

# Readiris™ PRO 10

## GEBRUIKERS- HANDLEIDING



Readiris Corporate

© 2005 I.R.I.S. Alle rechten voorbehouden

OCR technologie door I.R.I.S.

Connectionist, AutoFormat en linguïstische technologie door I.R.I.S.

ICR en streepjescodelectuurtechnologie door I.R.I.S.

BCR en veldanalysetechnologie door I.R.I.S.

© 2005 I.R.I.S. Alle rechten voorbehouden

## **BESPAAR TIJD, VERTIK HET TE HERTIKKEN!**

---

Gelukwensen bij de aanschaf van Readiris. Deze software zal u ongetwijfeld belangrijke hulp bieden bij het hernemen van uw teksten, tabellen, grafisch materiaal, visitekaartjes, streepjescodes en zelfs handgeschreven teksten!

Hoe efficiënt computers ook mogen zijn, u moet de informatie altijd wel nog eerst invoeren. Als u ooit een rapport van 15 pagina's of een grote tabel heeft moeten overtypen, dan weet u ongetwijfeld hoe vervelend en tijdrovend dat kan zijn. Gebruik echter dit geavanceerde OCR-pakket om teksten automatisch in uw toepassingen in te voeren en u zult een ongekend niveau van efficiëntie en comfort halen!

Scan een gedrukt of getyped document, geef de te hernemen zones aan - of laat het systeem deze klus voor u klaren - en voer de karakterherkenning uit. Documenten die meerdere bladzijden bevatten worden in één beweging verwerkt. Enkele eenvoudige muisklikken overtreffen urenlang werk naarmate Readiris uw papieren documenten omzet in bewerkbare computerbestanden; dit gebeurt wel 40 keer sneller dan handmatige invoer!

De OCR wizard gidst u doorheen de herkenning: antwoord op enkele eenvoudige vragen en u haalt snelle resultaten met Readiris. U kunt de herkenningsresultaten tevens rechtstreeks naar uw tekstverwerker en uw spreadsheet sturen. Om faxen te herkennen en PDF documenten om te zetten kunt u de beeldbestanden van de Windows Verkenner naar het Readiris venster slepen. Of klik met de rechter muisknop op een beeld om het ogenblikkelijk met Readiris te openen.

Readiris herkent tabellen en hercreëert die als werkbladen of als tabelobjecten binnen uw tekstverwerker; uw cijfermateriaal is onmiddellijk klaar voor verdere verwerking met uw rekenblad.

Gebaseerd op de Connectionist technologie van I.R.I.S., heeft Readiris het beste wat OCR te bieden heeft. Omnifont herkenning wordt aangevuld met zelflerende technieken die afgeleid zijn van neuraal netwerk. Het systeem kan



door de contextuele analyse nieuwe karakters aanleren: taalkundige kennis over lettergrepen en woorden verbetert de OCR-prestatie.

Readiris ondersteunt wel 117 talen: alle Amerikaanse en Europese talen worden herkend, inclusief de Centraal-Europese talen, de Baltische talen, het Grieks en de Cyrillische (“Russische”) talen. (Optioneel kunt Hebreeuwse documenten en vier Aziatische talen inlezen - Japans, vereenvoudigd en traditioneel Chinees en Koreaans.) Readiris verwerkt ook gemengde alfabetten: de software detecteert “Westerse” woorden die opduiken in Griekse, Cyrillische, Hebreeuwse en Aziatische documenten - vele niet vertaalbare eigennamen, merknamen enz. zijn immers geschreven in het Latijnse alfabet.

Readiris maakt gebruik van taalkundige informatie *tijdens* de herkenningsfase, en niet nadien. Dit laat Readiris toe documenten van allerlei soort te herkennen met uiterste nauwkeurigheid, en dat geldt ook voor documenten met een lage drukkwaliteit, faxen en dot matrix afdrucken. Readiris herkent eveneens slecht gescande en slecht gekopieerde documenten met te lichte of te donkere lettertekens. Verbonden lettertekens (“ligaturen”) worden goed gelezen, gefragmenteerde vormen zoals dot matrix symbolen worden opnieuw samengesteld.

Controle door de gebruiker in “popup” stijl markeert de twijfelachtige karakters en verhoogt de precisie van het systeem. Alle oplossingen die door de gebruiker zijn bevestigd worden opgeslagen, wat de snelheid en betrouwbaarheid bij verder gebruik nogmaals verhoogt. Readiris gebruiken betekent de software telkens intelligenter maken! Deze krachtige leerfunctie laat u toe Readiris te trainen op allerhande speciale lettertekens zoals de wiskundige symbolen, “dingbats” en de vervormde, beschadigde lettertypes die u in echte documenten kunt terugvinden.

Om uw productiviteit nog meer te verhogen, herkent Readiris niet alleen uw teksten, maar hij kan ze ook in uw plaats *opmaken*! Maak gebruik van “autoformattering” en Readiris hercreëert een facsimile kopie van het gescande document: de woord-, alinea- en paginaformattering van het oorspronkelijke document blijven behouden.

Er worden gelijkaardige lettertypes, puntgroottes en stijlen gebruikt zoals in het brondocument. De positie van de kolommen, tekstblokken en grafieken is identiek aan het oorspronkelijke document. Readiris kan zelfs een foto die in de

achtergrond van een gescande bladzijde voorkomt in het herkend document opnemen! En aangezien Readiris moeiteloos grijswaarden en kleurscans ondersteunt kunt u tegelijk al uw grafisch materiaal hernemen - het maakt niets uit of het zwart-wit illustraties, zwart-wit foto's of kleurfoto's betreft. Wanneer een document tabellen bevat, zal Readiris die "herorganiseren" in heuse cellen en de celgrenzen herscheppen naar het voorbeeld van de oorspronkelijke tabel.

Met andere woorden, met Readiris kunt u een conforme kopie van uw documenten archiveren, zij het als editeerbare, compacte tekstbestanden en niet als gescande beelden! Verscheidene niveau's van formattering zijn beschikbaar, de gebruiker maakt zijn keuze.

Streepjescodes die voorkomen op een gescande bladzijde kunt u ook lezen, en hetzelfde geldt voor handgeschreven tekst - zulke teksten kunt u herkennen zolang u netjes gescheiden "blokletters" op papier zet.

U kunt zelfs visitekaartjes gaan inlezen met Readiris: scan uw visitekaartjes, herken ze en zet ze zo om in een adresdatabank. Denk aan de laatste keer dat u terug kwam van een vakbeurs met een hele stapel visitekaartjes, en uw secretaresse twee dagen bezig was om de gegevens in te voeren!

De gegevens die op een kaartje voorkomen worden automatisch uit het beeld gehaald en de herkende data worden verdeeld over specifieke databankvelden. Readiris maakt hierbij extensief gebruik van een "kennisdatabank" die de software de vereiste intelligentie geeft om voor- en achternamen, steden en staten, telefoon- en faxnummers enz. te onderscheiden. De resulterende gegevens kunnen rechtstreeks worden doorgestuurd naar uw contact manager zoals Microsoft Outlook (Express) en naar elke toepassing die de vCard standaard ondersteunt.

Readiris ondersteunt een breed gamma aan populaire scanners: talrijke vlakbedscanners, sheetfed scanners, "all-in-one" machines of "MFPs" ("multi-functional peripherals") en digitale camera's worden ondersteund. Readiris ondersteunt tevens de Twain scanner standaard en enkele scanplatformen. Dankzij het intervalscannen kunt u meerbladige documenten efficiënt verwerven wanneer uw scanner niet is uitgerust met een automatische doorvoer. (Readiris Corporate ondersteunt snelle scanners en voert reeksherkenning uit op grote verzamelingen beelden: met blanco bladzijden worden beeldreeksen opgedeeld in aparte



documenten, de automatische lectuur van streepjescodes zorgt voor een correcte indexering van de “gedematerialiseerde” documenten.)

## **INHOUDSTAFEL**

---

Bespaar tijd, vertik het te hertikken! .....	III
Inhoudstafel .....	VI
Auteursrechten .....	VIII

### **Hoofdstuk 1: Installatie**

Systemvereisten .....	1-1
De Readiris software installeren .....	1-1
De Readiris software verwijderen .....	1-3
Het Readiris verwijderprogramma .....	1-3
De Windows (un)installwizard .....	1-4
Software-opties installeren .....	1-5
Geïnstalleerde bestanden .....	1-8
“Lees Mij” bestand en documentatie .....	1-8
Handschriftformulier .....	1-8
Scannerstuurbestanden .....	1-9
Registreer om mee te spreken! .....	1-9
Beroep doen op productondersteuning .....	1-11
I.R.I.S. contacteren .....	1-12

### **Hoofdstuk 2: Een rondleiding**

Readiris opstarten .....	2-1
De eerste sessie is uniek .....	2-2
De Readiris gebruikersinterface verkennen .....	2-3
Van start gaan met een eerste oefening .....	2-5
Inzoomen op beelden .....	2-10
Eén, een gescand beeld analyseren .....	2-13
Anderhalf, vensters sorteren .....	2-16
Twee, beelden manueel bevensteren .....	2-19
Drie, venstersjablonen opslaan .....	2-23
Met Readiris kunt u de hele wereld rond! .....	2-25
Nog meer taalkunde met gebruikerslexicons .....	2-30

Readiris verandert tussendoor van taal .....	2-33
Documenten met gemengde talen lezen .....	2-35
De kenmerken van het document aangeven .....	2-37
Readiris wordt telkens intelligenter! .....	2-39
Leren .....	2-41
Niet Leren .....	2-41
Wissen .....	2-42
Annuleren .....	2-42
Einde .....	2-43
Afbreken .....	2-43
De rol van woordenboeken .....	2-43
Tekst direct naar een toepassing sturen .....	2-46
Het resultaat in een tekstbestand opslaan .....	2-50
Universele bestanden aanmaken... ..	2-53
... of ze lezen! .....	2-61
Meerdere pagina's herkennen .....	2-64
Meerbladige documenten bewerken .....	2-74
Met een nieuw document beginnen .....	2-77
Tekstzones herkennen .....	2-77
De uitvoer opmaken .....	2-79
Uw scanner instellen .....	2-80
Geef uw scans wat kleur mee! .....	2-83
Andere apparaten, andere resolutie .....	2-86
Standaard instelling bewaren .....	2-90
Specifieke instellingen bewaren .....	2-91
Documenten scannen .....	2-92
De gescande beelden aanpassen .....	2-95
Laat de wizard voor u werken .....	2-100
Readiris herschept de originele lay-out .....	2-101
Kolommen aub, geen kaders! .....	2-106
Tekstformattering, deel 2 .....	2-110
Tekst meerdere keren uitvoeren .....	2-111
Figuren apart opslaan .....	2-111
Gekleurde achtergronden opslaan .....	2-114
Figuren maximaal benutten .....	2-117
Faxen herkennen en uitgestelde herkenning .....	2-119
Documentreeksen herkennen .....	2-120



Een bewaakte map instellen .....	2-122
Documentreeksen organiseren .....	2-124
Tabellen herkennen .....	2-128
Handgeschreven tekst herkennen .....	2-133
Balken en spaties inlezen .....	2-137
Visitekaartjes inlezen .....	2-139
Visitekaartjes scannen .....	2-140
En dan is de leesmodus voor visitekaartjes.....	2-145
Visitekaartjes herkennen .....	2-148
On-line hulp raadplegen .....	2-150

## AUTEURSRECHTEN

---

De Readiris software werd ontworpen en ontwikkeld door I.R.I.S. De OCR, ICR, BCR, streepjescodelectuur-, Connectionist, AutoFormat en linguïstische technologie werd ontwikkeld door I.R.I.S. I.R.I.S. bezit de auteursrechten op de Readiris software, de OCR technologie, de ICR technologie, de streepjescodelectuurtechnologie, de BCR technologie, de linguïstische technologie, het on-line hulpsysteem en deze handleiding.

AutoFormat, Connectionist, de IBCR-II, de I.R.I.S. linguïstische technologie, het I.R.I.S. logo en Readiris zijn handelsmerken van I.R.I.S.

De XML parser werd ontwikkeld door Apache. Dit product bevat software die werd ontwikkeld door de Apache Software Foundation ([www.apache.org](http://www.apache.org)).

Acrobat en Reader zijn (geregistreerde) handelsmerken van Adobe. Excel, Windows en Word zijn geregistreerde handelsmerken van Microsoft. Intel is een geregistreerd handelsmerk van Intel.



# Hoofdstuk 1

## INSTALLATIE

Dit hoofdstuk bespreekt de systeemvereisten en de installatie van de Readiris software.

### **SYSTEEMVEREISTEN**

---

Dit zijn de minimale systeemvereisten om Readiris te gebruiken:

- een 486 Intel PC of compatibele computer. Een Pentium PC is aanbevolen.
- 64 MB RAM. 128 MB RAM is aanbevolen om beelden in grijswaarden en kleuren te verwerken.
- 120 MB vrije ruimte op de harde schijf. 105 MB schijfruimte volstaat wanneer u de voorbeeldbestanden op de CD-ROM laat staan.
- het besturingssysteem Windows XP, Windows ME, Windows 2000, Windows 98 of Windows NT 4.0.

Readiris Corporate vereist een beeldscherm met een 1.024 x 768 resolutie.

Noteer dat sommige **scannerstuurbestanden** mogelijk niet werken onder de laatste versie(s) van Windows! Raadpleeg de documentatie die bij uw scanner geleverd werd om na te gaan welke besturingssystemen ondersteund worden.

### **DE READIRIS SOFTWARE INSTALLEREN**

---

De Readiris software wordt uitsluitend geleverd op een **zelfstartende CD-ROM**. Om de software te installeren laadt u eenvoudig de CD-ROM in uw CD-

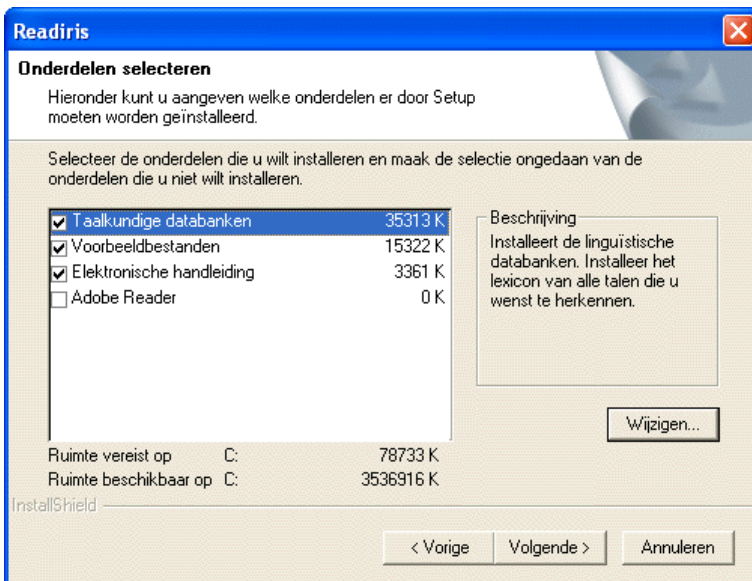


ROM station en wacht u tot het installatieprogramma begint te lopen. Voer de instructies uit die op het scherm verschijnen.

Mocht de installatie niet opstarten wanneer u de CD-ROM in de CD-ROM lezer laadt, start dan het installatieprogramma MENU.EXE op om de software te installeren.

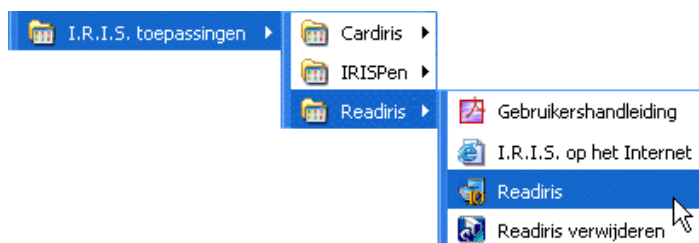
Gebruikers van Windows XP, Windows 2000 en Windows NT moeten nagaan of zij over de nodige **toegangsrechten** beschikken - contacteer zo nodig de systeembeheerder.

Er komen enkele opties kijken bij de installatie. Installeer in elk geval de **taalkundige databanken** van alle talen die u wenst in te lezen. Standaard worden alle woordenboeken geïnstalleerd. Wij raden u tevens aan de **voorbeeldbestanden** te installeren; die worden namelijk in deze handleiding gebruikt als lesbestanden.



Installeer tevens, mocht dat nodig blijken, de Adobe Reader software die toelaat de softwaredocumentatie te raadplegen. De **elektronische handleiding** wordt standaard op uw harde schijf gekopieerd, maar u kunt die ook op de CD-ROM laten staan.

Het installatieprogramma creëert automatisch het submenu "I.R.I.S. toepassingen - Readiris" onder het menu "Programma's".



Datzelfde geldt voor een **snelkoppeling** naar Readiris op het Windows bureaublad. Eén en ander betekent dat u de Readiris software onmiddellijk vanaf uw **bureaublad** kunt opstarten.

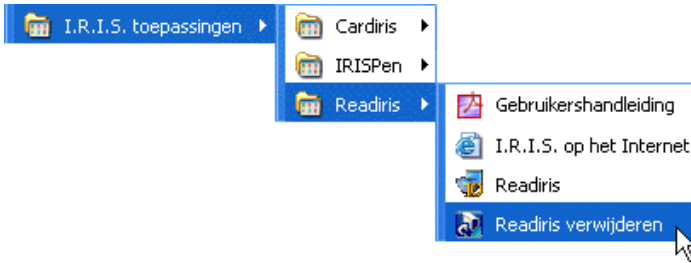
## **DE READIRIS SOFTWARE VERWIJDEREN**

---

Er zijn slechts twee juiste methodes om de Readiris software te verwijderen: gebruik maken van de "uninstall" toepassing en gebruik maken van de Windows (un)installassistent. Het wordt ten stelligste afgeraden om de Readiris software of zijn softwaremodules te verwijderen door de bestanden van de toepassing manueel te verwijderen.

### **Het Readiris verwijderprogramma**

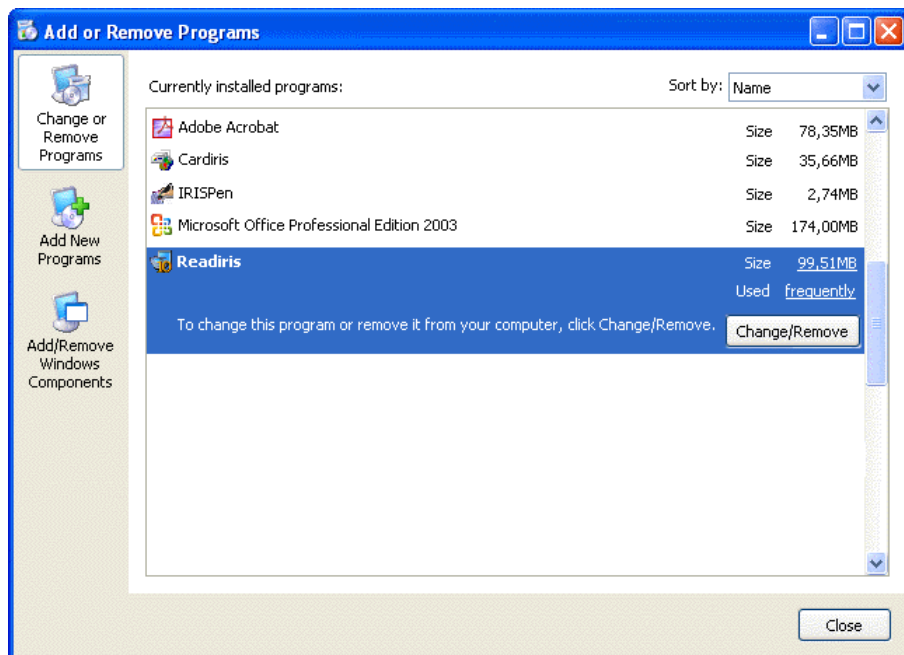
Selecteer "Readiris Verwijderen" in het submenu "I.R.I.S. toepassingen - Readiris" om de Readiris "uninstall"-toepassing op te starten. Volg de instructies die op het scherm verschijnen.



## De Windows (un)installwizard

Voer de volgende stappen uit om gebruik te maken van de Windows (un)installassistent.

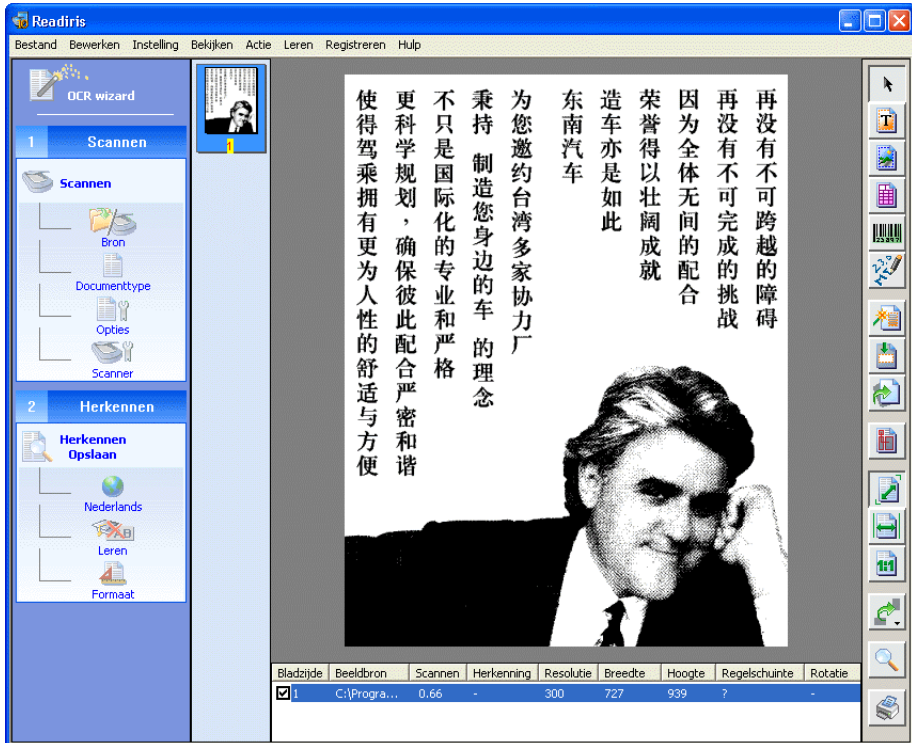
- Klik op "Instellingen" onder het "Start" menu van Windows en ga naar het "Configuratiescherm".
- Klik op het pictogram "Toepassingen Toevoegen/Wissen" in het "Configuratiescherm".



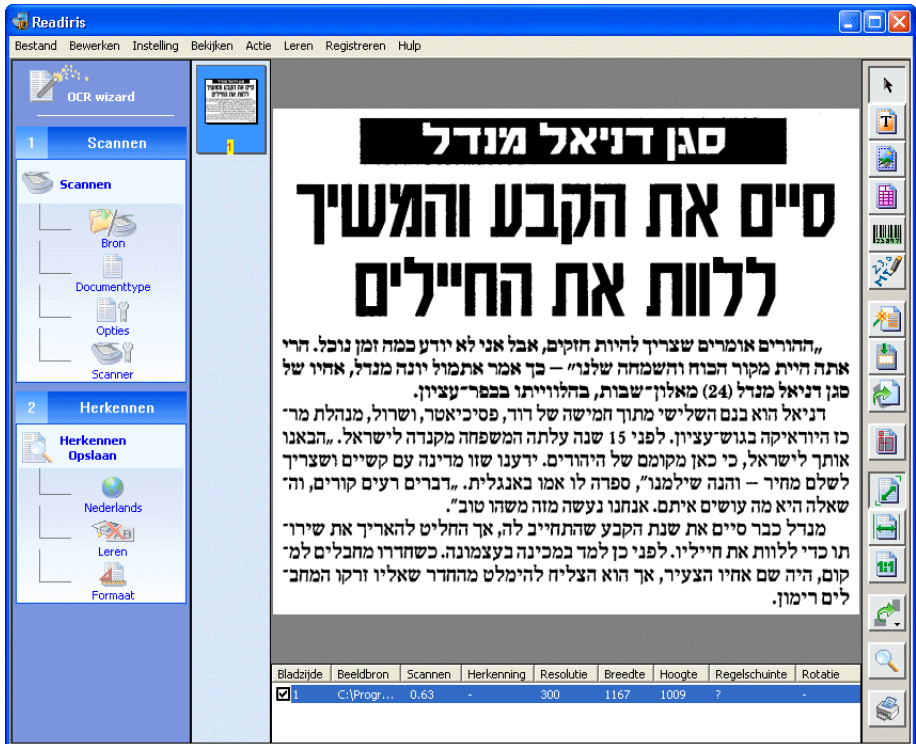
- Volg de instructies die op het scherm verschijnen om de Readiris software te verwijderen.

## SOFTWARE-OPTIES INSTALLEREN

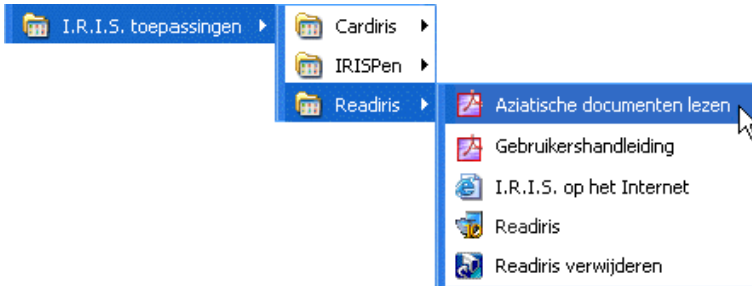
Er zijn twee software-opties beschikbaar voor Readiris: de complementen “Aziatische OCR-module” en “Hebreeuwse OCR-module”. De “Aziatische OCR-module” laat u toe het Japans, het traditioneel en vereenvoudigd Chinees en het Koreaans te herkennen.



De “Hebreeuwse OCR-module” houdt geen verrassingen in: ze laat u Hebreeuwse documenten herkennen.



Ook die extra software wordt op een zelfstartende CD-ROM geleverd. Wanneer u de optie “Aziatische OCR-module” installeert, wordt er specifieke documentatie toegevoegd die uitlegt hoe u Aziatische documenten kunt herkennen.



## **GEÏNSTALLEERDE BESTANDEN**

---

Het installatieprogramma maakt een map aan waarin de Readiris bestanden zich bevinden. Probeer nooit Readiris of sommige van zijn modules te verwijderen door handmatig programmabestanden te wissen. Gebruik hiervoor het Readiris “uninstall”-programma of de Windows (un)installwizard. Zie hierboven.

### **“Lees Mij” bestand en documentatie**

- |            |  |
|------------|--|
| README.HTM | “Lees Mij” bestand (in HTML formaat)             |
| MANUAL.PDF | Gebruikershandleiding (in Adobe Acrobat formaat) |

### **Handschriftformulier**

- |              |  |
|--------------|--|
| TEMPLATE.PDF | Blanco formulier voor handschriftherkenning<br>(af te drukken exemplaar) |
| TEMPLATE.DOC | Blanco formulier voor handschriftherkenning<br>(editeerbaar exemplaar)   |



## **Scannerstuurbestanden**

Aarzel niet uw scannerfabrikant of zijn vertegenwoordiger te contacteren mocht u problemen ondervinden met uw stuurbestand. Bij de meeste fabrikanten kunt u de laatste versies van de scannerstuurbestanden van de webstek halen.

## **REGISTREER OM MEE TE SPREKEN!**

Vergeet niet uw Readiris licentie te registreren! Dit stelt ons in staat om u op de hoogte te houden van verdere productontwikkelingen en aanverwante I.R.I.S. producten. De voordelen van de registratie, zoals gratis **productondersteuning** en **speciale offertes**, zijn exclusief voorbehouden aan geregisteerde gebruikers.

U kunt op vele verschillende manieren registreren: door uw registratiekaart op te sturen of door het elektronisch evenbeeld te faxen, door I.R.I.S. telefonisch te contacteren tijdens de kantooruren of door een registratieformulier in te vullen op de I.R.I.S. website!



**Readiris Hulp**

Hide Back Forward Home Print Options

Contents Index Search

- Welkom bij de Readiris hulp
- OCR inleiden
- Documenten herkennen
- Handgeschreven tekst herkennen
- Streepjescodes herkennen
- Hoe...?
- Referentie-informatie
- Software-versies en opties
- Productregistratie
  - Uw Readiris licentie registreren**
  - Readiris registratieformulier
  - Registreren per e-mail
- Productondersteuning
- I.R.I.S.

## Uw Readiris licentie registreren

### Waarom registreren?

- Registreren laat ons toe u op de hoogte te houden van toekomstige **productontwikkelingen** en **aanverwante I.R.I.S. producten**.
- De registratie geeft u toegang tot gratis **productondersteuning** en **speciale offertes**.
- Afhankelijk van de software-versie ontvangt u de **softwaresleutel** die u kan nodig hebben om Readiris te blijven gebruiken na één maand.

### Hoe registreren?

#### Post

Stuur ons uw **registratiekaart**.

#### Fax

Indien uw registratiekaart verloren raakte of indien de software-documentatie enkel in **elektronische vorm** geleverd werd, kan u het Readiris [registratieformulier](#) afdrukken en naar I.R.I.S. sturen of faxen.

Europa: 32-10-45 34 43

De Readiris **registratiewizard** die u vindt onder het menu "Registreren" van de Readiris software loodst u in luttele ogenblikken door het registratieproces.



Naargelang de softwareversie, ontvangt u in ruil de **softwaresleutel** die u kunt nodig hebben om de Readiris software na één maand te blijven gebruiken.

## **BEROEP DOEN OP PRODUCTONDERSTEUNING**

Het commando "Productondersteuning" in het "Hulp" menu van Readiris beschrijft hoe u beroep kunt doen op de technische ondersteuning. Beschrijf duidelijk het probleem dat u ervaart en verschaf ons alle nuttige informatie betreffende Readiris, uw scanner en uw computersysteem.



**Productondersteuning verkrijgen**

De **geregistreerde klanten** krijgen toegang tot **gratis** technische **ondersteuning**. ([Registratie](#) geeft u ook recht op onze speciale voorwaarden.)

**Europa**

Hotline: 32-10-45 13 64 (kantooruren) (alle hoofdtalen)  
Fax: 32-10-45 34 43

**VSA**

Hotline: 1-561-921-0847 / 800-477-4744 (kantooruren)  
Fax: 1-561-921-0854

**WWW**

[www.irislink.com/support.html](http://www.irislink.com/support.html) ("troubleshooting" informatie)  
Klik [hier](#) voor directe toegang tot de troubleshooting informatie.

**E-mail**

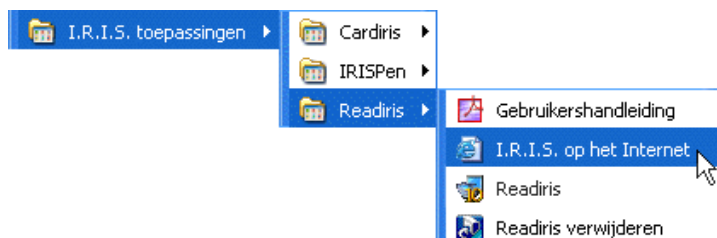
[support@irislink.com](mailto:support@irislink.com), [support@irisusa.com](mailto:support@irisusa.com)

## **I.R.I.S. CONTACTEREN**

U kunt tevens contact opnemen met I.R.I.S. om meer te vernemen over de andere oplossingen uit het productgamma; het commando "I.R.I.S. Contacteren" van het menu "Hulp" van Readiris vertelt u hoe u contact kunt opnemen met I.R.I.S.



Een pictogram in het submenu "I.R.I.S. toepassingen - Readiris" onder het menu "Programma's" brengt u rechtstreeks naar de I.R.I.S. website. Dat kan ook vanuit het startscherm van Readiris en aan de hand van het commando "I.R.I.S. op het Internet" onder het menu "Hulp" van Readiris.



# Hoofdstuk 2

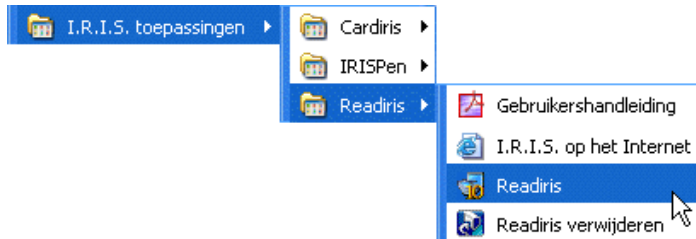
## EEN RONDLEIDING

Readiris is een geavanceerd OCR-pakket dat talrijke krachtige functies bevat. We zullen in dit hoofdstuk deze mogelijkheden doorlopen en vele tips betreffende het gebruik van Readiris toevoegen.

### READIRIS OPSTARTEN

---

Klik op de Readiris applicatie in het submenu "I.R.I.S. Toepassingen - Readiris" of klik op de snelkoppeling van de Readiris applicatie op uw bureaublad.



Het Readiris startscherm en applicatievenster verschijnen op het scherm. Het startscherm geeft de versie en de auteursrechten van de Readiris software aan. Het verleent tevens rechtstreekse toegang tot de I.R.I.S. **website**: klik op de URL om de webstek van I.R.I.S. te bezoeken. Door met de muis elders te klikken verdwijnt het startscherm.

Het volgende venster betreft de OCR assistent; klik deze keer op "Annuleren" om die te doen verdwijnen.



## DE EERSTE SESSIE IS UNIEK

---

Naargelang de versie van de software die u heeft aangekocht, kan de eerste sessie wat apart verlopen. U kunt namelijk uitgenodigd worden om uw licentie te registreren.

Als dat het geval is, is het gebruik van uw Readiris softwarelicentie beperkt tot één maand. Door te registreren ontvangt u van I.R.I.S. gratis de **softwaresleutel** die u toelaat om het programma na 30 dagen te blijven gebruiken.

