

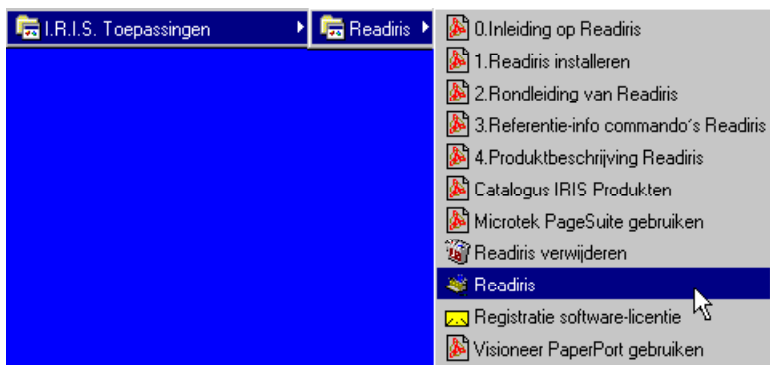
# Hoofdstuk 2

## EEN RONDLEIDING

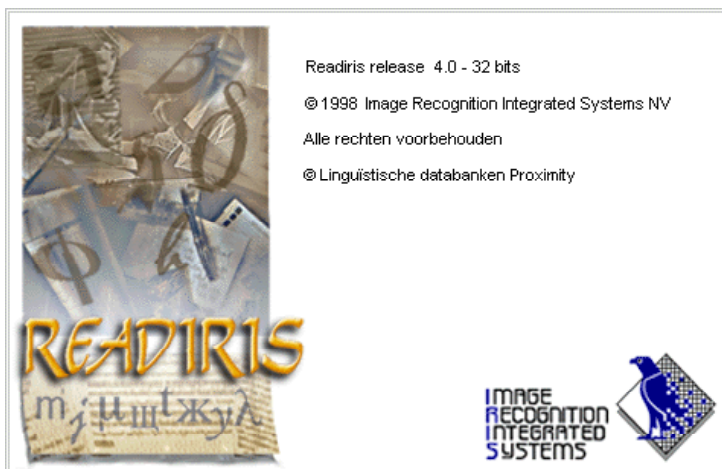
Readiris is een geavanceerd OCR-pakket dat ontelbare krachtige functies bevat. We zullen in dit hoofdstuk deze mogelijkheden doorlopen en vele tips betreffende het gebruik van Readiris toevoegen.

### 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. Door met de muis om het even waar te klikken verdwijnt dat scherm.



## 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 software-licentie 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.

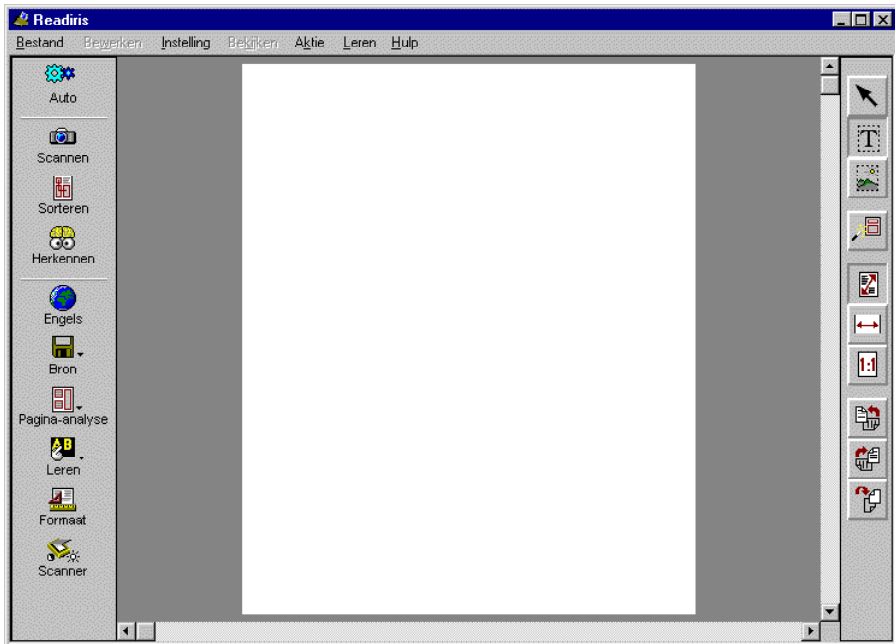
U heeft uw **identificatienummer** nodig om de softwaresleutel te kunnen genereren; zorg ervoor dat u dit nummer onder uw bereik hebt of vermeldt wanneer u uw licentie registreert.



## **DE READIRIS GEBRUIKERSINTERFACE VERKENNEN**

---

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



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

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



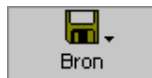
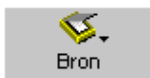
Tenslotte is er de **beeldzone** waar de gescande beelden worden getoond. U kunt beeldbestanden naar deze zone slepen om ze te herkennen.

## 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 laat u toe om beelden in TIFF formaat (niet gecomprimeerd of packbits, Groep 3 en Groep 4 gecomprimeerd), Paintbrush beelden (PCX) en Windows bitmaps (BMP) te openen. Alleen zwart-wit beelden worden ondersteund. Deze mogelijkheid is overigens bijzonder nuttig om **faxen** te converteren in editeerbare tekstbestanden.

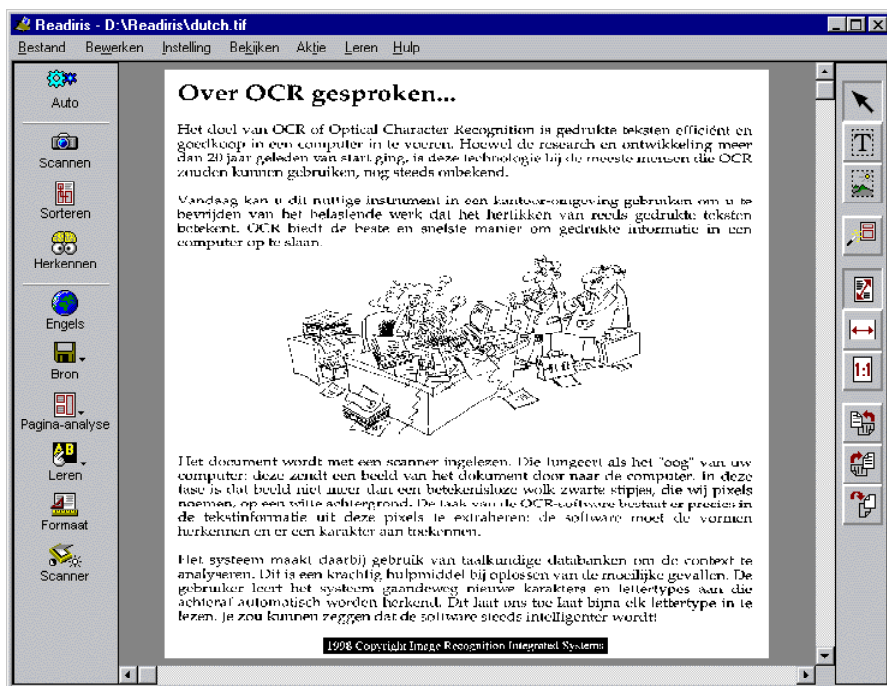
Om een vooraf gescand beeld te openen moet u de schijf, en niet de scanner, als beeldbron kiezen met de knop "Bron".



Daarna klikt u op knop "Scannen". U kunt ook het commando "Openen" onder het menu "Bestand" gebruiken om een vooraf gescand beeld te openen – dat werkt zelfs wanneer uw scanner fungeert als beeldbron.



U wordt uitgenodigd om een beeldbestand te kiezen. Selecteer het bestand DUTCH.TIF in de Readiris map. Dit beeld wordt van de harde schijf gelezen en afgebeeld in de beeldzone van Readiris.



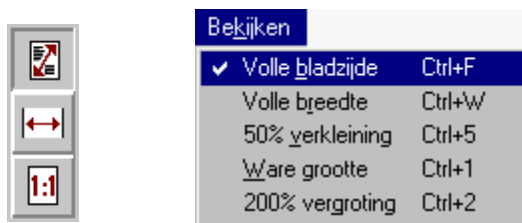
Een derde manier om vooraf gescande beelden te openen is het gebruik van “drag en drop”. Sleep een beeldbestand van de Windows Verkenner naar de beeldzone van Readiris en het beeld wordt onmiddellijk geopend.

De rechter knoppenbalk bevat alle opdrachten 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 beeld knoppenbalk 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 opdrachten en voegt twee extra zoomniveau's toe: u kunt het beeld laten afbeelden op 50 % en 200 % van zijn werkelijke grootte. Bij de werkelijke grootte beantwoordt één beeldpixel aan één scherpixel. (Er zijn versnellingstoetsen beschikbaar voor alle de zoomniveau's!)



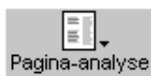
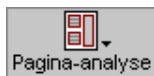
U kunt ook met de *rechter* muisknop klikken om op een deel van het gescande beeld in te zoomen op werkelijke grootte. Klik de rechter muisknop een tweede keer om opnieuw uit te zoomen.



