I.R.I.S.
Document to Knowledge™
A CANON COMPANY

Readiris™ 17

17

No **retyping**. No **paper**.
Just **smart documents**.

for **Mac**®

#1 Conversion Software

Table of Contents

Readiris 17

Welcome to the Readiris documentation. The aim of this documentation is to help you get up to speed with using Readiris. It covers Readiris' installation and setup, the basic processing steps and the application's commands and working areas. At the end of the documentation you find a 'How to' section in which particular questions are addressed.

The procedures described in this documentation are based on the **Mac OS X High Sierra** operating system. The information in this documentation is subject to change without prior notice.

Introduction

Readiris is I.R.I.S.' flagship document recognition software. With Readiris you can scan paper documents, open image files, PDF files and digital camera images and turn them into fully editable text documents and text-searchable PDF documents you can save, share and send to online Storage Systems. Readiris supports a wide range of output formats, the most important ones being PDF, DOCX and XLSX.

The documents Readiris generates can either be saved on your computer, directly sent via e-mail or exported to online Storage Systems such as Evernote, Dropbox, Google Drive, Box and Microsoft OneDrive. With Readiris Corporate you can also send documents to Microsoft SharePoint.

Readiris recognizes documents in 130 languages. All American and European languages, including the Central-European, Baltic and Cyrillic languages as well as Greek and Turkish are supported. Readiris also recognizes Hebrew, as well as four Asian languages: Japanese, Simplified and Traditional Chinese and Korean.

Besides text recognition, Readiris offers powerful compression features for PDF files, thanks to I.R.I.S.' compression technology iHQC.

What's new in Readiris 17

Adding Annotations that can be exported into the PDF output format

An annotation associates an object such as a note, sound, or text with a location on a page of a document.

Readiris 17 offers an **Annotate** Option that allows you to add annotations in your documents before you export them into a PDF format. That function can be useful in a revision process.

Other Output formats than PDF do not support annotations. Those annotations are fully compatible with Adobe Reader. The best result cannot be guaranteed with any another PDF viewer.

PDF Signature

Before exporting your document as PDF, a digital signature can be included via the **Export Options / PDF.**

Legal Notices

Copyright

Trademarks

The Readiris logo and Readiris are trademarks of Image Recognition Integrated Systems S.A.
OCR and barcode technology by I.R.I.S.
AutoFormat and Linguistic technology by I.R.I.S.
BCR and field analysis technology by I.R.I.S.
iHQC compression technology by I.R.I.S.

All other products mentioned in this documentation are trademarks or registered trademarks of their respective owners.

Patent Statements

Readiris is protected under these issued patents: US8068684; EP2143039; JP5501958B2; US8331706; US8666185; US8995780; BE1022635; US8913836; BE1021013; US9058517; US9311558; BE1022630; US8897600; US9183636; BE1022636; US8811751; BE1022166; BE1022562; US9798943; BE1024194; US9836646; applied for by I.R.I.S. SA.

Third-Party Software Acknowledgements

Portions of this I.R.I.S. software uses the following copyrighted material, the use of which is hereby acknowledged.

Evernote

Evernote SDK Copyright © 2007-2012 by Evernote Corporation. All rights reserved.

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

Apache

Portions of this product include software developed by the Apache Software Foundation.

- Google APIs Client Library for Objective-C, Copyright 2011 Google Inc., licensed under the Apache License, Version 2.0
- Live SDK for iOS, Copyright 2014 Microsoft Corporation, licensed under the Apache License, Version 2.0
- JSONKit, Copyright 2011 John Engelhart, licensed under the Apache License, Version 2.0

Licensed under the Apache License, Version 2.0 (the 'License'); you may not use this file except in compliance with the License. You may obtain a copy of the License at: http://www.apache.org/licenses/LICENSE-2.0

JPEG

This software is based in part on the work of the Independent JPEG Group.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

- Neither the name of the Brandon Walkin nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE AUTHOR "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

TIFF library

Copyright © 1988-1997 Sam Leffler

Copyright © 1991-1997 Silicon Graphics, Inc.

Permission to use, copy, modify, distribute, and sell this software and its documentation for any purpose is hereby granted without fee, provided that (i) the above copyright notices and this permission notice appear in all copies of the software and related documentation, and (ii) the names of Sam Leffler and Silicon Graphics may not be used in any advertising or publicity relating to the software without the specific, prior written permission of Sam Leffler and Silicon Graphics.

THE SOFTWARE IS PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EXPRESS, IMPLIED OR OTHERWISE, INCLUDING WITHOUT LIMITATION, ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

IN NO EVENT SHALL SAM LEFFLER OR SILICON GRAPHICS BE LIABLE FOR ANY SPECIAL, INCIDENTAL, INDIRECT OR CONSEQUENTIAL DAMAGES OF ANY KIND, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER OR NOT ADVISED OF THE POSSIBILITY OF DAMAGE, AND ON ANY THEORY OF LIABILITY, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

PoDoFo

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (http://www.openssl.org/)

PoDoFo v 0.9.5

Published under the LGPL, version 2

The PoDoFo library is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.

How to obtain the source code

The software included in this product contains copyrighted software that is licensed under the Lesser Gnu Public License. This license is requiring us to provide the source code of that software. You may obtain the complete corresponding source code for such copy- righted software from us for an unlimited period of time and at no charge. In this case, please contact our technical support (www.irislink.com/support).

Lame

LAME is a high quality MPEG Audio Layer III (MP3) encoder.

This code is distributed under the GNU LIBRARY GENERAL PUBLIC LICENSE (LGPL), version 2.

Disclaimer

The content of this document was created and arranged with the utmost care.

I.R.I.S. SA assumes no responsibility for the accuracy, completeness and timeliness of the provided content.

Section 1: Installation and Activation

System requirements

This is the **minimal system configuration** required to use Readiris:

- A Mac computer with a 64 bits Intel processor.
- The operating system Mac OS X 10.11 or higher. Earlier versions of the Mac OS operating system are not supported.
- 400 MB of free hard disk space.

Installation and Activation

Installation

- Download the Readiris package from www.irislink.com/softwaredownload.
- Run the Readiris installation package and follow the on-screen instructions.
- Agree with the terms of the license agreement.
- Then click **Install** to start the actual installation.

    You might be asked to enter an Administrator Username and Password to install Readiris.

Activation

- You are prompted to **activate** Readiris.
- The **activation code** consists of 34 characters. Depending on how you purchased Readiris, the code can be found:
    - On a license sheet called "License sheet for SN for IRIScan..." located inside the product box
    - In your confirmation email after electronic purchase
- Enter the **activation key**, and then click **Activate**.

    Note that an internet connection is required to activate. If you don't have an internet connection, see Manual Activation below.

    You can also choose to start the trial if you do not want to activate Readiris at this time. Fill in the form and click Start trial. Next time you start Readiris you will be prompted again to activate the software.

Manual Activation

If you don't have internet access or when an error occurs during the activation process, an error message is displayed allowing you to do a manual activation.

- To do so, click **Manual Activation** in the error message.

- Click **Save as text file** to save all the info I.R.I.S. needs to continue with the Manual Activation. It is recommended to keep the default file name and default storage location.
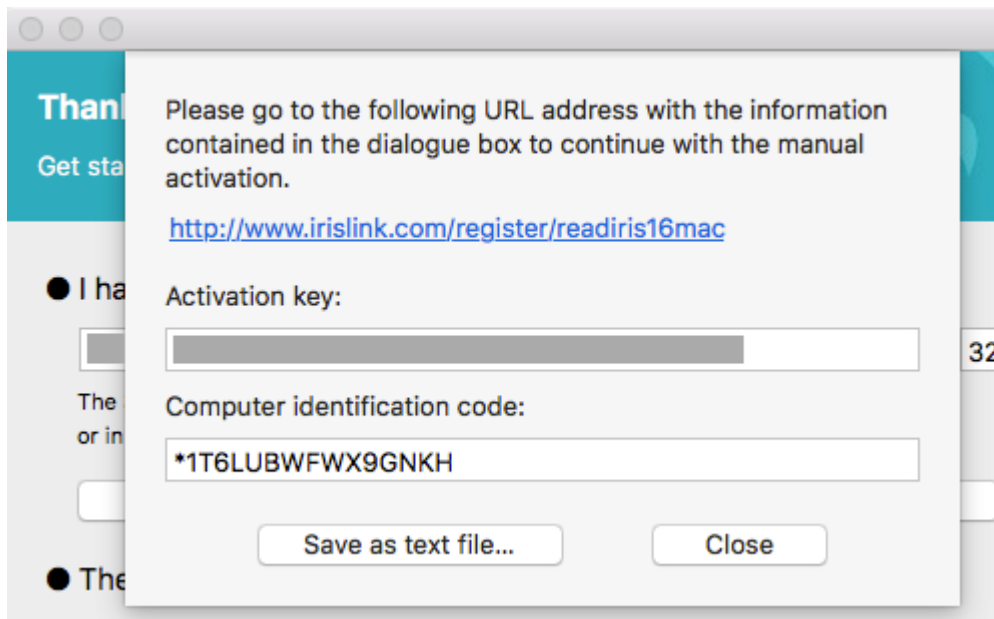
- On another computer on which you do have internet access: open the Web page indicated in the dialog box: http://www.irislink.com/register/readiris16mac
- On the Web page click the **Submit a ticket** link.

8

- Fill in all required fields and click **Add File** at the bottom of the page.

- Select the Text file you saved.
  The default file name is **Readiris Activation.text**. The default storage location is the Desktop.

Attach Files [➕ Add File]

- Click **Submit** to submit the ticket.
  You will receive the license file by email within 24 hours.

- Return to Readiris and click **Import a license file** in the **Activation** dialog box.

Thank you for choosing Readiris 17

Get started by choosing the option below

IRIS
Document to Knowledge
A CANON COMPANY

● I have an activation key:

The 34 digit activation key is located inside the product box
or in the confirmation email sent while purchasing the product.

| Activate | Import a license file... | Get a license... |
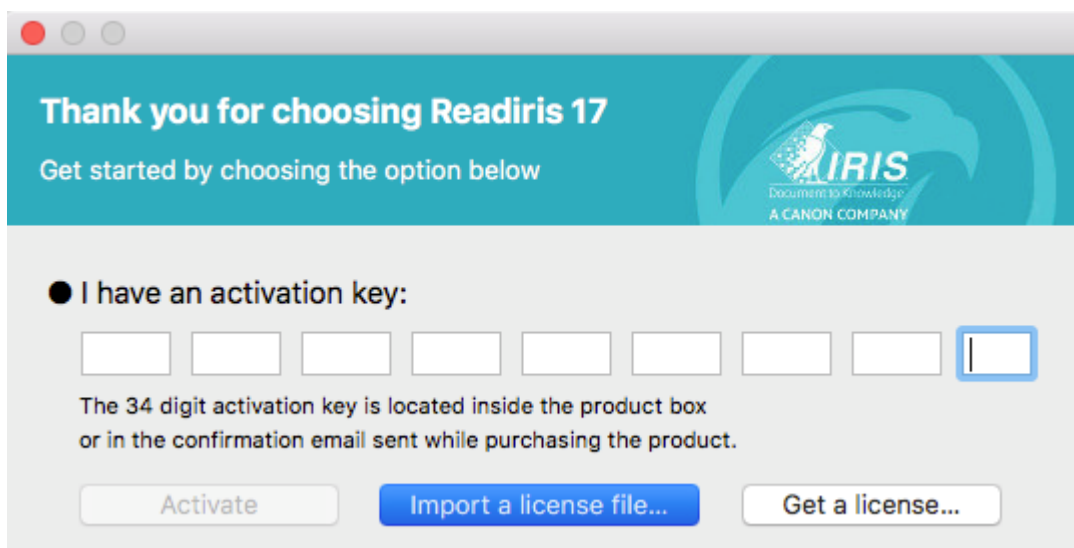
- When the installation is finished, click **Close**.