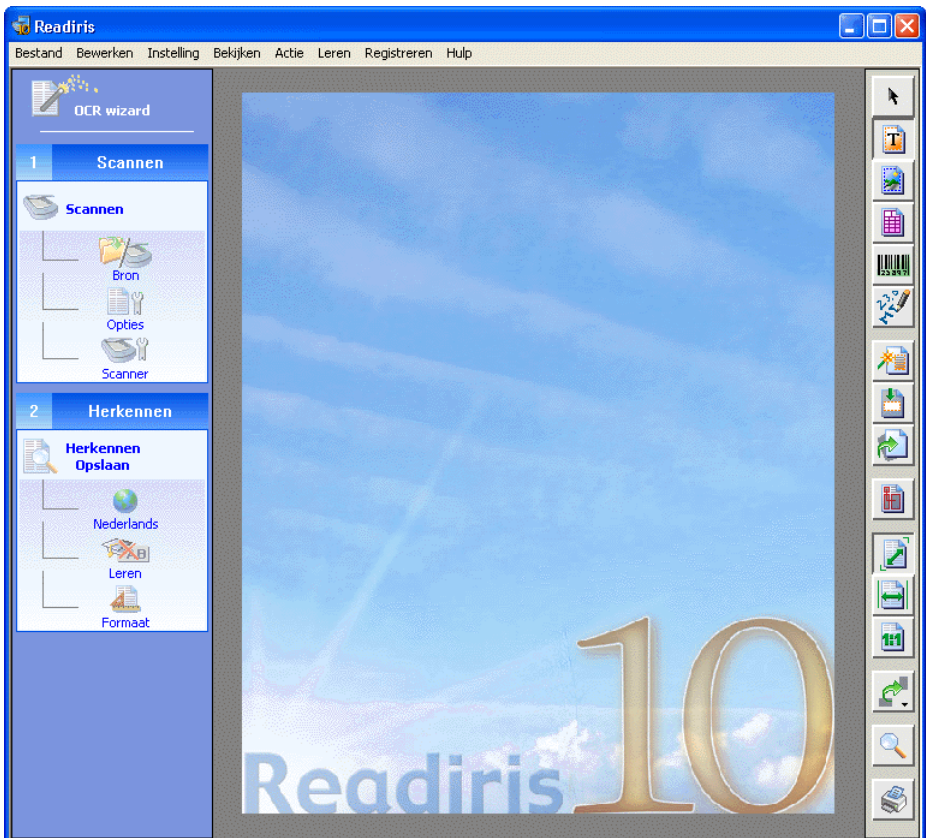
U heeft uw **identificatienummer** nodig om de softwaresleutel te kunnen genereren; zorg ervoor dat u dit nummer onder uw bereik hebt of vermeldt wanneer u uw licentie registreert.

The screenshot shows a dialog box titled "Readiris" with a red close button in the top right corner. The text inside the dialog box reads: "Het identificatienummer op deze PC is:" followed by a text box containing the number "31C0142535C08803888032535444808080" and a "Hulp" button. Below this, it says "U hebt een sleutel nodig om deze software te activeren." and "Contacteer I.R.I.S. om deze sleutel te verkrijgen". At the bottom, it asks "Voer uw softwaresleutel in:" followed by an empty text box. At the very bottom are two buttons: "OK" and "Ik heb deze sleutel niet".



## DE READIRIS GEBRUIKERSINTERFACE VERKENNEN

Het Readiris programmavenster bevat niet alleen de **commandomenu**'s, maar ook de knoppenbalken die snelle toegang geven tot alle frequent gebruikte bevelen. Aanvankelijk zijn sommige commando's grijs: ze betreffen de preview. Zolang er geen beeld geladen is, zijn ze niet voorhanden.





Dat geldt ook voor de **beeldknoppenbalk** aan de rechterkant van het programmavenster: deze balk bevat alle commando's die u tijdens de voorvertoning nodig heeft. De **algemene toolbalk** links geeft snelle toegang tot de frequente bevelen van algemene aard.

Om te vernemen welk bevel achter welke knop schuilt, kunt u uw muis een tijdje boven een knop houden: een **tooltip** zal u vertellen welke functie de knop vervult.



De **beeldzone** is de plaats waar de gescande beelden worden getoond. U kunt beeldbestanden naar deze zone (en naar het Readiris pictogram) slepen om ze te herkennen.

Van zodra een bladzijde wordt verwerkt, verschijnt een extra toolbalk, de **paginaknoppenbalk** links van de beeldzone: deze knoppenbalk stelt de verschillende bladzijden van het document voor en verleent met de rechter muisklik (het "Context" menu) toegang tot de bevelen die specifiek de bladzijden betreffen.



Daarnaast wordt het **documentpaneel** afgebeeld onder het gescand beeld. Het geeft **statistische data** weer en geeft informatie over alle gescande beelden - de beeldbron en beeldresolutie, de duur van het scannen en herkennen enz. (Het documentpaneel is ook uitgerust met tooltips...)

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuimte	Rotatie
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\multipage.tif	1.30	-	300	2000	2388	?	-
<input checked="" type="checkbox"/> 2	C:\Program Files\Readiris\multipage.tif	1.13	-	300	2000	1888	?	-
<input checked="" type="checkbox"/> 3	C:\Program Files\Readiris\multipage.tif	1.06	-	300	1912	2004	?	-

## VAN START GAAN MET EEN EERSTE OEFENING

De beste manier om vertrouwd te raken met het gebruik van Readiris is ongetwijfeld de software gebruiken. Een aantal **vooraf gescande beelden**



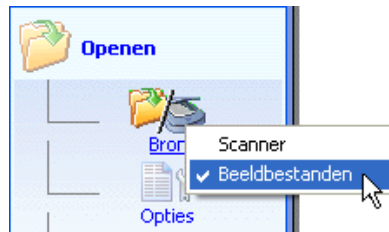
wordt bij de software geleverd; ze laten u toe om van start te gaan, zelfs als u nog geen scanner op uw PC aangesloten hebt. Laten we ermee van start gaan.

Met de knop "Bron" op de algemene knoppenbalk geeft u aan of u de scanner of een vooraf gescand beeld gebruikt als beeldbron.

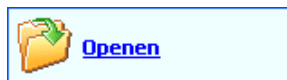
Readiris ondersteunt kleurbeelden, beelden in grijswaarden en zwart-wit beelden. Met Readiris kunt u Adobe Acrobat PDF documenten, DCX faxen (een meerbladige versie van het Paintbrush formaat), DjVu beelden (\*.djv, \*.djvu), JPEG beelden, JPEG 2000 beelden (\*.j2c, \*.jp2), PNG beelden, (niet gecomprimeerde of LZW, PackBits, Groep 3, Groep 4 en JPEG gecomprimeerde) TIFF beelden, meerbladige TIFF beelden, Windows bitmaps (\*.bmp) en ZSoft Paintbrush (\*.pcx) beelden openen.

Deze mogelijkheid is bijzonder nuttig om **faxen** om te zetten in editeerbare tekstbestanden.

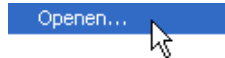
Om een vooraf gescand beeld te openen moet u "Beeldbestanden", en niet de scanner, als beeldbron kiezen met de knop "Bron".



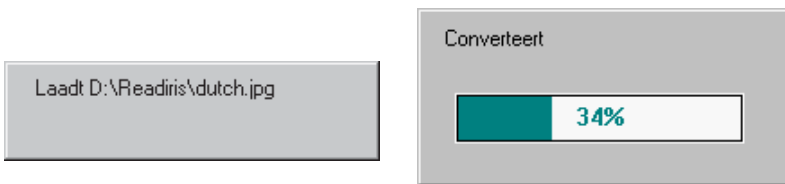
Daarna klikt u op knop "Openen". (Wanneer u de schijf als beeldbron kiest, wordt de knop "Scannen" vervangen door de knop "Openen" en wordt het corresponderende commando "Scannen" onder het menu "Actie" vervangen door het commando "Openen".)



U kunt ook het commando "Openen" onder het menu "Bestand" gebruiken om een vooraf gescand beeld te openen - dat werkt zelfs wanneer uw scanner als beeldbron fungeert.



U wordt uitgenodigd om een beeldbestand te kiezen. Selecteer het bestand DUTCH.JPG in de Readiris map. Daar dit een kleurbeeld is, wordt het niet enkel van de harde schijf gelezen; er wordt tevens een zwart-wit versie aangemaakt voor het OCR-proces.



Tenslotte wordt het beeld getoond in de beeldzone. De paginaknoppenbalk en het documentpaneel geven aan dat er één bladzijde in het geheugen geladen is.



**Readiris**  
Bestand Bewerken Instelling Bekijken Actie Leren Registreren Hulp

OCR wizard

1 Scannen

Scannen

Bron

Opties

Scanner

2 Herkennen

Herkennen Opslaan

Nederlands

Leren

Formaat

**Al jaren hebben vliegtuigen en schepen het aan boord. De laatste jaren zie je het ook steeds vaker in auto's: GPS of Global Positioning System. Voor wie een hekel heeft aan kaartlezen, is het een droom van een hulpmiddel.**

**G**PS in de wagen. Steeds meer mensen laten het inbouwen. Meer en meer automerken leveren het standaard op sommige modellen. Wie ooit met dit technologisch hoogstandje onderweg is geweest, wil nooit nog een wegeenkaart in zijn auto zien.

**Red je relatie**

Gedraai met het gezinnetje met je lief of je echtgenoot op weg naar de jaarlijkse vakantiebestemming. Wie kent het tafereel niet: man achter het stuur (kan natuurlijk ook de vrouw zijn), de koters op de achterbank die daar uiteraard hand liggend te dreinen van verveling en dan met die bepaalde afslag genomen worden. Man kijkt zenuwachtig opzij waar zijn halve trouwboek ingespannen zit te tuimelen op een hopeloos verfrommelde wegeenkaart. Het roept bij vertwijfel. De afrit komt maar rotsaart dichterbij, maar het enige wat hij hoort optuigen uit de berg papier die oogden moet opgevoeren wegeenkaart voorstelt: "Wacht en gesteren. Ein?" over hij steeds wanhopiger. Het wordt dood stil naast hem. De auto is bijna in de hoogte van de afrit.

Wat doe ik? schreeuwde de echtgenoot nu met een moorddadige blik in zijn roem. Op de achterbank stonden de koters onafgebroken zuchten naar het zieleke stukje papier boven de wegeenkaart. Het enige antwoord dat ze uiteindelijk kregen, tegen doffe, hopenloze blik, voor ze het beseffen, is: deze familie al weer op weg naar huis, nog voor hun vakantie goed eraan is begonnen is. Ach, hoeveel litewijgen zijn bijna niet berekenen aan het onvermogen kaart te lezen?

Nieuw: dan GPS in de wagen.

Een paar keer draaien aan een knop en je rijdt fluitend naar elke bestemming. Want zo eenvoudig werkt het: je geeft plaats en straat van je bestemming in en vervolgens berekent het systeem de route. Dat duurt hooguit enkele seconden. Je rijdt weg en een vriendelijke stem geeft je instructies over de te volgen reitroute. Synchron met de stem verschijnen er op het display van je autoradio of op een ander display, aanwijzingen met richtingspijlen of zelfs een heuse wegeenkaart.

Technisch ligt een en ander natuurlijk iets ingewikkelder. In feite maakt het navigatiesysteem gebruik van drie componenten: een satellietantenne, een navigatiecomputer en een digitale wegeenkaart op cd-rom of dvd-rom. De satellietantenne ontvangt signalen van het NAVSTAR Global Positioning System (zie kaderstuk). Op basis van deze satellietgegevens berekent de processor van de navigatiecomputer supersnel je positie en zoekt die op de digitale wegeenkaart die in de cd-speler of dvd-speler steekt. Als je daarna een bestemming ingeft, kan het systeem de route berekenen op basis van de gegevens op de cd-rom of dvd-rom. Begin je te rijden, dan herberekent het navigatiesysteem in je wagen je positie om de zoveel seconden aan de hand van de informatie van de GPS-satellieten die het koppelt aan metingen van een groescanvor (zeg maar een kompas) en een snelheidspeilsensor, die op basis van je snelheid de afgelegde afstand meet. Op die manier is de computer in staat je beweging nauwkeurig te volgen en je op basis van de digitale kaartgegevens te waarschuwen als je op een kruispunt bijvoorbeeld moet afslaan of recht door rijdt.

De *Narman-olieve (Gertvrijmoude) schuif je over de PDA, in elf geval de PDA van Compaq.*

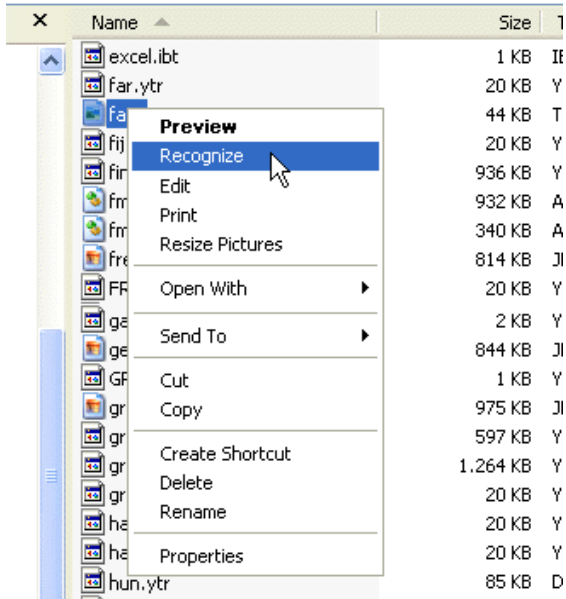
Er zijn uiteraard ver-

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuinte	Rotatie
1	C:\Progr...	4,27	-	300	1573	2464	-	-

Een derde manier om vooraf gescande beelden te openen is het gebruik van “**drag and drop**”: sleep beeldbestanden van de Windows Verkenner naar de beeldzone of naar het pictogram van Readiris en ze worden onmiddellijk geopend.



En er is nog een manier om beelden te openen vanuit de Windows Verkenner: **klik** op een beeldbestand **met de rechter muisknop** en selecteer het commando "Herkennen" van het "Context" menu. (Dit bevel verschijnt enkel indien het bestandstype door Readiris ondersteund wordt.)



Eén en ander betekent niet dat de herkenning ogenblikkelijk wordt uitgevoerd: de gebruiker behoudt de volledige controle over het OCR-proces. Readiris wordt gewoon opgestart en het beeld wordt in het geheugen geladen.

De beeldknoppenbalk bevat alle commando's die nuttig zijn tijdens de beeldpreview: u vindt er tools om de zones van belang aan te duiden, om het beeld te roteren, om in- en uit te zoomen enz.

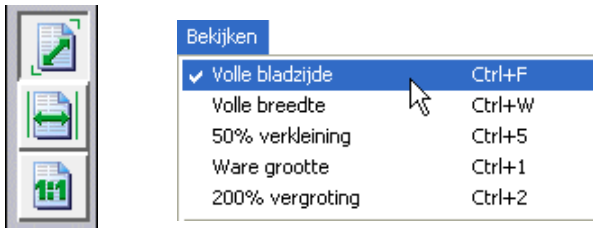
## INZOOMEN OP BEELDEN

---

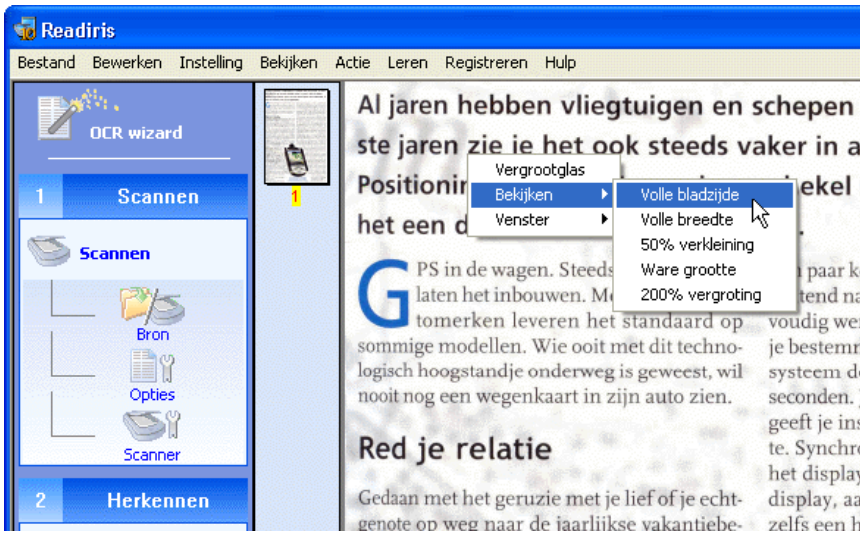
Readiris beschikt over enkele commando's waarmee u kunt **inzoomen** op het gescande beeld, bijvoorbeeld om de scankwaliteit te controleren.



De beeldknoppenbalk bevat knoppen waarmee u kunt inzoomen tot de werkelijke grootte, waarmee u het beeld kunt aanpassen aan de paginabreedte en waarmee u het volle beeld op het beeldscherm kunt brengen. Het menu "Bekijken" bevat diezelfde bevelen en voegt twee extra zoomniveau's toe: u kunt het beeld laten afbeelden op 50% en 200% van zijn ware grootte. Bij de werkelijke grootte beantwoordt één beeldpixel aan één scherpixel. (Er zijn overigens versnellingstoetsen beschikbaar voor alle de zoomniveau's.)



De zoomniveau's zijn overigens tevens beschikbaar via de rechter muisknop. Klik om het even waar met de rechtse muisknop om het "Context" menu op te roepen en selecteer er het gepaste zoomniveau.



U kunt tevens met de rechter muisknop *dubbel*klikken om op een deel van het gescande beeld in te zoomen op werkelijke grootte. Klik de rechter muisknop een tweede keer om opnieuw uit te zoomen.

Tenslotte kunt u van het **vergrootglas** gebruikmaken om in te zoomen op details van het gescand document. Het vergrootglas is ook beschikbaar op het "Context" menu wanneer u met de rechter muisknop boven het beeld klikt.

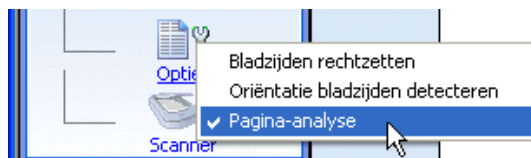




## EÉN, EEN GESCAND BEELD ANALYSEREN

Nu het beeld gescand is, kunt u aangeven welke delen u wilt omzetten in bewerkbare tekst door kaders, zogenaamde “vensters”, rond de zones van belang te tekenen.

Overigens kan Readiris dit automatisch voor u doen wanneer u de optie "Pagina-analyse" activeert met de knop "Opties" op de hoofdknoppenbalk (of onder het menu "Instelling").





De automatische pagina-analyse is bijzonder nuttig wanneer u in kolommen geschikte tekst moet herkennen en documenten met een ingewikkelde lay-out die mogelijkwerwijze illustraties en tabellen bevatten.

Bladzijde	Beeldbron	Scannen	Herkennen	Resolutie	Breedte	Hoogte	Regelschuimte	Rotatie
1	C:\Program ...	4,27	-	300	1573	2464	?	-

De pagina-analyse maakt gebruik van drie venstertypes: er zijn tekstvensters, grafische vensters en tabelvensters. Readiris onderscheidt de tekstblokken van de tabellen en de grafische zones die foto's, illustraties enz. bevatten. (We komen verder terug op de opslag van figuren en de herkenning van tabellen.)

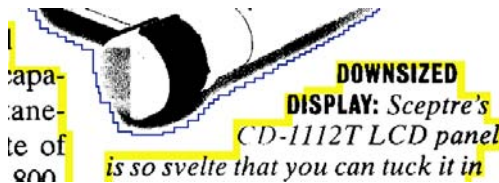
Twee extra venstertypes moeten altijd handmatig worden getekend: streepjescodevensters en handschriftvensters. (Straks meer over het lezen van streepjescodes en handgeschreven “blokletters”.)

Een **kleurcode** geeft aan met welk venstertype u te maken hebt: tekstvensters hebben een oranje rand, grafische vensters een purperen rand en tabellen een roze rand. Streepjescodezones zijn groen en handschriftzones zijn blauw.

Het aantal vensters is altijd aangegeven in de tooltip van de venstertools.



De pagina-analyse verloopt snel, springt tolerant om met lijnkromming en is erg accuraat: de analyse vindt complexe, “onregelmatige” vormen terug.



De pagina-analyse zal zelfs zones detecteren met **witte tekst op een zwarte achtergrond**. Zulke tekstkaders herkennen vormt geen probleem: terwijl de voorvertoning het gescande document getrouw op het beeldscherm brengt, zal Readiris tijdens de herkenningfase het beeld “negatief maken” om het kader te kunnen herkennen. (U kan ook de scanner *volledig* “negatieve” beelden laten aanmaken om zo bladzijden te verwerken met witte tekst op een zwarte achtergrond. Zie verderop.)



## ANDERHALF, VENSTERS SORTEREN

---

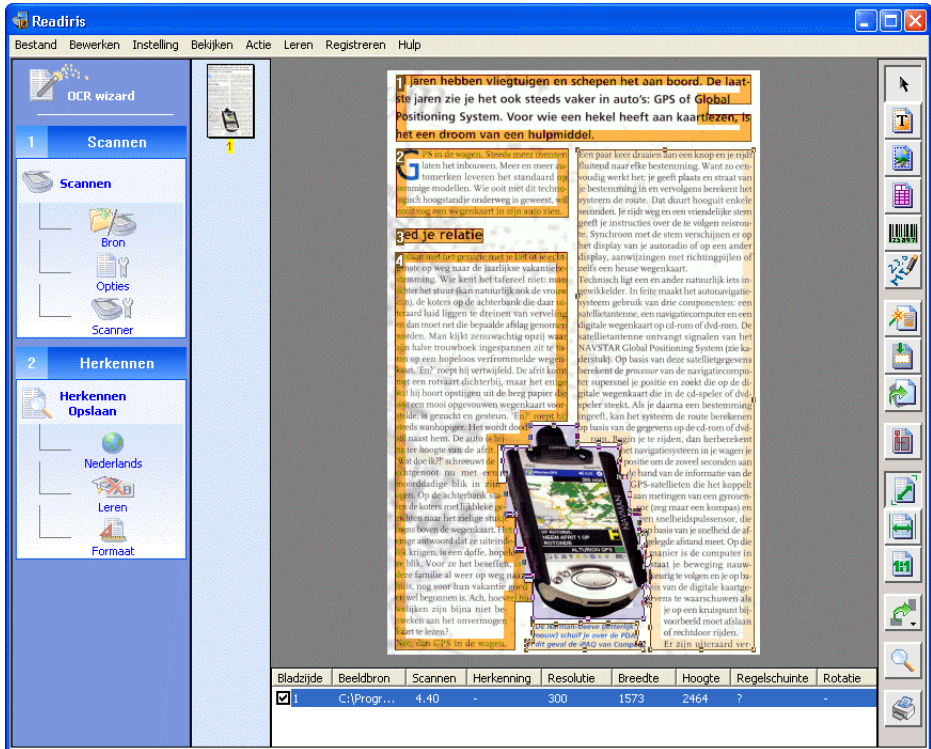
Readiris detecteert niet alleen de verschillende zones, maar *sorteert* die ook: standaard worden de vensters van boven naar onder, van links naar rechts gesorteerd zodat in kolommen geschikte documenten goed verwerkt worden.

Natuurlijk kunt u de **sorteervolgorde** wijzigen. Om dit te doen klikt u op de knop "Sorteren" op de beeldknoppenbalk. Van zodra de "sorteermodus" actief is verandert de muiscursor in een hand met wijzende vinger.



Klik op de vensters die u wenst over te nemen. De vensters die u *niet* aanklikt zullen simpelweg niet herkend worden. Vernemen welke vensters wel en niet geselecteerd zijn is erg makkelijk: de geselecteerde vensters hebben hun volle kleur, de niet geselecteerde vensters zijn lichter van kleur en hebben geen nummer.





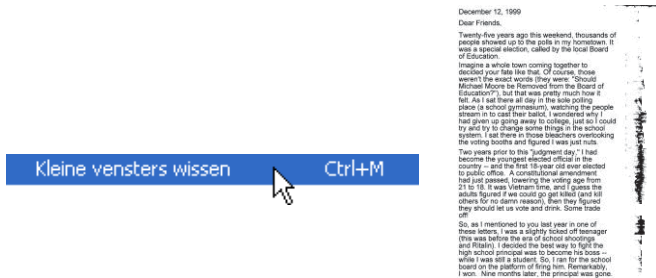
Standaard is de pagina-analyse geactiveerd. Om Readiris te dwingen de huidige pagina te analyseren - u hebt bijvoorbeeld de pagina-analyse per ongeluk uitgeschakeld, u heeft sommige vensters gewist en nu wilt u de pagina-analyse herhalen enz. -, klikt u op de knop "Pagina Analyseren" in de beeldknoppenbalk.





Selecteer de taal van het document *eer* u de pagina-analyse uitvoert bij Aziatische en Hebreeuwse documenten. Voor die talen worden er speciale routines gebruikt: de spatie tussen twee lijnen is meestal groter dan het geval is in Westerse documenten, de tekst bestaat uit kleine ikoontjes (“ideogrammen”) die in Westerse documenten wel eens als grafische zone kunnen worden gezien, en de tekst kan van rechts naar links, van boven naar beneden lopen. In Hebreeuwse documenten loopt de tekst van rechts naar links. En mocht u vergeten zijn de correcte taal te selecteren, selecteer die dan achteraf: Readiris voert de pagina-analyse vanzelf opnieuw uit!

Sommige documenten bevatten vlekjes op de bladzijde, zorgen voor een zwarte rand rond het eigenlijke beeld enz. U kunt alle kleine vensters wissen - we veronderstellen dat die geen tekst bevatten - en de overblijvende vensters opnieuw hersorteren met het bevel "Kleine Vensters Wissen" onder het menu "Bewerken".



Een gelijkaardige routine werkt automatisch: de detectie van vensters op de randen van de bladzijden. Wanneer deze routine uitgeschakeld is, zal de pagina-analyse alle zones negeren die de randen van de gescande bladzijden raken. Wanneer uw scanner zwarte randen toevoegt rond het eigenlijke beeld, zal de pagina-analyse vaak vensters detecteren die eigenlijk alleen “ruis” bevatten. Grafische vensters die de randen raken worden behouden: documenten bevatten vaak foto’s die de boorden van het blad raken of er is een achtergrondfoto die de hele bladzijde bedekt enz.





## TWEE, BEELDEN MANUEEL BEVENSTEREN

De pagina-analyse is de automatische manier om gescande pagina's te bevensteren. Maar u kunt de zones van een beeld ook handmatig aangeven met de **venstertools** van Readiris!

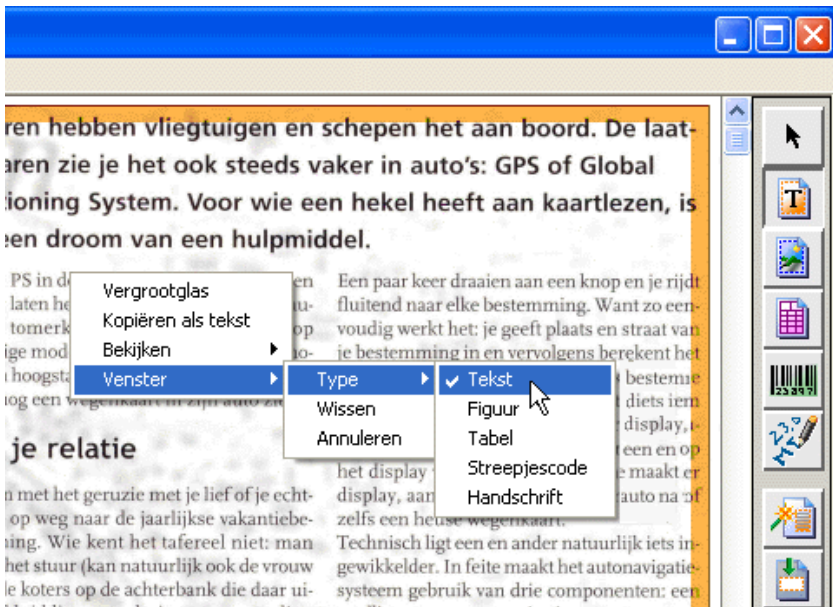
- Venster selecteren
- Tekstvensters tekenen: 0
- Grafische vensters tekenen: 0
- Tabelvensters tekenen: 0
- Streepjescodevensters tekenen: 0
- Handschriftvensters tekenen: 0

(We gaven al aan dat streepjescode- en handschriftvensters altijd manueel worden getekend door de gebruiker: de pagina-analyse detecteert dergelijke zones immers niet!)



Om een rechthoekig kader te **tekenen** rond een te hernemen zone, selecteert u de corresponderende tool in de beeldknoppenbalk. Sleep de cursor van de linker boven hoek van het venster naar de tegenovergestelde hoek rechts onder. (Zijden kleiner dan 1 mm zijn niet toegelaten, ze zouden overigens niet één karakter bevatten.)

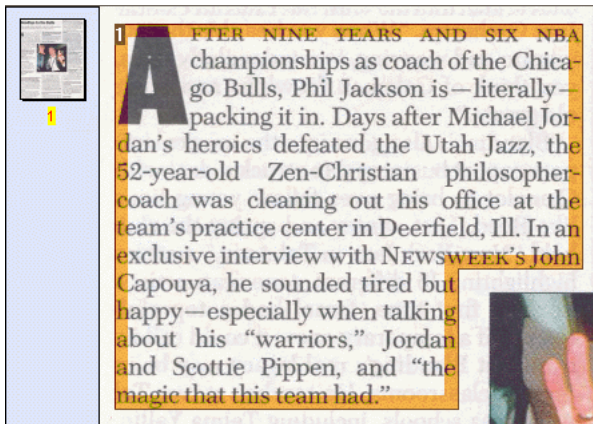
Maak u geen zorgen mocht u het verkeerde venstertype geselecteerd hebben: u kunt het zonetype snel wijzigen door met de rechter muisknop boven een venster te klikken. Selecteer het bevel "Venster - Type" in het "Context" menu.



De vensters worden automatisch gesorteerd naarmate ze gecreëerd worden: nummers geven de volgorde aan.

U kunt ook “onregelmatige” tekstblokken tekenen door **polygone vensters** rond dergelijke tekstzones te trekken. Veelhoekige vensters creëert men door

rechthoekige vensters te versmelten: van zodra twee rechthoeken (van hetzelfde type) elkaar raken, worden ze automatisch omgezet in één venster! Het is alsof u een huis bouwt door de ene kamer na de andere toe te voegen... (Veelhoekige tabelvensters en streepjescodevensters creëren heeft geen zin.)



Bovendien kunt u manuele bevenstering combineren met het sorteren van vensters; u kunt namelijk ook nieuwe vensters aanmaken wanneer de "sorteermodus" actief is. U sorteert dan een aantal vensters dat door de pagina-analyse werd gedetecteerd en u creëert nieuwe vensters waar de pagina-analyse geen bevredigend resultaat opleverde. Van zodra u een nieuw venster tekent in de "sorteermodus", worden alle vensters gewist die op dat ogenblik nog niet geselecteerd waren!

Om vensters te wijzigen, te verplaatsen en te wissen, moet u ze eerst **selecteren**. Om dit te doen, kiest u de knop "Vensters Selecteren" (de "pijl" knop) in de beeldknoppenbalk en klikt u op het venster. Vierkanten markeerpunten verschijnen op elke hoek en in het midden van de vensterzijden.



een droom van een hulpmiddel.

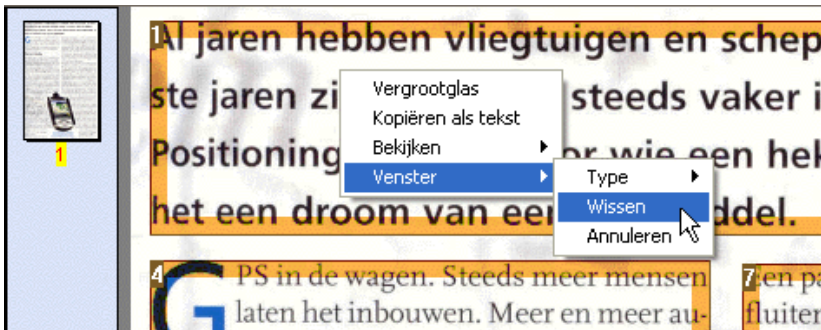


Klik de muisknop ergens anders om de **selectie** van vensters **ongedaan te maken**. Om **aanvullende vensters** te selecteren, houdt u de Shift knop ingedrukt terwijl u op de bijkomende vensters klikt. Om samen met een venster tegelijk de ingesloten vensters te selecteren (van een ander type), houdt u de Ctrl toets ingedrukt terwijl u op het hoofdvenster klikt.

Tot zover het selecteren van de vensters. Om een tekstvenster te **wijzigen**, selecteert u het, plaatst u uw muiscursor boven een markeerpunt en sleept u de vensterzijde zodat de afmetingen veranderen.

Om een venster te **verplaatsen**, selecteert u het en sleept u het naar een andere locatie.

Om vensters te **wissen**, selecteer u ze, klikt u met de rechter muisknop en selecteert u het bevel "Venster - Wissen" in het "Context" menu. Deze operatie wist alle geselecteerde venster én de zone onder de muiscursor.



U kunt tevens zones selecteren en het commando "Knippen" of "Wissen" van het menu "Bewerken" gebruiken. Het commando "Knippen" plaatst het venster(s) in een interne buffer, het commando "Wissen" wist het (de) venster(s) onherroepelijk. Wanneer u vensters plakt, worden ze ingevoegd worden op hun originele plaats; het is de taak van de gebruiker ze te verplaatsen naar hun nieuwe locatie.

Merk op dat *alle* klassieke opdrachten van het menu "Bewerken" van toepassing zijn op de vensters: vensters kunt u wissen, knippen, kopiëren en plakken! Het

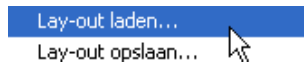
commando "Annuleren" speelt ook zijn rol: u komt op uw stappen terug wanneer u per ongeluk vensters heeft gewist, verplaatst, vergroot enz.

Bewerken	
Annuleren	Ctrl+Z
Knippen	Ctrl+X
Kopiëren	Ctrl+C
Plakken	Ctrl+V
Wissen	Delete
Kleine vensters wissen	Ctrl+M
Alles selecteren	Ctrl+A

Merk tevens op dat er versnellingstoetsen beschikbaar zijn voor alle bevelen! Laten we een voorbeeld geven: om alle bestaande vensters te wissen, kiest u het commando "Alles Selecteren" onder het menu "Bewerken" of de versnellingstoets Ctrl+A en klikt u op het comando "Wissen" of zijn versnellingstoets Del. U bent nu klaar om een nieuwe vensterlay-out aan te maken. Om de vorige lay-out te herstellen, kunt u "Annuleren" kiezen of zijn versnellingstoets Ctrl+Z.

## **DRIE, VENSTERSJABLONEN OPSLAAN**

De resulterende vensterlay-outs kunnen opgeslagen worden als venstersjablonen met het commando "Lay-out Opslaan" onder het menu "Bestand" om in de toekomst opnieuw gebruikt te worden. U laadt een sjabloon in het geheugen met het commando "Lay-out Laden".

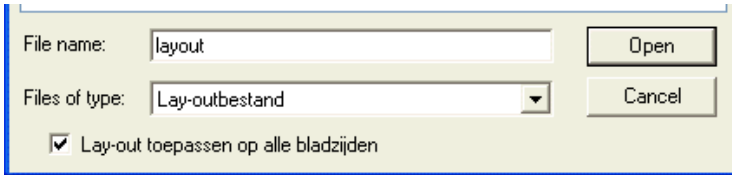


Als u documenten met een identieke opmaak wilt herkennen - bijvoorbeeld een rapport van 50 bladzijden waar de hoofding en de voetnoten om evidente redenen worden uitgesloten -, kunt u één sjabloon hanteren voor alle 50 pagina's.