## EÉN, EEN GESCAND BEELD ANALYSEREN

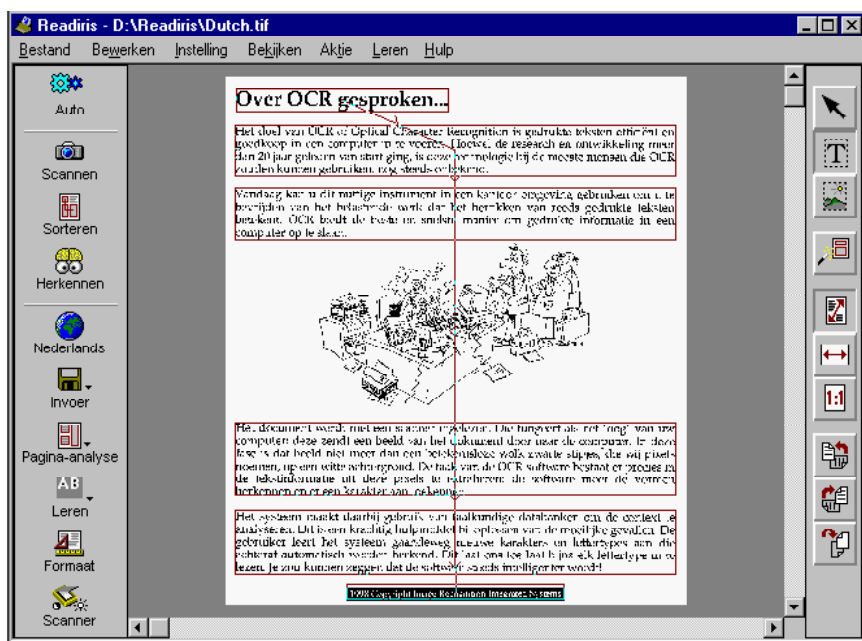
Nu het beeld gescand is, kunt u aangeven welke delen u wilt omzetten in editeerbare tekst door kaders, zogenaamde “vensters”, rond de zones van belang te tekenen.

Overigens kan Readiris dit automatisch voor u doen wanneer u de optie "Pagina-analyse" activeert op de hoofdknoppenbalk.

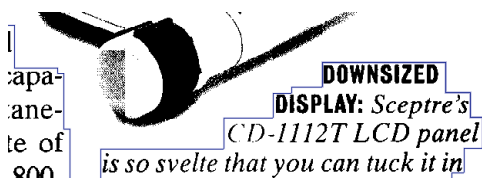




De automatische pagina-analyse is bijzonder nuttig wanneer u **in kolommen** geschikte tekst en documenten met een ingewikkelde layout moet herkennen.



De pagina-analyse verloopt snel, springt tolerant om met lijnkromming en is erg akkuraat: de analyse vindt complexe, “onregelmatige” vormen terug maar vermijdt de grafische zones die fotos, illustraties, enz. bevatten.





Merk op dat de pagina-analyse zelfs die zones detecteert met **witte tekst op een zwarte achtergrond**. Zulke tekstkaders herkennen vormt geen probleem: terwijl de preview het gescande document getrouw op het beeldscherm toont, 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.)

Het aantal vensters is altijd aangegeven in de tooltip van de tools "Tekstvenster" en "Grafisch Venster".



Readiris detecteert niet alleen de tekstblokken, maar *sorteert* die ook: standaard worden de venster 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 hoofdknoppenbalk en klikt u op de vensters die u wenst in te voegen. De vensters die u *niet* aanklikt zullen simpelweg niet herkend worden.



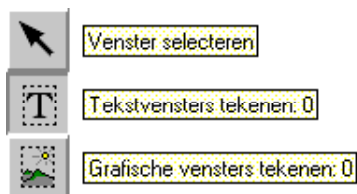
Standaard is de pagina-analyse geactiveerd. Om Readiris te dwingen de huidige pagina te analyseren – u hebt bijvoorbeeld de pagina-analyse per ongeluk uitgeschakeld, u heeft sommige vensters gewist en nu wilt u de pagina-analyse herhalen enz. -, klikt u op de knop "Pagina Analyseren" in de beeld knoppenbalk.



## **TWEE, BEELDEN MANUEEL BEVENSTEREN**

---

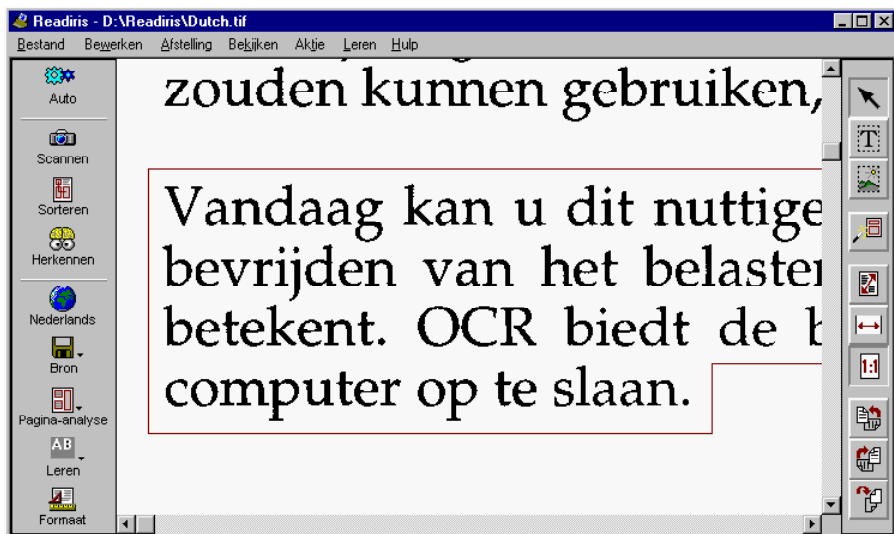
De pagina-analyse is de automatische manier om gescande pagina's te bevensteren. Maar u kunt de zones van een beeld ook handmatig aangeven met de **venstertools** van Readiris!



Om een vierkanten kader te **creëren** rond een tekstkolom, selecteert u het "Tekstvenster" in de beeldknoppenbalk. Klik de cursor in de linker boven hoek van het venster, trek het venster open door de muis te verschuiven naar de tegenovergestelde hoek rechts onder en klik nogmaals. (Zijden kleiner dan 1 mm zijn niet toegelaten, ze zouden overigens niet één karakter bevatten.)

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

U kunt ook "onregelmatige" tekstblokken tekenen door **polygone vensters** rond dergelijke tekstzones te trekken. Niet-vierkante vensters creëert men door vierkanten vensters te versmelten: van zodra twee vierkanten (van hetzelfde type) elkaar raken, worden ze automatisch omgezet in één venster! Het komt hierop neer dat u een huis bouwt door de ene kamer na de andere toe te voegen...



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 beeld knoppenbalk en klikt u op het venster. Vierkanten markeerpunten verschijnen op elke hoek en in het midden van de vensterzijden.



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 het andere type), houdt u de Ctrl toets ingedrukt terwijl u op het hoofdvenster klikt.

Merk op dat u het venstertype, de cöördinaten in pixels en de sorteervolgorde van een venster kunt zien door op de Alt toets te drukken terwijl u het venster selecteert!



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 om de afmetingen te veranderen.

Om een tekstvenster te **verplaatsen**, selecteert u het en sleept u het naar een andere lokatie.

Om vensters te **wissen**, selecteer u ze en kiest u het commando "Knippen" of "Wissen" van het menu "Bewerken". Het commando "Knippen" plaatst het venster(s) in een interne buffer, het commando "Wissen" wist 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 lokatie.



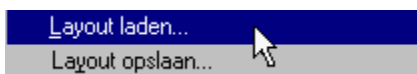
Merk op dat *alle* klassieke opdrachten van het menu "Bewerken" van toepassing zijn op de vensters: vensters kunt u wissen, knippen, kopiëren en plakken! Het commando "Annuleren" speelt ook zijn rol: u komt op uw stappen terug wanneer u per ongeluk vensters heeft gewist, verplaatst, vergroot, enz.



Merk tevens op dat er versnellingstoetsen beschikbaar zijn voor alle opdrachten! 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 vensterlayout aan te maken. Om de vorige layout te herstellen, kunt u "Annuleren" kiezen of zijn versnellingstoets Alt+Backspace.

## DRIE, VENSTERSJABLONEN OPSLAAN

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



Als u dokumenten met een zelfde layout gaat 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 vensterlayout blijft van kracht totdat u de pagina-analyse opnieuw activeert in de hoofd knoppenbalk.

## **MET READIRIS KUNT U DE HELE WERELD ROND!**

---

In de veronderstelling dat de vensters correct gedefiniëerd zijn, bent u nu bijna klaar om de karakterherkenning uit te voeren. We zeggen “bijna”, want u heeft de taal en de instellingen van het dokument nog niet gecontroleerd.

De taalinstelling komt voor in de hoofdknoppenbalk.