Software Registration

Registering is required to get **Technical Support**. It also offers other benefits, such as free updates, free trial downloads, video tutorials, discount on new products, and so on.

To register Readiris:

- On the **Readiris** menu, click **Register** to be directed to the Registration page.
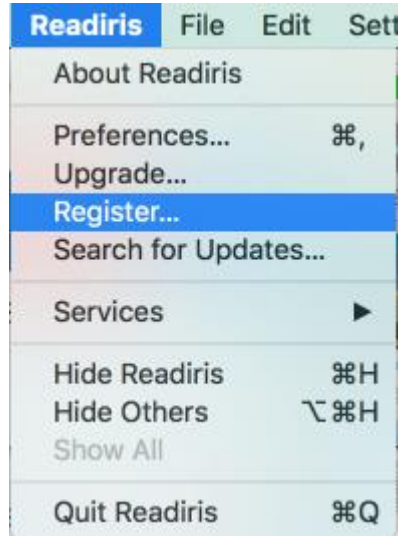
- Fill in your data and click **Submit**.

Note that an internet connection is required to complete the registration.

Search for updates

When you register your copy of Readiris you are entitled to free software updates.

To check if updates are available:

- On the **Readiris** menu, click **Search for Updates**.

Note: Readiris automatically checks for updates every 20 days.
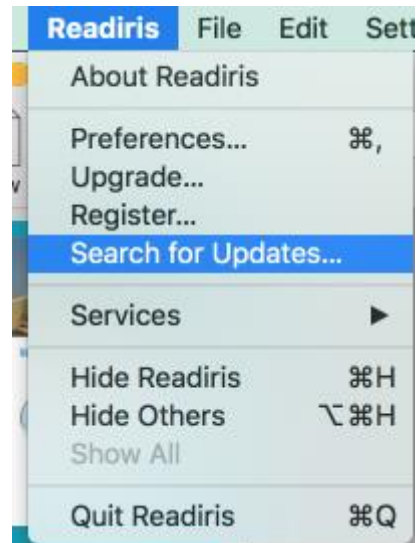
Uninstalling Readiris

To remove Readiris from your computer:

- Click **Finder** and open the **Applications** folder.

- Drag **Readiris** to the **Trash**.

    Readiris will be removed from your machine.

Note: the Readiris **preferences** are *not* removed by dragging **Readiris** to the trash can, in case you should want to re-install the software later on.

To remove the preferences:

- In the Finder's **Go** menu, press and hold **Alt**.

- Click **Library** > **Library** > **Application Support.**

- Drag the **Readiris** folder to the trash.

Technical Support

If you register your copy of Readiris you are entitled to free technical support. In the **Technical Support** section of our website www.irislink.com/support you can find F.A.Q.'s (frequently asked questions), Video Tutorials, and much more information to solve common issues.

If you still cannot find an answer to your question, click **Technical Support Contact** and fill in the **Web Contact Form**.

Section 2: Getting Started

Exploring the interface

The Readiris interface looks as follows (default mode):
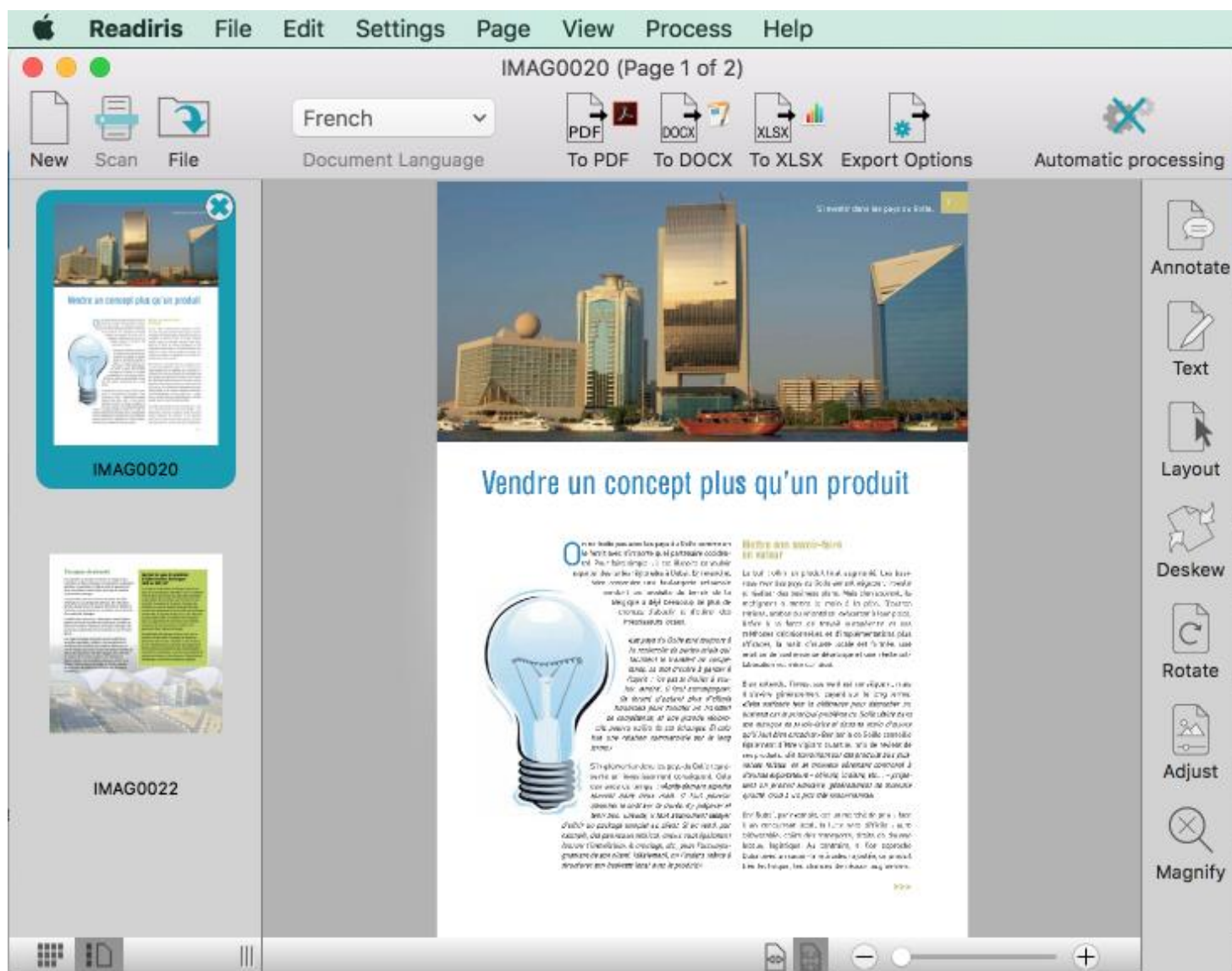
The interface consists of the following components and panels:

1. The Readiris menu bar
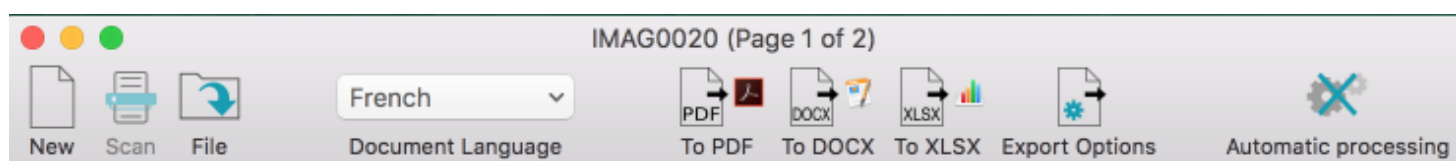
The Readiris menu bar regroups all the commands and options available in Readiris, including advanced settings.

2. The top toolbar

The top toolbar is by far the most important tool in Readiris. It contains all the basic commands you need to scan and export your documents.

Three default **Export** buttons are available to export your documents as **PDF**, **DOCX** and **XLSX**.

To change the default options of an **Export** button, click and hold it until its **Options** menu appears. Then move the cursor to the **Options** menu to open it.
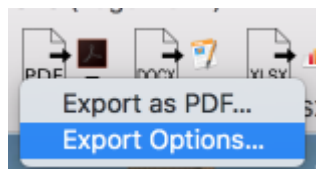
If the default Export buttons don't suit your needs and you want to select a different output format, click **Export Options**. Then drag the required output format to the **Favorites** list and configure its options. For more information see Select the Output Format**.**

Note that you can also customize the top toolbar by adding or removing commands. To do so:

- **Ctrl**-click the toolbar and click **Customize Toolbar**.
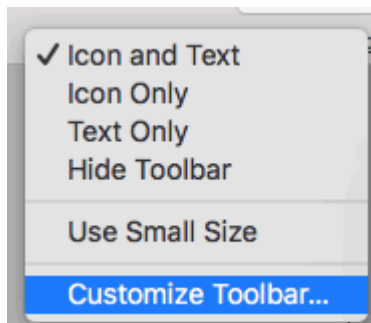
- Drag-and-drop the commands you want to move from or to the toolbar.

3. The Scanned image

When a page is scanned or opened into Readiris, its image is analyzed and displayed. The different recognition zones are not shown by default.

- When clicking on the image, the zone located at the mouse click position is shown.
- When clicking and dragging to make a selection rectangle, all the zones inside the selection rectangle are shown.
- When **Edit / Select All Zones** (Cmd-A) is selected, all the zones are shown

The content of the selected zone(s) can be copied to the clipboard or to the Evernote.

All the zones are shown and can be edited with the tools in the Layout mode.

4. The Pages panel

Each page you scan or open in Readiris has its thumbnail displayed in the Pages panel. The pages are processed top-down. To change the order of the pages, you drag a page from one location to another.

To have a better overview of the page thumbnails you can switch to Thumbnail view. In this view only the page thumbnails are displayed. The analyzed image is not displayed anymore. In Thumbnail view the thumbnails run from left to right, and from top to bottom.

To switch to Thumbnail view, click the Thumbnail view icon at the bottom of the Pages panel.

5. The Image and Layout Editing toolbar

The **Image and Layout Editing toolbar** allows to modify the current page: you can add **Annotations**, open the **Text Corrector**, change the **Layout** of the page, **deskew** pages that were scanned at an angle, **rotate** them, **adjust** the image quality, and so on. Point to the different buttons to display their tooltips.

Also see the sections Optimizing the Scanned Documents and Editing the Recognized Documents for more information.
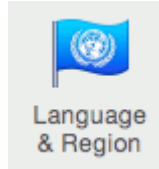
Changing the user interface language

Readiris opens in the user interface language that is currently activated in your Mac System Preferences.

To change the user interface language in Readiris:

- Click the **System Preferences** icon.

- Click the **Language & Region** icon.

- Drag the language of your choice to the top of the list.