Wanneer u een sjabloon in het geheugen laadt, wordt de pagina-analyse automatisch uitgeschakeld. De vensterlay-out blijft van kracht totdat u de pagina-analyse opnieuw activeert in de beeldknoppenbalk. Wanneer u een sjabloon laadt,



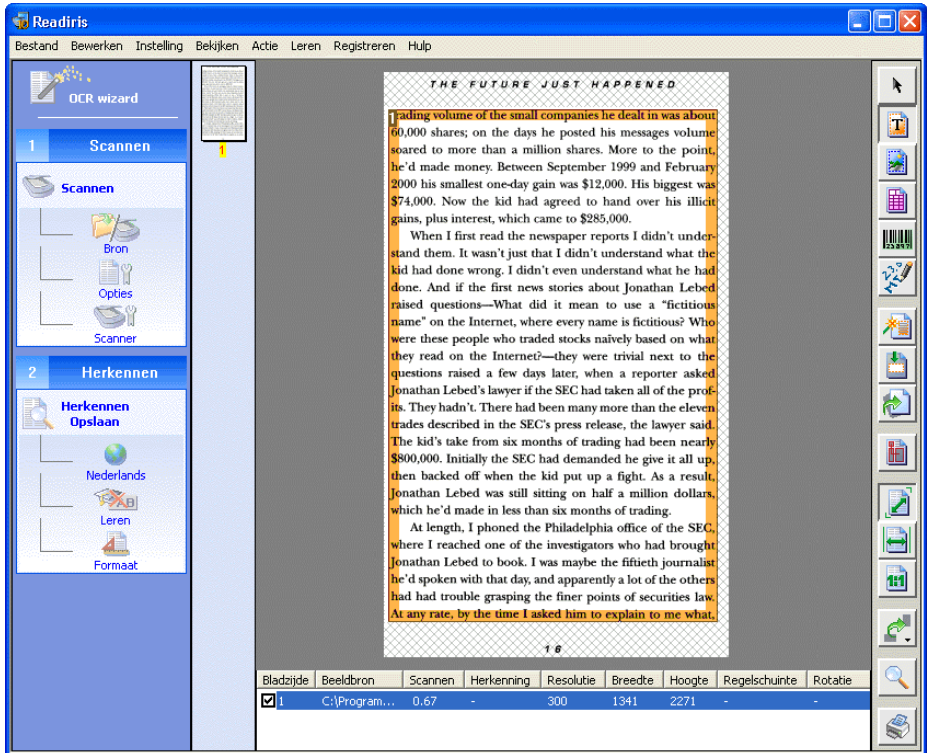
kunt u de optie "Lay-out Toepassen op Alle Bladzijden" activeren op de lay-out meteen op alle bladzijden van het huidige document toe te passen.



Overigens is er een leuk alternatief voorhanden voor zonesjablonen: de previewtool "Buitenkant van Zone Negeren" beperkt de pagina-analyse tot het "omkaderde" deel van het beeld.



Selecteer deze tool en trek een kader rond dat deel van het beeld dat u wilt verwerken. Wanneer u met een meerbladig document te maken hebt, kunt u dezelfde buitengrens hanteren voor alle bladzijden. (Voer de pagina-analyse opnieuw uit om het "bijsnijden" te annuleren, of pas de zones handmatig aan.)



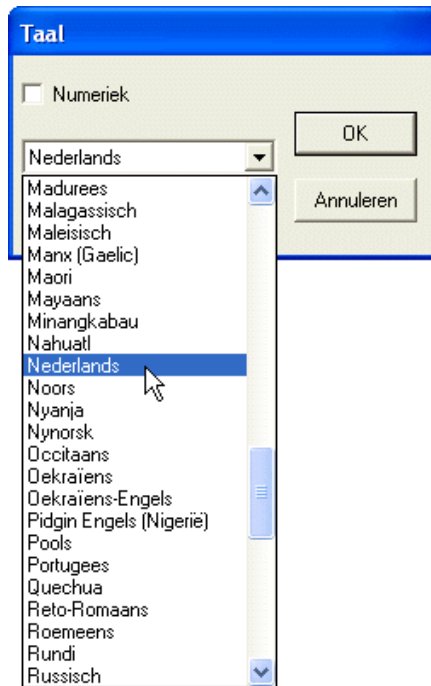
## MET READIRIS KUNT U DE HELE WERELD ROND!

Nu de vensters correct getekend zijn, bent u bijna klaar om de karakterherkenning uit te voeren. We zeggen “bijna”, want u heeft de taal en de instellingen van het document nog niet gecontroleerd!

De taalinstelling vindt u op de hoofdknoppenbalk terug.



Klik op de knop "Taal" om de taal van het document aan te passen.



U kunt een lettertoets indrukken om onmiddellijk naar een taal te gaan: als u momenteel Nederlands geselecteerd hebt en u wilt Quechua selecteren drukt u op de "Q" toets op uw klavier om onmiddellijk naar het Quechua te gaan. Wanneer verschillende talen dezelfde beginletter hebben drukt u herhaalde malen op deze toets om de verschillende opties te doorlopen. Bijvoorbeeld: Readiris leest Deens



en Duits. Door éénmaal op "D" te drukken, selecteert u het Deens, door een tweede maal op "D" te drukken selecteert u het Duits en door een derde maal op de "D" toets te drukken keert u terug naar het Deens. (Om naar een *andere* letter te gaan, bijvoorbeeld T, drukt u op de toets BackSpace voor u de "T" toets indrukt.)

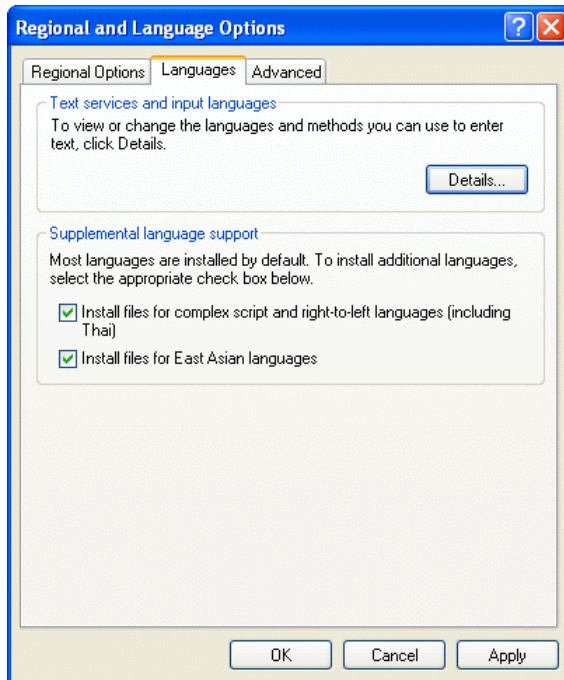
Readiris is helemaal niet beperkt tot het Nederlands: er worden wel 117 **talen** ondersteund! Alle Europese en Amerikaanse talen worden herkend, en daar horen ook de Centraal-Europese talen, het Grieks, het Turks, de Cyrillische talen (de "Russische" talen) en de Baltische talen bij.

Optioneel herkent Readiris **Hebreeuwse** en **Aziatische documenten**: de extra module voor Hebreeuwse herkenning leest zoals verwacht Hebreeuwse documenten in, de software-optie voor Aziatische herkenning leest het Japans, het vereenvoudigd en traditioneel Chinees en het Koreaans. (Het vereenvoudigd Chinees wordt gebruikt op het Chinees vasteland en in Singapore, het traditioneel Chinees in Hong Kong, Taiwan, Macau en de overzeese Chinese gemeenschappen.)

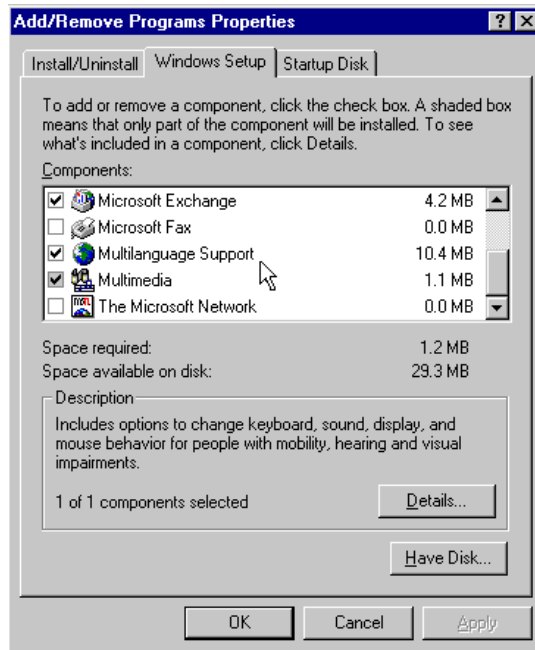
Merk ook op dat de Britse en Amerikaanse varianten van de Engelse taal - of zullen we eerder over "internationaal" Engels spreken in plaats van Amerikaans? - onderscheiden worden. Dat geldt ook voor het Spaans en het Mexicaans.

Merk op dat u een aangepaste configuratie van Windows nodig hebt om de Centraal-Europese, Griekse, Turkse, Cyrillische en Baltische karakters te ondersteunen. U dient de **veeltalige ondersteuning** van **Windows** te installeren opdat uw Windows systeem met deze talen kan omspringen.

Op een machine die uitgerust is met het besturingssysteem Windows XP, 2000 en Windows NT 4.0, selecteert u het pictogram "Regionale Instelling (en Talen)" onder het "Controlepaneel".



Indien u daarentegen met Windows ME of 98 werkt, selecteert u het pictogram "Software Toevoegen/Wissen" onder het "Controlepaneel" om uit te zoeken of de Windows module "Veeltalige Ondersteuning" al op uw PC geïnstalleerd is.



Om Aziatische en Hebreeuwse documenten op het scherm af te beelden en te bewerken, kunt u een Aziatische en Hebreeuwse versie van het Windows besturingssysteem installeren. Maar u kunt ook Word 2003, Word 2002 en Word 2000 gebruiken om dergelijke documenten te openen en bewerken: Office 2003 System, Office XP en Office 2000 werden specifiek ontworpen om te kunnen omspringen met documenten uit vele verschillende talen! Raadpleeg het “**Read Me**” bestand van Readiris als u hierover meer wilt vernemen.

De correcte taal kiezen is van essentieel belang. Op basis van de geselecteerde taal, zal het programma weten welke **karacterst** dient herkend te worden. De



veeltalige ondersteuning garandeert dat u “exotische” karakters zoals  $\beta$ , ñ,  $\gamma$  en  $\sigma$  correct herkent.

Ten tweede maakt het programma uitgebreid gebruik van **linguïstische databanken** om de resultaten te valideren. Stel dat u het woord "president" moet lezen waarbij een inktvlek de "r" in een "f" heeft veranderd. Door het voorlopige resultaat op te zoeken in het lexicon zal Readiris autonoom ontdekken dat het woord "president" gelezen wordt en dat het dus geen zin heeft om het symbool als "f" te herkennen. Deze **zelflerende techniek** hangt natuurlijk sterk af van de linguïstische context.

De taalkunde vormt een belangrijk hulpmiddel om **ambiguë gevallen** op te lossen zoals een "O" die als een '0' kan worden gelezen. Of denk aan de letter "l" en het cijfer '1' die in heel wat lettertypes dezelfde vorm hebben - denk aan teksten die afkomstig zijn van een oude schrijfmachines! De linguïstische context helpt bepalen of u te maken heeft met een "l" of een '1'.

De onderstaande illustratie laat verschillende exemplaren zien van ééntjes en l-symbolen. De vormpjes op de eerste lijn zijn niet dubbelzinnig, de vormen op de tweede lijn zijn dat wel, maar de taalkunde biedt een uitweg. Wanneer de context geen uitsluitel biedt, moet de gebruiker tussenkomen.

193 1950s. 1hr  
Well, Rossellini

## NOG MEER TAALKUNDE MET GEBRUIKERSLEXICONS

U kunt de linguïstische “feedback” nog een stap verder brengen door die te personaliseren: hoe krachtig de standaard lexicons ook mogen zijn, gebruikers van Readiris Corporate kunnen de **OCR-precisie** nogmaals verhogen door gebruikerslexicons in het geheugen te laden met het commando "Gebruikerslexicon" onder het menu "Instelling".

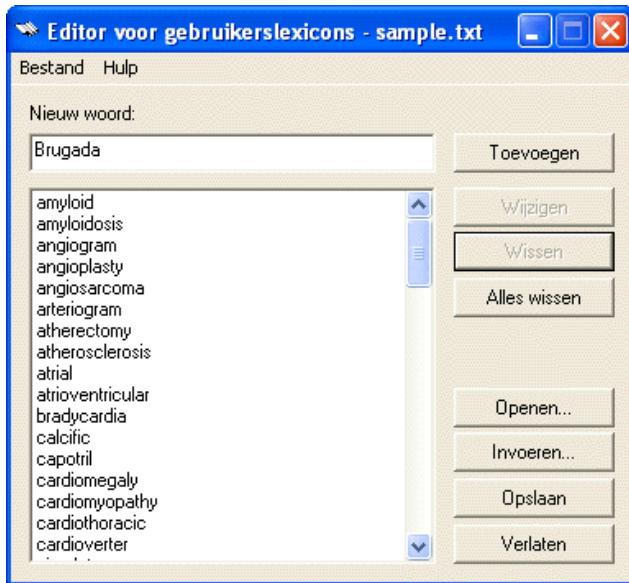


De tooltip van de knop "Taal" geeft aan welk gebruikerslexicon momenteel actief is.

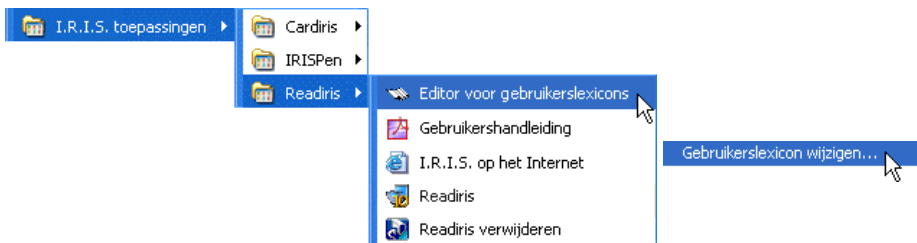


Gebruikerslexicons zijn woordenlijsten met termen die niet voorkomen in de basislexica van algemene aard. Denk bijv. aan technische, wetenschappelijke, legale terminologie of aan andere **termen die de activiteit van een onderneming** weerspiegelen.

Readiris wordt vergezeld van de Editor voor Gebruikerslexicons, een dienstprogramma waarmee u deze gebruikerslexica creëert en onderhoudt. Deze utility is erg gebruiksvriendelijk; raadpleeg de on-line hulpsysteem mocht u vragen hebben bij het gebruik.



U vindt deze editor terug in het sub-menu "I.R.I.S. toepassingen - Readiris" en onder het menu "Instelling" van Readiris.



## **READIRIS VERANDERT TUSSENDOOR VAN TAAL**

---

Maar het blijft hier niet bij: Readiris kan middenin een zin van taal veranderen zonder tussenkomst van de gebruiker! Wanneer Westerse woorden opduiken in Griekse, Cyrillische, Hebreeuwse en Aziatische documenten - vele eigennamen, merknamen enz. worden geschreven in het Latijns alfabet - zal Readiris automatisch naar het correcte alfabet grijpen. Met andere woorden, u kunt een **gemengd alfabet** activeren van Griekse, Cyrillische, Hebreeuwse of Aziatische karakters in combinatie met Latijnse karakters.

Wees er zeker van dat u "Grieks-Engels" geselecteerd hebt of een gepaste Cyrillische taalinstelling - bijvoorbeeld "Oekraïens-Engels". Met andere woorden: selecteer niet gewoon het "Grieks" of "Oekraïens" als taal van het document in de hoop dat de Westerse symbolen wel goed door de herkenning zullen komen.



Dit is een voorbeeld waar een Russische tekst enkele Engelse woorden bevat - open het beeld ALPHABETS.TIF en herken de corresponderende bladzijde mocht u het zelf willen proberen!



Readiris

Bestand Bewerken Instelling Bekijken Actie Leren Registreren Hulp

OCR wizard

1 Scannen

Scannen

Bron

Opties

Scanner

2 Herkennen

Herkennen Opslaan

Russisch-E

Leren

Formaat

1

2

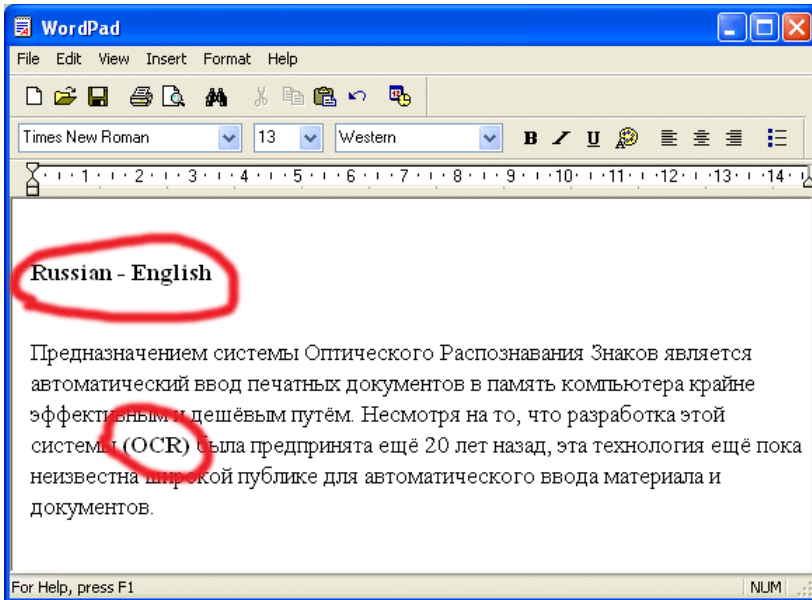
## Russian - English

Предназначением системы Оптического Распознавания Знаков является автоматический ввод печатных документов в память компьютера крайне эффективным и дешёвым путём. Несмотря на то, что разработка этой системы (OCR) была предпринята ещё 20 лет назад, эта технология ещё пока неизвестна широкой публике для автоматического ввода материала и документов.

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuimte	Rotatie
<input checked="" type="checkbox"/> 1	C:\Program ...	1.04	-	300	1889	1223	?	-
<input checked="" type="checkbox"/> 2	C:\Program ...	0.49	-	300	906	626	?	-
<input checked="" type="checkbox"/> 3	C:\Program ...	0.48	-	300	907	643	?	-

Wanneer u het resultaat met de tekstverwerker opent ziet het er zo uit - het kan gebeuren dat u een Cyrillisch **lettertype** moet activeren om de Russische tekst correct weer te geven op het beeldscherm.





## DOCUMENTEN MET GEMENGDE TALEN LEZEN

Readiris Corporate is veel krachtiger in het lezen van documenten waarin meerdere talen voorkomen: met die versie kunt u een hoofdtal en wel 4 **secundaire talen** (van dezelfde taalgroep) activeren!



Om een document te herkennen waar de Franse vertaling naast de Nederlandse tekst staat, moet u het Nederlands selecteren als hoofdtal en het Frans als secundaire taal of vice versa. Op die manier breidt u niet enkel de karakterset uit zodat de geaccentueerde karakters zoals ç, é en ù correct herkend worden. Bovendien zullen beide lexica taalkundige “feedback” leveren aan het herkenningproces, en dit laat Readiris zelfs toe om automatisch te detecteren welke taal waar in het document gebruikt wordt!

U kunt maximaal 5 talen tegelijk inschakelen. Enkel talen die tot dezelfde taalgroep behoren kunnen worden gecombineerd: talen die onder een andere Windows code page vallen dan de “hoofdtal” kunnen *niet* tegelijk actief zijn. Maar houd vooral hier rekening mee: selecteer geen talen die niet op uw docu-

menten van toepassing zijn. Hoe groter de karakterset, hoe trager de herkenning wordt en hoe groter het risico op leesfouten!

Indien u snelt wenst te vernemen welke talen momenteel actief zijn, weet dat de tooltip van de "Taal" knop ze voor u opsomt...



## **DE KENMERKEN VAN HET DOCUMENT AANGEVEN**

---

Nu de taal is ingesteld, gaan we over naar de andere kenmerken van het document. U kunt de herkenning aanpassen door enkele karakteristieken van het document aan te geven: het type en de karakterafstand van de lettertekens. (Deze commando's zijn niet van toepassing op Aziatische documenten.) Laten we verhelderen wat dit alles betekent.

We vangen aan met commando "Type Lettertekens" onder het menu "Instelling". Het "type" lettertekens onderscheidt de "normale" documenten van de **matrixgedrukte** documenten. "Draft" dot matrix symbolen - ook wel "9 pin" matrixdruk genoemd - zijn opgemaakt uit aparte vlekken, en er zijn sterk gespecialiseerde herkenningstechnieken nodig om dergelijke teksten te lezen.

**ape-descended life**

Matrixdruk met "briefkwaliteit", ook wel "25 pin" matrixdruk genoemd, vereist de "normale" instelling, en dat is ook het geval voor de **drukkwaliteiten** gedrukt, getikt, laserdruk en inkjetdruk.

De optie "Automatisch" betekent dat Readiris het lettertype automatisch zal terugvinden. Laat Readiris onder alle omstandigheden de drukkwaliteit automatisch detecteren - tenzij u zeker bent dat er enkel matrixgedrukte documenten worden ingelezen. ("Automatisch" is natuurlijk de standaard waarde.)



De drukkwaliteit wordt aangegeven in de tooltip van de knop "Herkennen-Op slaan": wanneer aan de tooltip geen bericht is toegevoegd, is de automatische detectie van de drukkwaliteit van toepassing, wanneer het bericht "Dot matrix" is toegevoegd, is de dot matrix leesmodus actief.



De **karakterafstand** kan gekozen worden met het bevel "Karakterafstand" onder het menu "Instelling".



Bij *vaste* lettertypes hebben alle symbolen van één lettertype dezelfde breedte. Een "l" neemt evenveel horizontale plaats op binnen een lijn als een "w", zoals in deze zin het geval is. Neem bijvoorbeeld de documenten die afkomstig zijn van een schrijfmachine: de wagen van de schrijfmachine schuift telkens op met een vaste afstand.

Bij een *proportioneel* lettertype hangt de breedte van elk symbool van zijn specifieke vorm af. Symbolen zoals "m" en "w" zijn duidelijk breder, nemen meer plaats in binnen een lijn dan de dunne karakters "l" of "j". Zowat alle boeken, tijdschriften en kranten gebruiken exclusief proportionele lettertypes.

Dit is de simpelste oplossing: laat deze optie altijd op "Automatisch" staan, zodat Readiris de karakterspatiëring automatisch zal detecteren.

## READIRIS WORDT TELKENS INTELLIGENTER!

Wanneer de taal en de kenmerken van het document aangegeven zijn, activeert u het interactief leren en klikt u op de knop "Herkennen-Opslaan".



De vooruitgang van het OCR-proces wordt op het scherm getoond. U kunt op de knop "Stop" klikken (en op Esc drukken) om de tekstherkenning af te breken.



Indien de leermodus ingeschakeld werd met de knop "Leren" op de hoofdknoppenbalk komt Readiris op het einde van de herkenning terecht in de interactieve fase.

(Het interactief leren is niet van toepassing op Aziatische documenten: het systeem trainen heeft geen zin bij talen die duizenden symbolen gebruiken - en dan gaan we er overigens van uit dat u de ideogrammen kunt intikken, geen makkelijke taak op een Westers toetsenbord!)



De leermodule kan in bepaalde gevallen de herkenning sterk verbeteren. Wanneer de gebruiker probeert om gebroken, misvormde letertekens in te lezen zoals men die terugvindt in echte documenten, of wanneer men op gestileerde letertekens botst die Readiris niet optimaal herkent, kan het **interactief leren** dit voorlopig “falen” overwinnen.

De leermodus wordt ook gebruikt om het systeem te trainen op **speciale symbolen** die Readiris initieel niet herkent, zoals wettenschappelijke symbolen en dingbats. Een paar voorbeelden: Readiris kan getraind worden om het symbool " $\pi$ " te herkennen als "pi" of de dingbat "☎" als "Tel". (Merk op dat de lijst van herkende symbolen echter niet kan uitgebreid worden met de symbolen " $\pi$ " en "☎"!)

De herkende tekst wordt progressief doorlopen en het systeem stopt op de twijfelachtige karakters of - als we het over elkaar rakende karakters (“ligaturen”) hebben - op de twijfelachtige karakterreeksen. De letertekens worden altijd in hun context getoond, met de nodige klemtoon op de twijfelachtige karakters. Niet herkende karakters worden voorgesteld met een tilde (het symbool "~").



Het eerste wat u moet doen is nagaan of het correcte woordenboek en de goede woordenboekmodus geactiveerd zijn - deze zijn altijd aangegeven in de titel van het leervenster. Mocht dat niet het geval zijn, klik dan op "Afbreken" - het beeld van het document wordt weer op het scherm gebracht mét zijn bevenstering -, laad het juiste woordenboek of de woordenboekmodus en herhaal de herkenning. (De werking van de woordenboeken komt zo meteen aan bod.)

Voer indien dat nodig is een karakter (of een karakterreeks) in voor de verkeerde of niet herkende tekens en klik op één van de volgende knoppen.

### **Leren**

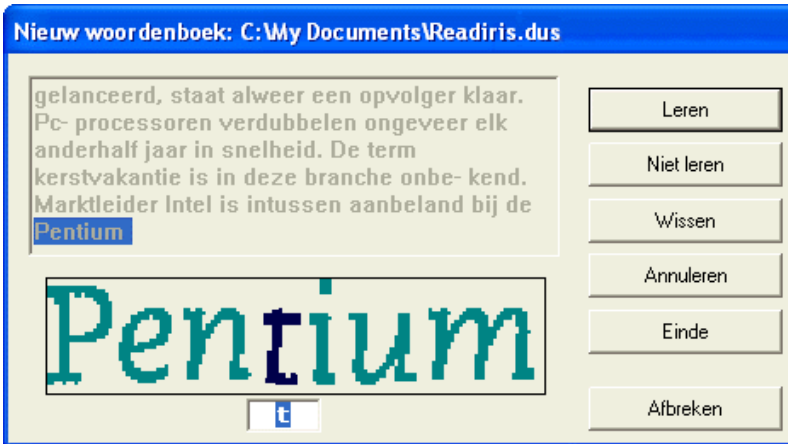
U gaat akkoord met de voorgestelde oplossing of u verbetert die. Het programma bewaart dit twijfelachtige karakter in het woordenboek als "zeker", definitief. De verdere herkenning ervan zal geen tussenkomst meer vergen, de vorm wordt als voor ééns en altijd als aangeleerd beschouwd.

In het bovenstaand voorbeeld stopt het systeem op een twijfelachtige vorm, en we klikken op "Leren" om een vorm aan te leren die niet kan verward worden met andere karakters.

### **Niet Leren**

U gaat akkoord met het voorgestelde oplossing of u verbetert ze. Het verschil met de knop "Leren" zit hierin dat de vorm de status "onzeker" krijgt in het woordenboek. Bij verdere herkenning zal het systeem u alvast wel de "aangeleerde" oplossing voorstellen, maar de herkenning zal nog steeds uw bevestiging vereisen.

Deze knop wordt gebruikt bij symbolen die kunnen verward worden met andere: een vervormde "e" die sterk op een "c" lijkt, een beschadigde "t" die veel weg heeft van een "r" enz.



De "t" hierboven is serieus beschadigd - in feite lijkt hij op een "r" -, en u klikt hier op "Niet Leren" om dit symbool niet met de "r" te verwarren.

### **Wissen**

De twijfelachtige vorm wordt uit de uitvoer verwijderd. Gebruik deze knop om "ruis" die op de documenten voorkomt - inkt- of koffievlekken die herkend raken als punten of komma's en dergelijke - uit de uitvoer te verwijderen, of om een willekeurig ongewenst symbool te wissen.

### **Annuleren**

U keert op uw stappen terug om fouten te verbeteren. Readiris houdt de 32 laatste operaties bij.



## **Einde**

Het leerproces wordt gestopt maar het OCR-proces gaat verder in de automatische modus. Alle oplossingen van het systeem worden aanvaard zonder validatie door de gebruiker.

Klik op deze knop wanneer u ziet dat de herkenning erg goed is en geen gedetailleerde verificatie behoeft.

## **Afbreken**

Verwar "Einde" niet met "Afbreken": bij "Afbreken" wordt er geen uitvoer gegenereerd en moet u de herkenning overdoen, met de knop "Einde" wordt de tekst wel aangemaakt, alleen neemt u die niet in detail door.

## **DE ROL VAN WOORDENBOEKEN**

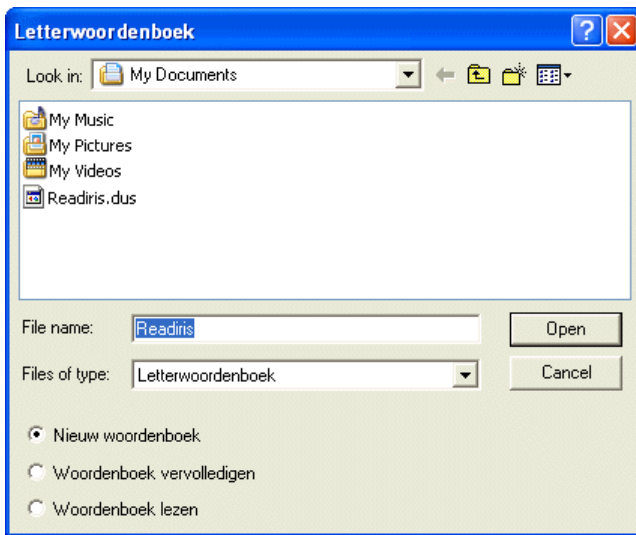
---

De resultaten van elke trainingssessie worden tijdelijk in het computergeheugen bewaard maar kunnen en moeten voor verder gebruik opgeslagen worden in bestanden die we “woordenboeken” noemen.

(Verwar letterwoordenboeken niet met lexicons! Woordenboeken bevatten de vormen van lettertekens waarop u het systeem trainde, lexicons zijn linguïstische databanken die de herkenning ondersteunen!)

De letterwoordenboeken worden in het geheugen geladen om bij gelijkaardige documenten van de extra intelligentie gebruik te maken. Readiris gebruiken betekent Readiris telkens intelligenter maken!

Hoe gaat dit in zijn werk? De werking van de woordenboeken wordt beheerd door het menu "Leren": u moet een woordenboek selecteren met het commando "Woordenboek" en zijn werkingsmodus bepalen.



**Woordenboeken** zijn beperkt tot 500 vormen; wij raden u aan aparte woordenboeken aan te maken voor specifieke applicaties - bijvoorbeeld per documenttype. Woordenboeken hebben standaard de extensie \*.dus. De training heeft geen effect meer wanneer het woordenboek vol is: de resultaten van het leerproces worden dan niet langer in het geheugen opgeslagen of naar het woordenboek weggeschreven.

U kunt de woordenboekmodus aanpassen met het commando "Woordenboek" of rechtstreeks onder het menu "Leren". Er zijn drie woordenboekmodi: nieuw, vervolledigen en lezen.



De optie "Nieuw Woordenboek" geeft aan dat de resultaten van de training zullen opgeslagen worden in een *nieuw* woordenboek. (Wanneer u een bestaand woordenboek selecteert wordt de inhoud gewist.)

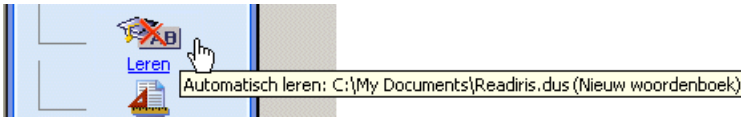
De uitbreidingsmodus geeft aan dat de trainingresultaten zullen bewaard worden in een *bestaand* woordenboek: de herkenning maakt al gebruik van de extra intelligentie die in het woordenboek aanwezig is, en u voegt nieuwe vormen aan het woordenboek toe. Het komt hierop neer dat u met deze optie een woordenboek opbouwt in verschillende stappen.

(Wanneer u een nieuwe bestandsnaam invult voor een woordenboek en u activeert de vervolledigingsmodus, zal een leeg woordenboek worden gecreëerd en gaandeweg ingevuld.)

Met de laatste optie, "Woordenboek Lezen", zal het woordenboek enkel gelezen, geraadpleegd worden: u maakt gebruik van het woordenboek *zonder* er wat dan ook aan toe te voegen ("read only").

Selecteer de modus "Nieuw" wanneer u één enkele bladzijde herkent. Om verschillende bladzijden van hetzelfde type - we bedoelen met hetzelfde lettertype en dezelfde drukwaliteit - te herkennen, kiest u de modus "Nieuw" voor de eerste bladzijde, de vervolledigingsmodus voor enkele volgende bladzijden en de leesmodus voor de rest van het document.

Noteer dat de tooltip van de knop "Leren" op elk ogenblik aantoont welk woordenboek actief is en in welke werkingsmodus dat woordenboek zich bevindt.



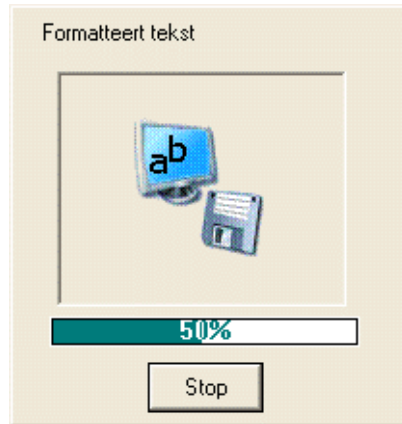
Wanneer u in het interactief leren terecht komt worden het woordenboek en zijn modus aangegeven in de titel van het leervenster, en u klikt op "Annuleren" om opnieuw te beginnen wanneer deze foutief blijken.



## TEKST DIRECT NAAR EEN TOEPASSING STUREN

Het interactief leren beëindigt de karakterherkenning. Daar Microsoft Word fungeert als standaard doeltoepassing voor de uitvoer, wordt uw tekstverwerker na de herkenning (indien nodig) automatisch opgestart en wordt de herkende tekst ingevoegd.

Het kan voorkomen dat u de vooruitgang van de formattering van het document op het scherm kunt volgen. (Of dat daadwerkelijk het geval is hangt van de grootte van het document en de complexiteit van de formattering.)

