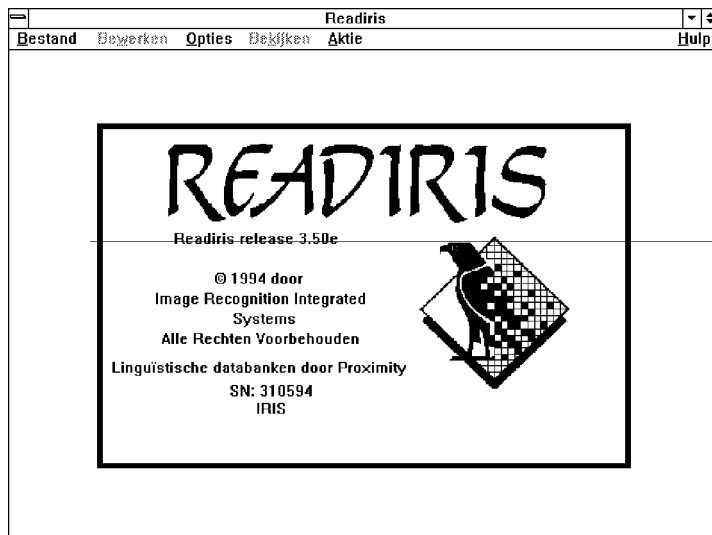
## 2 Een eerste verkenning

De beste methode om met Readiris vertrouwd te raken bestaat er uiteraard in, Readiris te gebruiken. Dit hoofdstukje laat je toe meteen van start te gaan. Een vooraf gescand beeld, DUTCH.TIF, wordt bij de licentie geleverd en kan dus worden gebruikt om met Readiris te oefenen, ook al heb je nog geen scanner geïnstalleerd.

---

### Readiris opstarten

Ga na of je in het Programmabeheer van Windows zit. Klik twee keer op het Readiris pictogram.



Het startscherm verschijnt na enkele seconden.





Het venster van Readiris bevat niet enkel commandomenu's maar ook een **knoppenbalk**: op deze knoppen klikken zal volstaan om de meeste operaties uit te voeren. Klik op het commando "Knoppenbalk" onder het menu "Opties" om de knoppenbalk te activeren en desactiveren.




---

## Startconfiguratie bepalen

- 1 Kies het menu "Opties". De volgende commando's verschijnen: "Scanner", "Tekstformaat", "Grafisch Formaat" en "Taal".
- 2 Doorloop deze commando's en stel de standaard instelling in.
- 3 Selecteer het commando "Startconfiguratie Opslaan" in het menu "Bestand". Klik "OK" om te bevestigen.

Het systeem is nu klaar voor het gebruik. Wij bevelen de onderstaande oefening aan, ook al is uw scanner al geïnstalleerd.

---

## Een vooraf gescand beeld inlezen

Deze kleine demo maakt gebruik van een vooraf gescand beeld over OCR. We zullen de tekst herkennen en de illustratie in een grafisch bestand opslaan.

### Readiris instellen

Laten we eerst de nodige parameters instellen.



- 1 Controleer of het Nederlands geselecteerd is in de "Taal" selectie in de knoppenbalk.
- 2 Selecteer met het commando "Tekstformaat" in het menu "Opties" de tekstfilter ASCII voor uw tekstbestand, en met het commando "Grafisch Formaat" het formaat PaintBrush voor de illustratie.
- 3 Ga naar het "Controlepaneel" onder het menu "Aktie".
- 4 Controleer of de karaktergrootte en het lettertype op normaal staan en de karakterafstand op proportioneel.

- 5 Ga na of het karakterwoordenboek READIRIS.DUS en het woordenboek in de modus nieuw staat.
- 6 Desactiveer zo nodig de "Pagina-analyse" in de knoppenbalk.
- 7 Sluit het Controlepaneel door op de knop "Sluiten" te klikken.

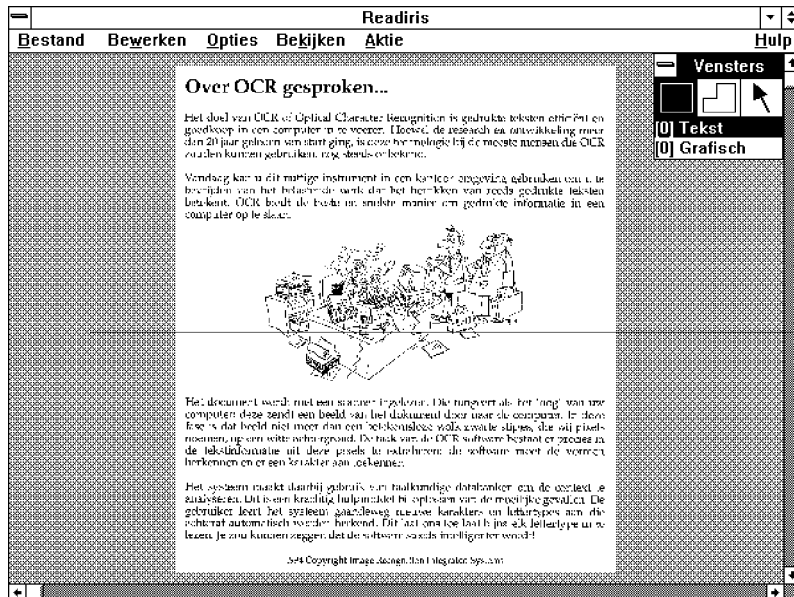
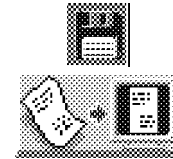
| Aktie               |    |
|---------------------|----|
| Controlepaneel...   | ^P |
| Scannen             | ^S |
| OCR                 | ^O |
| Interactief Lereren |    |
| ✓ Pagina-analyse    |    |
| Bekijken na Scannen |    |

## Een beeld openen

- 1 Zorg ervoor dat de optie "Bekijken na het Scannen" in het menu "Aktie" geactiveerd is.

Deze optie is actief indien er een vinkje naast staat. Klik op "Bekijken na het Scannen" als dit niet het geval is.

- 2 Klik op het schijf pictogram om de correcte invoerbron aan te geven.
- 3 Klik op de knop "Scannen" in de knoppenbalk. Je moet nu het beeldbestand gaan selecteren. Selecteer het voorafgescand beeld DUTCH.TIF in de Readiris directory.



- 4 Het beeld wordt van de schijf gelezen en verschijnt op het scherm. De venstertoolbox verschijnt eveneens.

## Inzoomen op een beeld

Je kan inzoomen op specifieke delen van het beeld.

- 1 Plaats je muis over dat deel van het beeld dat je wil uitvergroten.
- 2 Klik op de rechter muisknop om in te zoomen. Wanneer je nog eens klikt krijg je weer het volledige beeld.

Je kan eveneens de commando's onder het menu "Bekijken" of hun versnellingstoetsen gebruiken. Die kunnen zelfs worden gebruikt terwijl je de muis beweegt. Dat is zeker zinvol wanneer je blokken selecteert in een complexe pagina-layout.

## Tekst- en grafische vensters tekenen

De tweede fase bestaat hierin dat je aangeeft naar welke tekstblokken je interesse uitgaat. We zullen een tekstvenster tekenen rond de tekst die we willen herkennen en een grafisch venster rond de illustratie.



- 1 Selecteer de rechthoek in de venstertoolbox. De muiscursor wordt een kruisje.
- 2 Selecteer het venstertype: tekst.
- 3 Plaats de muis in de linkerbovenhoek van de tekstblok en klik.
- 4 Trek het venster open door de muis te verplaatsen naar de rechter onderhoek van de tekst en klik.

Het tekstvenster is getekend.

Een grafisch venster tekenen gebeurt op precies dezelfde manier, zij het dat je dan het grafisch venstertype selecteert in de venstertoolbox.

## Vensters wijzigen



- 1 Selecteer het venster door op het pijltje in de venstertoolbox te klikken en daarna op de rand van een venster te klikken.
- 2 Klik en houd de muisknop ingedrukt.
- 3 Verschuif de muis. Het venster volgt de muisbewegingen.

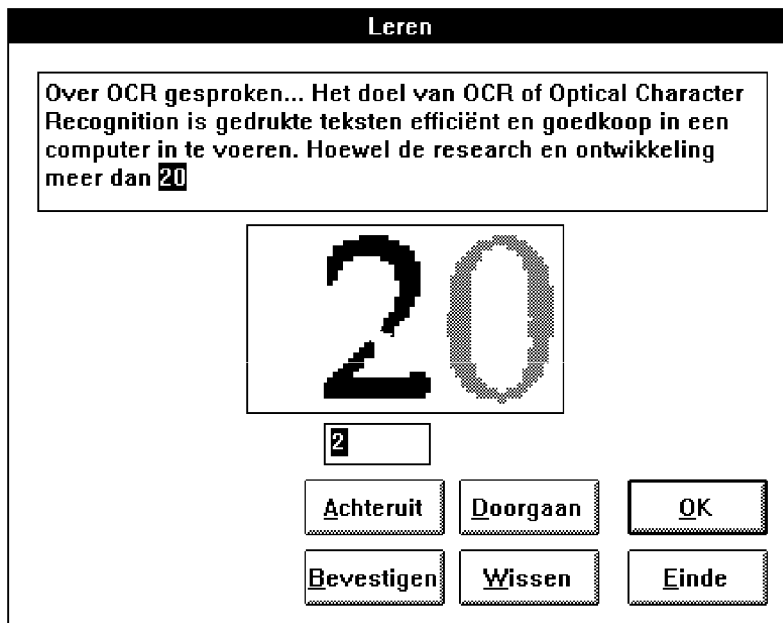
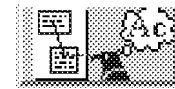
- 4 Laat de muisknop los wanneer het venster de juiste afmetingen heeft.

## Herkenning uitvoeren

Nu de vensters correct gedefinieerd zijn, bent u klaar voor de eigenlijke karakterherkenning.

- 1 Klik op de "OCR" knop in de knoppenbalk.

Aan het einde van de herkenning komt het programma terecht in de leermodule. De herkende tekst verschijnt op het scherm en het systeem stopt bij alle twijfelgevallen. Het systeem toont de twijfelachtige karakters in hun context en stelt een oplossing voor.



- 2 Klik op "OK" of druk op de ENTER-toets om een goed voorstel te aanvaarden; indien het systeem een fout maakt, tik je de correcte oplossing in eer je op "OK" klikt of op ENTER drukt.

Met "Doorgaan" verbeter je zonder dat het systeem bijleert. Negeer voorlopig de andere knoppen.

(Er is eigenlijk veel meer te vertellen over de leerfase. De precieze betekenis van deze knoppen zullen we pas in het volgend hoofdstuk diepgaand bespreken.)

- 3 Verlaat Readiris wanneer je klaar bent.

Aangezien de "OCR" knop ook grafieken opslaat, werd de grafische zone onzichtbaar opgeslagen.

## **De resultaten bekijken**

Je moet nu over een tekstbestand READIRIS.TXT en een grafisch bestand READIRIS.PCX beschikken in de huidige werkdirectory.

- 1 Start de Notepad op en open READIRIS.TXT.

De herkende tekst verschijnt nu op het scherm. Wat eerst een papieren dokument was is nu een editeerbare tekst geworden!

- 2 Verlaat Notepad.
- 3 Start Paintbrush en open READIRIS.PCX.

De grafische module van Windows toont nu het overgenomen beeld. Het staat je vrij de illustratie te wijzigen.