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



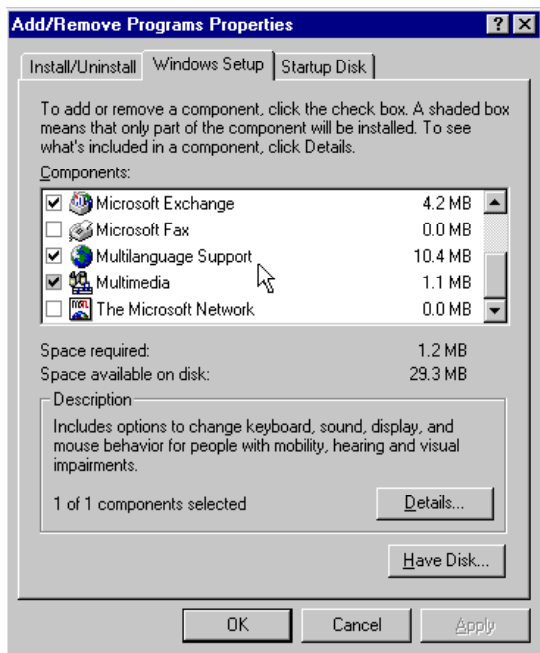
U kunt een lettertoets indrukken om onmiddellijk naar een taal te gaan: als u momenteel Nederlands geselecteerd hebt en u wilt Hongaars selecteren dan drukt u op de "H" toets op uw klavier om onmiddellijk naar het Hongaars te gaan. Wanneer verschillende talen dezelfde beginletter hebben drukt u herhaalde malen op deze toets om de verschillende opties te doorlopen. Bijvoorbeeld: Readiris leest Portugees en Pools. Door éénmaal op "P" te drukken, selecteert u het Pools, door een tweede maal op "P" te drukken selecteert u het Portugees en door een derde maal op de "P" toets te drukken keert u terug naar het Pools.

Readiris is geenszins beperkt tot het Nederlands: er worden wel **52 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.

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 zijn.



Merk op dat u een aangepaste configuratie van Windows nodig hebt om de Centraal-Europese, Griekse, Turkse, Cyrillische en Baltische karakters te ondersteunen. U kunt één en ander installeren met de Windows "service pack" "Veeltalige Ondersteuning". Selecteer het pictogram "Software Toevoegen/Wissen" onder het "Controlepaneel" om uit te vinden of deze Windows module al op uw PC geïnstalleerd is. Naargelang de software-bundle die u heeft aangekocht, bevat de Readiris CD-ROM mogelijkerwijze gedetailleerde informatie over de installatie van deze Windows module.



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



veeltalige ondersteuning garandeert dat u “exotische” karakters ß, ñ, γ en ø 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 waar omwille van een inktvlek de "r" eerder op een "f" lijkt. 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. Dit **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 "I" 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 "I" of een '1'.

De onderstaande illustratie laat verschillende exemplaren zien van ééntjes en 1-symbolen. De vormpjes op de eerste lijn zijn niet dubbelzinnig, de vormpjes 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

## **READIRIS VERANDERT TUSSENDOOR VAN TAAL**

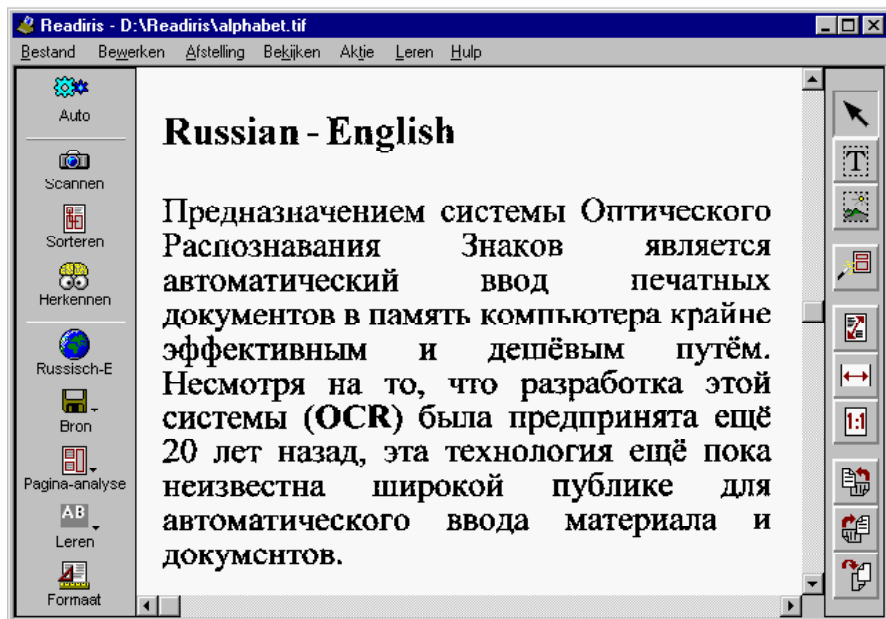
---

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 en Cyrillische 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 of Cyrillische karakters in combinatie met Latijnse karakters.

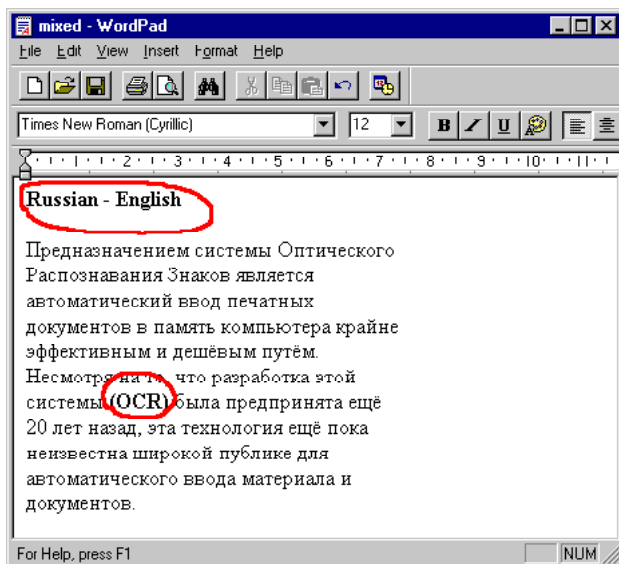
Wees er zeker van dat u "Grieks-Engels" geselecteerd hebt of een gepaste Cyrillische taalinstelling - bijvoorbeeld "Oekraïens-Engels". Met andere woorden: selecteer niet gewoon het "Grieks" of "Oekraïens" als taal van het dokument in de hoop dat de Westerse symbolen wel goed door de herkenning zullen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



Dit is een voorbeeld waar een Russische tekst enkele Engelse woorden bevat – open het beeld ALPHABET.TIF mocht u het zelf willen proberen!



Wanneer u het resultaat met de tekstverwerker opent ziet het er zo uit - het kan gebeuren dat u een Cyrillisch **lettertype** moeten activeren om de Russische tekst correct af te beelden.



Om **andere talen** te **mengen** moet u die taal selecteren met de meest uitgebreide karaktertabel. Stel dat u een dokument heeft waar de Franse vertaling naast de Nederlandse tekst staat: u moet het Frans selecteren om ervoor te zorgen dat de geaccentueerde karakters zoals ç, é en ù correct herkend worden.

## DE KENMERKEN VAN HET DOKUMENT AANGEVEN

Nu de taal is ingesteld, gaan we over naar de andere kenmerken van het dokument. U kunt de herkenning aanpassen door enkele karakteristieken van het dokument aan te geven: het type lettertekens en de karakterafstand. Laten we verhelderen wat dit alles betekent.

Eerst en vooral geeft u met het commando "Type Lettertekens" onder het menu "Instelling" aan of u "normale" of **matrixgeprinte** documenten wilt herkennen.

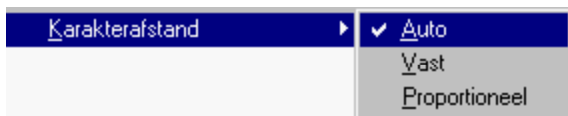


Het is noodzakelijk deze parameter correct in te stellen: "draft" dot matrix symbolen - ook wel "9 pin" matrixdruk genoemd - zijn opgemaakt uit aparte vlekken, en er zijn sterk gespecialiseerde herkenningstechnieken nodig om dergelijke tekst te lezen.

**ape-descended life**

Matrixdruk met "letterkwaliteit", ook wel "25 pin" matrixdruk genoemd, vereist de optie "Normaal", en dat is ook het geval voor de **drukkwaliteiten** gedrukt, getikt, laserdruk en inkjetdruk. "Normaal" is overigens de standaard waarde.

De **karakterafstand** kan gekozen worden met de opdracht "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.



Courier, een vast lettertype  
Letter Gothic, een vast lettertype

Century Schoolbook, een proportioneel lettertype  
Arial, een proporioneel lettertype

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

## **READIRIS WORDT TELKENS INTELLIGENTER!**

Wanneer de taal en de kenmerken van het dokument aangegeven zijn klikt u op de knop "Herkennen".



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



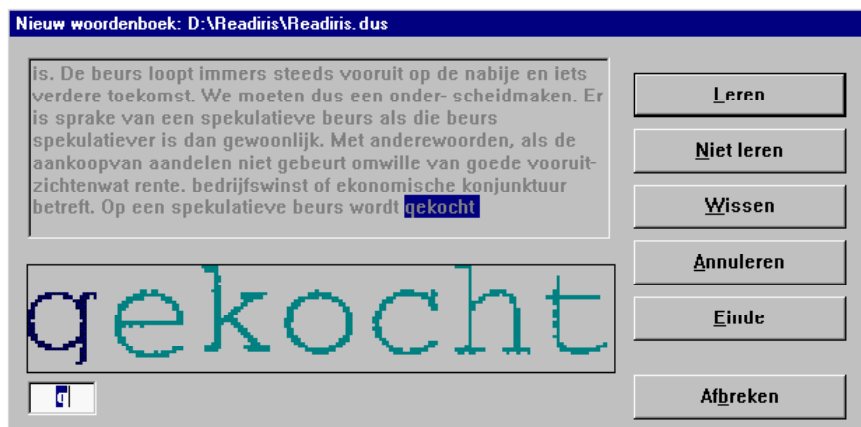
Indien de leermodus geactiveerd werd met de knop "Leren" op de hoofdknoppenbalk komt Readiris tegen het einde van de herkenning aan terecht in de interactieve fase i. Het interactief leren is standaard geactiveerd.



