

Background

package

READIRIS OCR

Save Time, No More Retyping!

IRIS
paper

Druckqualität

trainable

učici modul

OCR

matrix

обработка

caractères accentués

omnifont

easy

recognition
fax

neue Symbole und Schriftarten

IRIS
DOCUMENT
MANAGEMENT
SYSTEMS

JUST READ IT!®



© 1997 I.R.I.S. Všechna práva vyhrazena
Technologie OCR firmy I.R.I.S.
Technologie Connectionist firmy I.R.I.S.

The Proximity/Merriam Webster Linguibase
© 1984 Všechna práva vyhrazena
Proximity Technology Inc.
© 1984 Všechna práva vyhrazena
Merriam Webster Inc

The Proximity/Merriam Webster Linguibase
© 1984 Všechna práva vyhrazena
Proximity Technology Inc.
© 1984 Všechna práva vyhrazena
William Collins Sons & Co. Ltd.

Kontrola pravopisu tohoto výrobku je založena na
technologii Proximity Linguistic
© 1997 I.R.I.S. Všechna práva vyhrazena

ÚVODEM O PROGRAMOVÉM VYBAVENÍ READIRIS

Blahopřejeme Vám ke koupi programového vybavení Readiris, které Vám nepochybně velice pomůže při záznamu textu i obrázků.

Jakkoli jsou dnešní počítače účinné, musíte je nejdříve naplnit daty. A pokud jste někdy přepisovali třeba patnáctistránkovou zprávu s množstvím tabulek tak víte, jaká to je namáhavá a časově náročná práce. Použitím tohoto moderního programového vybavení OCR získáte možnost automatizovat zadávání textů do aplikací programů s vysokou mírou účinnosti.

Některými kliknutími na tlačítko myši ušetříte hodiny práce: program Readiris převede vaše dokumenty vytištěné na papíře do podoby souborů pro počítač 10 až 25 krát rychleji než při ručním přepisování. Díky prostředku Connect můžete používat funkce OCR přímo v prostředí aplikací pro Windows.

Readiris dokáže také rozpoznat a přečíst tabulky s čísly a převést je do souborů, které můžete později zpracovat v kterém z tabulkových programů (spreadsheets).

Program Readiris, založený na technologii Connectionist od firmy I.R.I.S., představuje nejlepší nabídku v oblasti vybavení pro OCR. Řešení nezávislé na typu písma a doplněné funkcemi samoučení je odvozeno z technologie neuronových sítí. Systém se dokáže učit nové znaky na základě kontextové analýzy: lingvistické znalosti o slabikách a stavbě slov ve 28 jazycích zvyšují možnosti a výkon OCR. To vše propůjčuje programu Readiris schopnost přesného řešení všech druhů dokumentů, včetně špatně čitelných faxů a výstupů z jehličkových tiskáren.

Přesnost systému dále zvyšuje verifikace uživatelem pomocí rozevíracích nabídek. Všechny uživatelem ověřené výsledky si systém zapamatuje a tím dále zvyšuje svoji rychlost a spolehlivost. Při každém použití programu Readiris se tak zvyšuje jeho inteligence. Tento účinný nástroj umožňuje uživateli "naučit" program rozpoznávat například matematické symboly a vypořádat se se zkrácenými písmeny v dokumentech.



NĚKOLIK SLOV O PŘÍRUČCE

Tato příručka je stručným úvodem k použití programového vybavení Readiris a obsahuje vše, co budete potřebovat pro započetí práce s tímto výrobkem firmy I.R.I.S. Příručka však není kompletní uživatelskou dokumentací programu Readiris. Pro využití všech jeho funkcí a možností je zapotřebí, abyste si podrobně prostudovali informace v kontextové nápovědě.

POŽADAVKY NA SYSTÉM

Minimální konfigurace počítače nutná pro práci s programem Readiris:

- ☐ Počítač PC kompatibilní s procesorem typu 386; doporučen je počítač s procesorem typu 486DX2.
- ☐ 4 MB paměti RAM. Použití prostředku Connect vyžaduje větší kapacitu paměti; doporučeno je 8 MB RAM.
- ☐ 8 MB volného prostoru na pevném disku.
- ☐ Operační systém MS-DOS 3.3; doporučen je operační systém MS-DOS 6.x.
- ☐ Prostředí Windows 3.1. Nespouštějte Windows 3.1 ve standardním režimu!