---

## **Een papieren dokument herkennen**



Wij gaan er nu van uit de scanner correct geïnstalleerd is.

## Scanner selecteren

- 1 Klik op de knop "Scanner" in de knoppenbalk. De ondersteunde scanners worden getoond.
- 2 Selecteer de naam van uw geïnstalleerde scanner en klik op "OK". Indien dat een vlakbedscanner is worden de beschikbare paginaformaten en resoluties getoond.
- 3 Voor OCR gebruiken we typisch een resolutie aan tot 300 dpi. Bij documenten van slechte kwaliteit kan je tot 400 dpi gaan.
- 4 Je kan de helderheid en het contrast desnoods aanpassen met een schuifbalk - indien deze opties beschikbaar zijn.  
  
Bij handscanners daarentegen wordt de resolutie en helderheid ingesteld met schakelaars op de scanner zelf.
- 5 Geef de correcte invoerbron aan door op de scanner in de knoppenbalk te klikken.

## Scannen met een vlakbedscanner

- 1 Plaats het dokument op het vlakbed van de scanner.
- 2 Selecteer de correcte scanoriëntatie in de knoppenbalk.
- 3 Klik op de knop "Scannen".

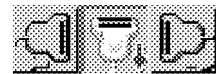


Het blad wordt gescand en op het scherm afgebeeld.

Indien de gekozen scanoriëntatie portret was, wordt het blad automatisch gerooteerd na het scannen.

## Scannen met een handscanner

- 1 Kies de correcte scanrichting in de knoppenbalk.
- 2 Klik op de knop "Scannen". De cursor verandert in een kleine handscanner.
- 3 Plaats uw scanner op dat deel van het document dat je wil inlezen. Zorg ervoor dat je document op een vlak oppervlak ligt.



- 4 Druk op de startknop van jouw scanner en schuif de scanner in de gekozen richting.

Beweeg de scanner langzaam en zo regelmatig mogelijk, zo niet kan het gescande beeld wel eens onbruikbaar zijn voor de OCR. Het beeld verschijnt op het scherm naarmate de scanner beweegt.

- 5 Druk na het scannen op de spatietoets om de controle aan Readiris terug te geven. Als de scanrichting van links naar rechts of vice versa was, wordt het beeld automatisch geroteerd.

---

## **Herkenning uitvoeren**

Nu kan je Readiris gebruiken net zoals we deden met het vooraf gescand beeld: inzoomen om het beeld even van naderbij te bestuderen, de tekstblokken selecteren, de herkenning uitvoeren enz.

## 3 Meer over Readiris

Dit hoofdstuk bevat meer details over de vele functies van Readiris.

Je zal leren hoe je een beeld scant en hoe je de kwaliteit van het gescande beeld kunt controleren. Eens het document correct gescand is kan je de tekstblokken aangeven die je wil overnemen of je kan Readiris je pagina-layout automatisch laten analyseren. Daarna voer je de karakterherkenning uit, waarbij een uitvoerbestand wordt gegenereerd.

We zullen ook specifiek ingaan op enkele geavanceerde mogelijkheden zoals automatische herkenning, de Connect mogelijkheid en het inlezen van tabellen.

---

### Startconfiguratie bepalen

Het eerste wat u moet doen, is uw startconfiguratie bepalen. Anders gesteld moet je de volgende zaken specificeren:

- de geïnstalleerde scanner en zijn parameters
- het bestandsformaat voor de tekst en illustraties
- de taal van de te herkennen tekst
- de parameters van het controlepaneel

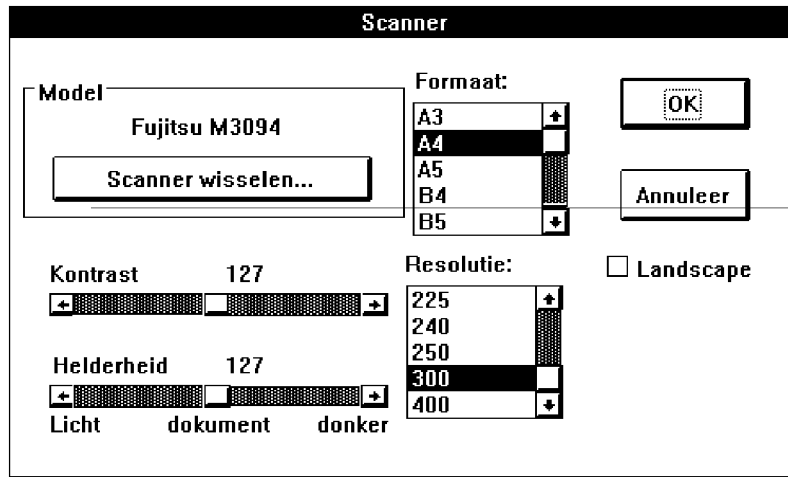
Deze parameters worden slechts zelden gewijzigd. Je kan die best opslaan als startconfiguratie, dan hoef je ze niet telkens opnieuw te definiëren.



## Een vlakbedscanner selecteren



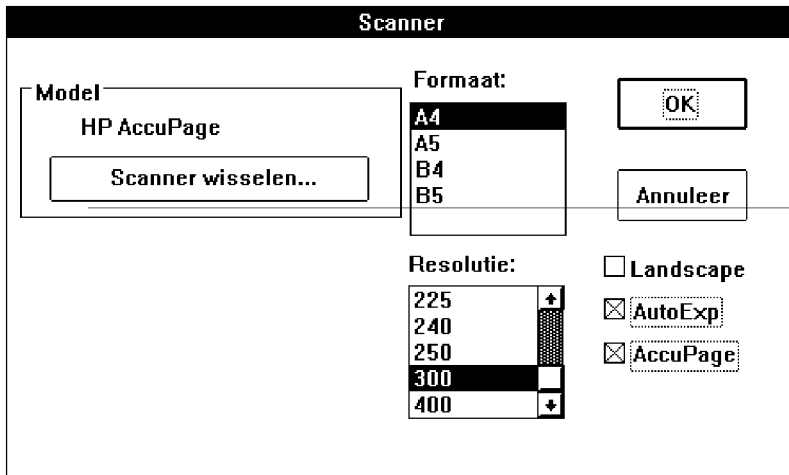
- 1 Klik op de knop "Scanner" in de knoppenbalk, en desnoods op de knop Scanner wisselen. De ondersteunde scanners worden getoond..



- 2 Selecteer de geïnstalleerde scanner en klik op "OK". De initiële resolutie, het paginaformaat, de helderheid en waar mogelijk het contrast verschijnen op het scherm.
- 3 Klik op het paginaformaat dat overeenstemt met de bladzijden die je normaal wil inlezen. Initieel is dat A4.
- 4 Selecteer de scanresolutie. Initieel is dat 300 dpi.
- 5 Stel de helderheid en mogelijkerwijze het contrast in met de schuifbalk. Initieel staat de schuifbalk op de middenpositie.

De optie contrast hoeft helemaal niet aanwezig te zijn op uw scanner. Door de helderheid en het contrast te regelen pas je het scannen aan donkere en lichte documenten aan. Bij HP scanners kan je de optie AutoExp activeren, waarbij de helderheid automatisch aan het document wordt aangepast.

Readiris ondersteunt eveneens de **HP AccuPage** technologie. Deze technologie past niet enkel de helderheid autonoom aan maar gebruikt ook geavanceerde beeldbewerkingstechnieken om



het gescande beeld te optimaliseren met het oog op de latere herkenning.

Deze technologie is enkel op HP scanners aanwezig. Selecteer de AccuPage "scanner" in plaats van een HP scanner en activeer de AccuPage optie.

De technische documentatie van jouw scanner geeft ongetwijfeld meer informatie over de afstelling en werking.

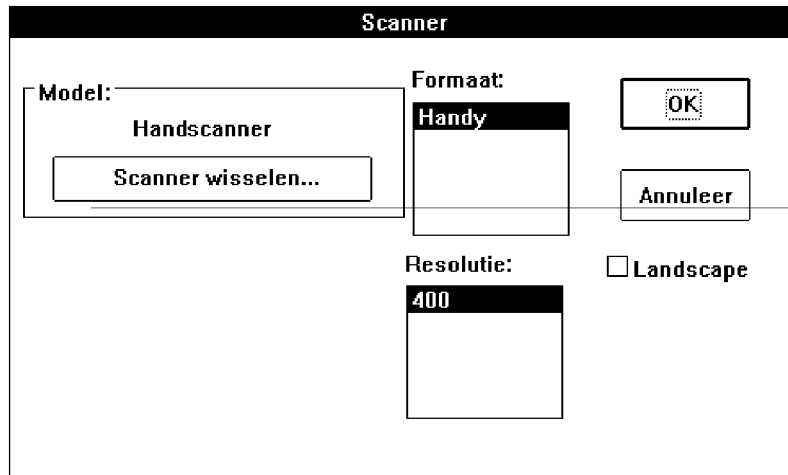
Het bestand READSCAN.TXT in de Readiris directory legt eveneens uit hoe je een scanner installeert.

## Een handscanner selecteren

- 1 Klik op de knop "Scanner" in de knoppenbalk.
- 2 Selecteer de geïnstalleerde handscanner en klik op "OK". De huidige resolutie wordt op het scherm afgebeeld. De resolutie en helderheid worden ingesteld met schakelaars op de scanner zelf.

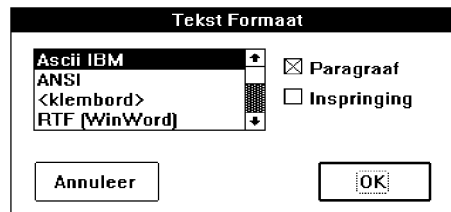
De technische documentatie van jouw scanner geeft ongetwijfeld meer informatie over zijn afstelling en werking.

Het bestand READSCAN.TXT in de Readiris directory legt eveneens uit hoe je een scanner installeert.



## Tekstformaten

- 1 Kies "Tekstformaat" onder het menu "Opties".



- 2 Selecteer het gewenste tekstformaat.

Readiris voert teksten rechtstreeks uit naar een groot aantal **tekstverwerkers** en **tekstformaten**. Gebruik het formaat ASCII of ANSI voor alle andere tekstverwerkers en editors. Readiris voert ook tabellen rechtstreeks uit naar **rekenbladen** in een generiek tabelformaat.

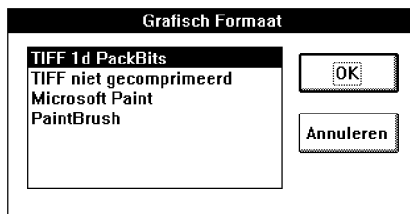
Het **klembord** "formaat" slaat de resultaten in het klembord op, waarna je die rechtstreeks in je Windows-toepassingen kan plakken.

Twee opties spelen in op de formattering van de uitvoer:

- **Paragraaf:** activeert de paragraafdetectie. Wanneer deze optie geactiveerd is, past Readiris de wordwrap toe die typisch is voor tekstverwerkers. U krijgt enkel een CR code op het einde van een paragraaf.
- **Inspringing:** activeert de codering van de tekstmarges. Deze optie is zeer nuttig om de vormgeving van de oorspronkelijke tekst te behouden. Activeer deze optie in geen geval wanneer de herkende tekst in een DTP-pakket moet worden ingevoerd.

## Grafisch formaat

- 1 Kies de functie "Grafisch Formaat" onder het menu "Opties".
- 2 Selecteer het grafisch formaat dat ondersteund wordt door jouw retoucheerprogramma.



