



**ScanSoft®**

# **OmniPage® SE**

**User's Guide**

## **LEGAL NOTICES**

Copyright © 2006 Nuance Communications, Inc. All rights reserved. No part of this publication may be transmitted, transcribed, reproduced, stored in any retrieval system or translated into any language or computer language in any form or by any means, mechanical, electronic, magnetic, optical, chemical, manual, or otherwise, without prior written consent from Nuance Communications, Inc., 1 Wayside Road, Burlington, Massachusetts 01803-4609. Printed in the United States of America and in Ireland. The software described in this book is furnished under license and may be used or copied only in accordance with the terms of such license.

### **IMPORTANT NOTICE**

Nuance Communications, Inc. provides this publication "As Is" without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability or fitness for a particular purpose. Some states or jurisdictions do not allow disclaimer of express or implied warranties in certain transactions; therefore, this statement may not apply to you. Nuance reserves the right to revise this publication and to make changes from time to time in the content hereof without obligation of Nuance to notify any person of such revision or changes.

### **TRADEMARKS AND CREDITS**

*Nuance, the Nuance logo, ScanSoft, OmniPage, PaperPort, True Page, Direct OCR, Logical Form Recognition, RealSpeak and ASR-1600* are registered trademarks or trademarks of Nuance Communications, Inc., in the United States of America and/or other countries. All other company names or product names referenced herein may be the trademarks of their respective holders.

### **THIRD PARTY LICENSES/NOTICES**

Please see acknowledgements/notices at the end of this guide.

### **Nuance Communications, Inc.**

1 Wayside Road  
Burlington, MA 01803-4609  
U.S.A.

### **Nuance Belgium BVBA**

Guldensporenpark 32  
BE-9820 Merelbeke  
Belgium

# CONTENTS

<b>WELCOME</b>	<b>5</b>
New features in OmniPage 15	7
<b>INSTALLATION AND SETUP</b>	<b>9</b>
System requirements	9
Installing OmniPage	10
Setting up your scanner with OmniPage	11
How to start the program	13
Registering your software	14
Activating OmniPage	15
Uninstalling the software	15
How to use OmniPage with PaperPort	16
<b>USING OMNIPAGE</b>	<b>17</b>
OmniPage Documents	17
The OmniPage Desktop	18
Basic Processing Steps	19
<b>PROCESSING DOCUMENTS</b>	<b>21</b>
Quick Start Guide	21
Processing methods	23
Manual processing	25
Processing with workflows	25
Processing from other applications	26
Processing with the Batch Manager	28
Defining the source of page images	29
Document to document conversion	31
Describing the layout of the document	32
Preprocessing Images	33
Image Enhancement Tools	35
Using Image Enhancement History	37
Saving and applying templates	37
Image Enhancement in Workflows	38

Zones and backgrounds	38
Table grids in the image	42
Using zone templates	42
<b>PROOFING AND EDITING</b>	<b>45</b>
The editor display and views	45
Proofreading OCR results	46
Verifying text	47
The Character Map	48
User dictionaries	49
Languages	50
Training	50
Text and image editing	52
On-the-fly editing	54
Reading text aloud	55
Working with Forms	56
<b>SAVING AND EXPORTING</b>	<b>59</b>
Saving and Exporting	59
Saving original images	60
Saving recognition results	61
Sending pages by mail	65
Other export targets	65
<b>WORKFLOWS</b>	<b>67</b>
Workflow Assistant	69
Batch Manager	73
Creating new jobs	74
Watched folders	78
Watched mailboxes	79
Barcode processing	80
Voice recognition	81
<b>TECHNICAL INFORMATION</b>	<b>83</b>
Troubleshooting	83
<b>INDEX</b>	<b>89</b>

# Welcome



Welcome to this OmniPage® 15 text recognition program, and thank you for choosing our software! The following documentation has been provided to help you get started and give you an overview of the program.

## This User's Guide

This guide introduces you to using OmniPage 15. It includes installation and setup instructions, a description of the program's commands and working areas, task-oriented instructions, ways to customize and control processing, and technical information.

This guide is written with the assumption that you know how to work in the Microsoft Windows environment. Please refer to your Windows documentation if you have questions about how to use dialog boxes, menu commands, scroll bars, drag and drop functionality, shortcut menus, and so on.

We also assume you are familiar with your scanner and its supporting software, and that the scanner is installed and working correctly before it is setup with OmniPage 15. Please refer to the scanner's own documentation as necessary.

## Online Help

OmniPage online Help contains information on features, settings, and procedures. It also has a comprehensive glossary, with its own alphabetical index and a table of contents. The online Help is provided as HTML help, and has been designed for quick and easy information retrieval. Online Help is available after you install OmniPage.



Press F1 as you are working with the program to see an online help topic relating to the current screen area, dialog box or warning message.

Comprehensive context-sensitive help aims to provide just enough assistance to let you keep working without delay. You can access the context-sensitive help in the following ways:



Click the Help tool in the Standard toolbar to get the help cursor. Click with this on any item on the desktop outside a dialog box or warning message.

Press Shift + F1 to get the same help cursor. Use Shift + F1 to get context-sensitive help for shortcut menu items.

Click the question mark button in the upper right corner of a dialog box and then click an item in the dialog box to see the popup window.

Some dialog boxes or warning messages have their own Help button, or a help text. Click the button or the text to get information on the dialog or message box.

Click anywhere to remove a context-sensitive popup Help window.

## **Readme File**

The Readme file contains last-minute information about the software. Please read it before using OmniPage. To open this HTML file, choose Readme in the OmniPage Installer or afterwards in the Help menu.

## **Scanning and other information**

The Nuance® web site at [www.nuance.com](http://www.nuance.com) provides timely information on the program. The Scanner Guide (<http://www.nuance.com/scannerguide/>) contains up-dated information about supported scanners and related issues; Nuance tests the 25 most widely used scanner models. Access Nuance's web site from the OmniPage 15 Installer or afterwards from the Help menu.

## **Tech Notes**

The web site at [www.nuance.com](http://www.nuance.com) contains Tech Notes on commonly reported issues using OmniPage 15. Web pages may also offer assistance on the installation process and troubleshooting.

# What's new in this release of OmniPage?

The new generation of the OmniPage product family offers a number of improvements and innovations compared to the previous version.

General improvements include:

- ◆ Faster OCR engine delivering higher accuracy
- ◆ Revised, intuitive user interface
- ◆ Improved character attributes and font matching
- ◆ Improved table conversion
- ◆ Better zoning and handling of colored backgrounds
- ◆ Better layout retention and document-level consolidation.



Features available in all versions of OmniPage - OmniPage SE 4.0, OmniPage 15 and OmniPage Professional 15

- ◆ Clearer and more precise suspect word and character display
- ◆ Ribbon Character map for easy insertion of non-keyboard characters during editing, and proofreading
- ◆ Image enhancement tools (SET tools) to improve OCR results and the quality of exported images.



Features available in OmniPage 15 and OmniPage Professional 15

- ◆ Workflow viewer enhancements make it easier to handle regular tasks.
- ◆ Extended PDF support up to version 1.5
- ◆ New supported file types including RTF 2000 ExactWord
- ◆ Better Workflow/Job distinction. A Job Wizard guides you through the job creating process.
- ◆ Folder and sub-folder input for specified file types when loading files and for workflows.
- ◆ Computer or OmniPage shutdown can be specified when a workflow or a job is finished.

- ◆ OmniPage Search Indexer plug-in with its own installer and help. It supports the Google Desktop Search utility for indexing image and PDF files.



### Features unique to OmniPage Professional 15

- ◆ Form recognition and handling: have form elements auto-detected or use form drawing tools for manual editing.
- ◆ Batch Manager provides more choice for job type, easier input of timing information and more precise schematic information for each job occurrence.
- ◆ Mailbox watching in Batch Manager jobs.
- ◆ Convert to PDF step in Batch Manager, using PDF Create! for direct conversion inside jobs.
- ◆ Support for sub-folders in folder watching.
- ◆ Customized actions can be set for your scanner Start button.
- ◆ Integrated programs PDF Converter 3 and PDF Create! 3 to unlock and create PDF files. Supports tagged, signed, encrypted PDF files, and much more. See their own documentation.
- ◆ Document-to-Document conversion to and from a wide range of formats.
- ◆ New and streamlined support for SharePoint 2003. See Loading image files from SharePoint and Saving to SharePoint.



This symbol denotes features not available in OmniPage SE, but available in OmniPage 15 and also in OmniPage Professional 15.



This icon is used throughout the guide to denote features that are available only in OmniPage Professional 15.

OmniPage 15 is supplied in Enterprise versions for network use. It is also supplied in Special Editions for selected scanner manufacturers and other resellers. The feature set in these editions may vary, in line with each vendor's requirements.

# Installation and setup



This chapter provides information on installing and starting OmniPage.

## System requirements

The minimum requirements to install and run OmniPage SE are:

- ◆ A computer with an Intel® Pentium® III processor or equivalent
- ◆ Microsoft® Windows® 98 (from second edition), Windows Me, Windows 2000 (from Service Pack 4), Windows XP or Windows Server 2003
- ◆ Microsoft Internet Explorer 5.5
- ◆ 128MB of memory (RAM), 256MB recommended
- ◆ 150MB of free hard disk space for application and sample files plus 60-65MB working space during installation.
- ◆ 5MB for Microsoft Installer (MSI) if not present (it is included in most Windows operating systems)
- ◆ Up to 5MB for system updates
- ◆ 800x600 pixel color monitor with 16-bit color or greater video card
- ◆ A CD-ROM drive for installation
- ◆ A Windows compatible pointing device
- ◆ A compatible scanner with its own scanner driver software, if you plan to scan documents. See the Scanner Guide at Nuance's web site ([www.nuance.com](http://www.nuance.com)) for a list of supported scanners
- ◆ Web access is needed for product registration, Scanner Wizard database updating and obtaining live updates for the program.

# Installing OmniPage

OmniPage 15's installation program takes you through installation with instructions on every screen.

## Before installing OmniPage:

- ◆ Close all other applications, especially anti-virus programs.
- ◆ Log into your computer with administrator privileges if you are installing on Windows 2000, XP or Server 2003.
- ◆ If you own a previous version of OmniPage, or if you are upgrading from demonstration software or an OmniPage Special Edition, the installer asks your consent to uninstall that product.

## ▼ To install OmniPage:

1. Insert the OmniPage CD-ROM in the CD-ROM drive. The installation program should start automatically. If it does not start, locate your CD-ROM drive in Windows Explorer and double-click the `Autorun.exe` program at the top-level of the CD-ROM.
2. Choose a language to use during installation. Accept the End-User License Agreement and enter the serial number shown on the CD envelope.
3. Choose a complete or a custom installation. A complete installation installs all RealSpeak™ Text-to-Speech language modules (currently 9). In OmniPage Professional 15, up to 7 ASR-1600™ Speech Recognition modules are installed. Custom installation lets you exclude or add modules. To exclude a module, click its down arrow and select 'This feature will be installed when required'. (Not applicable to OmniPage SE.)

4. Follow the instructions on each screen to install the software. All files needed for scanning are copied automatically during installation.

## Setting up your scanner with OmniPage

All files needed for scanner setup and support are copied automatically during the program's installation, but no scanner setup occurs at installation time. Before using OmniPage 15 for scanning, your scanner should be installed with its own scanner driver software and tested for correct functionality. Scanner driver software is not included with OmniPage.



Scanner setup is done through the Scanner Setup Wizard. You can start this yourself, as described below. Otherwise, it appears when you first attempt to perform scanning. Proceed as follows:

- ◆ Choose Start ▶ All Programs ▶ ScanSoft OmniPage 15.0 ▶ Scanner Setup Wizard
  - or click the Setup button in the Scanner panel of the Options dialog box.
  - or choose Scan in the Get Page drop-down list in the OmniPage Toolbox and click the Get Page button.
- ◆ The Scanner Setup Wizard starts. If you have a web connection, the first panel invites you to update the scanner database supplied with the wizard. Choose Yes or No and click on Next.
- ◆ Choose 'Select and test scanner or digital camera', then click Next. If you have a single installed scanner, it appears, along with any scanners previously set up with OmniPage. If the required scanner is not listed, click Add Scanner....
- ◆ You see a list of all detected scanner drivers in the checkmarked categories. This can include network devices. Select one and click OK. To install a second device, you must run the Scanner Wizard again.

- ◆ The wizard reports whether the chosen scanner model already has settings in the scanner database. If it does, you do not need to test it. If it does not, you should test it. Click on Next.
- ◆ If you chose not to test, click Finish. If you chose testing, click Next to have the scanner connection tested. If the connection is in order, you see a menu of further tests. Choose which testing steps you want to run. The Basic test scan is recommended.
- ◆ By default OmniPage uses its own scanning interface, located in the Scanner panel of the Options dialog box. If you want to use your scanner's own interface instead, choose Advanced settings... and select this. Click Hint editor... and choose Edit hints... only if you are experienced in configuring scanners or have been advised by Technical Support to do so.
- ◆ Click Next to start the tests. For the Basic scan test, insert a test page into your scanner. The wizard will scan using your scanner manufacturer's software. Click on Next. Your scanner's native user-interface will appear.
- ◆ Click on Scan to begin the sample scan.
- ◆ If necessary, click on Missing Image... or Improper Orientation... and make the appropriate selections.
- ◆ Once the image appears correctly in the window, click on Next.
- ◆ Move through the remaining requested tests, following the instructions on the screen.
- ◆ When all the requested tests have been completed successfully, the Scanner Wizard reports and invites you to click on Finish.
- ◆ You have successfully configured your scanner to work with OmniPage!

To change the scanner settings at a later time, or to setup or remove a scanner, reopen the Scanner Setup Wizard from the Windows Start menu or from the Scanner panel of the Options dialog box.

To test and repair an improperly functioning scanner, open the wizard and select 'Test the current scanner or digital camera' in the second panel, then work through the procedure described above, maybe using advice received from Technical Support.

To specify a different default scanner, open the wizard to reach the list of setup scanners. Move the highlight to the desired scanner and be sure to close the wizard with Finish.

To get updated settings for your current scanner, open the wizard, request a fresh database download in the first screen, then choose ‘Use current settings with current device’, click Next and then Finish.

## How to start the program

To start OmniPage do one of the following:

- ◆ Click Start in the Windows taskbar and choose All Programs ▶ ScanSoft OmniPage SE 4.0 ▶ OmniPage SE 4.0.
-  Double-click the OmniPage icon in the program's installation folder or on the Windows desktop if placed there.
-  Double-click an OmniPage Document (OPD) icon or file name; the clicked document is loaded into the program. See “OmniPage Documents” on page 17.
- ◆ Right click one or more image file icons or file names for a shortcut menu. Select Open With... OmniPage application. The images are loaded into the program.

On opening, OmniPage's title screen is displayed and then its desktop. See “The OmniPage Desktop” on page 18. It provides an introduction to the program's main working areas.

There are several ways of running the program with a limited interface:

- ◆ Use the Batch Manager program. Click Start in the Windows taskbar and choose All Programs ▶ ScanSoft OmniPage 15.0 ▶ OmniPage Batch Manager. See page 28. (Not applicable to OmniPage SE.)
- ◆ Click Acquire Text from the File menu of an application registered with the Direct OCRTM facility. See “How to set up Direct OCR” on page 26.
- ◆ Right-click on one or more image file icons or file names for a shortcut menu. Select OmniPage 15 and choose a target format or a workflow from its sub-menu. The files will be processed according to the workflow instructions. See “Workflows” on page 67. (Not applicable to OmniPage SE.)
- ◆ Click the OmniPage Agent icon on the taskbar. Choose a workflow to start the program and run the workflow. In OmniPage Professional 15, voice selection of workflow is possible. (Not applicable to OmniPage SE.)
- ◆ Use OmniPage with Nuance’s PaperPort® document management product, to add OCR services. See “How to use OmniPage with PaperPort” on page 16.