De leermodule kan in bepaalde gevallen de herkenning sterk verbeteren. Wanneer de gebruiker probeert om gebroken, misvormde lettertekens in te lezen zoals men die terugvindt in echte documenten, of wanneer men op gestileerde lettertekens 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 wiskundige 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 karakterstrings. De lettertekens 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 -, activeer 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 karakterstring) in voor de verkeerde of niet herkende tekens en klik op één van de volgende knoppen.

## Leren

U gaat akkoord met het voorgestelde oplossing of u verbetert ze. 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 op te nemen 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 "onzekeer" 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 "e" hierboven is serieus beschadigd - in feite lijkt hij op een "c"-, en u klikt hier op "Niet Leren" om dit symbool niet met de "c" te verwarren.



## **Wissen**

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

## **Annuleren**

U keert terug om fouten te verbeteren. Readiris houdt de negen 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.

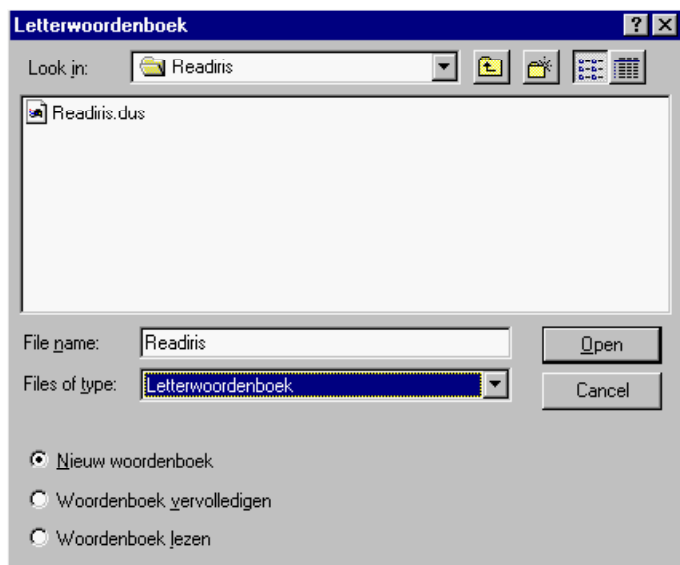
Verwar "Einde" niet met "Afbreken": bij "Afbreken" wordt er geen uitvoer gegenereerd en moet u de herkenning overdoen, met de knop "Einde" wordt de tekst wel aangemaakt, alleen neemt u die niet in detail door.

## **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. Deze woordenboeken 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 dokumenttype. Om het allemaal duidelijk te houden geeft u het best heldere namen aan deze woordenboeken, bijvoorbeeld RAPPORT.DUS, PALATINO.DUS enz. 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.)

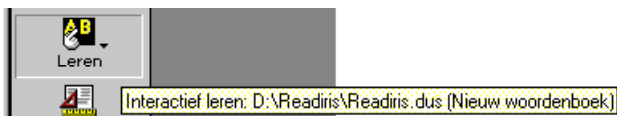
De uitbreidingsmodus geeft aan dat de trainingresultaten zullen bewaard worden in een *bestaand* woordenboek: de herkenning maakt al gebruik van de extra intelligentie die in het woordenboek aanwezig is, en u voegt nieuwe vormen aan het woordenboek toe. Eenvoudig gesteld kunt u met deze optie een woordenboek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

(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 drukkwaliteit - 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. Wanneer meerbladige documenten herkend worden, wordt het woordenboek automatisch in de uitbreidingsmodus gezet van zodra u een volgende bladzijde inscant.

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

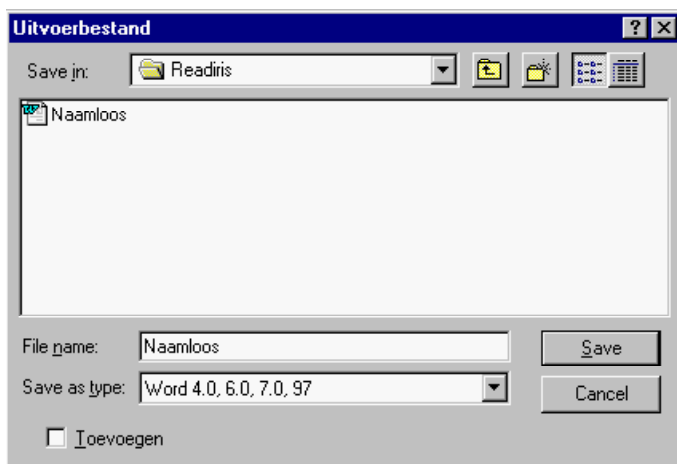


Wanneer u in het interactief leren terecht komt worden het woordenboek en zijn modus aangegeven in de titel van het leervenster, en u klikt op "Annuleren" om opnieuw te beginnen wanneer deze foutief blijken.



## EN WAAR MOET HET HEEN MET DE TEKST?

Het interactief leren beëindigt de karakterherkenning. U zult uitgenodigd worden om de herkende tekst te bewaren. (Als u de tekst naar het klembord stuurt, gebeurt dat uiteraard niet.)



U kunt een nieuw bestand aanmaken of de uitvoer toevoegen aan een bestaand tekstbestand – vandaar de knop "Toevoegen". U kunt bijvoorbeeld twee gescande bladzijden tekst toevoegen aan een studie die u zelf geschreven heeft. Wanneer u tekst toevoegt aan een bestaand bestand, wees er dan zeker van dat die tekst op dat ogenblik niet geopend is want zoiets belet u in dat bestand te schrijven!

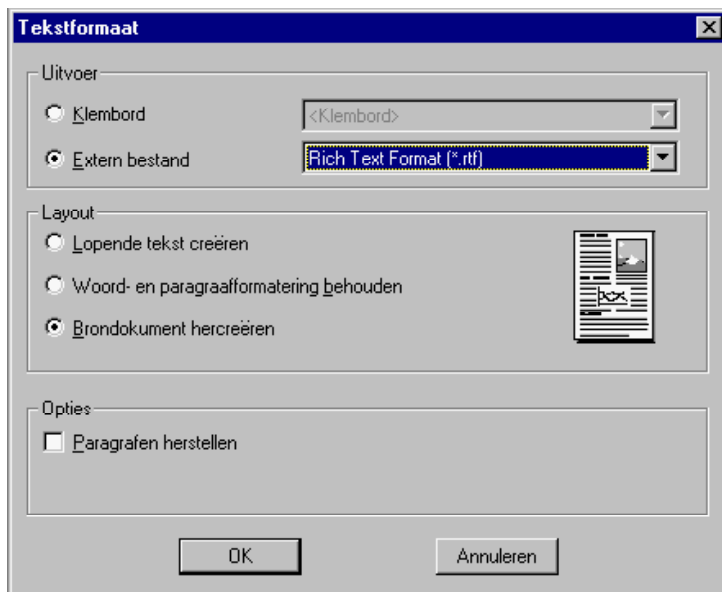


Een tekst opslaan houdt meer in dan een bestandsnaam opgeven voor tekstuitvoer. U moet ook een tekstformaat selecteren en de layout van de herkende tekst gaan bepalen. Kortom, u moet beslissen waar u met de tekst heen wilt eer u de herkenning uitvoert.

Het **tekstformaat** kunt u niet selecteren in dit venster, op het ogenblik dat u een tekst bewaart – u moet dit doen met de "Formaat" knop (of met het commando "Tekstformaat" onder het menu "Instelling") *vooraleer* u de OCR start.



Dit commando laat u het formaat kiezen voor de uitgevoerde tekst, maar vergeet niet dat u de uitvoer kunt kopiëren naar het klembord van Windows. Er is dus geen strikte noodzaak om het resultaat op te slaan in een extern bestand!



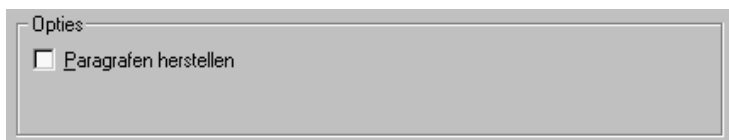


## DE UITVOER OPMAKEN

---

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

De “**tekstflow**” van het uit te voeren dokument wordt onmiddellijk beïnvloed door de optie "Paragrafen Herstellen" die u vindt onder de "Formaat" knop.



Houdt deze optie geactiveerd 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 vlote, continue tekst.

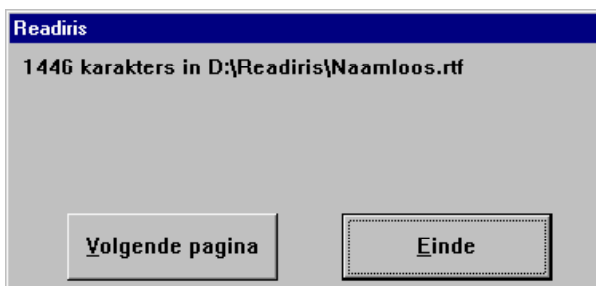
Was de paragraafdetectie *niet* geactiveerd, dan zou de originele layout behouden zijn, met een carriage return (Enter) aan het einde van elke zin.