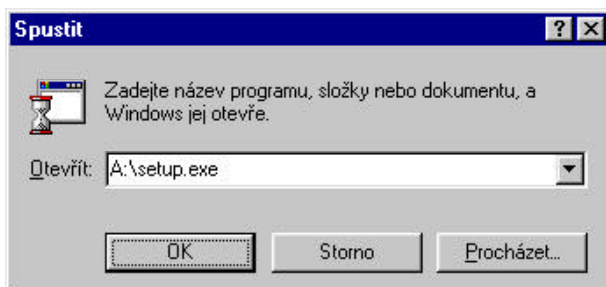
Program Readiris a prostředek Connect pracuje bez problémů také pod Windows 95.

INSTALACE PROGRAMOVÉHO VYBAVENÍ READIRIS

Programové vybavení Readiris je dodáno v komprimovaném formátu a proto je nutné použít instalace programu. Postupujte podle dále uvedených pokynů. Instalace ve Windows 95 a ve Windows 3.1 se poněkud liší.

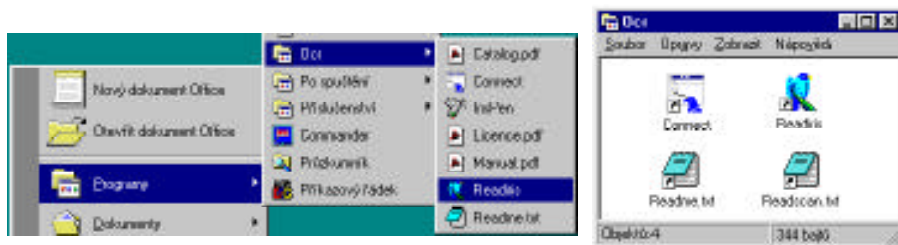
Instalace ve Windows 95

Nové programové vybavení se ve Windows 95 instaluje buď povelom " Spustit" z nabídky "Start" nepo pomocí volby " Poidat/Odstranit programy" v "Ovládacím panelu". V uvedené ukázce je použit povel "Spustit". Zadejte `a:\setup` a klepněte na tlačítko "OK".



Tím se spustí instalační program. Postupujte podle pokynů na obrazovce a vkládejte do disketové mechaniky další diskety, když o to budete požádáni.

Během instalace programu se v nabídce "Programy" automaticky vytvoří vedlejší nabídka "OCR".

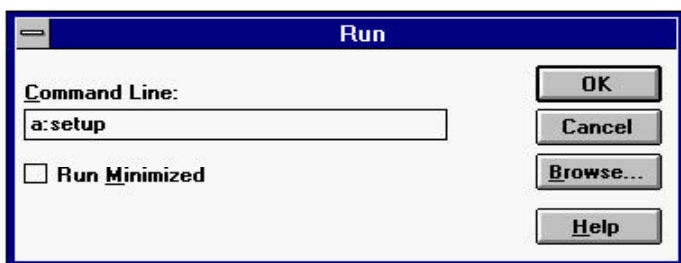


Instalace ve Windows 3.1

Instalace programového vybavení je velice podobná instalaci ve Windows 95. Jediným rozdílem je způsob spuštění instalačního programu.



Zkontrolujte, zda jste v prostředí Správce programů (Program Manager). Zadejte povel "Run" z nabídky "File", napište `a:setup` a klepněte na tlačítko "OK".



Další postup je stejný jako při instalaci ve Windows 95.

Pohodlí není lenost!

Pro uživatele Windows 95 je užitečné, když si vytvoří zástupce. V dokumentaci Windows 95 a v kontextové nápovědě je popsáno několik způsobů jak vytvořit zástupce. Jedním z jednoduchých způsobů je odtážení programu Readiris z aplikace Průzkumník na pracovní plochu.

Nezapomeňte na registraci!