## Registering your software

Nuance’s online registration runs at the end of installation. Please ensure web access is available. We provide an easy electronic form that can be completed in less than five minutes. When the form is filled, click Submit. If you did not register the software during installation, you will be periodically invited to register later. You can go to [www.nuance.com](http://www.nuance.com) to register online. Click on Support and from the main support screen choose Register in the left-hand column. For a statement on the use of your registration data, please see Nuance’s Privacy Policy.

# Activating OmniPage

You will be invited to activate the product at the end of installation. Please ensure that web access is available. Provided your serial number is found at its storage location and has been correctly entered, no user interaction is required and no personal information is transmitted. If you do not activate the product at installation time, you will be invited to do this each time you invoke the program. OmniPage 15 can be launched only five times without activation. We recommend Automatic Activation.



Activation is not applicable to OmniPage SE.

## Uninstalling the software

Sometimes uninstalling and then reinstalling OmniPage will solve a problem. The OmniPage Uninstall program will not remove files containing recognition results or any of the following user-created files:

- Zone templates (\*.zon)
- Training files (\*.otn) (Not applicable to OmniPage SE.)
- User dictionaries (\*.ud)
- OmniPage Documents (\*.opd)
- Job files (\*.opj) (Not applicable to OmniPage SE.)
- Workflow files (\*.xwf) (Not applicable to OmniPage SE.)

To uninstall from Windows 2000, XP or Windows Server 2003 you must be logged into your computer with administrator privileges.

### ▼ To uninstall or reinstall OmniPage:

- ◆ Close OmniPage.
- ◆ Click Start in the Windows taskbar and choose the Control Panel and then Add/Remove Programs.
- ◆ Select OmniPage SE 4.0 and click Remove.
- ◆ Click Yes in the dialog box that appears to confirm removal.
- ◆ Select Yes to restart your computer immediately, or No if you plan to restart later.

- ◆ Follow instructions until the process is finished.

When you uninstall OmniPage, the link to your scanner is also uninstalled. You must setup your scanner again with OmniPage if you reinstall the program.

All RealSpeak and ASR modules and the that were installed with the program will also be uninstalled. (Not applicable to OmniPage SE.) ScanSoft PDF Create! 3 and ScanSoft PDF Converter 3 need to be uninstalled separately. (Not applicable to OmniPage SE.)

## How to use OmniPage with PaperPort



The PaperPort® program is a paper management software product from Nuance. It lets you link pages with suitable applications. Pages can contain pictures, text or both. If PaperPort exists on a computer with OmniPage, its OCR services become available and amplify the power of PaperPort. You can choose an OCR program by right-clicking on a text application's PaperPort link, selecting Preferences and then selecting OmniPage as the OCR package. OCR settings can be specified, as with Direct OCR.



PaperPort provides the easiest way to turn paper into organized digital documents that everybody in an office can quickly find and use. PaperPort works with scanners, multifunction printers, and networked digital copiers to turn paper documents into digital documents. It then helps you to manage them along with all other electronic documents in one convenient and easy-to-use filing system.

PaperPort's large, clear item thumbnails allow you to visually organize, retrieve and use your scanned documents, including Word files, spreadsheets, PDF files and even digital photos. PaperPort's Scanner Enhancement Technology tools ensure that scanned documents will look great while the annotation tools let you add notes and highlights to any scanned image.

# Using OmniPage

OmniPage 15 uses optical character recognition (OCR) technology to transform text from scanned pages or image files into editable text for use in your favorite computer applications.

In addition to text recognition, OmniPage can retain the following elements and attributes of a document through the OCR process.

*Graphics* (photos, logos)

*Form elements* (checkboxes, radio buttons, text fields) - available in OmniPage Professional 15 only

*Text formatting* (character and paragraph)

*Page formatting* (column structures, table formats, headings, placing of graphics).

## Documents in OmniPage

A document in OmniPage consists of one image for each document page. After you perform OCR, the document will also contain recognized text, displayed in the Text Editor, possibly along with graphics, tables and form elements.

## OmniPage Documents



An OmniPage Document (.opd) contains the original page images (optionally pre-processed) with any zones placed on them. After recognition, the OPD also contains the recognition results.

When saving, you have two file type choices: OmniPage Document or OmniPage Document (Extended). The latter allows you to embed a user dictionary, training file or zone template file in the OPD. This can increase file size considerably but makes the OPD more portable.

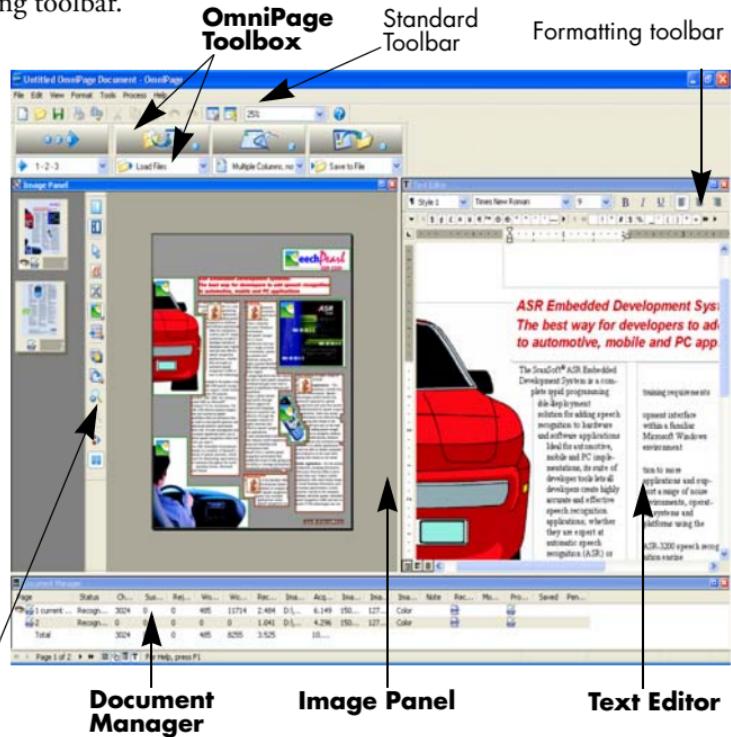


OmniPage SE does not support Extended OmniPage Documents.

When you open an OmniPage Document, its settings are applied, replacing those existing in the program.

## The OmniPage Desktop

The OmniPage Desktop has three main working areas, separated by splitters: the Document Manager, the Image Panel and the Text Editor. The Image Panel has an Image toolbar and the Text Editor has a Formatting toolbar.



**OmniPage toolbox:** This Toolbox lets you drive the processing.

**Document Manager:** This provides an overview of your document with a table. Each row represents one page. Columns present statistical or status information for each page, and (where appropriate) document totals.

**Image Panel:** This is displaying the image of the current page, together with its zones. The image panel can display the current page, thumbnails, or both. When this displays the current page image, the Image toolbar is available.

**Text Editor:** This is displaying the recognition results from the current page. The illustration shows True Page view.

## The Toolbars

The program has five main toolbars. Use the View menu to show, hide or customize them. The status bar at the bottom edge of the OmniPage program window explains the purpose of all tools.

**Standard toolbar:** Performs basic functions.

**Image toolbar:** Performs image, zoning and table operations.

**Formatting toolbar:** Formats recognized text in the Text Editor.

**Verifier toolbar:** Controls the location and appearance of the verifier.

**Reorder toolbar:** Modifies the order of elements in recognized pages.

**Form Drawing toolbar:** Creates new form elements.

**Form Arrangement toolbar:** Arranges and aligns form elements.



The Form toolbars appear only in OmniPage Professional 15.

## Basic Processing Steps

There are three ways of handling documents: with automatic, manual or workflow processing (latter not in SE). The basic steps for all processing methods are broadly the same:



**1. Bring a set of images into OmniPage.** You can scan a paper document with or without an Automatic Document Feeder (ADF) or load one or more image files.



**2. Perform OCR to generate editable text.** After OCR, you can check and correct errors in the document using the OCR Proofreader and edit the document in the Text Editor.



**3. Export the document to the desired location.** You can save your document to a specified file name and type, place it on the Clipboard, send it as a mail attachment or publish it. You can save the same document repeatedly to different destinations, different file types, with different settings and levels of formatting.

Using OmniPage, you can choose from the following processing methods: Automatic, Manual, Combined, or Workflow. You can start recognition from other applications, using the Direct OCR feature of OmniPage; and can also schedule processing to run at a later time.

Processing methods are detailed in the next chapter and in Online Help.

## Settings



The Options dialog box is the central location for OmniPage settings. Access it from the Standard toolbar or the Tools menu. Context-sensitive help provides information on each setting.

# Processing documents

This tutorial chapter describes different ways you can process a document and also provides information on key parts of this processing.

## Quick Start Guide

This topic takes you step-by-step through the basic OCR process.

You will process the document automatically and save the recognition results to a file. You will proof the document but will not edit it inside the Text Editor.

	What you do:	What happens:
1.	Set up your scanner using the Scanner Wizard, if this is not already done.	Configures OmniPage to work with your scanner.
2.	Select Start ▶ All Programs ▶ ScanSoft OmniPage SE 4.0 ▶ OmniPage SE 4.0	Opens OmniPage on your computer.
3.	Place the document correctly in your scanner.	
4.	From the Get Page drop-down list, select a scan option for your document: black-and-white, grayscale or color.	Allows you to determine how pictures or colored texts and backgrounds will look in the exported document. Color scanning needs a color scanner.
5.	From the Layout Description drop-down list, check Automatic is selected. For a wide range of documents, this is the best choice.	Configures the program how to place zones on the page and decide their properties automatically.

	What you do:	What happens:
6.	From the Export Results drop-down list, check that Save to File is selected.	This means you will be able to name your export file after you have proofed the document.
7.	Make sure 1-2-3 is selected in the Workflow drop-down list. Click the Start button.	OmniPage will start to scan in your document. A thumbnail appears with a progress indicator. The OCR Proofreader appears.
8.	Use the OCR Proofreader to modify words that the program suspects have not been recognized correctly.	The OCR Proofreader operates like a spell checker in a word processing program, but with added OCR-specific features. It removes markings from words you proof.
9.	Click in the Text Editor. Select Text Editor views one after another, to see how the page appears in each view.	Each Text Editor view defines a formatting level. This guides you which level to choose at saving time.
10.	Click Resume to restart proofing. When the message <i>OCR Proofreading is complete</i> appears, click on OK.	This ends the OCR Proofreader process. The Save to File dialog box will appear.
11.	Choose a file name, file type, path and a formatting level to save your recognized document. Click on OK.	By default, Save and Launch is enabled, so your document will be automatically opened in the word processing program associated with the file type that you selected.
12.	Inspect the document in your word processing program.	You have successfully used OmniPage to recognize your document and open it in your target application!

If you succeeded in getting good results from the sample image files, but not from the scanned page, check your scanner installation and settings: in particular brightness and image resolution. See “Input from scanner” on page 30. This provides a model of optimum brightness. See also the online Help topics *Setting up your scanner* and *Scanner troubleshooting*.

# Processing methods

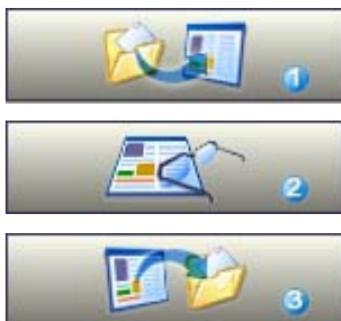
Using OmniPage, you can choose from the following processing methods:



## Automatic

A fast and easy way to process documents is to let OmniPage do it automatically for you. Select settings in

the Options dialog box and in the OmniPage Toolbox drop-down lists and then click *Start*. It will take each page through the whole process from beginning to end, when possible running in parallel. It will typically auto-zone the pages.



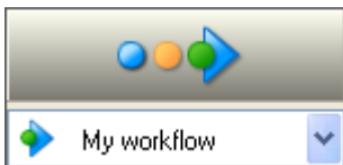
## Manual

Manual processing gives you more precise control over the way your pages are handled. You can process the document page-by-page with different settings for each page. The program also stops between each step: acquiring images, performing recognition, exporting. This lets you, for instance, draw zones manually or change

recognition language(s). You start each step by clicking the three buttons on the OmniPage Toolbox.

## Combined

You can process a document automatically and view results in the Text Editor. If most pages are in order, but a few have not turned out as expected, you can switch to manual processing to adjust settings and re-recognize just those problem pages. Alternatively, you can acquire images with manual processing, draw zones on some or all of them, and then send all pages to automatic processing.



## Workflow

A workflow consists of a series of steps and their settings. Typically it will include a recognition step, but it does not have to. Workflows are listed in the Workflow drop-down list – sample

workflows plus any you create. You can choose to place the OmniPage Agent icon on your taskbar. Its shortcut menu lists your workflows. Click a workflow to launch OmniPage and have it run.

Let the Workflow Assistant guide you in creating new workflows. It provides a choice of steps and the settings they need. After each step icon is selected and its settings defined, you get a new set of step icons to choose from. You can use the Assistant just to get more guidance when doing automatic processing. See “Workflow Assistant” on page 70.



Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15.

## In other applications

You can use the Direct OCR feature to call on the recognition services of OmniPage while working in your usual word-processor or similar application. See “How to set up Direct OCR” on page 26. OmniPage is automatically linked to the PaperPort document management program.

### At a later time



You can schedule OCR jobs or other processing jobs in

OmniPage Batch Manager to be performed automatically at a later time, when you may not even be present at your computer. This is done through the Batch Manager. When you choose New Job, first the Job Wizard, and then the Workflow Assistant appears - the latter with a slightly modified set of choices and settings. In the first panel of the Job Wizard, you define your job type and name your job; next you are to specify a starting time, a recurring job or watched folder instructions.

A job incorporates a workflow with timing instructions added. See “Batch Manager” on page 74.



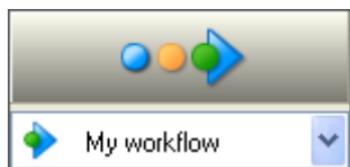
Batch Manager is only available in OmniPage 15 and its advanced features are offered only in OmniPage Professional 15.

## Manual processing

1. Manually zone pages where you want to process only part of the page or if you want to give precise zoning instructions. Use ignore backgrounds or zones to exclude areas from processing. Use process backgrounds or zones to specify areas to be auto-zoned.
2. Click the Start button, then choose *Finish Processing Existing Pages* in the Automatic Processing dialog box.
3. After proofing (if requested) you can save or export the document.

The default for manual processing is to have all entered pages automatically selected. This way you can have all new pages recognized by a single mouse click. You can remove this default in the Process panel of the Options dialog box.

## Processing with workflows



A workflow consists of a series of steps and their settings. It does not have to conform to the 1-2-3 pattern of traditional processing. Workflows allow you to handle recurring tasks more efficiently, because all the steps and their settings are pre-defined.



Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15.



### To run a workflow with OmniPage closed

Right-click on the OmniPage Agent icon in your taskbar. Select a workflow from its shortcut menu. OmniPage will start and immediately run the workflow. If you do not see the icon, enable it in the General panel of the Options dialog box.

The taskbar icon is not available in OmniPage SE.

## To run a workflow with OmniPage open

You can use the taskbar icon as described above, or you can select the workflow in the Workflow drop-down list and click Start. When a workflow is running, program settings are not accessible.



## To modify a workflow

Select the workflow in the Workflow drop-down list and press the Workflow Assistant button on the Standard toolbar, or choose Workflows... in the Tools menu, select the workflow and click Modify.

## To make a new workflow

There are sample workflows supplied with the program. You can modify these, or use them as the source for new workflows. New workflows are made with the Workflow Assistant. See page 70.

