Kofax OmniPage Ultimate

User's Guide

Version: 19.2

Date: 2019-10-11

KOFAX

© 2010 - 2019 Kofax. All rights reserved.

Kofax is a trademark of Kofax, Inc., registered in the U.S. and/or other countries. All other trademarks are the property of their respective owners. No part of this publication may be reproduced, stored, or transmitted in any form without the prior written permission of Kofax.

Table of Contents

Chapter 1: Welcome....6
Key features....7
Chapter 2: Installation and setup....9
System requirements....9
Install OmniPage....10
Install Kofax Power PDF Create....10
Register OmniPage....11
Activate OmniPage....11
Set up your scanner with OmniPage....11
Start the program....13
Uninstall the software....15
Chapter 3: Using OmniPage....16
OmniPage documents....16
OmniPage desktop and views....16
Program panels....17
Classic View....17
Flexible Views....19
Quick Convert View....21
Toolbars....22
Basic processing steps....23
Settings....23
Use OmniPage with Kofax PaperPort....23
Chapter 4: Processing documents....25
Processing methods....25
Processing from other applications....26
How to set up Direct OCR....27
How to use Direct OCR....27
Defining the source of page images....27
Input from image files....28
Input from the Cloud....28
Input from digital camera....28
Input via Easy Loader....28
Input from scanner....29
Brightness and contrast....30

Scan with an ADF.... 30
Scan without an ADF.... 30
Scan to OmniPage and workflows.... 30
Describe the layout of the document.... 31
Preprocess images.... 31
Image Enhancement tools.... 33
Using Image Enhancement history.... 37
Save and apply templates.... 37
Image Enhancement in workflows.... 37
Zones and backgrounds.... 38
Automatic zoning.... 38
Zone types and properties.... 39
Working with zones.... 40
Table grids in the image.... 41
Using zone templates.... 41
Chapter 5: Proofing and editing.... 43
The Text Editor display and formatting levels.... 43
Proofread OCR results.... 44
Verify text.... 44
Character Map.... 45
User dictionaries.... 45
Languages.... 46
Asian language recognition.... 46
Training.... 47
Text and image editing.... 48
On-the-fly editing.... 49
Marking and redacting.... 50
Reading text aloud.... 50
Creating and editing forms.... 52
Form Drawing toolbar.... 52
Form Arrangement toolbar.... 53
Editing Form object properties.... 53
Form data extraction.... 53
Chapter 6: Saving and exporting.... 55
Saving and exporting.... 55
Save original images.... 55
Save recognition results.... 56
Selecting a formatting level.... 56

Select converter options....57
Multiple converters....57
Save to PDF....57
Convert from PDF....58
eDiscovery Assistant for searchable PDF....59
Creating PDF files from other applications....59
Send pages by mail....59
Send to eBook readers....59
Send to Kindle....60
Send to ePub....60
Other export targets....61
Chapter 7: Workflows....62
Running workflows....62
Workflow Assistant....63
Create workflows....64
Modify workflows....65
Workflow to Kindle....65
DocuDirect....65
Create new jobs....66
Modify jobs....67
Managing and running jobs....67
The Workflow viewer....68
Watched folders....69
Watched mailboxes....69
Barcode processing....70
File-it Assistant....71
Single-step PDF jobs....72
Chapter 8: Technical information....73
Troubleshooting....73
Solutions to try first....73
Testing OmniPage....73
Text does not get recognized properly....74
Problems with fax recognition....75
System or performance problems during OCR....75
Supported file types....75
Third-party licenses and notices....76

Chapter 1

Welcome

Welcome to the Kofax OmniPage Ultimate 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.

User's Guide

This guide introduces you to using Kofax OmniPage Ultimate. 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.

Note

Throughout the documentation, the Kofax OmniPage Ultimate icon is used to denote features that are tailored to support intensive use in office environments.

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. Throughout the document. references to newer versions of Microsoft Office output file types (2007 and above) are written with the year numbers omitted.

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 used with Kofax OmniPage Ultimate. Please refer to the documentation from the scanner manufacturer as necessary.

How-to guides

The How-to guides are available from the Help menu. They are a series of mini-guides that help you get started easily by providing concise overviews of key program areas, such as getting input, image improvement, zoning, recognition, editing, proofreading, new features, and more.

Help

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

Comprehensive context-sensitive help aims to provide just enough assistance to let you keep working without delay. It is available from dialog boxes. Press F1 in any dialog box to access help, or click the Help button, if applicable.

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.

Other information

The Kofax website at www.kofax.com provides timely information on the program. Access the Kofax website from the Kofax OmniPage Ultimate installer or afterwards from the Help menu.

Key features

Customize Windows Explorer shortcut menus
The OmniPage items in the Windows Explorer shortcut menus for input files support direct conversion to standard file formats, and the addition of user-defined workflows to the menu; the Convert Now Wizard makes it easy to customize the conversion process.

Multiple document handling
Use multiple document handling to work on more than one document at a time. You can display page thumbnails from another open document, and then copy or move pages among documents.

Cloud support
Download input files from Evernote and return recognition results there.

Scanner enhancement (SET) tools
Recent innovation includes more control over despeckling, better margin cleaning, and a control for when whiteboard content is captured by a digital camera; the text and diagrams can be enhanced for maximum readability.

Asian recognition
OCR services are provided for Japanese, Korean, Simplified Chinese and Traditional Chinese, with support for both horizontal and vertical text flow and embedded English texts. You can view and verify results in the Text Editor.

Automatic Language Detection
Allow the program to assign a single language to each incoming page during unattended processing. The program chooses from the languages with dictionary support based on a Latin-based alphabet. When this feature is enabled, no manual language selection is possible.

Easy Loader
Easy Loader provides a display similar to Windows File Explorer to view the file system in one of the OmniPage windows, to keep files visible during your work and deliver full Explorer functionality, yielding quick file selections; a dialog box with a lock facility lets a file set be built up before loading starts. With Quick Convert View, Easy Loader allows not only fast file loading, but also one-click total processing: Load > Recognize > Save.

Linking workflows to scanner buttons
You can associateOmniPage functions and workflows with scanner buttons, so that pre-processing, recognition and storage of documents are managed entirely from the scanner.

Launchpad
This program lets you quickly set up and run recognition tasks. Make just three basic choices to create a Go-flow, and the fine-tune key settings as desired. The Go-flow runs with minimal need for your intervention. This approach is ideal for quick completion of similar recurring tasks without using the full Kofax OmniPage Ultimate or the Workflow Assistant.

DocuDirect
This is a powerful workflow management tool that was previously known as the Batch Manager. Technical improvements make large-scale processing more robust, with better reporting and separation of problematic documents and improved recovery from critical situations. Differing default settings for Workflow Assistant and DocuDirect are introduced to better match their purposes.

Make PDF file searchable
A new workflow step is available in DocuDirect. Input is a set of PDF files of any type. This step discovers image-only parts or pages, runs OCR and adds the text to the PDF, leaving notes and annotations intact. This functionality exists in the eDiscovery Assistant for searchable PDF, and it is available for a job, allowing for timed, recurring and unattended job running, plus the use of watched folders.

Electronic book support
The widely used ePub file type is now supported with three output converter choices. Export your scanned documents or image files to your favorite portable devices. This augments the existing ability to save texts to the Amazon Kindle book reader.

Automated handling of digital camera images
OmniPage detects whether an image came from a digital camera by reading the EXIF data that digital cameras generate. The auto-deskew option can be turned on or off. When on, the program applies 2D or 3D deskewing (for normal or camera images). Resolution enhancement and the straightening of text lines are applied to images coming from a digital camera.

Features in Kofax OmniPage Ultimate only

Enterprise Content Management (ECM)
Enterprise Content Management (ECM) links are available to eDOCS (Open Text) and iManage (Interwoven). When using SharePoint, the server, login and password information must be provided only once per session, and it is offered in each subsequent session.

Extracting data from filled forms
A workflow step allows data to be extracted from sets of forms and exported to databases, based on a PDF form template. The forms to be processed can be active PDF forms, static forms in a range on image formats or scanned paper forms.

File-it Assistant
Use this feature as an efficient aid for creating and using barcode cover page workflows, which allow for automatic processing and storage of documents driven by pushing just one scanner button.

Marking and redacting
Text can be highlighted, stricken out, or redacted (made unreadable) in the Text Editor. This can be done by selection or by searching for specified words. Redacting is useful for legal or confidential documents.

Chapter 2

Installation and setup

This chapter provides information on installing and starting OmniPage.

System requirements

You need the following minimum system requirements to install and run Kofax OmniPage Ultimate:

- A computer with Intel Pentium® 4 or higher processor or its equivalent. Dual-core or Quad-core support recommended.
- Supported operating systems:
  - Windows 10 (32-bit and 64-bit editions)
  - Windows 7 (32-bit or 64-bit Editions) with Service Pack 1
  - Windows 8.1 (32-bit and 64-bit Editions)
- 1 GB of memory (RAM), 2 GB recommended for advanced performance.
- 1.0 GB of free hard disk space, including Power PDF Create, plus 50 MB working space during installation.
- 1024 x 768 pixel color display with 16-bit color or greater video card.
- A sound card and speaker for reading text aloud.
- Web access suitable for download the installer.
- Microsoft Internet Explorer 10 or above.
- Using OmniPage features in Microsoft Office applications requires a 32-bit or 64-bit version of Microsoft Office 2010 (SP2), 2013 (SP1), 2016 or 2019.
- Using OmniPage features in WordPerfect requires Corel WordPerfect X5, X6, X7, X8 or X9.
- Kofax OmniPage Ultimate supports Microsoft SharePoint Server 2010, 2013, 2016, and 2019.
- A compatible scanner with its own scanner driver software for scanning documents (WIA, TWAIN or ISIS scanner driver).
- 2-megapixel digital camera or higher, with auto-focus, for digital camera text capture.
- Web access needed for online Activation, Registration, Live Update, and Scanner Wizard database updating.
- East Asian language handling must be installed in the operating system to view Japanese, Chinese or Korean documents. (Control Panel / Regional and Language Options).

If the size of the program files is an issue for you, perform a custom installation.

Performance and speed are enhanced if your computer's processor, memory, and available disk space exceed minimum requirements. This is especially true when converting very large color PDF files. OmniPage is specifically engineered to take advantage of multi-core processors.

Install OmniPage

The Kofax OmniPage Ultimate 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 own a previous version of OmniPage, or if you are upgrading from demonstration software or an OmniPage Special Edition, you must uninstall that product first.

To install OmniPage:

1. Download the program file and choose **Run** when the download is completed.
   The installation program is started automatically.
2. Choose a language to use during installation.
3. Click **Start installation** to continue.
   The InstallShield Wizard is started.
4. Read and accept the License Agreement by clicking **I accept the terms in the license agreement**.
5. Type **User Name** and **Organization**.
6. Accept the default program folder location or click **Change** to browse to a specific folder.
   All files are copied automatically during installation.
7. The Kofax Power PDF Create installer is started. See Install Kofax Power PDF Create for details.
8. The registration dialog box appears. See Register OmniPage for details.
9. Kofax OmniPage Ultimate is started with the activation wizard. See Activate OmniPage for details.

- Kofax OmniPage Ultimate is supplied with a complimentary copy of the PaperPort Professional document management product, which has different system requirements and must be installed separately.