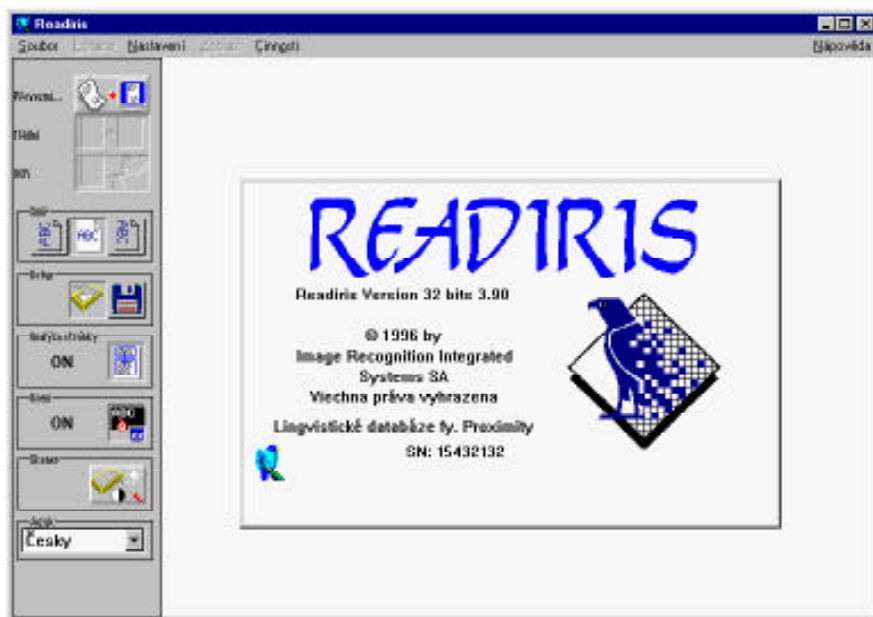
Nezapomeňte odeslat svoji registrační kartu, protože nám tím umožníte, abychom Vás průběžně informovali o vývoji programu a dalších potřebných výrobců firmy I.R.I.S. Výhody vyplývající z registrace, včetně podpory výrobku, jsou přesně omezeny pouze na registrované uživatele.

SPUŠTĚNÍ PROGRAMU

Poklepejte na aplikaci Readiris ve vedlejší nabídce či ve skupině "OCR". Ve Windows 95 můžete také poklepat na zástupce aplikace Readiris na pracovní ploše. Tím se zobrazí úvodní obrazovka a okno aplikace Readiris.

Uživatelská příručka k programovému vybavení

Na úvodní obrazovce je uvedena verze programu Readiris a odpovídající autorská práva. Klepnutím myši kamkoli do úvodního okna toto okno zmizí.



Okno aplikace Readiris neobsahuje pouze nabídku povelů, ale také pás s tlačítky, která zpřístupňují často používané povely. Na ploše okna se zobrazují skenované obrazy.

STRUŽNÁ VÝUKA

Nejlépeším způsobem, jak se seznámit s obsluhou programu Readiris je používat jej. Součástí programového vybavení je také naskenovaný obraz, který Vám



umožní pracovat s programem, i když ještě nemáte připojený skener. Můžete tedy začít ihned.

Klepněte v tlačítku na ikonu diskety. Tím vyberete jako zdroj dat naskenované obrázky a nikoli vlastní skener. Potom klepněte na tlačítko "Převzmi".



Budete vyzváni k volbě obrazového souboru. Označte soubor *czech.tif* ve složce Readiris. Obraz se načte z diskety a zobrazí se v okně programu. V pravém horním rohu okna se objeví okénko s nástroji pro práci v oknech programu Readiris.



ZVĚTŠOVÁNÍ OBRÁZKŮ

Program Readiris disponuje několika povely, které umožňují přiblížit skenovaný obraz (např. pro kontrolu kvality skenování). Nabídka "Zobrazení" obsahuje povely umožňující zobrazit naskenovaný obraz v polovičním měřítku (50%), ve skutečné velikosti nebo ve dvojnásobném měřítku (200%).



