

Readiris Pro - Corporate 8.0

Contents

- Software distribution
- Registering your licence
- New features version 8.0
- Readiris Pro vs. Readiris Corporate
- Last-minute change: process images as 300 dpi
- The software documentation
- The on-line help and Microsoft Internet Explorer
- Supported platforms
- Important note concerning foreign languages
- Reverse images with Twain compatible scanners
- Scanner drivers
- Recognizing the Euro (€) currency symbol
- Recognizing the Asian languages
- Getting technical support
- Creating Adobe Acrobat PDF files
- Contacting I.R.I.S.

1. Software distribution

This release of the Readiris software is contained on a **CD-ROM** or can be downloaded from the web site of **electronic resellers**.

A trial version can be downloaded from the I.R.I.S. web site (www.irislink.com).

The Readiris **CD-ROM** contains the following software packages:

- Readiris 8.0
- Cardiris 2.5 LE (in the folder “Cardiris”)

*** Note ***

Cardiris is only shipped with Readiris Pro, not with Readiris Corporate! (Readiris Corporate reads business cards as does Cardiris.)

- Microsoft Internet Explorer 5.0
- Adobe Acrobat Reader 5.1 (several languages)
- Some scanner drivers
- Free companion software – see below

Note above all that not all software mentioned above is necessarily included in your release. With some software bundles and with the downloadable versions, all “**goodies**” other than the essential Readiris software are *not* included.

To install the **Readiris 8.0** software, simply insert the **autorunning CD-ROM** into your CD-ROM drive and follow the on-screen instructions.

With some **software bundles**, you may have to install the additional software packages by going to the corresponding folder and running the “Setup” program. The downloadable versions equally come with an installer.

Before installing these scanning applications, you should install the software supplied with the **scanner**.

For the **trial version** downloadable from the I.R.I.S. web site, the operation is limited to 30 days. However, you can also purchase an “**ESD**” **version** (“Electronic Software Delivery”) on the I.R.I.S. web site.

2. Registering your licence

Don't forget to **register** your product! There are many good reasons for doing so:

- Registering allows us to keep you informed of future **product developments** and **related I.R.I.S. products**.
- Registering entitles you to free **product support** and **special offers**.
- Depending on the software bundle, you'll receive the **softkey** in return as may be needed to continue using Readiris after one month.

You can register in many ways. Send in your registration card, access the Readiris registration form (<http://register.irislink.com/register/index/index.html>) on the I.R.I.S. web site (www.readiris.com and www.irislink.com) or call I.R.I.S. during the working hours:

USA: 1-561-395 7831 / 800-447-4744
 Europe: 32-10-45 13 64 (all major languages)

3. New features version 8.0

Refer to the I.R.I.S. web site (www.readiris.com) for full details on the new features of Readiris 8.0. We will limit ourselves here to a short discussion of the “highlights”.

- A new, more powerful **recognition module** yields unparalleled OCR accuracy and a higher speed.

New preprocessing and recognition routines are used to process **faxes**, resulting in significantly higher **OCR accuracy** for this document types. (The OCR also makes use of extended and new **linguistic databases**.)

- Newer, more powerful “**autoformatting**” recreates your source documents with higher accuracy.

With Readiris 8.0, the OCR software detects columns in your document and can **recreate** them in the output file. Scan a columnized document and you'll get a Word document with **editable columns**. As you edit the text, the text “flows” naturally from one column to another!

- The code of Readiris 8.0 was **speed-optimized**.

Recognizing a color page can now easily take a few seconds less - and Readiris was already the fastest OCR package on the market!

- The **user interface** (GUI) was redesigned with **ergonomy** on our mind.

- Readiris 8.0 recognizes lots of extra **languages**. Readiris now supports up to 104 languages!
- Readiris 8.0 opens and recognizes **PDF documents** – even when they are “read-only”!

You can open image-only PDF files and convert them into text documents (in any supported text format). You can also convert image-based PDF files into text-based PDF files... or even in image-text PDF files that contain both the text and the scanned page!