    The user interface of Readiris is available in the following languages: English, French, German, Italian, Spanish, Dutch, Russian, Japanese, Traditional Chinese, Simplified Chinese, Korean, Danish, Finnish, Norwegian, Portuguese, Brazilian Portuguese, Swedish and Turkish.

- Close Readiris and restart **the computer** to apply the new language settings.

Configuring your scanner in Readiris

Readiris supports all **Twain 1.9** compliant scanners, all **Image Capture** compliant scanners, and **IRIScan 3** and higher scanners.

Most Image Capture scanners are plug and play. Before you can use a Twain scanner, however, its drivers need to be installed on your Mac.

Twain scanners

- Connect your scanner to your Mac and power it on.

- Install your scanner's Twain driver.

  In general, drivers can be found on the CD-ROM provided with your scanner and on the website of your scanner manufacturer. They are not provided by I.R.I.S. Note that some scanner drivers may not work under the latest versions of Mac OS. See the documentation supplied with your scanner to find out which platforms are supported. If necessary, contact your scanner manufacturer.

IRIScan Express scanners

- Connect your IRIScan Express scanner to your Mac and power it on.

- If Readiris is part of your IRIScan product, the driver has been installed during the installation of Readiris.

If that is not the case, visit our website www.irislink.com/support to download and install the required driver.

IRISCard Anywhere / IRIScan Anywhere / IRIScan Book scanners

You do not need to install any drivers when using these scanners. They are not Twain driver-based scanners that can be used from within any application to scan documents. Instead, they are designed to be used on their own, disconnected from any computer. The documents you scan are stored in the scanner's internal memory (or optional SD card / USB flash drive). From there, they can be loaded directly into Readiris when you connect the scanner or insert the SD card or USB flash drive into your Mac.

To check if the scanner driver has been installed successfully:

- In the **Readiris** menu, click **Preferences**.

- Click the **Scanner** tab.

- You should see your scanner listed*.

  *This is not the case for the IRISCard Anywhere, IRIScan Anywhere and IRIScan Book scanners.
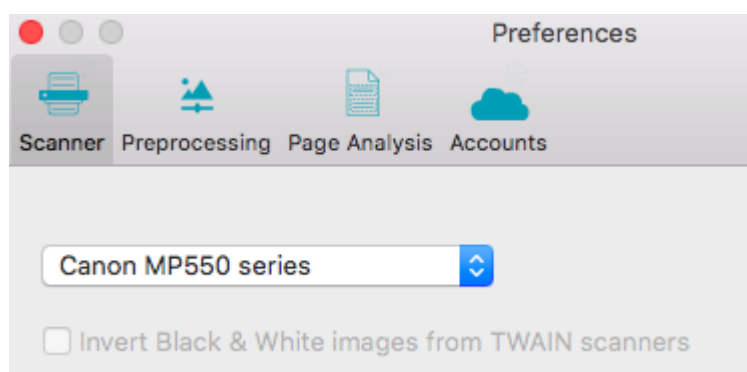
  If your scanner is not listed, its driver has not been installed successfully.

- Once the scanner has been configured you can start scanning, using the default scanner settings.

Section 3: Basic Document Processing

Processing documents with Readiris is fairly straightforward when using a basic configuration. In this section we discuss which steps to take when doing such a basic configuration. For detailed information about all the configuration steps, see sections 4 to 10.

Step 1: Scan, Open or Drag-and-Drop your images

In order to scan paper documents your scanner must be correctly configured. See the section Configuring your scanner if necessary.

- Click **Scan** to scan paper documents with your scanner.

  A preview window opens and displays your scanner settings. With TWAIN scanners, these settings differ from one scanner to another.

    - Select **Color** and **300 DPI** as preferences to obtain optimal results.

    - Select the **Page Analysis Options** of your choice:

      Page Analysis is selected by default. This option divides your documents into recognition zones. Without recognition zones, Readiris cannot process your documents. If you choose to disable this function, you will need to draw recognition zones manually on your documents.

      Readiris can **deskew** (straighten) documents and **rotate** them by means of **Page orientation detection**.

      When your scanner generates black borders around your documents, or when your documents contain information that is partially cut from the page, Readiris tends to create recognition zones on the borders of the document, even though they do not contain any useful content. In such cases it is recommended to use the option **Ignore border zones**.

or

- Click **File** to open existing image files.

    - Select the **Page Analysis Options** of your choice.

      The same **Page Analysis Options** are available here.

Or drag and drop PDF and image files to Readiris. You can do so in several ways:

- You can drag and drop files to the Readiris icon in the Dock.
  The files will be added **after** the images that are already loaded into Readiris.

- You can drag and drop files into the Readiris interface that is already opened.

  > 💡**Tip**: when you drag files to the Pages panel, you can choose where they will be inserted.

- You can drag and drop image files from the **Photos** app into the Readiris interface that is already opened.

> **Note**: it is not supported to drag images files from the Photos app to the Readiris icon in the Dock.

The recognition zones are visible in the **Layout** mode, available from the
The Image and Layout Editing toolbar.

Step 2: Modifying the pages or zones

The images you scan and load into Readiris can be modified easily:

- To change the order of the pages, drag the thumbnails to a different position in the Pages panel.
- To delete a page, select a page and click the delete icon.

If OCR errors need to be corrected, use the Text Corrector. See the section Editing the Recognized Documents for detailed information.
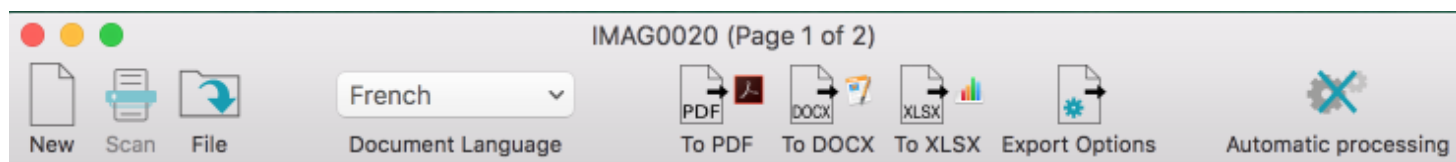
Step 3: Select the Recognition Language

Selecting the right document language is essential to obtain good recognition results with Readiris.
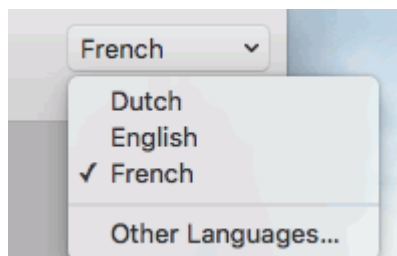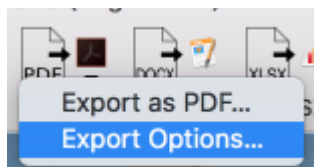

Step 4: Configure the Output

To export your documents directly in one of the default output formats and destinations, click one of the default **Export** buttons.
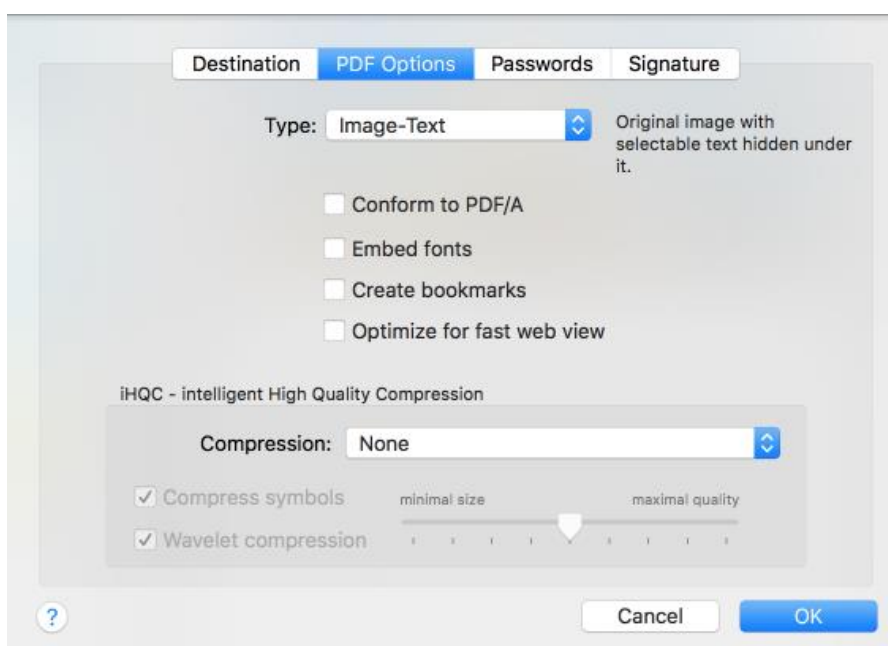

To change the default settings of an Export button:

- Click and hold an **Export** button until its **Options** menu appears.

- Now you can change the formatting and destination options.

To select a different output format than the 3 default ones:

- Click **Export Options**.

- Select the format of your choice and drag it to the **Favorites** list.

- Configure its formatting options.
  For detailed information about the different output formats and the different formatting options, see the sections Select the Output Format.

- Configure its destination.
  You can save documents as file, open them directly in an application or send them to a configured online Storage System.

> **Note**: the connection to the online Storage System must be configured correctly to be able to export. See the section Configure an online Storage System to learn how to do so.

- When you select **Open in application**, a default application is suggested for the selected output format.
  When you select PDF for instance, the output file will be opened in the Preview application. DOCX and RTF documents will be opened in the installed text editing applications. XLSX documents will be opened in the installed worksheet editor.

- To change the application, click it and then click **Select Application**.
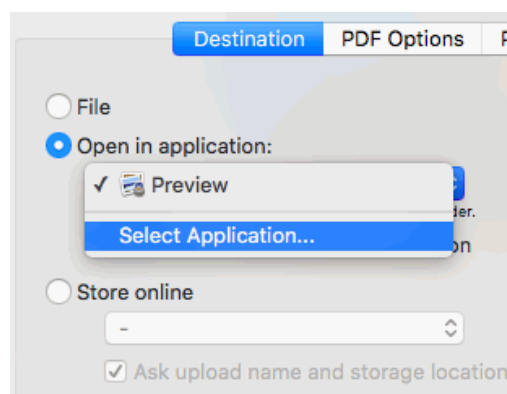
- When you are done, click **OK**.

When you have gone through these steps, click the **Export** button of your choice to start the text recognition and generate output documents.

Undoing an operation

Most of the operations you do or settings you configure can be undone with the **Undo** option.

The number of Undo operations is unlimited. Go to **Edit** menu and click **Undo** (the last action is shown between brackets). You can also use the keyboard shortcut: command-Z.
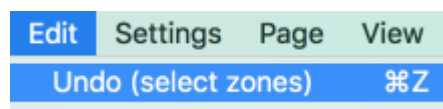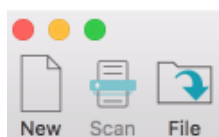

Starting a new project

When you are done saving your scanned documents, or when you want to start over:

- Click **New** on the top toolbar.

  This command erases your documents from Readiris so you can start a new project.

Automatic Processing

If you want to speed up the processing in Readiris or need to process batches of scanned documents, you can use the **Automatic Processing** feature. This feature automatically processes the documents you scan and open in Readiris, converts them into the output format you defined, and exports them either to a specific folder you defined or to the configured online Storage System.

To start automatic processing:

- Click **Automatic processing** on the top toolbar.