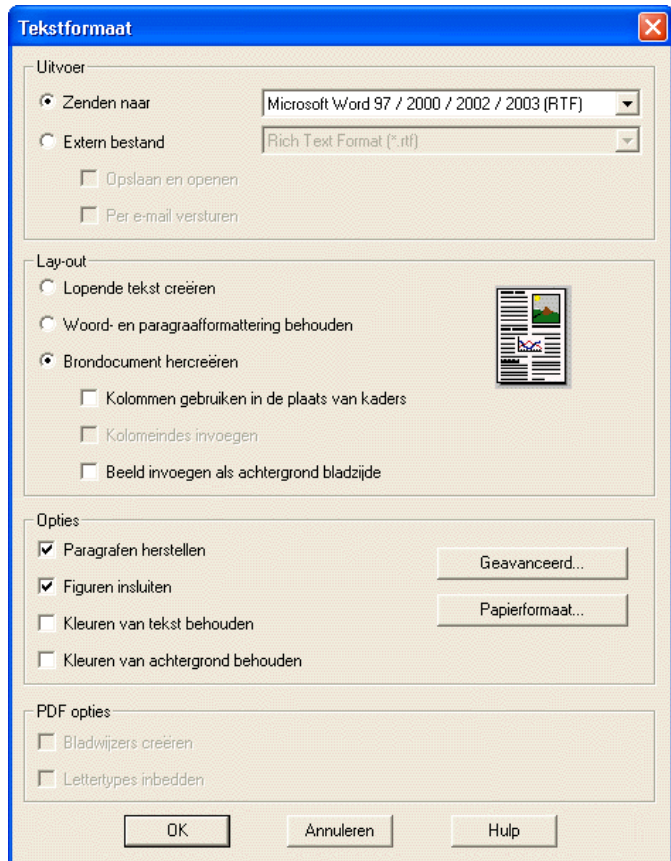


Het gescande beeld verschijnt opnieuw op het scherm mét de bevenstering. Het beeld is klaar voor verdere verwerking, het blijft afgebeeld tot u een nieuwe bladzijde inscant.

U heeft inderdaad een document omgezet in een editeerbaar computerbestand, en dit wel 40 maal vlugger dan indien u de tekst had hertikt! Ga uw gang en vergelijk het tekstresultaat met het beeld binnen het Readiris venster.

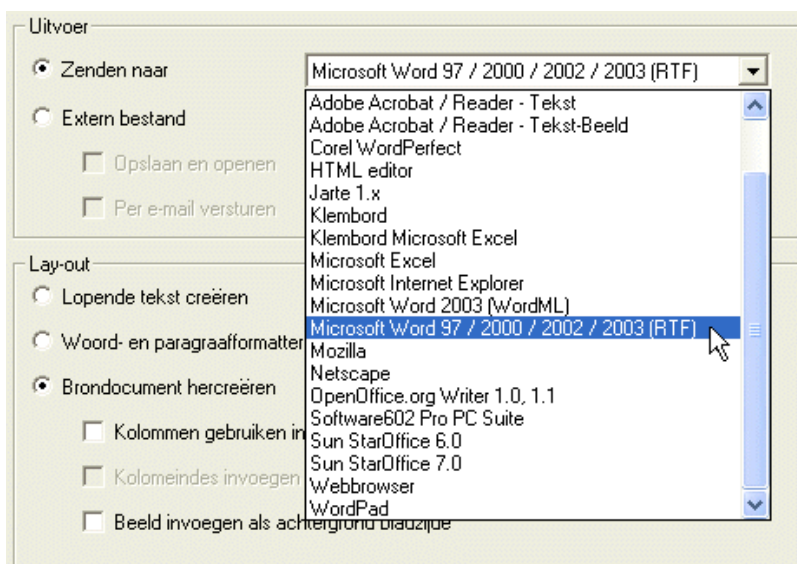
In feite biedt Readiris u drie verschillende methoden om de herkenningresultaten te bewaren: u kunt het herkende document rechtstreeks naar een doeltoepassing zenden, u kunt het resultaat opslaan in een extern bestand en in het Windows klembord stoppen.

De **uitvoermodus** wordt bepaald met de "Formaat" knop (of met het commando "Tekstformaat" onder het menu "Instelling").



De mogelijkheid om documenten “door te **zenden**” biedt u een rechtstreekse OCR link tussen uw scanner en uw Windows applicaties: u voert documenten

onmiddellijk uit naar uw tekstverwerker, rekenblad of webbrowser, naar Adobe Reader enz.!



Na de herkenning wordt de doeltoepassing opgestart en worden de resultaten automatisch ingevoegd.



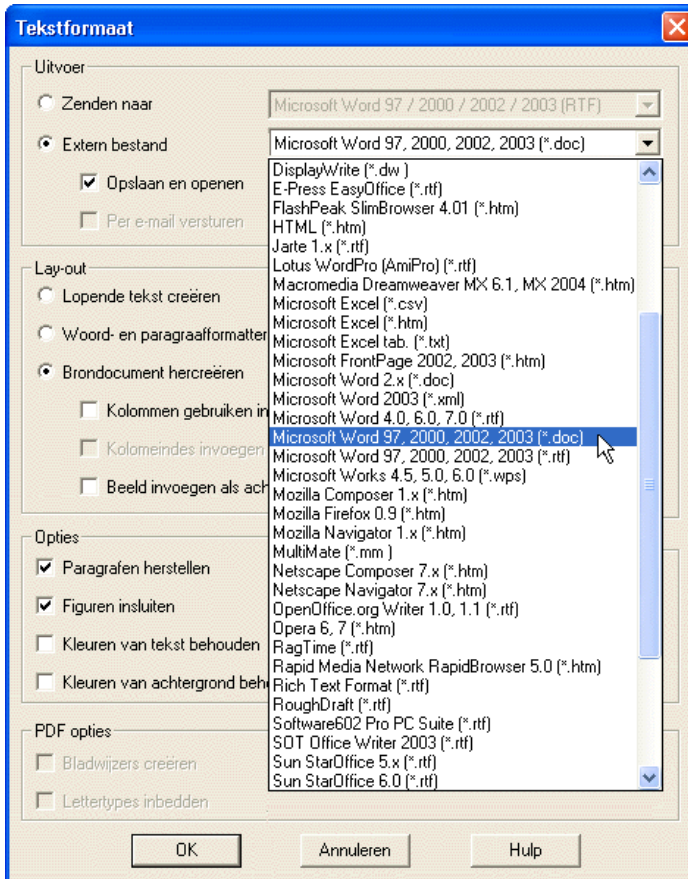
Vergeet echter niet dat de optie "Zenden naar" u ook teksten in het Windows **klembord** laat stoppen, er is dus geen strikte noodzaak om ze uit te voeren... of in een bestand op te slaan!

## **HET RESULTAAT IN EEN TEKSTBESTAND OPSLAAN**

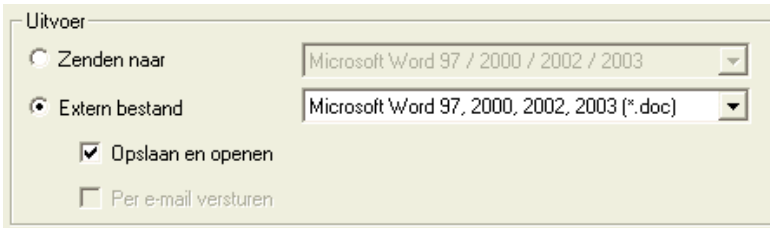
---

U kunt het herkenningresultaat inderdaad bewaren in een "extern" bestand. Ook hier ondersteunt Readiris een wijd gamma aan bestandsformaten dat alle populaire tekstverwerkers, rekenbladen, webtoepassingen enz. omvat. (Zo ondersteunt Readiris onder meer WordML, het nieuwe tekstformaat van Microsoft Office 2003!)

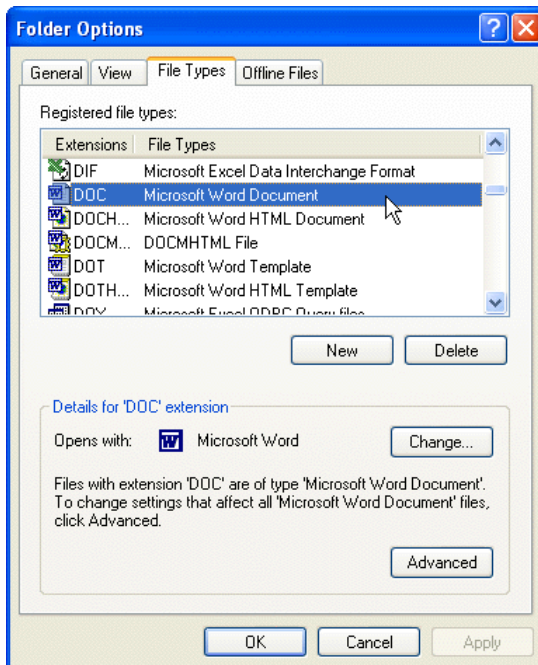




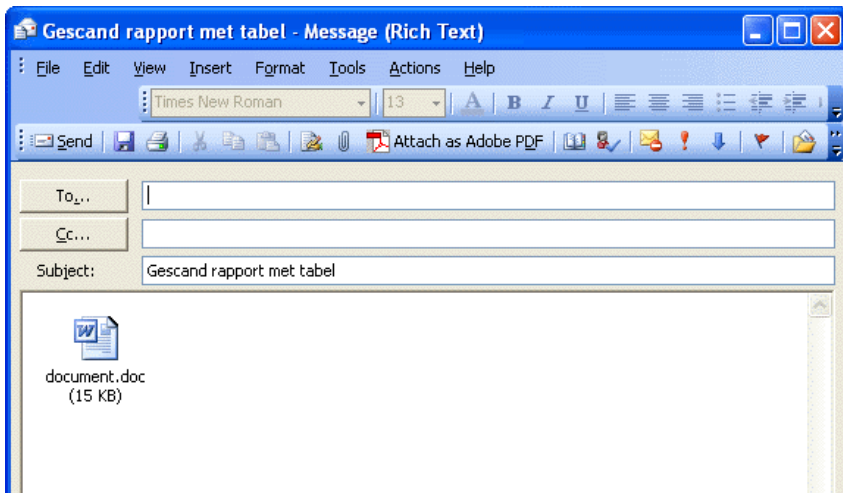
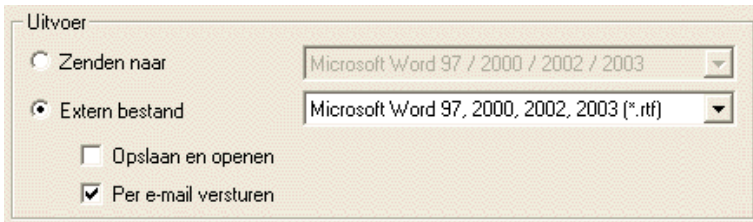
De Optie "Opslaan en Openen" is in grote lijnen gelijkwaardig met de optie "Zenden naar": u opent het herkende document eens u het opgeslagen hebt.



De manier die wordt aangewend om de doeltoepassing te doen lopen is echter fundamenteel verschillend. Hier bepaalt het **Windows bestandstype** welke toepassing zal opgestart worden. Het is net alsof u het uitvoerbestand in de Windows Verkenner twee keer aangeklikt hebt... (De optie "Zenden naar" richt zich rechtstreeks naar specifieke doeltoepassingen.)

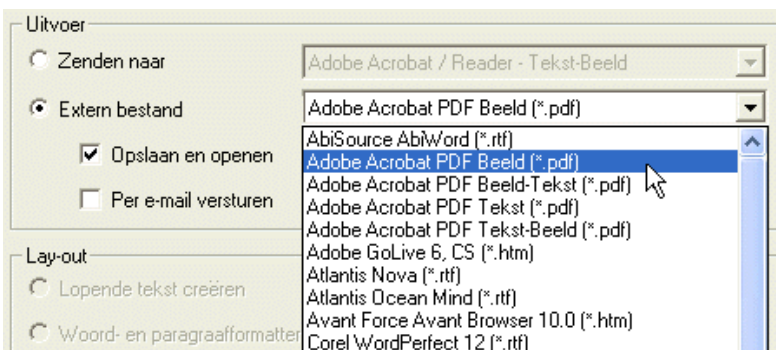
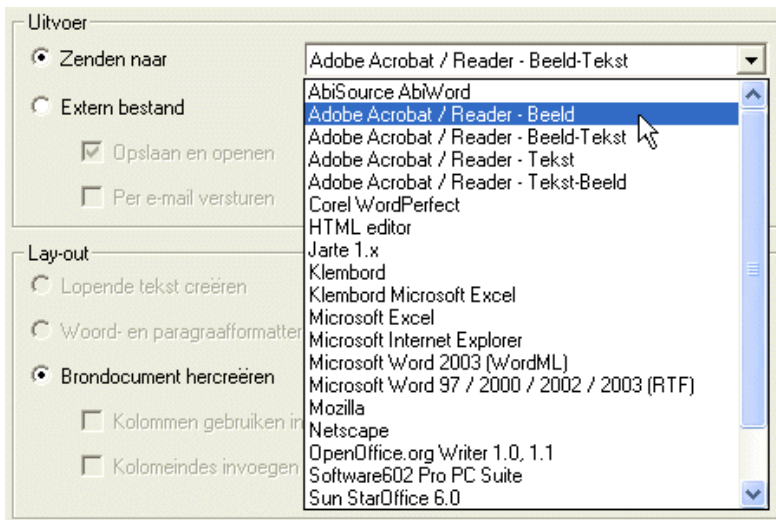


De optie "Per E-mail Versturen" creëert een nieuw e-mailbericht en voegt het herkend document toe als aangehecht bestand. Kent u een snellere manier om een papieren document vlug te verspreiden...?



## UNIVERSELE BESTANDEN AANMAKEN...

We gaan wat dieper in op één formaat: **Adobe Acrobat PDF**. Met Readiris kunt u PDF documenten aanmaken die op tekst en beeld gebaseerd zijn.



Waar zit het verschil tussen deze formaten? Wanneer u het formaat "PDF Tekst" kiest, maakt Readiris een PDF bestand aan dat het tekstresultaat bevat. (Er kunnen wel figuren in het bestand voorkomen maar enkel als die ook op de

bladzijde voorkomen - foto's, illustraties enz.) Met andere woorden: het beeld van de volle bladzijde is *niet* opgenomen in het PDF bestand en het PDF bestand bevat slechts één laag! Het formaat "PDF Beeld" maakt ook PDF bestanden aan van één laag, maar die bevatten natuurlijk het gescande beeld, niet het OCR resultaat!



Adobe Reader - [autoform.pdf]

Bestand Bewerken Beeld Document Opties Venster Help

Openen

46%

eBooks

# Autoformatting

The aim of "autoformatting" is to recreate a facsimile copy of the original document.

The OCR process does more than just recognize your text, it can format it for you too!

In a way, text recognition is becoming more and more page recognition or document recognition ...

Whether your OCR software reformats the recognized text or not is up to the user. You can perform OCR because you just need the text, in which case you will edit and format it yourself, and you can ~~recreate the source document, including its formatting.~~

The various levels of formatting are: creating body text, retaining the word and paragraph formatting and creating a facsimile copy.

Creating body text means no formatting is applied: you get a continuous, running text. All formatting, if any, is done afterwards by the user.

If you retain the word and paragraph formatting, the font type, size and typeface are maintained across the recognition. The justification of the paragraphs is also detected. However, no graphics are captured and the columns aren't recreated - the paragraphs just follow each other etc.

"Autoformatting" recreates a facsimile copy of the original document: the text blocks, graphics and tables are recreated in the same place and the word and paragraph formatting are maintained across the recognition.

Cell 1A	inmate
Cell 2A	Warden
Cell 3A	\$100,000

As a result, you get a true copy of your source document, be it a compact and editable text file, no longer a scanned image of your document!

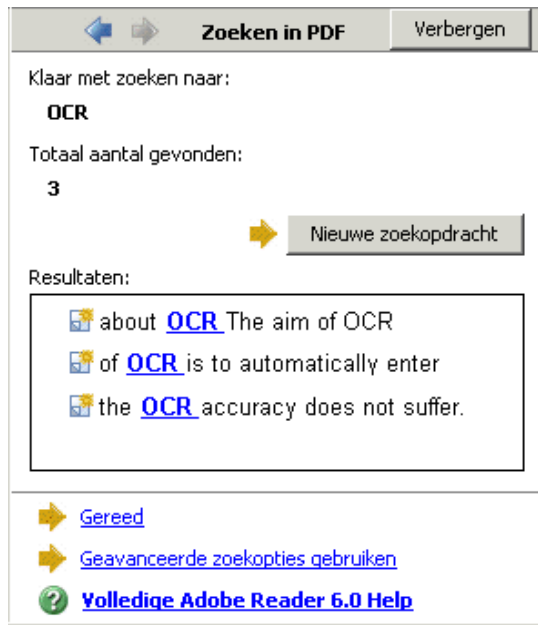
Copyright: Neep Technologies Incorporated 2006

7,03 x 8,71 in

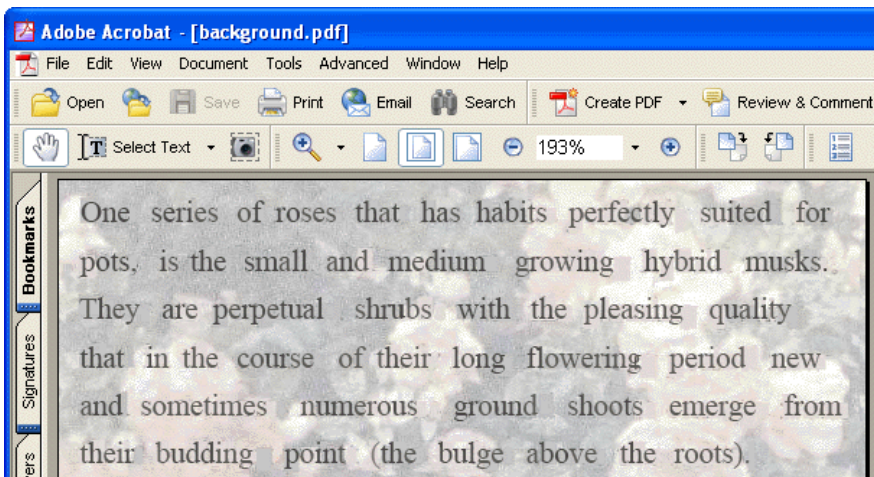
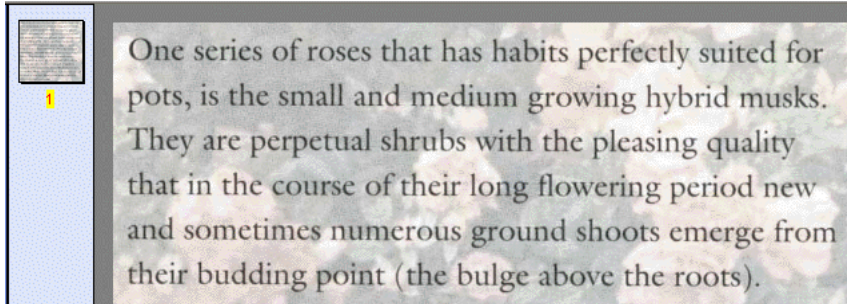
1 van 1

De formaten "PDF Tekst-Beeld" en "PDF Beeld-Tekst" leveren andere resultaten op: Readiris creëert een doorzoekbaar PDF bestand dat de herkende

tekst *en* het beeld van de bladzijde bevat. Bij bestanden van het type “tekst-beeld” is de tekst *boven* het beeld van de bladzijde opgenomen in het PDF bestand dat nu twee lagen bevat. Bij PDF bestanden van het type “beeld-tekst” bevindt de tekst zich *onder* het beeld van de bladzijde. Maak gebruik van de zoekfunctie van de Adobe Reader software en u zult snel het verschil begrijpen!



PDF bestanden van het type “tekst-beeld” zijn overigens vrij gesofistikeerd: de pixels van de herkende tekst worden gewist om een leesbaar document te garanderen! De herkende tekst afbeelden in bijv. het zwart boven zwarte bitmaps van karakters zou tekst opleveren met een zware schaduw... U kunt het alvast proberen met het voorbeeldbestand BACKGROUND.JPG. (Readiris Corporate biedt dezelfde functionaliteit aan voor andere tekstformaten...)

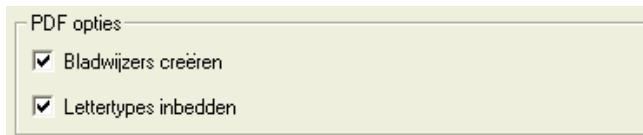


Alle op tekst gebaseerde PDF bestanden nemen de URLs van websites op als zichtbare links: klik op zo'n link en u bezoekt de vermelde webstek!

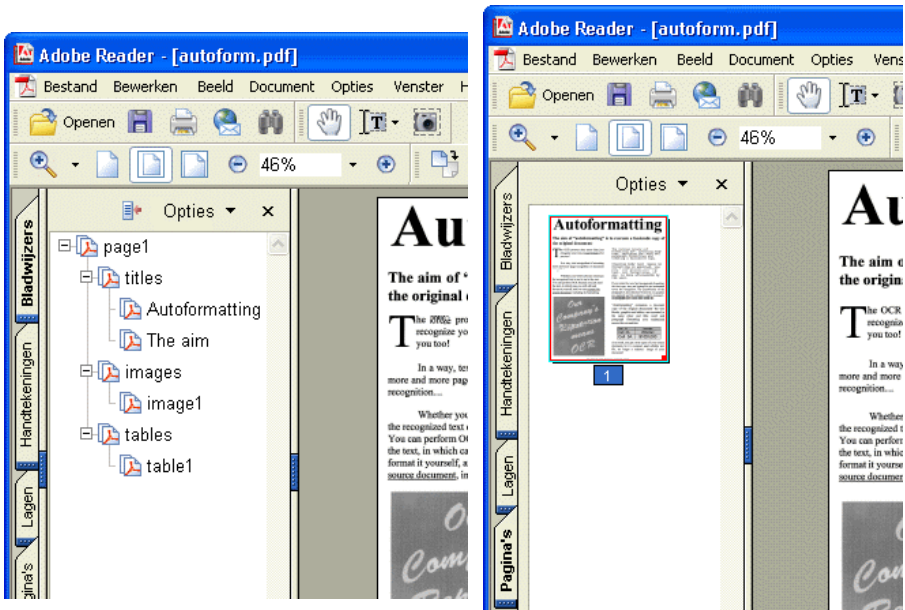




Klik op de knop "Formaat" om enkele opties te ontdekken die het Acrobat PDF formaat betreffen: "Bladwijzers Creëren" en "Lettertypes Inbedden".

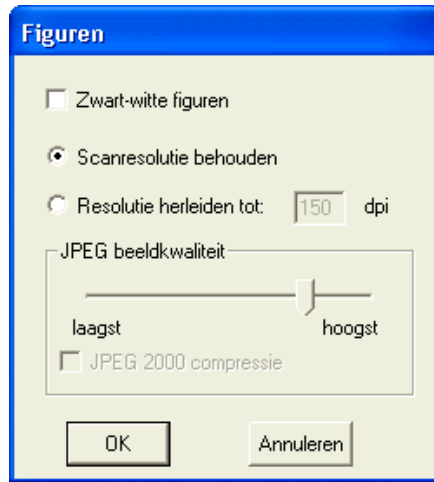


De optie "Bladwijzers Creëren" zorgt ervoor dat er een **bladwijzer** wordt aangemaakt voor elk element van het document - en de figuren zijn evengoed elementen als de tekstblokken en de tabellen. Voor de tekstvensters past Readiris een intelligent algoritme toe dat per venster een titel vindt, zeg maar een samenvatting maakt. De tabellen en figuren worden gewoon genummerd. Een ander navigatiemiddel van PDF documenten, **miniaturen** ("thumbnails"), kunt u dynamisch aanmaken met uw Adobe Reader software!



De optie "Lettertypes Inbedden" sluit de lettertypes in de PDF bestanden in. Lettertypes inbedden belet de vervanging van lettertypes wanneer de lezer het herkende document op het scherm bekijkt en afdrukt. Het zorgt ervoor dat de lezer - welke computer hij ook moge gebruiken - de tekst te zien krijgen met de correcte, oorspronkelijke lettertypes. Lettertypes inbedden verhoogt wel (enigzins) de bestandsgrootte van de herkende documenten!

Een aanvullende optie van PDF bestanden zit verborgen onder de geavanceerde grafische opties in het dialoogvenster "Formaat". Wanneer u PDF bestanden aanmaakt, kunt u een compressiemethode kiezen voor de figuren die in het PDF bestand zijn opgeslagen - JPEG of JPEG 2000. (JPEG 2000 is de nieuwste, meer compacte versie van de JPEG standaard.)



## ... OF ZE LEZEN!

---

We kijken even de andere richting uit. Aangezien Readiris het documentformaat Adobe Acrobat PDF volledig ondersteunt, kunt u dergelijke bestanden niet enkel aanmaken maar ook *inlezen* met Readiris!

**PDF documenten “hieroriënteren”** is één van de belangrijke objectieven van Readiris, en daar bestaan verschillende redenen voor. Het is bijvoorbeeld een manier om beelden in tekst om te zetten: laad op beeld gebaseerde PDF documenten in het geheugen, voer de herkenning uit en sla het herkenningsresultaat op in een tekstdocument (in het tekstformaat van uw keuze). Tekstbestanden zijn natuurlijk editeerbaar en beeldbestanden zijn dat niet.



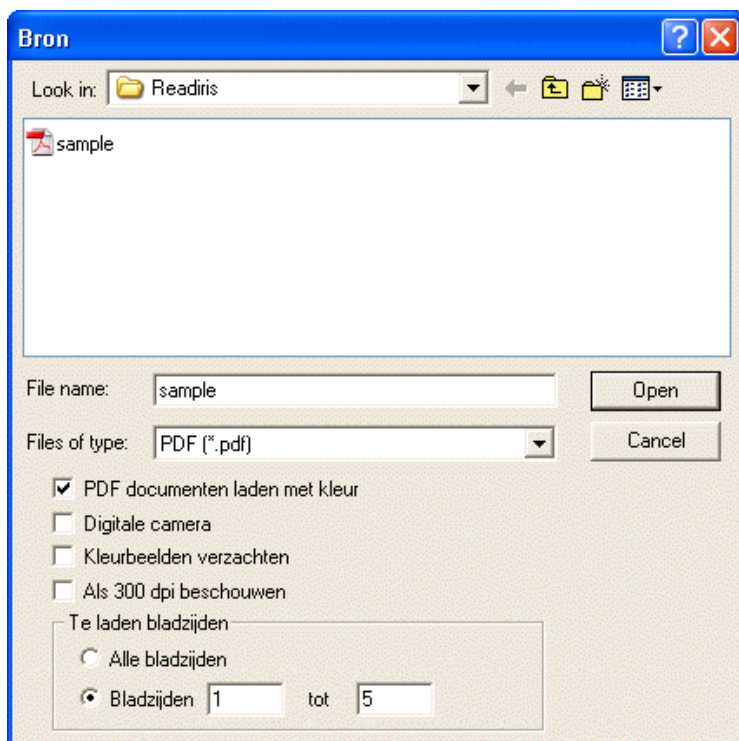
Tweede scenario: u zet op beeld gebaseerde PDF documenten om in PDF documenten die op tekst zijn gebaseerd. U voert dan de herkenning uit op PDF bestanden die enkel beelden bevatten en slaat de OCR resultaten op... in PDF *tekst*documenten! Moeten we het nog herhalen? Op tekst gebaseerde PDF bestanden zijn doorzoekbaar en bewerkbaar, PDF bestanden die enkel beelden bevatten zijn dat niet.

Tot slot is de conversie van PDF documenten een manier om hun inhoud toegankelijk te maken. U kunt namelijk “read-only” PDF bestanden waarvan de inhoud normaal gezien buiten uw bereik valt, herkennen. Bij PDF documenten wiens inhoud niet beschermd is, kan u bij de inhoud komen (om die naar het klembord te kopiëren en in een tekstbestand op te slaan). Bij “read-only” bestanden kunt u geen gebruik maken van de inhoud. Dergelijke documenten kunnen namelijk enkel op het scherm worden bekeken en afgedrukt!

We voegen twee belangrijke nuances toe: Readiris opent geen PDF documenten die met een wachtwoord beschermd zijn, ook al worden alle andere beveiligingen door Readiris gesloopt! (Laten we specifiek zijn: “meesterwachtwoorden” die de toelatingen van PDF bestanden regelen storen Readiris niet, “gebruikerswachtwoorden” die nodig zijn om een PDF bestand te openen doen dat wel.) Ten tweede zet Readiris geen PDF documenten om die figuren met JPEG 2000 compressie bevatten.

Ga te werk zoals gewoonlijk: laad PDF bestanden in het geheugen zoals u vooraf gescande beelden - faxen, snapshots die u met uw digitaal fototoestel nam enz. - inlaadt. Klik op de knop "Stoppen" of druk op de toets Escape om het laadproces tussen twee bladzijden te onderbreken.

Er is één optie die specifiek PDF bestanden betreft: u kunt die openen als kleurdocument of als zwart-wit document. Het nut van deze optie bestaat hierin dat het veel meer tijd kost om kleurdocumenten in bitmaps om te zetten dan het geval is bij zwart-wit documenten!



Bovendien kunt u aangeven welke bladzijden u wenst om te zetten. Als het uw bedoeling is om, zeg maar, één specifiek hoofdstuk te converteren uit een lange PDF publicatie, heeft het weinig zin om het hele boek in Readiris te laden... Geef de om te zetten **bladzijden** aan en u zult heel wat tijd besparen! (U kunt dat ook met meerbladige TIFF bestanden.)



## MEERDERE PAGINA'S HERKENNEN

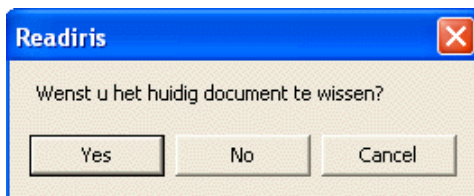
---

Na de herkenning wordt het gescand document opnieuw afgebeeld met zijn bevenstering; het document is beschikbaar voor verdere verwerking.

U kunt nu de herkende tekst openen met uw tekstverwerker of teksteditor, in uw DTP-software of een andere tekstapplicatie invoeren. Ga uw gang en vergelijk het tekstresultaat met het beeld dat door Readiris afgebeeld wordt.

Onze oefening van één bladzijde eindigt hier. Maar hoe slaat u de tekst op van de volgende bladzijden? Met andere woorden: hoe verwerkt u documenten die meerdere pagina's beslaan? Het is eigenlijk heel simpel: ga rustig verder met de herkenning van de volgende bladzijden en sla de resultaten op in hetzelfde tekstbestand. (Wees er dan zeker van dat die tekst op dat ogenblik niet geopend is want zoiets belet u in dat bestand te schrijven!) Vergeet ook niet het letterwoordenboek in de uitbreidingsmodus te plaatsen zodat u de training kunt verderzetten.

Van zodra u bladzijden scant (of beeldbestanden opent) *binnen* een document, moet u beslissen of u met een nieuw document wilt aanvangen of het huidig document wilt aanvullen.

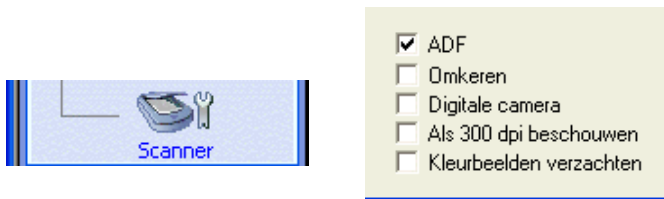


Kies "nee" om bladzijden toe te voegen aan het huidig document, kies "ja" om een nieuw document aan te maken. Dit laatste antwoord heeft hetzelfde effect als wanneer u het bevel "Nieuw Document" onder het menu "Bestand" gebruikt.



Maar er is een veel efficiëntere manier om meerdere bladzijden over te nemen dan die één voor één te herkennen. U kunt namelijk gewoon meteen **meerbladige documenten** verwerken.

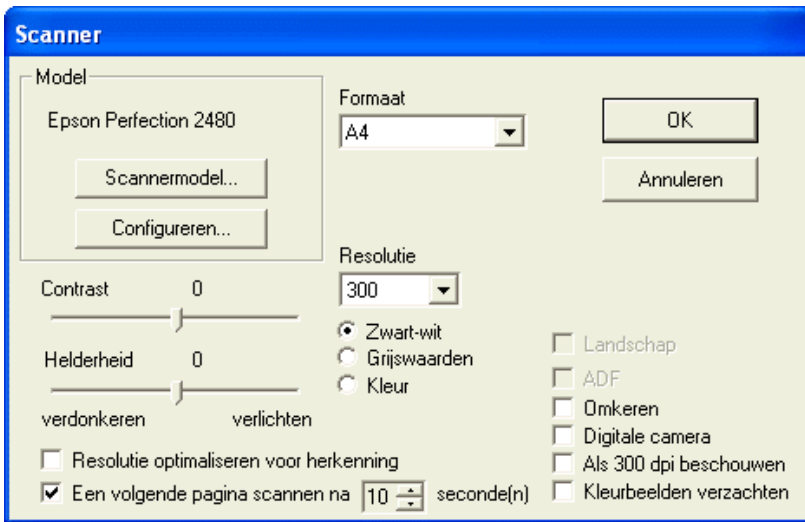
Om een document van meerdere bladzijden te scannen, schakelt u de documentinvoer van uw scanner in met de optie "ADF" onder de "Scanner" knop.



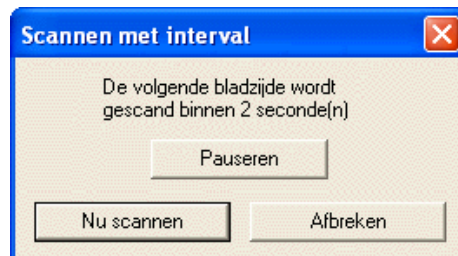
Leg de bladzijden in de documentinvoer en start het scannen: Readiris blijft scannen tot de documentdoorvoer leeg is.

Maak gebruik van **intervalscannen** om meerbladige documenten efficiënt te verwerven wanneer uw vlakbedscanner *niet* is uitgerust met een documentdoorvoer. De scanner scant automatisch een nieuwe bladzijde na een door de gebruiker gekozen aantal seconden; die pauze laat u de tijd om de bladzijde die op het venster van uw scanner ligt te vervangen.

Geef in de scannerinstelling aan welk interval u nodig heeft om een andere bladzijde in uw scanner te plaatsen; klik op de knop "Scanner" en kies een gepaste waarde voor de optie "Een Volgende Bladzijde Scannen na x Seconde(n)".



Klik nu op de knop "Scannen" om het scannen aan te vangen. Klik op "Afbreken" in het dialoogvenster voor intervalscannen om het automatisch scannen stop te zetten.