Install Kofax Power PDF Create

The Kofax OmniPage Ultimate installation starts the Power PDF Create installer.

1. Choose a language to use during installation.
2. Click **OK** to continue.
   All files are copied automatically during installation.
3. Click **Finish**.

Register OmniPage

When finished, the Kofax OmniPage Ultimate installer starts the Product Registration dialog box. Registration requires a web browser and Internet connection.

1. Select one of the following options.
   - **Register online**

     The online registration form appears in your default browser application. Fill the form and click **Submit**.
   - **Remind me in 7 days**

     If you did not register the software during installation, you will be periodically invited to register later.
2. Click **OK** to continue.

Activate OmniPage

Kofax OmniPage Ultimate starts the Product Activation wizard as launched. Activation requires a valid serial number and internet connection.

Note Refer to the *OmniPage 19 Release Notes* for details on how to install OmniPage 19 on a computer without an Internet connection.

1. Select one of the following options.
   - **Use in Trial Mode**

     Starts OmniPage without activation, commencing a 15-day trial period. During this period the activation wizard will appear each time you start OmniPage. Kofax OmniPage Ultimate can be launched any number of times within the trial period.
   - **Activate product**

     Provide your serial number, which may contain both letters and numbers. (Dashes are placed automatically as you type.) A green check mark on the right indicates whether the number is correctly entered. Activation requires no user interaction, and no personal information is transmitted.
2. Click **OK** to continue.

Set up your scanner with OmniPage

All files needed for scanner setup and support are copied automatically during the program installation, but no scanner setup occurs at installation time. Before using OmniPage 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, which you can start yourself, as described below. Otherwise, the wizard appears when you first attempt to perform scanning. Proceed as follows:

1. Do one of the following to start the **Scanner Setup Wizard**:
   - Choose **Start** > **Kofax OmniPage Ultimate** > **Scanner Setup Wizard** .
   - Click the **Setup** button in the **Scanner** panel in the **Options** dialog box.
   - Choose **Scan** in the **Get Page** drop-down list in the OmniPage Toolbox, and then click the **Get Page** button.

   The **Scanner Setup Wizard** is started.
2. If you have a web connection, the first panel prompts you to update the scanner database supplied with the wizard. Choose **Yes** or **No** and click on **Next**.
3. Choose **Select and test scanner or digital camera**, and 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**.
4. You see a list of all detected scanner drivers in the selected categories, which may include network devices. Select one and click **OK**. To install a second device, you must run the scanner wizard again.
5. 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. Otherwise, you should test it. Click **Next**.
6. 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 to run. The Basic test scan is recommended.
7. By default, OmniPage uses its own scanning interface, located in the **Scanner** panel in the **Options** dialog box. To use your scanner's own interface instead, choose **Advanced** settings. Click **Hint editor** and choose **Edit hints** only if you are experienced in configuring scanners or have been advised to do so by technical support.
8. Click **Next** to start the tests. For the Basic scan test, insert a test page into your scanner. The wizard starts to scan using the software provided by your scanner manufacturer. Click **Next** to display the native user interface for your scanner.
9. Click **Scan** to begin the sample scan.
10. If necessary, click **Missing Image** or **Improper Orientation** and make the appropriate selections. Once the image appears correctly in the window, click **Next**.
11. Move through the remaining requested tests, following the instructions on the screen.
12. When all requested tests are completed successfully, the **Scanner Setup Wizard** displays a confirmation and prompts you to click **Finish**.

    You have successfully configured your scanner to work with Kofax OmniPage Ultimate.

- To change the scanner settings at a later time, or to set up or remove a scanner, reopen the Scanner Setup Wizard from the Windows Start menu or from the Scanner panel in 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, and then work through the procedure described above, based on any 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, choose **Use current settings with current device**, and then click **Next** and **Finish**.

Start the program

Kofax OmniPage Ultimate features OmniPage Launchpad, a clear-cut metro-style start page for simplified, faster conversions. Click **Start** in the Windows taskbar, then locate and select **Kofax OmniPage Ultimate** > **OmniPage Launchpad** .

1. Build panel, Convert column: Choose a page type that best describes the layout of the input document.
2. Build panel, To column: Choose the output file type.
3. Build panel, Save column: Choose a destination for the recognition results.
4. Currently selected Convert tile.
5. Currently selected To tile.
6. Currently selected Save tile. These three form the Go-flow in the fourth slot.
7. Go-flow: Consists of the currently selected Convert, To and Save tiles.
8. Currently selected Go-flow, compiled from the selected Build Panel tiles.
9. Last unfilled Go-flow slot.

10. Runs the selected Go-flow.
11. Settings bar: Collection of eight buttons (six of them with two different states) for managing the prepared Go-flows. From left to right: Run Go-flow, File separation, Zoning (on or off), Proofing, Language, Display results (on or off), Unlocked / Locked, Clear Go-flow.

To start Kofax OmniPage Ultimate, do one of the following:

-

  Click **Start** in the Windows taskbar. Then locate and click **Kofax OmniPage Ultimate**.
-

  Double-click the OmniPage icon in the program folder, or if applicable, on the Windows Desktop.
-

  Double-click an OmniPage Document (OPD) icon or file name; the clicked document is loaded into the program.
-

  Right-click one or more image file icons or file names on a shortcut menu. Select **Open With OmniPage application**. The images are loaded into the program.

On opening, the OmniPage title screen is displayed with a view selection panel. OmniPage has three basic view types. 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 DocuDirect program. Click **Start** in the Windows taskbar, then locate and click **Kofax OmniPage Ultimate** > **OmniPage DocuDirect** .
-

  Click the **Acquire Text** button on the Kofax OCR toolbar for a supported application registered with the Direct OCR™ facility.

-

  Right-click one or more image file icons or file names in Windows Explorer for a shortcut menu. Select Kofax OmniPage Ultimate and choose a target format, or the Convert Now Wizard or a workflow from the submenu. The files are processed according to the workflow instructions.
-

  Click the OmniPage Agent icon on the taskbar. Choose a workflow to start the program and run the workflow.
- Use Kofax OmniPage Ultimate with the PaperPort document management product to add OCR services.

Uninstall the software

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

- Zone templates (*.zon)
- Image enhancement templates (*.ipp)
- Training files (*.otn)
- User dictionaries (*.ud)
- OmniPage documents (*.opd)
- Job files (*.opj)
- Workflow files (*.xwf)

To uninstall, you must be logged into your computer with administrator privileges.

To uninstall or reinstall OmniPage:

1. Close OmniPage.
2. Click **Start** in the Windows taskbar, locate and start the **Control Panel**, and then locate and start **Uninstall a program**.
3. Select **Kofax OmniPage** and click **Uninstall**.
4. Click **Yes** in the dialog box to confirm removal.
5. Select **Yes** to restart your computer immediately, or **No** if you plan to restart later.
6. Follow instructions until the process is finished.

When you uninstall OmniPage, the link to your scanner is also uninstalled. You must set up your scanner again with OmniPage if you reinstall the program. With Kofax OmniPage Ultimate, PaperPort must be uninstalled separately.

Chapter 3

Using OmniPage

Kofax OmniPage Ultimate 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)
- *Text formatting* (character and paragraph)
- *Page formatting* (column structures, table formats, headings, placement of graphics)

Documents in OmniPage

A document in OmniPage consists of one image for each document page. After you perform OCR, the document also contains 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.

An OmniPage Document can contain an embedded user dictionary, training file, zone template file, or an image enhancement template file. This can increase file size considerably but makes the OPD more portable. To embed a file, open the relevant dialog box from the Tools menu, select the desired file and click Embed. Use the Extract button to get a local copy of an embedded file inside an OPD you have received.

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

OmniPage desktop and views

OmniPage comes with three different views to suit your task.

- **Classic view**: Offers a look and feel similar to previous versions of OmniPage.

- **Flexible View**: Provides an alternate layout of the OmniPage function panels stacked in a tabbed view to give each panel more space.
- **Quick Convert View**: Offers simplified, quick and easy document conversion without having to become an expert. The most important conversion options are clearly visible on one screen.

Use the Window menu to switch between views and save your own custom view. On starting a new session, you receive the view and screen arrangement that was in effect when the program was last closed.

All three views can be reset to default values using Reset Current View in the Window menu.

Program panels

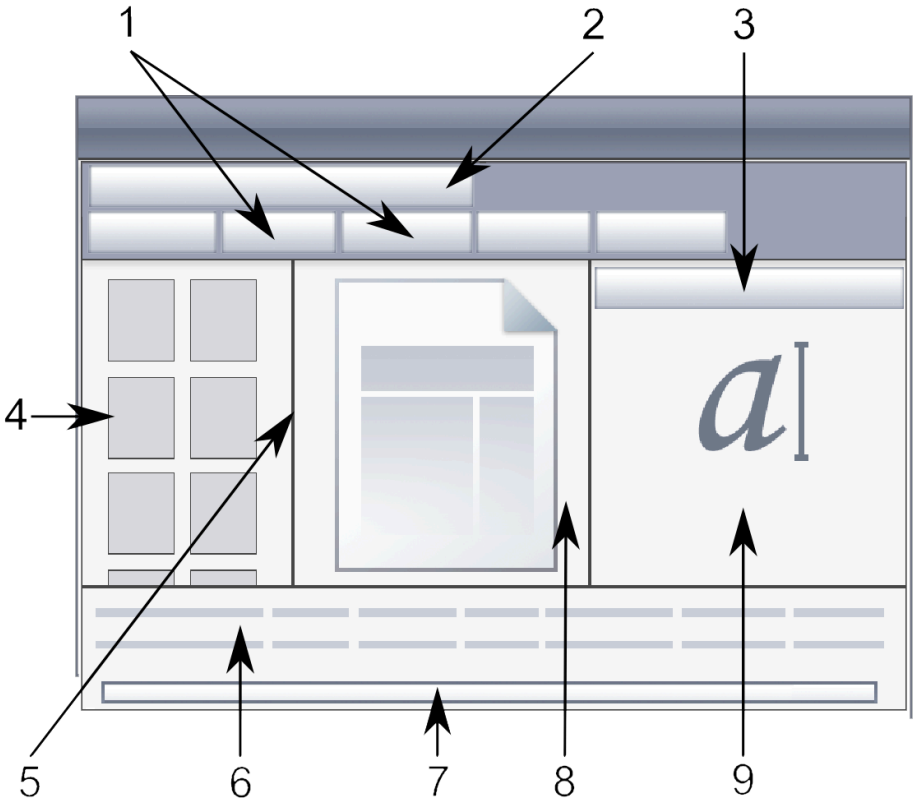
OmniPage has a set of panels that can be docked (tabbed or tiled), floated, resized, minimized and restored separately:

- Thumbnails
- Page Image
- Text Editor
- Document Manager
- Easy Loader
- Workflow Status
- Help

To float a panel, double-click its title bar or tab. To restore the floating panel to its previous docked position, double-click its title bar. To dock it to a new location, drag it to that location. A colored rectangle shows the docking position; release the mouse button to dock it. To see all possible docking positions one after the other (tiles and tabs), drag the panel over the OmniPage main window, holding down the left mouse button and pressing the spacebar repeatedly. When the desired location is indicated by coloring, release the mouse button. To move a floating panel without docking displays, keep Ctrl pressed while dragging.