- Select **Automatically process all imported images**.

- Select the required Output format.

- Select an **Output folder or account** to which the processed documents must be exported.
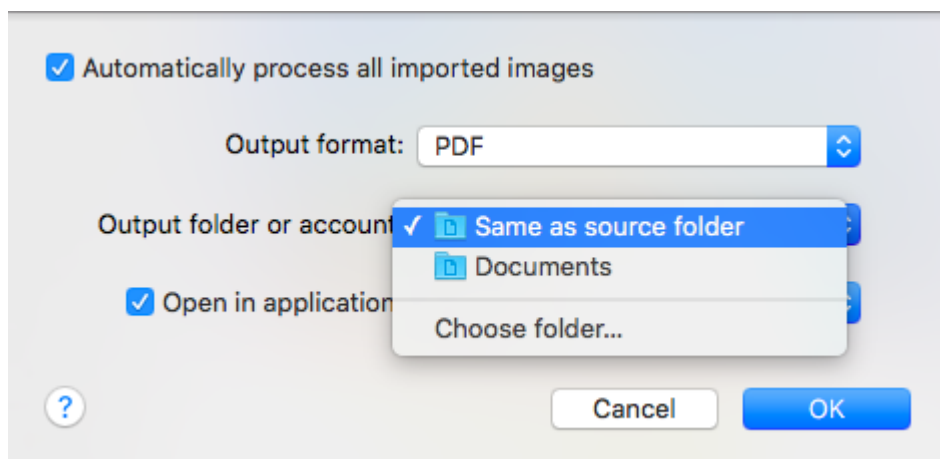
- Select in which **application** the processed files must be opened.

- Then start scanning or opening documents in Readiris. Or drag series of image files to Readiris.

> **Important notes:**
> When you scan multiple documents one after the other - on a scanner equipped with a document feeder for instance - all the documents are bundled into a single output file. The same goes for opening multiple image files at a time.
> When you scan documents on a flatbed scanner, each page will be saved in a separate output file.

💡**Tip:** Besides Automatic Processing, **Readiris Corporate** also offers Batch Processing and Watched Folder features to process larger batches.

Saving and loading settings

Note that Readiris saves the processing settings automatically when closing the application. Next time you start Readiris it applies the same settings.

In case you want to save a specific settings profile:

- First, select the required settings.

- On the **Settings** menu, click **Save Settings**.

- Name the settings, select an location and click **Save**.

The following settings will be saved: primary and secondary languages, favor recognition accuracy over speed, output format and any selected output format options, including PDF passwords, target application, online storage system settings, page sizes, page separation and indexing settings, page analysis, despeckling and deskewing options and training options.

Loading Settings

To load a settings file you saved:

- On the **Settings** menu, click **Open Settings**.

- Browse for the settings file you saved.

Restoring Factory Settings

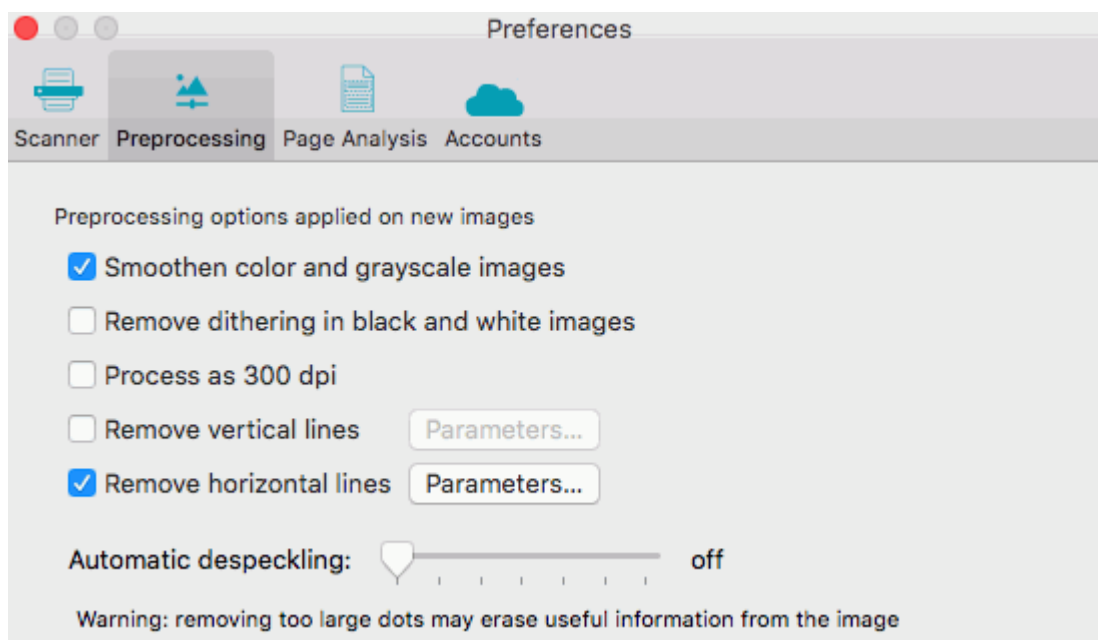- On the **Settings** menu, click **Restore Factory Settings**.

Section 4: Scanning and Opening Documents

Preprocessing Options

When scanning and opening documents, Readiris applies several **Preprocessing** options.

To access these options:

- On the **Readiris** menu, click **Preferences**.
- Click the **Preprocessing** tab.

The options you set here are applied as default to all the documents you'll open or scan. Note that changing these options does not affect the images that are already open in Readiris.

Note that you can also access these options via the **Settings** menu.

Smoothen color and grayscale images

This option is activated by default. It renders grayscale and color images more homogeneous by smoothening out differences in intensity. As a result, a stronger contrast is created between the foreground (text) and background (artwork). Sometimes smoothening is the only way to separate text from a colored background.

Remove dithering in black and white images

This option removes unwanted dither (noise) from black-and-white images.

Process as 300 dpi

Select the option **Process as 300 dpi** when you are processing images of an incorrect or unknown resolution. The images will be processed as if they had a 300 dpi resolution.

Remove vertical lines

To remove vertical lines:

- Select **Remove vertical lines**.
- Enter values for the 3 parameters:
    - **Minimum length**: minimum length (in pixels) of the lines to remove.
    - **Maximum thickness**: maximum thickness (in pixels)of the lines to remove.
    - **Maximum gap**: maximum line gap (in pixels) to close, in case of broken, interrupted lines.

Remove horizontal lines

To remove horizontal lines:

- Select **Remove horizontal lines**.
- The same parameters are available.

Note that lines are only removed to improve the recognition. They are not removed from the output files.

Automatic despeckling

Sometimes documents contain a lot of "noise" - little black spots that occur when scanning low-quality documents or using bad scanner settings.

To remove these little black spots, use the Despeckle slider. The more you move it to the right, the larger the spots that will be removed.
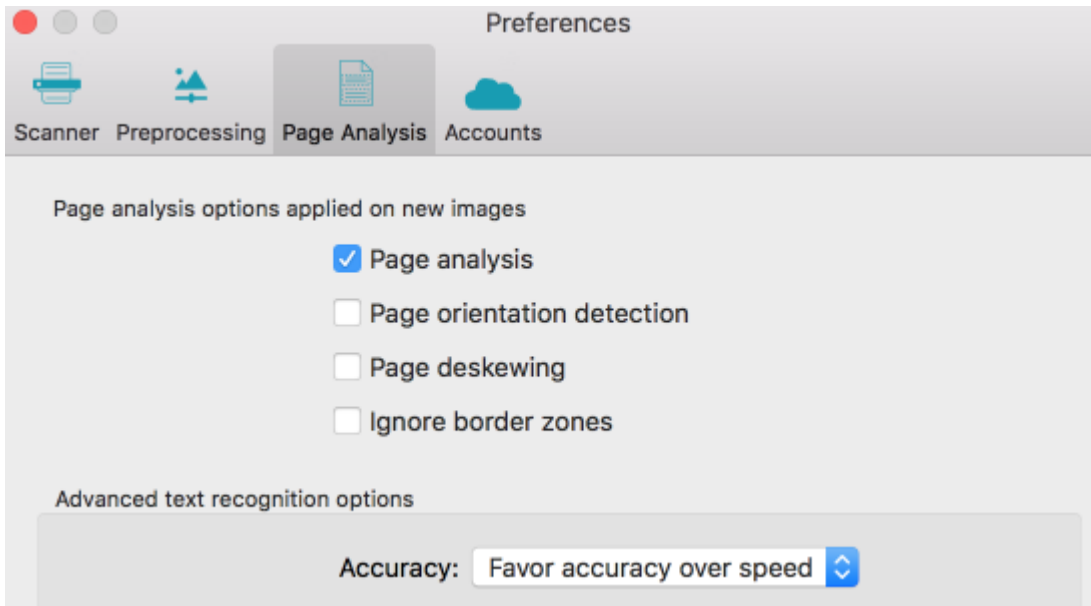
When you are done, close the **Preferences** window. The settings are saved automatically.

Page Analysis Options

When scanning and opening documents, Readiris applies several **Page Analysis** options.

To access these options:

- On the **Readiris** menu, click **Preferences**.
- Click the **Page Analysis** tab.

The options you set here are applied as default to all the documents you'll open or scan. Note that changing these options does not affect the images that are already open in Readiris.

Note that you can also access these options via the **Settings** menu.

Page Analysis Options applied on new images

Page Analysis

Page Analysis is selected by default. This option divides your documents into recognition zones. Without recognition zones, Readiris cannot process your documents. If you choose to disable this function, you will need to draw recognition zones manually on your documents.

Page Orientation Detection

This option detects in which direction your documents are oriented and rotates them automatically.
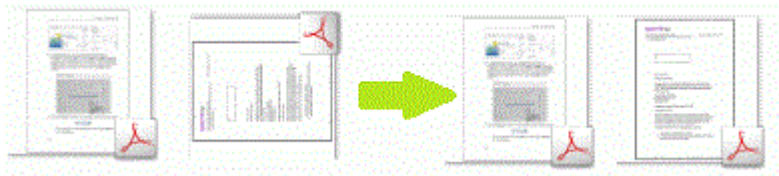
Page Deskewing

This option staightens - or deskews - pages that were scanned at an angle.

Ignore Border Zones

When your scanner generates black borders around your documents, or when your documents contain information that is partially cut from the page, Readiris tends to create recognition zones on the borders of the document, even though they do not contain any useful content. In such cases it is recommended to use the option **Ignore Border Zones**.

Advanced Text Recognition Options

Accuracy *(Readiris Corporate only)*

In Readiris you can choose whether to favor recognition **speed** over **accuracy** and the other way around.

- Click the **Accuracy** list and then select the required option.
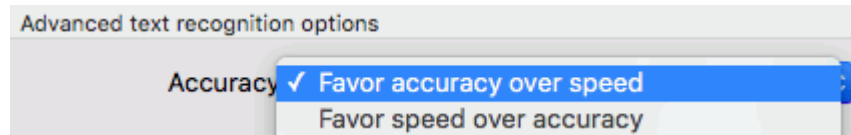
When you are done, close the **Preferences** window. The settings are saved automatically.

Scanning Documents

Readiris is Image Capture and Twain compliant. It supports almost any flatbed and sheet-fed scanner, all-in-one device, digital camera and scanner currently available.