Readiris zal de illustraties in dit grafisch formaat opslaan.

## Taal

Het is belangrijk dat de gebruiker de correcte taal van het te herkennen document selecteert, willen de **taalkundige databanken** een positieve rol spelen in het herkenningsproces.

- 1 Klik op de knop "Taal" in de knoppenbalk of kies het commando "Taal" onder het menu "Opties".
- 2 Selecteer die taal die je het meest herkent.

Readiris is uitgerust met linguïstische modules die de herkenning positief beïnvloeden. Je kan Readiris ook instellen op een **numerieke modus** om tabellen te herkennen die enkel cijfermateriaal bevatten.



## Startconfiguratie bewaren

Deze afstelling kan je meteen opslaan zodat je die niet telkens moet herdefiniëren wanneer je Readiris opstart.

Kies het commando “Startconfiguratie Opslaan” onder het menu “Bestand”. Readiris zal deze instelling laden zodra je het programma opstart.

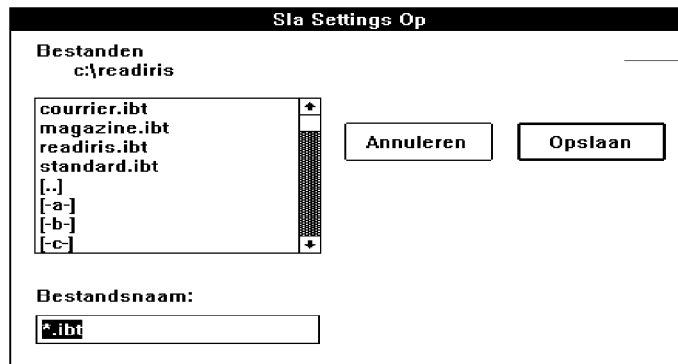
---

## Instellingen opslaan en laden

Je kan ook specifieke instellingen opslaan voor andere dan je basistoepassing. Deze parameters worden dan opgeslagen met het commando “Instellingen Opslaan” en geladen met het commando “Instellingen Laden” van het menu “Bestand”.

### Instelling opslaan

- 1 Kies “Instellingen Opslaan” onder het menu “Bestand”. Er verschijnt een lijst met alle settingsbestanden in jouw huidige directory.
- 2 Voer een nieuwe bestandsnaam in om de afstelling op te slaan in een nieuw bestand of beschrijf een bestaand bestand.



- 3 Klik op “Opslaan”.

De huidige configuratie van Readiris is opgeslagen worden in het vermelde bestand. Deze bestanden hebben initieel de extensie \*.ibt.

## Instellingen laden

- 1 Kies “Instellingen laden” onder het menu “Bestand”.  
De instellingen in jouw huidige directory worden getoond.
- 2 Selecteer het correcte bestand of ga naar een andere directory.
- 3 Klik op “Laden”.

---

## Alles over het controlepaneel

Eer je de karakterherkenning uitvoert op een document, moet je de kenmerken van het document aangeven. Het betreft voornamelijk:

- de taal en kwaliteit van het document
- het lettertype
- de letterafstand en -grootte

Je kan ook aanstippen of je al over een gebruikerswoordenboek beschikt voor het lettertype in kwestie.

Het controlepaneel geeft je controle over alle parameters die je moet aanpassen om de te herkennen documenten in te lezen. Het is onderverdeeld in drie delen:

- Dokument
- Aktie
- Bestanden

**Controlepaneel**

|  |   |
|--|---|
| <p><b>Dokument</b></p> <p><b>Pitch</b></p> <p><input type="radio"/> Vast</p> <p><input checked="" type="radio"/> Proportioneel</p> <p><b>Lettergrootte</b></p> <p><input type="radio"/> Klein</p> <p><input checked="" type="radio"/> Normaal</p> <p><b>Font</b></p> <p><input checked="" type="radio"/> Normaal</p> <p><input type="radio"/> Dot Matrix</p> | <p><b>Bestanden</b></p> <p>Invoer Scanner</p> <p>Uitvoer Tekst: (N) READIRIS.TXT<br/>Beeld: READIRIS.TIF</p> <p>Layout Default</p> <p>Dico (D) READIRIS.DUS</p> <p><b>Aktie</b></p> <p>1 Te lezen pagina's</p> <p><input checked="" type="checkbox"/> Pagina-analyse</p> <p><input checked="" type="checkbox"/> Bekijken na Scannen</p> <p><input checked="" type="checkbox"/> Interactief Lereren</p> <p><input type="checkbox"/> Strenge Controle</p> <p style="text-align: right;"><b>Scannen</b></p> <p style="text-align: right;"><b>Sluiten</b></p> |
|--|---|

## Dokument

Dit onderdeel van het controlepaneel bepaalt de kenmerken van het te herkennen document dat je wil herkennen. Je vindt er de volgende aspecten terug:

- de letterafstand: proportioneel of vast
- lettertype: normaal of dot matrix
- de lettergrootte: normaal of klein

Bij een **vaste letterafstand** hebben alle symbolen van het lettertype dezelfde breedte en een gelijke spatiering. Denk bijvoorbeeld aan de documenten die getikt werden met een mechanische schrijfmachine.

Bij een **proportioneel lettertype** daarentegen varieert de breedte van de symbolen naargelang hun vorm. Een "m" is bijvoorbeeld veel breder dan een "i". De overgrote meerderheid aan boeken en dagbladen zijn gedrukt in een proportioneel lettertype.

We noemen karakters **klein** wanneer ze kleiner zijn dan 10 punt en wanneer de spatiering tussen twee lijnen te klein is. In dit geval zou Readiris er niet in slagen twee opeenvolgende lijnen correct te segmenteren.

De gebruiker moet ook het lettertype aangeven: je kan geen **matrix-geprinte teksten** inlezen zonder dat je hier de correcte

parameter hebt ingesteld. Deze lettertekens zien er erg verbrokkeld uit, zijn gewoon opgeknipt of samengesteld uit individuele stippen. Neem het onderstaande voorbeeld.

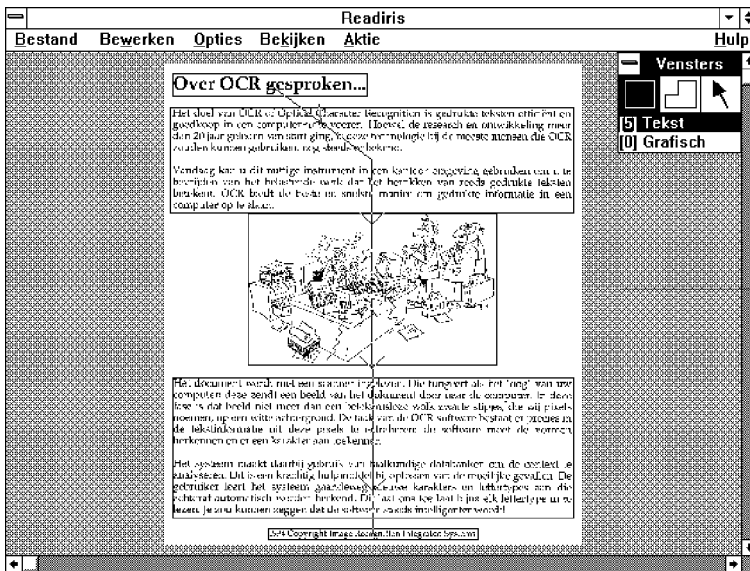
ape-descended life

## Aktie

Bepaalt de werkingsmodus van Readiris.

Wanneer de **pagina-analyse** geactiveerd is, hoef je de gescande bladzijden niet langer manueel te voorzien van vensters! De software detecteert en sorteert automatisch de tekst- en grafische blokken op de bladzijde. Dit is bijzonder nuttig bij in kolommen geschikte teksten en documenten met een complexe layout.

De pagina-analyse gaat er oorspronkelijk van uit dat elk gedetecteerd blok een tekstblok is. Grijp niet in wanneer er grafieken op het document voorkomen: in een tweede ziet de software dat daar geen karakters te herkennen zijn. (Evenmin worden deze vensters grafisch opgeslagen.)







De blokken worden initieel van boven naar onderen, van links naar rechts gesorteerd maar je kan de sorteerorde erg makkelijk wijzigen. Klik op de knop "Sorteren" en klik op een markeerder van de te herkennen vensters. De vensters die je dan niet aanklikt worden gewoon genegeerd!

Vermelden we eveneens dat de pagina-analyse snel en nauwkeurig werkt en tolerant is qua kromming van documenten.

Met de optie "Bekijken na het Scannen" activeer en desactiveer je de preview. Bij de **preview** worden de gescande documenten op het beeld getoond voor je ze verwerkt. (Bij handscanners heeft het geen zin deze optie te desactiveren.)

De derde parameter is het **interactief leren**. Deze activeert de leerfase die Readiris toelaat nieuwe informatie over lettertypes aan te leren. Desactiveer deze optie om meerdere bladzijden automatisch te herkennen en wanneer de herkenning uitstekend verloopt.

Je kan ook de graad van interactie beïnvloeden. Met de optie "Strenge Controle" geactiveerd, handhaaf je een hoge verificatiegraad: Readiris confronteert de gebruiker met zelfs de minste twijfel, neemt geen autonome beslissingen. Wanneer Readiris een tekst met groot gemak herkent, kan je je beperken tot het overlopen van die gevallen die heuse twijfelgevallen zijn.

## Bestanden

Hier vind je alle bestanden terug waarvan Readiris mogelijk gebruik maakt tijdens zijn werking. Er zijn vier soorten bestanden.

**Invoer:** bepaalt de invoerbron - je scant een bladzijde of opent een voorafgescand beeld..

**Uitvoer:** bepaalt waar je de uitvoer wegschrijft en welke uitvoermodus je hanteert - je kan bestaande bestanden beschrijven of aanvullen. Readiris creëert maximaal twee bestanden: één tekst- en één grafisch bestand.

**Layout:** bepaalt het layoutbestand. Je start met de initiële, lege layout of met een specifiek layoutbestand. Layouts zijn vensterbestanden, zij bevatten de vensterstructuur zoals die gedefinieerd werd gedurende de preview.

**Woordenboek:** geeft het woordenboek en zijn modus aan.

Het gebruikers- of letterwoordenboeken bevatten de karakterinformatie die het systeem aanleert tijdens de interactieve bevestigingsfase.

Er zijn drie woordenboekmodi:

- **Nieuw:** je creëert een nieuw letterwoordenboek. Readiris draait zonder al karakterinformatie aangeleerd te hebben.
- **Bijvoegen:** je maakt gebruik van aangeleerde karakterinformatie en je blijft Readiris trainen op bepaalde lettertypes.
- **Lezen:** je gebruikt de karakterinformatie die Readiris al opgeslagen heeft maar voegt geen nieuwe informatie toe.

Het is van essentieel belang de woordenboeken efficiënt te gebruiken om tot een optimale en snelle herkenning van documenten te komen. Dit impliceert dat je telkens de correcte modus activeert voor de gebruikerswoordenboeken.

Wanneer je slechts één pagina wil inlezen, selecteer je de modus nieuw. Om meerdere pagina's van hetzelfde type - we bedoelen documenten met hetzelfde lettertype en dezelfde drukwaliteit - in te lezen, ga je als volgt te werk: je kiest de modus nieuw voor de eerste pagina, de modus bijvoegen voor de tweede en de derde pagina en de leesmodus voor de resterende pagina's.