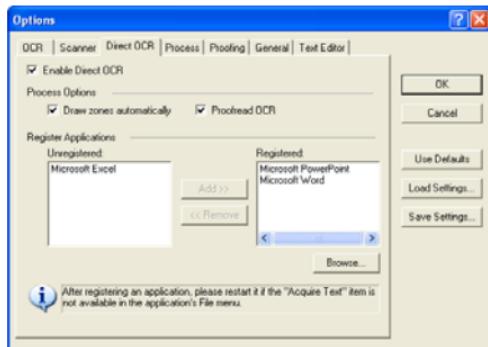
## Processing from other applications

You can use the Direct OCR™ feature to call on the recognition services of OmniPage while you work in your usual word-processor or other application. First you must establish the direct connection with the application. Then, two items in its File Menu open the door to OCR facilities.

## How to set up Direct OCR

1. Start the application you want connected to OmniPage. Start OmniPage, open the Options dialog box at the Direct OCR panel and select *Enable Direct OCR*.
2. Select process options for proofing and zoning. These function for future Direct OCR work until you change them again; they are not applied when OmniPage is used on its own.

3. The Unregistered panel displays running or previously unregistered applications. Select the desired one(s) and click Add. You can browse for an unlisted application.



## How to use Direct OCR

1. Open your registered application and work in a document. To acquire recognition results from scanned pages, place them correctly in the scanner.
2. Use the target application's File Menu item Acquire Text Settings... to specify settings to be used during recognition. Any settings not offered take their values from those last used in OmniPage. Settings changed for Direct OCR are also changed in OmniPage.

Acquire Text (OmniPage SE 4.0)...  
Acquire Text Settings (OmniPage SE 4.0)...  
▼

3. Use the File Menu item Acquire Text to acquire images from scanner or file.
4. If you selected *Draw zones automatically* in the Direct OCR panel of the Options dialog box, or under Acquire Text Settings..., recognition proceeds immediately.

5. If *Draw zones automatically* is not selected, each page image will be presented to you, allowing you to draw zones manually. Click the Perform OCR button to continue with recognition.
6. If proofing was specified, this follows recognition. Then the recognized text is placed at the cursor position in your application, with the formatting level specified by Acquire Text Settings... .

## Processing with the Batch Manager



Batch Manager is only available in OmniPage 15 and its advanced features are offered only in OmniPage Professional 15.



You can schedule processing jobs to be performed automatically at a specified time in the future. Unscheduled jobs can be activated manually. The job pages can come from a scanner with an ADF or from image files. You do not have to be present at your computer at job start time, nor does OmniPage have to be running. It does not matter if your computer is turned off after the job is set up, so long as it is running at job start time. If you are scanning pages, your scanner must be functioning at job start time, with the pages loaded in the ADF. Here is how to set up your first job:

1. Click Batch Manager... in the Process menu or in the Windows Start menu: select All Programs ▶ ScanSoft OmniPage 15.0 ▶ OmniPage Batch Manager. The Batch Manager window appears. Click the Create Job button to start the Job Wizard.
2. Select the type of your job in the next panel: Normal, Barcode driven, Folder Watching, Outlook mailbox watching, or Lotus Notes mailbox watching. The mailbox watching job types are only available if you have the given mail system configured properly on your computer.
3. Name your job in the same panel. Click Next.
4. Use the Start and Stop Options panel to specify your job timing and schedule. When the job is complete, you can choose to have the

input image file deleted or an e-mail notification sent to a given address (latter available in OmniPage Professional 15 only).

5. Define a starting point for the new job. This can be a fresh start, or an existing workflow. Click Next to finish each step.
6. The upcoming panels allow you to build the workflow for the job, as described in Chapter 6.
7. Click Finish to confirm job creation.

For more information, please see *Batch Manager* in the online Help and “Batch Manager” on page 74.

## Defining the source of page images

There are two possible image sources: from image files and from a scanner. There are two main types of scanners: flatbed or sheetfed. A scanner may have a built-in or added Automatic Document Feeder (ADF), which makes it easier to scan multi-page documents. The images from scanned documents can be input directly into OmniPage or may be saved with the scanner’s own software to an image file, which OmniPage can later open.

### Input from image files

You can create image files from your own scanner, or receive them by e-mail or as fax files. OmniPage can open a wide range of image file types. Select *Load Files* in the Get Pages drop-down list. Files are specified in the Load Files dialog box. This appears when you start automatic processing. In manual processing, click the Get Page button or use the Process menu. The lower part of the dialog box provides advanced settings, and can be shown or hidden.

The minimum width or height for an image file is 16 by 16 pixels; the maximum is 8400 pixels (71cm; 28 inches at the resolution 201 to 600 dpi). See online Help for pixel limits.



In OmniPage Professional 15, files can also be imported from FTP locations, Microsoft SharePoint, SharePoint 2003, or ODMA sources.

## **Input from scanner**

You must have a functioning, supported scanner correctly installed with OmniPage. You have a choice of scanning modes. In making your choice, there are two main considerations:

- ◆ Which type of output do you want in your export document?
- ◆ Which mode will yield best OCR accuracy?



### **Scan black and white**

Select this to scan in black-and-white. Black-and-white images can be scanned and handled quicker than others and occupy less disk space.



### **Scan grayscale**

Select this to use grayscale scanning. For best OCR accuracy, use this for pages with varying or low contrast (not much difference between light and dark) and with text on colored or shaded backgrounds.



### **Scan color**

Select this to scan in color. This will function only with color scanners. Choose this if you want colored graphics, texts or backgrounds in the output document. For OCR accuracy, it offers no more benefit than grayscale scanning, but will require much more time, memory resources and disk space.

## **Brightness and contrast**

Good brightness and contrast settings play an important role in OCR accuracy. Set these in the Scanner panel of the Options dialog box or in your scanner's interface. After loading an image, check its appearance. If characters are thick and touching, lighten the brightness. If characters are thin and broken, darken it. Then rescan the page.

If your scanning results are still not satisfactory, open the scanned image in the Image Enhancement window to edit it using a range of different tools.

## Scanning with an ADF

The best way to scan multi-page documents is with an Automatic Document Feeder (ADF). Simply load pages in the correct order into the ADF. You can scan double-sided documents with an ADF. A duplex scanner will manage this automatically.

## Scanning without an ADF

Using OmniPage's scanner interface, you can scan multi-page documents efficiently from a flatbed scanner, even without an ADF. Select *Automatically scan pages* in the Scanner panel of the Options dialog box, and define a pause value in seconds. Then the scanner will make scanning passes automatically, pausing between each scan by the defined number of seconds, giving you time to place the next page.

## Document to document conversion



A major new feature of OmniPage Professional 15 is that it can open not only image files, but also documents created in word-processing and similar applications. Supported file types include .doc, .xls, .ppt, .rtf, .wpd and others. Click the Load Files button in the OmniPage Toolbox or select the Load Files command under Get Page, in the File menu. In the Load Files dialog box, choose Documents.

When you are finished, you can use a variety of document file formats to save your files in.

# Describing the layout of the document

Before starting recognition you are requested to describe the layout of the incoming pages to assist the auto-zoning process. When you do automatic processing, auto-zoning always runs unless you specify a template that does not contain a process zone or background. When you do manual processing, auto-zoning sometimes runs. See online Help: *When does auto-zoning run?* Here are your input description choices:



## Automatic

Choose this to let the program make all auto-zoning decisions. It decides whether text is in columns or not, whether an item is a graphic or text to be recognized and whether to place tables or not.



## Single column, no table

Choose this setting if your pages contain only one column of text and no table. Business letters or pages from a book are normally like this.



## Multiple columns, no table

Choose this if some of your pages contain text in columns and you want this decolumnized or kept in separate columns, similar to the original layout.



## Single column with table

Choose this if your page contains only one column of text and a table.



## Spreadsheet

Choose this if your whole page consists of a table which you want to export to a spreadsheet program, or have treated as single table.

### Form



Choose this if your whole page consists of a form and you want form elements auto-recognized. After recognition, you can modify form element properties, create new ones, or edit form layout. This option is available in OmniPage Professional 15 only.

### Custom



Choose this for maximum control over auto-zoning. You can prevent or encourage the detection of columns, graphics and tables. Make your settings in the OCR panel of the Options dialog box.

### Template



Choose a zone template file if you wish to have its background value, zones and properties applied to all acquired pages from now on. The template zones are also applied to the current page, replacing any existing zones.

If auto-zoning yielded unexpected recognition results, use manual processing to rezone individual pages and re-recognize them.

## Preprocessing Images

To improve OCR results, you can enhance your images before zoning and recognition using the Image Enhancement tools. To open the Image Enhancement window, click the Enhance Image button in the Image Toolbar, or click Tools and choose Enhance Image.



You can also build Image Enhancement steps into your workflows by choosing the Enhance Images step. Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15.

The input for Image Enhancement is the Primary image.

We must distinguish three types of image:

**Original image:** The image created by your scanner or contained in a file before it enters the program.

**Primary image:** The state of the original image after it has been loaded into OmniPage, possibly modified by automatic or manual pre-processing operations.

**OCR image:** A black-and-white image derived from the primary image, optimized for good OCR results.

Some tools affect the Primary image, others the OCR image. Be sure you know which image you are editing.

Good brightness and contrast settings play an important role in OCR accuracy. Set these in the Scanner panel of the Options dialog box or in your scanner's interface. The diagram illustrates an optimum brightness setting. After loading an image, check its appearance. If characters are thick and touching, lighten the brightness. If characters are thin and broken, darken it. Use the OCR Brightness tool to optimize the image.

extensive research	Unsuitable
extensive research	Tolerable
extensive research	Good
extensive research	Best
extensive research	Good
extensive research	Tolerable
extensive research	Unsuitable

## Image Enhancement Tools

The Image Enhancement tools can also be used to edit images to save and use them as image files. Note that some tools of OmniPage work only on this, so-called Primary image, others on the one used for OCR (OCR image). Click the Primary/OCR Image button in the Image Enhancement window, to see the current state of either image.

The Image Enhancement window has two panels. The left panel shows the starting image. Your changes are shown in the right preview panel. When you click Accept, the right image is moved to the left panel to become the new starting image for further enhancement.

The following tools are accessible on the toolbar:



**Pointer** (F5) - the Pointer is a neutral tool carrying out different operations under different circumstances (for example, to pick a color for the Fill operation, or to catch the deskew line.)



**Zoom** (F6) - click the tool then use the left mouse button to zoom in on your image or the right mouse button to zoom out. You can also use the mouse wheel for zooming in and out - even in the inactive view. In the active view the "+" and "-" buttons serve the same purpose.



**Select Area** (F7) - click and draw your selection on the image to use a tool only on the selected area. (Image Enhancement Tools, by default, work on the whole page.) Selection has three modes (in the View menu):

*Normal* - you can select rectangular areas on the page, then move or resize the selection.

*Additive* - this mode enables you to make irregular selections by drawing overlapping rectangles that will be added to each other.

*Subtractive* - use this mode to cut out parts from your existing selections by drawing overlapping new areas.



**Primary/OCR Image** - click this tool to switch between the primary and the OCR image in the active view. Primary images can be of any image mode, while an OCR image is its black and white version, generated purely for OCR purposes.



**Synchronize Views** - click this tool to zoom and scroll the inactive view to the same zoom value and scroll position as the active view. To make the inactive view dynamically follow the focus of the active one, click View then choose the Keep Synchronized command.



**Brightness and Contrast** - click this tool to adjust the brightness and contrast of your primary image or a selected part of it. Use the sliders in the tool area to achieve the desired effect.



**Hue / Saturation / Lightness** - click this tool then use the sliders to modify the hue, saturation and lightness of your primary image.



**Crop** - if you decide to use only a given part of your image, click the Crop tool then select the area to keep and the rest of the image will be removed.



**Rotate** - click this tool to rotate (by 90, 180 or 270 degrees) and/or flip your image, or its selected area.



**Despeckle** - click this tool to remove stray dots from your image. Despeckle works on the OCR image at 4 levels. You can also use this tool not to remove noise from the page but to strengthen letter outlines: to do this mark the checkbox Inverse despeckling.



**OCR Brightness** - use this tool the set Brightness and Contrast of your OCR image. See the diagram on page 34.



**Dropout color** - click this tool and pick a color. Sections of the scanned image in this color will be set transparent. The tool has its effect on the OCR image.



**Resolution** - use this tool to decrease the resolution of your primary image in percentages. Note that you cannot adjust a resolution higher than that of the original one.



**Deskew** - sometimes pages are scanned crookedly. To straighten the lines of text manually, use the Deskew tool. (Auto-deskew is also available in the Process panel of Options.)



**Fill** - use this tool to apply uniform coloring to selected areas.

## Using Image Enhancement History

To commit or undo your image edits (one by one or all the steps), use the History panel in the Image Enhancement window. Once you have modified the original image, its preview displays the changes, but they are not done until you click the Apply button next to the History list. Modifications not added to the History by clicking the Add button will not be applied.

Any time you want to see what output a certain step resulted in, double click it in the History list.

To discard changes you have performed with a given tool, but before applying it, select the step in the list, then click the Reset button.

To restore the image as it was before you started the current enhancement session, click the Discard all changes button.

## Saving and applying templates



This feature is not available in OmniPage SE.

If you have a number of similar images to enhance, you can build up a list of enhancement steps to apply to all of them.

To create and store an image enhancement template, first bring an image file into the Image Enhancement window, then carry out your preprocessing steps and add them to the History clicking the Apply button. When you are done, choose Save Enhancement Template from

the File menu. Browse to your preferred destination and save the template file (with the extension .ipp).

To carry out the set of modifications saved in the template file on another image, simply open the new image in the Image Enhancement window and choose Load Enhancement Template from the File menu.

## Image Enhancement in Workflows



Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15.



To incorporate image enhancement in a workflow choose its icon in the Workflow Assistant. The following options are available:

**Display images for manual enhancement** - during the execution of a workflow, each loaded image will be displayed for manual editing.

**Apply enhancement template** - an already saved enhancement template will be applied automatically to the image while being processed by the workflow.

**Apply enhancement template and display** - the workflow will apply the selected image enhancement template, and will also display the image so that you can make further edits to it.

## Zones and backgrounds

Zones define areas on the page to be processed or ignored. Zones are rectangular or irregular, with vertical and horizontal sides. Page images in a document have a background value: process or ignore (the latter is more typical). Background values can be changed with the tools shown. Zones can be drawn on page backgrounds with the tools shown under Zone Types and Properties (see later).

**Process** areas (in process zones or backgrounds) are auto-zoned when they are sent to recognition.

**Ignore** areas (in ignore zones or backgrounds) are dropped from processing. No text is recognized and no image is transferred.

## Automatic zoning

Automatic zoning allows the program to detect blocks of text, headings, pictures and other elements on a page and draw zones to enclose them.

You can Auto-zone a whole page or a part of it. Automatically drawn zones and template zones have solid borders. Manually drawn or modified zones have dotted borders.



### Auto-zone a page background

Acquire a page. It appears with a process background. Draw a zone. The background changes to ignore. Draw text, table or graphic zones to enclose areas you want manually zoned. Click the Process background tool (shown) to set a process background. Draw ignore zones over parts of the page you do not need. After recognition the page will return with an ignore background and new zones round all elements found on the background.

## Zone types and properties

Each zone has a zone type. Zones containing text can also have a zone contents setting: alphanumeric or numeric. The zone type and zone contents together constitute the zone properties. Right-click in a zone for a shortcut menu allowing you to change the zone's properties. Select multiple zones with Shift+clicks to change their properties in one move.

The Image toolbar provides six zone drawing tools, one for each type.