- The “**Send to**” feature was extended substantially. After the recognition, Readiris can automatically start up the application that allows you to edit the recognized documents!

You create a direct link between the OCR software – the tool that convert paper into computer-editable files – and the application that edits these files... These “target applications” include “traditional” word processors such as Microsoft Office 97, 2000 and 2002 but also HTML editor. If you want to post paper documents on the web, Readiris is the proper tool...

- The support of **output formats** was broadened.

Readiris 8.0 supports the newest text applications – **web browsers** such as Netscape 7.0, **office suites** such as Sun StarOffice 6.0, AbiSource AbiWord, Software602 Pro PCSuite, **HTML editors** etc. Even the new free **wordprocessor** Jarte 1.x is supported!

HTML support is based on format 4.0 and fully “**WYSIWYG**”!

- Readiris 8.0 supports the latest **scanner models** from all major manufacturers and supports new **image formats** (on the input side) such as PNG and the fax format DCX.
- Readiris **Corporate** 8.0 introduces lots of extra **novelties** – see below.

4. Readiris Pro vs. Readiris Corporate

Readiris Corporate is a special version for the **corporate market**. It combines the “Pro” package with **extra features**:

- Readiris Corporate includes the **BCR** functionality (“Business Card Reading”).

*** Note ***

Depending on your software version, **Cardiris** “Lite”, a business card organizer, may be shipped with Readiris Pro. This software also reads business cards!

*** Note ***

Readiris Corporate comes with free extra software to maintain your business contacts – GoodContacts. See the section “Free Companion Software” below.

- Readiris Corporate activates **secondary languages** alongside the “main” language. You can read documents that **mix** several languages and **alphabets**!
- Readiris Corporate “boosts” the **OCR accuracy** further by loading **user lexicons** that reflect your company’s activity.
- Readiris Corporate maintains the **colors** of the **text** in the recognized documents.
- Readiris Corporate is not limited to **multipage documents** of 50 pages.
- Readiris Corporate supports **duplex scanning**.
- Readiris Corporate offers improved control over the **Adobe Acrobat PDF** output: influence the file size directly by determining the color mode and image resolution.
- Readiris Corporate offers advanced options for the **saving** of **graphics**.

5. Last-minute change: process images as 300 dpi

A last-minute change was applied to the Readiris software: the image option “**Force to 300 dpi**” was renamed “**Process as 300 dpi**”.

Its **operation** however has not changed! A small recap.

Thanks to this option, the images will be processed “normally”, as if they had a 300 dpi resolution. This option never changes the image resolution in any way! (You also avoid a warning that you’re submitting images with a resolution lower than 200 dpi or higher than 800 dpi.)

This image option has specific relevance for the **(auto)formatting** of the recognized documents! Autoformatting means that you recreate a true copy of your source documents: the document has the same size, the point sizes of your titles, text blocks etc. are recreated and so forth.

Readiris obviously needs to know the correct image resolution to be able to do this: for instance, Readiris “knows” that when a symbol is 10 pixels high in an image with the resolution y , the letter should have point size z in the output. But double the image

resolution to y^2 and the ratio pixel dimensions vs. point size for any given character changes dramatically...

However, images generated with **digital cameras** don't indicate any resolution. And there is the the unfortunate fact that some image files actually indicate an **erroneous resolution in the file header**.

Let's investigate a real example for a while. A, A4 300 dpi image gets presented as a 72 or 100 dpi image by its file header. Readiris will try and make the recognized document and the point sizes of the titles and text blocks 3 times bigger than they should be. Add to this the fact that Microsoft Word (and many other text applications) don't handle any documents bigger than A3 (twice the size of an A4 page) and you begin to see how things can go wrong here.

Which leads us to this conclusion: when you generate recognition with chaotic formatting results, the first thing to do is to check this option! Should it have been enabled or not?

6. The software documentation

Whether you receive a paper manual for Readiris depends on the version-bundle of the software. In all cases, the software documentation is supplied in electronic format - hence the Adobe Acrobat **PDF files** on the Readiris CD-ROM.