## Scannen met een vlakbedscanner

- 1 Selecteer de correcte scanoriëntatie. (Je kan ook de optie "Landschap" activeren door op de knop "Scanner" in de knoppenbalk te klikken.)
- 2 Klik op de knop "Scannen".



Het document wordt dan meteen gescand. Indien de preview geactiveerd is verschijnt het beeld gaandeweg op het scherm.

Indien de oriëntatie "Landschap" actief was, wordt het beeld meteen geroteerd. Je kan ook beelden achteraf roteren met het commando "Roteren" van het menu "Bekijken".

Besteedt enige aandacht aan de **schuinte**. Hoewel de pagina-analyse en de herkenning tolerant zijn qua kromming van documenten, wordt het wel eens moeilijk om een bladzijde correct te bevensteren en te herkennen wanneer de kromming te groot is.

Klik op de knop "Scanner" om de instelling van de scanner zoals resolutie, pagina-formaat e.d. te wijzigen.

## Scannen met een handscanner

Wanneer de nodige parameters ingesteld zijn ben je klaar om een document te scannen.

- 1 Selecteer de correcte scanrichting in de knoppenbalk.
- 2 Klik op de knop "Scannen" in de knoppenbalk. De cursor verandert in een kleine handscanner.
- 3 Helderheid en resolutie worden ingesteld met schakelaars op de handscanner. Gebruik een resolutie van 300 dpi bij documenten van goede kwaliteit, en een resolutie van 400 dpi bij documenten van slechtere kwaliteit.
- 4 Leg het te scannen document op een vlak oppervlak en plaats de scanner over dat deel van het document dat je wil herkennen.
- 5 Druk op de startknop van je scanner en beweeg die in de gewenste richting. Het beeld verschijnt op het scherm.

De beweging moet langzaam en zo gelijk mogelijk zijn, zoniet is het gescande beeld soms onbruikbaar voor de karakterherkenning.

De gebruikersinstructies van de fabrikant bevatten ongetwijfeld meer informatie over de manier waarop je de scanner moet gebruiken.

- 6 Druk wanneer je klaar bent met scannen op de spatietoets om de controle aan Readiris terug te geven.

Indien de gekozen scanrichting van links naar rechts was of vice versa, wordt het beeld direkt geroteerd. Je kan ook beelden achteraf roteren met het commando "Roteren" van het menu "Bekijken".

Besteed enige aandacht aan de **schuinte**. Hoewel de pagina-analyse en de herkenning tolerant zijn qua kromming van documenten, wordt het wel eens moeilijk om een bladzijde correct te bevensteren en te herkennen wanneer de kromming te groot is.

## Gescande beelden van nabij bekijken

Een volledige pagina ineens bekijken is handig als je de tekstblokken en illustraties wil overzien. Dit zicht noemen we “Hele Pagina”.

Om de kwaliteit van de scanning te controleren, moet je echter het beeld van nabij bekijken. Readiris biedt u drie zoomniveau's:

- Ware Grootte: benadert de oorspronkelijke grootte van de gedrukte pagina
- 50% Verkleining: de helft van de ware grootte
- 200% Vergroting: tweemaal de ware grootte

Je kan heen en weer gaan tussen “Ware Grootte” en “Hele Pagina” door op de *rechter* muisknop te klikken.

| Bekijken            |    |
|---------------------|----|
| ✓ Hele Pagina       | ^W |
| 50% Verkleining     | ^5 |
| Ware Grootte        | ^1 |
| 200% Vergroting     | ^2 |
| ✓ Alle Vensters     |    |
| Geselecteerd Type   |    |
| Sorteren            |    |
| ✓ Sorteren Bekijken |    |
| Sorteren Reset      |    |
| Roteren...          | ^R |
| Functies...         |    |
| Coördinaten...      | ^C |

## De kwaliteit van het gescande beeld controleren

De precisie van de optische karakterherkenning hangt sterk af van de kwaliteit van het scannen en het document. Even inzoomen op het gescande beeld is altijd nuttig om de kwaliteit te controleren. Readiris presteert het best op documenten van “goede” kwaliteit. We bedoelen hiermee documenten waarvan de karakters nauwkeurig gevormd zijn en vrij goed van elkaar gescheiden.

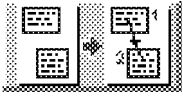
Bij te lichte documenten worden de karakters soms in twee delen gebroken: een 'O' kan dan makkelijk '()' worden. Bij te donkere documenten daarentegen worden de karakters erg zwaar. Open letters raken gesloten - een "c" wordt bijvoorbeeld een "o" - en beginnen alle karakters aan elkaar te kleven, wat uiteraard een zeer nadelige invloed heeft op de herkenningsresultaten.

Je kan de kwaliteit van het gescande beeld corrigeren door de helderheid van je scanner bij te stellen.

Eens het document opnieuw gescand is, kan je de verbetering op het scherm verifiëren.

## Bevenstering

Wanneer het document op het scherm is afgebeeld, bestaat de volgende stap hierin dat je de zones die van belang zijn, aangeeft.



Er zijn meerdere manieren om dat te doen: de pagina-analyse detecteert en ordent de verschillende blokken automatisch. Je kan de resulterende layout wijzigen met met de knop "Sorteren". De andere strategie houdt in dat je manueel vensters tekent. Ten derde kan je specifieke vensterstructuren opslaan in bestanden, "layouts", voor toekomstig gebruik.

Meerdere kolommen detecteren of tekenen is noodzakelijk wanneer een tekst in **kolommen** is geschikt. Om tabellen correct te herkennen, moet je vensters trekken om elke kolom, zodat de uitvoer netjes in tabelcellen kan worden geschikt.

Nu volgt de beschrijving van de Readiris venstertools.

### Rechthoekige vensters creëren



- 1 Selecteer de rechthoek door in de venstertools te klikken.
- 2 Selecteer het venstertype: tekst of grafisch.
- 3 Plaats de muiscursor op de hoek van het venster dat je wil creëren en klik.
- 4 Trek het venster open door de muis te verschuiven naar de diagonaal tegenovergestelde hoek en klik nogmaals.

Het venster is getekend.

Zijden kleiner dan 1 mm zijn niet toegestaan!

### Veelhoekige vensters creëren



- 1 Selecteer de veelhoek door in de venstertools te klikken.
- 2 Plaats de muiscursor waar je het venster wil creëren en klik.
- 3 Verplaats de muis. Eén zijde van de veelhoek wordt getekend naarmate de muis opschuift.
- 4 Klik om de eerste zijde te voltooien.

- 5 Verplaats de cursor en klik eender waar om een tweede zijde te voltooien.
- 6 Klik op de *rechter* muisknop om de veelhoek te sluiten.

Je kan geen zijde kruisen die je al getekend hebt!

Je kan ook veelhoeken creëren door rechthoeken te versmelten. Wanneer twee vensters van hetzelfde type elkaar raken, versmelten zij tot één enkel venster.

## Vensters selecteren

- 1 Selecteer de pijl door in de venstertools te klikken.
- 2 Klik op een vensterzijde.



Er verschijnen markeerders op elke hoek en in het midden van de zijden.

- 3 Om een aanvullend venster te selecteren, zonder de selectie van het eerste venster te verliezen, houd je de Shift-toets ingedrukt wanneer je op het tweede venster klikt.
- 4 Om een venster te selecteren met alle vensters die erin zitten, houd je de Ctrl-toets ingedrukt terwijl je op het grootste venster klikt.
- 5 Om vensters te deselecteren, klik je ergens buiten de vensters.

Je kan het type en de **coördinaten** van een venster op het scherm tonen door op de Alt-toets te drukken terwijl je het venster selecteert.

## Vensters wijzigen

- 1 Selecteer het venster. Zie hierboven.
- 2 Plaats de muis op een markeerder.
- 3 Klik en houdt de muisknop ingedrukt.
- 4 Verplaats de muis.

De afmetingen van het venster wijzigen in overeenstemming met de muisbewegingen.

- 5 Laat de muisknop los wanneer het venster de juiste afmetingen heeft.

### Vensters verplaatsen

- 1 Selecteer het venster. Zie hierboven.
- 2 Plaats de muis op een vensterzijde, tussen twee markeerders in.
- 3 Klik en houdt de muisknop ingedrukt.
- 4 Verplaats de muis. Het venster wordt verplaatst in overeenstemming met de muisbewegingen.
- 5 Laat de knop los wanneer het venster zich op de juiste plaats bevindt.

### Vensters wissen

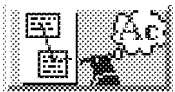
- 1 Selecteer het venster. Zie hierboven.
- 2 Selecteer het commando "Knippen" of "Wissen" van het menu "Bewerken".

Bij "Knippen" slaat u een interne copie op van het venster, kunt u het venster weer plakken, bij "Wissen" is het venster onherstelbaar verloren..

| Bewerken         |          |
|------------------|----------|
| Ongedaan maken   | Alt+BkSp |
| Knippen          | Sh Del   |
| Kopiëren         | *Ins     |
| Plakken          | Sh Ins   |
| Wissen           | Del      |
| Alles Selecteren | *A       |

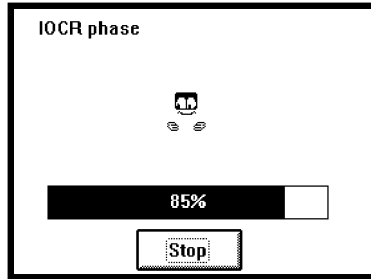
---

## Karakterherkenning



Klik op de knop "OCR" in de knoppenbalk om de karakterherkenning uit voeren op die getekende of gedetecteerde vensters, ook wanneer je enkel illustraties wil opslaan.

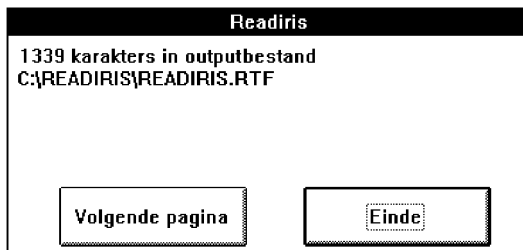
Het verloop van de herkenning wordt aangegeven; je kan ten allen tijde onderbreken door op de "Stop"-knop te klikken.



Daar Readiris ook **multitasking** ondersteunt, kan je de herkenning ook in de achtergrond laten verlopen. Hierleidt het Readiris venster tot een pictogram en voer een andere taak uit op je PC; de interactieve fase meldt zich wanneer het systeem klaar is.

Tijdens die **interactieve fase**, verschijnt er een dialoogvenster waarin de herkende tekst afgebeeld wordt. Wanneer het systeem twijfels heeft over een bepaalde oplossing, verschijnt het beeld van het woord uitvergroot op het scherm. Het karakter waarover twijfel bestaat wordt geaccentueerd, de software stopt en wacht op een beslissing van de gebruiker.

Na de beslissingsfase wordt de herkende tekst opgeslagen in het uitvoerbestand of klembord en vraagt het systeem je of een volgende bladzijde wil inlezen.



Bij meerbladige documenten, zorgt de knop "Volgende bladzijde" er voor dat het gebruikerswoordenboek en uitvoerbestand in de modus aanvullen worden gezet. Indien u een vlakbedscanner heeft, wordt de volgende bladzijde meteen gescand, als u daarentegen een handscanner hebt, bent u klaar om te scannen.

Je zult zeker merken dat de leerfase veel minder inspanning vraagt bij de volgende bladzijden wanneer die hetzelfde lettertype hebben.