### Process zone

Use this to draw a process zone, to define a page area where auto-zoning will run. After recognition, this zone will be replaced by one or more zones with automatically determined zone types.



### Ignore zone

Use this to draw an ignore zone, to define a page area you do not want transferred to the Text Editor.



### Text zone

Use this to draw a text zone. Draw it over a single block of text. Zone contents will be treated as flowing text, without columns being found.



### Table zone

Use this to have the zone contents treated as a table. Table grids can be automatically detected, or placed manually.



### Graphic zone

Use this to enclose a picture, diagram, drawing, signature or anything you want transferred to the Text Editor as an embedded image, and not as recognized text.



### Form zone

Use this to enclose an area of your document containing form elements such as a checkbox, radio button, text field or anything you want transferred to the Text Editor as a form element.



Afterwards, in True Page view, you can edit form layout, and modify the properties of form elements. Form zones are available in OmniPage Professional 15 only.

## Working with zones



The Image toolbar provides zone editing tools. One is always selected. When you no longer want the service of a tool, click a different tool. Some tools on this toolbar are grouped. Only the last selected tool from the group is visible. To select a visible tool, click it.

To **draw a single zone** select the zone drawing tool of the desired type, then click and drag the cursor.

To **resize a zone**, select it by clicking in it, move the cursor to a side or corner, catch a handle and move it to the desired location. It cannot overlap another zone.

To **make an irregular zone by addition** draw a partially overlapping zone of the same type.

To **join two zones of the same type** draw an overlapping zone of the same type (drawn zones on the left, resulting zone on the right).



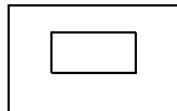
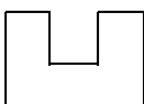
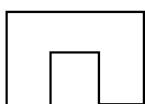
To make an irregular zone by subtraction draw an overlapping zone of the same type as the background.

To **split a zone** draw a splitting zone of the same type as the background.

A full set of zoning diagrams appear in the Online Help.

When you draw a new zone that partly overlaps an existing zone of a different type, it does not really overlap it; the new zone replaces the overlapped part of the existing zone.

The following zone types are prohibited:



**Speed zoning** lets you do manual zoning quickly. Activate the zone selection cursor, then move the cursor over the page image. Shaded areas

will appear showing the auto-detected zones. Double-click to transform a shaded area into a zone.

## Table grids in the image



After automatic processing you may see table zones placed on a page. They are denoted with a table zone icon in the top left corner of the zone. To change a rectangular zone to or from a table zone, use its shortcut menu. You can also draw table type zones, but they must remain rectangular.



You draw or move table dividers to determine where gridlines will appear when the table is placed in the Text Editor. You can draw or resize a table zone (provided it stays rectangular) to discard unneeded columns or rows from the outer edges of a table.



Using the table tools you can insert row and column dividers; move and remove dividers. Click the Place/Remove all dividers tool to have dividers in a table auto-detected and placed.



You can specify line formatting for table borders and grids from a shortcut menu. You will have greater choice for editing borders and shading in the Text Editor after recognition.

## Using zone templates

A template contains a page background value and a set of zones and their properties, stored in a file. A zone template file can be loaded to have template zones used during recognition. Load a template file in the Layout Description drop-down list or from the Tools menu. You can browse to network locations to load templates created by others.

When you load a template, its background and zones are placed:

- ◆ on the current page, replacing any zones already there
- ◆ on all further acquired pages
- ◆ on pre-existing pages sent to (re-)recognition without any zones.

With manual processing the template zones in the first two cases can be viewed and modified before recognition.

With automatic processing the template zones can be viewed and modified only after recognition.

With workflow processing, use the zone images step. This combines two steps: load templates and manual zoning. To use a zone template, click the Add button in the appropriate panel of the Workflow Assistant, and select the zone template file to use. Then make your choice between displaying images for manual zoning; applying the zone template; or applying it and display the images.



Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15.

Templates accept ignore and process zones and backgrounds. They can therefore be useful to define which parts of the pages to process with auto-zoning, and which parts to ignore. Process zones or process background areas from a template may be replaced during recognition by a set of smaller zones; specific zone types will be assigned to these zones.

## How to save a zone template

Select a background value and prepare zones on a page. Check their locations and properties. Click Zone Template... in the Tools menu. In the dialog box, select [zones on page] and click Save, then assign a name and optionally a different path. Choose a network location to share the template file. Click OK. The new zone template remains loaded.

## How to modify a zone template

Load the template and acquire a suitable image with manual processing. The template zones appear. Modify the zones and/or properties as desired. Open the Zone Template Files dialog box. The current template is selected. Click Save and then Close.

## **How to unload a template**

Select a non-template setting in the Layout Description drop-down list. The template zones are not removed from the current or existing pages, but template zones will no longer be used for future processing. You can also open the Zone Template Files dialog box, select [none] and click the Set As Current button. In this case, the layout description setting returns to Automatic.

## **How to replace one template with another**

Select a different template in the Layout Description drop-down list, or open the Zone Template Files dialog box, select the desired template and click the Set As Current button. Zones from the new template are applied to the current page, replacing any existing zones. They are also applied as explained above.

## **How to remove a template file**

Open the Zone Template Files dialog box. Select a template and click the Remove button. Zones already placed by this template are not removed. Template files can be deleted only from the operating system.

## **How to include a template file in an OPD**

Load the template, then click the Save button in the Standard toolbar and choose the file type OmniPage Document (Extended). That means the template will travel with the OPD if it is sent to a new location. When the extended OPD file is opened later, the included zone template will be shown in the Zone Template dialog box as [embedded] and can be saved to a new named template file at the new location.



OmniPage SE does not support Extended OmniPage Documents.

# Proofing and editing



Recognition results are placed in the Text Editor. These can be recognized texts, tables, forms and embedded graphics. This WYSIWYG (What You See Is What You Get) editor is detailed in this chapter.

## The editor display and views

The Text Editor displays recognized texts and can mark words that were suspected during recognition with red, wavy underlines. They are displayed with red characters in the OCR Proofreader.

A word may be suspect because it was not found in any active dictionary: standard, user or professional. It may also be suspect as a result of the OCR process, even if it is found in the dictionary. If the uncertainty stems from certain characters in the word, these are shown with a yellow highlight, both in the Editor and the OCR Proofreader.

Choose to have non-dictionary words marked or not in the Proofing panel of the Options dialog box. All markers can be shown or hidden as selected in the Text Editor panel of the Options dialog box. You can also show or hide non-printing characters and header/footer indicators. The Text Editor panel also lets you define a unit of measurement for the program and a word wrap setting for use in all Text Editor views except Plain Text view.

OmniPage can display pages with three levels of formatting. You can switch freely between them with the three buttons at the bottom left of the Text Editor or from the View menu.



### Plain Text view

This displays plain decolumnized left-aligned text in a single font and font size, with the same line breaks as in the original document.



### Formatted Text view

This displays decolumnized text with font and paragraph styling.



### True Page view

True Page® view tries to conserve as much of the formatting of the original document as possible. Character and paragraph styling is retained. Reading order can be displayed by arrows.

## Proofreading OCR results

After a page is recognized, the recognition results appear in the Text Editor. Proofreading starts automatically if that was requested in the Proofing panel of the Options dialog box. You can start proofing manually any time. Work as follows:

1.  Click the Proofread OCR tool in the Standard toolbar, or choose Proofread OCR... in the Tools menu.
2. Proofing starts from the current page, but skips text already proofed. If a suspected error is detected, the OCR Proofreader dialog box colors the suspect word in its context, adds a yellow highlight to any suspect characters and provides a picture of how the word originally looked in the image. The explanation says 'Suspect word' or 'Non-dictionary word'.
3. If the recognized word is correct, click Ignore or Ignore All to move to the next suspect word. Click Add to add it to the current user dictionary and move to the next suspect word.
4. If the recognized word is not correct, modify the word in the Edit panel or select a dictionary suggestion. Click Change or Change All to implement the change and move to the next suspect word. Click

Add to add the changed word to the current user dictionary and move to the next suspect word.

5. Color markers are removed from words in the Text Editor as they are proofread. You can switch to the Text Editor during proofing to make corrections there. Use the Resume button to restart proofing. Click Page Ready to skip to the next page and Document Ready or Close to stop proofreading before the end of the document is reached.
6. A page is marked with the proofed icon  on its thumbnail and in the Document Manager if proofing ran to the end of the page.



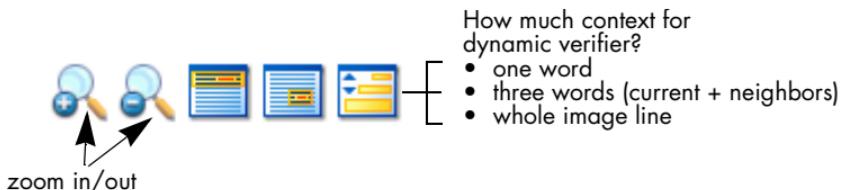
Voice-driven proofing is available in OmniPage Professional 15. See “Voice recognition” on page 82. The proofreader’s suggestions are numbered. Speak the number of the suggestion you want to accept.

## Verifying text

After performing OCR, you can compare any part of the recognized text against the corresponding part of the original image, to verify that the text was recognized correctly.



The verifier tool is in the Formatting toolbar. The verifier can also be controlled from the Tools menu. Hover the cursor over a verifier display to obtain the verifier toolbar. Use it as follows:



To turn the Verifier on, click the Verifier tool or press F9. To turn it off, click the Verifier tool again, press F9 again, or press Esc.

A full list of verifier keyboard shortcuts is available in the Online Help.

## The Character Map



The Character Map is a dockable tool giving you aid in proofing. It is used for essentially two purposes:

- ◆ to insert characters during proofing, and editing that are not or not easily accessible from your keyboard. In this respect, it is very similar to the system Character Map.
- ◆ to show all characters validated by the current recognition languages. (Not applicable to OmniPage SE.)

To access the Character Map, click its button in the Formatting Toolbar, or choose Character Map from the View menu and click Show.

Under the Character Map menu item, you have additional options:

- ◆ Recent Characters Only: click this option to display only the 36 recently used characters in the formatting toolbar. This is useful if you work with a limited set of characters to be inserted.
- ◆ Character Sets: choose this, then select all the character sets that you want displayed in the character map.

You can access the Character Map in other ways, such as:

- ◆ Click Tools > Options and choose the OCR tab. Click the Additional Characters button to select characters to be included in proofing. Similarly, you can modify the Reject Character by using the Character Map.
- ◆ Select Train Character under the Tools menu. The Character Map will display when you click the (...) button beside the Correct field.
- ◆ Select Train Character from the shortcut menu of a suspect, or non-dictionary word in the Text Editor.



The above three ways to access the Character Map are not available in OmniPage SE.

## User dictionaries

The program has built-in dictionaries for many languages. These assist during recognition and may offer suggestions during proofing. They can be supplemented by user dictionaries. You can save any number of user dictionaries, but only one can be loaded at a time. A dictionary called Custom is the default user dictionary for Microsoft Word.

### Starting a user dictionary

Click Add in the OCR Proofreader dialog box with no user dictionary loaded or open the User Dictionary Files dialog box from the Tools menu and click New.

### Loading or unloading a user dictionary

Do this from the OCR panel of the Options dialog box or from the User Dictionary Files dialog box.

### Editing or removing a user dictionary

Add words by loading a user dictionary and then clicking Add in the OCR Proofreader dialog box. You can add and delete words by clicking Edit in the User Dictionary Files dialog box. You can also import words from OmniPage user dictionaries (\*.ud). While editing a user dictionary, you can import a word list from a plain text file to add words to the dictionary quickly. Each word must be on a separate line with no punctuation at the start or end of the word. The Remove button lets you remove the selected user dictionary from the list.



OmniPage SE does not support importing and exporting User Dictionaries.

To embed a user dictionary in an OmniPage Document, load it and save to the file type OmniPage Document (Extended).



OmniPage SE does not support Extended OmniPage Documents.

## Languages

The program can read over 50 languages with three alphabets: Latin, Greek and Cyrillic. See the list in the OCR panel of the Options dialog box. It shows which languages have dictionary support. A listing is also provided on the Nuance web site.

In addition to user dictionaries, specialized dictionaries are available for certain professions (currently medical, legal and financial) for some languages. See the list and make selections in the OCR panel of the Options dialog box.

Legal and Medical dictionaries are only available in OmniPage 15. Financial dictionaries are only available in OmniPage Professional 15.0.

## Training



Intelligent proofing (IntelliTrain), character training and training files are only available in OmniPage 15.

Training is the process of changing the OCR solutions assigned to character shapes in the image. It is useful for uniformly degraded documents or when an unusual typeface is used throughout a document. OmniPage 15 offers two types of training: manual training and automatic training (IntelliTrain). Data coming from both types of training are combined and available for saving to a training file.

When you leave a page on which training data was generated, you will be asked how to apply it to other existing pages in the document.

### Manual training

To do manual training, place the insertion point in front of the character you want to train, or select a group of characters (up to one word) and choose Train Character... from the Tools menu or the shortcut menu. You will see an enlarged view of the character(s) to be trained, along with the current OCR solution. Change this to the desired solution and click OK. The program takes this training and examines the rest of the page. If it

finds candidate words to change, the Check Training dialog box lists these. Incorrect words should be re-trained before the list is approved.

## **IntelliTrain**

IntelliTrain is an automated form of training. It takes input from the corrections you make during proofing. When you make a change, it remembers the character shape involved, and your proofing change. It searches other similar character shapes in the document, especially in suspect words. It assesses whether to apply the user correction or not.

You can turn IntelliTrain on or off in the OCR panel of the Options dialog box.

IntelliTrain remembers the training data it collects, and adds it to any manual training you have done. This training can be saved to a training file for future use with similar documents.

For examples of IntelliTrain, see the Online Help.

## **Training files**

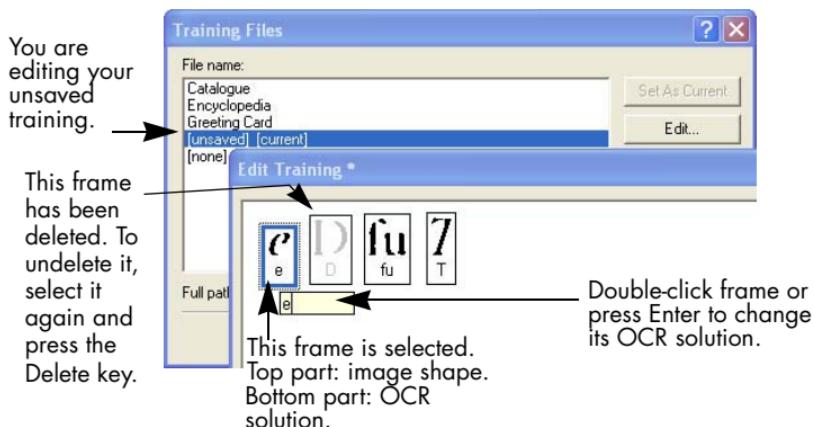
If you want to be prompted to save your unsaved training data when you close OmniPage, select that option in the Proofing panel of the Options dialog box. Unsaved training data is stored in an OmniPage Extended Document. If you do not save the training data, it is discarded when OmniPage is closed. To save a training file into an OPD, load it and save to the file type OmniPage Document (Extended).

Saving training to file, loading, editing and unloading training files are all done in the Training Files dialog box.

Unsaved training can be edited in the Edit Training dialog box, an asterisk is displayed in the title bar in place of a training file name. Save it in the Training Files dialog box.

A training file can be also edited; its name appears in the title bar. If it has unsaved training added to it, an asterisk appears after its name. Both the unsaved and the modified training are saved when you close the dialog box.

The Edit Training dialog box displays frames containing a character shape and an OCR solution assigned to that shape. Click a frame to select it. Then you can delete it with the Delete key, or change the assignation. Use arrow keys to move to the next or previous frame.