Before you begin: make sure your scanner is connected to your computer and configured correctly. To do so, go to the **Readiris** menu and click **Preferences**. If you see your scanner listed on the **Scanner** tab it is correctly installed. If you don't see it, refer to the topic Configuring your scanner.

To scan documents:

- Click **Scan**.

- A preview window opens, displaying the settings available for your scanner.

  Note that with TWAIN scanners these settings differ from one scanner to another.

- Select **Color** and **300 DPI** as preferences to obtain optimal results.

  When you are scanning business cards or Asian languages it is recommended to use a scanning resolution of 400 dpi.


- Then select the **Page Analysis options**.

  See Page Analysis Options for more information.

- Click the **Scan** button in the dialog box to start scanning documents.

Notes:

- Pay attention to line skew. Line skew over 0.5° increases the risk of OCR errors.

- When an image has been analyzed in Readiris and its recognition accuracy is not sufficient, a warning sign is displayed both in the Page thumbnail and below the top toolbar.
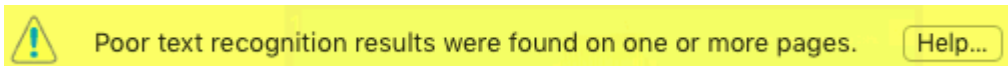
Tips:

- Sometimes Twain scanners display white text on a black background when scanning in black-and-white. To avoid this, click **Preferences** on the **Readiris** menu. Then click the **Scanner** tab and select **Invert Black & White images from TWAIN scanners**.

- To scan batches of documents, see the sections Automatic Processing and Watched Folder.

- Readiris can be set as target application of your scanner, if your scanner supports that feature. That way you scan documents on your scanner and transfer them to Readiris, just by clicking your scanner buttons. See your scanner documentation for more information.

Scanning with I.R.I.S. scanners

### *IRISCard / IRIScan Anywhere / IRIScan Book*

Preliminary note

The IRISCard and IRIScan Anywhere, and IRIScan Book scanners are designed to be used on their own, disconnected from any computer. The documents you scan are stored in the scanner's internal memory (or optional SD card / USB flash drive).

From there they can be loaded into Readiris.

So, note that these scanners are not Twain-driver scanners that can be used from within any application to scan documents. Instead, think of them as USB storage devices that are able to scan independently.

To load scanned documents into Readiris:

- Connect your scanner to your computer with the provided USB cable and power it on.

- OR insert the SD card or USB flash drive you used in the scanner into your Mac.

- Readiris detects if card images were scanned to the storage medium. Click **Yes** to select them. Readiris then opens the right folder.
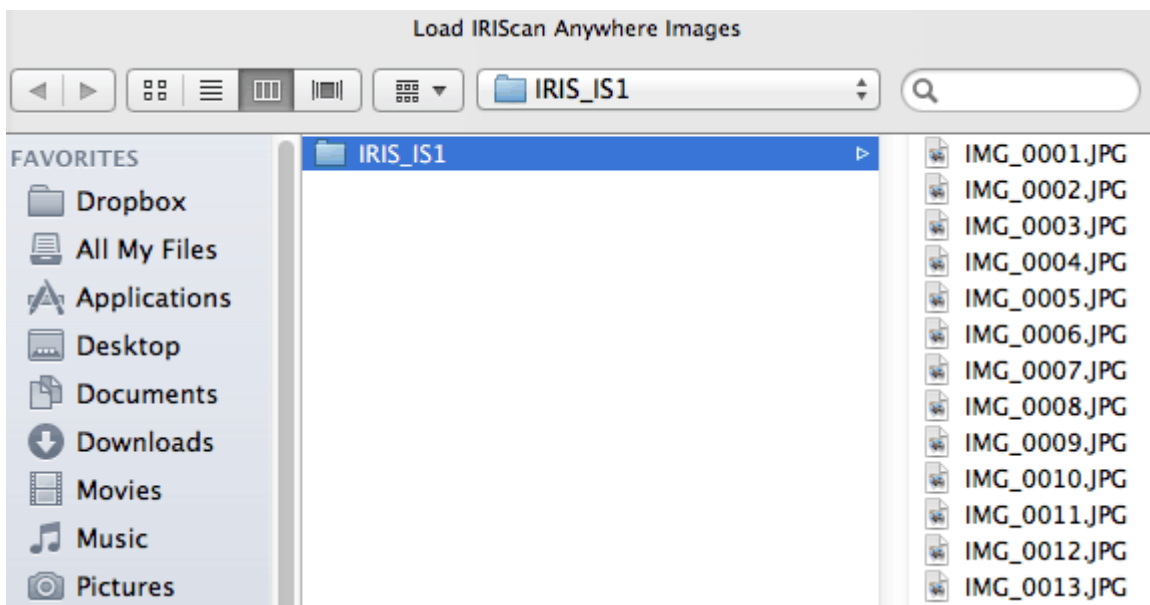
Opening Image Files

Besides scanning documents, you can also open image files and PDF files with Readiris.
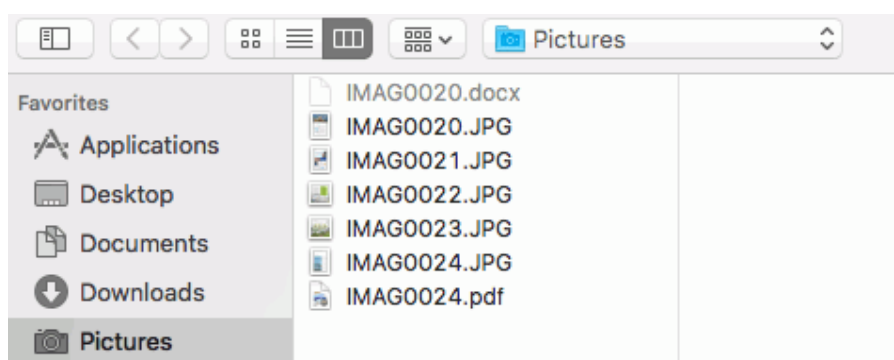
Readiris supports the following image file formats: GIF images, JPEG images, JPEG2000 images, Photoshop images, PNG images, (uncompressed, packbits and Group 3 compressed) TIFF images, multipage TIFF images, Windows bitmaps (BMP) and PDF documents.

To open existing image files:

- Click the **File** button.
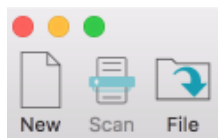
- Then select the images files you want to process, and click **Open**.

You can also open multiple images files at a time:

- Select the first image file and hold down the **Cmd**-key as you select additional images or;
- Select a continuous range of image files by clicking the first image and holding down the **Shift** key as you select the last image.

Should you want to abort the loading process, click the x icon in the top right corner of the progress bar.

Tip: when loading multipage image files (TIFF images) and PDF documents, you can define the page range (in case you only need a certain chapter of a document for instance)

Notes:

- When you click the **File** button on the top toolbar **after** you saved your current document, you are prompted whether you want to delete the current document. Click **No** to add image files to the recognized document or click **Yes** to start a new document.

- When you try to open a password-protected PDF document you will be asked to enter the password.

- When an image has been analyzed in Readiris and its recognition accuracy is not sufficient, a warning sign is displayed both in the Page thumbnail and below the top toolbar.
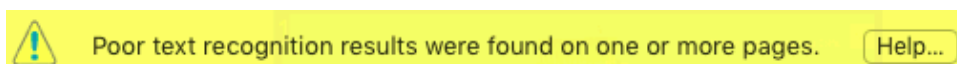
Other ways of opening image files

- You can drag and drop files to the Readiris icon in the Dock.
  The files will be added **after** the images that are already loaded into Readiris.

- You can drag and drop files into the Readiris interface that is already opened.

  💡 **Tip**: when you drag files to the Pages panel, you can choose where they will be inserted.

  > **Note**: when you have split documents, you cannot insert pages before the first page of Document 1. See How to separate documents for more information.

- You can drag and drop image files from the **Photos** app into the Readiris interface that is already opened.

  > **Note**: it is not supported to drag images files from the Photos app to the Readiris icon in the Dock.

- **Ctrl**-click any image file you want to open, point to **Open With** and click **Readiris**. The Readiris software will open and display the image.

- **Ctrl**-click any image file you want to open, point to **Services** and select the conversion option of your choice: **Convert to XLSX**, **Convert to PDF** or **Convert to DOCX**.
  Or select any image file you want to open, click the **Finder** menu > **Services** and select the conversion option of your choice.

  This option converts your files without opening the Readiris interface, but uses the settings that have been configured in the interface.

  For more information see Converting Image Files from Finder.

- Readiris can be set as target application of your scanner, if your scanner supports that feature. That way you scan documents on your scanner and transfer them to Readiris, just by clicking your scanner buttons. See your scanner documentation for more information.
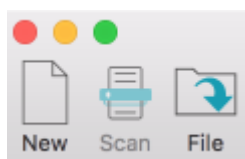
💡 **Tip**: to process batches of images files, see the topics Automatic Processing and Batch Processing.

Opening Digital Camera Images

Readiris automatically detects when images have been made by a digital camera and uses special recognition algorithms to process them.

To open digital camera images:
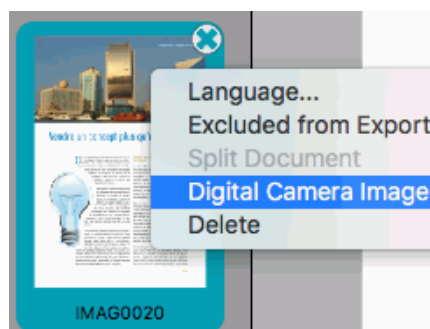
- Click **File** on the top toolbar.

- Select your digital camera images and click **Open**.

- In the page thumbnail, you see a little camera icon, which indicates Readiris recognizes the image as digital camera image.

If Readiris does not detect your image is a digital camera image:

- **Ctrl**-click the page thumbnail.

- Select **Digital Camera Image**.

Note: when photographing documents with a digital camera, it may happen that the documents are shot from an angle, which gives them a 3D effect and renders the recognition impossible. To solve this issue, see Optimizing the Scanned Documents > Deskewing digital camera images.

Tips for making digital camera images

- Always select the highest image resolution.

- Enable the macro mode of the camera to take close-ups.

- Enable the document mode of the camera (if any).

- Only use optical zoom, not digital zoom.

- Hold the camera directly above the document. Avoid photographing documents at an angle.

- Produce stable images.

- Disable the flash when photographing glossy paper.

- Avoid opening compressed camera images.

33

Converting Image Files from Finder

In Readiris you can now convert Image files and PDF files directly from Finder. You may select a mix of Image and PDF files.

When you select multiple files for conversion, a multipage document will be created, each page corresponding to a file. The output file takes the name of the first input file you selected. The order of the files in Finder will also be the order in the multipage document. So the order in which you *select* the files in Finder does not matter.