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



## MEERDERE PAGINA'S HERKENNEN

Wanneer u op "Opslaan" klikt om een tekst te bewaren, zal het systeem u uitnodigen om een volgende pagina in te lezen. (Wanneer u de tekst naar het klembord gestuurd heeft, zult u natuurlijk niet uitgenodigd worden om de tekst op te slaan.)



Er wordt nu een **statistisch gegeven** vermeld - het aantal karakters dat u heeft ingelezen. Selecteer "Einde" om de eerste oefening te beëindigen. 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 kunt nu de herkende tekst openen met uw tekstverwerker of teksteditor, in uw DTP-software of een andere tekstapplicatie invoeren. U heeft inderdaad een dokument omgezet in een editeerbaar computerbestand, en dit wel 10 tot 25 maal vlugger dan indien u de tekst had hertikt! Ga uw gang en vergelijk het resultaat met het beeld binnen het Readiris venster.

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: klik op "Volgende Pagina" wanneer u uitgenodigd wordt om een volgende bladzijde te lezen. Als uw scanner als beeldbron fungeert scant u een ander blad. Tevens worden het uitvoerbestand en het woordenboek



automatisch in de uitbreidingsmodus geplaatst zodat u de tekst kunt aanvullen en de training comfortabel kunt verderzetten.

U kunt rustig verdergaan met het herkennen van andere pagina's: de nieuwe tekst zal telkens aan het bestand worden toegevoegd na de herkenning.

## UW SCANNER INSTELLEN

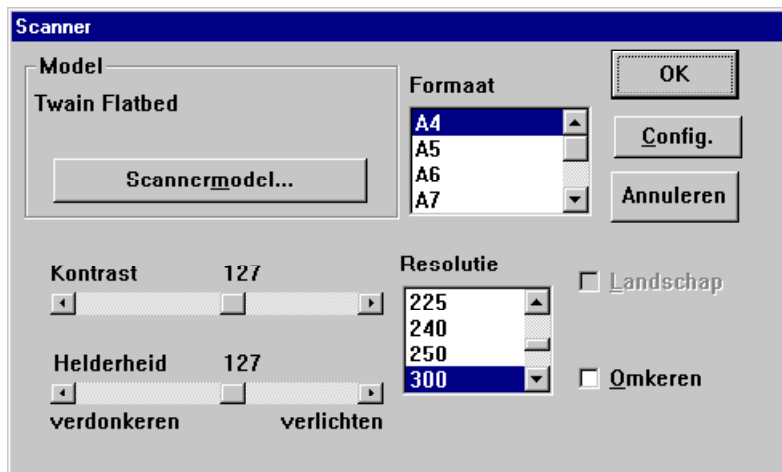
Laten we uw scanner gaan instellen. We veronderstellen nu wel dat uw scanner en de nodige stuurbestanden goed werden geïnstalleerd.

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.

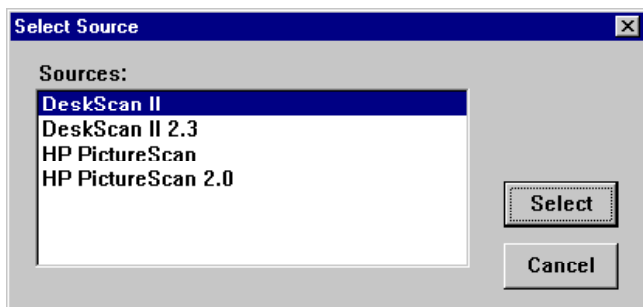


Klik op de knop "Scannermodel" om uw **scannermodel** te kiezen.



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 "Config." 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".)



Uw scanner kan ook vergezeld zijn van een **scanplatform** zoals Visioneer Paperport of HP Document Assistant. Afhankelijk van de software-versie die u aangeschaft heeft, zult u gespecialiseerde handleidingen betreffende de scanplatformen terugvinden op de Readiris CD-ROM.

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 dokument feeder wilt gebruik maken. Bij Twain scanners worden alle instellingen meestal binnen de Twain interface geregeld.

Kies een **resolutie** van 300 dpi voor normale applicaties en de hogere resolutie van 400 dpi voor kleine druk (kleiner dan 10 punt) en wanneer het dokument erg gedegradéerd is.

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



6 punt

72

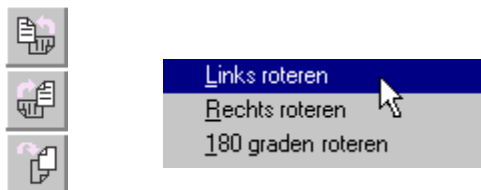
Readiris leest eveneens “**dropletters**”, hoofdletters die meerdere lijnen beslaan. (Die kunnen uiteraard niet groter zijn dan 72 punt!)

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

Door de optie "Landschap" te selecteren geeft u aan dat de pagina-oriëntatie breed is en niet lang (“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 landschaporientatie scannen, maar volle A4 bladzijden kunt u natuurlijk alleen in één richting inscannen!



Wanneer u de pagina-oriëntatie moet aanpassen, kunt u gebruik maken van de rotatieknoppen op de beeld knoppenbalk. (Overeenkomstige opdrachten bevinden zich in het menu "Bekijken".)



Er zijn drie rotatiemogelijkheden: rotatie naar links, naar rechts en 180°. De rotatie neemt een paar seconden in beslag: het achterliggende beeld wordt herschreven, en niet enkel de afbeelding ervan op het scherm.

Stel de **helderheid** in, en, wanneer dat mogelijk is, **het contrast**. De contrastinstelling is enkel beschikbaar op sommige modellen. Sleep de glijbalk om het scannen donkerder of helderder te maken; de geselecteerde waarden worden ook numeriek voorgesteld.

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

## 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.



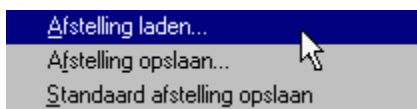
Configuratiebestanden bevatten meer dan enkel de scanopties: ze bepalen ook of u het interactief leren gaat gebruiken, welk woordenboek actief is (en zijn modus), welke taal en lettertype – bijvoorbeeld normale, proportionele letters – de documenten bevatten enz. Kortom, *alle* operationele parameters van Readiris worden in de configuratiebestanden bewaard.



## SPECIFIEKE INSTELLINGEN BEWAREN

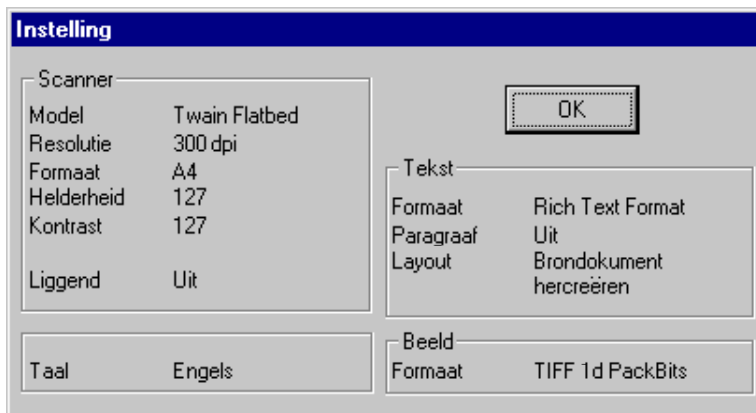
---

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.



Laten we een voorbeeld geven: als u regelmatig Nederlandse documenten moet herkennen met een specifieke layout, dan wordt u aangeraden om een configuratiebestand aan te maken voor dit dokumenttype. U zult dan met name "Nederlands" als dokumenttaal selecteren, een aangepaste vensterlayout 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 instellingen, 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.



## BLADZIJDEN AUTOMATISCH HERKENNEN

Nu de scanner ingesteld is, kunnen we van start gaan met het scannen van documenten. Vergeet niet uw scanner te selecteren als beeldbron met de "Bron" knop!



In plaats van de OCR stap per stap uit te voeren, zullen we een bladzijde volautomatisch herkennen. We klikken hiervoor op de "Auto" knop in de plaats van op de knop "Scannen".



Automatische herkenning houdt in dat de pagina achtereenvolgens wordt gescand, bevensterd door de pagina-analyse of aan de hand van een venstersjabloon en herkend zonder interactief leren. De gebruiker beperkt er zich



toe het scannen te starten en de herkende tekst op te slaan, alle tussenliggende stappen neemt Readiris voor zijn rekening.