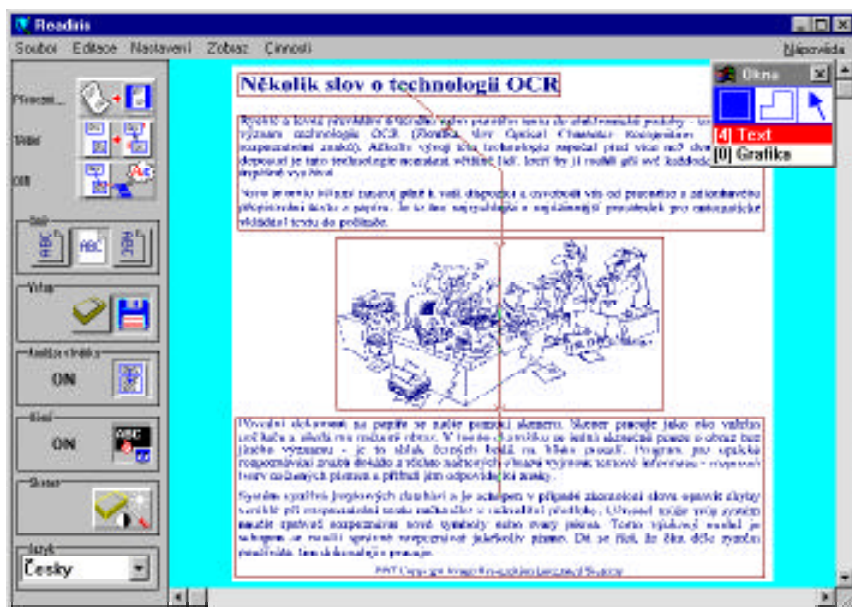
Další možností je klepnout pravým tlačítkem myši na určitou oblast skenovaného obrazu a tím ji okamžitě zvětšit na skutečnou velikost. Následujícím klepnutím pravým tlačítkem se vrátíte k původní velikosti zobrazení.



ZOBRAZENÍ NASKENOVANÉHO OBRAZU

Když je obraz naskenovaný, musíte určit, které jeho části chcete převést na běžný zpracovatelný text. To provedete nakreslením rámečků (tzv. oken) kolem bloků textu, které vás zajímají.

Pokud je v pásu tlačítek aktivována funkce "Analýza stránky", provede program Readiris tuto činnost za vás automaticky.



Program Readiris nejenom zjistí bloky textu, ale také je setřídí. Chcete-li toto posádit zmínit, klepněte v pásu tlačítek na tlačítko "Třídění".



Uživatelská příručka k programovému vybavení

Analýza stránky je metoda, která automaticky rozčlení naskenovanou stránku na textové rámečky. Příslušné nástroje programu Readiris umožňují rozčlenit obraz na zóny také ručně.



Chcete-li kolem požadovaného odstavce textu *nakreslit* rámeček, označte v okénku s nástroji "Obdélník" a položku "Text". Potom klepněte kurzorem do levého horního rohu označovaného odstavce, tažením myši roztáhněte rámeček přes celý text, a potom klepněte do pravého dolního rohu odstavce.

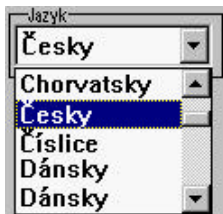
Chcete-li již vytvořený textový rámeček *upravit*, označte v okénku s nástroji "Šipku" a klepněte na jednu stranu rámečku. V rozích rámečku a v polovině každé jeho strany se objeví značky. Nyní můžete nad kteroukoli značkou umístit kurzor a tažením myši změnit velikost rámečku.

Když chcete rámeček s textem *posunout*, umístěte kurzor na ni kterou ze stran rámečku mezi značky a odtáhněte jej na požadovanou pozici.

Odstavce textu s nepravidelným tvarem lze "zarámovat" nakreslením mnohoúhelníku, který je možno uložit jako předlohu pro další použití.

ČTENÍ TEXTU NA STRÁNCE

Za předpokladu, že jsou všechny rámečky správně definovány, můžete nyní spustit rozpoznávání písma. Klepněte na tlačítko "OCR" a nezapomeňte předtím označit správný jazyk tlačítkem "Jazyk" v pásu s tlačítky.