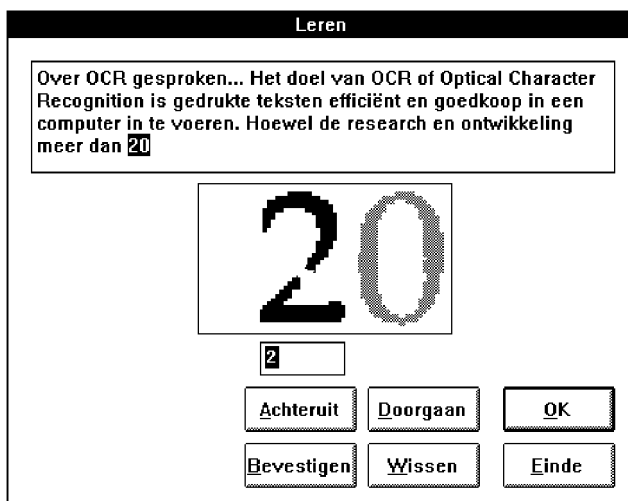


---

## Lettertypes aanleren

De **leerfase** is de laatste stap van het herkenningsproces. (De graad van interactie is te beïnvloeden met de optie "Strengere Controle" in het "Controlepaneel".)

Indien Readiris twijfels heeft over een herkend resultaat, wordt het twijfelachtige karakter in zijn context op het scherm gebracht. Dit karakter - het kan ook om een karakterstring gaan wanneer de karakters niet gescheiden zijn - wordt beklemtoond en de voorgestelde oplossing verschijnt. Er verschijnt een "~" als het karakter niet herkend werd.



Geef als dat nodig is de correcte karakters in en selecteer één van de volgende knoppen.

### OK

Je aanvaardt of corrigeert de voorgestelde oplossing. Het resultaat wordt in het karakterwoordenboek opgeslagen als zeker, definitief: in de toekomst zal de herkenning van deze vorm automatisch verlopen.

Alternatief: druk op de ENTER-toets.

## **Doorgaan**

Je aanvaardt of corrigeert de voorgestelde oplossing. Het verschil met "OK" is dat Readiris het resultaat in het gebruikerswoordenboek opslaat als onzeker, als voorlopig. Toekomstige herkenning zal nog steeds uw tussenkomst vergen.

Gebruik deze knop met name bij beschadigde karakters die wel eens verward kunnen worden met andere karakters indien je "OK" kiest.

## **Bevestigen**

Het cijfer '1' en de letter "l" hebben bij veel lettertypes dezelfde vorm. In dat geval voert Readiris een contextanalyse uit om ze uit elkaar te houden. Waar deze analyse niet volstaat, moet de gebruiker ingrijpen.

Gebruik deze knop indien het cijfer '1' en de letter "l" duidelijk verschillen, zoniet klik je op "Doorgaan".

## **Achteruit**

Je kan terugkeren om fouten te corrigeren. Readiris houdt de laatste negen tussenkomsten bij.

## **Wissen**

Wist karakters uit de uitvoer. Er zitten bijvoorbeeld stippen op het document die herkend worden als punten, komma's enz.

## **Einde**

Onderbreekt het interactieve leerproces. De beslissingfase wordt automatisch afgewerkt. Je aanvaardt alle oplossingen van de software.

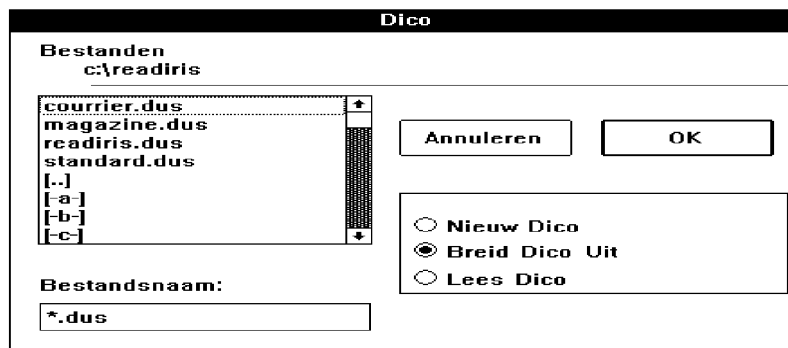
## De efficiëntie van woordenboeken

Wanneer je Readiris gebruikt om veel documenten van hetzelfde type in te lezen, zal je allicht hetzelfde leerproces niet telkens opnieuw doorlopen. Je kan inderdaad de aangeleerde karakters in een karakterwoordenboek opslaan en van die verworven intelligentie later handig gebruik maken wanneer je gelijkaardige documenten inleest.

Een gebruikerswoordenboek creëren of activeren gaat als volgt:

- 1 Kies het "Controlepaneel" uit het menu "Aktie".
- 2 Klik op de knop "Dico".

De woordenboeken in de huidige directory worden getoond. Hun normale extensie is \*.dus.



- 3 Tik een bestandsnaam in of selecteer een bestaand bestand.
- 4 Klik op een woordenboekmodus: "Nieuw" om een nieuw woordenboek te creëren, "Uitbreiden" om een bestaand woordenboek verder aan te vullen en "Lezen" om gebruik te maken van een woordenboek zonder er nieuwe gegevens in op te slaan.
- 5 Klik op "OK".

Het actieve woordenboek wordt vermeld naast de "Dico"-knop in het "Controlepaneel"; de geselecteerde modus wordt tussen haakjes aangegeven.

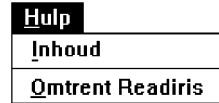
---

## Hulpinformatie opzoeken

Het algemeen overzicht van de werking van Readiris is hiermee afgesloten. Readiris bevat overigens een **on-line hulpsysteem** dat elk aspect van de software uitgebreid documenteert. U kunt er ook wel eens wat informatie van het laatste ogenblik terugvinden.

Selecteer het commando "Inhoud" onder het menu "Hulp" om toegang te krijgen tot die informatie. Versnellingstoets: druk op F1. "Omtrent Readiris" toont uw startscherm met uw registratienummer.

Daarnaast bevat Readiris ook een reeks **geavanceerde functies** die we meer in detail moeten behandelen. We zullen met name uitleggen hoe je tabellen herkent, hoe je automatische herkenning uitvoert en hoe de Connect mogelijkheid werkt.



---

## Tabellen herkennen

Je kan Readiris gebruiken om cijfertabellen rechtstreeks in jouw rekenblad in te voeren.

Je moet wel een venster tekenen rond elke kolom om de uitvoer netjes in overeenkomstige cellen te kunnen ordenen, anders zul je er nooit in slagen tabellen correct in te voeren. De onderstaande illustratie toont je hoe dat moet.

Selecteer één van beide tekstformaten: Excel ASCII of Excel klembord. Hoewel beiden een generieke tabelfilter bevatten, zal het formaat Excel ASCII een tabelbestand creëren dat rechtstreeks kan worden ingevoerd in je rekenblad, terwijl het formaat Excel klembord de herkende tabellen in het klembord van Windows plaatst. Gebruik achteraf het commando "Plakken" van je rekenblad om de tabel in te voegen.

De detectie van paragrafen en de inspringing zijn niet van kracht wanneer je een tabelformaat selecteert.

| Readiris                               |                |                 |                |                 |
|--|----------------|-----------------|----------------|-----------------|
| Bestand Bewerken Opties Bekijken Aktie |                |                 |                | Hulp            |
| MB                                     | First Platform | Second Platform | Third Platform | Fourth Platform |
| 0.0                                    | E2,857         | E2,143          | E2,143         | E2,143          |
| 0.1                                    | E2,921         | E2,207          | E2,207         | E2,207          |
| 0.5                                    | E3,179         | E2,464          | E2,464         | E2,464          |
| 1.0                                    | E3,500         | E2,786          | E2,786         | E2,786          |
| 2.0                                    | E4,143         | E3,429          | E3,429         | E3,429          |
| 3.0                                    | E4,786         | E3,929          | E3,929         | E3,929          |
| 4.0                                    | E5,429         | E4,286          | E4,286         | E4,286          |
| 5.0                                    | E6,071         | E4,643          | E4,643         | E4,643          |
| 6.0                                    | E6,714         | E4,957          | E4,957         | E4,957          |
| 7.0                                    | E7,357         | E5,271          | E5,271         | E5,271          |
| 8.0                                    | E7,921         | E5,500          | E5,500         | E5,500          |
| 9.0                                    | E8,571         | E5,714          | E5,714         | E5,714          |
| 10.0                                   | E9,214         | E5,929          | E5,929         | E5,929          |
| 11.0                                   | E9,857         | E6,143          | E6,143         | E6,143          |
| 12.0                                   | E10,500        | E6,357          | E6,357         | E6,357          |
| 13.0                                   | E11,143        | E6,571          | E6,571         | E6,571          |
| 14.0                                   | E11,786        | E6,786          | E6,786         | E6,786          |
| 15.0                                   | E12,429        | E7,000          | E7,000         | E7,000          |
| 16.0                                   | E13,071        | E7,214          | E7,214         | E7,214          |
| 17.0                                   | E13,714        | E7,429          | E7,429         | E7,429          |
| 18.0                                   | E14,357        | E7,643          | E7,643         | E7,643          |
| 19.0                                   | E15,000        | E7,857          | E7,857         | E7,857          |
| 20.0                                   | E15,643        | E8,071          | E8,071         | E8,071          |
| 25.0                                   | E16,286        | E8,286          | E8,286         | E8,286          |
| 30.0                                   | E16,929        | E8,500          | E8,500         | E8,500          |
| 35.0                                   | E17,571        | E8,714          | E8,714         | E8,714          |
| 40.0                                   | E18,214        | E8,929          | E8,929         | E8,929          |
| 45.0                                   | E18,857        | E9,143          | E9,143         | E9,143          |
| 50.0                                   | E19,500        | E9,357          | E9,357         | E9,357          |
| 55.0                                   | E20,143        | E9,571          | E9,571         | E9,571          |
| 60.0                                   | E20,786        | E9,786          | E9,786         | E9,786          |

Numeriek



Readiris kan de herkenning beperken tot een numerieke karakterset wanneer cijfermateriaal zonder tekst wordt ingelezen. Selecteer de **numerieke modus** met de keuzelijst "Taal" in de knoppenbalk.

## Automatische herkenning

Je kan Readiris automatische herkenning laten uitvoeren. Hoe automatisch die herkenning precies verloopt hangt af van het scannertype.

| Aktie               |    |
|---------------------|----|
| Controlepaneel...   | ~P |
| Scannen             | ~S |
| OCR                 | ~O |
| Interactief Leren   |    |
| √ Pagina-analyse    |    |
| Bekijken na Scannen |    |

Met **flatbedscanners** ligt volautomatische herkenning van vele bladzijden in het bereik. Desactiveer de opties "Bekijken na Scannen" en "Interactief Leren" en activeer de optie "Pagina-analyse" in het menu "Aktie". (Je kan ook een layout gebruiken om de zones van belang aan te geven.) Voer het aantal te herkennen bladzijden in het "Controlepaneel" in. Plaats een pak bladzijden in de documentaanvoer van de scanner, klik op de "Scan"-knop en laat de rest over aan je PC. De bladzijden zullen worden gescand, bevestigd en herkend zonder enige tussenkomst van de gebruiker! Als de aanvoer leeg raakt eert het

vermelde aantal bladzijden herkend werd, wordt een foutmelding getoond.

Bij **handscanners** liggen de zaken anders. Die vergen nu eenmaal altijd enige inspanning van de gebruiker bij het scannen. Volautomatische herkenning is er dus niet bij, maar niets belet je om de bevenstering en herkenning automatisch te laten verlopen.