## Text and image editing

OmniPage has a WYSIWYG Text Editor, providing many editing facilities. These work very similarly to those in leading word processors.

### Editing character attributes

In all views except Plain Text view, you can change the font type, size and attributes (bold, italic, underlined) for selected text.

### Editing paragraph attributes

In all views except Plain Text view, you can change the alignment of selected paragraphs and apply bulleting to paragraphs.

### Paragraph styles

Paragraph styles are auto-detected during recognition. A list of styles is built up and presented in a selection box on the left of the Formatting toolbar. Use this to assign a style to selected paragraphs.

## Graphics

You can edit the contents of a selected graphic if you have an image editor in your computer. Click Edit Picture With in the Format menu. Here you can choose to use the image editor associated with BMP files in your Windows system, and load the graphic. Alternatively, you can use the Choose Program... item to select another program. This will replace the Default Image Editor item. Edit the graphic, then close the editor to have it re-embedded in the Text Editor. Do not change the graphic's size, resolution or type, because this will prevent the re-embedding. You can also edit images before recognition using the Image Enhancement tools.

## Tables

Tables are displayed in the Text Editor in grids. Move the cursor into a table area. It changes appearance, allowing you to move gridlines. You can also use the Text Editor's rulers to modify a table. Modify the placement of text in table cells with the alignment buttons in the Formatting toolbar and the tab controls in the ruler.

## Hyperlinks

Web page and e-mail addresses can be detected and placed as links in recognized text. Choose Hyperlink... in the Format menu to edit an existing link or create a new one.

## Editing in True Page

Page elements are contained in text boxes, table boxes and picture boxes. These usually correspond to text, table and graphic zones in the image. Click inside an element to see the box border; they have the same coloring as the corresponding zones. The online Help topic *True Page* provides details on the operations summarized here.

**Frames** have gray borders and enclose one or more boxes. They are placed when a visible border is detected in an image. Format frame and table borders and shading with a shortcut menu or by choosing Table... in the Format menu. Text box shading can be specified from its shortcut menu.

**Multicolumn areas** have orange borders and enclose one or more boxes. They are auto-detected and show which text will be treated as flowing columns when exported with the Flowing Page (not in OmniPage SE) formatting level.



**Reading order** can be displayed and changed. Click the Show reading order tool in the Formatting toolbar to have the order shown by arrows. Click again to remove the arrows.



Click the Change reading order tool for a set of reordering buttons in place of the Formatting toolbar. A changed order is applied in Plain Text and Formatted Text views. It modifies the way the cursor moves through a page when it is exported as True Page.

## On-the-fly editing

This allows you to modify a recognized page through re-zoning, without having to re-process the whole page. When on-the-fly editing is enabled, zone changes (deleting, drawing, resizing, changing type) immediately make changes in the recognized page. Conversely, when you modify elements in the Text Editor's True Page view, this changes the zones on that page.

Two linked tools on the Image toolbar control on-the-fly zoning. One of these tools is always active whenever no recognition is in progress.



Click this to activate on-the-fly editing. The red signal shows there are no stored zoning changes.



Click this to turn on-the-fly editing off. Your zoning changes are stored; the on-the-fly tool displays a green signal to show there are stored changes. To activate these changes, do one of the following:



Click the on-the-fly tool with a green signal. The zoning changes will cause changes in the Text Editor.



Click the Perform OCR button to have the whole page (re)recognized, including your zone changes.

For details on how changes are handled in on-the-fly zoning and their effects in the Text Editor views, see *On-the-fly processing* in online Help.

## Reading text aloud



The Text-to-Speech facility and the saving to WAV audio files are not included in OmniPage SE. They are available in OmniPage 15.

The ScanSoft RealSpeak™ speech facility is provided for the visually impaired, but it can also be useful to anyone during text checking and verification. The speaking is controlled by movements of the insertion point in the Text Editor which can be mouse or keyboard driven.

To hear text:	Use these keys:
One character at a time, forward or back	Right or left arrow. Letter, number or punctuation names are spoken.
Current word	Ctrl + Numpad 1
One word to the right	Ctrl + right arrow
One word to the left	Ctrl + left arrow
A single line	Place the insertion point in the line
Next line	Down arrow
Previous line	Up arrow
Current sentence	Ctrl + Numpad 2
From insertion point to end of sentence	Ctrl + Numpad 6
From start of sentence to insertion point	Ctrl + Numpad 4
Current page	Ctrl + Numpad 3
From top of current page to insertion point	Ctrl + Home
From insertion point to end of current page	Ctrl + End

Previous, next or any page	Ctrl + PgUp, PgDown or navigation buttons
Typed characters	Each typed character is pronounced separately.

The Text-to-Speech facility is enabled or disabled with the Tools menu item Speech Mode or with the F5 key. A second menu item Speech Settings... allows you to select a voice (for example, male or female for a given language), a reading speed and the volume. You must ensure the language selection is appropriate for the text you want to hear.

You also have the following keyboard controls:

To do this:	Use this:
Pause/Resume	Ctrl + Numpad 5
Set speed higher	Ctrl + Numpad +
Set speed lower	Ctrl + Numpad -
Restore speed	Ctrl + Numpad *

All speech systems will be installed with OmniPage 15 if you choose a complete installation. If you perform a custom installation, you can choose the languages you need.

## Working with Forms



You can bring paper or electronic forms (distributed mainly as PDF in an office environment) into OmniPage Professional 15, recognize them and edit their content, layout or both - in True Page view. Draw form zones over the relevant areas of your image before recognition, or choose

Form as recognition layout, then use the two toolbars: Form Drawing and Form Arrangement to make modifications and produce a fillable form and save it in the following formats: PDF, RTF, or XSN (Microsoft Office InfoPath 2003 format). Static forms can be saved to HTML. OmniPage

Professional 15 uses the Logical Form Recognition™ technology to process forms.

Please note that OmniPage supports form creation and editing, however the tools available here are not designed to fill in forms.

## The Form Drawing Toolbar

This is a dockable toolbar, displayed in the Text Editor that allows you to create a range of form elements using the following tools:



**Selection:** Click this tool to be able to select, move, or resize elements in your form.



**Text:** Use the text tool to add fixed text descriptions on your form such as titles, labels and headers.



**Line:** The Line tool is mainly used in layout design: click it and draw lines to separate distinct sections in your form.



**Rectangle:** Click this tool to create rectangles in your form for design purposes.



**Graphic:** Use this tool to select areas of your form that are to be treated as graphics.



**Fill text:** Click this tool to create fillable text fields. These are fields where you want people to enter text.



**Comb:** Use this tool to create a text field consisting of boxes. This is typically used for information such as ZIP codes.



**Checkbox:** Click this tool and draw Checkboxes - typically for Yes/No questions and marking one or more choices.



**Circle text:** Its function is similar to the Checkbox element (above): the Circle text tool creates elements that get encircled when selected.



**Table:** This tool creates tables in your form.

You can also create form elements by right-clicking an existing form element in your recognized form, and choose the Insert Form Object menu item.

## The Form Arrangement Toolbar

The tools on this toolbar can be used to line up form elements or to set which one is on top of the others when they overlap. This latter function is useful for example if you want to create a background graphic design for your form.

 To set the order of overlapping elements, use the “Bring to Front” and “Send to Back” buttons.

To align the right/left, top/bottom edges or the centers of the selected form elements

 horizontally - use the horizontal alignment tools

 vertically - use the vertical arrangement tools.

The commands of the Form Arrangement toolbar are also accessible from the shortcut menu of any form element.

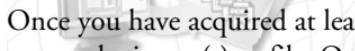
## Editing Form object properties

To edit a form object directly select it then right-click the given element to display its shortcut menu. You can edit the appearance or the properties of any form element here. Use the following commands:

Form Object Appearance - use the tabs Borders, Shading and Shadow to design the look of your form elements in a similar way as you would do in a text-editing application.

Form Object Properties - this command gives you access to the element properties such as size, position, name. Note that properties dynamically vary depending on what type of an element you select.

# Saving and exporting



Once you have acquired at least one image for a document, you can export the image(s) to file. Once you have recognized at least one page, you can export recognition results – a single page, selected pages or the whole document – to a target application by saving to file, copying to Clipboard or sending to a mailing application. Saving as an OmniPage Document is always possible.

A document remains in OmniPage after export. This allows you to save, copy or send its pages repeatedly, for example with different formatting levels, using different file types, names or locations. You can also add or re-recognize pages or modify the recognized text.

With automatic processing and in Batch Manager jobs, you specify where to save first before processing starts.

A workflow may contain one or more saving steps, even to different targets (for instance, to file and to mail). A Batch Manager job must contain at least one saving step. See Chapter 6, “Workflows”.



Since neither workflows nor the Batch Manager are supported in OmniPage SE, all references to these in this chapter should be ignored.

## Saving and Exporting

If you want to work with your document again in OmniPage in a later session, save it as an OmniPage Document. This is a special output file type. It saves the original images together with the recognition results, settings and training.

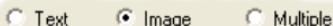
Exporting is done through button 3 on the OmniPage Toolbox. It lists available export targets. The picture on the left shows all possible targets.

The last four appear only in OmniPage Professional 15. Some appear only if access to the target is detected on your computer. Select the desired target then click the Export Results button to begin export. You can also perform exporting through the Process menu.

## Saving original images

You can save original images to disk in a wide variety of file types with or without image enhancement (using the Image Enhancement Tools).

1. Choose Save to File in the Export Results drop-down list. In the dialog box that appears, select Image under Save as.



2. Choose a folder location and a file type. Type in a file name.
3. Select to save the selected zone image(s) only, the current page image, selected page images or all images in the document. For multiple zones or multiple pages, you can have all images in a single multi-page image file, providing you set TIFF, MAX, DCX, JB2 or Image-only PDF as file type. Otherwise each image is placed in a separate file. OmniPage adds numerical suffixes to the file name you provide, to generate unique file names.
4. Click Converter Options... if you want to specify a saving mode (black-and-white, grayscale, color or 'As is'), a maximum resolution and other settings. For TIFF files, you specify the compression method here.
5. Click OK to save the image(s) as specified. Zones and recognized text are not saved with the file.

## Saving recognition results

You can save recognized pages to disk in a wide variety of file types.

1. Choose Export Results... in the File menu, or click the Export Results button in the OmniPage Toolbox with Save to File selected in the drop-down list.
2. The Save to File dialog box appears. Select Text under Save as.
3. Select a folder location and a file type for your document. Select a page range, file options, naming options and a formatting level for the document. See “Selecting a formatting level” on this page.
4. Type in a file name. Click Converter Options... if you want to specify precise settings for the export. See “Selecting converter options” on page 63.
5. Click OK. The document is saved to disk as specified. If *Save and Launch* is selected, the exported file will appear in its target application; that is the one associated with the selected file type in your Windows system or in the advanced saving options for your selected file type converter.

## Selecting a formatting level

The formatting level for export is defined at export time, in the saving dialog box (Save to File, Copy to Clipboard, Send in Mail or other dialog box). Three of the levels correspond to the format views of the same name in the Text Editor. However, the level to be applied for saving is independent of the formatting view displayed in the Text Editor. When exporting to file or mail, first specify a file type. This determines which formatting levels are available.

The formatting levels are:



### **Plain Text**

- This exports plain decolumnized left-aligned text in a single font and font size. When exporting to Text or Unicode file types, graphics and tables are not supported.

You can export plain text to nearly all file types and target applications; in these cases graphics, tables and bullets can be retained.



- This exports decolumnized text with font and paragraph styling, along with graphics and tables. This is available for nearly all file types.



### **Flowing Page**

This keeps the original layout of the pages, including columns. This is done wherever possible with column and indent settings, not with text boxes or frames. Text will then flow from one column to the other, which does not happen when text boxes are used.



- This keeps the original layout of the pages, including columns. This is done with text, picture and table boxes and frames. This is offered only for target applications capable of handling these. True Page formatting is the only choice for XML export and for all PDF export, except to the file type 'PDF Edited'.



### **Spreadsheet**

- This exports recognition results in tabular form, suitable for use in spreadsheet applications. This places each document page onto a separate worksheet.

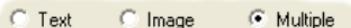
When exporting to Microsoft Excel, 'Spreadsheet' is good for saving whole-page tables. Prefer 'Formatted Text' if your document contains smaller tables: each table will be placed on a separate worksheet with non-table parts placed in an index worksheet with hyperlinks to each relevant worksheet.

## Selecting converter options

Click the Converter Options... button in a saving dialog box to have precise control over the export. This brings up a dialog box with the name of the converter associated with the current file type. It presents a series of options tailored to this file type. First, confirm or change the formatting level, because this influences which other options are presented. Select options as desired. Online Help details how to do this.

## Using multiple converters

Multiple converters allow you to export to two or more file types in one export step. Choose Multiple in the saving dialog box:



To make your own multiple converter, open the Export Converters dialog box from the Tools menu. Choose the heading Multiple converters. Select a converter and click Create from.... This will make a copy of the selected converter that you can freely modify without overwriting the original one.

The new converter appears in the list. Select it and click Options... to specify its settings. You receive a list of all text converters, followed by all image converters. Checkmark the desired ones. Optionally specify sub-folder paths for each file type.

You can save pages with different formatting levels or file options to the different file types, as defined in their simple converters. A few saving operations cannot be done with multiple converters. These are:

### Saving OmniPage Documents

Use a workflow with two saving steps, or perform two separate saves.

### Saving to two targets

For instance, you cannot use a multiple converter to save a document to

file and also send it in mail. Use a workflow with two saving steps, or perform two separate saves.

### **Saving different page ranges**

You cannot save different page ranges to different file types, because only one set of selected pages can exist at saving time. For the same reason, a single workflow cannot be used either. Perform two separate saves or use two workflows.

## **Saving to PDF**



OmniPage SE does not support opening or saving of PDF files. You have five choices when saving to Portable Document Format (PDF) files. The first four are presented as Text converters, the last one is listed among the Image converters.

### **PDF (Normal):**

Pages are exported as they appeared in the Text Editor in True Page view. The PDF file can be viewed and searched in a PDF viewer and edited in a PDF editor.

### **PDF Edited:**

Use this if you have made significant editing changes in the recognition results. You have three formatting level choices, including True Page. The PDF file can be viewed, searched and edited.

### **PDF with image on text:**

The PDF file is viewable only and cannot be modified in a PDF editor. The original images are exported, but there is a linked text file behind each image, so the text can be searched. A found word is highlighted in the image.

### **PDF with image substitutes:**

As for PDF (Normal), but words containing reject and suspect characters have image overlays, so these uncertain words display as they were in the original document. The PDF file can be viewed, searched and edited.

### **PDF, image only:**

The original images are exported. The PDF file is viewable only and cannot be modified in a PDF editor and text cannot be searched.

## Converting from PDF



OmniPage Professional 15 is supplied with a separate program from Nuance: the PDF Converter. This allows you to convert PDF files into Word, WordPerfect documents, RTF files or Excel spreadsheets quickly and easily. Once OmniPage is installed, PDF becomes available as a file type in the Microsoft Word File Open dialog box. In most cases the conversion can be done without invoking OmniPage.

## Sending pages by mail

You can send page images or recognized pages as one or more files attached to a mail message if you have installed a MAPI-compliant mail application, such as Microsoft Outlook.

### ▼ To send pages by e-mail:

- ◆ With automatic processing, select Send in Mail as the setting in the Export Results drop-down list on the OmniPage Toolbox. The Export Options dialog box appears as soon as the last available page in the document is recognized or proofed.
- ◆ With manual processing, select Send in Mail as the setting in the Export Results drop-down list and then click its button. The dialog box appears immediately.
- ◆ Workflows and jobs accept a Send in Mail export step. (Not applicable to OmniPage SE.)