Program Readiris není omezen pouze na angličtinu, ale pracuje až s 28 jazyky. Aplikace podporuje všechny americké a evropské jazyky, včetně jazyků východoevropských, řecké abecedy i azbuky. Volba správného jazyka je proto nutná. Podle jeho volby totiž program pozná, jaké znaky a symboly má rozpoznávat. Jedná se zejména o jazykově specifické akcenty, které jsou na základě volby jazyka přesně rozpoznány.

Další věcí je rozsáhlé použití jazykových databází pro ověření správného výsledku. Dejme tomu, že v textu je slovo "president", ve kterém se písmeno "r" rozpilo tak, že vypadá jako "f". Prohledáním anglického lexikonu program nezávisle na obsluze zjistí, že správné znění přetiskového slova je "president", bez ohledu na to, že rozpoznal písmeno "f". Tato technika "samovýuky" pochopitelně velice závisí na jazykových kontextech.

VOLBA UMÍSTĚNÍ TEXTU

Další věc, kterou musíte před spuštěním programu udělat, je volba umístění zpracovaného textu. Zadejte povel "Formát textu" z nabídky "Předvolby".

Uživatelská příručka k programovému vybavení



Tento povel slouží pro volbu formátu souboru do něhož bude rozpoznaný text uložen, případně pro umístění textu do Schránky v prostředí Windows.

Vzhled výsledného textu lze ovlivnit některými volbami. Dbejte na to, aby byla označena volba "Odstavec", umožňující programu rozpoznávat odstavce: přečtený text bude formátován do odstavců typických pro textové programy. Pokud uvedenou volbu neoznačíte, bude každý řádek textu ukončen znakem CR!

Pojmenování výstupního souboru se provádí v "Ovládacím panelu" vyvolaném z nabídky "Činnosti".



Ovládací panel

Dokument Typ písma <input type="radio"/> Fixní <input checked="" type="radio"/> Proporcionální Velikost <input type="radio"/> Malé písmo <input checked="" type="radio"/> Normální písmo Font <input checked="" type="radio"/> Normální <input type="radio"/> Maticový	Soubory <table border="1"><tr><td>Vstup</td><td>[G] polish.tif</td></tr><tr><td>Výstup</td><td>Text: [N] READIRIS.TXT Grafika: READIRIS.TIF</td></tr><tr><td>Rozlolení</td><td>Standardní</td></tr><tr><td>Slovník</td><td>[P] READIRIS.DUS</td></tr></table> Činnost 1 Stránka(y) ke čtení <input checked="" type="checkbox"/> Analýza stránky <input checked="" type="checkbox"/> Zobrazit po skenování <input checked="" type="checkbox"/> Interaktivní učení <input type="checkbox"/> Podrobná kontrola	Vstup	[G] polish.tif	Výstup	Text: [N] READIRIS.TXT Grafika: READIRIS.TIF	Rozlolení	Standardní	Slovník	[P] READIRIS.DUS	Skenuj Zavři
Vstup	[G] polish.tif									
Výstup	Text: [N] READIRIS.TXT Grafika: READIRIS.TIF									
Rozlolení	Standardní									
Slovník	[P] READIRIS.DUS									

Bez ohledu na to, zda jste se rozhodli pro umístění výsledného textu do Schránky, musíte klepnutím na tlačítko "Výstup" definovat cestu a název pro výstupní soubor.

(Ovládací panel také umožňuje podrobné nastavení postupu rozlišování: můžete zde například určit, že čtený text pochází z maticové tiskárny; rozpoznávání znaků v takovém případě používá speciální rutiny.)

INTERAKTIVNÍ VÝUKA

Klepněte v pásu tlačítek na tlačítko "OCR".



Uživatelská příručka k programovému vybavení

Postup zpracování OCR je indikován na obrazovce. Pokud je tlačítkem "Učení" v pásu tlačítek aktivována fáze interaktivního učení, program Readiris ji spustí po dokončení rozpoznávání textu.



Rozpoznaný text je postupně zobrazován v okně programu a systém se bude zastavovat u nejasných znaků; v případě nejasných interpunkčních znamének bude označen celý textový řetězec.