## **READIRIS HERSCHEPT DE DOKUMENT LAYOUT**

---

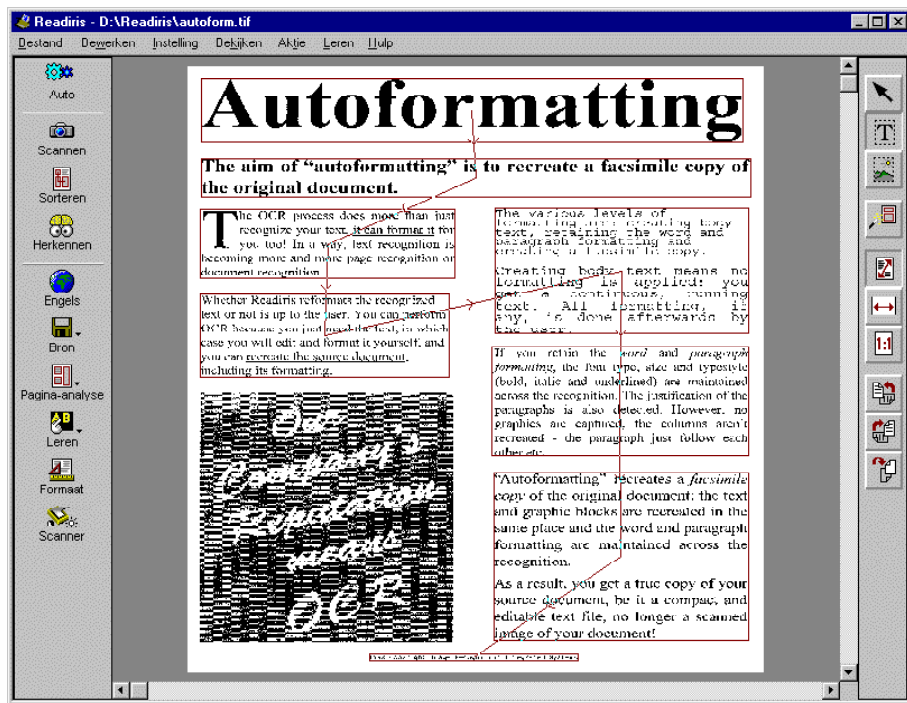
Automatische herkenning, waarbij het herkenningsproces automatisch doorlopen wordt, mag u geenszins verwarren met automatische formattering! “Autoformattering” houdt met name in dat Readiris een **getrouwe copie** aanmaakt van het gescande document: de woord-, paragraaf- en bladformattering van het oorspronkelijk document keren terug in de uitvoer.

Gelijkaardige lettertypes als in het brondokument (met schreef of schreefloos, proportioneel of vast) worden gebruikt, de puntgroottes en stijlen (vet, cursief en onderlijnd) worden behouden tijdens de herkenning. De Tab-codes en de uitlijning (links en uitgelijnd) van elke tekstblok worden gerecreëerd. De plaats van de kolommen, tekstblokken en grafieken baseert zich op het origineel document. Met andere woorden, Readiris laat u toe om een copie te archiveren van uw dokumenten, zij het dat u dit keer te maken hebt met een editeerbaar, compact tekstbestand, en niet langer met een gescand beeld!

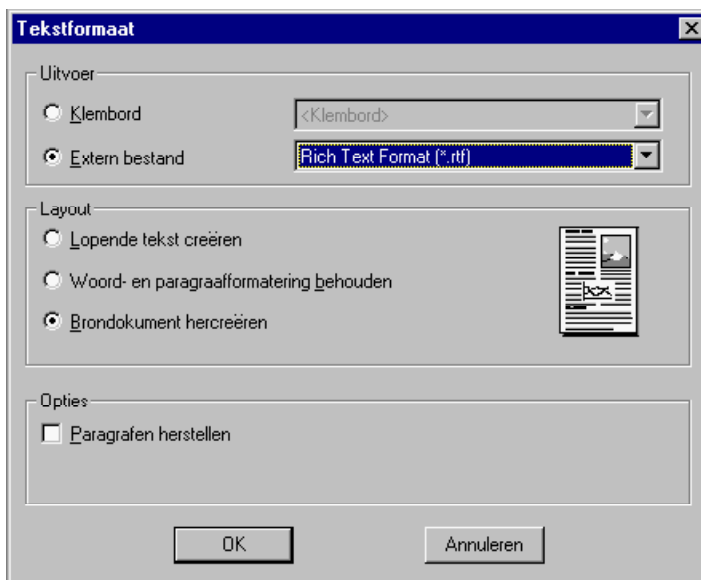
Dit alles houdt in dat bij “autoformattering” de sortering van vensters *niet* geldt: als u een manuele sorteerorde hebt aangebracht op het resultaat van de pagina-analyse, wordt die genegeerd!

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





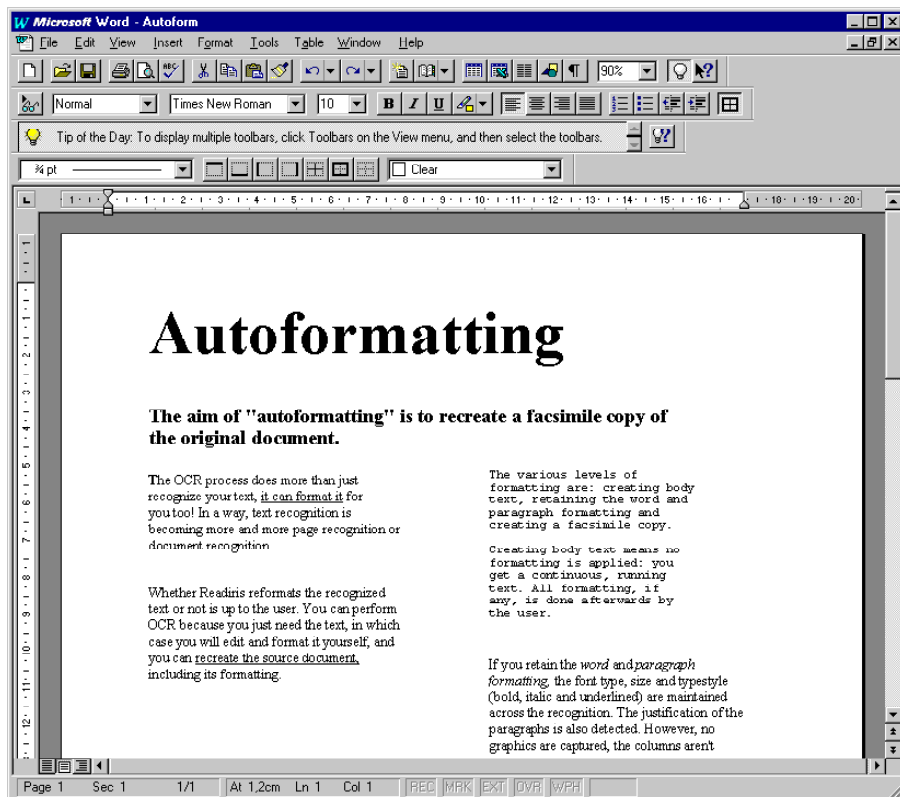
Klik op de "Formaat" knop op de hoofd knoppenbalk en selecteer het uitvoerformaat "RTF" (Rich Text Format) en de layoutoptie "Brondokument Hercreëren". (Het is uw vrije keuze of u de optie "Paragrafen Herstellen" al dan niet activeert om binnen de paragrafen wordwrap toe te passen.)



Noteer dat de layoutreconstructie alleen met het RTF formaat mogelijk is. Andere, “arme” tekstformaten die “platte” tekst genereren zoals ASCII, ANSI enz. ondersteunen geen geavanceerde formatteringscodes en kunnen dus ook *geen* “autoformattering” bieden.

En nu het goede nieuws: RTF is een wijd verspreid tekstformaat dat door alle populaire tekstverwerkers kan geopend worden. (Besef wel dat WordPad een sterk gereduceerde teksteditor is, eerder dan een volledig uitgeruste tekstverwerker; WordPad opent wel RTF bestanden maar negeert de meeste formatteercodes zoals tekstblokken, alignering en zo meer.)

Wanneer de herkende tekst geopend wordt met de tekstverwerker ziet hij er als volgt uit zonder *enige* interventie van de gebruiker. Om de layout correct te beoordelen moet u wel de “WYSIWIG”-modus van uw tekstverwerker activeren; meestal is er sprake van een “pagina-layout” modus.



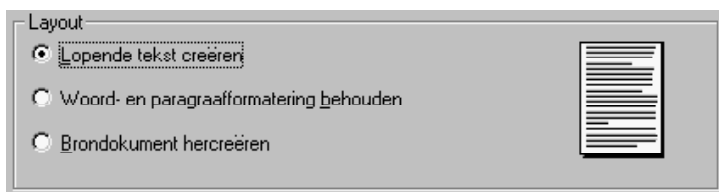
Samengevat, Readiris herkent niet alleen uw teksten, maar kan ze ook formatteren. OCR is tegenwoordig meer dan tekstherkenning, OCR wordt meer en meer blad- of **dokumentherkenning**!



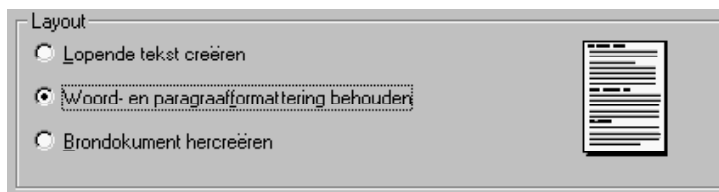
## TEKSTFORMATTERING, DEEL 2

De andere layoutopties 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 creëert. De tekst wordt hernomen, maar van zijn layout wordt abstractie gemaakt. Gebruik deze optie wanneer u gewoon een tekst wilt overnemen maar geen belang hecht aan de layout.



