

Background

package

# READ IRIS OCR

*Save Time, No More Retyping!*

IRIS  
paper

Druckqualität

trainable

yamanāt  
real  
text  
typed

učici modul

OCR



matrix

допра

caractères accentués

omnifont

easy

recognition  
fax

neue Symbole und Schriftar

CR



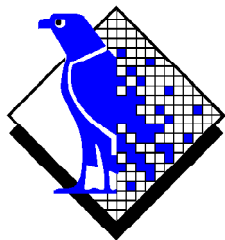
**JUST READ IT!**



# READIRIS



## USER'S GUIDE



© 1998 I.R.I.S. All rights reserved  
OCR technology by I.R.I.S.  
Connectionist technology by I.R.I.S.

The Proximity/Merriam Webster Linguibase  
Copyright 1995 All rights reserved  
Proximity Technology Inc.  
Copyright 1995 All rights reserved  
Merriam Webster Inc.

The Proximity/Merriam Webster Linguibase  
Copyright 1995 All rights reserved  
Proximity Technology Inc.  
Copyright 1995 All rights reserved  
William Collins Sons & Co. Ltd.

The spelling of this product is based on  
Proximity Linguistic Technology  
© 1998 I.R.I.S. All rights reserved

## **SAVE TIME, NO MORE RETYPING!**

---

Congratulations on acquiring Readiris. This software package will undoubtedly be of great help in recapturing your texts and tables.

As efficient as computers are, you have to key in your information first. If you have ever retyped a 15 page report or a large table of figures, you know how tedious and time-consuming it can be. Use this state-of-the-art OCR package to automatically enter text in your applications and you'll acquire an unprecedented level of efficiency and comfort!

Scan a printed or typed document, indicate the zones of interest - or have the system detect them for you - and execute the character recognition. A few mouse clicks beat long hours of work as Readiris converts your paper documents into editable computer files: it's 10 to 25 times faster than manual retyping.

With the automatic mode of operation, the user's effort is reduced to a single click: he initiates the scanning and saves the text result, all intermediate steps are taken care of by Readiris. And thanks to the "Connect" capability, you can access the OCR function from within your favourite Windows applications! To recognize faxes, you can drag the fax files from the Windows Explorer to the Readiris application window.

Readiris recognizes tables of figures and recreates the table structure by assigning your figures to separate cells; your numeric data are immediately ready for further processing with your spreadsheet.

Based on the Connectionist technology from I.R.I.S., Readiris represents the best OCR has to offer. Font-independant feature extraction is complemented by self-learning techniques derived from a proprietary neural network. The system can learn new characters through context analysis: linguistic knowledge about syllables and words improves the OCR performance.

Readiris supports up to 52 languages: all American and European languages are supported, including the Central-European languages, Greek and the Cyrillic ("Russian") languages. Readiris even copes with mixed alphabets: the software detects "Western" words that pop up in Greek and Cyrillic documents - many



untranscribable proper names, brand names etc. are written using the Western symbols.

Readiris uses linguistics *during* the recognition phase, not after it. As a direct result, Readiris recognizes documents of all kinds with top accuracy, including low-quality documents, faxes and dot matrix printouts. It copes beautifully with badly scanned and copied documents containing too light or dark font shapes. Joined characters (“ligatures”) are resolved and fragmented forms, such as dot matrix symbols, are recomposed.

User verification in pop-up style not only flags doubtful characters but also increases the system's precision. All solutions confirmed by the user are memorized, increasing speed and confidence as you go along. Using Readiris means rendering it more intelligent each time! This powerful learning tool allows you to train Readiris on special characters such as mathematic symbols and dingbats but also to handle distorted fonts as you will find in real documents.

To increase your productivity further, Readiris not only recognizes your texts, but can *format* them for you as well! Make use of “autoformatting” and Readiris recreates a facsimile copy of the scanned document: the word, paragraph and page formatting of the original document are retained. Similar typefaces are used, the point sizes and typestyles as used in the source document are maintained across the recognition. The placement of columns, text blocks and graphics follows your original documents.

In other words, Readiris allows you to archive a true copy of your documents, be it editable and compact text files instead of scanned images! Various levels of formatting are available, the choice is up to the user.

Readiris supports a wide range of popular scanners: numerous flatbed scanners, sheetfed scanners, hand-held scanners and “all-in-one” multifunctional peripherals (“MFPs”) can be used. Readiris also supports the Twain scanning standard and some scanning platforms such as Visioneer PaperPort or HP Document Assistant.