Note that the Readiris interface is allowed to be open while converting images from Finder, but this is not required. Also note that converting files from Finder is even possible when Readiris is busy with file processing.
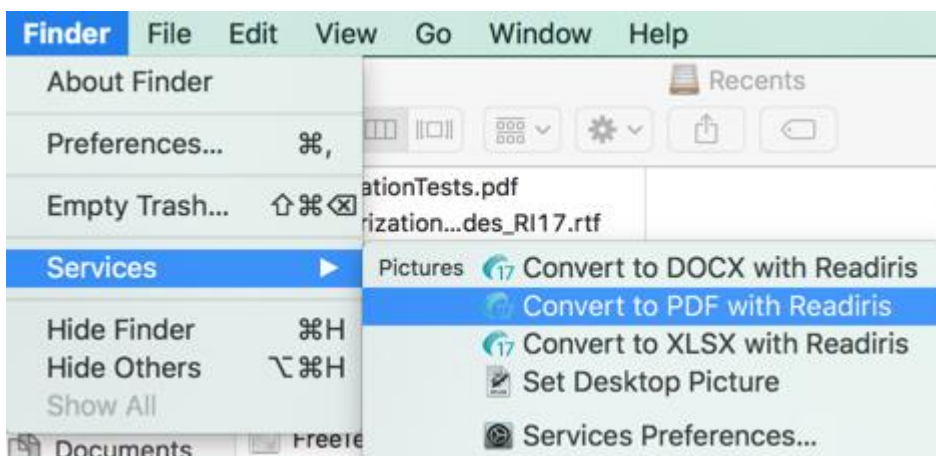
To convert Image Files from Finder:

- Ctr-click or Right-click the file(s) you want to convert.

- Point to **Services** and select the required conversion option:

    - **Convert to XLSX with Readiris**

    - **Convert to PDF with Readiris**
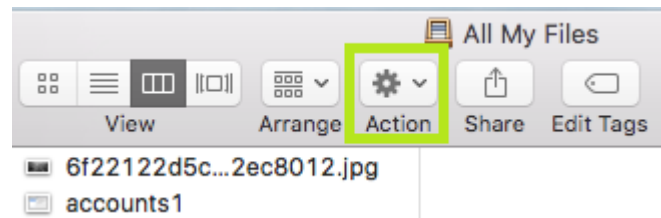
    - **Convert to DOCX with Readiris**
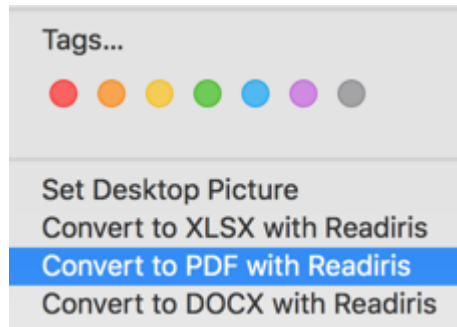
Or

- Select the file(s) you want to convert.

- Click **Finder > Services**.

- Then select the required conversion option.

Or

- Open Finder.

- Select the file(s) you want to convert.

- Click the down arrow in the **Action** button.

- Point to **Services** and select the required conversion option

Note that the **Training Mode options**, **Page Separation options** and **Cloud Storage Options** are not taken into account.

The file is converted into the format you selected, keeping the same file name and storage location.
If a file with the same name already exists you are prompted whether to overwrite it or not.

⬥ Warning

- The following Dialog boxes must not be open in Readiris in order for the right-click conversion to work:

    - **Settings** > **Document Language**

    - **Settings** > **Selected Pages Language**

    - **Settings** > **Export Options**

    - **Settings** > **Open Settings**

- When you have not activated Readiris yet and you are using it in trial mode, you can only convert up to 10 pages at a time, and a maximum of 150 pages.

Section 5: Selecting the Document Language

Readiris converts scanned images, image files and PDF files into editable text documents and text-searchable PDF documents. In order for Readiris to recognize the text in your images, you need to select the correct **recognition options**.

By far the most important recognition option is the **document language**.

To select the document language:

- Click the language list, and select your document language.

- To access the complete list of available languages, click **Other Languages**.
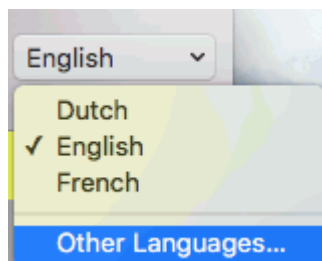
💡 **Tip Readiris Pro**: in case you want to recognize documents in multiple languages, make sure to select the language with the largest character set. E.g. if you want to recognize a document that contains both English and French text, select French as document language. This way, the accents will be recognized correctly.

Recognizing Numeric Documents

When you are processing documents that only contain numbers and almost or no text, then it is recommended to select the **Numeric** option:

- Click the language list and select **Other Languages**.

- Then select **Numeric** at the top of the list.
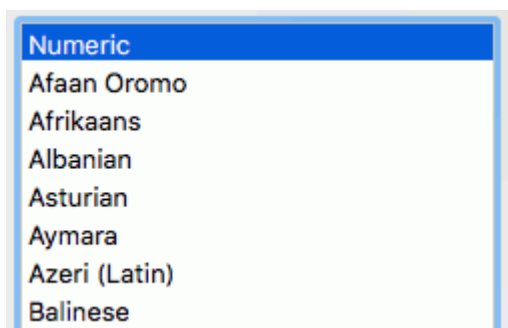
When this option is selected Readiris only recognizes the numerals 0-9 and the following series of symbols:

| +	*	/	%	,
plus sign	asterisk	slash	percentage sign	comma
.	(	)	-	=
period	opening parenthesis	closing parenthesis	hyphen	equation sign
$	£	€	¥	
dollar sign	pound sign	euro sign	yen sign	

Recognizing Western words in non-Latin alphabets

When you are processing Cyrillic, Slavic, Greek or Asian documents that also contain "Western" words written in the Latin alphabet - such as proper nouns, then it is recommended to select one of the available **Language pairs**.

Language pairs are always combined with the English language and are available for Russian, Byelorussian, Ukrainian, Serbian, Macedonian, Bulgarian and Greek.

Note: when processing Asian or Hebrew documents, mixed characters sets are used automatically.

To select a Language pair:

- Click the language list and select **Other Languages**.
- Select the required language pair in the (primary) language list.
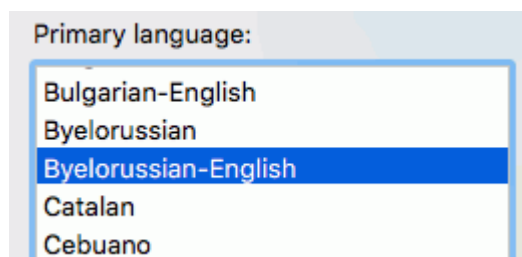
Selecting the language per page

When specific pages use a different language than the overall document, you don't need to define a secondary language. You can apply a different language to those pages.

Select the pages in the Pages panel, **Ctrl**-click them and use the command **Language** to assign another language than the overall document language to that/those page(s).
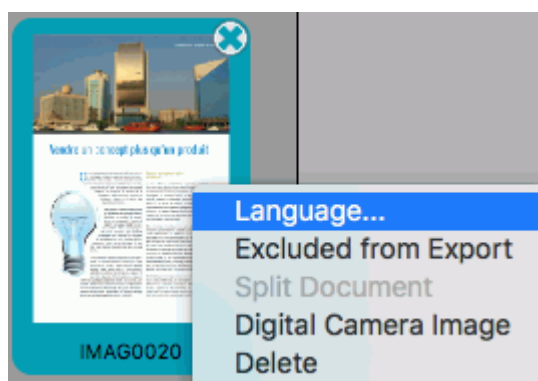
Pages with a different language than the overall language are **marked in red** in the Pages panel.

Unlike secondary languages, there are no limitations here.

Note: the tooltip of each page in the Pages panel indicates which language applies to that page.

Recognizing Secondary Languages inside a single document (Readiris Corporate only)

When your documents contain text in multiple languages, it is recommended to select a main recognition language, combined with several **Secondary languages**. You may select up to 4 secondary languages:

- Click the language list and select **Other Languages**.

- Select the **Primary language**.

- Cmd-click to select the **Secondary languages**.

  The list of secondary languages changes depending on the selected primary languages.

Note: do not select languages that do not apply; the bigger the character set, the slower the recognition and the higher the risk of OCR errors.

Section 6: Using Training Mode

(This feature is not available for Asian languages)

If you are processing documents with a non standard font and you notice that Readiris is systematically having trouble recognizing characters, use the **Training Mode** to train the recognition system on the fonts and character shapes you are using.

During the training process, any characters the recognition system isn't sure of are displayed in a preview window, in combination with the word in which they were spotted and the result suggested by Readiris.

1	A character Readiris is not sure of.
2	The word in which the character was spotted.
3	The solution how Readiris suggests to recognize it.

Training can substantially enhance the accuracy of the recognition system and is particularly useful when recognizing distorted, defaced forms. Training can also be used to train Readiris on special symbols it is unable to recognize initially, such as mathematical and scientific symbols and dingbats.

ATTENTION: training occurs *during* recognition. The training results are temporarily stored in the computer memory, for the duration of the recognition. Readiris will no longer display the trained characters when OCRing the rest of the document. When a new document is OCRed, the training results are erased. To save training results permanently, save them in a Training Results set. Once Training Results have been saved, you can also use them without activating Training Mode.

When is it better not to use Training Mode?

- When you are processing documents with a misleading or nearly unreadable content - such as a picture showing an opened book or showing a user interface. In this case it is better to change the Zone Type and set it to **Graphic**. This way you avoid Readiris trying to recognize those parts of the document.

- When you are processing documents that are too dark or to bright, it is better to Adjust the Image Quality so that the characters are better isolated.

- When processing Digital Camera Images the final result is often easier to improve by correcting the lighting conditions, focus, camera stability and taking a new picture instead of trying to process a low quality image.

- When you want to inspect the recognized text and make necessary corrections, it is better to use the Text Corrector instead of the Training Mode.

Using Training Mode

- Go to **Settings** > **Training Mode**.

- Select **Training mode**.
  The option **Start to create a new training results set** is now selected.

- Click **OK** to close the Training Mode settings for now.
  The other options are explained in Combining Training Mode with Training Results sets below.

- Scan or Open your document.

- At the end of the recognition, Readiris enters the Training Mode.
 The characters the recognition system is not sure of are displayed.

If the results are correct:

- Click the **Learn** button to save the result as sure.
  The training results are temporarily stored in the computer memory, for the duration of the recognition. When you scan a new document, you will have to go through the same steps again.
  To avoid having to go through the same steps, you can combine Training Mode with Training Results sets.

- Click **Finish** to accept all solutions the software offers.

If the results are incorrect:

- Type in the correct characters and click the **Learn** button.