De optie "Woord- en Paragraafformattering Behouden" is de midden weg: de **woordformattering** – het lettertype (met schreef - schreefloos, proportioneel – vast), de puntgrootte en stijl (vet, cursief en onderlijnd) – wordt behouden gedurende de herkenning. Dat is ook het geval met de **paragraafformattering** – de tabs en de uitlijning (links en gealigneerd).



Verwar deze formatteringsoptie echter niet met de “volle” autoformattering: deze optie plaatst de paragrafen gewoon na elkaar, ze creëert geen kolommen of de relatieve positie van de verschillende tekstblokken. Evenmin zullen de grafieken automatisch ingesloten worden binnenin het RTF bestand zoals de volle “autoformattering” dat wel kan.

## GRAFIEKEN AAN HET TEKSTBESTAND TOEVOEGEN

In het vorig voorbeeld was de grafiek niet opgenomen in het tekstbestand: we kregen een lege ruimte waar de illustratie voorkwam. Dat is normaal: de pagina-analyse negeert alle grafieken aangezien die geen tekst bevatten. Maar aangezien het onze bedoeling was een conforme copie te produceren van het oorspronkelijk dokument zullen we de herkenning herhalen, alleen zullen we deze keer de grafiek insluiten.

Teken hiertoe een grafisch venster rond de cartoon. Dat gebeurt op dezelfde wijze als wanneer u tekstvensters tekent, alleen selecteert u nu de tool "Grafisch Venster".



Zoals de tooltip van het grafisch venster ons vertelt, hebben we één grafisch venster op het beeld.



U kunt de Alt knop ingedrukt houden terwijl u een venster selecteert om zijn type en coördinaten te tonen. (In tegenstelling tot de tekstvensters hebben de grafische vensters geen sorteerorde.)



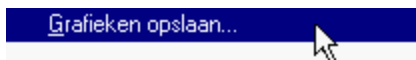
Voer de tekstherkenning opnieuw uit met de layoutoptie "Brondokument Hercreëren" geactiveerd. Wanneer u de herkende tekst opent, zult u zien dat de zwart-witte cartoon opgenomen is in het tekstbestand.

## GRAFIEKEN APART OPSLAAN

Grafieken bewaren binnen de tekst is alleen mogelijk met de "volle" autoformattering – en dat impliceert het RTF tekstformaat en de layoutoptie "Brondokument Hercreëren" -, niet bij "arme" tekstformaten zoals ASCII, ANSI enz. Maar u kunt met Readiris ook de omgekeerde weg bewandelen: grafieken opslaan zonder aan tekstherkenning te doen.

Aangezien een OCR-pakket **zwart-wit beelden** genereert, geen grijswaarden of kleuren, heeft het geen zin om foto's te scannen met Readiris, maar u kunt probleemloos zwart-wit materiaal scannen.

Hoe? Teken een grafische zone rond de illustratie, de tekening enz. die u wilt opslaan. Kies daarna het commando "Grafieken Opslaan" onder het menu "Bestand".



U zult gevraagd worden om een bestandsnaam op te geven. Bepaal welk grafisch formaat u wilt gebruiken. Selecteer een bestandsformaat dat uw fotoretouchesoftware ondersteunt. U hebt de keuze tussen TIFF en Paintbrush (PCX). (Paintbrush bestanden kunt u in elk geval openen met het interne Windows programma Paint.)

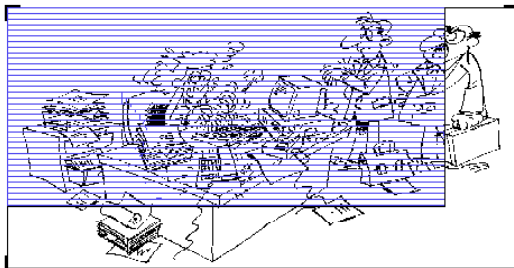
De grafieken worden altijd in één enkel bestand opgeslagen. U kunt meerdere grafische vensters tekenen, maar wanneer u dat doet, worden die samen opgeslagen in één bestand. (U kunt het "crop"-commando van uw fotoretouchesoftware gebruiken om ze te scheiden.)



De zijden mogen niet kleiner zijn dan 1 mm. – bitmaps met dergelijke afmetingen bevatten toch bijna geen informatie. “Onregelmatige”, niet-vierkanten vensters worden op onzichtbare manier omgezet in vierkanten vensters: Readiris beslaat de maximale oppervlakte op basis van de grootste coördinaten. In het onderstaand voorbeeld wordt het grafisch venster (met zwarte rand) door de software aangevuld met de gearceerde zone.



ur computer automatically.



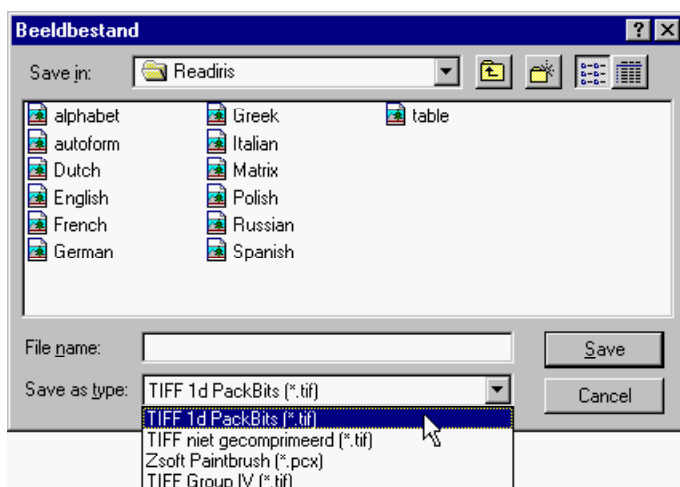
ent is read by your scanner. This device acts as the "eye" of yo

### **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 in onze oefeningen.

Teken een grafisch venster rond de hele pagina en sla de grafiek op als een TIFF bestand.



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

Readiris opent zwart-witte TIFF bestanden (niet gecompriemd of packbits, Groep 3 en Groep 4 gecompriemd), Paintbrush (PCX) beelden en Windows bitmaps (BMP). Deze mogelijkheid is vooral handig om **faxen** om te zetten in editeerbare tekstbestanden!

Als u enige invloed heeft op uw correspondenten, vraag hen dan om 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 gebruik kunt maken van "**drag and drop**" om beelden te openen: sleep een beeldbestand van de Windows Verkenner naar de beeldzone van Readiris en het wordt onmiddellijk geopend!

## TABELLEN HERKENNEN

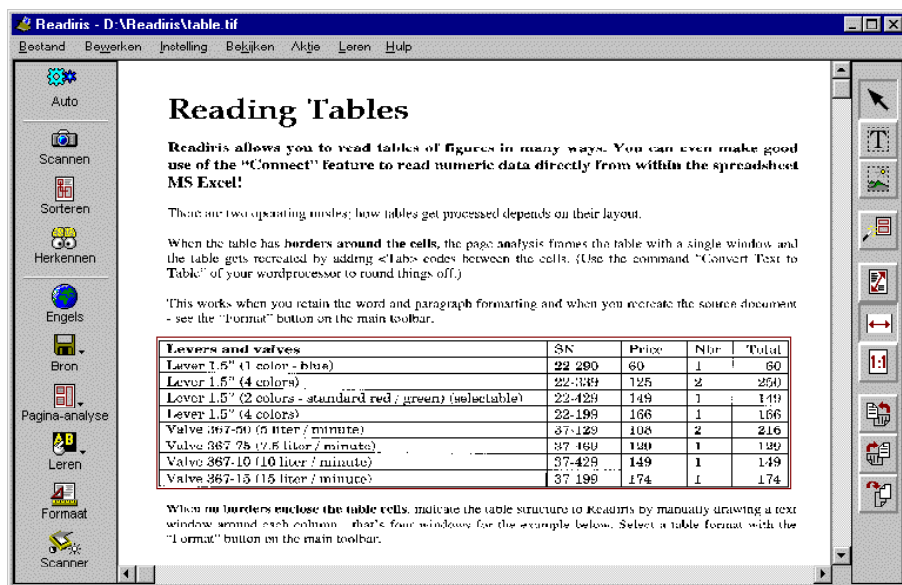
---

Tot nog toe hebben we teksten en faxen herkend en grafieken bewaard. Laten we nu een tabel gaan verwerken. Neem een tabel en scan die of open het voorbeeld TABLE.TIF in uw Readiris map.

In feite bevat het beeld TABLE.TIF twee tabellen en dat is geen toeval! Hoe u tabellen verwerkt hangt namelijk van hun layout af.