Omdat bij automatische herkenning de bevestigingsfase overgeslagen wordt, verdient het aanbeveling Readiris eerst op enkele bladzijden te trainen. Zoniet kan het herkenningspercentage wel eens wat minder zijn.

---

## De Connect-mogelijkheid gebruiken

De Connect-mogelijkheid opent het perspectief op een rechtstreekse OCR-brug tussen je scanner en je Windows-toepassing. Je kan inderdaad vanuit een Windows-toepassing teksten herkennen!

Na de installatie van de Connect-functie verschijnt een nieuw commando in de menu's van je Windows-toepassing. Klik deze functie aan en Readiris wordt opgestart. Je scant en herkent de tekst en de uitvoer wordt in je Windows-toepassing ingevoerd op de huidige cursor-positie.

### WordPerfect, Word en Excel

Ga na of Readiris niet in de achtergrond draait.

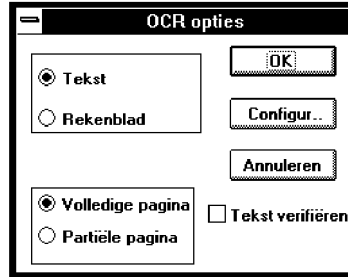
Klik op het nieuw commando "Tekst Scannen" in het menu van jouw toepassing om Readiris op te starten. Bij WordPerfect 5.x maakt dit commando deel uit van het menu "Macro", bij WordPerfect 6.0 van het commando "Macro" onder het menu "Extra" en bij Word en Excel van het menu "Bestand".

De Connect interface kan je enkele vragen stellen eer Readiris wordt opgestart. Ga te werk zoals gewoonlijk en verlaat de Connect interface weer wanneer je klaar bent. De resultaten van de herkenning worden ingevoerd op de huidige cursor-positie.

### Andere Windows-toepassingen

Klik twee keer op het Connect pictogram om de Connect task server in de achtergrond te activeren eer je de Windows-toepassing opstart.

Klik op het nieuw commando "OCR" dat aan het menu "Bestand" van uw toepassing is toegevoegd. Een dialoogvenster confronteert u met enkele parameters.



Geef aan of je een lopende tekst ("Tekst") of een cijfertabel ("Rekenblad") wenst in te lezen. Selecteer "Volledige Pagina" om de pagina-analyse uit te voeren en "Gedeeltelijk" om het gescande document manueel te bevensteren.

Activeer de optie "Tekst Verifiëren" om de interactieve fase te doorlopen. Initieel is die niet geactiveerd.



De knop "Config" geeft toegang tot enkele geavanceerde parameters. Selecteer "Aanslagen simuleren" om de herkende tekst in te voeren alsof die op het klavier werd ingetikt of "Plakken vanuit klembord" om de tekst onzichtbaar via het klembord over te brengen.

Ga na of de correcte versnellingstoets is vermeld voor het commando Plakken van jouw Windows-toepassing. Enkele toetsen worden voorgesteld door een speciaal karakter.

| <b>TOETS</b> | <b>VOORSTELLING</b> |
|--------------|---------------------|
| Shift        | +                   |
| Ctrl         | ^                   |
| Alt          | %                   |
| Insert       | {INSERT}            |

Enkele voorbeelden: wanneer de versnellingstoets voor het commando Plakken Ctrl+V is, tik je hier **^V** in, wanneer de versnellingstoets Shift+Insert is geef je **+{INSERT}** in enz.





## 4 Commando's

Dit hoofdstuk geeft een kort overzicht van alle menu's van Readiris. Je bent zeker al bekend met deze commando's indien de vorige hoofdstukken aandachtig gelezen hebt.

---

### Bestand

Geeft je de mogelijkheid om instellingen en vensterlayouts op te slaan en te activeren. Laat toe vooraf gescande beelden te openen.

#### Openen

Opent een beeldbestand met TIFF-formaat. Enkel actief in de previewmodus.

#### Info

Toont de huidige Readiris-configuratie.

#### Bron Selecteren

Readiris ondersteunt de Twain scanner-standaard. Dit commando laat toe een Twain scanner te selecteren en de scanparameters in te stellen.



## **Verwerven**

Start het scannen. Equivalent met het commando "Scannen" in het menu "Aktie" en de "Scan"-knoppen in de knoppenbalk en het controlepaneel.

## **Layout opslaan**

Slaat een vensterstructuur zoals die op een document wordt gedefinieerd tijdens de preview. Layouts worden opgeslagen in een bestand met de default extensie .wdw. Enkel actief tijdens de preview.

## **Layout laden**

Laadt een vensterstructuur die werd opgeslagen in een layoutbestand met default extensie .wdw. Enkel actief tijdens de preview.

## **Instellingen opslaan**

Slaat de huidige Readiris configuratie op in een bestand met default extensie .ibt.

## **Instellingen laden**

Laadt de huidige Readiris configuratie zoals opgeslagen in een configuratie bestand.

## **Startconfiguratie opslaan**

Slaat de huidige Readiris configuratie op als startconfiguratie.

## **Verlaten**

Sluit Readiris af.

## Bewerken

Deze commando's zijn enkel actief tijdens de preview: zij hebben betrekking op de vensters die worden gedefinieerd wanneer een document op het scherm wordt afgebeeld.

Om een venster te selecteren klik je op één van de zijden. Om meerdere vensters tegelijk te selecteren, houd je de Shift toets ingedrukt terwijl je op de vensterzijden klikt.

| Bewerken         |          |
|------------------|----------|
| Ongedaan maken   | Alt+BkSp |
| Knippen          | Sh Del   |
| Kopiëren         | *Ins     |
| Plakken          | Sh Ins   |
| Wissen           | Del      |
| Alles Selecteren | *A       |

### Ongedaan maken

Annuleert het laatste commando.

### Knippen

Wist de geselecteerde vensters maar slaat ze op in een interne buffer.

### Copiëren

Copieert de geselecteerde vensters in een interne buffer.

### Plakken

Voegt de inhoud van de interne buffer toe. De vensters worden daar ingevoegd waar zij oorspronkelijk geknipt of gecopieerd werden. Om ze te verplaatsen, klik je op een zijde en sleep je de muis: de vensters volgen de muisbewegingen.

### Wissen

Wist alle geselecteerde vensters zonder mogelijkheid tot herstel.

### Alles Selecteren

Selecteert alle vensters ineens.

| Opties  |    |
|---|----|
| Scanner...                                      | ^Q |
| Tekstformaat...                                 |    |
| Grafisch Formaat...                             |    |
| Taal...   | ^L |
| <input checked="" type="checkbox"/> Knoppenbalk |    |

---

## Opties

Bepaalt de basisinstelling van Readiris: welk tekst- en grafisch formaat worden gebruikt, geeft de taal van het document aan. Activeert ook de knoppenbalk.

## Scanner

Toont het scannermodel en zijn afstelling. Equivalent met de knop "Scanner" op de knoppenbalk.

## Tekstformaat

Bepaalt het tekstformaat voor de herkende karakters.

Activeert de paragraafdetectie en de detectie van de insprong. Om tabellen te herkennen moet je wel het Excel formaat selecteren.

## Grafisch Formaat

Bepaalt het grafisch formaat voor de grafische vensters. Ga na welk bestandsformaat uw grafische toepassingen kunnen bewerken.

## Taal

Selecteert de taal van het document. Readiris is uitgerust met vele taalkundige databanken. Je kan ook de herkenning beperken tot numerieke karakters, bijv. om cijfertabellen die geen tekst bevatten optimaal en erg snel te herkennen.

## Knoppenbalk

Activeert de knoppenbalk. De knoppenbalk verschijnt aan de linkerkant van het Readiris venster. Wij bevelen het gebruik van de knoppenbalk aan omdat die het makkelijkst toegang geeft tot alle frequente commando's.

## Bekijken

Deze commando's zijn enkel beschikbaar tijdens de preview van een document.

### Hele Pagina

Past het zoomniveau zodanig aan dat het beeld van uw document volledig op uw scherm wordt getoond.

### 50% Verkleining

Gebruikt een zoomniveau van 50%.

### Ware Grootte

Zoomt in op de reële grootte van uw document: één scherpixel stemt overeen met één beeldpixel. Klik op de *rechter* muisknop om heen en weer te schakelen tussen "Ware Grootte" en "Hele Pagina".

### 200% Vergroting

Gebruikt een zoomniveau van 200%.

### Alle Vensters

Toont alle vensters die deel uitmaken van de vensterlayout .

### Geselecteerd Type

Toont enkel de vensters van het geselecteerd type.

### Sorteren

Sorteert de tekstvensters. Bepaalt in welke volgorde de verschillende tekstvensters in de uitvoer zullen voorkomen. Dit

| Bekijken            |    |
|---------------------|----|
| ✓ Hele Pagina       | ^W |
| 50% Verkleining     | ^5 |
| Ware Grootte        | ^1 |
| 200% Vergroting     | ^2 |
| ✓ Alle Vensters     |    |
| Geselecteerd Type   |    |
| Sorteren            |    |
| ✓ Sorteren Bekijken |    |
| Sorteren Reset      |    |
| Roteren...          | ^R |
| Functies...         |    |
| Coördinaten...      | ^C |

commando is erg nuttig bij teksten die in kolommen zijn geschikt. Klik op een zijde van de vensters om de orde aan te geven.

### **Sorteren Bekijken**

Toont aan in welke volgorde de vensters gesorteerd zijn. Pijlen verbinden de verschillende tekstvensters.

### **Sorteren Reset**

Annuleert de huidige sorteerorde. Alternatief: gebruik de "Sorteer"-knop in de knoppenbalk.

### **Roteren**

Roteert het gescande beeld. Kies een rotatierichting: met de klok mee of tegen de klok in. We bevelen jouw aan om voor het scannen al de goede richting in te stellen.

### **Funkties**

Toont de venstertools, waar je het venstertype (tekst of grafisch) en de venstervorm (rechthoek, veelhoek of pijl) kan kiezen.

### **Coördinaten**

Displayt het coördinatenvenster. De coördinaten van de cursor worden afgebeeld in document pixels.

---

## **Aktie**

Bevat de commando's die te maken hebben met de eigenlijke herkenning: scannen en OCR, en alle parameters die daarbij een rol spelen.

## Controlepaneel

Toont het Controlepaneel.

| Akte                |    |
|---------------------|----|
| Controlepaneel...   | ^P |
| Scannen             | ^S |
| OCR                 | ^O |
| Interactief Leren   |    |
| √ Pagina-analyse    |    |
| Bekijken na Scannen |    |

## Scannen

Start het scannen van een document of opent een vooraf gescand beeld, afhankelijk van de invoerbron. Is equivalent aan de knop "Scannen" in de knoppenbalk.

## OCR

Start de karakterherkenning en slaat de illustraties op. Equivalent met de knop "Scannen" in de knoppenbalk.

## Interactief Leren

Activeert de verificatiefase. Bepaalt of je de twijfelgevallen zult controleren of gewoon alle voorstellen van het systeem aanvaardt. (De optie "Strengere Controle" bepaalt de graad van verificatie.)

De leermodus is ook beschikbaar op de knoppenbalk. Desactiveer deze modus om automatische OCR uit te voeren.

## Pagina-analyse

Activeert de pagina-analyse. Bepaalt of het scannen al of niet wordt gevolgd door een automatische layout-analyse.