**TABLE OF CONTENTS**

Save Time, No More Retyping! .....	III
Table of Contents .....	V
Credits and Copyrights .....	VIII

**Chapter 1: Installation**

System Requirements .....	1-1
Installing the Readiris Software .....	1-1
Comfort Isn't Laziness! .....	1-3
Installing the "Connect" Capability .....	1-5
Uninstalling the Readiris Software .....	1-7
Readiris "uninstall" program .....	1-7
Windows (un)install wizard .....	1-7
Installing Related Products .....	1-9
Installed Files .....	1-11
Read Me files and documentation .....	1-11
User files .....	1-11
Register to Vote! .....	1-11
Getting Product Support .....	1-14

**Chapter 2: Guided Tour**

Starting the Software up .....	2-1
The First-Time Startup .....	2-2
Discovering the Readiris Interface .....	2-3
Getting Started with a First Tutorial .....	2-5
Zooming in on Images .....	2-7
One, Decomposing a Scanned Image .....	2-8
Two, Windowing a Scanned Image Manually .....	2-10
Three, Saving Windowing Templates .....	2-14
Readiris Takes You around the World .....	2-15
Readiris Changes Languages As Needed .....	2-18
Defining the Document Characteristics .....	2-20
Readiris Gets More Intelligent Each Time! .....	2-22
Learn .....	2-24
Don't Learn .....	2-24
Delete .....	2-25



Undo .....	2-25
Finish .....	2-25
The Role of Font Dictionaries .....	2-26
Where to Take the Text .....	2-29
Organizing the Text Output .....	2-31
Recognizing Multiple Pages .....	2-32
Setting up Your Scanner .....	2-34
Saving Default Settings .....	2-37
Saving Specific Settings .....	2-38
Recognizing Pages Automatically .....	2-39
Readiris Recreates Your Document Layout .....	2-39
Text Formatting, Part 2 .....	2-43
Placing the Graphics inside the Text .....	2-44
Saving Graphics Separately .....	2-46
Reading Faxes and Deferred Recognition .....	2-48
Recognizing Tables .....	2-50
The “Connect” Feature Offers a Direct Link .....	2-55
Getting On-line Help .....	2-58

## **Chapter 3: Command Reference**

Main Toolbar .....	3-1
Auto button .....	3-1
Scan button .....	3-2
Sort button .....	3-2
Recognize button .....	3-3
Language button .....	3-4
Source button .....	3-5
Page Analysis button .....	3-6
Learn button .....	3-7
Format button .....	3-8
Scanner button .....	3-11
Image Toolbar .....	3-13
Select Window button .....	3-13
Text Window button .....	3-14
Graphic Window button .....	3-15
Analyze Page button .....	3-16
Fit to Window button .....	3-17
Fit to Width button .....	3-18
Actual Size button .....	3-18
Rotate Left button .....	3-19

---

Rotate Right button .....	3-19
Turn Upside Down button .....	3-19
<b>File Menu .....</b>	<b>3-20</b>
Open .....	3-20
Save Graphics .....	3-21
Info .....	3-22
Select Source .....	3-23
Acquire .....	3-23
Load Layout .....	3-23
Save Layout .....	3-24
Load Settings .....	3-24
Save Settings .....	3-25
Save Default Settings .....	3-25
Exit .....	3-25
<b>Edit Menu .....</b>	<b>3-26</b>
Undo .....	3-26
Cut .....	3-26
Copy .....	3-27
Paste .....	3-27
Clear .....	3-27
Select All .....	3-28
<b>Settings Menu .....</b>	<b>3-28</b>
Scanner .....	3-28
Language .....	3-28
Font Type .....	3-28
Character Pitch .....	3-29
Text Format .....	3-30
View after Scan .....	3-30
Page Analysis .....	3-30
<b>View Menu .....</b>	<b>3-30</b>
Fit to Window .....	3-30
Fit to Width .....	3-30
50% Actual Size .....	3-31
Actual Size .....	3-31
200% Actual Size .....	3-31
Rotate Left .....	3-31
Rotate Right .....	3-31
Turn Upside Down .....	3-31
Coordinates .....	3-32
<b>Process Menu .....</b>	<b>3-32</b>
Automatic OCR .....	3-32
Scan .....	3-32
Sort .....	3-32
Recognize .....	3-32





Analyze Page .....	3-32
<b>Learn Menu .....</b>	<b>3-33</b>
Font Dictionary .....	3-36
New, Append, Read Font Dictionary .....	3-38
Interactive Learning .....	3-39
<b>Help Menu .....</b>	<b>3-39</b>
Help Topics .....	3-39
Register .....	3-40
Product Support .....	3-42
Contact I.R.I.S. ....	3-43
About Readiris .....	3-44

## **Appendix: Product description**

## **CREDITS AND COPYRIGHTS**

The Readiris software is designed and developed by I.R.I.S. OCR and Connectionist technology by I.R.I.S. I.R.I.S. retains the copyrights to the Readiris software, the OCR technology, the on-line help system and this manual.

Cardiris, Connectionist, the I.R.I.S. logo, Paper2File and Readiris are trademarks of I.R.I.S.

The spelling of this product is based on Proximity Linguistic Technology. Merriam Webster Inc. and Williams Collins Sons & Co. retain all rights to the Proximity/Merriam Webster Linguibase.

Acrobat Reader is a (registered) trademark of Adobe. Document Assistant is a (registered) trademark of Hewlett-Packard. Excel, Windows and Word are registered trademarks of Microsoft. Intel is a registered trademark of Intel. JetSuite is a (registered) trademark of JetFax. Lotus 1-2-3 is a registered trademark of IBM. PageSuite is a (registered) trademark of Microtek. PaperEase is a (registered) trademark of Primax. Visioneer is a (registered) trademark of Visioneer. WordPerfect is a registered trademark of Corel.