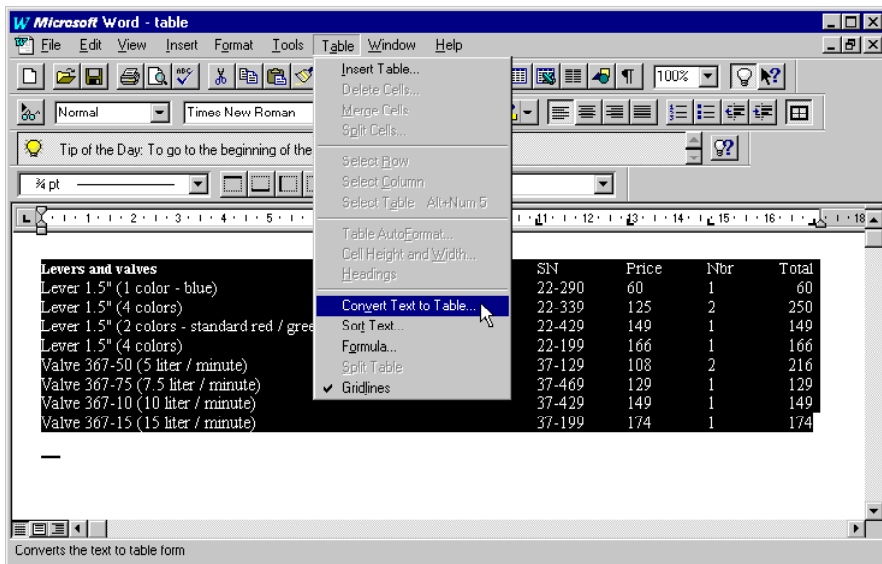
Wanneer de tabelcellen omgeven zijn door randen, zal de pagina-analyse de volledige tabel omkaderen en kan Readiris deze tabel hercreëren. Wanneer er geen randen rond de cellen getrokken zijn, moet de gebruiker vensters trekken rond ieder kolom van de tabel om zo de uitvoer netjes in cellen te schikken.

Laten we twee voorbeelden geven, we starten met de omkaderde tabel. Aangezien *elke cel omgeven is door een rand*, plaatst de pagina-analyse één (tekst)venster rond de volledige tabel.

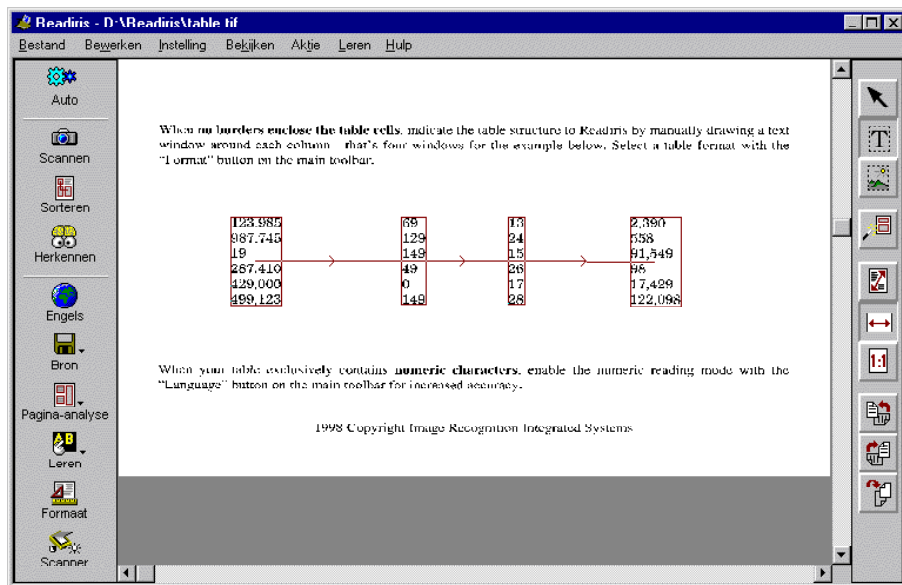


Voer de herkenning uit met de layoutoptie "Woord- en Paragraaformattering Behouden" of "Brondokument Hercreëren" geactiveerd en de tabel wordt gerecreëerd door Tab-codes toe te voegen tussen de cellen.

Open uw tekstverwerker om het resultaat even te bekijken – u kunt gebruik maken van het commando "Tekst in Tabel Omzetten" van uw tekstverwerker om één en ander af te ronden. (U had trouwens de overige tekstblokken in de uitvoer kunnen insluiten.)

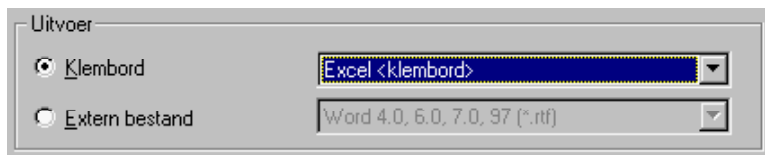


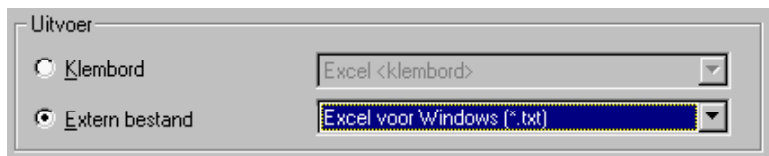
Nu het voorbeeld waar er *geen randen de cellen omkaderen*. Aangezien de pagina-analyse de tabel niet als dusdanig, niet als een geheel kan detecteren, moet de gebruiker een tekstvenster tekenen rond iedere kolom om de uitvoer netjes te schikken in aparte cellen.



Deze keer moet u een tabelformaat kiezen met de "Formaat" knop om Tab-codes toe te voegen tussen de cellen. (De "algemene" tekstformaten volstaan *niet*: u hebt een specifiek formaat nodig om de rijen en kolommen aan te maken!)

Zoals met teksten het geval is, kunt u een extern bestand creëren dat u kunt importeren in uw spreadsheet, en kunt u de tabel in het klembord kopiëren en onmiddellijk plakken in uw rekenblad.



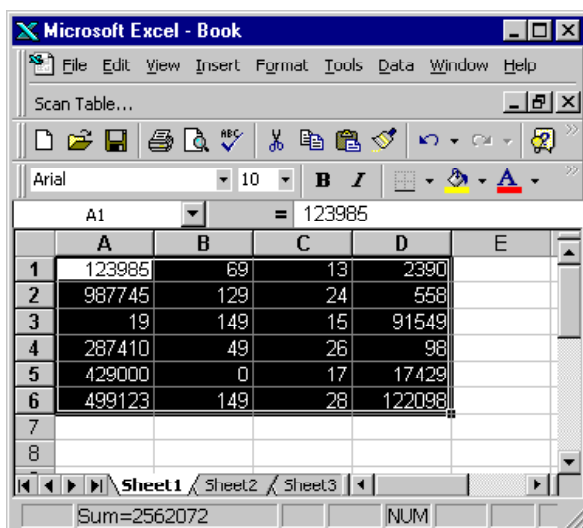


Voor een optimale OCR precisie, beperkt u de herkenning best tot de **numerieke symbolen** met de "Taal" knop. (De numerieke modus bevat niet enkel de getallen, maar ook de symbolen +, \*, /, %, , (komma), . (punt), (, ), -, =, \$ en £.)



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.

Het resultaat, geopend met of geplakt in uw spreadsheet, ziet er zo uit: de typische tabelstructuur met rijen en kolommen werd herschapen en u bent onmiddellijk klaar om de data te verwerken.



## DE “CONNECT” FUNCTIE BIEDT U EEN DIRECTE LINK

Nu we het over tabellen hebben: er is er een andere manier om ze te herkennen, namelijk *vanuit* uw rekenblad!

De “Connect” mogelijkheid biedt u een rechtstreekse OCR link tussen uw scanner en uw Windows applicaties. Installeer de “Connect” mogelijkheid en u kunt documenten onmiddellijk vanuit uw favoriete Windows applicaties scannen – en dat kan evengoed uw rekenblad als uw tekstverwerker zijn!

Er zijn enkele handmatige stappen vereist die geen automatisch **installatieprogramma** voor u kunt verzorgen; u kunt de on-line hulp of het bestand CONNREAD.DOC raadplegen om te leren hoe u één en ander installeert. (Dit “Lees Mij” bestand wordt trouwens automatisch op het scherm gebracht op het einde van de installatieprocedure.)

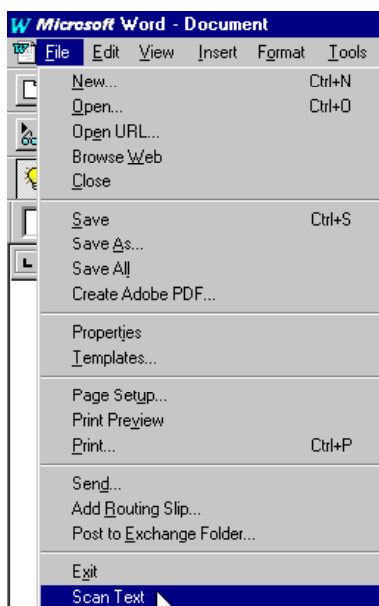




Wanneer de "Connect" mogelijkheid correct geïnstalleerd is, voegt deze een nieuw commando toe de menu's van uw applicaties. In de onderstaande voorbeelden is het commando "Tekst Scannen" toegevoegd aan het "Bestand"



menu van Word en is het menu "Tabel Scannen" toegevoegd aan de menubalk van Excel.



Hoe het werkt: Readiris mag nog niet in de achtergrond lopen op het ogenblik dat u de “Connect” macro oproept. Klik op het nieuw commando in uw applicatiemenu's om Readiris te laten lopen. De “Connect” interface kan u enkele vragen stellen eer Readiris opstart. Ga te werk zoals u dat gewoonlijk doet en verlaat de “Connect” interface wanneer u klaar bent. Het herkenningsresultaat wordt ingevoegd op de plaats van de cursor.

## 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 "Inhoud" en zijn versnellingstoets F1 laten u navigeren door de vele hulpbladzijden.