> **Note**: if you are dealing with documents that contain special characters that are not available on your keyboard, click the browse button to open the Character Palette. Double-click the character you want to insert.
> You can also drag and drop a character from the Character Palette to the character field in Training Mode.
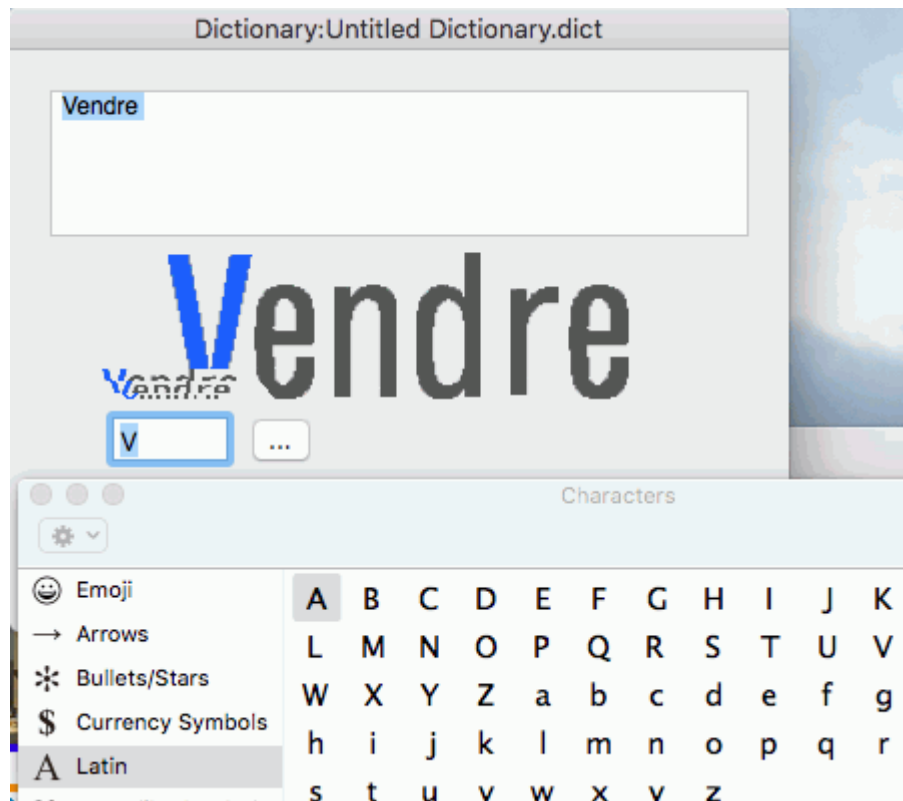
or

- Click **Don't learn** to save the result as unsure.
  Use this command for damaged characters which could be confused with other characters if trained. E.g. the number 1 and the letter I, which have an identical form in many fonts.

- Click **Delete** to delete characters from the output.
  Use this button to prevent document noise from appearing in the output file.

- Click **Undo** to correct mistakes.
  Readiris keeps track of the last 32 operations.

- Click **Abort** to abort Training mode.
  All training results will be deleted. Next time you process a document, Training Mode will start again.

When Training Mode has finished, you can now save the Training Results sets.

Combining Training Mode with Training Results Sets

As stated above, you can use the Training Mode combined with a Training Results set, in order to store the training results permanently. Using Training Results Sets is recommended when you are processing multiple documents that have the same typographic characteristics.

- Go through Training Mode as explained above.

- Return to **Settings** > **Training Mode**.

- Click **Save** to save the Training Results set.

- Name the Results set and save it in a location of your choice.

> **Note**: Results sets are limited to 500 shapes. You are recommended to create separate sets for specific applications.
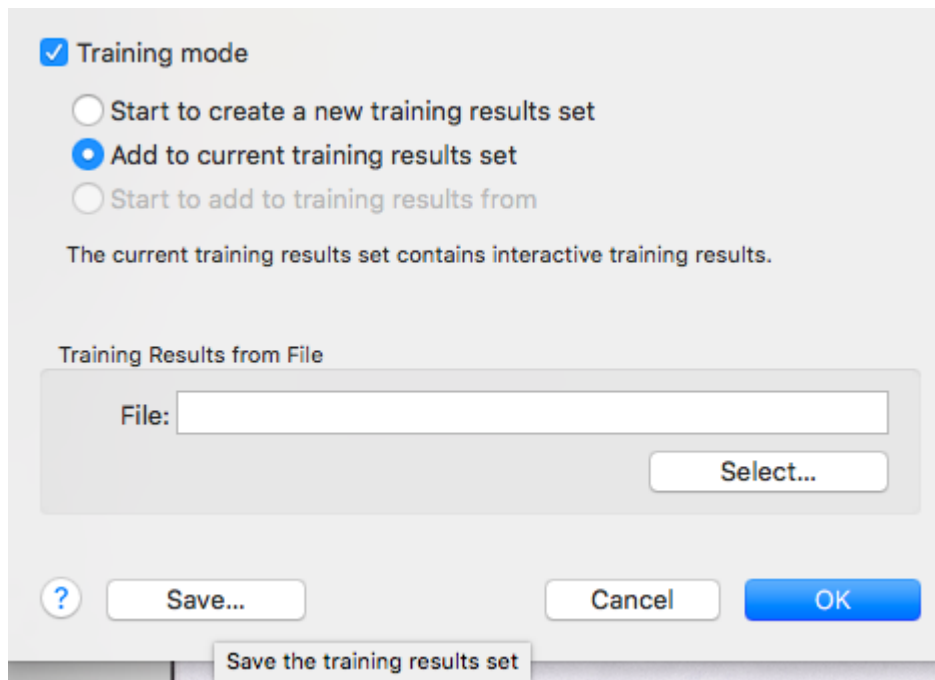
Now you have several possibilities:

- You can select **Start to create a new training results set**.
  In this case a new Training Mode process starts.

- You can select **Add to current training results set**.
  In this case the new Training results are only stored temporarily and they are not stored in the Results set you saved.

- You can click **Select**, browse for the Results set you saved and select **Start to add to training results from <YourFileName.dict>.**
  In this case the new Training results are added to the Results set you saved.

You can also choose to deactivate Training Mode and:

- **Don't use any training results**.

- **Use the current training results set**.
  In this case only the temporary training results are used and you don't have to go through Training Mode again.

- Click **Select**, browse for the Results set you saved and select **Use training results from <YourFileName.dict>.**
  In this case you use the results from the stored Results set without going through Training Mode.

Section 7: Optimizing the Scanned Documents

The documents you scan and open in Readiris can be optimized in several ways with the Image and Layout Edting toolbar on the right side of the interface.

Deskewing pages

When a page has been scanned crookedly it can be **deskewed** - or straightened.

To deskew a page:

- Select the page in the Pages panel:
- Click the **Deskew** icon.

Deskewing digital camera images

Readiris automatically detects if an image was made by a digital camera. Notice the camera icon on the page thumbnail.

> **Tip**: if Readiris should not detect that the image was made by a digital camera, Ctrl-click it and select **Digital Camera Image**.

If the image was made from an angle, you can correct its perspective with the **Deskew** icon.

> Notice that the Deskew icon looks different for digital camera images.

To deskew a digital camera image:

- Select the image in the Pages panel.
- Click the **Deskew** icon.

Automatic detection

- Readiris tries to detect the borders of the image automatically. If it succeeds, a frame showing the detected borders is displayed and the option **Automatic detection** is selected in the **Borders** list. This means the borders cannot be moved manually.
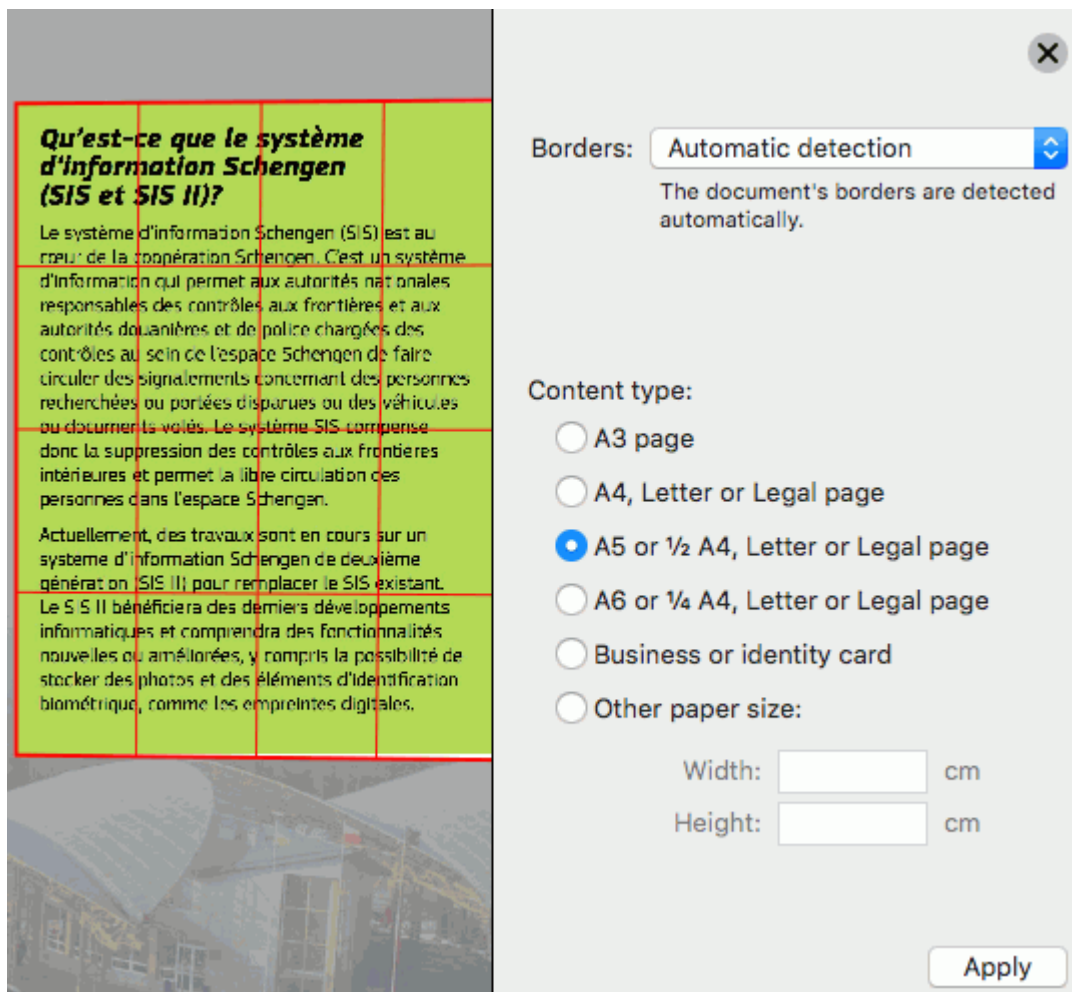
- Select the appropriate **Content type**.

> **Important**: Readiris computes the image resolution based on the **Content type** you select. A correct resolution is required to obtain satisfactory OCR results.

Note: the fields **Width** and **Height** are only available when you select **Other paper size**.

- Then click **Apply** to apply the automatic deskew.

If Readiris fails to detect the borders automatically, Readiris tries to find the best alternative **Borders** option.

Interactive definition

When **Interactive definition** is selected, you can drag the corners of the frame to a new position in order to align the frame with the page borders.

You can also move the horizontal borders vertically and the vertical borders horizontally.

Rectangular frame

When **Rectangular frame** is selected, you can move the horizontal borders vertically and the vertical borders horizontally.

No cropping

When **No cropping** is selected, no frame is displayed and no cropping can be done.

Note that the **Borders** options still work even if the borders of the document are outside of the image.

Rotating images

To rotate an image, click the **Rotate** icon. Then click the required option.
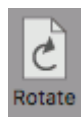
Adjusting the Image Quality

The recognition results depend a great deal on the **image quality** of your documents. If you're scanning very low-quality documents, Readiris will not be able to recognize them properly.

The reason for this is that Readiris works by converting **pictures** of text characters into **actual text characters**. If your documents are too dark, the characters fade into one another and create black spots. If your documents are too light, the text characters might be broken. Consequently, Readiris will not be able to recognize them.

Example 1: text is too dark

Here, the letter shapes begin to fill in and become fuzzy. Letters like a, e, and o are almost completely blurred.

Example 2: text is too bright

Here, the text characters start to show gaps and incomplete shapes.