The Readiris installer prompts you to install the "**User's Manual**" in the Readiris folder and creates a **program icon** under "I.R.I.S. Applications - Readiris" that allows you to access the manual directly. (Alternatively, you can leave the manual on the CD-ROM: when you click the program icon, you'll be asked to insert the Readiris CD-ROM in your CD-ROM drive.)

You'll also find a "**Getting Started Guide**" on the Readiris CD-ROM. This shorter, reduced version of the "User's Manual" (available only in English) is never installed on your hard disk, but you can find it on the Readiris CD-ROM in the folder "Getting Started".

The **Adobe Acrobat Reader** software (version 5.1) is used to view and print the electronic documentation. The Readiris installer again proposes to install that software for you. (The installer selects the same language as the language you selected for Readiris.)

However, know that the Adobe Acrobat Reader software comes with a reduced manual (and on-line help) of just a few pages. To download the full **manual**, click the hot zone

on the first page of the on-line help/manual. Or don't bother doing that: you'll find the appropriate files on the Readiris CD-ROM!

7. The on-line help and Microsoft Internet Explorer

Readiris is equipped with an **HTML** based **on-line help** system. The HTML Help Viewer, part of the Windows operating system, uses portions of the Internet Explorer software. In other words, a version 4.x or higher of Internet Explorer must be installed if the Readiris on-line help system is to display correctly! (The browser is included automatically in all recent versions of Windows.)

I.R.I.S. is entitled to ship the Microsoft web browser **Internet Explorer** free of charge. Depending on the software version you received, you may find the English version of this software under the folder "IE50" of the Readiris CD-ROM. Feel free to install it!

Run the **installation** program IE5SETUP.EXE to install the web browser on your PC. Version 5.0 of Internet Explorer is supplied, so, we recommend you *not* to install the web browser if a more recent version is already installed on your PC. The folder "IE50" of the Readiris CD-ROM contains "ReadMe" files that issue detailed installation instructions for Internet Explorer 5.0.

Other files needed by the Windows component HTML Help Viewer are verified-installed by the Readiris installer. Which files get installed depends entirely on the version of Windows you dispose of. Windows XP, ME, 2000 and 98 have their built-in components to display HTML help files while Windows 95 requires the installation of extra files. In any case, the Readiris installer handles all these issues for you!

8. Supported platforms

Readiris is a **32-bit** application that runs on the Windows XP, ME, 2000, 98, 95 and Windows NT 4.0.

That doesn't mean the drivers of your **scanner** support all those platforms: check the documentation that accompanies your scanner to detect which platforms are supported!

9. Important note concerning foreign languages

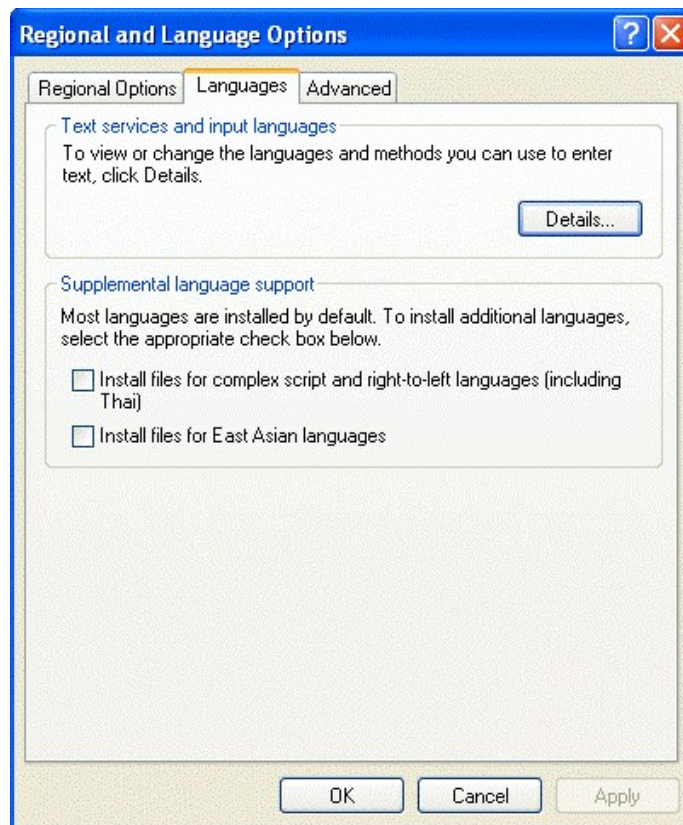
Because of internal reasons, Windows ME, 98 and 95 use only 8 bit character tables, not so-called “Unicode” tables. (No special steps are required on Windows XP, Windows 2000 and Windows NT 4.0 systems.) As a result, the document language you select with Readiris must be supported by your **localized version** of Windows.