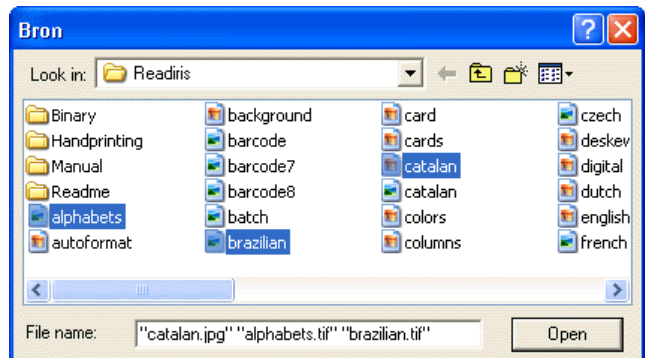


Readiris is overigens erg flexibel: u kunt het scaninterval afbreken om onmiddellijk te scannen en het interval bevriezen om een telefoontje te beantwoorden! Klik op "Pauseren" in het dialoogvenster voor intervalscannen

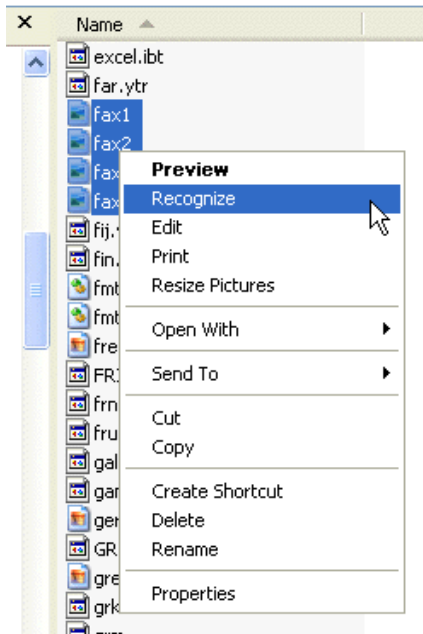


om de scanpauze te bevroeven; klik op "Verdergaan" wanneer u opnieuw klaar bent om verder te gaan. Of klik op "Nu Scannen" in het dialoogvenster om een bladzijde meteen te scannen: de pauze tussen twee bladzijden wordt afgebroken!

U kunt ook meerdere vooraf gescande beeldbestanden *openen*. Om meerdere beeldbestanden in het geheugen te laden, selecteert u het eerste beeldbestand en houdt u de Ctrl toets ingedrukt terwijl u extra beeldbestanden selecteert. Om een ononderbroken reeks van beelden te selecteren, selecteert u het eerste beeldbestand en houdt u de Shift toets ingedrukt terwijl u het laatste bestand selecteert.



U kunt hetzelfde resultaat bereiken vanuit de Windows Verkenner: selecteer meerdere bestanden, klik met de rechter muisknop en kies het bevel "Herkennen" in het "Context" menu. U kunt deze operatie overigens herhalen: alle beeldbestanden die u zo in Readiris invoert vullen het huidig document aan tot u op het bevel "Nieuw Document" klikt!



U kan zelfs meerdere vooraf gescande beelden van de Windows Verkenner naar het Readiris venster *slepen!* En ook hier geldt: alle beeldbestanden die u naar Readiris sleept vullen het huidig document aan tot u het commando "Nieuw Document" aanklikt.

Readiris sorteert de beelden automatisch - beeld 001.tif gaat vooraf aan beeld 002.tif dat vooraf gaat aan beeld 003.tif enz.

Het **documentpaneel** geeft **statistische data** weer en geeft informatie over alle gescande bladzijden - de beeldbron en beeldresolutie, de duur van het scannen en herkennen enz. U kunt het documentpaneel verbergen en tonen met het overeenkomstig bestand onder het menu "Bekijken".



U kunt ook leren welk beeld in het geheugen werd geladen. Indien u een meerbladig bestand in het geheugen heeft geladen, is er natuurlijk slechts één bestand voor alle beelden. En wanneer u documenten *scant*, geeft het documentpaneel aan welk scannermodel u hiervoor gebruikt.

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuimte	Rotatie
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\multipage.tif	1.30	-	300	2000	2388	?	-
<input checked="" type="checkbox"/> 2	C:\Program Files\Readiris\multipage.tif	1.13	-	300	2000	1888	?	-
<input checked="" type="checkbox"/> 3	C:\Program Files\Readiris\multipage.tif	1.06	-	300	1912	2004	?	-

Sleep de cursor in de titelbalk van het documentpaneel om de breedte van een kolom te wijzigen. (U kunt de *volgorde* van de kolommen in het documentpaneel echter niet wijzigen.) En dan is er nog dit: houd uw muiscursor boven een kolom wanneer die niet breed genoeg is om alle data af te beelden: een tooltip geeft de data helemaal weer.

Beeldbron	Scannen
C:\Program Files\...	3.07

Bladzijde	Beeldbron	Scannen
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\dutch.jpg	

U kunt dezelfde informatie afbeelden voor alle bladzijden tegelijk met het commando "Informatie" uit het menu "Bestand". En u kunt deze informatie ook afbeelden voor een specifieke bladzijde door uw muiscursor boven het pictogram van een bladzijde te houden in de **paginaknoppenbalk**. Deze knoppenbalk op de linkerkant van de gebruikersinterface verschijnt van zodra er een bladzijde in het geheugen geladen wordt. Een klik van de rechter muisknop roept de bevelen op die specifiek de bladzijden betreffen.



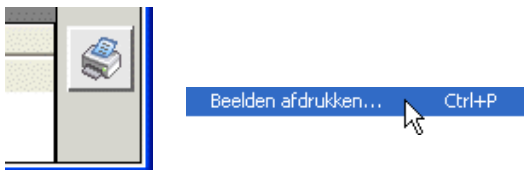
Bladzijde insluiten	ja
Beeldbron	C:\Program Files\Readiris\dutch.jpg
Scannen	4.75
Herkenning	-
Resolutie	300
Breedte	1573
Hoogte	2464
Regelschuine	-
Rotatie	-

- Bladzijde 1 uitsluiten
- Bladzijde 1 selecteren
- Bladzijde 1 wissen
- Bladzijde 1 omhoog verplaatsen
- Bladzijde 1 omlaag verplaatsen

De huidige bladzijde wordt beklemd in het documentpaneel en in de paginaknoppenbalk; ze wordt tevens vermeld in de titelbalk van Readiris.

Om een bladzijde te wissen, selecteert u die in het documentpaneel en drukt u op de Del toets! (Of u selecteert de pagina in de paginaknoppenbalk, u klikt met de rechter muisknop en kiest het bevel "Bladzijde Wissen" van het "Context" menu.)

Mocht u een overzicht willen van uw document, dan kunt u vlug de gescande **beelden afdrukken** met de knop "Afdrukken" op de beeldknoppenbalk (of met het bevel "Beelden Afdrukken" onder het menu "Bestand").

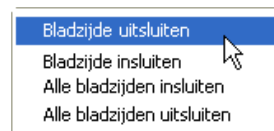


U kunt de huidige bladzijde of alle bladzijden afdrukken. Kies het aantal bladzijden of miniatures dat u op één pagina wenst af te drukken.

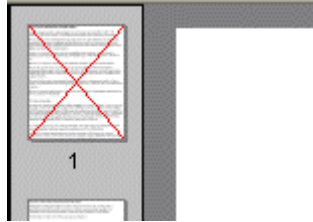


Niet dat u alle bladzijden moet afdrukken: met het documentpaneel (en met de overeenkomstige bevelen in het menu "Bewerken" en de contextuele commando's op de paginaknoppenbalk) kunt u bepaalde bladzijden (tijdelijk) negeren. Klik gewoon op het paginanummer in het documentpaneel om een bladzijde uit te sluiten van het druk- (en OCR-)proces. Klik nogmaals om ze weer in te sluiten. Omwille van de flexibiliteit biedt het menu "Bewerken" nog enkele handige bevelen die op *alle* bladzijden tegelijk slaan.

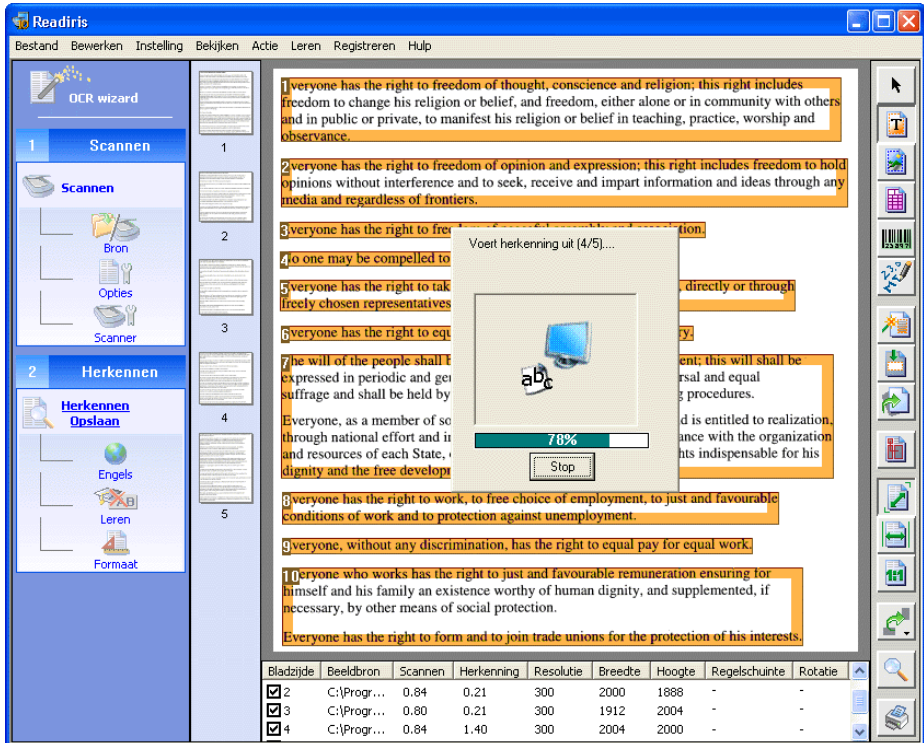
Bladzijde	Beeldbron	Scannen
<input type="checkbox"/> 1	C:\Progr...	1.23
<input checked="" type="checkbox"/> 2	C:\Progr...	0.84



De miniaturen van de uitgesloten bladzijden zijn doorgedaald. Let op: de huidige bladzijde drukken lukt altijd, zelfs al is die pagina op het ogenblik uitgesloten!



Open het beeldbestand MULTIPAGE.TIF en voer de herkenning uit. De bladzijden worden één na één afgebeeld, de titelbalk van Readiris geeft het paginanummer aan.



Als het interactief leren ingeschakeld is, doorloopt u de herkenning en de leerfase voor elke bladzijde. De woordenboekmodus "nieuw" wordt gebruikt voor de eerste bladzijde, de toevoegingsmodus voor de volgende bladzijden.

Door op de knop "Einde" te klikken worden alle oplossingen van het systeem aanvaard zonder validatie door de gebruiker. Met andere woorden, het interactief leren wordt dan afgebroken voor *alle* bladzijden, het OCR-proces gaat verder in de automatische modus.



Het herkenningsresultaat van meerbladige documenten wordt opgeslagen in één enkel uitvoerbestand. (Wanneer u de uitvoer naar een doeltoepassing stuurt, worden meerdere pagina's aangemaakt binnen één document.)

Tenminste, dat is het geval wanneer u de optie "Eén Document Creëren per Bladzijde" uitgeschakeld is op het ogenblik dat u het herkende document opslaat. Deze optie zorgt er namelijk voor dat elk blad van een meerbladig document wordt opgeslagen in een apart bestand. Indien de gebruiker het bestand tekst.doc noemt, zullen de bestanden tekst-1.doc, tekst-2.doc enz. heten. (Deze optie is enkel beschikbaar wanneer u de OCR resultaten in een extern bestand opslaat, niet wanneer u de leesresultaten rechtstreeks naar een doeltoepassing zendt.)

File name:  Save

Save as type:  Cancel

Opslaan en openen

Per e-mail versturen

Eén bestand creëren per bladzijde  
(multipage-1, multipage-2...)

## MEERBLADIGE DOCUMENTEN BEWERKEN

---

De gebruiker kan meerbladige documenten bewerken, hoofdzakelijk om scanfouten te corrigeren: hij kan bladzijden uit het document verwijderen en de volgorde van de pagina's binnen het document wijzigen.

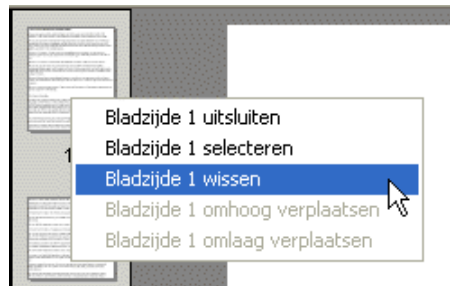
Maar we beginnen met de navigatie. Om *naar een bladzijde te gaan*, klikt u erop in het documentpaneel of in de paginaknoppenbalk. (Of u houdt de muiscursor boven de miniatuur van de bladzijde, u roept het "Context" menu op door met de rechter muisknop te klikken en u gebruikt het commando "Bladzijde Selecteren".) Om naar de vorige bladzijde te verspringen kan u op de toets PageUp duwen, om naar de volgende bladzijde te gaan, op de toets PageDn. Druk op Home op naar



de eerste bladzijde te gaan, druk op End op naar de laatste bladzijde te gaan. Of gebruik de overeenkomstige bevelen uit het menu "Bekijken".

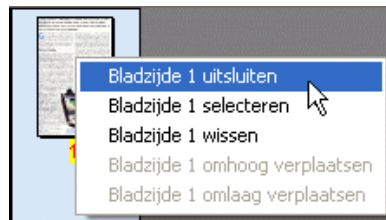
Eerste bladzijde	Home
Vorige bladzijde	PageUp
Volgende bladzijde	PageDown
Laatste bladzijde	End

Het document wijzigen gaat als volgt. Om een *bladzijde* te *wissen*, selecteert u die in het documentpaneel en drukt u op de Del toets. (Of houd de muiscursor boven de miniatuur van de bladzijde, roep de "Context" menu op door met de rechter muisknop te klikken en gebruik het commando "Bladzijde Wissen".



We herinneren u eraan dat u specifieke bladzijden tijdelijk kunt uitsluiten van het herkenningsproces (en van het afdrukken van de beelden) - zonder ze te wissen: het documentpaneel, de paginaknoppenbalk en het menu "Bewerken" bevatten de nodige bevelen.

Bladzijde	Beeldbron	Scannen
<input type="checkbox"/> 1	C:\Progr...	1.23
<input checked="" type="checkbox"/> 2	C:\Progr...	0.84





Om een *bladzijde naar boven te schuiven*, een hoger paginanummer te geven, gebruikt u het commando "Bladzijde Omhoog Verplaatsen", om een *bladzijde naar benedente verschuiven*, gebruikt u het commando "Bladzijde Omlaag Verplaatsen".

Om een *bladzijde te verhuizen*, namelijk ze een heel andere plek te geven in het document, sleept u haar pictogram naar een andere locatie.

The screenshot shows the Readiris software interface. The title bar reads "Readiris". The menu bar includes "Bestand", "Bewerken", "Instelling", "Bekijken", "Actie", "Leren", "Registreren", and "Hulp".

The main window is divided into several sections:

- OCR wizard**: Located at the top left, featuring a magic wand icon.
- 1 Scannen**: A section with a "Scannen" button and a "Scanner" icon. Below it are options for "Bron" (Source) and "Opties" (Options).
- 2 Herkennen**: A section with a "Herkennen Opslaan" button and a magnifying glass icon.

On the right side, there is a preview of the scanned document with four numbered sections (1, 2, 3, 4) corresponding to the steps in the wizard. The text in the preview includes:

- Section 1: national or social origin, property. Furthermore, no distinction shall be made on the basis of the international status of the country of origin, birth, nationality, race, sex, language, religion, political or other opinions, wealth, independent, trust, non-self-governing territories, etc.
- Section 2: Everyone has the right to life, liberty and security of person.
- Section 3: No one shall be held in slavery or servitude; no one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment. No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment. No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.
- Section 4: Everyone has the right to recognition as a person before the law. All are equal before the law and are entitled to equal protection of the law. All are entitled to equal protection of the law. All are entitled to equal protection of the law. Declaration and against any incitement to such discrimination.

A mouse cursor is shown hovering over the "3" section in the preview, and a small rectangular box highlights the "3" section in the wizard's "Herkennen" step.

## MET EEN NIEUW DOCUMENT BEGINNEN

---

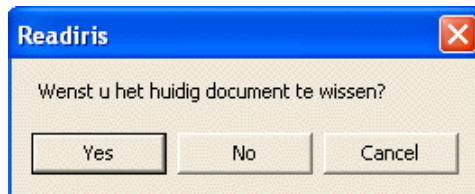
Met het bevel "Nieuw Document" onder het menu "Bestand" sluit u het huidige document af.



Dit commando wist alles uit het geheugen. Het in het geheugen geladen document (dat één of meerdere bladzijden bevat) wordt gewist. U bent nu klaar om een nieuw document aan te maken.

Maar u kunt ook met een nieuw document beginnen vanuit het huidige document. Zolang de herkenning niet werd uitgevoerd, veronderstelt het systeem dat u nog bladzijden aan het document wilt toevoegen. U kunt bijvoorbeeld alle bladzijden in de documentinvoer inscannen, de documentdoorvoer bijvullen en opnieuw beginnen. Of u scant enkele documenten en voegt daar enkele beeldbestanden, zeg maar faxen, aan toe. Deze bladzijden vormen één document, het enige wat u moet doen is tussendoor de beeldbron wijzigen met de knop "Bron".

Wanneer de herkenning wel al was uitgevoerd en u begint opnieuw te scannen (of beelden te laden), zal het systeem u vragen of u met een nieuw document wilt beginnen of met het huidige document wilt verdergaan.



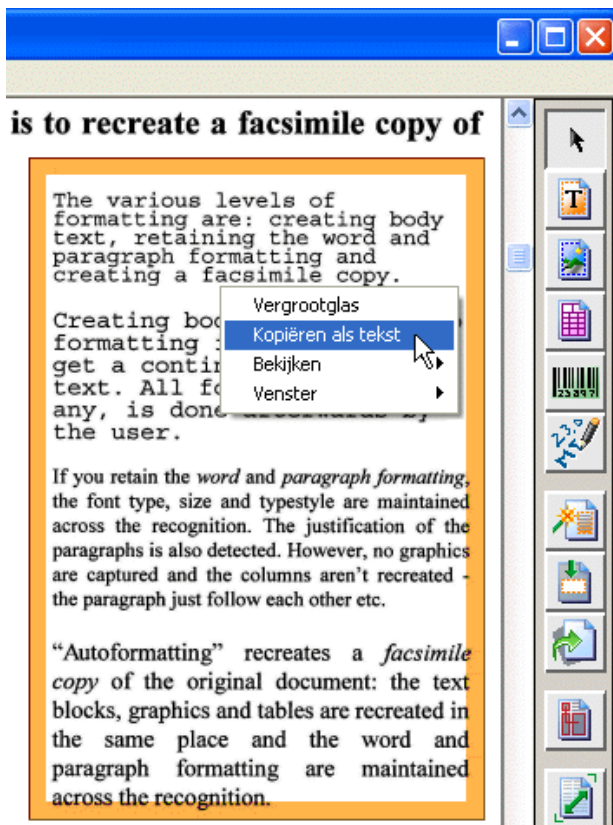
## TEKSTZONES HERKENNEN

---

We weten nu hoe men een bladzijde herkent en hoe men meerbladige documenten verwerkt. Maar kunnen we ook makkelijk delen van een tekstbladzijde



herkennen? Dat kan: klik met de rechter muisknop en selecteer het bevel "Kopiëren Als Tekst" in het "Context" menu: het tekstvenster onder de muiscursor wordt herkend en het tekstresultaat wordt in het klembord gestopt.



De huidige systeeminstelling - taal, type letterteken enz. - is van toepassing. Het OCR-resultaat wordt naar het klembord gestuurd als lopende, niet geformatteerde tekst.

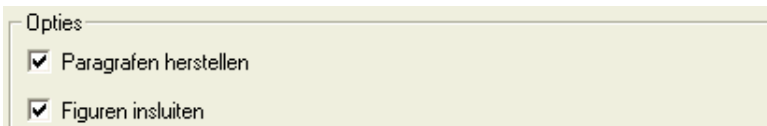
## DE UITVOER OPMAKEN

---

Een tekst opslaan of uitvoeren houdt meer in dan een uitvoermethode kiezen of een bestandsnaam opgeven voor de tekstuitvoer. U moet ook een tekstformaat selecteren en de lay-out van de herkende tekst gaan bepalen. Kortom, u moet beslissen waar u met de tekst heen wilt eer u de herkenning uitvoert.

Sommige opties van de "Formaat" knop laten u toe om het uitzicht van de uitgevoerde tekst te beïnvloeden.

De "tekstflow" van de uitvoer wordt onmiddellijk beïnvloed door de optie "Paragrafen Herstellen".



Schakel deze optie in om Readiris de **paragrafen** te laten **detecteren**: Readiris zal dan de normale "wordwrap" toepassen die typisch is voor tekstverwerkers, zoniet wordt een carriage return (Enter) toegevoegd na iedere lijn en blijven gesplitste woorden gesplitst! De paragraafdetectie is standaard actief.

Laten we een voorbeeld geven om één en ander te verduidelijken. Wanneer de eerste drie lijnen van een kolom "De nieuwe presi-", "dent wuift van op het terras." en "Zijn vrouw vergezelt hem." zijn, zorgt de paragraafdetectie voor het volgende resultaat: "De nieuwe **president** wuift van op het terras. **Zijn** vrouw vergezelt hem." De twee stukken van het woord "president" die door een koppelteken gesplitst zijn, zijn opnieuw aaneengelijmd, en er werd een spatie ingevoegd op het einde van de eerste zin. Kortom, u krijgt een vlotte, continue tekst.



Was de paragraafdetectie *niet* geactiveerd, dan zou de originele lay-out behouden zijn, met een carriage return (Enter) aan het einde van elke zin.

Deze optie is *niet* beschikbaar wanneer u een PDF formaat kiest: Adobe Acrobat bestanden slaan tekst altijd lijn per lijn op.

(De "Formaat" knop bevat tevens formatteeropties die we nog niet besproken hebben - we zullen dit straks doen.)

## UW SCANNER INSTELLEN

---

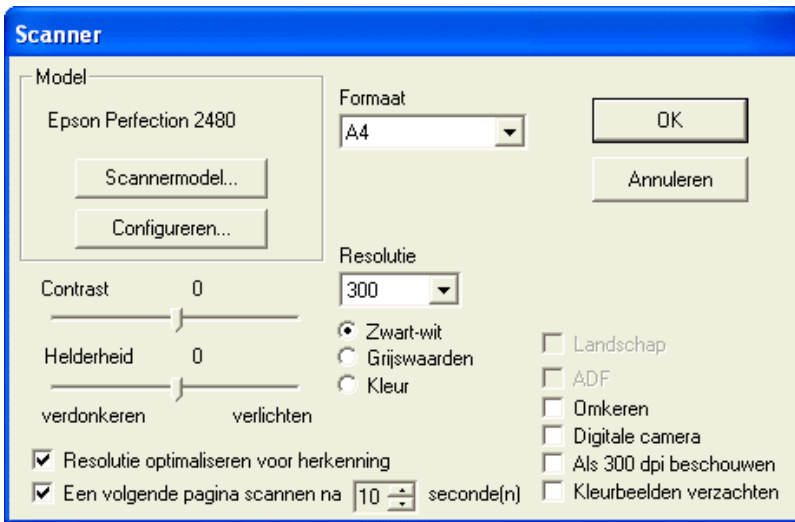
Laten we uw scanner gaan instellen. We veronderstellen nu wel dat uw scanner en de nodige stuurbestanden goed werden geïnstalleerd.

Als uw Readiris licentie gebundeld is met een scanner, kan deze stap overbodig blijken omdat uw scanner dan allicht al onder Readiris geconfigureerd is.

Klik op de "Scanner" knop in de hoofdknoppenbalk.

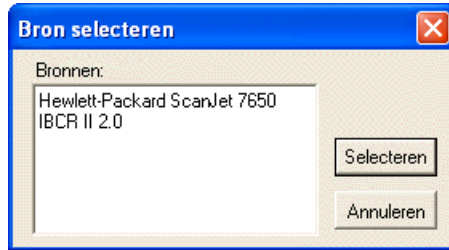


Klik op de knop "Scannermodel" om uw **scannermodel** te kiezen.



Wanneer u de optie "<Image>" (beeld) als "scanner" selecteert, fungeren vooraf gescande beelden ten allen tijde als beeldbron – u hoeft zelfs de schijf niet te selecteren met de knop "Bron" op de hoofdknoppenbalk.

De knop "Configureren" is enkel beschikbaar wanneer uw scanner het toelaat. Hij geeft toegang tot sommige geavanceerde scanparameters; bij Twain scanners kunt u deze knop gebruiken om de Twain bron te selecteren. (U kunt dat ook met het commando "Bron Selecteren" onder het menu "Bestand".)



Wanneer de scanner geselecteerd is, zal hetzelfde venster u toelaten om de scanresolutie, het papierformaat en -oriëntatie, de helderheid en het contrast in te stellen. U kunt tevens aangeven of u van de documentinvoer wilt gebruik maken. Bij Twain scanners worden alle instellingen meestal binnen de Twain interface geregeld.

Stel de **helderheid** in en, wanneer dat mogelijk is, het **contrast**.

Door de optie "Landschap" te selecteren geeft u aan dat de pagina-oriëntatie liggend is en niet staand ("portret"). Deze pagina-oriëntatie is eigenlijk enkel van toepassing op gereduceerde formaten: bij een A4 flatbed scanner kunt u bijvoorbeeld A5 bladzijden (de helft van een A4 blad) in portret- en in landschaporiëntatie scannen, maar volle A4 bladzijden kunt u natuurlijk alleen in één richting inscannen!



De optie "Omkeren" tenslotte laat de scanner volledig "**negatieve**" beelden laten aanmaken in de zwart-wit scanmodus - activeer deze optie om bladzijden te verwerken met witte tekst op een zwarte achtergrond.

Het automatisch scannen met een interval hebben we al besproken. Dit is een erg efficiënte manier om meerbladige documenten te verwerven wanneer uw vlakbedscanner niet is uitgerust met een documentdoorvoer.



## **GEEF UW SCANS WAT KLEUR MEE!**

---

Readiris ondersteunt zwart-witte scans, scans in grijswaarden en in kleur. U kan dus de **kleurmodus** kiezen die u het best bevalt. Om zwart-wit illustraties in de documenten in te sluiten, scant u in zwart-wit, om zwart-wit foto's over te nemen, scant u in grijswaarden en om kleurfoto's over te nemen, scant u in kleur.

Maar waarom zou u de bitdiepte van de beelden bij het scannen herleiden? Het spreekt voor zich dat het scannen van grijswaarden en kleur trager is en meer RAM-geheugen vereist dan zwart-wit beelden.

Grijswaarden en kleur gebruiken zijn niet enkel nuttig om de figuren met voldoende kwaliteit te hernemen; in een aantal gevallen hebt u grijswaarden of kleuren gewoon nodig voor een goede herkenning! Bij teksten die afgedrukt zijn op een gekleurde achtergrond kan een kleurenscaan de kleurverschillen creëren die in het zwart-wit beeld afwezig zijn. Wanneer er slechts weinig contrast is tussen de tekst en de achtergrond, kan de achtergrond "ruis" opleveren die de herkenning bemoeilijkt of gewoon onmogelijk maakt!

Nemen we bijvoorbeeld een zwarte tekst op een donkere achtergrond. Hoezeer u de helderheid ook bijstelt, wanneer u zo'n document in zwart-wit scant zult u er allicht niet in slagen de achtergrond te filteren zonder dat de tekst eveneens wegvalt.



**MASAYOSHI SON**, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet companies in the U.S., Japan, Europe, and other Asian countries. Today, Softbank manages about \$4 billion in venture capital funds for global investments.

**YASUMITSU SHIGETA**, 35, has invested in more than 70 Web or mobile Net-based ventures in Japan and the U.S., including Tumbleweed Communications and Phone.com. Shigeta is also developing new businesses that take advantage of the growth of the Internet and mobile communications.

**MASAYOSHI SON**, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet companies in the U.S., Japan, Europe, and other Asian countries. Today, Softbank manages about \$4 billion in venture capital funds for global investments.

**YASUMITSU SHIGETA**, 35, has invested in more than 70 Web or mobile Net-based ventures in Japan and the U.S., including Tumbleweed Communications and Phone.com. Shigeta is also developing new businesses that take advantage of the growth of the Internet and mobile communications.

Readiris creëert een zwart-witte kopie van elk kleur- en grijswaardenbeeld. De software maakt daarbij gebruik van intelligente routines, zodat zelfs de moeilijke gevallen worden opgelost. Hieronder ziet u hoe ons “lastig” beeld wordt omgezet...

**MASAYOSHI SON**, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet companies in the U.S., Japan, Europe, and other Asian countries. Today, Softbank manages about \$4 billion in venture capital funds for global investments.

**YASUMITSU SHIGETA**, 35, has invested in more than 70 Web or mobile Net-based ventures in Japan and the U.S., including Tumbleweed Communications and Phone.com. Shigeta is also developing new businesses that take advantage of the growth of the Internet and mobile communications.

Om de zwart-witte versie van een gescand beeld af te beelden, schakelt u de optie "Document in Kleur Afbeelden" onder het menu "Bekijken" uit.

✓ Document in kleur afbeelden Ctrl+O



Maar eigenlijk krijgt u geen zwart-wit beelden op het scherm te zien - zelfs niet wanneer u inderdaad zwart-witbeelden scant! Dat komt omdat Readiris de beelden optimaliseert voor een uitstekende leesbaarheid. I.R.I.S.' gespecialiseerde hoge-resolutie displaytechniek zet zwart-wit beelden om in grijswaardenbeelden.

### **Reading dot matrix documents**

You can read dot matrix document without changing the font mode. The software detects whether "normal" text or dot matrix printouts are being read.

Far out in the uncharted backwaters of the unfashionable end of the Western Spiral arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-two million miles is an utterly insignificant little blue green

### **Reading dot matrix documents**

You can read dot matrix document without changing the font mode. The software detects whether "normal" text or dot matrix printouts are being read.

Far out in the uncharted backwaters of the unfashionable end of the Western Spiral arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-two million miles is an utterly insignificant little blue green

Grijswaarden- en kleurbeelden daarentegen worden zachter gemaakt.

### **A word about OCR**

The aim of OCR is to automatically enter printed text documents in a very effective and low cost way. Although the first research and development on Optical Character Recognition (OCR) began more than 30 years ago, this technology is still unknown by most of the people who could use it for their document entry applications.



## A word about OCR

The aim of OCR is to automatically enter printed text documents in a very effective and low cost way. Although the first research and development on Optical Character Recognition (OCR) began more than 30 years ago, this technology is still unknown by most of the people who could use it for their document entry applications.

Dit alles zorgt ervoor dat u niet hoeft in te zoomen, zelfs op een draagbare PC met een LCD scherm of op een desktop-PC met een bescheiden 15" scherm.

Zoom in op ware grootte (of hoger) om het "ruwe" beeld te bekijken zoals het werd gescand.

## ANDERE APPARATEN, ANDERE RESOLUTIE

Wat uw kleurmodus ook moge wezen, kies een **resolutie** van 300 dpi voor normale toepassingen en de hogere resolutie van 400 dpi voor kleine druk (kleiner dan 10 punt) en wanneer het document erg gedegradeerd is.

Readiris leest **puntgroottes** van 6 tot 72 punt (0,21 tot 2,54 cm.).

6 punt

72

Readiris leest eveneens **kapitalen** ("dropletters"), hoofdletters die meerdere lijnen beslaan. Zelfs negatieve kapitalen worden herkend... (Die kunnen uiteraard niet groter zijn dan 72 punt!)

**R**eadiris leest "drop letters" die meerdere lijnen bedekken en kent die toe aan de eerste lijn van de paragraaf.

**L**e Festival de Wallonie est dans sa phase ultime et nale avec l'ouverture de ses dernières sections.

Cette semaine, c'est la branche

Daar optimale OCR een resolutie vereist tussen 300 dpi en 400 dpi, waarschuwt Readiris u wanneer u beelden poogt te herkennen met een resolutie die lager ligt dan 200 dpi of hoger dan 800 dpi. Noteer dat de beeldresolutie van dergelijke beelden rood is gemarkeerd in het documentpaneel.

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte
<input checked="" type="checkbox"/>	C:\Documents .. 1.24	-	-	96	400	447

Readiris kan scans met teveel details automatisch corrigeren! Activeer hiervoor de optie "Resolutie Optimaliseren voor Herkenning" in de scannerinstelling. Van zodra de beeldresolutie van uw scans 600 dpi overschrijdt, wordt hun resolutie automatisch verminderd voor de herkenning.

Resolutie optimaliseren voor herkenning

Er zijn andere manieren om deze waarschuwing te vermijden: u leest bijvoorbeeld **faxen** - die hebben een resolutie van 100 of 200 dpi -, u herkent beelden die afkomstig zijn van een digitale fotoestel - de resolutie is dan onbekend -, of u opent beeldbestanden waarvan de hoofding een verkeerde resolutie aangeeft. Om zulke beelden succesvol te verwerken activeert u de optie "Als 300 dpi Beschouwen". Deze parameter geldt evengoed voor rechtstreekse scans als voor vooraf gescande beelden die u in het geheugen laadt.

ADF  
 Omkeren  
 Digitale camera  
 Als 300 dpi beschouwen  
 Kleurbeelden verzachten

PDF documenten laden met kleur  
 Digitale camera  
 Kleurbeelden verzachten  
 Als 300 dpi beschouwen

Wanneer de beelden werden aangemaakt met een **digitale camera** in plaats van een scanner, moet u een speciale optie inschakelen. Ook deze optie is van toepassing op rechtstreekse scans en beelden die u opent.



<input type="checkbox"/> ADF	<input type="checkbox"/> PDF documenten laden met kleur
<input type="checkbox"/> Omkeren	<input checked="" type="checkbox"/> Digitale camera
<input checked="" type="checkbox"/> Digitale camera	<input type="checkbox"/> Kleurbeelden verzachten
<input checked="" type="checkbox"/> Als 300 dpi beschouwen	<input checked="" type="checkbox"/> Als 300 dpi beschouwen
<input type="checkbox"/> Kleurbeelden verzachten	

Door dit te doen verbetert u de kwaliteit van het beeld eer de herkenning plaatsvindt. Beelden die afkomstig zijn van digitale fotoestellen stellen OCR-programma's op proef: digitale camera's produceren beelden met een lage resolutie - zelfs wanneer de u camera erg dicht boven uw document houdt - en de beeldresolutie is bovendien onbekend.

Er zijn enkele "details" waarvan u hoogte moet nemen om beelden die met een digitale camera werden genomen succesvol te herkennen.

Selecteer eerst en vooral de hoogst mogelijke resolutie. Schiet bijvoorbeeld plaatjes van 2.048 x 1.536 pixels wanneer u ook beelden van 1.024 x 768 pixels en van 640 x 480 pixels kunt schieten. Activeer tevens de "macromodus" van uw camera zodat u closeups neemt. Documenten fotograferen betekent immers altijd dat u closeups neemt. (Deze modus werd eigenlijk ontworpen om bloemen, insecten enz. te fotograferen.) Als u dat niet doet zijn de beelden zo onscherp dat ze onleesbaar worden.