Is ook beschikbaar op de knoppenbalk. Activeer deze optie om automatische herkenning uit te voeren.

## Bekijken na het Scannen

Activeert de preview: het gescande document wordt op jouw scherm afgebeeld. Zo kan je de tekstblokken isoleren die herkend moeten worden en de illustraties die Readiris moet opslaan. Deze optie desactiveren heeft geen zin bij handscanners.



---

## Controlepaneel

Alle elementen van het controlepaneel worden in detail beschreven in hoofdstuk 3, "Readiris beter leren kennen".

### Letterafstand

Selecteert de letterafstand: vast of proportioneel.

### Lettergrootte

Selecteert de karaktergrootte: klein of normaal. We noemen een karakter "klein" wanneer de karaktergrootte kleiner is dan 10 punt of wanneer de software er niet in slaagt twee opeenvolgende lijnen te segmenteren.

### Lettertype

Geeft de drukkwaliteit van de tekst aan: je leest "normale" teksten of teksten die van een matrix printer afkomstig zijn.

### Te Lezen Pagina('s)

Bepaalt het aantal automatisch in te lezen pagina's. Deze optie is niet beschikbaar bij handscanners.

### Pagina-analyse

Activeert de pagina-analyse. Bepaalt of het scannen al of niet wordt gevolgd door een automatische layout-analyse.

Is ook beschikbaar op de knoppenbalk. Activeer deze optie om automatische herkenning uit te voeren.

### Bekijken na het Scannen

Activeert de preview. Equivalent aan het "Bekijken na het Scannen" in het menu "Aktie".

## **Interactief Leren**

Activeert de verificatiefase. Bepaalt of je de twijfelgevallen zult controleren of gewoon alle voorstellen van het systeem aanvaardt.

De leermodus is ook beschikbaar op de knoppenbalk. Desactiveer deze modus om automatische OCR uit te voeren. De volgende optie bepaalt de graad van verificatie.

## **Strengere Controle**

Wanneer deze optie geactiveerd is, is de verificatiegraad hoog: Readiris zal de gebruiker confronteren met elk twijfelgeval en nooit autonoom beslissingen nemen. Wanneer Readiris een tekst met groot gemak herkent, kan de gebruiker zich beperken tot de verificatie van die gevallen die een reële twijfel inhouden.

## **Invoer**

Selecteert de invoerbron. Je gebruikt een scanner of opent vooraf gescande beelden. Beide invoerbronnen zijn eveneens beschikbaar op de knoppenbalk.

## **Uitvoer**

Bepaalt de naam van het uitvoerbestand. Als het bestand al bestaat, kies je tussen de modus vervangen of aanvullen. De extensie van het uitvoerbestand wordt automatisch toegevoegd op basis van het gekozen uitvoerformaat.

## **Layout**

Selecteert het layoutbestand dat de vensterstructuur bevat die aangepast is aan het documenttype. De default extensie is .wdw.

## **Dico**

Selecteert het gebruikerswoordenboek en de woordenboekmodus (nieuw, bijvoegen of lezen). De default extensie is .dus.

## Scannen

Dit commando start het scannen. Is ook beschikbaar op de knoppenbalk.

## Sluiten

Sluit het controlepaneel.

---

## Hulp

Dit menu bevat het on-line hulpsysteem en het venster "Aangaande Readiris".

|                          |
|--------------------------|
| <b>Hulp</b>              |
| <u>I</u> nhoud           |
| <u>O</u> mtrent Readiris |

## Inhoud

Geeft toegang tot het on-line hulpsysteem dat uitgebreid elk aspect van de software behandelt. Het on-line hulpsysteem kan eveneens informatie van het laatste uur bevatten.

Versnellingstoets: druk op F1.

De beschikbare informatie kan niet enkel op het scherm worden geraadpleegd maar ook uitgeprint! Raadpleeg je Windows handleiding om meer te vernemen over Windows hulpsystemen.

## Aangaande Readiris

Toont het startscherm met uw registratienummer.

## 5 Meer over OCR

Het doel van OCR is gedrukte en getikte teksten efficiënt en goedkoop in te voeren. Hoewel de eerste research rond Optical Character Recognition (OCR) meer dan 20 jaar geleden van start ging, is deze technologie bij de meeste mensen die ervan gebruiken zouden kunnen en moeten maken, nog steeds onbekend.

Je kan dit nuttige instrument op kantoor gebruiken en je bevrijden van het belastende werk dat het hertikken van bestaande teksten met zich meebrengt.

OCR-systemen werken prima op vele documenten, maar niet op alle... Dit hoofdstuk beschrijft kort de verschillende stappen van het OCR-proces op een document en de invloed die bepaalde parameters kunnen hebben op het resultaat.

---

### Het OCR-proces

Het document wordt door uw scanner gelezen. Deze scanner fungeert als het "oog" van de computer en zendt het beeld door. In deze fase is het beeld van het document niet meer dan een betekenisloze wolk zwarte punten - die we "pixels" noemen - op een witte achtergrond. Wat de OCR-software precies doet is uit deze pixels de tekstinformatie halen: de software moet de vormen herkennen en een symbool toekennen.

Er zijn twee types OCR-systemen: automatische en lerende systemen. Het eerste type herkent automatisch teksten met een breed scala aan lettertypes, maar in geen geval alle. Bij het tweede type moet je het systeem nieuwe karakters aanleren telkens je een nieuw type document aanbiedt. Deze methode

vereist meer interactie van de gebruiker, maar kan bijna alle lettertypes aan.

Beide types systemen gebruiken gelijkaardige processen om tekst te herkennen. De belangrijkste fases zijn:

- lijnsegmentatie
- woord- en karaktersegmentatie
- karakterherkenning
- tekstuitvoer.

De **lijnsegmentatie** verdeelt een bladzijde op in haar verschillende lijnen tekst. Deze fase analyseert ook de kromming en spatiëring van de lijnen. De moeilijkste documenten om te segmenteren hebben weinig of geen spatie tussen twee lijnen en een vrij grote kromming.

De **woord-** en **karaktersegmentatie** isoleren respectievelijk de individuele woorden en karakters. Bij vaste lettertypes is dit vrij eenvoudig. Dit is bijvoorbeeld het geval bij documenten die met een klassieke schrijfmachine geschreven werden.

Het probleem wordt interessanter wanneer de breedte van een symbool afhangt van haar vorm (proportionele letterafstand). De moeilijkste gevallen zijn die documenten waar proportionele karakters elkaar willekeurig raken. Matrixgeprinte karakters zijn samengesteld uit kleine vlekjes en vereisen zeer gespecialiseerde segmentatietechnieken.

De eigenlijke **karakterherkenning** extraheert de kenmerken uit elke geïsoleerde vorm en kent er een symbool aan toe.

De tekstuitvoer is de conversie van de herkende tekst naar een tekstbestand dat kan ingevoerd worden in een tekstverwerker, rekenblad, DTP-pakket enz.

---

## Het bereik van Readiris

Readiris is een hoogstaand en gebruiksvriendelijk OCR-pakket. De functies zijn gebaseerd op alle geavanceerde OCR-technieken die I.R.I.S. ontwikkeld heeft. Dit product heeft de volgende kenmerken:

- Het maakt gebruik van automatische leertechnieken waardoor het een open systeem is. Readiris is een lerend systeem dat er automatisch uitziet: de software vindt autonoom de juiste oplossingen voor het overgrote deel van een document.
- Het maakt uitgebreid gebruik van de taalkundige context om oplossingen te vinden, maar kan ook zonder enige context werken. Dat is bijvoorbeeld het geval bij cijfertabellen.
- Het leidt de gebruiker door de karakterherkenning in vijf fasen: scannen, preview, bevenstering, herkenning en validatie van de twijfelgevallen.
- Zoals elk OCR produkt, kan Readiris niet alle documenttypes aan, maar wat het doet, doet het ook goed. Dit betrouwbaar en rendabel systeem is ontworpen om originele documenten en copieën van goede kwaliteit te lezen.

De automatische leermethode is zeer krachtig en geeft de beste resultaten wanneer aan de volgende voorwaarden voldaan is.

### **Aantal karakters en taalkundige context**

Readiris heeft een redelijk aantal karakters en taalkundige context nodig per weergegeven letterteken om het leerproces quasi-automatisch te doorlopen. Wanneer de taalkundige context beperkt is of gewoon afwezig, zal het systeem natuurlijk ook wel werken maar meer interactie vereisen van de gebruiker.

### **Karakterisolatie**

Readiris maakt gebruik van beperkte segmentatie-algoritmen om de werktijd en geheugeneisen te beperken.

De eerste fase van de karakterisolatie wordt uitgevoerd op karakters die redelijk netjes gescheiden zijn. In een tweede fase daarentegen gebruikt Readiris een zeer krachtig algoritme om groepjes gekartelde karakters en willekeurig aaneengedrukte lettertekens te segmenteren.

Niettemin moet Readiris erin slagen een minimaal aantal elementaire modellen te isoleren om de karakters die elkaar raken correct te kunnen segmenteren.

Het rechtstreekse gevolg van deze techniek is dat elk lettertype herkend kan worden op voorwaarde dat het systeem erin slaagt een voldoende aantal elementaire karakters te isoleren.

### **Karaktergrootte**

Readiris leest lettertekens met een karaktergrootte die varieert van 8 punt tot 20 punt. Alle lettergroottes die in courante tekst worden gebruikt vallen daaronder. Kleinere lettertypes zijn soms overigens ook leesbaar bij een scanresolutie van 400 dpi.

Wij bevelen aan grote koppen niet te vermengen met de te herkennen tekst. In de meeste gevallen tikt de DTP-operator overigens zelf zijn titels in.

### **Onleesbare documenten**

Readiris leest geen teksten waarvan de letters te dicht op elkaar staan. Dit is bijvoorbeeld het geval wanneer bijna alle letters van een woord elkaar raken. Dit verklaart waarom documenten van te lage kwaliteit uitgesloten zijn van herkenning. We denken bijvoorbeeld aan zeer donkere, vette fotocopieën.