Simply put, Readiris can read Greek, Cyrillic etc. comfortably, but your Windows licence may not be able to handle Greek, Cyrillic etc. characters.

This becomes clear in two ways: (1) Windows may be unable to represent these special characters on your **screen**, even if Readiris recognized them correctly, (2) the learning phase may prompt you to respond to recognized “special” characters, and here again it only works when Windows is able to display and accept **keyboard input** of these special characters.

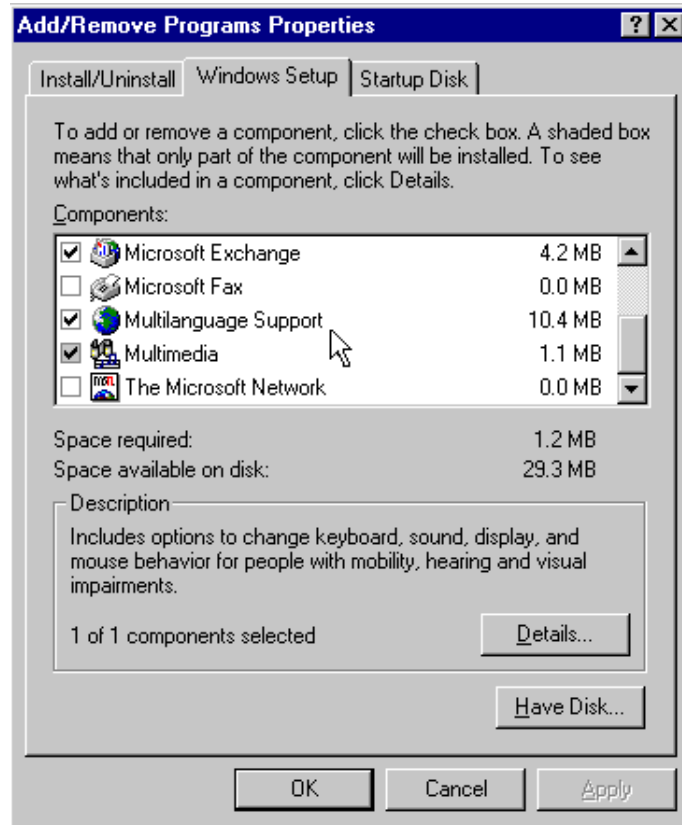
Windows XP, 2000 and NT 4.0

No special steps are required on **Windows XP, 2000 and NT 4.0** systems. To make sure that your configuration supports the required languages, you can check the control panel “Regional Settings (and Languages)”.



Windows ME, 98 and 95

Windows ME, 98 and 95 can be easily adjusted to support extra languages. Simply put, you need to ensure that the Windows module “**Multilanguage Support**” is installed on your computer system. (To do this, select “Settings” under the “Start” menu and go to the “Control Panel”. Now select “Add/Remove Programs” and click the tab “Win(dows) Setup”.



You’ll find the item “Multilanguage Support” in the list of Windows components. Click “OK” to execute, you will be prompted for the Windows CD-ROM.

10. Reverse images with Twain compatible scanners

Readiris and Cardiris have been designed to interpret correctly the image of the Twain source.

However, in some cases, the **image** received by the I.R.I.S. application is “**negative**” or “**inverted**”: black pixels on the page are white on the screen and the white background shows up black on the screen.

In this case, activate the option “Invert” as you find under the “Scanner” button on **Readiris**’ main toolbar.