Classic View

In Classic View, the default OmniPage Desktop has four main tiled working areas, separated by splitters: the Document Manager, the Page Image, Thumbnails and the Text Editor. The Page Image has an Image toolbar and the Text Editor has a Formatting toolbar.



1. OmniPage toolbox: This toolbox lets you drive the processing.
2. Standard toolbar
3. Formatting toolbar
4. Thumbnails panel: Displays page thumbnails.
5. Image toolbar
6. Document Manager: 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.
7. Status bar
8. Page Image: Displays the image of the current page with its zones. When a page is displayed, the Image toolbar is available.
9. Text Editor: Displays recognition results from the current page.

Panels can be rearranged freely, either horizontally or vertically; use the Window menu to open the Easy Loader, Workflow Status or Help panels. Panels can be minimized or closed, but not tabbed. To restore the default Classic View appearance, choose Reset Current View in the Window menu.

Flexible Views

Use this view to set up the OmniPage workspace so that it fits your task optimally. By default, all panels appear. There are five tabs:

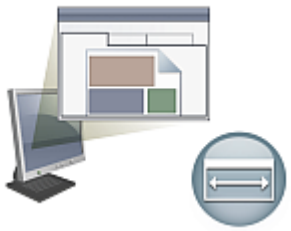
- Page Image (including Thumbnails)
- Text Editor
- Easy Loader
- Workflow Status
- Help

The Document Manager appears in a horizontal panel at the base of the working area. You can undock, move, minimize, group or close panels as already described. Drag a tab onto the working area to convert it to a classic-type tiled panel. Drag it back to the tab bar to revert to a tabbed panel, or use the Spacebar as already described. If panels are grouped, the tab name shows the active one. To restore the default Flexible View appearance, choose Reset Current View in the Window menu.

Easy Loader provides a Windows Explorer type file listing and functionality that can remain open during the session, allowing quick file selection and assembly.

Suggested scenarios:

Maximizing workspace (single screen)



Load a document. Open the panels you want to use. Grab them by their captions one by one, and drag them so that they dock beside the active one as tabs. You can also dock Help to avoid handling two separate windows.

Working with recognition results (single screen)

Load a document and have it recognized. Close all panels except the Document Manager and the Text Editor. Maximize both horizontally, scale down the Document Manager and dock it to the top or bottom. You can now step through the pages, double-clicking them one by one in the Document Manager, and inspecting recognition results in the Text Editor. The number of suspect words and reject characters in the Document Manager help you identify problematic pages.

Handling large documents (dual-screen)

Load the document you want to work on. Move its Thumbnail View to your second monitor and maximize it for a large scale overview of your document and far more space for thumbnail operations.

Verifying (dual-screen)

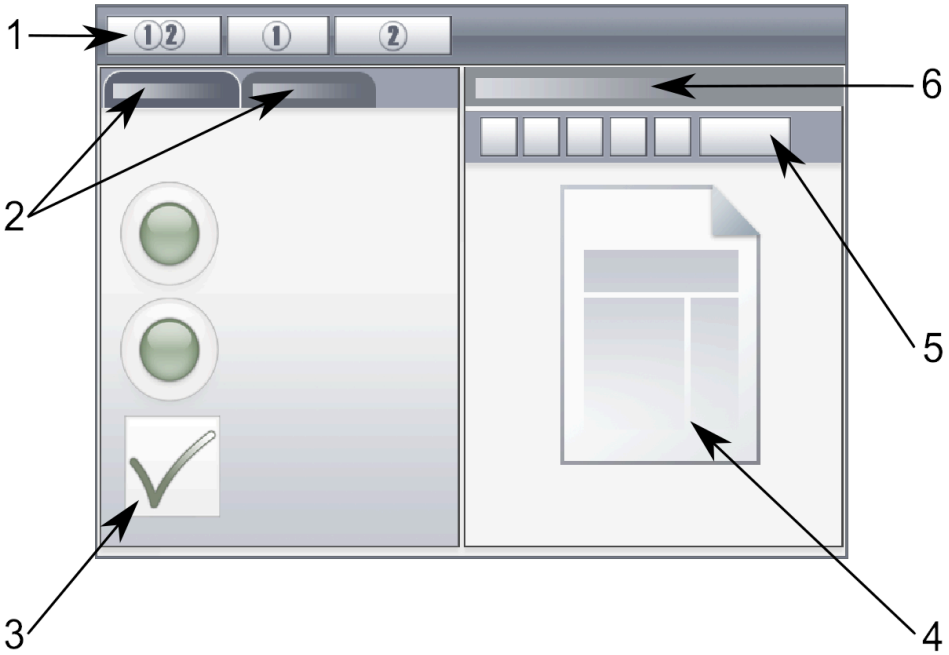
Place the Page Image on one screen and the Text Editor on the other to gain more space for editing and proofing.

The Page Image is always available for verifying recognition and for performing on-the-fly zoning and editing.

The scenarios presented are only examples to give you an idea of what you can do in Flexible view.

Quick Convert View

Use the Quick Convert View for fast recognition and saving. You can switch to Quick Convert View only when you have no opened document and it can handle only one input file and one output document at a time. The picture shows the default appearance.



1. Processing buttons
2. Quick Convert Options on toggled tab with Easy Loader
3. Quick Convert Options: Document source and layout output text format, formatting level output folder and file name saving options page range
4. Page Image
5. Quick Convert toolbar
6. Page Image panel title

The Easy Loader is by default on a tab that toggles with the Quick Convert Options panel. A Help panel can be added, but further panels are not available in this view. You can change tabs to separate panels and minimize them, as in other views.

After loading a file, you should convert it before loading the next file. When an image conversion is finished, you do not need to explicitly close the image; just load a new file.

The Easy Loader in Quick View provides an additional feature: one-click processing. Choose the Easy Loader submenu in the Process menu and choose either Load Files or Get and Convert. When the latter

is chosen, multiple files can be selected. These files are loaded, recognized and saved using the current settings. For this, set the output file names to be the same as the source file names.

The Page Image panel in Quick Convert View includes the Quick Convert toolbar, offering the most useful image handling operations. To access advanced functionality, such as image file saving, SET tools, on-the-fly zoning, zone reordering and manual zone drawing for vertical text, a different view should be used.

Custom views

For a custom view, arrange the panels and toolbars as you wish, then choose **Window** > **Custom Views** > **Manage** . Click **Add** and name your view. Your screen layouts are displayed in the Custom Views submenu with a checkmark beside the active one. Resetting to a default is not available for custom views.

Changing views

Use the Window menu to change views. Panels are shown or hidden and arranged as they were when the chosen view was last used. The Help topic on display remains unchanged regardless of view. Easy Loader retains its file location regardless of view and the Workflow Status continues to display information on the last workflow run. On program restart, Help displays the Welcome topic, with Easy Loader as the default folder location with the Workflow status as empty.

Toolbars

The program has eleven main toolbars. Use the View menu to show, hide, or customize them. Status bar texts at the bottom edge of the OmniPage program window explain the purpose of all tools.

- Standard toolbar: Performs basic functions.
- Image toolbar: Performs image, zoning and table operations. Three of its tool groups can now be handled separately (mini-toolbars):
  - Zones toolbar: Offers zoning tools.
  - Rotate toolbar: Provides rotating tools.
  - Table toolbar: Inserts, moves and removes row and column dividers.
- 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.
- Mark Text toolbar: Performs text marking and redacting.
- Form Drawing toolbar: Creates new form elements.
- Form Arrangement toolbar: Arranges and aligns form elements.

All toolbars can be moved and customized in each view to your particular needs, including use of a secondary monitor.

The Form toolbars and the Mark Text toolbar appear only in Kofax OmniPage Ultimate.

Basic processing steps

There are three ways of handling documents: with automatic, manual, or workflow processing. 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 from your file system, storage sites in the cloud, and more.
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 Direct OCR or schedule processing to run at a later time.

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

Use OmniPage with Kofax PaperPort

The Kofax PaperPort program is a paper management software product from Kofax. 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 a text application’s PaperPort link, selecting Preferences and then selecting OmniPage Professional 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 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.

Kofax PaperPort is included in the Kofax OmniPage Ultimate package. For application information, refer to the PaperPort documentation. PaperPort must be installed and uninstalled separately from OmniPage.

When PaperPort is available, its folder structure is offered in the OmniPage Load Files and Save to File dialog boxes.

Chapter 4

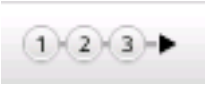
Processing documents

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

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 languages. You start each step by clicking the three buttons on the OmniPage Toolbox.

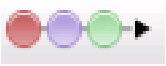
1. Use button one to get a set of images.
2. Manually zone pages where you want to process only part of the page or if you want to give precise zoning instructions. Use "Ignore backgrounds or zones" to exclude areas from processing. Use process backgrounds or zones to specify areas to be auto-zoned.
3. Use button two to have the pages recognized.
4. Do proofing and editing as desired.
5. Use button three to save your results.

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.

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 by pressing the Start button and choosing to process existing pages.

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. It does not have to conform to the 1-2-3 pattern of traditional processing. Workflows are listed in the Workflow list – sample workflows plus any you create. Workflows allow you to handle recurring tasks more efficiently, because all the steps and their settings are pre-defined. 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. Click Next after each step to add another one. You can use the Assistant just to get more guidance when doing automatic processing.

At a later time

You can schedule OCR jobs or other processing jobs in OmniPage DocuDirect to be performed automatically at a later time, when you may not even be present at your computer. This is done through DocuDirect. 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.

When you choose Create 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.

Processing from other applications

You can use the Direct OCR feature to call on the recognition services of OmniPage while you work in the following applications: Microsoft Office, Corel WordPerfect. First you must select the Enable Direct OCR check box under **Tools** > **Options** > **General** . Then, two buttons in the Kofax OCR tab of the Office application, or in an OmniPage toolbar open the door to OCR facilities.

How to set up Direct OCR

Start the application you want connected to OmniPage. Start OmniPage, open the **Options** dialog box at the **General** panel, and select **Enable Direct OCR**.

In the target application, use the Acquire Text Settings button in the OmniPage toolbar (in Office applications go to the Kofax OCR tab). Select options in the following panels:

- OCR: Languages, dictionaries, layout, fonts.
- Process: Image pre-processing, choices for PDF opening, feature retention.
- Output Format: Set a formatting level.
- Direct OCR: Automatic or manual zoning, perform or skip proofing, image source.
- Scanner: Set up or change scanner settings.

These functions for Direct OCR work until you change them again; they are not applied when OmniPage is used on its own.

How to use Direct OCR

1. Open your application and work in a document. To acquire recognition results from scanned pages, place them correctly in the scanner.
2. Use the **OmniPage** toolbar button **Acquire Text Settings** or the same item in the target application's File menu (or the **Kofax OCR** tab in Office) to review your recognition settings, if necessary; the **Direct OCR** panel lets you specify input from scanner, image file or digital camera image files.
3. Use the **OmniPage** toolbar button **Acquire Text** or the same item in the **File** menu (use the **Kofax OCR** tab in Office) to acquire images from the specified source.
4. If you selected **Draw zones automatically** in the **Direct OCR** panel of the **Options** dialog box, 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 in the **Output Format** panel under **Acquire Text Settings**.

Defining the source of page images

There are three possible image sources: image files, a digital camera, and 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 saved with the scanner's own software to an image file, which OmniPage can later open.