## Other export targets



Turn recognized text into an audio wave file for later listening, using ScanSoft RealSpeak. (Not applicable to OmniPage SE.) A multiple converter is useful for this, allowing you to save the document to file and generate the wave file in one saving step. You must specify the reading language in the converter options for the wave file type.



In OmniPage Professional 15 you can export files to other targets. You can save files to a central server (an FTP site) or to Microsoft SharePoint and SharePoint 2003. Exporting choices are made in the Export Options dialog box. When you click OK you are directed to FTP or SharePoint log-in and invited to specify the required path.

If an ODMA-compliant Document Management System (DMS) is detected in your computing environment, it will be offered. If you have access to more than one DMS, the system default will apply. The ODMA server must be pre-configured to accept the file types to be exported from OmniPage Professional, as defined by their extensions.

See the Online help for more information on these targets.

# Workflows



Workflows, Workflow Assistant and Workflow Viewer are supplied only with OmniPage 15. Batch Manager is only available in OmniPage 15 and its advanced features are offered only in OmniPage Professional 15.

A workflow contains a series of processing steps and their settings. It can be saved for repeated use whenever you have a task needing the same processing. Workflows usually begin with a scanning or loading step, but they can also start from the document currently open in OmniPage. After that, they do not have to conform to the traditional 1-2-3 processing pattern. Usually a workflow will include a recognition step, but this is not compulsory. For instance, page images can be saved to image files in a different file type or to an OmniPage Document. With or without OCR, any number of saving steps are possible, even to different targets, each with their own export settings.

Workflows are designed for efficient whole-document processing. They cannot handle recognizing or saving single or selected pages from a document. You should use manual processing for such cases.

Some workflows run without user interaction. Workflows needing interaction are those with a manual image enhancement step, a manual zoning step, a proofing/editing step, or when run-time prompting is requested for input or output file names and paths.

Batch Manager jobs are closely related to workflows. Jobs are created in the Job Wizard which uses the Workflow Assistant in the creation process. Jobs run workflows according to the job parameters and it is more typical of them to run unattended.

## Sample workflows

Sample workflows are provided with OmniPage 15 to offer you typical work processes. They are available in the Workflow drop-down list. Choose one then click the Workflow Assistant button to see its steps and settings.

## Running workflows

Here is how to run a sample workflow or one you have created:

1. If your workflow takes input from scanner, place your document in its ADF or its first page on the scanner bed.



2. Select the desired workflow from the Workflow drop-down list.
3. Press the Start button. The OmniPage Toolbox displays the steps in the workflow and acts as a progress monitor. You do not have access to most program functions while the workflow is running. To stop the workflow before it completes, press the Stop button.
4. If run-time input selection is specified, the Load Images dialog box awaits your choice of files.
5. If you requested a step requiring interaction (image enhancement, manual zoning, or proofing) the program presents pages for attention.
6. When a page is enhanced, zoned or proofed, click the Page Ready button in the Toolbox to move to the next page.
7. When the last page is enhanced, zoned or proofed, or when you no longer want to do zoning or proofing, press the appropriate Document Ready button on the Toolbox. Any pages without zones will be auto-zoned.
8. The After Completion menu under Process / Workflows gives you three options to end a workflow. You can choose to close the document, close OmniPage, or shut down your computer. These settings are typically applied if the workflow runs unattended - if your workflow is so, remember including a saving step.

You can also run workflows from an OmniPage Agent icon on the Windows taskbar. Right-click it for a shortcut menu listing your

workflows. Select one to run it. OmniPage will be launched if necessary. If it is running with a document loaded, you will be invited to either start with an empty document, or start with the current one.

If you do not see the OmniPage Agent icon, enable it in the General panel of the Options dialog box or choose Start ▶ All Programs ▶ ScanSoft OmniPage 15.0 ▶ OmniPage Agent.



When the Agent is running in OmniPage Professional 15 with the ASR-1600 voice recognition system in operation, you can launch workflows by voice commands.

You can launch some workflows from your desktop. Right click on an image file icon or file name for a shortcut menu. Multiple file selection is possible. Choose OmniPage 15.0 and a workflow name from the sub-menu. This sub-menu also provides quick access to five target formats using default settings: Word, Excel, PDF, TXT and WordPerfect. Only workflows with run-time prompting for input files are listed here.

Pressing Stop while a workflow is running pauses it. Click Start to resume processing. If you pause a workflow, maybe do some manual processing, and then save the document as an OmniPage Document, when you later open that OmniPage Document, the interrupted workflow will use the OPD as input and finish the processing.

## **Workflow Assistant**

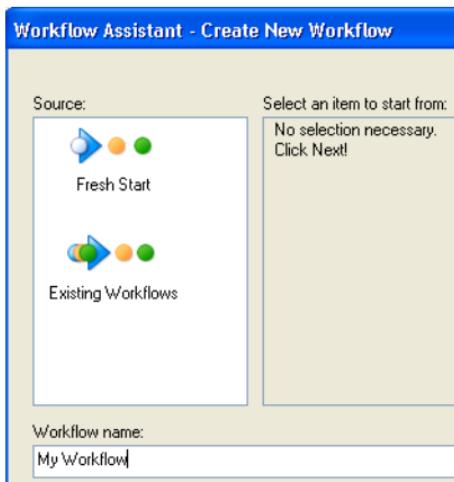
This allows you to create and modify workflows. The Job Wizard also uses this to create or modify workflows that jobs execute - see the next section. The Assistant offers a selection of steps, each represented by an icon. When you choose a step with settings, clicking Next brings up a dialog box allowing you to check and change them. You click Next again to receive a new set of step icons. At any moment in the process, the Assistant offers icons for all steps that are logically possible at that point.

## Creating workflows



Select New Workflow... in the Workflow drop-down list, or from the Process menu. Or click the Workflow Assistant button in the Standard toolbar when no workflow is selected.

The opening Assistant panel offers two starting points.

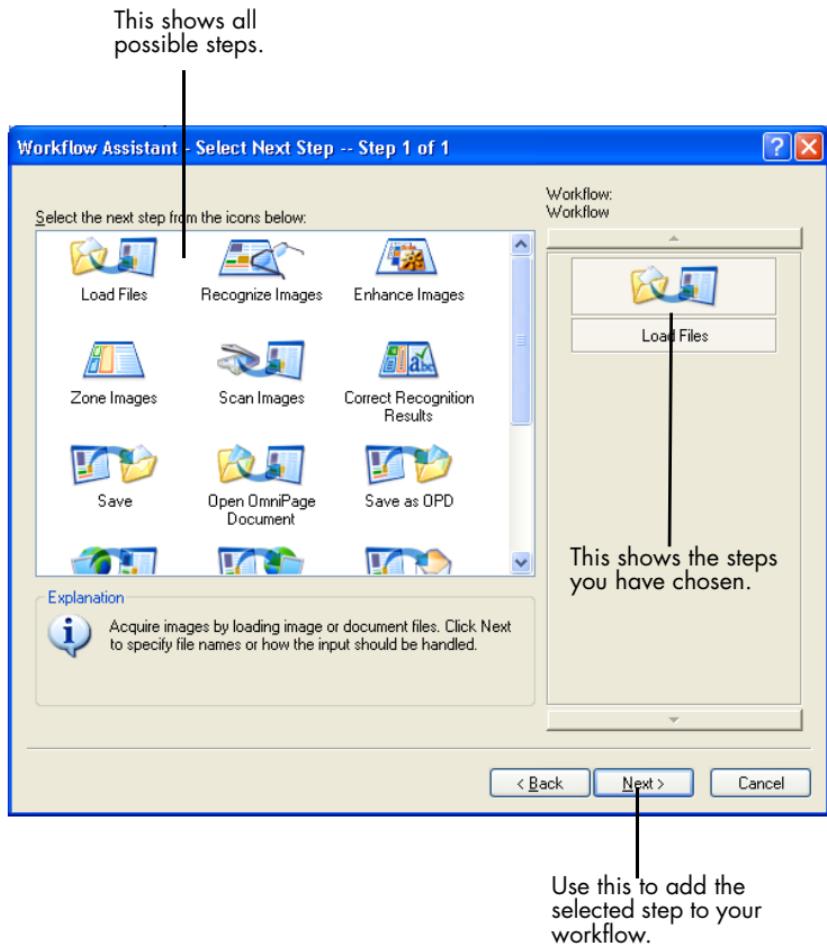


Choose **Fresh Start** to begin with no steps in the workflow diagram on the right. Name your workflow. Then click Next to choose your first step.

Choose **Existing Workflows** to see a list of existing workflows. These are the sample workflows plus any you have created. Select one as source. Its steps will appear in the workflow diagram on the right.

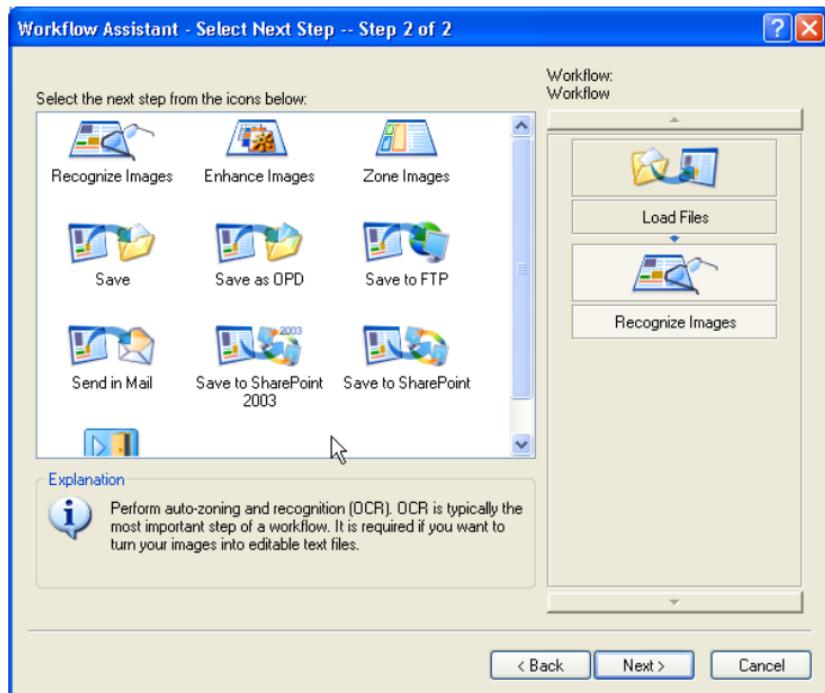
If you selected a workflow as source, you proceed by modifying its steps and settings. See the next section. If you save the workflow to a new name, the changed settings apply to the new workflow only and are not written back to the workflow used as the source. Similarly, when you make a new workflow with Fresh Start as source, its panels present settings as they were last set in OmniPage. Any changed settings enter the new workflow, but do not affect the settings in the program.

We now describe the creation of a workflow from a fresh start. Click Next to proceed to the panel where the input is defined:



After defining the input settings, define the path for your input files, automatic image preprocessing operations (e.g. rotate, deskew, despeckle) and PDF security settings. Then click Next to choose your second step.

The screen looks like this:



Use Next to move to further steps. Select steps and settings as requested, always using Next to confirm settings and proceed. Use Back to return to earlier steps and modify their settings. If you select a different step, all the steps following it will be deleted. To place multiple saving steps, always use Next. After each saving step is chosen and its settings specified, you still have a full choice of saving icons.

Finally, select Finish Workflow and click Next.

## Modifying workflows



Select the workflow you want to modify in the Workflow drop-down list and click the Workflow Assistant button in the standard toolbar. Or choose Workflows... in the Tools menu, select the desired workflow and click Modify... . The first panel of the Workflow Assistant appears with the workflow loaded. Click the icon in the workflow diagram that represents the step you want to modify. Click Next to move to the settings panel relating to that step. Make the desired changes.

When you add or replace a step, all following steps are deleted. To add a step, select the step in the workflow diagram below the place you want to insert the new step. Click the icon of the new step. To replace a step, select its icon in the workflow diagram and click the replacement step icon. Then recreate the subsequent steps and settings.

When all changes are made, move to the Finish Workflow step.

You can create a new workflow by modifying an existing one. Click New Workflow... in the Workflow drop-down list, choose the existing workflow as the starting point, and give it a different name as described in the previous section.

## Batch Manager



The Batch Manager is a separate but integrated program to let you create jobs to be processed immediately, or at some time in the future. By choosing steps carefully, you can set up jobs that can run unattended. A job executes a workflow according to the job settings. Jobs are created in the Job Wizard.



In OmniPage Professional 15 you have the following additional Batch Manager capabilities:

- ◆ Setting job timing and recurrence
- ◆ Folder watching for incoming image files
- ◆ E-mail inbox watching for incoming attachments (Outlook and Lotus Notes)
- ◆ E-mail notification of job completion to specified recipients
- ◆ Driving workflows with barcodes.

## Creating new jobs

Open the Batch Manager from the Process Menu or from your system, by choosing Start ▶ All Programs ▶ ScanSoft OmniPage 15.0 ▶ OmniPage Batch Manager or from the OmniPage Agent on the taskbar.



Creating a job is basically timing a workflow. To do this, start the Batch Manager (as described above) and click the Create Job icon or choose Create Job from the File menu.

The Job Wizard starts. First you need to define your job type. You can create five different types, instances of two basic categories: Normal and Watch type.

Normal and Watch type jobs may have a recurrence pattern. The latter are tailored to monitor a specified folder or e-mail inbox for incoming images to be processed in OmniPage. A specific type within this category is Barcode driven jobs, where barcode cover pages are used to identify which workflow to carry out.



**Normal job:** Set starting time and specify or create the Workflow to be run. If you select 'Do not start now' use the Activate button in the Batch Manager to start it.



Job types available in OmniPage Professional 15 only:



**Barcode driven job:** This job type can use either scanner or image file input and carries out workflows that are identified by the user-supplied barcode cover pages.



**Folder watching job:** Select this job type and browse to the folder(s) to be watched for incoming image files.



**Outlook mailbox watching job:** This job watches an Outlook e-mail inbox for incoming image attachments of a specified type.



**Lotus Notes mailbox watching job:** Same as above, but a Lotus Notes inbox is watched.

Name your job and click Next.

The next panel shows Start and Stop Options. Specify Start and End Time here, recurrency pattern (for recurrent jobs) and set if the input files are to be deleted when the job is completed. If you wish, you can set e-mail notification as well.

From the next panel onwards, you can construct your job as you normally do with Workflows. Set your starting point (Fresh Start or Existing Workflows) and proceed as described in "Creating workflows" on page 70.

The Options dialog box in the Batch Manager is in the Tools menu. Its General panel has an option *Enable OmniPage Agent on system tray at system startup*. By default it is on. It must remain selected for jobs to run at their scheduled time. The option is provided so it is possible to prevent all jobs from running without having to disable them individually. Its state also governs the running of barcode driven jobs.

The General panel lets you limit the number of pages allowed in an output document, even if the file option *Create one file for all pages* is selected. When the limit is reached, a new file is started, distinguished by a numerical suffix.

## Modifying jobs



Jobs with an inactive status can be modified. Select the job in the left panel of the Batch Manager and choose Modify from the Edit menu or click the Modify Job button. First, modify timing instructions as desired. Then the Workflow Assistant appears with the workflow steps and settings loaded. Make the desired changes as already described for workflows. See “Modifying workflows” on page 74.

## Managing and running jobs

This is done with the Batch Manager. It presents two panels. The left panel lists each job, its next run, status and history. The status will be:

**Waiting:** Scheduled but job start time is in the future.

**Running:** Processing is currently underway.

**Watching:** Watching is in progress but there is no processing.

**Inactive:** Created with timing instruction: *Do not start now*; or any deactivated jobs.

**Expired:** Scheduled job but start time is in the past.

**Collecting:** Watching in progress but the job is waiting for all incoming files to arrive.

**Paused:** User has paused the job and not yet resumed it.

**Closing:** Watch type job is saving its result.

**Starting:** The status right before Running. Displays when a job is just being started or when more jobs are about to run than the number of jobs Batch Manager can simultaneously run.