For **Cardiris** (“LE”), you have to edit the file WIN.INI file under the Windows folder, search for the entry “[Twain]” and add the following lines:

```
[Twain]
invert = 1
```

This information will instruct Cardiris (and other scanning software) to reverse the image when it is being transferred from the Twain source.

11. Scanner drivers

If you experience problems with the scanner driver, you may find alternative third party drivers on the Readiris CD-ROM under the folder “Drivers”.

I.R.I.S. offers no guarantee that drivers are supplied for your scanner model or that the drivers supplied on the Readiris CD-ROM will work (well) with your scanner model.

Don’t hesitate to contact your scanner manufacturer or its representative should problems with scanner drivers continue. Most manufacturers allow you to download the latest versions of the scanners drivers from their web site.

12. Recognizing the Euro (€) currency symbol

Although Readiris 8.0 recognizes the **Euro symbol (€)** comfortably, the currency symbol may not show in your text file when you study the recognition result.

This is not due to Readiris, but to your computer’s **operating system** (keyboard drivers) and the fonts as are used when you display the text result.

Windows XP, ME, 2000 and 98 are equipped to represent the Euro symbol, but it takes a **software “patch”** to represent the Euro symbol under Windows 95 and Windows NT 4.0.

Contact your reseller to obtain the necessary files or **download** the software patch from the Microsoft web site at the URL <http://www.microsoft.com/windows/euro.asp>.

Consult the Microsoft web site to obtain more **information** on how to install the Euro product update.

13. Recognizing the Asian languages

Readiris 8.0 optionally offers recognition of 4 **Asian languages**: Japanese, Simplified Chinese, Traditional Chinese and Korean.

Note that it does not take an Asian, Japanese, Chinese or Korean version of the Windows operating system to make good use of such recognized texts. You can also run specialized **“emulating” software** on a Western version of Windows to correctly represent the **“exotic” symbols** of these Asian languages.

Secondly, if you are running Word 2002 (Office XP), Word 2000 (Office 2000), this step is superfluous: **Word 2002** and **2000** have what it takes to view and edit documents in over 80 languages, and that includes Arabic, Hebrew, Japanese, Korean, Traditional and Simplified Chinese.

So much for the language support of Windows. But the Adobe Acrobat (Reader) software also plays a role! Indeed, you must have the appropriate version of Acrobat (Reader) to correctly display the files Readiris generates! To view and edit Asian **PDF documents**, you can use an Asian version of the Adobe Acrobat (Reader) software or install the Asian font pack over a “Western” version of Adobe Acrobat (Reader). The Asian versions of the Adobe Acrobat Reader and the Asian font pack can again be found on the Readiris **CD-ROM!**

Contact I.R.I.S. to purchase the **“add-on” recognition module** for these Asian languages.

This extra CD-ROM complements your installed Readiris 8.0 licence automatically - it does not work on earlier versions than 8.0 of the Readiris software - and explains all you need to know to successfully recognize Asian documents.

14. Creating Adobe Acrobat PDF files

One of the many “hot” features of Readiris 8.0 is the generation of **Adobe Acrobat PDF output**. Readiris will even generate the bookmarks for you if you enable the right options!

There are actually two PDF **formats** on offer: you can generate “PDF Text” (text only PDF files, with possibly graphic zones) and “PDF Image-Text” (where the text is placed under the page image in a two-layered file).

Both type files yield searchable, editable PDF files. Acrobat PDF files indeed have many **advantages**:

- “Text only” PDF files are much more **compact** than image files!
- Text-based PDF files are **searchable**. (Bitmap images - “image only” PDF files - can be viewed but not searched.)
- Text-based PDF files are **editable**. (Bitmap images - “image only” PDF files - can be viewed but not edited.)

When the Adobe Acrobat PDF format is selected, the layout option is limited to “**autoformatting**” (no body text or word and paragraph formatting) and the option “Create Bookmarks” becomes available. For the text zones, Readiris applies an intelligent algorithm to come up with a title, a “summary” for each item, the tables and graphics are simply numbered.