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

You can govern how PDF files are opened under **Tools** > **Options** > **Process** : open with the text layer or as image, import tag information to assist layout retention and whether to use PDF fonts or the mapped system fonts. See the eDiscovery Assistant for searchable PDF section on how to make image-only PDF files searchable.

Input from image files

You can create image files from your own scanner, or receive them by email or as fax files. OmniPage can open a wide range of image file types. Select **Load Files** in the **Get Page** 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.

Input from the Cloud

The Get Pages list offers direct connections to the Evernote web-based storage site.

When taking files from the Cloud, you may have to provide login information.

In Kofax OmniPage Ultimate, files can also be imported from Microsoft SharePoint, eDOCS, iManage and ODMA-compliant Enterprise Content Management sources.

Input from digital camera

Digital camera files are auto-detected in Kofax OmniPage Ultimate, hence there is no need to use the Load Digital Camera Files button. Auto-detection of camera files means that now they can be processed as camera files from any source, even from the cloud. However, in case a non-camera file whose content is similar to a camera file is to be processed, the Load Digital Camera Files button can be used. For tips and advice on working with digital camera images see the How-to-Guides and the Help.

Input via Easy Loader

This provides the Windows Explorer interface in an OmniPage window. In Flexible and Quick Convert views, it appears by default. Click **Window** > **Easy Loader** in the menu to add it to **Classic View** or to show or hide it in other views. It functions as an alternative to the File Open dialog box; letting you browse your whole file system and efficiently select files to be loaded into OmniPage. Click **Process** > **Easy Loader** > **Process** to view files as Lists, Thumbnails, Tiles, Icons (arranged as desired) or Details, as you do in Explorer. Easy Loader can remain displayed as you work.

Easy Loader is driven from the Process menu. Instead of selecting files to send them straight to OmniPage you can choose Queue Window to get a dialog box with a lock. Turn the lock on to build up and reorder a list of files, which may come from different folders. The lock applies to all files collected to enter the currently open document. When the list is ready, turn the lock off to start loading. If the lock is off from the start, files are listed only if they are selected faster than OmniPage can load them. Practically, you can load a few files, send them to recognition and while that is underway, build up the rest of the input list.

Turning on the menu item Show/hide Queue Window automatically causes the window to appear whenever files are listed but not yet loaded and to be closed as soon as the list is empty.

Easy Loader can be used in Classic View and Flexible View to compile files for multiple documents. Engage the lock, make document 1 active and collect files. Then make document 2 active and collect its files, and so on. When all is ready, remove the lock. Each document has its own lock, but the Process menu offers Lock all and Unlock all to lock or release all files destined for all documents. You can remove selected files with Delete, or all files in the current document's list with Delete All or Clear in the Process menu. Use Clear All to clear all files destined for all open documents. See a tutorial in Help on loading files for multiple documents.

Easy Loader is available as a panel in Quick Convert View. The Process menu has two commands unique to Quick Convert View.

- Get and Convert offers one-button processing: files are loaded, passed through recognition and saved to files using existing settings. Only in this case, multiple file selection is allowed with Quick Convert View; the result is one output document for each input file. Before starting, you should choose Same as the source file name under Output file name.
- Load Files performs file loading without recognition, as in other views. In Quick Convert View it allows only one file to be loaded at a time. Each file should be processed before selecting a new input file. In this case the Queue Window and its lock play no useful role.

Easy Loader can process digital camera images. Set this in the Quick Convert Options panel before invoking Easy Loader. If Scan is set as input, this setting is temporarily ignored and pages are loaded as normal (non-camera) images.

All Windows Explorer functionality is available in Easy Loader. For instance, you can also select files and use the shortcut menu item Kofax OmniPage Ultimate to send them via background processing to MS Excel, MS Word, PDF, RTF, Text and WordPerfect. Existing settings are used and by default, generated files are placed in the input folder. Use the Convert Now Wizard to access basic settings, such as whether to view results in the target application. This wizard lets you do immediate conversions or call the Workflow Assistant to access all settings, such as to change target file names and locations. This shortcut menu item also offers all workflows that have image file input.

Input from scanner

You must have a functioning, supported scanner correctly installed with Kofax OmniPage Ultimate. 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 to scan in black-and-white. Black-and-white images can be scanned and handled more quickly than others and occupy less disk space.

Scan grayscale
Select to use grayscale scanning. For best OCR accuracy, use this setting 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 to scan in color. This setting functions only with color scanners. Choose this setting if you want color graphics, text or backgrounds in the output document. For OCR accuracy, it offers no more benefit than grayscale scanning, but requires 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 in the Options dialog box or in your scanner 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 apply adjustments using a range of different tools.

Scan 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.

Scan without an ADF

Using the OmniPage 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 in the Options dialog box, and define a pause value in seconds. Then the scanner makes scanning passes automatically, pausing between each scan by the defined number of seconds, giving you time to place the next page.

Scan to OmniPage and workflows

Go to **Tools** > **Options** > **Scanners** to choose an action to be performed when a button on your local scanner is pressed. This can be simple scanning resulting in images loaded into OmniPage. It is also possible to select a scanner-based workflow from those you have created, or select to be prompted to select a workflow whenever the button is pressed. Use the Control Panel button to associate OmniPage with a scanner event (a scanner button being pressed). Then a button press launches OmniPage, runs the workflow and sends the results to the defined target, with or without interaction.

In Kofax OmniPage Ultimate, this feature can also be used to initiate barcode-driven workflows.

Note Document-to-document conversion:

OPU ULTIMATE

In Kofax OmniPage Ultimate you 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 choose from a wide variety of document file types for saving. These conversions require Kofax Power PDF Create to be installed.

Describe 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. The table lists your input description choices.

Option	Description
Automatic	Select 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	Select this setting if your pages contain only one column of text and no table. Typical examples are business letters or pages from a book.
Multiple Columns, no Table	Select if some of your pages contain text in columns and you want them decolumnized or kept in separate columns, similar to the original layout.
Single Column with Table	Select if your page contains only one column of text and a table.
Spreadsheet	Select if your whole page consists of a table that you want to export to a spreadsheet program, or treat as single table.
Form	Select 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. OPU ULTIMATE This option is available in Kofax OmniPage Ultimate only.
Legal Pleading	Select to recognize legal documents. Legal headers are detected and removed. Choose to have pleading numbers retained or dropped.
Custom	Select 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 in the Options dialog box.
Template	Select 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 yields unexpected recognition results, use manual processing to rezone individual pages and re-recognize them.

Preprocess images

To improve OCR results, you can enhance images before zoning and recognition using the Image Enhancement tools.

Click the SET - Enhance Image button in the Image toolbar to open the Image Enhancement window. This window consisted of the following:

1. Starting image panel displays the original
2. Result panel shows the preview of the output
3. Settings for the selected tool
4. Apply saves the settings.
5. Discard last change revokes the last modification only
6. Discard all changes withdraws all modifications.
7. History panel
8. Page Ready
9. Document Ready

When you click Apply, the result image moves to the left panel to become the new starting image for further enhancement. Changes are listed in the History panel (7). When all changes are in order, click Page Ready (8) to have the next page loaded or Document Ready (9) to finish enhancing.

We must distinguish three types of images:

- 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.

This tool lets you switch between the Primary and the OCR image.

Some tools affect the Primary image, while others affect 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 in the Options dialog box or in your scanner 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. Use the OCR Brightness tool to optimize the image.

The three examples in the middle give optimal OCR results. The two light samples at the top and the two dark samples at the bottom are unsuitable for OCR.

Image Enhancement tools

The Image Enhancement tools can also be used to edit primary images to save and use them as image files. The following tools are accessible on the toolbar from left to right; their usage is detailed as follows:

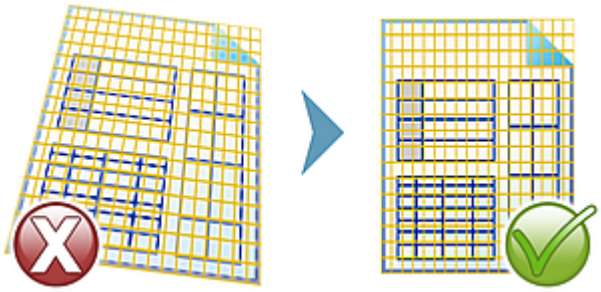
- P: Affects Primary image only.
- O: Affects OCR image only.
- PO: Applies to either the Primary or OCR image (or both)
- P+O: Single action is applied to both the Primary and OCR image.
- P/O: Affects both images.
- WH: Applies to whole images only.
- AR: Applies to selected image areas.

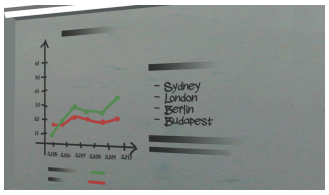
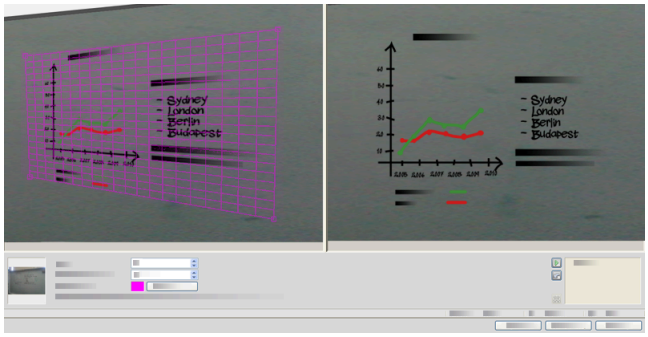
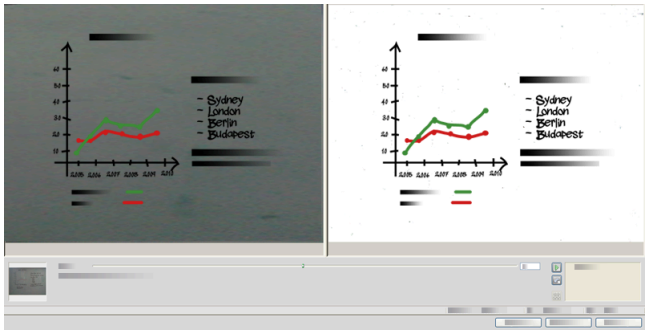
Tool	Description
Pointer (F5)	The Pointer is a neutral tool carrying out different operations under different circumstances (for example, to pick a color for the Fill operation, or to catch the deskew line.) PO.

Tool	Description
Zoom (F6)	Click the tool, and 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. P+O. WH.
Select area (F7)	Click this tool, and then on a tool that can work on a page area (marked AR) and draw your selection on the image. Image enhancement tools by default work on the whole page. Selection has three modes (in the View menu): Normal, Additive, and Subtractive. PO. AR.
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. P/O. WH.
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. PO. WH.

The following SET tools allow you to modify image contents.