V uvedeném příkladu se systém zastavil na zkrácené čárce mezi slovy. Klepnutím na "OK" nebo stisknutím klávesy Enter potvrdíte navrhovaný znak v



systému učení; klepnutím na tlačítko "Pokračuj" přijmete navrhanou opravu bez učení.

Možností interaktivního učení je ve skutečnosti mnohem víc. Modul učení je složitý a velice výkonný nástroj umožňující systému naučit se rozpoznávat zvláštní, ale také zkreslené znaky v podobě, v jaké se vyskytují v běžných dokumentech.

Výsledkem procedury učení se jsou uložené soubory, kterým se říká "slovníky". Tyto katalogy písem rozšiřují schopnosti systému a při čtení podobných dokumentů je třeba uvedené soubory zavést do paměti.

Interaktivní učení také představuje závěr procesu rozpoznávání znaků, při němž jste běžný papírový dokument převedli do formy počítačového textového souboru 10 až 25 krát rychleji než při ručním přepisování!

NASTAVENÍ SKENERU

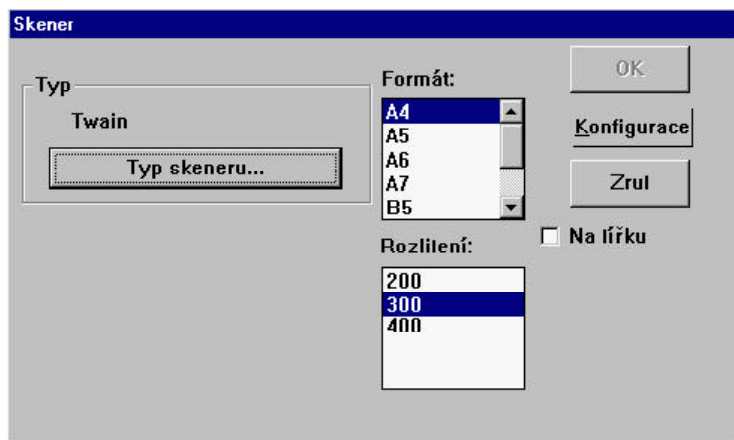
Nyní je třeba provést nastavení skeneru. Předpokladem je správná instalace všech technických i programových prostředků skeneru. Pokud jste vybavení Readiris zakoupili společně se skenerem, nebude zřejmě tento krok potřebný, protože váš skener byl nastaven v rámci instalace programu Readiris.

V pásu tlačítek klepněte na tlačítko "Skener".



Klepnutím na tlačítko "Typ skeneru" určíte typ použitého skeneru. V tomto okně můžete dále nastavit formát stránky, rozlišení skeneru, jas a kontrast.

Uživatelská příručka k programovému vybavení

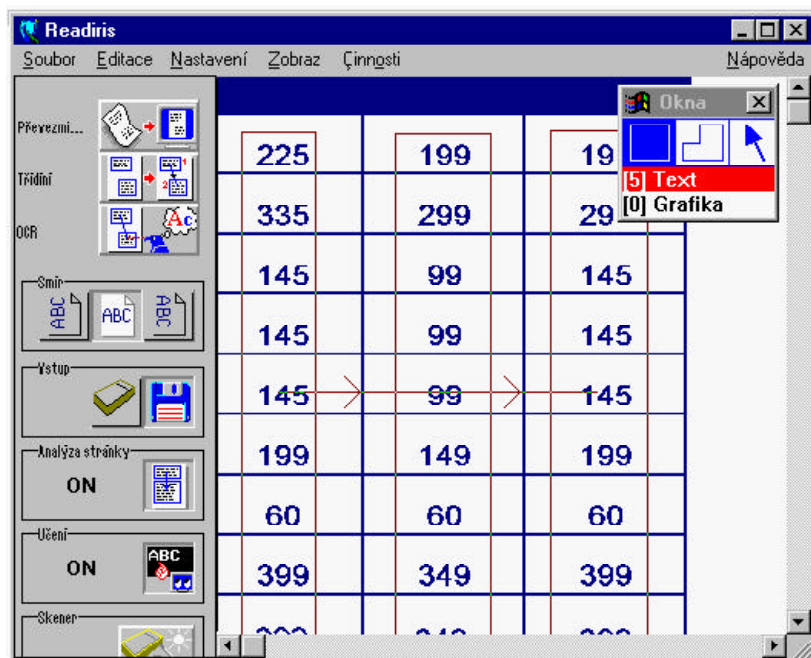


Nastavte správně všechny parametry a potom zadejte povel " Ulož standardní nastavení" z nabídky "Soubor". Tím uložíte provedené nastavení jako výchozí pro další použití.