Readiris generates text-based PDF files in all supported languages (while Adobe Capture only supports 15 languages). Readiris even generates PDF output for the 4 **Asian languages** (Japanese, Simplified Chinese, Traditional Chinese and Korean) and **Greek**. With **Cyrillic** PDF files, there’s a minor limitation: you can’t get the typestyle “italic”.

You must have the appropriate version of Acrobat (Reader) to correctly display the files Readiris generates. To view and print Central-European texts (such as Czech and Polish), Baltic texts, Turkish and Cyrillic texts in the PDF format, you must have the special “**CE**” (Central-European) **version** of the Acrobat (Reader).

The “CE” Acrobat Reader software can be downloaded for free from the Adobe web site (www.adobe.com), but you can also find it on the Readiris CD-ROM.

Finally, a small comment on the **compression** methods. Graphic zones first: black-and-white images are TIFF G4 compressed. Greyscale and color images are JPEG files (with (0.8) high quality). The text is compressed using the Gzip mode. This applies to both graphic zones inside “text-only” PDF files and “text-image” PDF files.

The recognized text can obviously be edited and re-used.

Editing the recognized text

Use the **TouchUp Text** tool of the Acrobat software to **correct** small recognition errors in the PDF file.

Exporting text to other applications

- You can isolate the text from a “image-text” PDF file. You can also convert text-only PDF files into RTF files. Open the file with Adobe Acrobat and use the command “**Save As**” to save it in an RTF text file.
- To re-use small **text portions** from a PDF file in other applications, select the **Text Select tool** of the Adobe Acrobat software, select the required text and **copy-paste** it to another application. (Select the tool Table/Formatted Text to maintain the text formatting.) The command **Select All** selects all text of the current page, not of the entire PDF file.

Intelligent searching

Use the **Find** command of your Acrobat (Reader) software for simple searches within a document and the **Search** command for advanced searching across several PDF documents.

Searching for words

The button **Find** of the Adobe Acrobat (Reader) software finds complete words or word parts in the **current PDF document**. Acrobat looks for the word by sequentially reading every word on every page in the file.

Searching on indexes

The button **Search** of the Adobe Acrobat (Reader) software allows you to perform **advanced** and fast **searching** on a **collection** of indexed PDF documents.

- You can search for a simple **word** or **phrase**.
- You can expand your search query by using **wildcard characters** and **Boolean operators**.
- You can use the **search options** to **refine** your search further.

Index-based searching implies that the **full-text index** was created for a collection of PDF files with the command **Catalog**. (A full-text index is an alphabetized list of every word used in a document or a series of documents. Index-based searching is much **faster** than the Find command: Acrobat goes right to the word in the list rather than progressively reading through the documents.)

Warning: not all versions of the Adobe Acrobat Reader software include the Search function!

15. Free companion software

Depending on the software version you received, extra “**goodies**” other than the essential Readiris software may be included. (With some software bundles and with the downloadable versions, no software other than the essential Readiris software is included.)

Neuratron PhotoScore MIDI - software for music recognition

Readiris comes with a “lite” version of the music recognition software Neuratron PhotoScore MIDI.

PhotoScore MIDI is the advanced **music scanning** solution from Neuratron (including I.R.I.S. OCR technology). It’s the musical equivalent of the text recognition software Readiris.

What does PhotoScore MIDI do?

PhotoScore MIDI scans and recognizes **printed music** (“sheet music”).

This software reads a wide range of musical markings - and we have to get a bit technical here:

- notes and chords (including tail direction, beams and flags), rests
- accidentals and articulation marks
- triplets and triplets
- clefs, key signatures and time signatures
- 5-line staves (normal and small) and barlines
- the page format (page size, staff size, margins and where systems end)
- slurs, ties and hairpins
- text including lyrics, dynamics, guitar chords, fingerings, instrument names, tempo and technique markings. (To recognize the text bits in sheet music, PhotoScore MIDI incorporates the powerful OCR kernel that drives Readiris Pro!)

The recognition results can be retrieved in and edited by MIDI sequencing and editing software.