Tool	Description
Brightness /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. P. AR.
Hue / Saturation / Lightness	Click this tool, and then use the sliders to modify the hue, saturation and lightness of your primary image. P. AR.
Crop Image	To use only a part of your image, click the Select Area tool, then the Crop tool and select the area to keep: the rest of the image is removed. P+O. WH > AR
Rotate	Click this tool to rotate (by 90, 180 or 270 degrees) and/or flip your image. P+O. WH.
Despeckle	Click this tool to remove stray dots from your image. Despeckle works on the OCR image at 4 levels of severity. You can also use this tool not to remove noise from the page but to strengthen letter outlines: to do this, select the checkbox Inverse despeckling. O. AR.
OCR Brightness	Use this tool the set Brightness and Contrast of your OCR image. See the diagram of optimum brightness under Preprocessing Images above. O. AR.
Drop-out Color	Click this tool and select Red, Green, Blue; or choose a color from the primary image with the Select Area tool. Sections of the scanned image in this color are set to transparent. The tool has its effect on the OCR image. This feature enables a chosen color to be dropped when preprinted color forms are scanned or loaded. Then the fixed texts, boxes and other elements can be dropped from the images, leaving only the respondent data visible and ready for OCR. P/O. WH.
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. P. WH.
Deskew	Sometimes pages are crooked when scanned. To straighten the lines of text manually, use the Deskew tool. (Auto-deskew is also available in the Process panel of Options.) P+O. WH.

Tool	Description
3D Deskew	Use this tool to remove perspective distortion from digital camera images. This is particularly useful when you want to check the results of automatic 3D Deskew or you prefer to do 3D deskew manually after a Load Files step. P+O. WH. 3D Deskew works by snapping the distorted image to a grid. All you need to do is to manually straighten this grid, and image coordinates will follow; see the illustration below.  **Before - after 3D Deskew**
Fill	Use this tool to apply a color to the image or a selected part of it. PO. AR.
Auto-crop	Automatically detects margin areas on the page and reduces them to a minimum. Use Auto-crop to unify the margins on a set of pages with different sized text areas. P+O. WH > AR
Clean borders	Removes scanning shadows, spots, and marginal notes from page edges. P+O. WH but relates only to the border area.
Punch Hole Remover	Replaces punch holes with the background page color. P+O. WH but relates only to the border area.

Tool	Description
Enhance whiteboard photo	Provides a slider control to let you improve the readability of text and diagrams on whiteboards or blackboards, when captured by digital camera. The following pictures show the possible difference when using this tool along with the 3D Deskew tool.  **A typical digital photo of a white board, taken from the side with low contrast**  **The 3D deskew is being applied, with the result on the right**  The Enhance whiteboard photo tool slider is used to improve the contrast of the image. **The starting image on the left; the result on the right**

Some of these tools are also available for automatic pre-processing of all incoming images. These are shown on the Process panel in the Options dialog box.

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 starting image, the result window displays the changes.

-

  Click **Apply** next to the History list to commit the change. Modifications not added to the History by clicking the Apply button are not performed.

-

  Click **Reset** to discard changes you have performed with a given tool, before they are applied.

-

  Click **Discard all changes** to restore the image as it was before you started the current enhancement session.

Any time you want to see what output a certain step resulted in, double-click it in the History list. The display shows the result of that action, removing all actions performed afterwards. If you apply a new change to the displayed image, that replaces all changes that were made in the History list after the chosen one.

Save and apply templates

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 by clicking the Apply button. When you are done, choose Save Enhancement Template from the Image Enhancement window 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 same File menu.

Image Enhancement in workflows

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 is displayed for manual editing.

Apply enhancement template
An already saved enhancement template is applied automatically to the image while being processed by the workflow.

Apply enhancement template and display
The workflow applies the selected image enhancement template, and also displays 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.

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 tool for setting a process background. Draw ignore zones over parts of the page you do not need. After recognition the page returns with an ignore background and new zones round all elements found on the background.

Auto-zoning vertical text

If you set Japanese, Korean or Chinese as the recognition language, auto-zoning will find text blocks and detect the text direction.

Vertical Asian text appears horizontally in the Text Editor, but can be exported as vertical.

Auto-zoning detects vertical texts in non-Asian languages in table cells and anywhere on Normal PDF or XPS pages. Multi-line detection is possible in these cases.

For image-only PDF and XPS files, and for all other image file or scanner input, auto-detection works with the following conditions:

- It must be only a single line of text
- It must be on the left or right of a diagram or picture
- It must be situated on the left or right edge of the page and does not have to extend over the full height of the page.

Vertical text outside tables can be manually zoned, as described below. This allows multiple vertical lines to be handled correctly.

Vertical text can be viewed and edited with a vertical cursor in the Text Editor using True Page. In other formatting levels the text is placed horizontally.

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 to open a shortcut menu with options to change the zone properties. Select multiple zones with Shift+clicks to change their properties in one step.

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

Tool	Description
Process Zone	Use to draw a process zone, to define a page area where auto-zoning will run. After recognition, this zone is replaced by one or more zones with automatically determined zone types.
Ignore Zone	Use to draw an ignore zone, to define a page area you do not want transferred to the Text Editor.
Text Zone	Use to draw a text zone. Draw it over a single block of text. Zone contents are treated as flowing text, without columns being found. Apply to text that uses the Latin, Greek or Cyrillic alphabets and for horizontal text in the Asian languages.
Vertical Asian Text Zone	Use to draw text zones for vertical text in Japanese or Chinese. Zones should be rectangular.
Vertical left-rotated Text Zone	Use to draw text zones for vertical text that is left rotated (non-Asian languages only). The zones should be rectangular.
Vertical right-rotated Text Zone	Use to draw a text zone for vertical text that is right rotated (non-Asian languages only). The zones should be rectangular.
Table zone	Use to have the zone contents treated as a table. Table grids can be automatically detected or placed manually. Table zones should be rectangular. Vertical texts in tables cannot be zoned manually – they can be auto-detected in table grids.
Graphic Zone	Use 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 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. Afterward, in True Page, you can edit form layout, and modify the properties of form elements. OPU ULTIMATE Form zones are available in Kofax OmniPage Ultimate only.

Working with zones

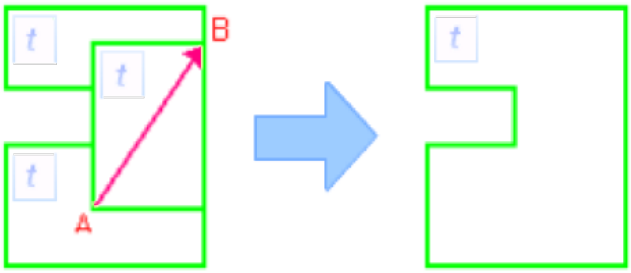
The Image toolbar provides zone editing tools. Grouped tools can be undocked/floated and re-docked as a separate mini toolbar for convenience. One is always selected. When you no longer want the service of a tool, click a different tool. Some tools on this toolbar are grouped. If docked as a single tool, 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).



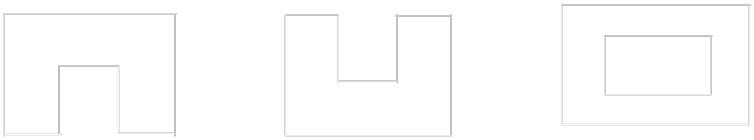
To make an irregular zone by subtraction, draw an overlapping zone of the same type as the background.

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

A full set of zoning diagrams appear in 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, and 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; or 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 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 Add 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.

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 are 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 are no longer 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
Open a document, then click Tools and choose Zone Template. Select the one you want to include and click Embed. Then save the document to the OPD format. This means the template will travel with the OPD if it is sent to a new location. When the OPD file is opened later, the included zone template is shown in the Zone Template Files dialog box as [embedded] and can be saved to a new named template file at the new location by using the Extract button.

Chapter 5

Proofing and editing

Recognition results are placed in the Text Editor. They can be recognized texts, tables, forms and embedded graphics. This WYSIWYG (What-You-See-Is-What-You-Get) editor is detailed in this chapter. Asian text handling is in some respects different from other languages.

The Text Editor display and formatting levels

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 Ultimate. 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, they are shown with a yellow highlight, both in the Text Editor and the OCR Proofreader.

Choose to have non-dictionary words marked or not in the Proofing panel in the Options dialog box. All markers can be shown or hidden as selected in the Text Editor panel in 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 formatting levels except Plain Text.

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

Plain Text

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

Displays decolumnized text with font and paragraph styling.

True Page

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.

Proofread OCR results

After a page is recognized, the recognition results appear in the Text Editor. Proofreading starts automatically if requested in the Proofing panel in 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. As an alternative to clicking a suggestion to select it and **Change** to accept it, hold down the Ctrl key and type the suggestion number.
6. 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.
7. 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. Choose **Recheck Current Page** from the **Tools** menu to re-proof a page.

Verify 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 the zoom icons to zoom in and out, and click any of the tree page icons to set the context range for the dynamic Verifier (you may switch between one word, three words or the whole image line).

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 Help.

Character Map

The Character Map is a dockable tool to aid you 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.

To access the Character Map, click the button in the **Formatting** toolbar, or click **Character Map** in the **View** menu and click **Show**.

Under the **Character Map** menu item, you can also choose to display recent characters only, or different character sets (by default only two are displayed). Asian characters are not supported.

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. Click the ellipsis (...) button next to the **Correct** field.
- Select **Train Character** from the shortcut menu of a suspect or non-dictionary word in the **Text Editor**.

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 in 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.

To embed a user dictionary in an OmniPage Document, load your input file, choose **Tools** > **User Dictionary** ; select the user dictionary you want to use, click **Embed**, and name it. Then save to the file type OmniPage Document.

Languages

The program can read over 120 languages with multiple alphabets: Latin, Greek, Cyrillic, Chinese, Japanese and Korean. See the full language list in the OCR panel in the Options dialog box. It shows which languages have dictionary support. Select the language or languages that will be in documents to be recognized. Selecting a large number of languages may reduce OCR accuracy.

A language listing is also provided on the Kofax website.

The option **Detect single language automatically** removes the need to select languages. It is designed for unattended processing when documents or forms in different languages are expected. OmniPage then examines each incoming page and assigns a single recognition language to the whole page. That means this feature is not suitable for pages containing multiple languages.

The program chooses from the languages with dictionary support that use a Latin-based alphabet (meaning Russian and Greek are excluded) plus optionally Asian languages. Choose from three language groups:

- Latin-alphabet languages (select to see the enabled languages)
- Asian languages (Japanese, Korean and Chinese – Traditional and Simplified)
- Latin-alphabet and Asian languages.

When this feature is enabled, no manual language selection is possible and the option Verify language choices (see below) is not available.

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 in the Options dialog box.

Asian language recognition

Four languages with Asian alphabets are supported: Japanese, Korean, Traditional Chinese and Simplified Chinese. The ideal font size for body text is 12 points, scanned at 300 dpi, resulting in characters with around 48x48 pixels. Minimum is 30x30, that is 10.5 points at 300 dpi. For smaller characters, 400 dpi should be used. Asian texts can be horizontal (left-to-right) or vertical (top-to-bottom, right-to-left). Operating systems supported by Kofax OmniPage Ultimate can handle Asian languages, but if East Asian language support was not selected during system installation, it must be added from Control Panel / Regional and Language Settings / Languages / Supplemental language support / Install files for East Asian languages. You may be required to insert a Windows system disk.

The four Asian languages are listed alphabetically with the others in the Options/OCR panel. You should select only one of these languages at a time and avoid a multiple selection with other languages. Asian

OCR can handle short embedded English texts without English being explicitly set; this is not designed for longer English texts or for texts in other Western languages. Vertical text is typical in Japanese and Chinese; English may be embedded in different orientations. The program can handle these; in the output they appear right-rotated.