SKENOVÁNÍ TABULKY

Naskenujte požadovanou tabulku s čísly.

Kolem každého sloupce tabulky nakreslete rámeček, čímž uspořádáte jednotlivá čísla do odpovídajících buněk; v opačném případě byste nebyli schopni správně interpretovat tabulku v tabulkovém programu.



Pro dosažení optimální přesnosti lze pomocí tlačítka "Jazyk" v pásu tlačítek omezit rozpoznávání pouze na čísla.

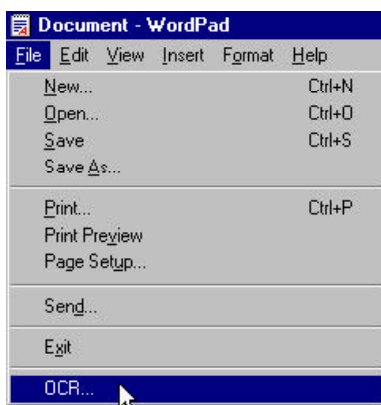
NĚKOLIK SLOV O PROSTŘEDKU CONNECT

Prostředek Connect umožňuje přímé propojení OCR mezi skenerem a aplikací pro Windows. Instalace uvedeného prostředku vám umožní přímé skenování dokumentů z prostředí vámi používané aplikace pro Windows.

Uživatelská příručka k programovému vybavení

Podrobnosti o instalaci prostředku Connect se dozvíte prostřednictvím kontextové nápovědy. Po instalaci Connect se v nabídce použité aplikace objeví nový povel.

V uvedeném příkladu je do nabídky "Soubor" v programu Word přidán povel "Skenovat text" a do téže nabídky v programu WordPad je přidán povel "OCR".



Klepnutím na uvedené povely spustíte program Readiris přímo z prostředí aplikace, ve které pracujete. Další postup je běžný. Rozpoznaný text nebo tabulka jsou umístěny za pozici kurzoru v okně spuštěné aplikace.

KONTEXTOVÁ NÁPOVĚDA

V této příručce jsou popsány pouze ty, které základní funkce programu Readiris. Tento program je však obsáhlým prostředkem OCR s množstvím funkcí, které zde nejsou popsány. Doporučujeme proto, abyste se se všemi možnostmi tohoto programu seznámili prostřednictvím informací v kontextové nápovědě programu Readiris.





REGISTRACE VÝROBKU

Nezapomeňte nám poslat svoji registrační kartu! V případě, že jste **registrační kartu** ztratili, vyplňte teno list a pošlete nebo odfaxujte jej firmě I.R.I.S.

Vaše registrace nám umožní informovat Vás o dalších výrobcích firmy I.R.I.S. a o vývoji **nových produktů**. Registrace Vás kromě toho opravňuje pro získání podpory a **speciálních nabídek**.

Výrobek / verze

Sériové číslo

Firma

Jméno

Titul

Ulice / číslo

PSČ

Země

Telefon

Fax

Kde a kdy jste výrobek koupili?

Datum

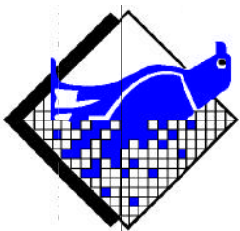
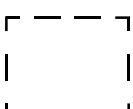
Image Recognition Integrated Systems

Product registration

Rue du Bosquet 10, 1348 Louvain-la-Neuve (Belgium)

Tel: 32-10-45 13 64 - Fax: 32-10-45 34 43





I.R.I.S.

Product registration
Rue du Bosquet 10,
1348 Louvain-la-Neuve
Belgium