Once the music has been “PhotoScored” it gets saved as a MIDI file and can therefore be opened in (virtually) any music editing program! After this conversion into an electronic music file, you can then edit the music, transpose it, re-arrange it, play it back, print it and extract parts.

Here are some applications: use this software as accompaniment and guidance when practicing instruments or to create MIDI files that give your web pages some pizzazz. Or scan in sheet music and use it in sequencers or other music editing programs where you’ll transpose it by bringing for instance a vocal part into range.

In other words, with PhotoScore, you can scan in your favorite music, have it played back to you, change the instruments and so on. Your imagination is the limit - as long as you don’t expect PhotoScore to bring Elvis back to life!

PhotoScore MIDI plays back scanned music

Not that it takes external MIDI software to play the music! PhotoScore MIDI directly plays back scanned music with astonishing realism using the Espressivo feature.

Trying the “Lite” software and moving beyond it

You can find the “**Lite**” version of PhotoScore MIDI on the Readiris CD-ROM in the folder “Goodies”.

You are entirely free to go ahead and make use of it. The “Lite” software does save the recognized sheet music to MIDI files. And there's more good news: the demo software is not limited in time, so there’s no hurry for you to try it quickly...

However, you don’t get the complete functionality of the full version. For instance, the “Lite” version does not recognize text (added lyrics), does not recognize triplets, hairpins etc. To get full details on the differences between the “Lite” and the full version, just run the PhotoScore MIDI Lite software. The startup details the many extras you get by acquiring the full version.

Click the button “More Details” in the startup screen of PhotoScore MIDI and you’ll visit the web page where you can order the full version from I.R.I.S. (Or contact I.R.I.S. by telephone, fax etc.)

GoodContacts – “Auto-updating” contact management software

GoodContacts complements **Readiris Corporate**, the high-end version of Readiris that recognizes business cards and sends them directly to Microsoft Outlook (Express).

Readiris Corporate acquires the business cards for you, and GoodContacts maintains them!