Beside the language list, the option **Verify language choices** invokes automatic language detection that warns of differences between a detected language and the language setting. It works at page-level and identifies four categories: Japanese, Chinese, Korean and non-Asian. It cannot distinguish between Traditional and Simplified Chinese or between non-Asian languages. The last category means Japanese, Chinese or Korean characters were not detected. Verification takes place during image pre-processing, so the required recognition language must be set before image loading.

Auto-layout and auto-zoning are recommended for Asian pages to place all detected text into text zones; by choosing an Asian recognition language you set Asian OCR to run in these zones and that can automatically detect and transmit the text direction, coping with mixed areas of horizontal and vertical texts on a page.

However, the zoning tool 讯 lets you force vertical Asian recognition by manual zoning. Draw rectangular zones with this tool. To manually zone horizontal Asian text, use the usual text zone type. Do not use the two other vertical-text tools on Asian texts. Drawing a vertical Asian zone does not automatically enable an Asian language, nor influence the language auto-detection.

Digital camera images are accepted for Asian languages. However, the automatic 3D deskew algorithm is unlikely to be useful, and certainly not for vertical texts. Preferably use the standard image loading command and perform manual 3D deskewing with the relevant SET tool if required. In general, SET tools can be used on Asian images.

Recognized Asian pages appear in the Text Editor, provided your system has support for East Asian languages, and always with horizontal text direction. There is no need to specify Asian fonts under Options/OCR; a default font is automatically applied: typically Arial Unicode MS. Other Asian-capable fonts on your system can be chosen in the Text Editor. Editor support allows text viewing and verifying, and Formatted Text is recommended as the formatting level. Large-scale editing and spell-checking are better done in the target application. Proofing, training and dictionary support are not available for Asian texts. Therefore, prior to performing Asian OCR, go to the Proofing panel under Options and disable dictionary word marking, automatic proofreading and IntelliTrain and ensure that no training file is loaded. Redaction can be applied to Asian texts, either by selection or searching. The workflow step Form Data Extraction should not be applied to Asian pages.

Typical output converters for Asian texts are RTF, Microsoft Word, Searchable PDF or XPS. The text direction will be as detected during pre-processing. Changes made in the Text Editor - where text is horizontal - will be exported, also to vertical text. Plain Text converters are available (Unicode TXT, Notepad) but here text direction is always horizontal.

Training

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 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 it 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 them. Incorrect words should be retrained 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 or not to apply the user correction.

You can turn IntelliTrain on or off in the Proofing panel in 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.

Training files

Whenever you close a document or switch to another one when unsaved training data exists, you are prompte to save it. To save a training file into an OPD, load it from **Tools** > **Training File** , click **Embed**, and save to the file type OmniPage Document.

Saving training to a file, or 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 formatting levels except Plain Text, you can change the font type, size and attributes (bold, italic and underlined) for selected text.

Editing paragraph attributes
In all formatting levels except Plain Text, 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 and 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 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 email 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.

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 formatting level.

Reading order can be displayed and changed. Click the **Show /Hide 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 the formatting levels Plain Text and Formatted Text. It modifies the way the cursor moves through a page when it is exported as True Page.

On-the-fly editing

Use on-the-fly editing to modify a recognized page through rezoning, without having to reprocess 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 True Page formatting level, 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 to activate on-the-fly editing. The red indicator shows there are no stored zoning changes.
- Click to turn on-the-fly editing off. Your zoning changes are stored; the on-the-fly tool displays a green indicator to show there are stored changes. To activate these changes, do one of the following:
  - Click the on-the-fly tool with a green indicator. 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.

Marking and redacting

The Mark Text toolbar gives you tools to mark (highlight or strike-out); and to redact text. Use the View menu to have this toolbar displayed. You can float or dock this tool group. Each tool has its equivalent menu item in the Format menu or the Text Editor shortcut menu.

Redacting is blacking out confidential information. It is unreadable and unsearchable. To mark and redact text manually, click the **Mark for redacting** tool and use its cursor to select all the text parts you want to redact. They appear with a gray highlight. When you are ready, click the **Redact Document** tool. Choose to do redaction in a copy (safer) or the original document. If you choose to redact a copy, both the copy and the original remain open in OmniPage, ready to be saved.

Note If you redact the original document, you cannot retrieve the information you have blacked out.

To find and redact text by searching, select **Find and Mark Text** from the **Edit** menu to display the **Find, Replace and Mark Text** dialog box. Search for text to be marked for redaction. Step through all occurrences and decide for each case whether to redact immediately or mark for redaction. In the latter case, perform the redaction by choosing **Close and Redact Document** in the **Mark Text** dialog box or later click the **Redact Document** button.

You can apply highlighting and striking out either by selection or searching.

Reading text aloud

OmniPage provides speech facility using Windows Speech (Windows SAPI), which is a great help for the visually impaired, but it can also be useful to anyone during text checking and verification. This feature

requires Windows 10 or better. 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 F10 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 *

Text-to-speech languages can be installed and configured in Windows Control Panel, using the Region and Speech settings.

Creating and editing forms

You can bring paper or static electronic forms (distributed mainly as PDF in an office environment) into Kofax OmniPage Ultimate, recognize them and edit their content, layout or both, in True Page. 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. Kofax OmniPage Ultimate uses the Logical Form Recognition technology to create fillable forms from static ones.

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

Form Drawing toolbar

Use this dockable toolbar in the Text Editor to create a range of form elements using the following tools.

Tools	Description
Select Objects	Click to select, move, or resize elements in your form.
Text	Use to add fixed text descriptions on your form such as titles, labels and headers.
Line	Use primarily in layout design: click the tool to draw lines to separate distinct sections in your form.
Rectangle	Click to create rectangles in your form for design purposes.
Graphic	Use to select areas of your form that are to be treated as graphics.
Fill Text	Click to create fillable text fields where you want people to enter text.
Comb	Use to create a text field consisting of boxes. This tool is typically used for information such as ZIP codes.
Checkbox	Click to draw checkboxes: typically for Yes/No questions and marking one or more choices.
Circle Text	Similar to the checkbox element above, use the Circle text tool to create elements that are encircled when selected.
Table	Use to create tables in your form.

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

Form Arrangement toolbar

Use the options on this toolbar to line up form elements, or to set which one is on top of the others when they overlap. For example, the latter function is useful 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 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. Properties dynamically vary, depending on what type of element you select.

Form data extraction

Form data extraction (FDE) is a workflow step. Data is extracted from elements such as fillable fields, check boxes, and option buttons. FDE is a simplified implementation of the full Logical Form Recognition technology.

To create a workflow that contains form data extraction:

1. Define the processing input and its settings. Input types include image PDF, PDF form, image files and forms scanned from paper.
2. Choose **Extract Form Data** in place of recognition, and specify settings, including a language choice. The option "Detect single language automatically" can be useful for unattended processing of forms when the language used to fill each of the forms cannot be determined beforehand.
3. Set an active PDF form as a template (single or multi-page, filled or unfilled). The program determines the location and type of the form fields based on this form template.
4. Finish the workflow with a save step.

OmniPage extracts data from incoming forms, using the specified template. Export is to a comma-separated value text file (.csv) ready to be loaded into a spreadsheet.

Once you select Form Data Extraction in a workflow, only save steps will follow.

Chapter 6

Saving and exporting

Once you have acquired at least one image for a document, you can export the image to file. Once you have recognized at least one page, you can export recognition results. After further recognition you can save a single page, selected pages or the whole document by saving to file, copying to the Clipboard or sending to a mailing application. Saving as an OmniPage Document is always possible. OmniPage provides comprehensive support for Office applications and formats.

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 DocuDirect 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 DocuDirect job must contain at least one saving step.

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. 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.

Save 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 Files** in the **Export Results** list. In the dialog box that appears, select the **Image** option under **Save as**.
2. Choose a folder location and a file type. Type in a file name.
3. Select to save the selected zone images 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 or XPS 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 **Options** 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 images as specified. Zones and recognized text are not saved with the file.

Save 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 Files** selected in the list.
2. The **Save to Files** 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.
4. Type in a file name. Click **Options** to specify precise settings for the export.
5. Click **OK**. The document is saved to disk as specified. If **View Result** 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 applicable dialog box (Save to Files, Copy to Clipboard, Send in Mail or other). 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, to determine which formatting levels are available.

The formatting levels are:

Plain Text
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.

Formatted Text
Exports decolumnized text with font and paragraph styling, along with graphics and tables. This level is available for nearly all file types.

Flowing Page
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 then flows from one column to the other, which does not happen when text boxes are used.

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

Spreadsheet
Exports recognition results in tabular form, which is 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. Formatted Text is better if your document contains smaller tables: each table is placed on a separate worksheet with non-table parts placed in an index worksheet with hyperlinks to each relevant worksheet.

Select converter options

Click Options in a save 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, which determines which other options are presented. Select options as desired. Help details how to do this.

To make changes apply to all future exports done with the given converter, select the check box "Make changes permanent". If not selected, changes are applied to the current export only and are not saved for future use. Export settings can be changed and saved without a document save: click **Tools** > **Saving Preferences** .

Multiple converters

Use multiple converters to export to two or more file types in one export step. Select the Multiple option in the save dialog box.

To make your own multiple converter, open the **Saving Preferences** 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. Select the desired ones. Optionally specify subfolder 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. The following save operations cannot be done with multiple converters:

Saving OmniPage Documents
OmniPage workflows cannot be saved via multiple converters. Use the **File** menu or a workflow with a step **Save OPD**.

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 save 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.

Save to PDF

You have five choices when saving to Portable Document Format (PDF) files. The first four are presented as Text converters, and 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 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 Searchable Image
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

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

Besides the above types, you can use other parameters in defining your PDF output by clicking Options.

PDF 1.6 or 1.7
Save to PDF version 1.6 or 1.7 for enhanced security, markup and attachment embedding functionality.

PDF/A
Choose to create PDF/A compliant files to be confident that files display identically regardless of the computer environment and remain readable even after many years of technological evolution.

Tagged PDF
Create a tagged PDF file to preserve its structure. This will ensure logical reading order, correct table structure and more.

PDF MRC
Use this high compression technology for good quality and smaller file size; available for color and grayscale PDF Images or PDF Searchable Images.

Linearized PDF
Choose this to create PDF files optimized for fast loading and display when embedded in web pages.

Password protection

In Kofax OmniPage Ultimate you can set a type and level of encryption and then define an Open password and/or a Permissions password for PDF files.

A smaller range of choices is available for saving to XPS files.

Convert from PDF

To extract text content from a PDF file, load it into OmniPage, recognize it, and save the results to a text format.

A variety of outputs is also available from a PDF file shortcut menu: Word, Excel, RTF, WordPerfect, or text. For more options, use the Convert Now Wizard.

eDiscovery Assistant for searchable PDF

Access this Assistant from the Tools menu or from a PDF file's shortcut menu in Windows Explorer. The Assistant is specially designed to create searchable PDF files from image-only PDF files, or files that already contain some text elements or text pages; it does this without altering or applying an OCR process to existing text. In other words, it limits its processing to the image-only parts of the input PDF. All text-based elements in a PDF remain untouched, including document metadata, annotations, mark-up, stamps and more. The process can run automatically or with interaction for zoning or proofing. The Assistant loads files you select from your file system and returns the results to the same location; choose whether to have the original files overwritten or retained as backup copies. Zoning and proofing occur in pop-up windows, with no connection to any documents open in OmniPage at the time.