**Police seek cyber**

Cyberfraud and electronic extortion is becoming a real problem for UK businesses, and the law

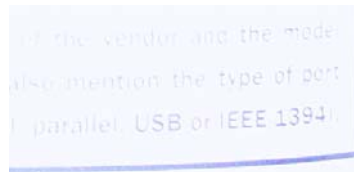
**Microsoft antitrust**

Microsoft seems to have a rosier future than expected after the court appeal against its antitrust case has been heard. The split of the software giant now seems to be in doubt as the US government's case against it came in for serious questioning.

Pas weinig of geen compressie toe: significante compressie herleidt de scherpte van de genomen foto. Zoom handmatig om uw document te omkaderen; sommige camera's zijn uitgerust met software die de aparte stukken van foto's aan elkaar "naait" ("photo stitching"), maar het loont niet de moeite dit te doen bij OCR-toepassingen.

Houd de camera recht boven het document, scan het document niet onder een hoek. Maar probeer wel om met uw handen geen schaduw op het document te werpen! Maak stabiele beelden aan. U kunt de camera desnoods op een statief plaatsen.

Schakel het flitslicht uit wanneer u glanzend papier fotografeert, zoniet kan het beeld veel te licht uitvallen. Door de bank genomen kunt de helderheid en het contrast van een digitale camera aan de omgeving aanpassen - er zijn vaak aparte instellingen voor daglicht, kunstlicht, neonlicht enz. (Sommige camera's kunt u zelfs calibreren door een wit document te fotograferen!)



Als u het even wilt proberen, kunt u het beeldbestand DIGITAL.JPG in de Readiris map openen en de herkenning uitvoeren.



Readiris

Bestand Bewerken Instelling Bekijken Actie Leren Registreren Hulp

OCR wizard

1 Scannen

Scannen

Bron

Opties

Scanner

2 Herkennen

Herkennen Opslaan

Engels

Leren

Formaat

General Information

COMDEX Center for International Commerce

On the Lower Level (Exhibit Hall G) of the Sands Expo and Convention Center, the newly renamed Center for International Commerce is supported by the U.S. Department of Commerce. The CIC offers many free services to international delegates and U.S. exporters, including daily seminars expertly presented by DOC specialists and international trade consulting.

Show Management Office and Information

Information concerning any aspect of the Conference and Exposition may be obtained at the Show Office located in Room 305 of the Sands Exposition and Conference Center and the registration tent of the Las Vegas Convention Center.

Bladzijde	Beeldbron	Scannen	Herkennen	Resolutie	Breedte	Hoogte	Regelschuimte	Rotatie
<input checked="" type="checkbox"/> 1	C:\Program Fil...	1.05	-	300	1519	635	-	-

## STANDAARD INSTELLING BEWAREN

Stel de scanparameters goed in en klik op het commando "Standaard Instelling Opslaan" onder het menu "Bestand" om de actuele instellingen te bewaren als standaard **instelling** voor verder gebruik.





Standaard instelling opslaan

Configuratiebestanden bevatten meer dan enkel de scanopties: ze bepalen ook of u het interactief leren gaat gebruiken, welke taal de documenten bevatten, welke uitvoermethode en lay-outopties u gebruikt - bijvoorbeeld lopende tekst naar WordPad sturen enz. Kortom, *alle* operationele parameters van Readiris worden in de instellingen bewaard.

## SPECIFIEKE INSTELLINGEN BEWAREN

---

De standaard instelling zal natuurlijk gebruikt worden telkens u Readiris opstart, maar u kunt specifieke configuraties bewaren om te vermijden dat u operationele parameters moet herdefiniëren. De opdrachten "Instelling Opslaan" en "Instelling Laden" onder het "Bestand" menu nemen deze taak voor hun rekening.



Instelling laden...

Instelling opslaan...

Standaard instelling opslaan

Laten we een voorbeeld geven: als u regelmatig Duitse documenten moet herkennen met een specifieke lay-out, dan wordt u aangeraden om een configuratiebestand aan te maken voor dit documenttype. U zult dan met name "Duits" als documenttaal selecteren, een aangepaste vensterlay-out laden om te vermijden dat u steeds dezelfde vensters moet hertekenen, het leren uitschakelen maar een woordenboek laden in de leesmodus omdat telkens dezelfde lettertypes terugkeren enz.

Als u twijfels heeft over de huidige instelling, hoeft u niet in elk menu en elk commando te duiken om ze te achterhalen. U kunt gewoon het commando "Informatie" onder het "Bestand" menu aanklikken om een praktisch overzicht te krijgen.



**Instelling**

Scanner		OK					
Model	HP Scanjet 7650	Formaat	Microsoft Word 97 / 2000 / 2002 / 2003				
Resolutie	300 dpi	Paragraaf	Aan				
Formaat	A4	Lay-out	Brondocument hercreëren				
Mode	Zwart-wit						
Liggend	Uit						
Document							
Lettertekens	Automatisch	Taal	Nederlands				
Bladzijde							
Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuinte
1	C:\Progr...	1.05	-	300	1519	635	-

Dit bevel geeft tevens de informatie die u op het documentpaneel vindt weer voor alle bladzijden.

## DOCUMENTEN SCANNEN

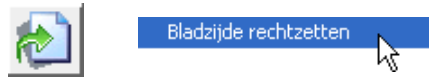
Nu de scanner is ingesteld, zijn we klaar om te gaan scannen. Wat het scannen betreft, moeten we wel op enkele punten wijzen.

Besteed enige aandacht aan de regelkromming. De pagina-analyse en de herkenning springen wel tolerant om met schuine regels, maar het kan fout aflopen wanneer het document te schuin werd ingescand. Beperkte regelkromming (die kleiner is dan  $0,5^\circ$ ) kunt u negeren omdat die geen invloed heeft op de herkenning.

De optie "Bladzijden Rechtzetten" onder de knop "Opties" (en onder het menu "Instelling") bepaalt of die documenten die schuin werden ingescand automatisch zullen **rechtgezet** worden. Beperkte regelkromming wordt daarbij genegeerd; deze optie is standaard uitgeschakeld.

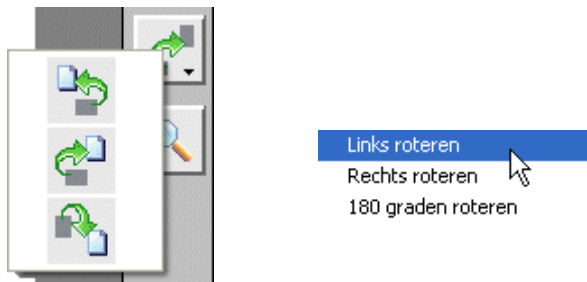


Indien u vergat deze optie in te schakelen, gebruikt u de knop "Bladzijde Rechtzetten" op de beeldknoppenbalk (of het overeenkomstig commando onder het menu "Actie") om de regelschuimte te corrigeren.



Een blad rechtzetten duurt enkele seconden. Eerst wordt het beeld geanalyseerd om de hoek van de regelschuimte te detecteren - als die er is tenminste -, daarna wordt het kleur- of grijswaardenbeeld én de zwart-witte versie rechtgezet en wordt de pagina-analyse opnieuw uitgevoerd.

Het kan ook nodig zijn dat u de pagina-oriëntatie moet aanpassen. Maak hiertoe gebruik van de rotatieknoppen op de beeldknoppenbalk. (Overeenkomstige opdrachten bevinden zich in het menu "Bekijken".) Er zijn drie rotatierichtingen: rotatie naar links, naar rechts en met 180°. De **rotatie** neemt een paar seconden in beslag: het achterliggende beeld wordt herschreven, en niet enkel de afbeelding ervan op het scherm.



Merk overigens op dat Readiris bladzijden met een verkeerde oriëntatie voor u kan rechtzetten. Schakel de optie "Oriëntatie Bladzijden Detecteren" in met de knop "Opties" (of onder het menu "Instelling") en Readiris zal wanneer dat nodig is de richting van de bladzijde corrigeren.



✓ Oriëntatie bladzijden detecteren



U kunt dit alles eens proberen met het beeldbestand DESKEW.JPG in de Readiris map. Schakel de opties "Bladzijden Rechtzetten" en "Oriëntatie Bladzijden Detecteren" in eer u het beeld laadt en Readiris zal de toren van Pisa “herstellen” zoals we die graag zien.

Readiris

Bestand Bewerken Instelling Bekijken Actie Leren Registreren Hulp

OCR wizard

1 Scannen

Scannen

Bron

Opties

Scanner

2 Herkennen

Herkennen Opslaan

Engels

Leren

Formaat

Detecting the page orientation

Under the option "Detect Page Orientation" and correct the page orientation for your OCR use the "examples" to have Readiris detect any skew angles and compensate for them. Try some attention to landscape. Although the page angles and compensations are skew-angle, it may become difficult to witness and OCR a page correctly when the skew is too significant.

Landscape because the OCR accuracy does not matter.

Readiris will recognize images rotated with a landscape automatically when the system "Page Orientation" is enabled by default. You can also set the desktop button on the image under the command "Detect Page Orientation" gets recognized after the following.

Copyright 2012

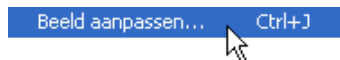
Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuine	Rotatie
<input checked="" type="checkbox"/> 1	C:\Prog...	3,79	-	300	1434	3348	?	-

Noteer ook dat het documentpaneel aangeeft welke regelschuinte werd gecorrigeerd en welke rotatie werd uitgevoerd!

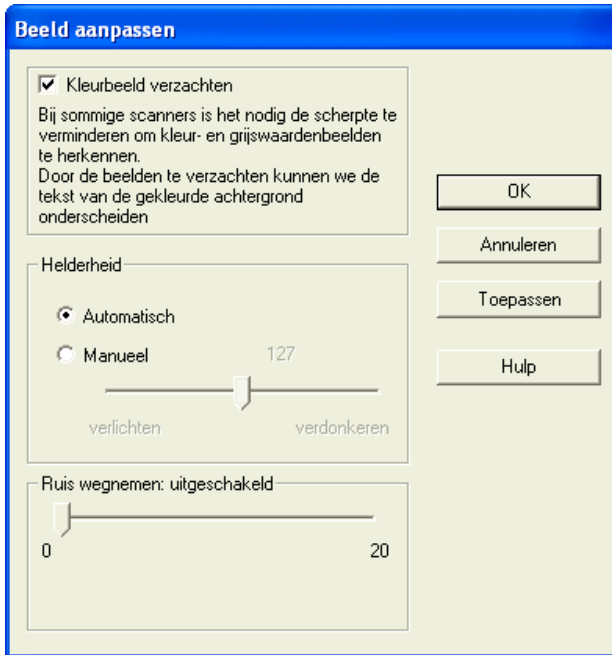
## **DE GESCANDE BEELDEN AANPASSEN**

---

Zoals al aangestipt werd, worden krachtige, intelligente routines gebruikt om kleur- en grijswaardenbeelden automatisch om te zetten in zwart-wit beelden. Maar de gebruiker kan, mocht dit nodig blijken, het beeld verder optimaliseren voor het OCR-proces. Dit gebeurt met het bevel "Beeld Aanpassen" onder het menu "Actie".

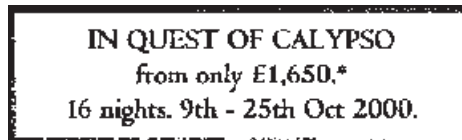
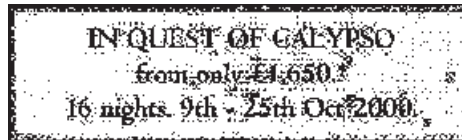
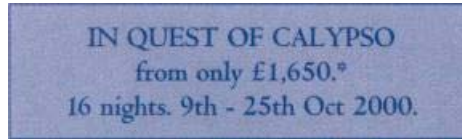


Wanneer u dit commando aanklikt, wordt het zwart-wit beeld automatisch op het scherm gebracht, het is alsof u de optie "Document in Kleur Afbeelden" uitschakelde. Dit bevel hanteert enkele moeilijke begrippen, en we zullen er in detail op ingaan.

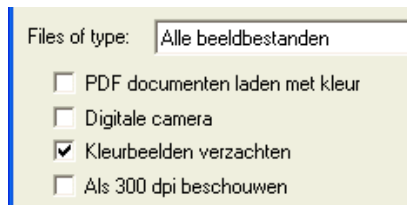


De optie "Kleurbeeld Verzachten" maakt de kleur- en grijsbeelden homogener door ze "af te vlakken", door beperkte verschillen in intensiteit te weg te nemen. Het resultaat van deze bewerking is dat er een groter contrast wordt gecreëerd tussen de voorgrond - de tekst - en de achtergrond - een steunkleur, foto enz.

Deze optie **optimaliseert** het beeld voor de herkenning. Deze functie ziet er misschien erg technisch uit en moeilijk om te begrijpen, maar ze heeft wel degelijk haar belang: bij sommige scannermodellen moeten we de scherpte verminderen om kleur- en grijswaardenscans succesvol te herkennen. De beelden verzachten is soms de enige manier om tekst te onderscheiden van de gekleurde achtergrond! Hieronder ziet u een voorbeeld dat zonder verzachting gewoon onleesbaar is.



Beelden verzachten kan ook wanneer u voorafgescande beeldbestanden in het geheugen laadt.



En wat bedoelen we met de **helderheid**? De helderheid bepaalt de algemene klaarte van een beeld: wanneer u het beeld helderder maakt of verduistert slaat dat op alle pixels. Het komt erop neer dat u de achtergrond van de bladzijde moet wegfilteren. Laten we twee voorbeelden geven. In het eerste voorbeeld zijn alle delen van het beeld donker. We verlichten dus het beeld om de achtergrond van de bladzijde weg te halen. De voorgrond - de tekst - blijft hierbij voldoende donker om door de binarisering te worden gedetecteerd. Voorbeeld 2: het beeld is zo licht dat zelfs de voorgrond niet zichtbaar is in het gebinariseerd beeld! We verduisteren het beeld in die mate dat de tekst leesbaar wordt.



**Verenigde Staten,  
een antwoord te vi  
maar met name oi**

**wyjscia kazdego  
brawia, ze nasze**



wyjscia kazdego  
brawia, ze nasze

**Verenigde Staten,  
een antwoord te vi  
maar met name oi**

**wyjscia kazdego  
brawia, ze nasze**

Het **contrast** bepaalt het lokaal contrast tussen de donkere en lichte delen van het beeld. (De tekst is gewoonlijk donkerder dan de achtergrond - tenzij u natuurlijk geïnverteerde tekst herkent.) Hier moet u de karaktervormen netjes uit hun (gekleurde) achtergrond doen springen. Hier hebt u een voorbeeld waar we het contrast moeten verhogen omdat de standaard waarde gebroken karakters oplevert.

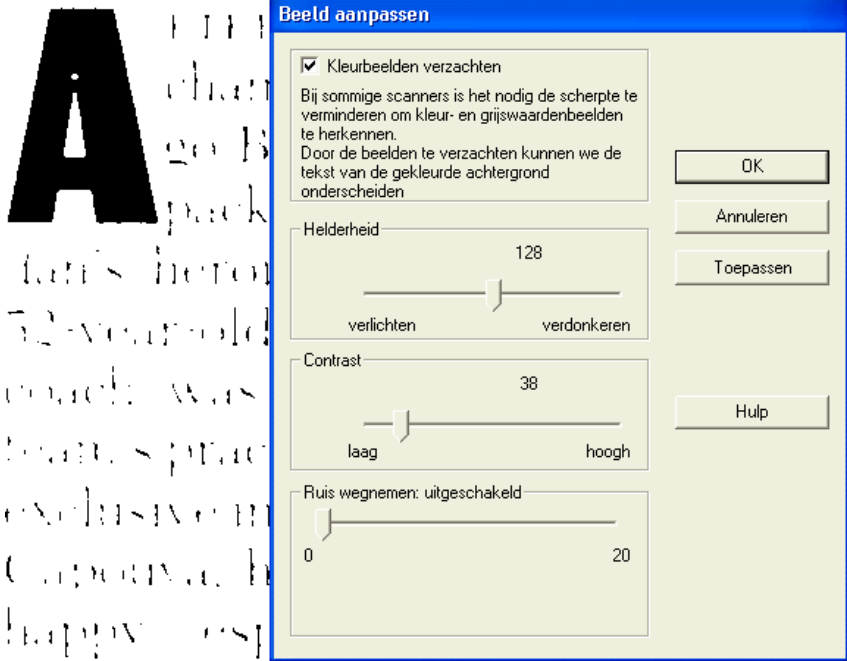
**A Look at International  
Planning the Future .....**

**A Look at International  
Planning the Future .....**

**A Look at International  
Planning the Future .....**

Let wel, het beeld wordt pas echt aangepast wanneer u de knop "Toepassen" aanklikt. Door op "OK" te klikken past u het beeld aan *en* sluit u het venster. Hier hebt u een voorbeeld waarbij het zwart-wit beeld dramatisch werd verlicht - zij het dat we geen goede OCR-resultaten verwachten van dit beeld!



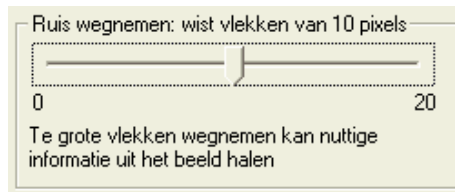


Deze opties betreffen kleur- en grijsbeelden. De laatste optie, "Ruis Wegnemen", slaat exclusief op de zwart-wit beelden. "Ruis" wegnemen betekent dat de toevallige, "parasitaire" pixels uit het zwart-wit beeld worden verwijderd.

**The Olympic Games**  
**The Olympic Games**



Zorg ervoor dat u geen te grote vlekken uit het beeld verwijdert, want het zou wel eens kunnen voorkomen dat u de punten op de "i" wegneemt of delen van matrix letters wist!



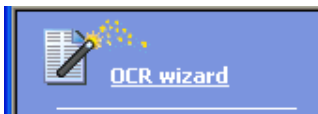
Dit is de beste manier om een beeld te optimaliseren: plaats het venster waarin u het beeld aanpast in een hoek van het scherm zodat u goed kunt beoordelen welk effect de aanpassingen die u uitvoert hebben op het beeld. Pas de parameters aan, klik telkens op "Uitvoeren" tot u een scherp en "proper" beeld op het scherm hebt.

## LAAT DE WIZARD VOOR U WERKEN

---

Laten we van start gaan met het scannen van documenten. In plaats van alle parameters te doorlopen zullen we echter gebruik maken van de **OCR assistent**, een erg makkelijke manier om bladzijden te herkennen.

Klik op de knop "OCR Wizard" op de hoofdknoppenbalk of selecteer het commando "OCR Wizard" onder het menu "Actie".



De assistent gidst u comfortabel doorheen de herkenning: het volstaat op enkele eenvoudige vragen te antwoorden om snelle resultaten halen met Readiris.



Merk op dat de OCR assistent begint te lopen telkens u Readiris opstart; u kunt dit vermijden door de optie "Wizard Activeren bij Opstarten" in het eerste scherm van de wizard uit te schakelen. (U kunt ook de gelijkwaardige optie onder het menu "Instelling" gebruiken.)

## READIRIS HERSCHEPT DE ORIGINELE LAY-OUT

De OCR assistent automatiseert de herkenning in hoge mate, maar u mag deze "automatische" herkenning niet verwarren met automatische formattering! "Autoformattering" houdt met name in dat Readiris een **getrouwe kopie** aanmaakt van het gescande document: de woord-, paragraaf- en bladformattering van het oorspronkelijk document keren terug in de uitvoer.

Gelijkaardige lettertypes als in het brondocument (met schreef of schreefloos, proportioneel of vast, normaal of samengedrukt) worden gebruikt, de puntgroottes

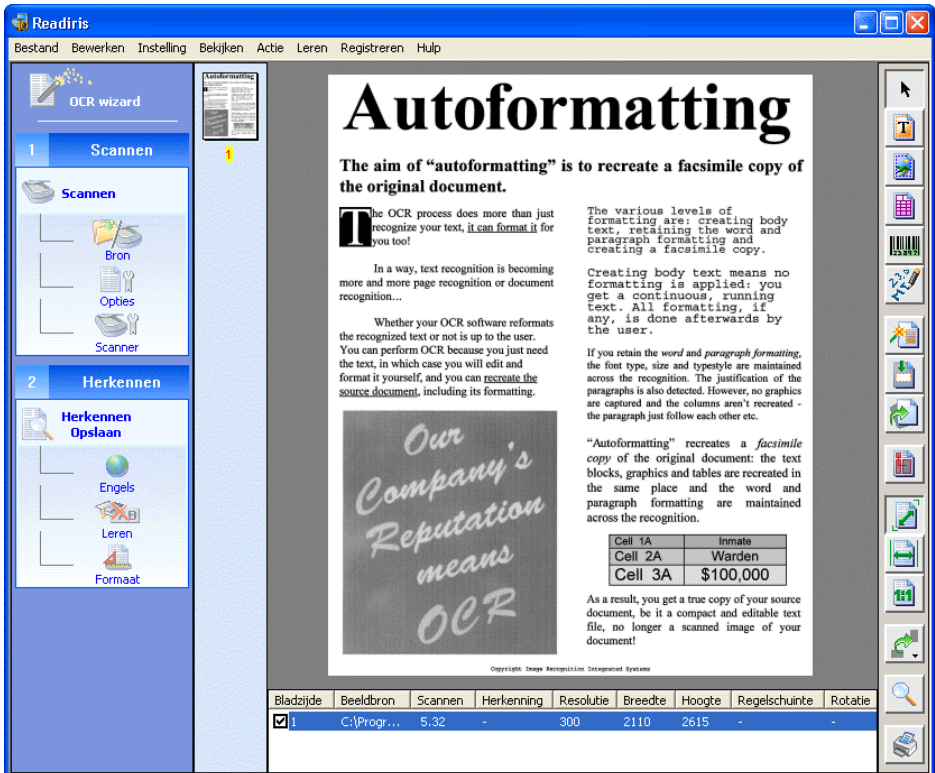


en stijlen (vet, cursief en onderlijnd, superscript en subscript) worden behouden tijdens de herkenning. De tabulatie en de uitlijning (links, gecentreerd, rechts en uitgelijnd) van elke tekstblok worden gerecreëerd. Ook de opsommingstekens (“bullets”) en de nummering worden herschapen. Alle e-mailadressen en URLs van webbladzijden worden gedetecteerd en als hyperlinks in de uitvoer geplaatst. De plaats van de kolommen, tekstblokken en figuren is gebaseerd op het origineel document.

Met andere woorden, Readiris laat u toe om een kopie te archiveren van uw documenten, zij het dat u dit keer te maken hebt met een editeerbaar, compact tekstbestand, en niet langer met een gescand beeld!

Dit alles houdt in dat bij “autoformatting” de sortering van vensters slechts *gedeeltelijk* geldt: u kunt zones insluiten en uitsluiten, maar als u de sorteerder wijzigt, wordt dat genegeerd!

Laten we bekijken hoe dit in zijn werk gaat. Om vertrouwd te raken met “autoformatting” opent u best het beeld AUTOFORMAT.JPG in de Readiris map.




Klik op de "Formaat" knop op de hoofdknoppenbalk en kies ervoor het herkenningresultaat naar Microsoft Word te zenden of selecteer het uitvoerformaat Word (\*.doc) of RTF (Rich Text Format). Activeer eveneens de lay-outoptie "Brondocument Hercreëren". (De optie "Paragrafen Herstellen" wordt standaard ingeschakeld om binnen de paragrafen "wordwrap" toe te passen.)



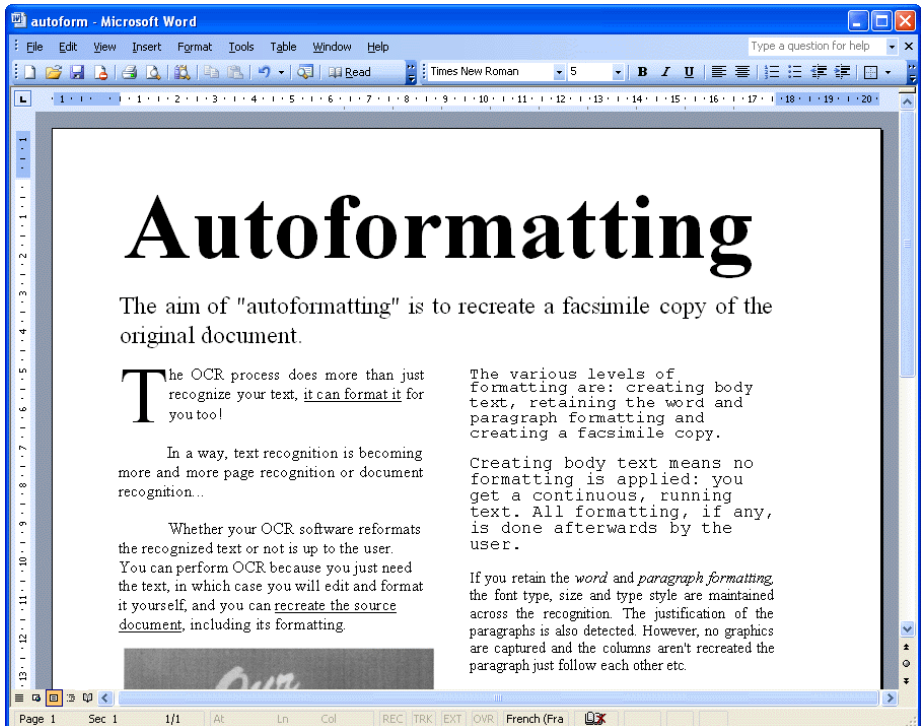
Lay-out

- Lopende tekst creëren
- Woord- en paragraafformattering behouden
- Brondocument hercreëren
  - Kolommen gebruiken in de plaats van kaders
  - Kolomgrenzen invoegen

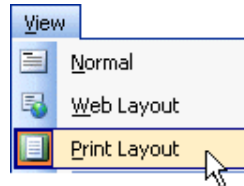


Of de opmaakreconstructie beschikbaar is, hangt af van de gekozen uitvoermodus. Sommige “arme” tekstformaten die “platte” tekst (“plain text”) genereren zoals Tekst (ANSI), MS-DOS Tekst (ASCII) enz. ondersteunen geen geavanceerde formatteercodes en kunnen dan ook geen “autoformattering” bieden. Het Adobe Acrobat PDF formaat aan de andere kant werd specifiek ontworpen om de lay-out van uw documenten te bewaren: PDF documenten impliceren dus van nature uit autoformattering!

Wanneer de herkende tekst geopend wordt met een heuse tekstverwerker ziet hij er als volgt uit zonder enige interventie van de gebruiker.



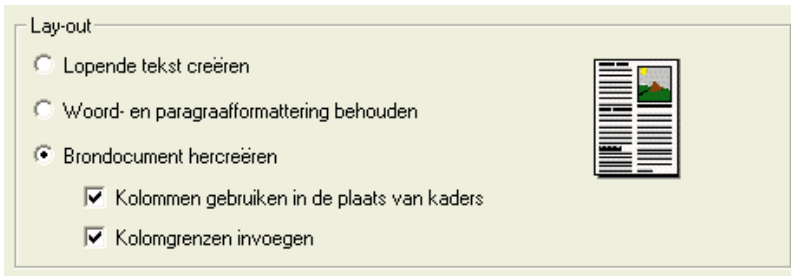
Om de lay-out correct te beoordelen moet u wel de “WYSIWIG”-modus van uw tekstverwerker activeren; meestal is er sprake van een “pagina-lay-out” optie. Let wel: wanneer u het herkende document rechtstreeks naar Microsoft Word zendt, wordt de “pagina-lay-out” optie automatisch geactiveerd!



Samengevat, Readiris herkent niet alleen uw teksten, maar kan ze ook formatteren. OCR is tegenwoordig meer dan tekstherkenning, OCR is eigenlijk **documentherkenning** geworden!

## KOLOMMEN AUB, GEEN KADERS!

De formatteeroptie "Kolommen Gebruiken in de Plaats van Kaders" bepaalt *hoe* de "autoformatting" plaatsvindt: de tekstblokken, tabellen en figuren kunnen in kaders of in editeerbare **kolommen** worden opgeslagen.



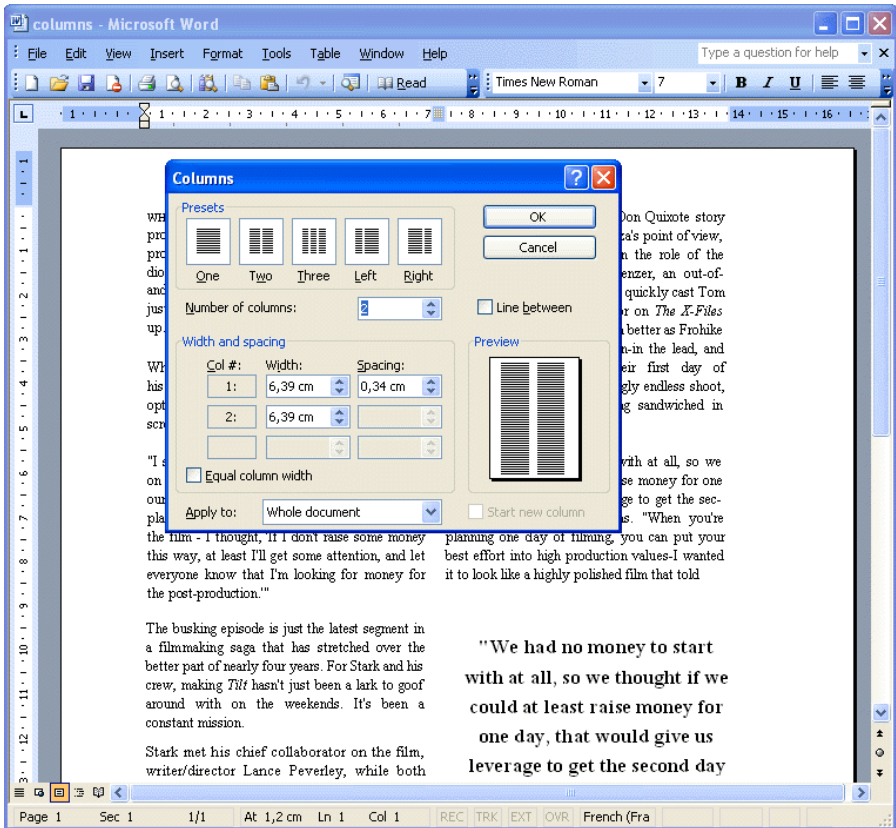
Kaders zijn aparte "containers" zodat je meerdere tekstblokken, figuren en tabellen op een bladzijde kan plaatsen. Bij kolommen ligt dat anders: de tekst vloeit natuurlijk van de ene kolom over in de volgende kolom. Tekst die in kolommen is geschikt is veel makkelijk te wijzigen dan tekst die over meerdere kaders is verdeeld.

We veronderstellen nu wel dat er wel degelijk kolommen voorkomen op het gescande document: wanneer het systeem niet in staat is kolommen terug te



vinden in het brondocument, valt deze lay-outmodus op het gebruik van kaders terug!

Laad het beeldbestand COLMUNS.TIF uit de Readiris map in het geheugen als u dit alles eens wilt proberen.





De optie "Kolomeindes Invoegen" verfijnt het gebruik van heuse kolommen verder: deze optie bepaalt of u een kolomgrens toevoegt aan het einde van elke kolom.

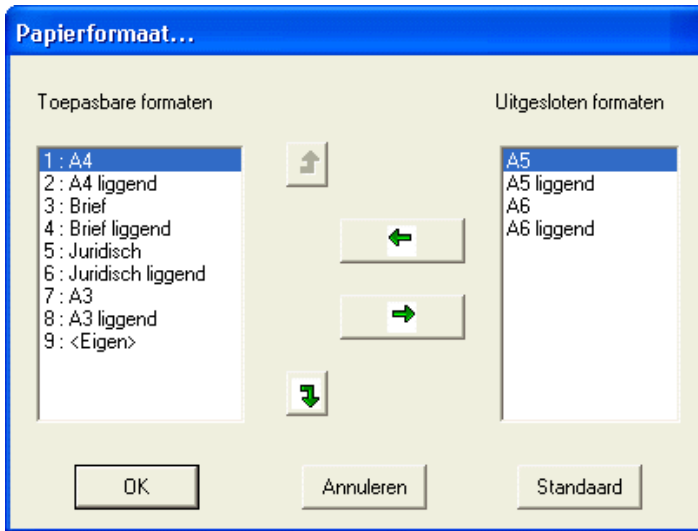
Wanneer er kolomgrenzen toegevoegd worden, blijft elke tekst u die u wijzigt, toevoegt of verwijdert binnen de kolom; tekst zal nooit vanzelf over een kolomgrens heen vloeien. Alle tekst die na een kolomgrens komt verschuift naar het begin van een volgende kolom!

Schakel deze optie in wanneer u de kolomgrenzen wilt behouden waar die in het brondocument werden gedetecteerd - welke wijzigingen u later ook aanbrengt na de herkenning. In kranten en tijdschriften behoren de kolommen vaak aan verschillende artikelen toe. De tekst stiekem van de ene kolom in de andere laten overlopen is dan niet zo'n goed idee!

Schakel deze optie uit wanneer u een doorlopende tekst hebt die in kolommen is geschikt; u zorgt er dan voor dat de tekst op natuurlijke wijze van de ene kolom naar de andere overloopt.

Er is één aspect waar u bewust kunt beslissen om het brondocument *niet* te herscheppen: het paginaformaat van uw uitvoerdocumenten. Wat bedoelen we hiermee? We zullen enkele voorbeelden geven: u scant bladzijden met het Letter formaat maar u slaat de uitvoer op in het A4 formaat: A4 bladzijden worden meteen afgedrukt, pagina's in Letter formaat vereisen dat u handmatig papier in de printer voert. Of u bent advocaat en u scant A4 bladzijden die u opslaat in het formaat Legal.

Vandaar dat u met Readiris de voor uw uitvoerdocumenten gewenste papierformaten kunt aangeven. Klik op de knop "Papierformaat" in het dialoogvenster "Tekstformaat".



Selecteer de toepasbare en uitgesloten papierformaten: de toepasbare papierformaten kunnen worden gebruikt om de herkende documenten te formatteren, de uitgesloten formaten worden nooit gebruikt. Orden de toepasbare formaten: Readiris doorloopt de verschillende papierformaten in de aangegeven volgorde en past het eerste papierformaat toe dat voldoende groot is om het gescande document te bevatten. De knop "Standaard" past de standaard waarden opnieuw toe. (Die standaard waarden hangen overigens af van uw Windows instelling!)

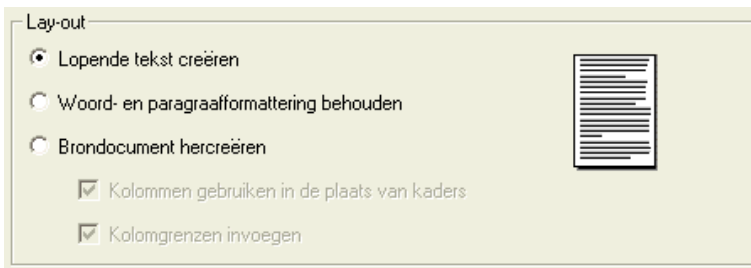
Deze optie is niet van toepassing op HTML bestanden - dat tekstformaat zonder papierformaten werd voor het Internet ontwikkeld! Het is evenmin van toepassing op PDF bestanden die een specifiek formaat gebruiken om het brondocument precies te herscheppen.