# Index

## !

|                   |        |
|-------------------|--------|
| 200% Ware grootte | 29, 47 |
| 50% Ware grootte  | 29, 47 |

## A

|                         |    |
|-------------------------|----|
| Aangaande Readiris      | 52 |
| Aanleren                | 34 |
| Achteruit               | 35 |
| Afstelling              |    |
| Startconfiguratie       | 22 |
| Aktie Menu              | 48 |
| Alles Selecteren        | 45 |
| Alles Vensters          | 47 |
| ASCII                   | 10 |
| Automatische herkenning | 38 |

## B

|                         |        |
|-------------------------|--------|
| Backup                  | 7      |
| Beeld                   |        |
| Bestand openen          | 43     |
| Hele pagina             | 12     |
| Kwaliteit               | 29     |
| Op schijf               | 12     |
| Zoom                    | 12, 47 |
| Bekijken                | 47     |
| 200% Ware Grootte       | 47     |
| 50% Ware grootte        | 29, 47 |
| Alle Vensters           | 47     |
| Geselecteerd Type       | 47     |
| Hele Pagina             | 29, 47 |
| Ware grootte            | 29, 47 |
| Bekijken na het scannen | 50     |
| Beslissingsfase         | 33     |
| Bestand                 |        |
| Dico                    | 51     |
| Formaat                 | 43     |

|               |    |
|---------------|----|
| Layoutbestand | 26 |
| Menu          | 43 |
| Openen        | 43 |
| Uitvoer       | 51 |
| Woordenboek   | 27 |

|                  |    |
|------------------|----|
| Bestand extensie |    |
| .wdw             | 44 |
| Bevestigen       | 35 |
| Bewerken         |    |
| Copiëren         | 32 |
| Plakken          | 32 |
| Wissen           | 32 |
| Bewerken Menu    | 45 |

## C

|                         |        |
|-------------------------|--------|
| Commando                |        |
| 200% Ware Grootte       | 47     |
| 50% Ware Grootte        | 47     |
| Aangaande Readiris      | 52     |
| Alle Vensters           | 47     |
| Alles Selecteren        | 45     |
| Bekijken na het scannen | 49     |
| Bron selecteren         | 43     |
| Controlepaneel          | 49     |
| Coördinaten             | 48     |
| Copiëren                | 45     |
| Funkties                | 48     |
| Geselecteerd Type       | 47     |
| Grafisch Formaat        | 21, 46 |
| Hele Pagina             | 47     |
| Info                    | 43     |
| Inhoud                  | 52     |
| Instellingen laden      | 23, 44 |
| Instellingen opslaan    | 22, 44 |
| Interactief Lereren     | 49, 51 |
| Knippen                 | 45     |
| Knoppenbalk             | 46     |
| Layout laden            | 44     |
| Layout opslaan          | 44     |



|                           |             |
|---------------------------|-------------|
| OCR                       | 49          |
| Ongedaan Maken            | 45          |
| Openen                    | 43          |
| Plakken                   | 45          |
| Roteren                   | 48          |
| Scannen                   | 27 - 28, 49 |
| Scanner                   | 46          |
| Sorteren                  | 47          |
| Sorteren Bekijken         | 48          |
| Sorteren Reset            | 48          |
| Startconfiguratie opslaan | 22, 44      |
| Strengte Controle         | 51          |
| Taal                      | 21, 46      |
| Tekstformaat              | 46          |
| Tesktformaat              | 20          |
| Verlaten                  | 44          |
| Verwerven                 | 44          |
| Ware Grootte              | 47          |
| Wissen                    | 45          |
| Commando's                | 43 - 52     |
| Configuratie              |             |
| Startconfiguratie         | 10, 17      |
| Connect                   |             |
| Gebruik                   | 39          |
| Context                   |             |
| Taalkundige               | 55          |
| Contextanalyse            | 35          |
| Controlepaneel            | 23, 49 - 50 |
| Akte                      | 25          |
| Bekijken na het scannen   | 50          |
| Dokument                  | 24          |
| Invoer                    | 51          |
| Invoerbron                | 26          |
| Layout                    | 26          |
| Letterafstand             | 50          |
| Lettergrootte             | 50          |
| Lettertype                | 50          |
| Output Bestand            | 26          |
| Pagina-analyse            | 49 - 50     |
| Scannen                   | 52          |
| Sluiten                   | 52          |
| Te lezen pagina('s)       | 50          |
| Uitvoer                   | 51          |
| Woordenboek               | 51          |
| Woordenboekestand         | 26          |
| Coördinaten               | 48          |
| Copiëren                  | 45          |

## D

|                    |    |
|--------------------|----|
| Desktop Publishing | 54 |
| Dico               | 51 |

|                         |        |
|-------------------------|--------|
| Dictionary              |        |
| Bijvoegen               | 36     |
| Lezen                   | 36     |
| Document                |        |
| Onleesbaar              | 56     |
| Dokument                |        |
| Bekijken na het scannen | 26     |
| Bepaald type            | 36     |
| Goede Kwaliteit         | 29     |
| Kenmerken               | 24     |
| Kwaliteit               | 29, 55 |
| Scannen                 | 15, 49 |
| Tussenlijnspatiëring    | 24     |
| Types                   | 55     |
| Doorgaan                | 35     |
| DOS                     |        |
| Versie                  | 7      |
| dpi                     | 15     |

## E

|       |    |
|-------|----|
| Einde | 35 |
|-------|----|

## F

|          |    |
|----------|----|
| Formaat  |    |
| Grafisch | 21 |
| Funkties | 48 |
| Pijltje  | 13 |

## G

|                         |        |
|-------------------------|--------|
| Gebruikerswoordenboeken | 36     |
| Geheugen                |        |
| Vereisten               | 7      |
| Geselecteerd Type       | 47     |
| Grafiek                 |        |
| Formaat                 | 21     |
| Grafisch Formaat        | 46     |
| Grootte                 | 24, 50 |

## H

|                |        |
|----------------|--------|
| Handscanner    | 15, 28 |
| Hele Pagina    | 29, 47 |
| Hulp 52        |        |
| Hulpinformatie | 37     |

## I

|                           |        |
|---------------------------|--------|
| Inhoud                    | 52     |
| Inspiringing              | 21, 46 |
| Installatie               | 7      |
| Instelling                |        |
| Instellingen laden        | 44     |
| Instellingen opslaan      | 44     |
| Laden                     | 22     |
| Oproepen                  | 43     |
| Opslaan                   | 22     |
| Startconfiguratie opslaan | 44     |
| Instellingen              |        |
| Opslaan                   | 43     |
| Interactief Leren         | 49, 51 |
| Invoerbestand             | 51     |

## K

|                         |            |
|-------------------------|------------|
| Karakter                |            |
| Dico                    | 27         |
| Dot matrix              | 24         |
| Gekartelde              | 55         |
| Grootte                 | 50         |
| Herkenning              | 13         |
| Isolatie                | 55         |
| Klein                   | 24         |
| Letterafstand           | 24, 50, 54 |
| Lettergrootte           | 24, 50, 56 |
| Lettertype              | 50         |
| Rakende                 | 55         |
| Schrijfmachine          | 24         |
| Segmentatie             | 54 - 55    |
| Willekeurig aangedrukte | 55         |
| Woordenboek             | 27         |
| Karakterherkenning      | 32         |
| Karakters               |            |
| Twijfel                 | 14         |
| Kleine karakters        | 24         |
| Knippen                 | 45         |
| Knoppenbalk             | 46         |
| Kolommen                | 48         |

## L

|            |        |
|------------|--------|
| Layout     | 26, 51 |
| Laden      | 44     |
| Opslaan    | 44     |
| Leermodule | 14     |
| Leren      | 14     |
| Achteruit  | 35     |

|               |    |
|---------------|----|
| Bevestigen    | 35 |
| Einde         | 35 |
| Wissen        | 35 |
| Letterafstand | 50 |
| Proportioneel | 24 |
| Proportionele | 54 |
| Vast          | 24 |
| Lettergrootte | 50 |
| Lettertype    | 50 |
| Lettertypes   |    |
| Aanleren      | 34 |
| Lijn          |    |
| Segmentatie   | 54 |
| Spatiëring    | 54 |

## M

|             |    |
|-------------|----|
| Markeerders | 31 |
| Menu        |    |
| Aktie       | 48 |
| Bekijken    | 47 |
| Bestand     | 43 |
| Bewerken    | 45 |
| Hulp        | 52 |
| Opties      | 46 |

## N

|                 |    |
|-----------------|----|
| Numerieke modus | 38 |
|-----------------|----|

## O

|                     |         |
|---------------------|---------|
| OCR                 | 15, 49  |
| Meer over OCR       | 53 - 56 |
| Meerdere Bladzijden | 27      |
| Proces              | 13, 53  |
| OK                  | 34      |
| Ongedaan Maken      | 45      |
| Openen              | 11      |
| Opties              | 46      |

## P

|                |         |
|----------------|---------|
| Pagina         |         |
| Aantal         | 50      |
| Controlepaneel | 23      |
| Scannen        | 27 - 28 |
| Pagina-analyse | 49 - 50 |
| PaintBrush     | 10      |

|                   |         |                       |        |
|-------------------|---------|-----------------------|--------|
| Paragraaf         |         | Bijvoeg-/Vervangmodus | 26     |
| Detectie          | 21      | Inspringing           | 21     |
| Paragraafdetectie | 46      | Tekstformaten         | 21, 37 |
| Pijl              | 31      | Paragraafdetectie     | 21     |
| Pixel             | 53      | Tekstvensters         | 12     |
| Plakken           | 45      | Tekstverwerking       | 46, 54 |
| Preview           | 49 - 50 | Toevoegingsmodus      | 51     |
| Proces            |         | Tussenlijnspatiëring  | 24     |
| Stoppen           | 32      |                       |        |

## R

|                    |             |
|--------------------|-------------|
| Readiris           |             |
| Configuratie       | 43          |
| Installatie        | 7           |
| Instelling         | 10, 46      |
| Pagina scannen     | 15          |
| Scanner Selecteren | 15, 18 - 19 |
| Taal               | 10          |
| Van start gaan     | 7 - 8       |
| Venstertools       | 30          |
| Werkingsmodus      | 25          |
| Rechthoek          | 30          |
| Resolutie          | 15, 18, 56  |
| Roteren            | 48          |

## S

|                    |            |
|--------------------|------------|
| Scannen            | 15, 49, 52 |
| Scanner            | 46         |
| Afstelling         | 29         |
| Instelling         | 46         |
| Resolutie          | 15, 18     |
| Segmentatie        | 54         |
| Sluiten            | 52         |
| Sorteren           | 47         |
| Sorteren Bekijken  | 48         |
| Sorteren Reset     | 48         |
| Startconfiguratie  | 17         |
| Stoppen            | 32         |
| Strengere Controle | 51         |
| Systeemvereisten   | 7          |

## T

|                     |        |
|---------------------|--------|
| Taal                | 21, 46 |
| Taalkundige context | 55     |
| Tabellen            | 37     |
| Te lezen pagina('s) | 50     |
| Tekstformaat        | 46     |

## U

|                |    |
|----------------|----|
| Uitvoerbestand | 51 |
|----------------|----|

## V

|                  |        |
|------------------|--------|
| Venster          |        |
| Coördinaten      | 48     |
| Sorteren         | 47     |
| sSorteren Reset  | 48     |
| Vensters         |        |
| Alles Selecteren | 45     |
| Bewerken         | 45     |
| Coördinaten      | 31     |
| Copiëren         | 45     |
| Creëren          | 30     |
| Funkties         | 48     |
| Knippen          | 45     |
| Layout laden     | 44     |
| Layout opslaan   | 44     |
| Markeerders      | 31     |
| Plakken          | 45     |
| Selecteren       | 31     |
| Sorteren         | 48     |
| Tekenen          | 12, 30 |
| Tools            | 30     |
| Type             | 31     |
| Verplaatsen      | 32     |
| Wijzigen         | 13, 31 |
| Wissen           | 32, 45 |
| Venstertoolbox   |        |
| Rechthoek        | 12     |
| Verkenning       | 9      |
| Verlaten         | 44     |
| Vervangingsmodus | 51     |
| Vlakbedscanner   | 15, 27 |

## W

|              |        |
|--------------|--------|
| Ware grootte | 29, 47 |
| Windows      |        |

|                    |        |                  |        |
|--------------------|--------|------------------|--------|
| Versie             | 7      | Modi             | 51     |
| Wissen             | 35, 45 | Woordenboekmodus | 36     |
| Woordenboek        |        |                  |        |
| Modus              | 27     |                  |        |
| Modus Bijvoegen    | 27     |                  |        |
| Modus Lezen        | 27     |                  |        |
| Modus Nieuw        | 27     | <b>Z</b>         |        |
| Selecteren         | 36     |                  |        |
| Woordenboekbestand | 51     | Zoom             | 12, 47 |
| Woordenboeken      |        | Zoomniveau's     | 29     |