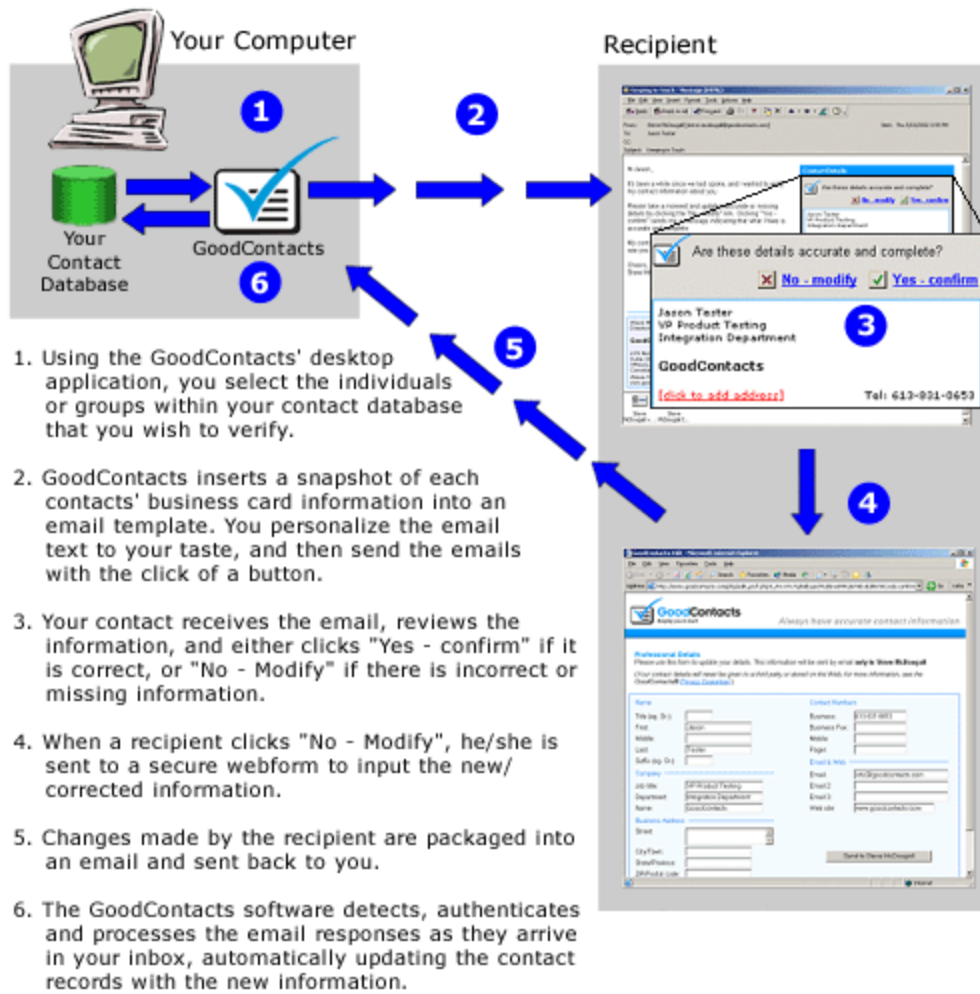
*** Note ***

Depending on your software version, **Cardiris** "Lite", a business card organizer, may be shipped with Readiris Pro. This software also reads business cards!

It's not just another contact manager...

Good Contacts is a specialized **contact manager** that goes to great lengths to help you keep your contacts **update**. The software offers valuable capabilities to verify and update contact information stored on your PC. It also includes a feature to automatically respond to GoodContacts messages from other people.

Sounds implausible? Think again! Here's how GoodContacts can **verify** the information in your contact database...



So much for keeping your info up-to-date. But GoodContacts also automates the *creation* of **new contacts** in your address database!



1. As new email arrives in your inbox, Goodcontacts automatically checks to see if you already have the sender's email address in your contact database.
2. At your convenience, you launch the GoodContacts application and review the email addresses flagged by GoodContacts as not already in your contact records. For each email address that you want to capture, you can create a new record, or add it as an alternate email address to an existing record.
3. When you create a new contact record, GoodContacts asks you if you would like to send a GoodContacts email asking the new contact to fill in any missing details, and sends the email with the click of a button.
4. Responses are automatically handled by GoodContacts, adding new information to your contact records, helping you to build a richer database of contacts.

GoodContacts maintains your **privacy** at all levels: your contact information is only exchanged with other people you choose. Any information you store in your Address Book is yours, it is not stored on the web. Nor can it ever be handed to any third party: there is no central server that organizes everybody's contacts!

Trying the GoodContacts software and maintaining it after a year

In short, GoodContacts is a great productivity tool for busy businessmen... Consult the web site www.goodcontacts.com for more information. Or try it this free companion software directly!

You can find a **full version** of GoodContacts 3.0 on the Readiris CD-ROM under the folder "Goodies", and you are entirely free to make use of it. The software is totally free for one year, and that the only limitation of this version.

16. Getting technical support

Free technical support is offered to all **registered customers** in many ways.

Europe

Hotline: 32-10-45 13 64 (working hours) (all major languages)

Fax: 32-10-45 34 43

USA

Hotline: 1-561-395 7831 / 800-447-4744 (working hours)

Fax: 1-561-347 6267

WWW

Consult the troubleshooting info on the I.R.I.S. web site.

E-mail

support@irislink.com

17. Contacting I.R.I.S.

Contact I.R.I.S. for more information.

I.R.I.S.

Image Recognition Integrated Systems

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. Inc.

Image Recognition Integrated Systems

1600 N.W. Boca Raton Bld Suite 20

Boca Raton, FL 33432 (USA)

Tel: 1-561-395-7831 / 800-447-4744

Fax: 1-561-347-6267

E-mail info: info@irislink.com

E-mail sales: sales@irislink.com

E-mail support: support@irislink.com

I.R.I.S. home page: <http://www.irislink.com>

Readiris web site: <http://www.readiris.com>
Cardiris web site: <http://www.cardiris.com>
On-line shop: <http://shop.irislink.com>

Tuesday, December 03, 2002