Click on a job and a step-by-step analysis of all pages in the job appears in the right panel. It shows where input was taken from, the page status and where output was directed to. Click on a plus icon to see more information about the page. Click on a minus icon to hide details. For jobs with the error or warning status, the listing shows which pages failed or what problems occurred.



**Activate Job** in the File menu serves to activate any inactive job immediately.



**Deactivate Job** in the File menu deactivates any active job. If the job is running, this will stop it before deactivating. Choose this to close a Watch type job immediately to save its result.



**Stop Job** in the File menu stops a job with status Starting, Running, or Paused.



**Pause Job** is available for jobs with status Running or Starting. To modify such a job's timing instructions you must stop it.



**Resume Job** lets the job continue from its state when it was paused.



**Delete Job** in the Edit menu serves to delete the currently selected job. Only Inactive jobs can be deleted.



**Rename Job** serves to modify the name of any job.

Use the Edit menu to send a copy of a job's status report to Clipboard. Use Save OPD As... in the File menu to save any intermediate result of a paused job to an OPD file.

To remove data files storing data of any previously run job, click Edit, then choose Clear Occurrence. Clear All Occurrences removes all data for all job occurrences. These two options are useful to free disk space, but cleared occurrences cannot be viewed any more, so use these with caution.

## The Workflow viewer

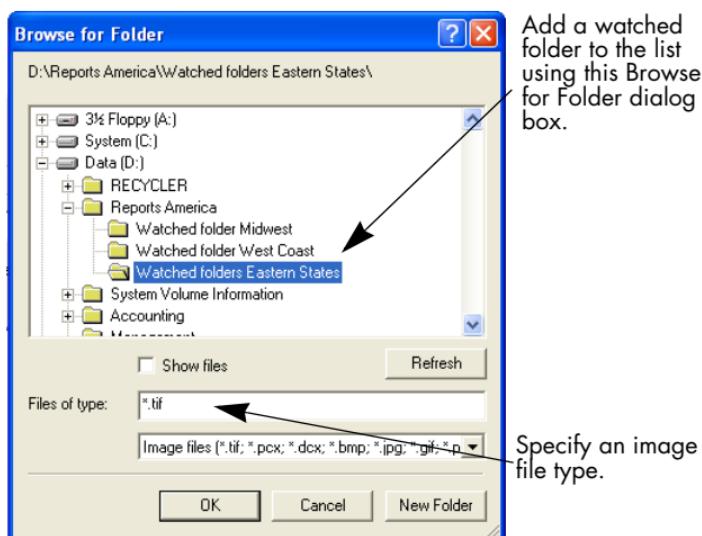
The Workflow viewer is integrated into the Batch Manager to the right of the list of your jobs. Use it to get comprehensive and detailed information about the processing of each occurrence of the job. The viewer shows the process in a step-by-step fashion - following the steps of the workflow. It displays input and output at each stage. Job results are marked by icons. Drop-down lists give you information about processing steps.

## Watched folders



In OmniPage Professional 15, you can specify watched folders and e-mail inboxes (Outlook and Lotus Notes) as job input. These allow processing to be started automatically whenever image files are placed in pre-defined folders or arrive into inboxes as e-mail attachments. This is useful to have sets of files with predictable content arriving from remote locations processed automatically on arrival, even if no-one is in attendance. Typically these are reports or form-like documents that are delivered repeatedly or at recurring intervals, for example each week or month.

To use this facility, prepare a set of folders or e-mail folders to be watched. You should not use these folders for other purposes, not even for barcode driven jobs. When setting up such a job, choose Folder watching job, name it and click Next. In the dialog box that appears, browse to the folders.



Add the desired folders and file types (one type or all types). Click the checkbox in front of your selected folder to include its subfolders as well. To enable a number of file types, add the Folder repeatedly, once for each type. Add a checkmark to watch subfolders of the selected folder as well.

When you reach the next panel of the Job Wizard, you set the timing instructions: a starting time and an end time for the watching to occur. You can specify recurrences, for instance to have the folder(s) watched only during your lunch hour (Start 12.15, End 13.05) every Monday, Wednesday and Friday, or overnight in the last three days of each month, when you keep your computer running to collect and process monthly reports arriving from afar.

When files enter a watched folder, the program waits approximately for the interval specified in Batch Manager Options for more files to arrive in order to process them together. When files cease to arrive, processing starts.

To finish the watching early, choose Deactivate Job. Then you can modify the job freely.

## Watched mailboxes



In OmniPage Professional 15, you can specify watched mailboxes as job input. These allow processing to be started automatically whenever image files of specified file types are placed in pre-defined e-mail folders. This is useful to have sets of files with predictable content arriving processed automatically on arrival, even if no-one is in attendance.

The program supports watching Microsoft Outlook and Lotus Notes mailboxes.

# Barcode processing



In OmniPage Professional 15, you can run workflows (series of steps and settings) using barcode cover pages that define which workflow should run. A barcode cover page identifies a workflow (with workflow identifier, workflow name and workflow steps) and contains information on workflow creation (name of the creator, date of creation, etc.). Note that barcode processing cannot be recurrent.

There are two ways of doing barcode processing:

**Scanner input:** Workflow processing is driven by placing the cover page on top of a document to be scanned and pushing the scanner's Start button.

**Image file input:** Job processing is driven by placing the barcode cover page image along with document images into a watched barcode folder.

For scanner input you have to

1. Create a workflow that contains the processing steps you need with Scan Images as first step.
2. Print a barcode page that identifies the workflow.
3. Start barcode processing from scanner.

To scan with a barcode page:

1. Place the barcode cover page on the top of the document in the ADF.
2. Press the Start button on the scanner.
3. If Prompt for workflow is selected in the Scanner panel, a dialog box appears with the available choices: Scanning, Barcode driven workflow, and all scanning workflows.

All available pages will be processed by the specified workflow, or until a new barcode page is encountered. The result will be saved.

For image input you must create a barcode driven job.

A barcode driven job uses a special kind of watched folder. Always use a separate folder for barcode processing. The starting time for the workflow is defined by the moment the barcode cover page enters a watched folder.

For a barcode driven job processing you need to

1. Create a workflow that contains the processing steps you need. Select Load Files as input with Prompt for input file names selected.
2. Save a barcode cover page that identifies the workflow, and
3. Define timing instructions for barcode folder watching in the Batch Manager by creating a barcode driven job.

To process with a barcode driven job:

1. Make sure that the job is running at the required time.
2. The folder is being monitored and the workflow will be started as soon as a barcode cover page is placed in the specified watched folder.
3. The workflow will process image files arriving in the folder after the cover page.
4. The workflow will be completed at the specified end time of the job, or each time a new barcode cover page is detected.

You can copy the barcode cover page image and the image files into the watched barcode folder yourself, or direct others to do this. You can also place just a barcode cover page image file in the watched folder, then have a network scanner make and send image files there.

## Voice recognition



ScanSoft ASR-1600<sup>TM</sup> voice recognition software is supplied with OmniPage Professional 15. By default this is installed during program installation for up to seven languages, depending on which interface languages are available in your version of the product. By performing a custom installation, voice modules can be added and removed. Please consult the ASR-1600 Help system for

guidance on microphone installation and handling. Go to the General panel of the Options dialog box and select *Enable voice assistant*. The voice recognition functions in the language set for the user interface in the General panel of the Options dialog box. You may need to change this temporarily if you have to use a different voice recognition language. Voice activation is applied to two fields.

### **Workflow activation**

To do this, the OmniPage Agent icon must be visible on the Windows taskbar. Use the Start menu or the General panel of the Options dialog box to place it there if it is not visible. When your ASR-1600 system is operational, just say **OmniPage workflows**. This will activate the shortcut menu. Speak the name of a workflow or say its number to have the workflow start, invoking OmniPage Professional 15 if necessary. The following commands are available:

OmniPage number <number>

OmniPage <workflow name>

OmniPage Batch Manager

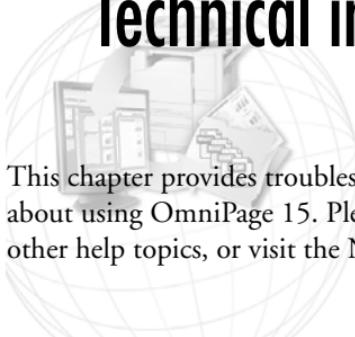
OmniPage Exit

The last command removes the OmniPage Agent icon from the taskbar. Use the Start menu or the General panel of the Options dialog box to restore it.

### **Proofing**

When you have OmniPage Professional 15 operational with the OCR Proofreader dialog box open, the voice recognition can be used to accept a suggestion from the list offered by the proofreader. The suggestions are numbered. Say 'Number', then the suggestion number in the appropriate language. For instance, say 'Number five'. Then suggestion five will be placed in the text and you will move to the next suspect word. You can also speak the word you want accepted. The buttons in the OCR Proofreader dialog box can also be voice activated.

# Technical information



This chapter provides troubleshooting and other technical information about using OmniPage 15. Please also read the online Readme file and other help topics, or visit the Nuance web pages.

## Troubleshooting

Although OmniPage is designed to be easy to use, problems sometimes occur. Many of the error messages contain self-explanatory descriptions of what to do – check connections, close other applications to free up memory, and so on.

Please see your Windows documentation for information on optimizing your system and application performance.

On supported file formats, see the Online Help.

## Solutions to try first

Try these solutions if you experience problems starting or using OmniPage:

- ◆ Make sure that your system meets all the listed requirements. See “System requirements” on page 9.
- ◆ Make sure that your scanner is plugged in and that all cable connections are secure.
- ◆ Visit the support section of Nuance’s web site at [www.nuance.com](http://www.nuance.com). It contains Tech Notes on commonly reported issues using OmniPage. Our web pages may also offer assistance on the installation process and troubleshooting.

- ◆ Use the software that came with your scanner to verify that the scanner works properly before using it with OmniPage.
- ◆ Make sure you have the correct drivers for your scanner, printer, and video card. Visit Nuance's web page through the Help menu and consult its scanner section for more information.
- ◆ Defragment your hard disk. See Windows online Help for more information.
- ◆ Uninstall and reinstall OmniPage, as described in the section, "Uninstalling the software" on page 15.

## **Testing OmniPage**

Restarting Windows 98, Me, 2000, XP or 2003 Server in its safe mode allows you to test OmniPage on a simplified system. This is recommended when you cannot resolve crashing problems or if OmniPage has stopped running altogether. See Windows online Help for more information.

### **▼ To test OmniPage in safe mode:**

1. Restart your computer in safe mode by pressing F8 immediately after you see the 'Starting Windows' message.
2. Launch OmniPage and try performing OCR on an image. Use a known image file, for instance one of the supplied sample image files.
  - If OmniPage does not launch or run properly in safe mode, then there may be a problem with the installation. Uninstall and reinstall OmniPage, and then run it in Windows safe mode.
  - If OmniPage runs in safe mode, then a device driver on your system may be interfering with OmniPage operation. Troubleshoot the problem by restarting Windows in Step-by-Step Confirmation mode. See Windows online Help for more information.

## **Increasing memory resources**

OmniPage may run poorly under low-memory conditions. This may be indicated by various error messages or if OmniPage works slowly and accesses the hard drive often. Try these solutions for low memory conditions:

- ◆ Restart your computer.
- ◆ Close other open applications to release memory.
- ◆ Close unnecessary OmniPage applications.
- ◆ Defragment your hard disk to free up contiguous blocks of disk space. See Windows online Help for instructions.
- ◆ Increase the amount of free hard disk space.
- ◆ Increase your computer's physical memory (RAM). More memory optimizes OCR performance. See "System requirements" on page 9.

## **Increasing disk space**

Problems may occur if your system runs low on free disk space. Try these solutions for low disk space problems:

- ◆ Empty the Windows Recycle Bin.
- ◆ Close all open applications and delete the \*.tmp files in the Temp folder. This folder is usually located in your Windows folder.
- ◆ Back up unneeded files and delete them from your hard disk.
- ◆ Remove Windows applications that you do not use.
- ◆ Defragment your hard disk.
- ◆ Clear the cache for your web browser and limit its size.

## Text does not get recognized properly

Try these solutions if any part of the original document is not converted to text properly during OCR:

- ◆ Look at the original page image and ensure that all text areas are enclosed by text zones. If an area is not enclosed by a zone, it is generally ignored during OCR. See the section on creating and modifying zones, “Working with zones” on page 40.
- ◆ Make sure text zones are identified correctly. Reidentify zone types and contents, if necessary, and perform OCR on the document again. See “Zone types and properties” on page 39.
- ◆ Be sure you do not have an unsuitable template loaded by mistake. If zone borders cut through text, recognition is impaired.
- ◆ Adjust the brightness and contrast sliders in the Scanner panel of the Options dialog box. You may need to experiment with different settings combinations to get the desired results.
- ◆ Use the Image Enhancement Tools to optimize your image for OCR.
- ◆ Check the resolution of the original image. Hover the cursor over a page thumbnail for a popup display. If the resolution is significantly above or below 300 dpi, recognition is likely to suffer.
- ◆ Make sure the correct document languages are selected in the OCR panel of the Options dialog box. Only languages included in the document should be selected.
- ◆ Turn IntelliTrain on and make some proofing corrections. This is most likely to help with stylized fonts or uniformly degraded documents. If IntelliTrain was running, try turning it off – on some types of degraded documents it may not be able to help.
- ◆ Do some manual training, or edit existing training to remove unsuccessful training.



The above advice on training does not apply to OmniPage SE.

- ◆ If you use True Page as the Text Editor view or for export, recognized text is put into text boxes or frames. Some text may be hidden if a text box is too small. To view the text, place the cursor in the text box and use the arrow keys on your keyboard to scroll to the top, bottom, left, or right of the box.
- ◆ Check the glass, mirrors, and lenses on your scanner for dust, smudges or scratches. Clean if necessary.

## **Problems with fax recognition**

Try these solutions to improve OCR accuracy on fax images:

- ◆ Ask senders to use clean, original documents if possible.
- ◆ Ask senders to select Fine or Best mode when they send you a fax. This produces a resolution of 200 x 200 dpi.
- ◆ Ask senders to transmit files directly to your computer via fax modem if you both have one. You can save fax images as image files and then load them into OmniPage. See “Input from image files” on page 29.

## **System or performance problems during OCR**

Try these solutions if a crash occurs during OCR or if processing takes a very long time:

- ◆ Resolve low memory and low disk space problems. See “Testing OmniPage” on page 84.
- ◆ Check image quality. Consult your scanner documentation on ways to improve the quality of scanned images.
- ◆ Break complex page images (lots of text and graphics or elaborate formatting) into smaller jobs. Draw zones manually or modify automatically created zones and perform OCR on one page area at a time. See “Working with zones” on page 40.
- ◆ Restart Windows 98, Me, 2000 or XP in safe mode and test OmniPage by performing OCR on the included sample image files.

If you are performing multiple tasks at once, such as recognizing and printing, OCR may take longer.

## Supported file types

Supported image file formats for loading are TIFF, PCX, DCX, BMP, JPEG, JB2, JP2, GIF, PNG, XIFF, MAX, PDF\*.