Kofax OmniPage Ultimate adds the ability to make PDF files searchable as a pre-programmed job in DocuDirect. This can be with a Normal job (starting immediately, at a fixed later time or with recurrence) or as a Folder watching job.

Creating PDF files from other applications

The Kofax Power PDF Create product supplied with Kofax OmniPage Ultimate provides the ability to create standard PDF files from documents in any print-capable application on your system. Click **File** > **Print** and select the printer **Power PDF** Adjust properties as desired, click **OK** and supply a file name and location. If "View resulting PDF" is selected, your default PDF viewer displays the result.

Send 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 email:

- With automatic processing, select **Send in Mail** as the setting in the **Export Results** 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. After export options are specified, an empty mail message appears with files attached; you can add recipients and message text as desired.
- With manual processing, select **Send in Mail** as the setting in the **Export Results** drop-down list and then click its button. The Export Options dialog box appears immediately and then the mail message with the attachments.
- Workflows and jobs accept a Send in Mail export step, but they require the recipients and message text to be specified as workflow settings, so the workflow can be run unattended.

Send to eBook readers

Kofax OmniPage Ultimate supports saving workflow results to two standard eBook formats: Kindle and ePub.

Send to Kindle

A Kindle reader is an electronic book product from Amazon. The Kindle Assistant in the Tools menu lets you create a simple workflow that sends recognition results optimized for Kindle reader display to a Kindle account at Amazon.

To prepare a Kindle workflow:

1. Have your Kindle reader and its associated email address on hand.
2. Choose **Kindle Assistant** in the Tools menu.
3. Type in a name for the new workflow.
4. Choose a document source: **Scan**, **Load Files**, or **Load Digital Camera Files**. With file input, you will be prompted to choose input files when the workflow starts running.
5. Enter the email address linked to your Kindle reader.
6. Provide a name for the output file. All recognition results enter a single file.
7. Choose **Save** to save the workflow for later use, or **Save and Run** to immediately run the workflow and transfer its results to your Kindle device.

This simple workflow has three steps: acquire images, perform OCR and send to Kindle. The recognition language can be selected. All other settings take either default values or values optimized for Kindle.

When you run the Kindle Assistant the first time, a customized output converter is created, called "Kindle Document". It converts colored items to grayscale, pictures to 72 dpi and sets Formatted Text to remove any columns. This converter is then available for later processing, with or without workflows.

You can modify the Kindle workflow using the Workflow Assistant, to add other steps and change settings. For instance, you can specify a page range or add more save steps, so the file is not only sent to Kindle, but also saved to a file with different settings (for instance with Flowing Page and color retention). Take care not to make modifications that are unsuitable for Kindle, such as creating multiple output files, setting non-supported languages, etc.

You can also compile workflows targeting Kindle with the Workflow Assistant; set a Send in Mail step, choose the Kindle output converter in its settings and enter the Kindle email address. You can do the same without using a workflow by choosing Send in Mail in the Export results drop-down list.

Send to ePub

ePub is a free, open-source electronic book standard that can be displayed on any of the widely popular devices capable of functioning as an ebook reader.

For better ePub results, the **Treat as book (ePub)** option should be selected under **Tools** > **Options** > **Process** > **Retain features** . This way processing steps are optimized for ePub output.

Three output file types are available:

- ePub: This retains as much formatting as possible and allows text to flow.
- ePub simple: This removes most formatting, but allows text to flow, so it can be resized by the mobile device. Many smart devices analyze incoming text and apply their own formatting.
- ePub for poems: This retains formatting but line breaks from the original are conserved.

Two ePub sample workflows are shipped with Kofax OmniPage Ultimate:

- ePub from PDF or Scanned Document: retains formatting
- ePub from PDF or Scanned poems: conserves line breaks

The simplest way to prepare an ePub workflow:

1. Choose a document source via **Workflow** button (1-2-3). With file input, you are prompted to choose input files when the workflow starts running (Load Files dialog).
2. Provide a name for the output file (Save to File dialog). All recognition results enter a single file.
3. Choose **Save** to save the workflow for later use, or **Save and Run** to immediately run the workflow.

Other export targets

Turn recognized text into an audio MP3 file for later listening, leveraging Windows Speech (Windows SAPI). You must specify the reading language in the Text-To-Speech settings. See *Converting Text to Speech* in the Help.

Kofax OmniPage Ultimate is delivered with Evernote cloud access. The cloud provider is available in the Export Results drop-down list.

In Kofax OmniPage Ultimate you can export files to other targets. You can save files to Microsoft SharePoint, to eDOCS (Open Text) or iManage (Interwoven). Exporting choices are made in the Export Options dialog box. When you click OK, you may be directed to log in and prompted to specify the required path.

When using SharePoint, the server, login and password information must be provided only once per session, and it is offered in each subsequent session.

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 Kofax OmniPage Ultimate, as defined by their extensions.

See Help for more information on these targets.

Chapter 7

Workflows

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 is possible, even to different targets, each with their own export settings.

Workflows are designed for efficient whole-document processing. They can also handle recognizing or saving single or selected pages from a document.

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, the ones when run-time prompting is requested for input or output file names and paths, or scanning workflows prompting for more pages.

DocuDirect jobs are closely related to workflows. Jobs are created in the Job Wizard which uses Workflow Assistant in the creation process. Jobs run workflows according to the job parameters (mostly timing instructions) and it is more typical for them to run unattended.

Click **Workflow Assistant** in the **Standard** toolbar to see the 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 a scanner, place your document in the ADF or the first page on the scanner bed.
2. Select the desired workflow from the Workflow drop-down list.
3. Press the Start (1-2-3) button. The OmniPage Toolbox displays the steps in the workflow and acts as a progress monitor. The Workflow Status panel shows progress in more detail. To stop the workflow before it completes, press the Stop button.
4. If run-time input selection is specified, the Load Files 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 or appropriate dialog box 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 unattended, remember to include a save 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, the Start Workflow dialog box displays where you can choose what to process from the current document: only the Workflow-defined pages, all pages, selected pages, or the current page.

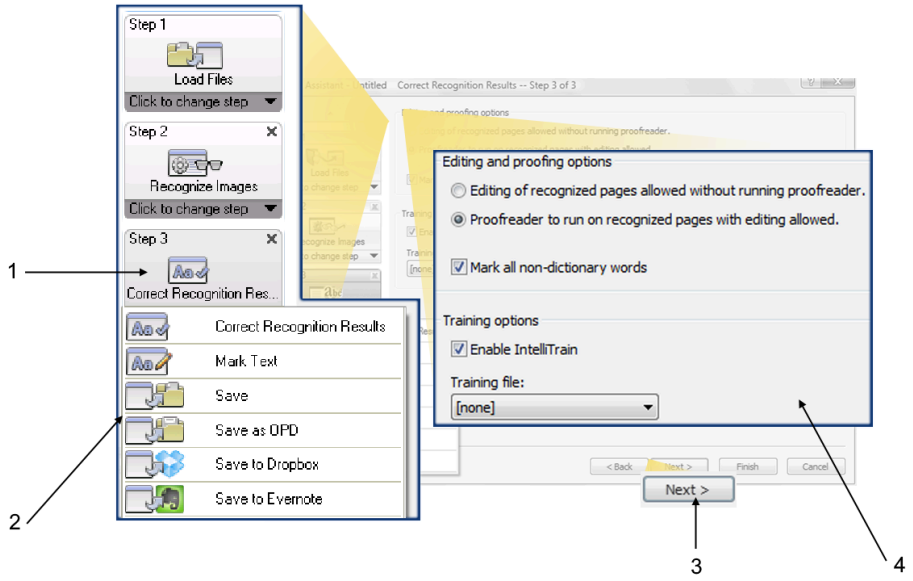
If you do not see the OmniPage Agent icon, enable it in the General panel of the Options dialog box or choose `C:\Program Files (x86)\Kofax\OmniPage19\OpAgent.exe`.

You can launch some workflows from your desktop, from Windows Explorer or the Easy Loader. Right-click an image file icon or file name for a shortcut menu. Multiple file selection is possible. Choose Kofax OmniPage Ultimate and a workflow name from the submenu, which also provides quick access to six target formats using default settings: Word, Excel, PDF, RTF, TXT and WordPerfect. To customize which workflows you would like to see here, click the Add and Remove Workflows menu item. 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 resume.

Workflow Assistant

Use the Workflow Assistant 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 one or more steps, each with a drop-down list. This left panel of the Workflow Assistant dialog box lets you build your workflow.



1. The step currently chosen. You can:
   - Click the close button (x) to delete this step together with all subsequent, dependent steps.
   - Click the triangle in the lower right to display the drop-down list with all available activities.
2. Drop box with the available activities
3. Click to add new step to the workflow
4. Specify settings for the current step here

At any moment in the process, the Assistant drop-down menu offers all steps that are logically possible at that point.

In Kofax OmniPage Ultimate, additional steps are available: Extract Form Data and Mark Text.

Create workflows

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

The opening Workflow Assistant panel offers two starting points:

- Choose **Fresh Start** to begin with no steps in the workflow diagram on the right. Accept or change the default workflow name. Then click **Next** and choose your first step. Choose an image loading step that can take input from file, scanner or digital camera files. Specify settings on the right. Then move on to build your workflow: it can include a variety of different steps. When done, click **Finish**.

- 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. Enter a name for your new workflow. Click **Next** to proceed; modify its steps and settings as described in the next section. The changed settings apply to the new workflow only and are not written back to the workflow used as the source. Any changed settings enter the new workflow, but do not affect the settings in the program. Finally, select **Finish** to complete your new workflow.

Modify workflows

Select the workflow you want to modify in the **Workflow** drop-down list and click **Workflow Assistant** in the **Standard** toolbar. Or click **Tools** > **Workflows** , 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 the downward pointing arrow under the icon to replace this step with another one. Continue modifying steps and/or settings as desired. Remember that deleting or modifying a step may result in later, dependent steps being removed. Click **Next** to replace removed steps or to add new ones. Click **Finish** to confirm the changes to your workflow.

After creating or modifying a workflow, you must either run a workflow or select the **1-2-3** item in the **Workflow** drop-down list, to return to normal processing.

Workflow to Kindle

The Kindle Assistant in the Tools menu helps you create a simple workflow that will accept input, perform OCR and send the results in a suitable format to a Kindle account at Amazon; it will then appear on the Kindle device registered to that account.

DocuDirect

DocuDirect 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 Kofax OmniPage Ultimate, you have the following additional DocuDirect capabilities:

- Setting job timing and recurrence
- Folder watching for incoming image files
- Email inbox watching for incoming attachments (Outlook and Lotus Notes)

- Email notification of job completion to specified recipients
- Driving workflows with barcodes.

Create new jobs

1. Do one of the following:
   - Select **Process** > **DocuDirect** from the menu.
   - Select **Start** > **Kofax OmniPage Ultimate** > **OmniPage DocuDirect** .
   - Start from the **OmniPage Agent** on the taskbar.
2. Creating a job is basically timing a workflow. To do this, start DocuDirect (as described above) and click the **Create Job** icon or click **File** > **Create Job** in the menu.
3. 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 email inbox for incoming images to be processed in OmniPage. A specific type within this category is **Barcode cover page job**, 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.