## TEKSTFORMATTERING, DEEL 2

De andere lay-outopties zijn "Lopende Tekst Creëren" en "Woord- en Paragraafformattering Behouden".

Zoals het pictogram aan de rechterkant illustreert, betekent **lopende tekst creëren** dat u een niet geformatteerde, "continue" tekst aanmaakt. De tekst wordt hernomen, maar van zijn lay-out wordt abstractie gemaakt. Gebruik deze optie wanneer u gewoon een document wilt overnemen als broodtekst maar geen belang hecht aan zijn opmaak.



Een lopende tekst is tevens wat u krijgt wanneer u snel een tekstzone herkent door er met de rechter muisknop op te klikken en het bevel "Kopiëren als Tekst" te selecteren: wanneer de herkenning van de zone klaar is, kunt u een lopende tekst in uw teksttoepassing plakken.

De optie "Woord- en Paragraafformattering Behouden" is de middenweg: de **woordformattering** - het lettertype, de puntgrootte en stijl - wordt behouden gedurende de herkenning. Dat is ook het geval met de **paragraafformattering** - de tabs en de uitlijning.

Verwar deze formatteroptie echter niet met de "volle" autoformattering: deze optie plaatst de paragrafen gewoon na elkaar, ze creëert geen kolommen en herschept evenmin de relatieve positie van de verschillende zones.

## TEKST MEERDERE KEREN UITVOEREN

Overigens kunt u de herkenningresultaten meerdere keren exporteren zonder de herkenning opnieuw uit te voeren! Verander het tekstformaat en de formatteeropties met de knop "Formaat" en klik opnieuw op de knop "Herkennen Opslaan". Deze keer wordt er geen herkenning uitgevoerd - tenzij u nieuwe vensters tekende of bestaande vensters wijzigde. Zoniet zal Readiris gewoon de herkende tekst opnieuw formatteren en opslaan in het nieuwe formaat of naar de nieuwe doeltoepassing sturen.

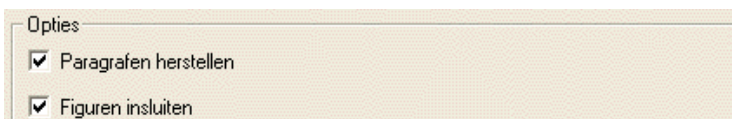


Datzelfde principe geldt voor enig ander element dat u wijzigt: wanneer u een bladzijde toevoegt aan uw document, zal enkel die pagina herkend worden. Wanneer u een nieuwe tekstzone toevoegt op eender welke bladzijde, zal enkel die zone worden herkend eer de resultaten worden opgeslagen.

Stel, u herkent een document van 10 bladzijden en slaat het op in een Word bestand. Daarna scant u de samenvatting op de coverbladzijde en u die stuurt die per e-mail naar een ongeduldige collega. Tenslotte scant u de bijlage - een tabel - en u slaat alle bladzijden op in een HTML bestand dat u op de webstek van uw onderneming plaatst.

## FIGUREN APART OPSLAAN

In het vorig voorbeeld was de figuur opgenomen in het uitvoerdocument; of dit al dan niet het geval is, hangt af van de formatteeroptie "Figuren Insluiten". Of het mogelijk is figuren binnen de tekst te bewaren hangt van de uitvoermodus af: "arme" tekstformaten zoals Tekst (ANSI) enz. slaan geen figuren op!



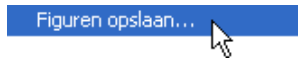


Maar u kunt met Readiris ook figuren en opslaan zonder aan tekstherkenning te doen! En aangezien Readiris zwart-wit beelden, grijswaarden en kleuren ondersteunt, kunt u probleemloos zwart-wit grafisch materiaal en foto's scannen.

Hoe? Teken een grafische zone rond de illustratie, de tekening enz. die u wilt opslaan. Dat gebeurt op dezelfde wijze als wanneer u tekst- en tabelvensters tekent, alleen selecteert u nu de tool "Grafisch Venster".

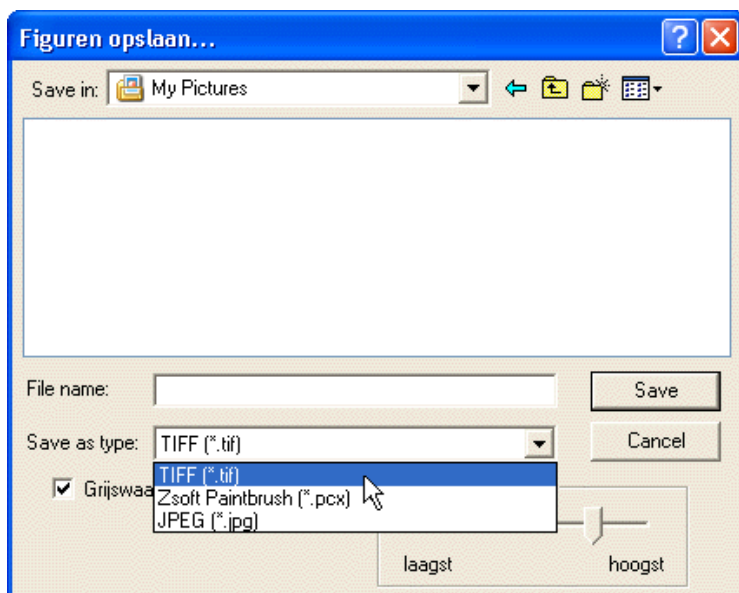


Kies daarna het commando "Figuren Opslaan" onder het menu "Bestand".

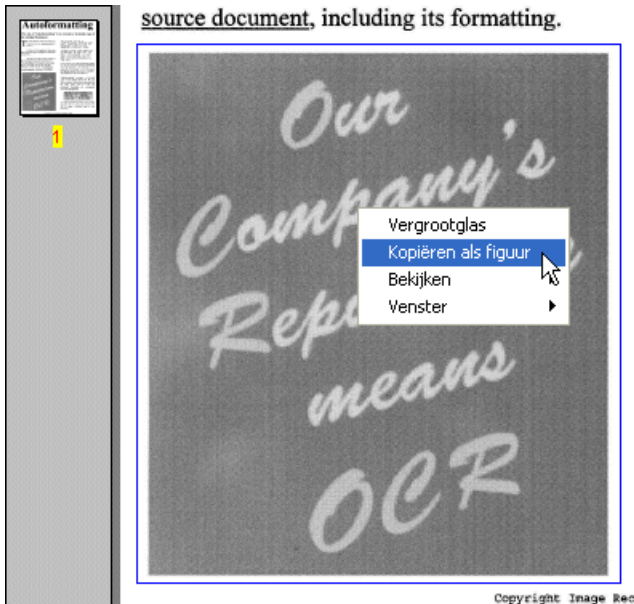


U moet nu een bestandsnaam opgeven. Bepaal welk grafisch formaat u wilt gebruiken. Selecteer een bestandsformaat dat uw fotoretouchesoftware ondersteunt. U hebt de keuze tussen de formaten JPEG, TIFF en ZSoft Paintbrush (\*.pcx). Readiris Corporate ondersteunt tevens het compact formaat JPEG 2000 (\*.j2c)!

Schakel de optie "Grijswaarden/Kleur" in om de figuur in kleur of grijswaarden op te slaan. Wanneer u zwart-wit figuren opslaat in het TIFF formaat, wordt Groep 4 compressie toegepast. Wanneer u figuren in grijswaarden of kleur opslaat in het TIFF formaat, wordt JPEG compressie toegepast en kunt u de JPEG kwaliteit aanpassen.



Om een figuur naar het klembord te sturen eerder dan die op te slaan in een beeldbestand, klikt u met de rechter muisknop boven een grafisch venster en selecteert u het bevel "Kopiëren als Figuur": de grafische zone onder de muiscursor is klaar om geplakt te worden!



## GEKLEURDE ACHTERGRONDEN OPSLAAN

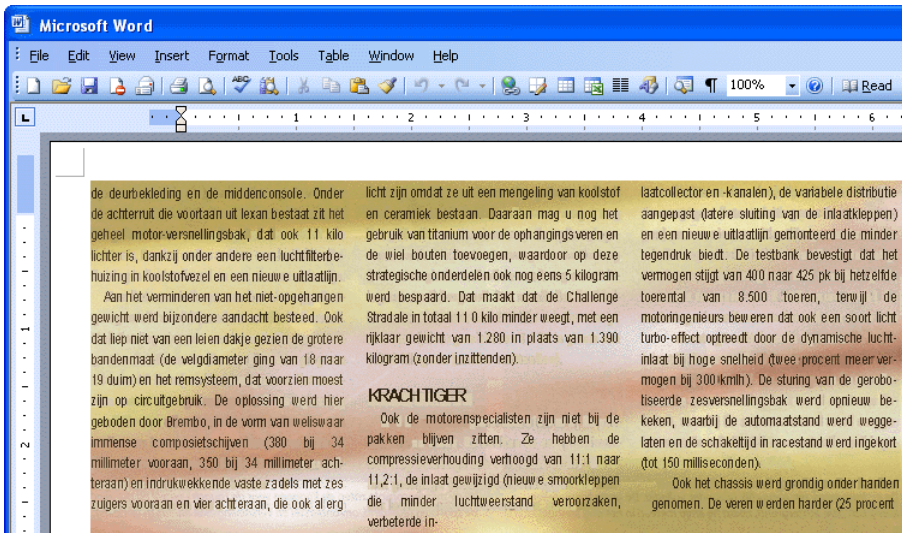
Met Readiris Corporate kunt u tevens de achtergrond bewaren binnen de herkende documenten! De optie "Beeld Invoegen als Achtergrond Bladzijde" plaatst het gescande beeld als achtergrond onder de herkende tekst; dit is mogelijk bij HTML, RTF, Word (\*.doc) en WordML (\*.xml) uitvoer. (Wanneer u PDF bestanden aanmaakt, kunt u hetzelfde effect verkrijgen door het formaat "PDF Tekst-Beeld" te selecteren.)

Het spreekt voor zich dat deze optie de bestandsgrootte van uw herkende documenten gevoelig verhoogt: deze optie voegt immers het gescande beeld met al zijn details in de achtergrond toe aan het uitvoerdocument. *Al zijn details? Nou*



ja: zoals ook gebeurt bij PDF uitvoer van het type “tekst-beeld” worden de pixels van de herkende tekst gewist om de uitvoer leesbaar te maken. Tekst in het zwart afbeelden boven zwarte bitmaps van karakters zou uw tekst een zware schaduw meegeven.... (Het voorbeeldbestand BACKGROUND.JPG illustreert hoe één en ander werkt.)



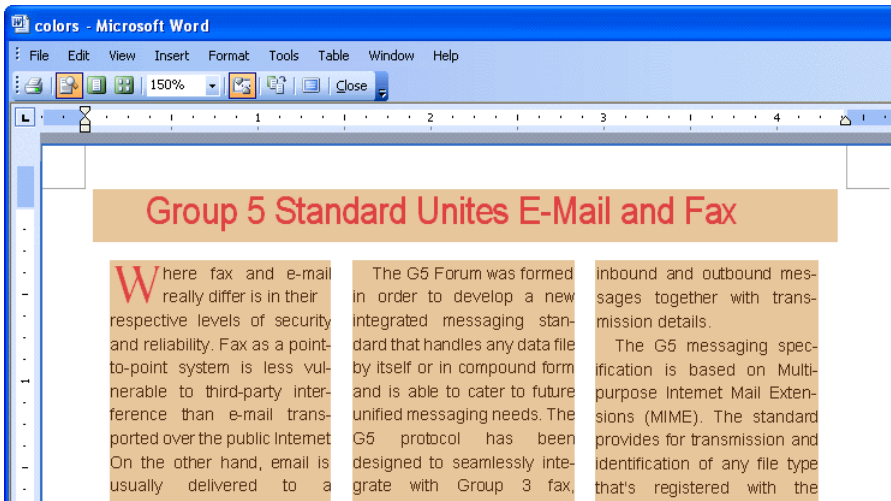


Maar Readiris biedt ook een minder ingrijpend, compacter alternatief aan: de optie "Kleuren van Achtergrond Behouden" behoudt de steunkleuren die op een bladzijde voorkomen na de herkenning. (Deze optie verplicht u er toe ook de kleuren van de tekst te bewaren.)



U krijgt één uniforme achtergrondkleur - als die tenminste in het gescande document voorkomt - per paragraaf. Ook deze optie is van toepassing op HTML, Word en WordML documenten. De details van bijv. een achtergrondfoto die de

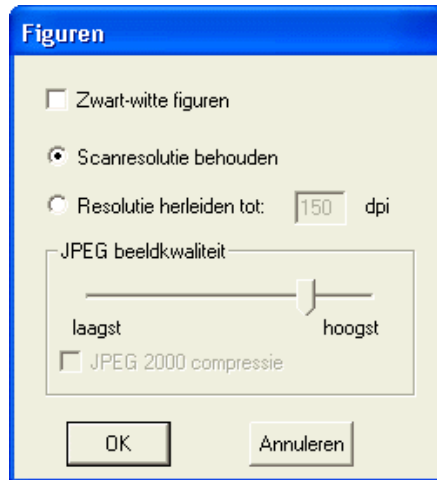
volle bladzijde bedekt worden deze keer *niet* bewaard; alleen de steunkleur van een tekstkader wordt wel behouden. (Ga uw gang met het voorbeeldbestand COLORS.JPG als u het eens zelf wil proberen...)



## FIGUREN MAXIMAAL BENUTTEN

Readiris Corporate biedt nog enkele geavanceerde opties voor de figuren. U vindt deze onder de knop "Geavanceerd" van het bevel "Tekstformaat".

Deze opties zijn op *alle* figuren van toepassing - de grafische zones die in herkende documenten worden opgenomen en de beelden van de bladzijden die u boven de tekst plaatst in een Adobe Acrobat PDF bestand van het type "beeldtekst".



Bepaal de kleurmodus: sla uw figuren op in kleur-grijswaarden of als zwart-witte figuren. Kies de resolutie van de figuren in de herkende documenten: behoud de scanresolutie of reduceer die. (U kunt de resolutie op deze manier echter niet *verhogen!*) Wanneer de herkende documenten als HTML documenten op een webstek worden geplaatst, zult u allicht de figuren met de schermresolutie willen opslaan. Vandaar dat de figuren in HTML bestanden standaard gereduceerd worden tot 72 dpi. (U kunt wel manueel een hogere waarde invullen.) Kies tenslotte de JPEG kwaliteit. (JPEG beelden worden gebruikt om kleur- en grijswaardenfiguren in PDF documenten, Word en RTF bestanden enz. op te slaan.)

(De optie "JPEG 2000 Compressie" hebben we al besproken: ze past JPEG 2000 compressie toe op alle figuren en beelden die in PDF bestanden worden opgeslagen.)

Met deze opties kunt u de **bestandsgrootte** van uw uitvoerdocumenten ingrijpend beïnvloeden. Laten we een voorbeeld geven: wanneer u PDF bestanden aanmaakt van het type "beeld-tekst" met zwart-wit-figuren, slaat u de beelden

op in TIFF Groep 4 gecomprimeerde bestanden. Sla diezelfde scans op als kleurbeelden en u bewaart standaard JPEG beelden met een hoge (0,8) kwaliteit.

## **FAXEN HERKENNEN EN UITGESTELDE HERKENNING**

Gescande documenten als beeldbestanden bewaren opent een uitdagende mogelijkheid: u kunt de *volledige* pagina bewaren en er later **uitgestelde OCR** op uitvoeren. Dat is eigenlijk wat we eerder deden met de vooraf gescande beelden van onze oefeningen.

Scan een document. Selecteer het commando "Volledige Bladzijde als Beeld Opslaan" onder het menu "Bestand" om een enkele bladzijde op te slaan. U wordt uitgenodigd de volledige bladzijde te bewaren als een JPEG, JPEG 2000 (\*.j2c), TIFF of ZSoft Paintbrush (\*.pcx) bestand.

A screenshot of a software menu. A blue rectangular button with white text "Volledige bladzijde als beeld opslaan..." is highlighted. A mouse cursor is positioned over the button.

Alle bladzijden als beeld opslaan...

Selecteer het commando "Alle Bladzijden als Beeld Opslaan" om een meerbladig document op te slaan. Er zijn nu slechts twee bestandsformaten beschikbaar: PDF en het meerbladig TIFF formaat.

U kunt nu de schijf selecteren als bronbeeld en het beeld openen met de knop "Openen" (of met het overeenkomstige bevel onder het menu "Actie"). (Als u het commando "Openen" onder het "Bestand" menu gebruikt moet u de beeldbron zelfs niet aanpassen.)

Daar Readiris beelden in kleur, grijswaarden en zwart-wit ondersteunt, kunt u Adobe Acrobat PDF bestanden, DCX faxen (meerbladige versie van het Paintbrush formaat), DjVu beelden (\*.djv, \*.djvu), JPEG beelden, JPEG 2000 beelden (\*.j2c, \*.jp2), PNG beelden, (niet gecomprimeerde of LZW, PackBits, Groep 3, Groep 4 en JPEG gecomprimeerde) TIFF beelden, meerbladige TIFF beelden, Windows bitmaps (\*.bmp) en ZSoft Paintbrush (\*.pcx) beelden openen.

Deze mogelijkheid is vooral handig om **faxen** om te zetten in editeerbare tekstbestanden! Readiris gebruikt overigens extra intelligentie om faxen te herkennen: de software detecteert de typische faxresoluties - 100 x 200 dpi



(“normale kwaliteit”), 200 x 200 dpi (“fijne kwaliteit”) en 200 x 400 dpi (“superfijne kwaliteit”) - en optimaliseert deze beelden vóór de herkenning automatisch om een zo goed mogelijk herkenningresultaat te behalen.

Niettemin is het een goed idee om uw correspondenten te vragen faxen te sturen met de “fijne” kwaliteit - deze faxen hebben de hogere resolutie van 200 dpi en zullen een beter OCR resultaat opleveren.

Vergeet overigens niet dat u met de rechter muisknop op beelden kunt klikken in de Windows Verkenner en het bevel "Herkennen" van het "Context" menu kunt aanklikken. U kunt ook gebruik maken van een “drag and drop”-operatie om beelden te openen: sleep beeldbestanden van de Windows Verkenner naar de beeldzone of naar het pictogram van Readiris om ze onmiddellijk te openen!

## **DOCUMENTREEKSEN HERKENNEN**

---

Readiris Corporate is veel krachtiger in het herkennen van vooraf gescande beelden: u kunt hele reeksen documenten automatisch verwerken en u kunt een bewaakte map instellen. We gaan hier dieper op in.

**Reeksherkenning** voert de herkenning uit op alle vooraf gescande beelden in een specifieke map. U kunt uw documenten bijvoorbeeld overdag scannen en ze 's nachts herkennen... Scan alle te herkennen documenten in; wanneer u klaar bent laat u de reeksherkenning lopen. Laat uw PC die klus klaren: wanneer u de volgende dag terugkomt zijn al uw documenten herkend!

De herkende documenten krijgen dezelfde bestandsnamen als de beeldbestanden. De extensie van het bestand hangt natuurlijk af van het gekozen uitvoerformaat: als Word het geselecteerde tekstformaat is wordt beeldbestand 001.TIF omgezet in bestand 001.DOC.



Kies uw beeldmap en uw tekstmap. De tekstmap mag maar moet niet verschillen van de beeldmap! Wanneer beide mappen identiek zijn, vindt u de tekstdocumenten naast de scans (tenzij u de optie "Beelden Wissen na Herkenning" inschakelde)!

De optie "Submappen Verwerken" bepaalt of de submappen van de beeldmap ook verwerkt dienen te worden. Dit is zinvol wanneer de vooraf gescande documenten in specifieke submappen zijn geplaatst. U beschikt bijvoorbeeld over een map "12-01-2004" die de dag aangeeft en een submap "1", "2" enz. die de eigenlijke documenten bevat.

Schakel deze optie in en alle submappen worden verwerkt: u kunt de herkenning niet beperken tot (één) bepaalde submap(pen). Wanneer de tekstmap verschilt van de beeldmap, worden de submappen als een spiegelbeeld gecreëerd in de uitvoermap!

Klik op "OK" om alles uit te voeren. er is geen verdere inspanning vereist om de documenten om te zetten; de herkenning verloopt volautomatisch. (Het interactief leren is hier niet van toepassing.)



U bent nu beperkt tot de “externe” tekstformaten. De herkenningresultaten worden op de harde schijf opgeslagen. De leesresultaten meteen naar een doeltoepassing zenden, ze per e-mail versturen of ze automatisch openen na de herkenning heeft immers geen zin wanneer de herkenning wordt uitgevoerd op een verlaten PC. Maar alle andere opties - de taal van het document, het type letterteken enz. - zijn wel van toepassing: zorg ervoor dat u de correcte opties heeft ingesteld eer u de herkenning opstart!

Reeksherkenning biedt nog enkele geavanceerde opties. De optie "Uitvoerbestanden Beschrijven" bepaalt of het OCR-proces eerdere herkenningresultaten kan vervangen of niet. Schakel deze optie uit wanneer u beeldbestanden toevoegt aan een map die beelden bevat die al herkend werden. (Zoniet worden de beeldbestanden een tweede keer herkend wanneer u deze map verwerkt.)

De optie "Beelden Wissen na Herkenning" bepaalt of de beeldbestanden na de herkenning gewist worden. Met andere woorden, Readiris kan de beeldmap voor u leegmaken.

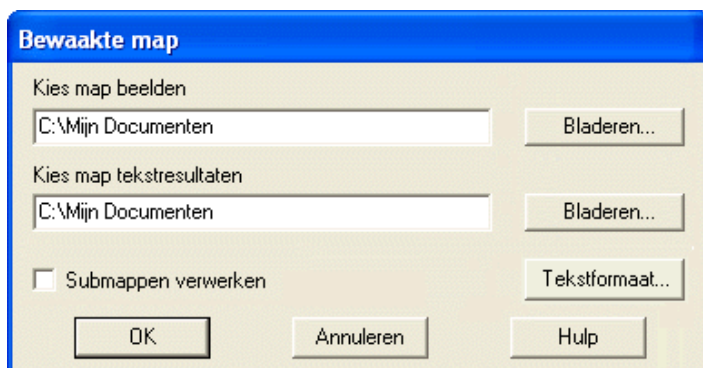
Schakel deze optie in wanneer u enkel de herkende documenten opslaat-archivert (waarbij u de beeldbestanden van “tijdelijke” aard negeert). Schakel ze in wanneer u de herkende documenten én de scans wilt bewaren. (We herhalen dat Readiris PDF documenten van het type “tekst-beeld” aanmaakt: dit uitvoerformaat slaat het gescand beeld *en* het herkend document op in één enkel bestand!)

## **EEN BEWAAKTE MAP INSTELLEN**

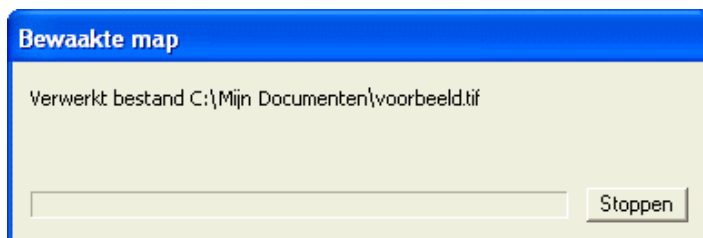
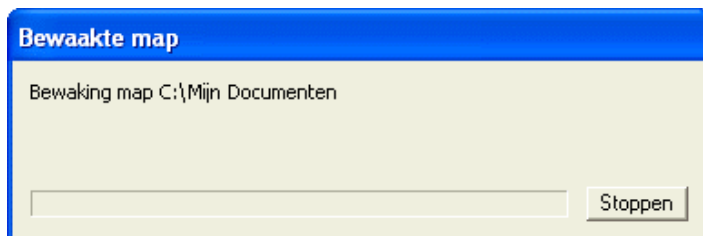
---

Het gebruik van een “bewaakte” map is grotendeels gelijk aan de werking van de reeksherkenning. Het grote verschil zit precies hierin dat Readiris deze keer de herkenning uitvoert op elk beeldbestand dat in een specifieke map wordt “gedropt”. U kunt de OCR software dag en nacht laten lopen... Scan nieuwe documenten in en ze worden prompt herkend.





Klik op "OK" om de bewaking van de gekozen map in te schakelen en Readiris zal progressief alle beelden verwerken van zodra die in de **bewaakte map** geplaatst worden. Klik op "Stoppen" om de bewaking op te schorten.

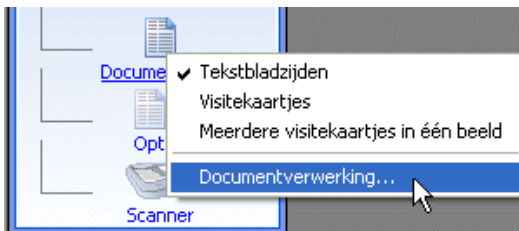




Net zoals bij de reeksherkenning het geval is, verwerkt Readiris de beelden in alle ondersteunde bestandsformaten. Het is niet mogelijk de herkenning te beperken tot één specifiek formaat: Readiris zet de beeldbestanden van *alle* ondersteunde grafische formaten om. (Alle bestanden met een ander bestandsformaat worden gewoon genegeerd.)

## DOCUMENTREEKSEN ORGANISEREN

Maar dit alles betekent helemaal niet dat u de voorgescande bladzijden noodzakelijkerwijze op een blinde manier, “recht-toe recht-aan” moet verwerken. Integendeel, Readiris Corporate is uitgerust met gesofistikeerde routines die u toelaten de gescande documentreeksen intelligent te verwerken. Ga naar het bevel "Documentverwerking" op de knop "Documenttype" (of onder het menu "Instelling") om de mogelijkheden te ontdekken.





U kunt blanco coverbladzijden invoegen om de **documenten te scheiden**. Een “blanco bladzijde” is een bladzijde die nauwelijks enige zwarte pixels bevat - met uitzondering van eventuele zwarte randen. De paginaknoppenbalk gebruikt een specifiek pictogram voor de coverbladzijden en het documentpaneel markeert ze.

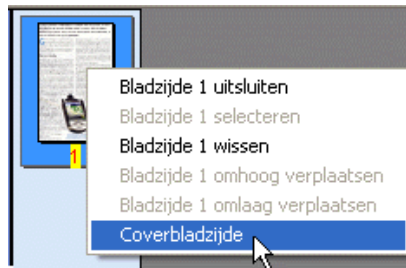


Bladzijde	Coverbladzijde	Streepjescode	Beeldbron	Scannen
<input checked="" type="checkbox"/> 1	ja		C:\Documents...	1.55
<input checked="" type="checkbox"/> 2	-	978905466618	C:\Documents...	1.27

Maar u kunt een “stroom” van gescande bladzijden ook manueel segmenteren: selecteer een bladzijde - die al dan niet blanco is - in de paginaknoppenbalk, klik met de rechter muisknop en activeer de optie "Coverbladzijde" op het "Context"



menu. (Weet dat de inhoud van een coverbladzijde altijd verloren raakt: coverpagina's fungeren als grens tussen twee documenten maar worden *niet* gelezen!)

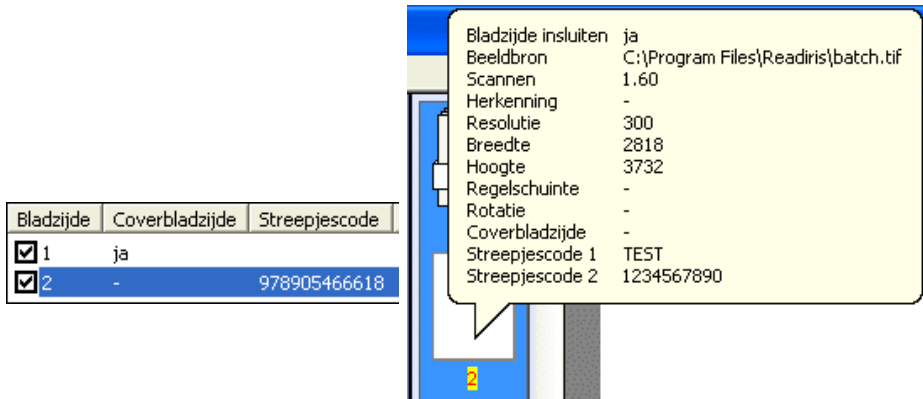


De verschillende documenten in een stroom van gescande bladzijden onderscheiden is slechts de eerste stap. Die documenten correct **indexeren** is de volgende stap. Hiervoor gebruiken we streepjescodelectuur. Alle streepjescodes die voorkomen op een pagina net na de blanco coverbladzijde worden automatisch gelezen. (Herken het voorbeeldbestand BATCH.TIF om het eens te proberen...)



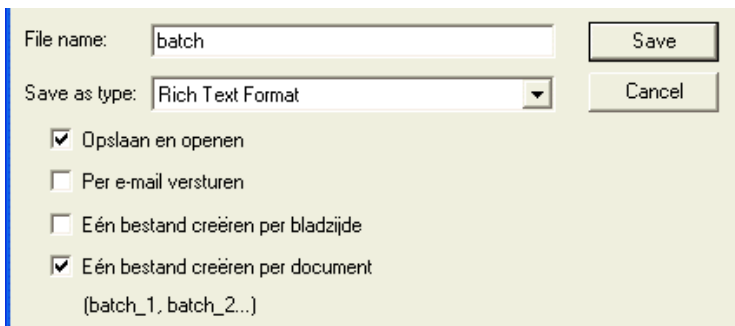
De optie "Duplex Scannen" geeft aan dat u de voor- en achterkant van documenten scant. De achterzijdjes worden genegeerd wanneer u naar streepjescodes zoekt op de volgende bladzijde: de volgende bladzijde is nu de volgende *voorkant*, niet het volgende beeld!

De gebruikersinterface van Readiris is ook op andere punten behulpzaam wanneer u documentreeksen gaat verwerken: het documentpaneel en de tooltips op de pictogrammen van de pagina's geven de waarde van de herkende streepjescodes aan!



De leesresultaten worden opgeslagen in de XML index, niet in het herkend document!

Wanneer u documentreeksen manueel verwerkt, zorgt de optie "Eén Bestand Creëren per Document" van het dialoogvenster "Opslaan" ervoor dat u aparte documenten bewaart in de plaats van volledige documentreeksen. (Dit kan enkel wanneer u externe bestanden aanmaakt, niet wanneer u de uitvoer rechtstreeks naar een doeltoepassing stuurt.)





## **TABELLEN HERKENNEN**

---

Tot nog toe hebben we teksten en faxen herkend en figuren opgeslagen. Laten we nu een tabel gaan verwerken. Neem een tabel en scan die of open het voorbeeld TABLES.JPG in uw Readiris map.

Het beeld TABLES.JPG bevat eigenlijk twee tabellen en dat is geen toeval! De pagina-analyse detecteert beide tabellen en Readiris zal ze hercreëren door ze cel per cel te reconstrueren in uw rekenblad of door een tabelobject in te voegen in uw tekstbestanden.

Laten we de verschillende mogelijkheden doorlopen. We starten met de “omkaderde” tabel, haar cellen zijn omgeven door een rand.

**Reading Tables**

Readiris recognizes tabular data and recreates them cell by cell in worksheets or as table objects inside wordprocessor files.

To insert tables as table objects, you must retain the word and paragraph formatting or recreate the source document; see the "Format" button on the main toolbar.

The page analysis detects "**gridded**" and "ungridded" tables. "Gridded" or "framed" tables have borders around the cells - as does the example below. The borders of the table cells get recreated.

Performance test optical media				
CD-ROM Digital Versatile Disk	Average access time (msec)	CPU utilization (%)	Video clip playbacks (frames dropped)	Sequential read 16 KB (K bps)
CD-ROM 4x speed	442	4.2	10 8	612
CD-ROM 12x speed	137	20.9	5 4	1,586
CD-ROM 24x speed	80	58.2	3 2	2,258
CD-ROM 32x speed	60	72.1	- -	2,987
DVD	58	78.9	- -	3,143

Tested on 333 MHz Pentium II PC with 64 MB RAM and 4 GB SCSI HD

"**Ungridded**" tables don't have any borders around the cells. When the columns of ungridded tables are too widely spaced, the page analysis may not detect a table window to avoid confusion with columnized text blocks.

When your tables exclusively contain **numeric characters**, enable the numeric reading mode with the "Language" button on the main toolbar for increased accuracy.

Bladzijde	Beeldbron	Scannen	Herkenning	Resolutie	Breedte	Hoogte	Regelschuine	Rotatie
<input checked="" type="checkbox"/> 1	C:\Progra...	5,44	-	300	2092	2517	-	-

Voer de herkenning uit terwijl de lay-outoptie "Woord- en Paragraafformattering Behouden" of "Brondocument Hercreëren" ingeschakeld is en de tabel wordt herschapen. Open uw tekstverwerker om het resultaat even te bekijken: de cellen en randen werden één voor één door Readiris herschapen! (U had trouwens de overige tekstblokken in de uitvoer kunnen insluiten.)



table - Microsoft Word

File Edit View Insert Format Tools Table Window Help

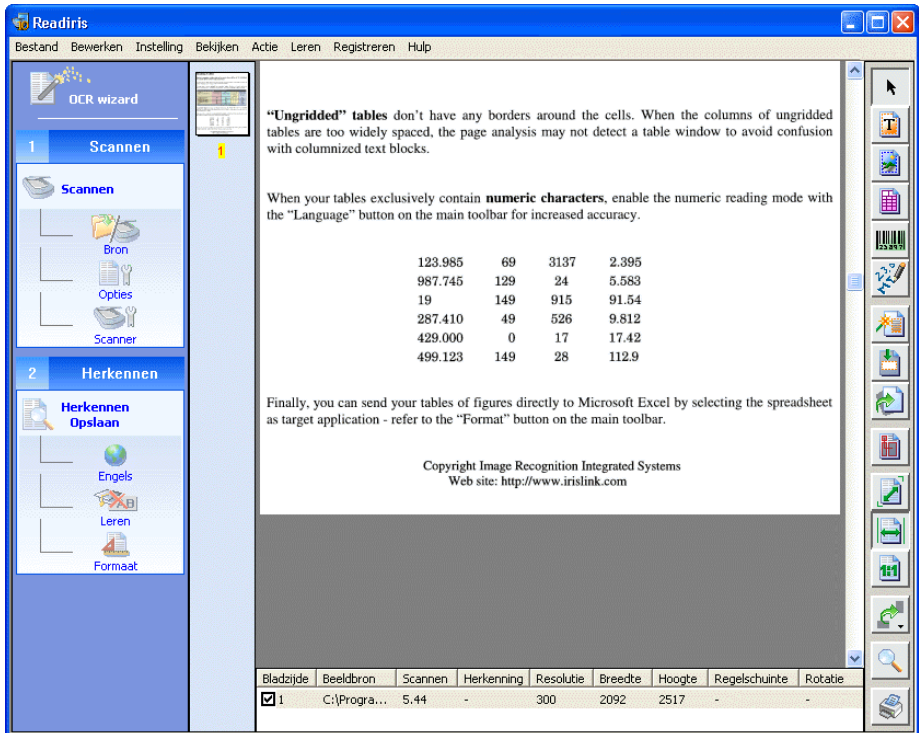
Type a question for help

Times New Roman 12 B I U

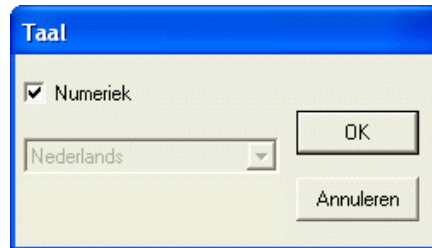
Performance test optical media					
CD-ROM Digital Versatile Disk	Average access time (msec)	CPU utilization (%)	Video clip playbacks (frames dropped)		Sequential read 16 KB (K bps)
CD-ROM 4x speed	442	4.2	10	8	612
CD-ROM 12x speed	137	20.9	5	4	1,586
CD-ROM 24x speed	80	58.2	3	2	2,258
CD-ROM 32x speed	60	72.1	-	-	2,987
DVD	58	78.9	-	-	3,143
Tested on 333 MHz Pentium II PC with 64 MB RAM and 4 GB SCSI HD					

Nu de “niet omkaderde” tabel. Die heeft geen randen rond de cellen, maar niettemin detecteert de pagina-analyse haar!