Supported file types for saving recognition results as text are:

HTML 3.2, 4.0\*

Microsoft Excel 97, 2000, XP, 2003

Microsoft PowerPoint 97

Microsoft Publisher 98

Microsoft Word 97, 2000, XP, 2003 (WordML)\*

OmniPage Documents

PDF\* (Normal), Edited, with image on text, with image substitutes

RTF Word 6.0/95, RTF Word 97, RTF Word 2000, RTF 2000

ExactWord\*

WordPad

WordPerfect 8, 9, 10

Text, Text with line breaks, Text - Formatted, Text - Comma Separated

Unicode Text, Unicode Text with line breaks, Unicode Text - Formatted, Unicode Text - Comma Separated

Wave Audio Converter\* (to save recognized text being read aloud)



File types with an asterisk are only supported in OmniPage 15.

In OmniPage Professional 15 there is also support for: eBook, Microsoft InfoPath (for forms), Microsoft Reader, and XML.

# Index

## A

Accuracy  
improvement 30, 50, 86  
influence of brightness 30  
influence of training 50  
scanning influence 30  
Acquire Text menu items 27  
Activating OmniPage 15  
Activation of workflows by  
voice 81  
Adding  
to zones 41  
training to training files 51  
words to user dictionary 46  
ADF 29, 31  
Advanced saving options 62  
Advice on problems 83  
Alphanumeric zones 39  
ASR-1600 81  
Attachments to mail 65  
Auto-detect layout 32  
Automatic Document Feeder  
(ADF) 29, 31  
Automatic training 51  
Auto-sending by mail 65  
Auto-zoning 32, 40

## B

Backgrounds for zoning 38  
Basic processing steps 18  
Batch Manager 28, 73  
Black-and-white  
images 60  
scanning 30  
Bold text 52  
Boxes 53  
Boxes for recognized text 87

Brightness 30, 86  
Brightness / Contrast (E) 36  
Bring to Front tool (F) 58

## C

Changing  
part of a page 54  
reading order 54  
Character attributes 52  
Character Map 48  
Characters, suspect 45  
Checkbox tool (F) 57  
Checking OCR results 47  
Circle text tool (F) 57  
Clipboard 65  
Color  
images 60  
markers 47  
scanning 30  
Comb tool (F) 57  
Comparing recognized words  
with originals 47  
Composition of workflows 67  
Contrast 30, 86  
Converters multiple 63  
Converting from PDF 65  
Converting image files 69  
Copying pages to Clipboard 65  
Creating  
Batch Manager jobs 29  
training data 51  
workflows 70  
Crop (E) 36  
Custom Layout 33  
Customizing export converters  
62

## D

Deleting  
jobs 77  
training files 51  
user dictionaries 49  
Describing document layout 32  
Deskew (E) 36  
Desktop 18  
Desktop launching of work-  
flows 69  
Despeckle (E) 36  
Dictionaries 46  
Direct OCR 26  
Disabling job running 75  
Disk space 9, 85  
Document Layout, Form 33  
Document Manager 18  
Document Ready button 68  
Document to document con-  
version 31  
Documents  
copying to Clipboard 65  
double-sided 32  
exporting 59  
in OmniPage 17  
layout description 32  
saving 59  
with varied layout 32  
Double-sided documents 31  
Drawing zones in Direct OCR  
28  
Dropout color (E) 36  
Dropping graphics from export  
61  
Duplex scanners 31  
Dynamic verifier 47

## **E**

Editing  
character attributes 52  
form objects 58  
graphics 53  
in True Page 53  
on-the-fly 54  
paragraph attributes 52  
PDF output 64  
recognized text 52  
tables 42, 53  
training files 51  
user dictionaries 49  
E-mail notification 72  
Embedding templates in OPD  
files 43

Enabling OmniPage taskbar  
icon 69  
Error messages for jobs 76, 77  
Export converters 62  
Export Results button 61  
Exporting  
graphics 61  
in Flowing Page 62  
in True Page 62  
repeated 59  
to Clipboard 65  
to file 61  
to mail 65  
to PDF 64

## **F**

Fax recognition 87  
Features, new 7  
Files  
as export target 60  
as image source 29  
retained on uninstall 15  
separation options 61  
types for export 61  
Fill (E) 37

Fill text tool (F) 57

## Finding

non-dictionary words 46  
suspect words 46

## Finishing

proofing in a workflow 68  
workflows 72  
zoning in a workflow 68

## Flowing Page 62

Form Arrangement Toolbar 58  
Form Drawing Toolbar 57  
Form zone 40  
Formatted Text view 46, 62  
Formatting levels 28, 45, 61  
Formatting toolbar 18  
Frames 53, 62, 87

## **G**

Graphic tool (F) 57

Graphic zone 40

## Graphics

editing 53  
in export 61

## Grayscale

images 60  
scanning 30

Grouping elements 53

## **H**

Header/footer indicators 45

Hearing texts read aloud 56

Hiding / showing markers 45

Horizontal Alignment tools (F)  
58

Hue / Saturation (E) 36

Hyperlinks 53

## **I**

Ignore backgrounds 38

Ignore zones 40

Image Enhancement

history 37

in workflows 38

tools 35

## Image files

conversion 69

input 29

reading order 29

samples 84

## Image Panel 18

## Image toolbar 18

## Images

backgrounds 38

black-and-white 60

color 60

editing 53

grayscale 60

quality 31

resolution 60, 86

saving 60

substitutes in PDF 64

## Improving accuracy 30, 51, 86

## Increasing disk space 85

## Increasing memory 85

## Input

for workflows 71

from image file 29

from PDF files 29

from scanner 30

## Installing

OmniPage 10

scanners 11

## IntelliTrain 51, 86

## Italic text 52

## **J**

## Jobs

disabling 75

error messages 76, 77

in Batch Manager 28

managing 76, 77

modifying 76  
page limit 75  
recurring 79  
running 76, 77  
status 76, 77  
timing instructions 79  
Joining zones 41

## L

Languages 50  
for recognition 86  
Launch  
target application 61  
workflows from desktop 69  
Layout description 32  
Layout retention 46  
Layout, auto-detect 32  
Legal dictionaries 46  
Line tool (F) 57  
Links to web pages 53  
Loading  
a user dictionary 49  
image files 29  
training files 51  
zone templates 33, 42

## M

Mail 65  
Mailbox watching 79  
Managing jobs 76, 77  
Manual training 50  
Manual zoning 38  
Marked words in Editor 45  
Markers 45, 47  
Medical dictionaries 46  
Memory requirements 9, 85  
Microsoft Word, opening PDF  
files in 65  
Minimum system requirements  
9

Modifying  
image quality 33  
jobs 76  
tables 42, 53  
workflows 73  
zone templates 43  
zones 40  
MS Outlook 65  
Multicolumn areas 53  
Multi-page image files 60  
Multiple column pages 32  
Multiple converters 63

## N

New features 7  
Non-dictionary words 45  
Non-printing characters 45  
Numeric zones 39

## O

OCR  
Batch Manager 28, 73  
checking OCR results 47  
Direct OCR 26  
poor performance in 87  
proofreading results 46  
settings for Direct OCR 26  
Workflows 25  
OCR Brightness (E) 36  
OCR image 34  
OmniPage  
activating 15  
documents in 17  
earlier versions 10  
installing 10  
new features 7  
registering 14  
reinstalling 15  
starting 11  
testing 84

uninstalling 15  
OmniPage Desktop 18  
OmniPage Documents 17  
extended 17  
saving as 59  
OmniPage Professional 7  
OmniPage Toolbox 18  
OmniPage Agent 69  
Online registration 14  
On-the-fly editing 54  
OPD Files 17  
extended 17  
template embedding 43  
Opening image files 29  
Optimizing brightness 30  
Options dialog box 20  
Options for proofing 46  
Options for saving 62  
Order of page elements 54  
Original image saving 60  
Overview  
of processing 21  
of processing steps 18

## P

Page limit for jobs 75  
Page Ready button 68  
Pages  
copying to Clipboard 65  
multi-page image files 60  
navigation 18  
sending as mail 65  
Paragraph  
editing attributes 52  
styles 52, 61  
Pausing workflows 69  
PDF Converter program 65  
PDF Edited 64  
PDF file input 29  
Pending pages 54  
Performance problems during

OCR 87  
 Plain Text view 62  
 Pointer (E) 35  
 Preprocessing images 33  
 Primary image 34  
 Primary/OCR Image (E) 35  
 Problems with faxes 87  
 Process backgrounds 38  
 Process zones 40  
 Processing  
     basic steps of 18  
     from other applications 26  
     manual 25  
     step-by-step 25  
     steps, overview 18  
     with Batch Manager 28  
     with workflows 25, 68  
 Professional dictionaries 46  
 Professional version 7  
 Proofing  
     by voice 81  
     in a workflow 68  
     options 46  
 Properties of zones 39  
 Purpose of training 50  
 Purpose of workflows 67

**Q**  
 Quality of images 31  
 Quick Start Guide 21

**R**  
 Reading order 54  
 Reading text aloud 55  
 RealSpeak 55  
 Recognition  
     accuracy 31, 50, 86  
     in workflows 72  
     languages 50, 86  
     problems with faxes 87

saving results 61  
 speeding up 87  
 Rectangle tool (F) 57  
 Recurring jobs 79  
 Registering  
     applications for Direct OCR 27  
     OmniPage 14  
 Reinstalling OmniPage 15  
 Removing zone templates 43  
 Repeated exporting 59  
 Replacing zone templates 43  
 Resolution 60, 86  
 Resolution (E) 36  
 Retaining paragraph styles 61  
 Re-training 50  
 Rotate (E) 36  
 Running  
     Batch Manager jobs 75  
     workflows 68  
     workflows by voice 81

**S**  
 Safe mode 84  
 Sample image files 21, 84  
 Saving  
     and launching 61  
     as OmniPage Document 59  
     documents 59  
     options 62  
     original images 60  
     recognition results 61  
     steps in workflows 72  
     text 61  
     to file 60  
     to multiple file types 63  
     training files 51  
     user dictionaries 49  
     zone templates 43  
 Saving and applying Image Enhancement templates

37  
 Scanners 30, 87  
 drivers 12  
 duplex 31  
 setting up 11  
 Scanning 30  
 input from 30  
 pictures 30  
 Wizard 11  
 Scheduled processing 73  
 Searching PDF output 64  
 Select Area (E) 35  
 Selection tool (F) 57  
 Send to Back tool (F) 58  
 Sending pages by mail 65  
 Setting up a scanner 11  
 Setting up Direct OCR 27  
 Settings  
     Acquire Text 27  
     for Direct OCR 27  
     Options dialog box 20  
     zone types 42  
 Single-column pages with tables 32  
 Slow recognition 87  
 Smart folders 78, 79  
 Solutions for poor performance 83  
 Speech recognition 81  
 Spreadsheet pages 32  
 Standard toolbar 18  
 Starting a user dictionary 49  
 Starting Batch Manager 74  
 Starting the program 11  
 Status of jobs 76, 77  
 Step-by-step processing 18  
 Stopping workflows 69  
 Storing zoning changes 54  
 Suggestions in proofing 46  
 Suspect words 45  
 Synchronize Views (E) 36  
 System or performance prob-

lems during OCR 87  
System requirements 9

## T

Table tool (F) 57  
Tables  
  editing 53  
  editing dividers 42  
  in single column pages 32  
  in Text Editor 53  
removing dividers 42  
rows in 42  
zones 40, 42  
Taskbar workflow icon 68  
Technical information 83  
Template zones 33, 42, 86  
Templates in OPDs 43  
Testing OmniPage 84  
Text Editor 18, 45  
Text saving 61  
Text tool (F) 57  
Text-to-Speech facility 56  
Thumbnails 18  
Timing of jobs 79  
Toolbar docking / floating 47  
Training 50  
  automatic 51  
  IntelliTrain 51  
  manual 50  
  training files 51  
Troubleshooting 83  
True Page editing 53  
True Page export 62  
True Page view 46  
TWAIN scanner drivers 12  
Types of zones 39

## U

Underlined text 52  
Ungrouping elements 53

Uninstalling the software 15  
Unloading  
  training files 51  
  user dictionaries 49  
  zone templates 43  
URLs 53  
User dictionaries 46, 49  
Using Direct OCR 27

## V

Verifying text 47  
Vertical Arrangement Tools (F)  
  58  
Views  
  Formatted text 45  
  Plain Text 45  
  True Page 46  
Voice recognition 81

## W

Watched folders 78, 79  
Watched mailboxes 79  
Web page links 53  
Wizard for scanner setup 11  
Workflow Assistant 24, 69  
Workflows  
  composition 67  
  creating 70  
  finishing 72  
  input 71  
  modifying 73  
  pausing and stopping 69  
  recognition 72  
  running 68  
  saving steps 72  
  taskbar icon 68  
  user interaction 68  
Working with zones 40

## Z

Zones  
  adding to 41  
  alphanumeric 39  
  changing types 40  
  deleting templates 42  
  graphic 40  
  ignore 40  
  in Direct OCR 28  
  irregular 41  
  joining 41  
  manual 38, 86, 87  
  modifying templates 42  
  numeric 39  
  process 40  
  properties 39  
  replacing templates 43  
  saving templates 43  
  table 40, 42  
  templates 33, 42, 86  
  types 39, 86  
  unloading templates 44  
  working with 40  
Zoning in a workflow 68  
Zoning on-the-fly 54  
Zoom (E) 35  
Zooming displays 18, 47

(E) = Image Enhancement tool  
(F) = Form Drawing or  
  Arrangement tool  
  (Professional only)

## THIRD PARTY LICENSES/NOTICES

The Independent JPEG Group's software, copyright © 1991-1995, Thomas G. Lane.

This software is based, in part, on the work of the Independent JPEG Group, Colosseum Builders, Inc., the FreeType Team, and Catharon Productions, Inc.

Zlib copyright © 1995-1998 Jean-loup Gailly and Mark Adler.

This product was developed using Kakadu software.

The word verification, spelling and hyphenation portions of this product are based in part on Proximity Linguistic Technology.

The Proximity Hyphenation System © Copyright 1988. All Rights Reserved. Proximity Technology Inc.

The Proximity/Merriam-Webster American English Linguibases. © Copyright 1982, 1983, 1987, 1988 Merriam-Webster Inc. © Copyright 1982, 1983, 1987, 1988 Proximity Technology Inc. Words are checked against the 116,000, 80,821, 92,641, 106713, 118,533, 91928, 103,792, 130,690, and 140,713 word Proximity/Merriam-Webster Linguibases. The Proximity/Collins British English Linguibases. © Copyright 1985 William Collins Sons & Co. Ltd. Legal and Medical Supplements © Copyright 1982 Merriam-Webster Inc. © Copyright 1982, 1985 Proximity Technology, Inc. Words are checked against the 80,307, 90,406, 105,785, and 115,784 word Proximity/Collins Linguibases. The Proximity/Collins French, German, Italian, Portuguese (Brazilian), Portuguese (Continental), Spanish Linguibases. © Copyright 1984, 1985, 1986, 1988 William Collins Sons & Co. Ltd. © Copyright 1984, 1985, 1986, 1988 Proximity Technology, Inc. Words are checked against the 136,771, 150,893, 178,839, 207,119, 212,565, and 194,393 word Proximity/Collins Linguibases. The Proximity/Van Dale Dutch Lingubase. © Copyright 1987 Van Dale Lexicografie bv. © Copyright 1987 Proximity Technology, Inc. Words are checked against the 119,614 word Proximity/Van Dale Lingubase. The Proximity/Munksgaard Danish Lingubase. © Copyright 1988 Munksgaard International Publishers Ltd. © Copyright 1988 Proximity Technology Inc. Words are checked against the 113,000 word Proximity/Munksgaard Lingubase. The Proximity/IDE Norwegian and Swedish Lingubases. © Copyright 1988 IDE a.s. © Copyright 1988 Proximity Technology Inc. Words are checked against the 126,123 and 150,000 word Proximity/IDE Lingubases.

© Nuance Communications, Inc., 2006. All rights reserved. Subject to change without prior notice.