   Job types available in Kofax OmniPage Ultimate only:
   - **Barcode cover page job**: This is a special type of folder watching job (see below). It monitors a folder for incoming barcode pages, and then processes subsequently incoming images with the workflow identified by the barcode.
   - **Folder watching job**: Select this job type and browse to the folders to be watched for incoming image files.
   - **Outlook mailbox watching job**: This job watches an Outlook email inbox for incoming image attachments of a specified type.
   - **Lotus Notes mailbox watching job**: Same as above, but a Lotus Notes inbox is watched.
4. Name your job and click **Next**.

   The next panel shows Start and Stop Options. Specify **Start time** and **End time**, set whether input files are to be deleted or saved when the job is completed. If you have a job requiring user interaction, choose whether to allow it or not with the checkmark **Run job without any prompts**. This lets you run such jobs in two ways, avoiding the need to create two jobs. If you plan to be at the computer as the job runs, clear the check box. If you want to run the job without being present, select

the check box. Then only automatic image enhancement will run, auto-zoning will replace manual zoning and proofing is skipped.

In this case, you must ensure that the input and saving file sets and locations are pre-defined.

In Kofax OmniPage Ultimate you can set a recurrence pattern and request email notification when the job is completed.

From the next panel onwards, you can construct your job (except for barcode cover page jobs) as you normally do with Workflows. Set your starting point (Fresh Start or Existing Workflows) and proceed as described in the Create workflows topic.

The **Options** dialog box in **DocuDirect** 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 cover page 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.

5. Click **Finish** to confirm job creation.

Modify jobs

Jobs with an inactive status can be modified. Select the job in the left panel of DocuDirect and click **Edit** > **Modify** in the 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.

Managing and running jobs

This is done with DocuDirect. It has two panels. The left panel lists each job, its next run, status and history. The status is:

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 has 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 DocuDirect 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 click **Edit**, then choose **Clear Occurrence**. This removes files storing the reporting data from the current occurrence of the current job. **Clear All Occurrences** removes all data for all job occurrences of the selected job. These two options are useful to free disk space, but cleared occurrences cannot be viewed anymore, so use these with caution.

The Workflow viewer

The Workflow viewer, as displayed in the Workflow Status panel, is integrated into DocuDirect 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 page information at each stage, allowing you to quickly view any page. Job results are marked by icons. Drop-down lists give you information about processing steps.

Watched folders

In Kofax OmniPage Ultimate you can specify watched folders and email 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 email 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 email folders to be watched. You should not use these folders for other purposes, not even for barcode cover page 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.

Incoming files are removed from the watched folders as soon as they are transferred to OmniPage for processing; you should therefore arrange additional storage elsewhere if you want to retain the incoming files.

Add the desired folders and file types (one type or all types). Click the check box 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 folders 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 for approximately the interval specified in DocuDirect 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 Kofax OmniPage Ultimate, 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 email 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 Kofax OmniPage Ultimate you can run workflows (sets of steps and their 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 copying the barcode cover page image into a watched folder that will receive the document images to be processed.

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 the 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. Select **Barcode cover page workflow** as the Scanner button default action on the **Scanner** tab in **Options**. You can also set it to Prompt for workflow. In this case, a dialog box appears with the available choices: Scanning, Barcode cover page workflow, and all scanning workflows.

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

For image input, you must create a barcode cover page job.

A barcode cover page 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 cover page job processing you need to:

1. Create a workflow that contains the processing steps you need. Select **Load Files** as input with **Select files for loading each time this workflow is started** selected.
2. Save a barcode cover page that identifies the workflow.

3. Define timing instructions for barcode folder watching in DocuDirect by creating a barcode cover page job.

To process with a barcode cover page job:

1. Make sure that the job is running at the required time.
2. The folder is being monitored and the workflow is started as soon as a barcode cover page is placed in the specified watched folder.
3. The workflow processes image files arriving in the folder after the cover page.
4. The workflow is 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.

File-it Assistant

The File-it Assistant lets you create scanning workflows for repeated document conversion tasks. The Assistant is for scanning jobs that require no user interaction during the processing. In a typical scenario, operators at a scanning station prepare documents, applying the appropriate barcode cover page to each, without needing to know anything about the later processing or destination of the documents, because all that is pre-determined. Associate a button on your scanner with OmniPage and print a barcode cover page to identify your workflow. As a result, you can scan, convert and save without interaction beyond pressing the scanner button.

Create the workflow:

1. Click **Tools** > **File-it Assistant** in the menu.
2. Name your workflow, choose an output file type, location and file name.
3. Review and optionally change the workflow settings.
4. Print the barcode cover page.
5. Associate OmniPage with a scanner button (must be done only once) in the Control Panel.

Use the workflow:

1. Place the printed barcode cover page on top of a document in your scanner.
2. Push the OmniPage-associated scanner button. The document is converted using steps and settings from the referenced workflow and sent to the location you defined.

It is possible to use barcode cover pages stored as image files to drive jobs from watched folders. Such jobs permit interactive steps like manual zoning and proofing that are not available via the File-it Assistant.

Single-step PDF jobs

Two special workflow steps are available for use inside DocuDirect. Both relate to PDF file output and can usefully be combined with local folders for automatic processing; Make PDF Searchable can be used with watched folders as well.

Convert to PDF Job
This allows input from document files (typically MS Office files plus .txt, .csv) provided their native applications are installed; output is one PDF file for each input file with the same name as the input file. The saving location can be specified. Typically, the resulting PDF files are both searchable and editable. Kofax Power PDF Create must be present for this job type.

Make PDF Searchable
This accepts input from image-only PDF files or PDF files that may contain image-only areas or pages. It results in the original PDF files becoming searchable. Kofax Power PDF Create is not required for this job type.

Both these jobs allow only a single PDF step. If selected, the Next button is not available and Finish must be chosen when all settings are applied as desired.

Chapter 8

Technical information

This chapter provides troubleshooting and other technical information about using OmniPage. Please also read the Readme file and other help topics, or visit the Kofax 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, such as check connections, close other applications to free up memory, and so on.

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

Supported file formats are listed here, and the Help provides more detail.

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.
- Make sure that your scanner is plugged in and that all cable connections are secure.
- Visit the support pages on the Kofax website at www.kofax.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. Access the Kofax website through the Help menu and consult the scanner section for more information.
- Defragment your hard disk. See Windows online Help for more information.
- Uninstall and reinstall OmniPage.

Testing OmniPage

Restarting Windows in safe mode allows you to test OmniPage on a simplified system. This approach is recommended when you cannot resolve failure 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.
   - In Windows 7: Press F8 immediately when Windows starts to boot, just after you see the Starting Windows message.
   - In Windows 8 or 10: Press and hold the Shift key while clicking the Start menu, then Power, followed by Restart.
2. Launch OmniPage and try performing OCR on an image. Use a known image file, such as one of the supplied sample image files.
   - If OmniPage does not launch or run properly in safe mode, 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.

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 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.
- Make sure text zones are identified correctly. Re-identify zone types and contents, if necessary, and perform OCR on the document again.
- 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. In particular, setting an Asian language for non-Asian texts (and vice versa) is likely to produce unusable results.
- Recognition results in Japanese, Korean and Chinese can be viewed and saved only if your system has East Asian language support.
- 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.
- If you use True Page as the Text Editor formatting level 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.

System or performance problems during OCR

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

- 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.
- Restart Windows 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, XPS, and HD Photo.

Supported file types for saving recognition results as text are:

- ePub (*.epub)
- ePub for poems (*.epub)
- ePub simple (*.epub)
- HTML 3.2 (*.htm)
- HTML 4.0 (*.htm)
- InfoPath (*.xsn)
- Kindle Document (*.doc)
- Microsoft Excel (*.xlsx)
- Microsoft Excel XP, 2003 (*.xls)
- Microsoft PowerPoint (*.pptx)
- Microsoft PowerPoint 97 (*.rtf)
- Microsoft Publisher 98 (*.rtf)
- Microsoft Word 2000, XP (*.rtf)
- Microsoft Word 2003 (WordML) (*.xml)
- Microsoft Word (*.docx)

- MP3 Audio (*.mp3)
- PDF (*.pdf)
- PDF Edited (*.pdf)
- PDF Searchable Image (*.pdf)
- PDF with image substitutes (*.pdf)
- Text (*.txt)
- Text - Comma Separated (*.csv)
- Text - Formatted (*.txt)
- Text with line breaks (*.txt)
- Unicode Text (*.txt)
- Unicode Text - Comma Separated (*.csv)
- Unicode Text - Formatted (*.txt)
- Unicode Text with line breaks (*.txt)
- WordPad (*.rtf)
- WordPerfect 12, X3-X9 (*.wpd)
- XML (*.xml)
- XPS (*.xps)
- XPS Searchable Image (*.xps)

Third-party licenses and notices

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. Franklin Electronic Publishers, Inc.

The Proximity/Merriam-Webster American English Linguibases. © Copyright 1982, 1983, 1987, 1988 Merriam-Webster Inc. © Copyright 1982, 1983, 1987, 1988 Franklin Electronic Publishers, 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 Franklin Electronic Publishers, 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 Franklin Electronic Publishers, 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 Franklin Electronic Publishers, Inc. Words are checked against the 119,614 word Proximity/Van Dale Linguibase. The Proximity/Munksgaard Danish Linguibase. © Copyright 1988 Munksgaard International Publishers Ltd. © Copyright 1988 Franklin Electronic Publishers, Inc. Words are checked against the 113,000 word Proximity/Munksgaard Linguibase. The Proximity/IDE Norwegian and Swedish Linguibases. © Copyright 1988 IDE a.s. © Copyright 1988 Franklin Electronic Publishers, Inc. Words are checked against the 126,123 and 150,000 word Proximity/IDE Linguibases.

INSO/Vantage Research dictionaries: International CorrectSpell ™ spelling correction system © 1993 by Lernout & Hauspie.

Slovenian Speller Database, copyright © 2002 Ambeis d.o.o.

Esperanto dictionary based on compilation by Toon Witkam and Stefan MacGill.

Asian OCR capabilities are jointly developed by the Beijing Wintone Information Technology Corporation Ltd and Nuance Communications, Inc. All rights reserved.

International Components for Unicode (ICU) project Copyright © 1995-2009 International Business Machines Corporation and others.

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

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

Portions of this software are copyright © 2006 The FreeType Project <www.freetype.org>. All rights reserved. FreeType 2.3.1, Turner, Wilhelm, Lemberg.

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

This product was developed using Kakadu software.

Export Options dialog controls from Allan Nielsen, Supergrid control, copyright © 1999.

This product includes software developed by the OpenSSL project <http://www.openssl.org/> with software written by Eric Young and Tim Hudson.

Part of this software is derived from the RSA Data Security Inc. MD5 Message-Digest Algorithm.

AES encryption/decryption for PDF © 2001, Dr Brian Gladman, Worcester, UK.

Amazon's Kindle 2 copyright ©1999-2011.

Components for Asian font handling: copyright © 2009 Adobe Systems Incorporated. All rights reserved.

Some integration and other components: © 2009 Microsoft Corp. All rights reserved.

PDF creation: ©1993-2019 Zeon Corporation. All Rights Reserved

© Kofax, Inc., 2010-2019.

All rights reserved. Subject to change without prior notice.