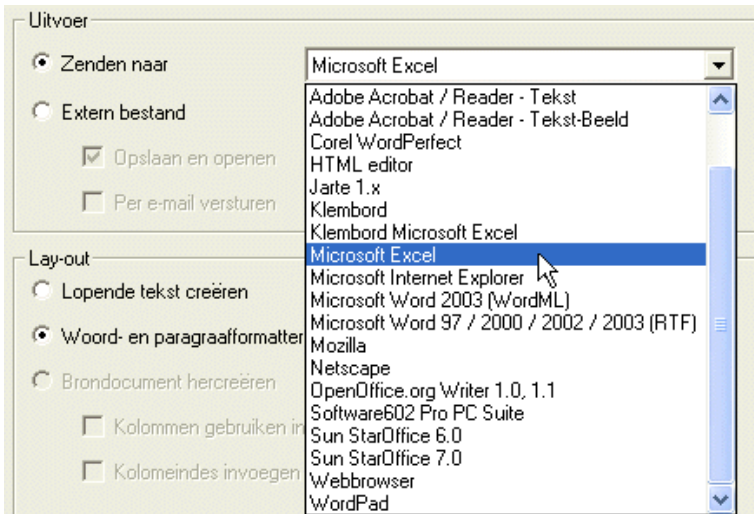


Voor een optimale OCR precisie, beperkt u de herkenning best tot de **numerieke symbolen** met de "Taal" knop. (De numerieke modus bevat niet enkel de getallen “0” tot “9”, maar ook de symbolen “+”, “\*”, “/”, “%”, “,” (komma), “.” (punt), “(”, “)”, “-”, “=”, “\$”, “£”, “¥” en het “€” symbool.)



Aangezien u dit alleen kunt doen wanneer de tabel geen alfabetische symbolen bevat - zoniet worden de tekstdelen niet goed herkend - kunnen we hier de numerieke modus activeren maar konden we dat eerder niet voor de eerste tabel.

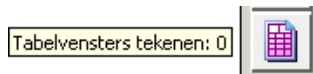
Deze keer zullen we het herkenningresultaat rechtstreeks naar het rekenblad Microsoft Excel sturen, we kiezen dus Excel als doeltoepassing met de "Formaat" knop.



Het rekenblad wordt automatisch opgestart en het resultaat ziet er zo uit: de typische tabelstructuur met rijen en kolommen werd herschapen, u bent dus onmiddellijk klaar om de data te verwerken.

	A	B	C	D	E
1	123.985	69	3137	2.395	
2	987.745	129	24	5.583	
3	19	149	915	91.54	
4	287.410	49	526	9.812	
5	429.000	0	17	17.42	
6	499.123	149	28	112.9	
7					

Het kan voorkomen dat u niet omkaderde tabellen ontmoet die de pagina-analyse niet als tabelvensters markeert omdat er teveel ruimte zit tussen de kolommen. Readiris tracht namelijk de verwarring met tekstkolommen te vermijden. Om tabelvensters manueel te creëren, klikt u op de tool "Tabelvenster" en gaat u zoals gewoonlijk te werk. De tooltip van de knop "Tabelvensters" geeft het aantal tabelvensters aan.



## HANDGESCHREVEN TEKST HERKENNEN

Tot nog toe hebben we documenten gescand, we hebben tabellen, faxen en plaatjes die we schoten met een digitaal fotoestel herkend, we hebben figuren opgeslagen en PDF bestanden omgezet. Readiris voegt nog een leuke leescapaciteit toe: de mogelijkheid om handgeschreven teksten te herkennen.



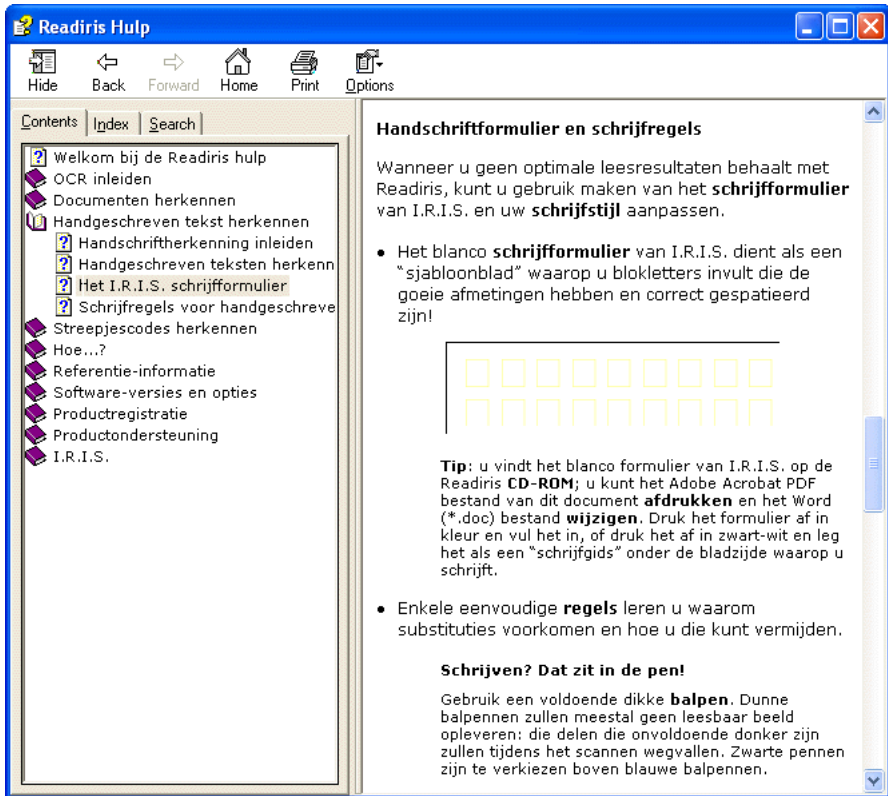
Eigenlijk zouden we over handgedrukte tekst moeten spreken, niet over handgeschreven tekst! Met *handschrift* bedoelt men gewoonlijk continue, “cursieve” handgeschreven tekst. De symbolen in een woord of karakterstring raken elkaar en het is onmogelijk te bepalen waar één symbool eindigt en het volgende symbool begint. Bij *handgedrukte* karakters spreken we over losse, aparte “blokletters” en de herkenningsoftware heeft het makkelijker om de individuele karakters te onderscheiden.

*Italiënskeie*  
*001-0405369-28*

Er is sterk gespecialiseerde software vereist - die heet “ICR” of “Intelligent Character Recognition” software - om handgedrukte symbolen te herkennen. De krachtige ICR technologie van I.R.I.S. is gebaseerd op méér dan één miljoen handschriften! Readiris ondersteunt alle natuurlijke Europese en Amerikaanse schrijfstijlen. Er is geen specifieke schrijfstijl vereist.

De herkenning van handgedrukte tekens is beperkt tot de cijfers (0-9), de hoofdletters (A-Z) en de leestekens “,” (komma), “.” (punt) en “-” (koppelteken).

Wanneer u leesresultaten behaalt die niet optimaal zijn, kunt u uw schrijfstijl aanpassen en I.R.I.S.’ schrijfformulier gebruiken. Raadpleeg het on-line hulpsysteem van Readiris om de schrijfregels te ontdekken. Enkele eenvoudige tips leren u waarom er substituties optreden en hoe u die kunt vermijden. Het blanco schrijfformulier van I.R.I.S. dient als een “sjabloonblad” waarop u blokletters invult met de correcte afmetingen en tussenafstand! U vindt het lege formulier op de CD-ROM van Readiris; u kan het formulier afdrukken en wijzigen.



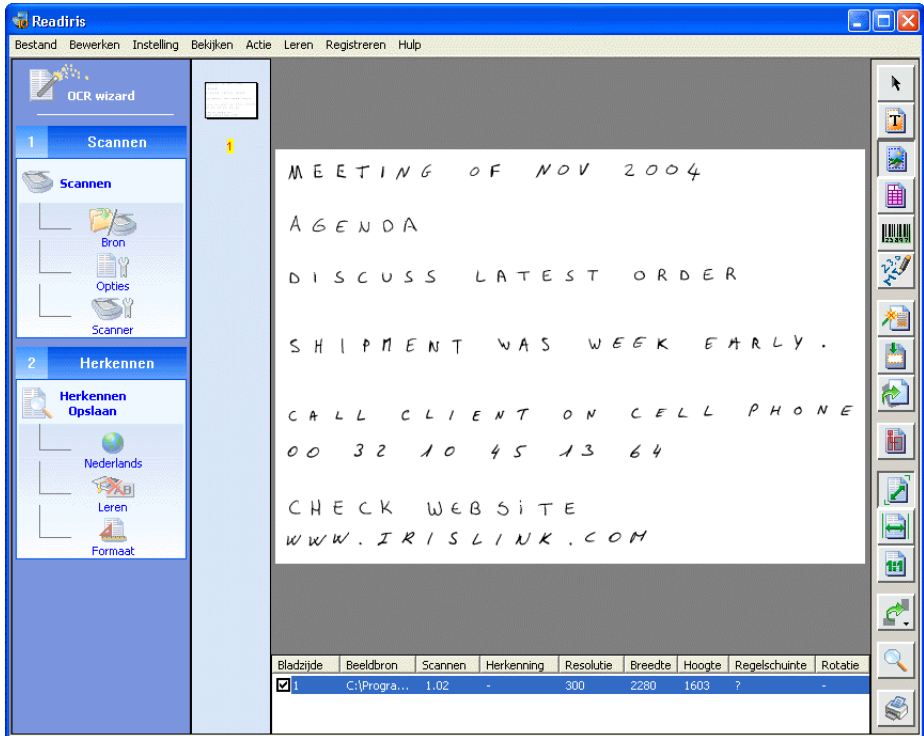
We weten nu dus hoe we tijdens een vergadering notities kunnen neerpennen die door de "machine" gelezen kunnen worden. Maar hoe kunnen we die notities achteraf inlezen met Readiris? Teken een handschriftvenster rond de handgeschreven tekst en voer de herkenning uit. (Ga eens uw gang met het voorbeeldbestand HANDPRINTING.TIF.)



Handschriftvensters tekenen: 0



De kenmerken van het document - de taal, het type letterteken en de karakterafstand - gelden niet voor handgeschreven tekst. U bent beperkt tot een elementaire Engelse - of moeten we zeggen “Latijnse”? - karakterset van blokletters. Interactief leren is evenmin van toepassing: het systeem trainen heeft geen zin in een context waarin iedereen zijn eigen handschrift heeft. (Zoals we al aanstipten is de ICR technologie gebaseerd op meer dan één miljoen schrijfstijlen...)



## BALKEN EN SPATIES INLEZEN

En Readiris leest zelfs streepjescodes...! Streepjescodes die voorkomen in gescande beelden kunnen worden gelezen en in de uitvoerdocumenten worden opgenomen als herkende data.



Streepjescodes zijn vastgesteld uit parallelle balken en spaties tussen die balken. Vooraf gedefinieerde combinaties van balken en spaties stellen specifieke karakters voor. Er zijn meerdere streepjescodestandaards, die heten in technische taal “symbolgieën”. Readiris ondersteunt alle populaire symbolgieën.

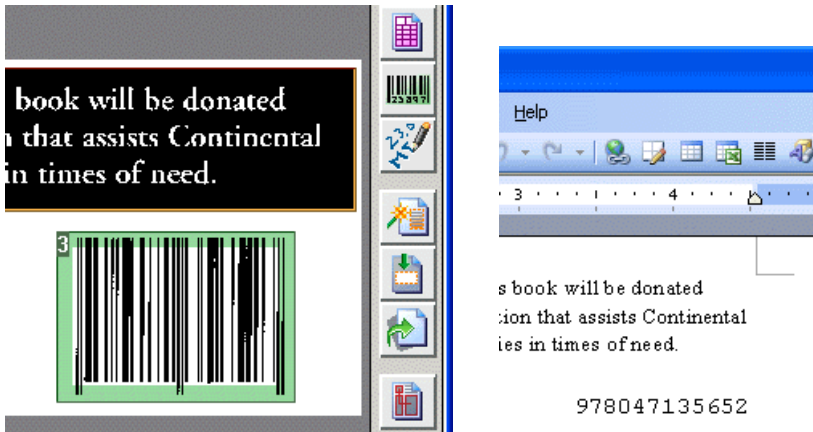


Enkel streepjescode die gedrukt werden met een laserprinter of een inkjet printer hebben voldoende druk kwaliteit. Streepjescodes die gedrukt werden met een matrixdrukker moet u uitsluiten: ze leveren niet voldoende contrast op en hun resolutie is meestal beperkt tot 60 dpi! Readiris herkent streepjescodes met een sterk contrast het makkelijkst; zwarte balken op een witte achtergrond geven de beste leesresultaten. De meeste types streepjescodes vereisen een “stille zone” rond de eigenlijke streepjescode. Streepjescodes produceren trouwens nooit gedeeltelijke resultaten: een afwezig start- of stopkarakter of een verkeerd controlekarakter leidt altijd tot een leesfout, een nul resultaat!



Teken een streepjescodevenster rond elke streepjescode - de pagina-analyse detecteert ze immers niet - en voer de herkenning uit. De streepjescodes worden gelezen en in de tekstuitvoer opgenomen. U kunt ook op een streepjescodevenster klikken met de rechter muisknop en het bevel "Kopiëren als Data" van het menu "Context" gebruiken: de streepjescode wordt dan herkend en naar het klembord gestuurd... (De controlekarakters van sommige streepjescodestandaards worden geverifieerd maar uit de leesresultaten gefilterd.) Het voorbeeldbestand BARCODE.TIF illustreert hoe het lezen van streepjescodes werkt.





## VISITEKAARTJES INLEZEN

Tot nog toe hebben we documenten gescand, we hebben tabellen, faxen en beelden die we schoten met een digitaal fotoestel herkend, we hebben figuren opgeslagen en PDF bestanden omgezet. Readiris Corporate voegt nog een leuke leescapaciteit toe: de mogelijkheid om **visitekaartjes** te **herkennen** (“BCR” - “Business Card Reading”).

Met Readiris kunt u visitekaartjes scannen, herkennen en zo omzetten in een **adresdatabank**. In deze context biedt OCR u de mogelijkheid kaartjes te verwerken zonder de tijdrovende taak ze eerst over te typen. Denk aan de laatste keer dat u terug kwam van een vakbeurs met een hele stapel visitekaartjes, en uw secretaresse twee dagen bezig was om de gegevens in te voeren!

De gegevens die op een kaartje voorkomen worden automatisch uit het beeld gehaald en de herkende data worden verdeeld over specifieke databankvelden. Readiris maakt hierbij extensief gebruik van een “kennisdatabank” die de software de vereiste intelligentie geeft om voor- en achternamen, steden en staten, telefoon- en faxnummers enz. te onderscheiden. Elk land heeft een eigen “stijl”



voor de opmaak van visitekaartjes: Nederlanders stellen een adres anders samen dan Fransen enz.

Dit kan voor wel **28 landen**: de Noord- en Zuid-amerikaanse visitekaartjes en kaartjes uit de Europese landen, inclusief de Oost-Europese landen, worden ondersteund. (Optioneel kunt u Aziatische kaartjes uit de volksrepubliek China, Japan, Korea en Taiwan herkennen.)

De resulterende gegevens zijn beschikbaar voor **uitvoer**. U kunt de contacten opslaan in een gestructureerd tekstbestand - bijvoorbeeld in het komma gedelimitteerd formaat of in een vCard bestand - zodat u de adressen in elke databank kan invoeren - bijvoorbeeld in Microsoft Access.

Maar u kunt de contacten ook rechtstreeks naar uw contact manager Microsoft Outlook (Express) en naar de **PDA** software Palm Desktop sturen. Readiris vult hier naadloos toepassingen aan zoals persoonlijke en relatiebeheersystemen, databanken of zelfs tekstverwerkers waarvan de mail merge-functie het drukken van brieven, enveloppen en etiketten mogelijk maakt.

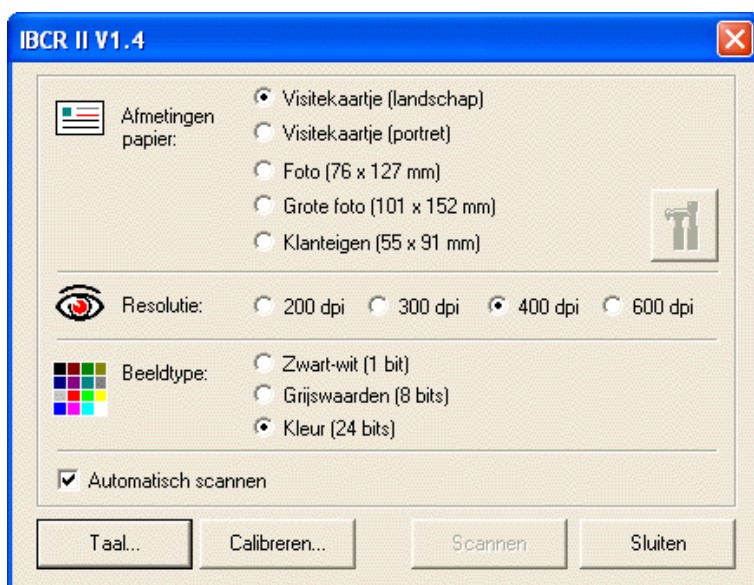
## **VISITEKAARTJES SCANNEN**

---

Hoe werkt dit alles? Om te beginnen moet de scanner goed ingesteld zijn.

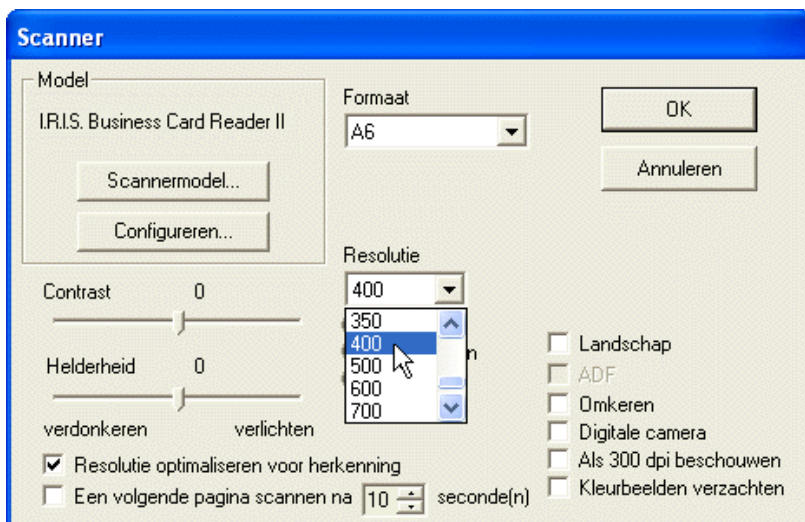
Wanneer het scannen van visitekaartjes voor u een belangrijke toepassing is - neem bijvoorbeeld dat u een pak visitekaartjes wilt inscannen op een vakbeurs - , kunt u gebruik maken van I.R.I.S.' gespecialiseerde visitekaartjesscanner **IBCR-II**. **Scanners voor visitekaartjes** zijn specifiek ontworpen om visitekaartjes in te scannen en hebben veel voordelen vergeleken met andere scannertypes: ze nemen nauwelijks plaats in op uw bureau, verwerken vlot kaartjes van verschillende papier- en drukkwaliteit en u hoeft zich nooit druk te maken over het selecteren van het juiste formaat.

Gebruik de Twain gebruikersinterface van de IBCR-II om de **automatische scanmodus** in te schakelen! Geef het scanbevel één keer en schakel de optie "Automatisch Scannen" in de Twain gebruikersinterface in.



U kunt nu het ene visitekaartje na het andere invoeren: het scannen vangt aan van zodra u een kaartje in de scanner legt...

Om visitekaartjes succesvol te herkennen, bevel we een **scanresolutie** aan van 450 dpi.



Wanneer u echter een **vlakbedscanner** gebruikt, kunt u meerdere visitekaartjes tegelijk op het vlakbed leggen en de individuele kaartjes door de software laten isoleren. De achtergrond moet wel zwart zijn opdat Readiris het “grote” beeld zou opdelen in de aparte visitekaartjes. Er is gelukkig een doodeenvoudige manier om hier voor te zorgen: laat het deksel van uw vlakbedscanner openstaan terwijl u uw visitekaartjes inscaneert...

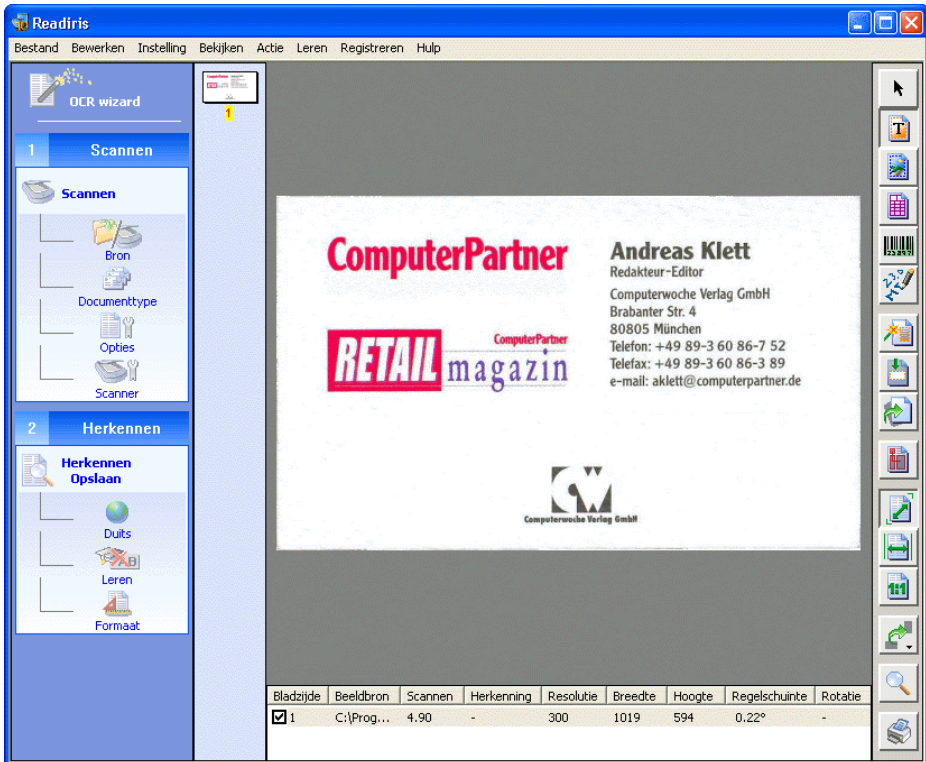


Schakel de optie "Meerdere Visitekaartjes in één Beeld" in met de knop "Documenttype" op de hoofdknoppenbalk (of met het bevel "Documenttype" onder het menu "Instelling") en scan uw visitekaartjes in.



Bij vlakbedscanners dient u op de knop "Scannen" te klikken om een beeld te verwerven! (U kunt natuurlijk ook het bevel "Scannen" uit het menu "Actie" gebruiken.)

De volle bladzijde zoals u die gescand hebt zal nooit als dusdanig op uw beeldscherm verschijnen: u krijgt enkel de "opgedeelde" visitekaartjes te zien.

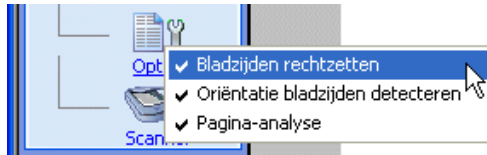


Indien u vergat de “segmentatie” van het beeld in te schakelen met de modus voor meerdere kaarten, gebruik dan het bevel "Visitekaartjes Uitsnijden" onder het menu "Actie" om het beeld op te delen in de individuele visitekaartjes en de overbodige zwarte randen weg te gooien.

Visitekaartjes uitsnijden



Maar welk scannertype u ook gebruikt, onzichtbaar worden enkele opties ingeschakeld om het leesproces te optimaliseren: het rechtzetten van de bladzijden en de detectie van de oriëntatie van de bladzijden. Klik op de knop "Opties" in de hoofdknoppenbalk om deze opties terug te vinden.



Het is moeilijk voorstelbaar dat de visitekaartjes die u op een vlakbedscanner legde (allemaal) netjes recht zullen staan. Laat de software dit dus voor u regelen... (Mocht u deze optie per ongeluk uitgeschakeld hebben, dan kunt u met de knop "Bladzijde Rechtzetten" de visitekaartjes achteraf rechtzetten, maar het wordt snel een vervelende klus als u een groot aantal kaartjes handmatig moet rechtzetten!)

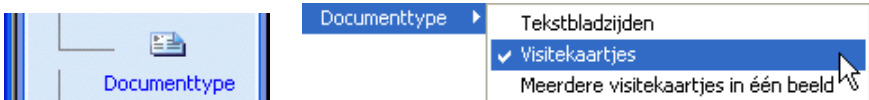
Hetzelfde geldt voor de detectie: laat de software deze klus voor u klaren, zoniet zult u alle visitekaartjes moeten roteren die ondersteboven of met een hoek van 90° op het vlakbed van de scanner werden gelegd. (Gebruik de rotatieknoppen op de beeldknoppenbalk mocht u deze optie uitgeschakeld hebben...)



Bladzijde rechtzetten

## EN DAN IS DE LEESMODUS VOOR VISITEKAARTJES...

Eens de uitgesneden kaartbeelden in het beeldvenster staan, moet u nog de lectuur van visitekaartjes mogelijk maken door de **kaartmodus** te activeren met de knop "Documenttype" (of met het equivalent bevel onder het menu "Instelling").





Kies de optie "Visitekaartjes" wanneer u de visitekaartjes één voor één scant, kies de optie "Meerdere Visitekaartjes in Eén Beeld" om meerdere visitekaartjes tegelijk te scannen op het vlakbed van uw scanner. Deze optie betreft niet enkel het segmenteren, het "opdelen" van uw scans; u moet één van beide opties inschakelen om visitekaartjes te kunnen *herkennen!* Selecteer de optie "Tekstbladzijden" om de leesmodus voor visitekaartjes terug uit te schakelen.

De knop "Documenttype" bevestigt dat de leesmodus voor visitekaartjes actief is.

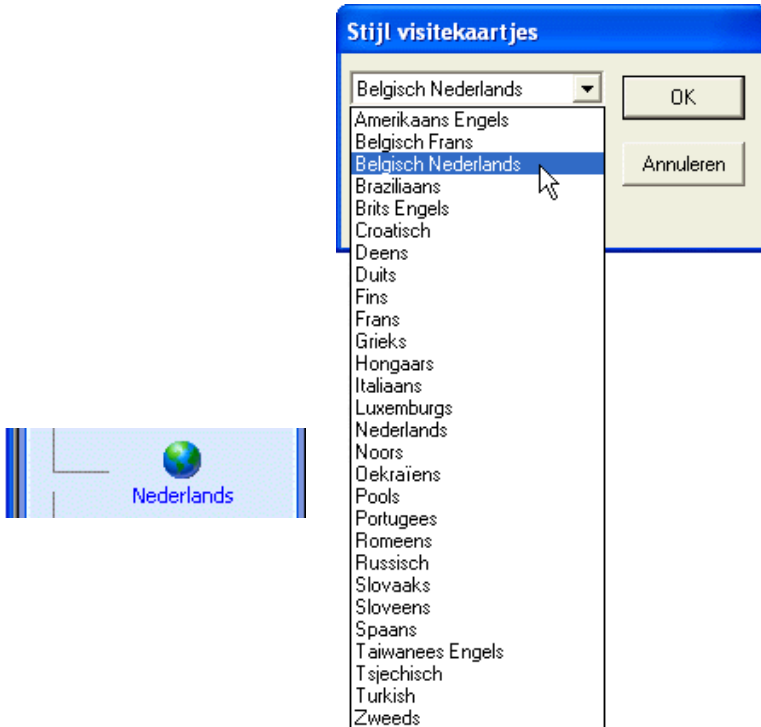


Deze speciale leesmodus is nodig omdat dan speciale OCR-routines worden gebruikt die toelaten de herkende gegevens te verdelen over de verschillende databankvelden - firmanaam, e-mailadres en website enz. Wanneer Readiris de tekenreeks "Leo Tielemans" leest, "weet" het dat Leo de voornaam is en Tielemans de achternaam. Op dezelfde wijze is het systeem in staat titels te herkennen, zoals "vice-voorzitter" of "ingenieur", steden zoals Zwolle en Antwerpen, provincies zoals Gelderland en Brabant enz.

Readiris ondersteunt wel 28 landen: de Noord- en Zuid-Amerikaanse visitekaartjes en kaartjes uit de Europese landen, inclusief de Oost-Europese landen, worden ondersteund. (Met de "Aziatische BCR-module" kunt u kaartjes uit de volksrepubliek China, Japan, Korea en Taiwan herkennen. Visitekaartjes uit deze Aziatische landen die in het Engels zijn opgesteld worden wel gelezen door de "standaard" Readiris Corporate software!)

Van zodra de kaartmodus ingeschakeld is, wijzigt de lijst van ondersteunde "talen". Kies altijd het correcte land.





Door het land van het visitekaartje aan te geven op de knoppenbalk, kiest u niet enkel de taal van de te herkennen tekst, maar ook de algemene lay-out, de stijl van het visitekaartje. Ziet het kaartje er Nederlands of Frans uit? Elk land heeft zijn eigen “stijl” om visitekaartjes samen te stellen. Amerikanen stellen een adres anders samen dan Belgen, Nederlandse postcodes en telefoonnummers hebben een andere structuur dan hun Duitse tegenhangers enz. (Sommige kaartstijlen stemmen overeen met meerdere talen: België en Canada hebben twee officiële talen, Zwitserland heeft er drie!)



De veldanalyse voert een derde taak uit: Readiris analyseert de gelezen tekst niet alleen maar *formateert* hem ook. Het systeem filtert alle irrelevante gegevens van het visitekaartje, zelfs al spelen deze een actieve rol! Als er op het gelezen visitekaartje "Telefoon: (508) 898-42 89" staat, komt de karakterstring "5088984289" in het telefoonveld terecht. Het woord "telefoon" valt weg, al maakt Readiris er wel gebruik van om het telefoonnummer op het kaartje terug te vinden. De haakjes, het streepje en de spaties in het telefoonnummer worden eveneens gewist.

## VISITEKAARTJES HERKENNEN

---

Verwerk de visitekaartjes zoals gewoonlijk door op de knop "Herkennen-Opslaan" te klikken.

We bevelen u echter aan het uitvoerformaat te bepalen eer u de herkenning uitvoert. Klik hiertoe op de knop "Formaat".



Een aantal populaire Personal Information Managers (“PIMs”) wordt rechtstreeks ondersteund: Microsoft Outlook (Express) en uw PDA software Palm Desktop. Om uw contacten naar andere toepassingen uit te voeren, gebruikt u de “universele” formaten vCard en komma gedelimeerde data. (vCard bestanden importeert u erg makkelijk in elke toepassing die de vCard standaard ondersteunt: dubbelklik het vCard bestand en de contacten worden automatisch aan uw Windows Adresboek toegevoegd!)

Het tekstresultaat ziet er bijvoorbeeld als volgt uit wanneer u de resultaten rechtstreeks naar Microsoft Outlook (Express) stuurt.



**Kirk Y. Matsuo - Contact**

File Edit View Insert Format Tools Actions Help

Save and Close

General Details Activities Certificates All Fields

Full Name... Kirk Y. Matsuo

Job title: Product Manager Business & Home/Office Proc

Company: Seiko Instruments USA Inc.

File as: Matsuo, Kirk Y.

E-mail... kirk.matsuo@seikosj.com

Display as: Kirk Y. Matsuo (kirk.matsuo@seikosj.com)

Web page address: <http://www.seikosmart.com>

IM address:

Phone numbers

Business... 40843210

Home...

Business Fax... 408 4333261

Mobile...

Addresses

Business... 1130 Ringwood Court  
95131 San Jose California

This is the mailing address

Contacts... Categories... Private

## ON-LINE HULP RAADPLEGEN

Hier eindigt onze rondleiding van Readiris. Het kan voorkomen dat bepaalde informatie van het laatste ogenblik niet in deze handleiding is opgenomen. We raden u dan ook aan om de on-line hulp te raadplegen als u meer wilt weten over Readiris.

Ga hiervoor naar het "Hulp" menu. Het commando "Hulphema's" en zijn versnellingstoets F1 laten u navigeren door de vele hulpbladzijden.

**Readiris Hulp**

Hide Back Forward Home Print Options

Contents Index Search

2 Welkom bij de Readiris hulp

- OCR inleiden
- Documenten herkennen
- Handgeschreven tekst herkennen
- Streepjescodes herkennen
- Hoe...?
- Referentie-informatie
- Software-versies en opties
- Productregistratie
- Productondersteuning
- I.R.I.S.

**Welkom bij de Readiris™ hulp...**

- Gebruik de on-line hulp om meer te vernemen over Readiris.
- Vind snel een antwoord op uw vragen.
- Verbindt u met de I.R.I.S. website voor de laatste info en product-updates.

©2005 Auteursrechten [I.R.I.S.](#) Alle rechten voorbehouden

De andere bevelen uit het menu "Hulp" vertellen u hoe u op onze productondersteuning kunt beroep doen, hoe u I.R.I.S. kunt contacteren, geven u rechtstreeks toegang tot de homepage van I.R.I.S. enz.