To improve the recognition results:

- Check the **Scanner Settings** and make sure your documents have been scanned in a resolution of 300 dpi and in color.
- If that's not the case, re-scan your documents using these settings.

Then access the Image Adjustment options:

- Click the **Adjust** icon on the Image and Layout Editing toolbar.
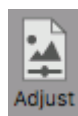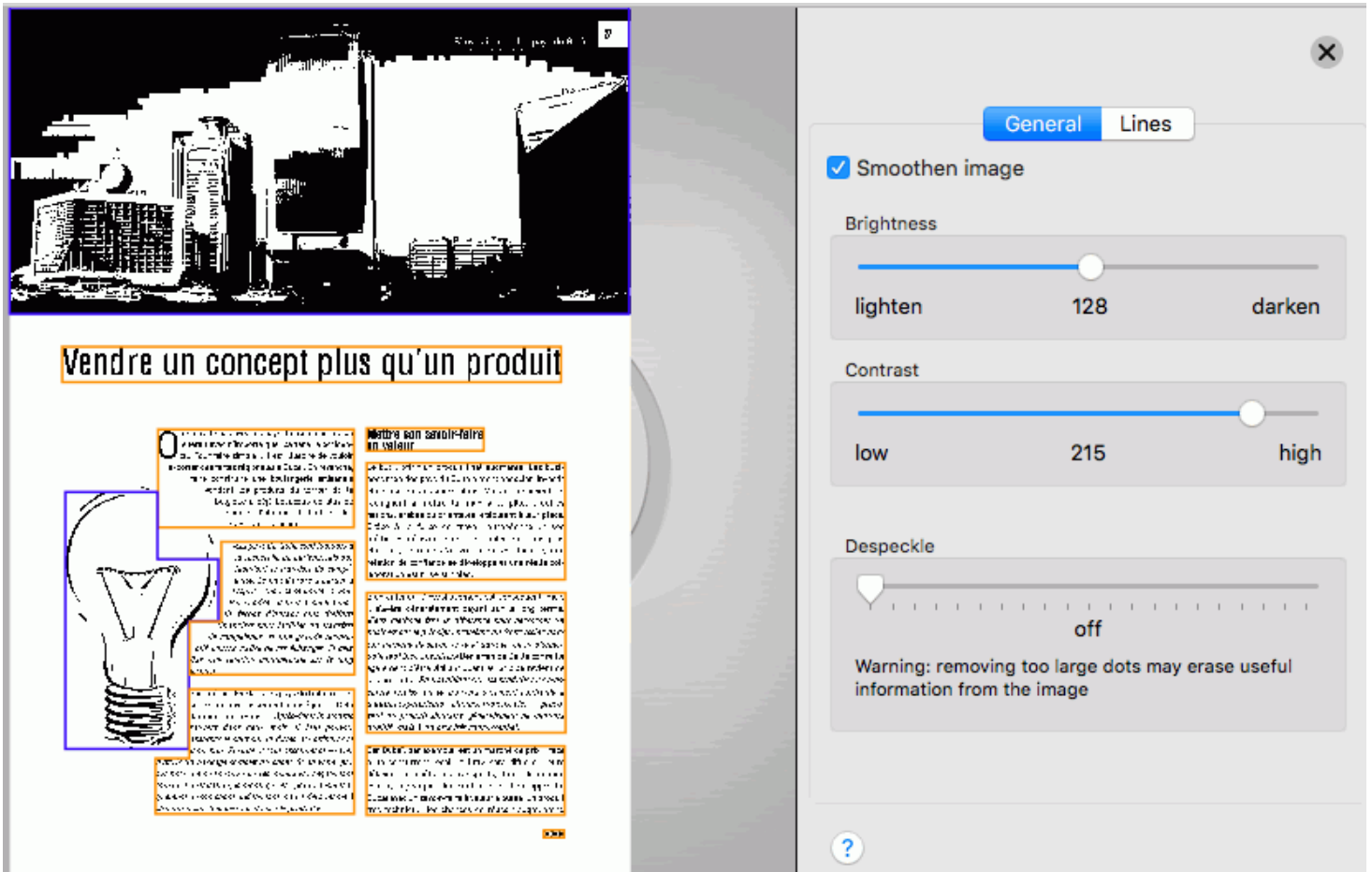
General

Smoothen (color or grayscale) image

This option is selected by default. It smoothens out differences in intensity, creating a greater contrast between the text and the background. Sometimes smoothening is the only way to separate text from a colored background.

Remove dither (black-and-white image)

This option removes unwanted dither (noise) from black-and-white images.

Brightness

Move the slider to increase or decrease the brightness of your documents.

The results on the binarized image are shown immediately. If the results are not satisfactory, use the **Undo** button to restore the settings.

Example 1: image is too dark

The image below is too dark and results in a plain black image. There is no text to be recognized.

In this case you need to brighten the image in order for the text to be displayed.

Example 2: image is too bright

The image below is too light and results in broken text characters. The text is hardy legible.

In this case you need to darken the image to obtain satisfactory results.

Contrast

Move the slider to increase or decrease the contrast between the text and the background.

The results on the binarized image are shown immediately. If the results are not satisfactory, use the **Undo** button to restore the settings.

Example

The image below results in broken text characters.

In this case you need to increase the contrast to obtain satisfactory results.

Despeckle

Sometimes documents contain a lot of "noise" - little black spots that occur when scanning low-quality documents or using bad scanner settings.

To remove these little black spots, use the Despeckle slider. The more you move it to the right, the larger the spots that will be removed.

The results on the binarized image are shown immediately. If the results are not satisfactory, use the **Undo** button to restore the settings.

Lines

Readiris is now able to remove vertical and horizontal lines to improve the text recognition.

> Note that the lines are not removed from the output documents, they are simply removed before doing the recognition.

To remove vertical lines:

- Select **Remove vertical lines**.
- Enter values for the 3 parameters:
    - **Minimum length**: minimum length (in pixels) of the lines to remove.
    - **Maximum thickness**: maximum thickness (in pixels) of the lines to remove.
    - **Maximum gap**: maximum line gap (in pixels) to close, in case of broken, interrupted lines.

To remove horizontal lines:

- Select **Remove horizontal lines**.

- The same parameters are available.
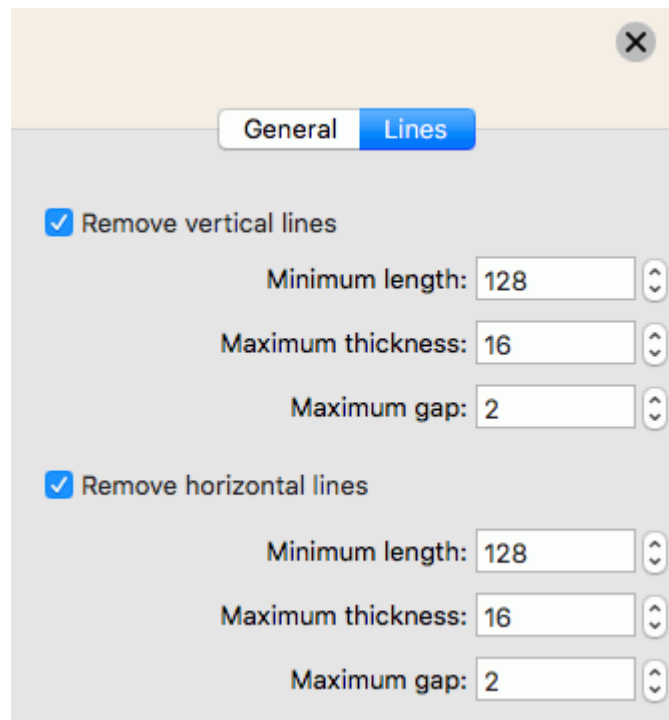
When you are done adjusting the image quality, click **X**.

Important note: the **Image Adjustment** options are applied the binarized image only. Once you have applied them, the original image is displayed again.

Section 8: Editing the Recognized Documents

Introduction

When you scan a document or open an image file in Readiris, each page is divided automatically into **recognition zones**.

To visualize the recognition zones, click the **Layout** icon in the **Image and Layout Editing** toolbar**.** Readiris is now in **Layout** mode.
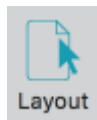
Readiris uses the zones to determine how each part of your documents must be converted. If your document does not contain any zones, no recognition will be done.

In Readiris there are 4 kinds of recognition zones. You have **text** zones, **image** zones, **table** zones and **barcode** zones.

Notice that each zone type has its own color.

Text zones	Orange
Image zones	Blue
Table zones	Purple
Barcode zones	Green

The zones are also numbered; notice the number in the top left corner of each zone. The numbers run top-down and from left to right. When you select an output format that does not preserve the original layout, this will be the order in which the zones are saved in the output document.

Note: Zones can also be selected when the **Layout** mode is not enabled: E.g. to copy as ...

Viewing the Recognition Zones

Several viewing options are available allowing you to inspect the opened images and their recognition zones.

Changing the View Size

- Use the size icons **Fit to width** and **Fit to window** in the bottom right corner of the interface.
- Or move the slider to zoom in and out.

Note that you can also use keyboard shortcuts to apply these viewing options.

Using keyboard shortcuts

- Shift-click on an image to toggle between **Zoom to 200%** (at cursor location) and **Fit to window**.
- Cmd-Shift-click to toggle between **Zoom to 200%** (at cursor location), **Fit to width** and **Fit to window**.
- Use Cmd + and Cmd - to zoom in/out progressively.

Using the Magnify icon

- Click the **Magnify** icon on the **Image and Layout Editing toolbar**.
  Then move the Magnifying glass over the image.

- Or Alt-Shift-click to display the Magnifying glass.

You can also use the keyboard to zoom in on images:

- Shift-click on an image to zoom in.

- Then Alt-click on the image to grab it and move to another section of the image.

- To zoom back out, Shift-click on the image again.

Or

- Shift-Alt-click on an image to display the Magnifying glass.

- Hold down the mouse button and move the Magnifying glass over the image.

Modifying the automatic Page Analysis

Readiris divides documents into **text**, **table** and **graphic** recognition zones when the option **Page Analysis** has been activated. This automatic page analysis can be modified in several ways. For instance, you can expand and reduce zones in size, move zones to a different location, manually create new zones, change the zone type of existing zones, change the sort order, and so on.

Tip: pay attention to the messages at the bottom of the interface when using the **Layout** toolbar. They provide useful information.
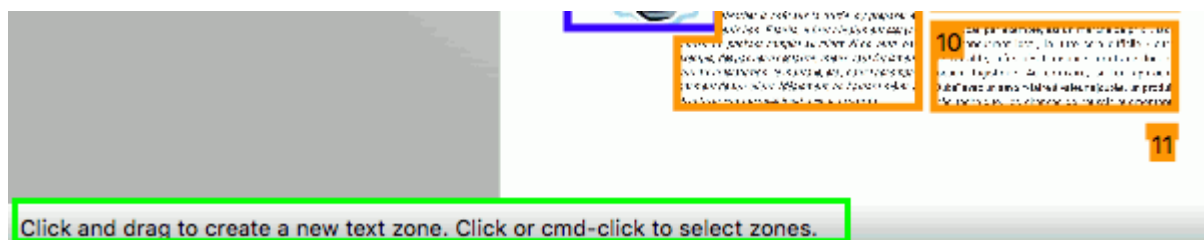
Note: if you need to do **text corrections** start the **Text Corrector** by clicking the corresponding button.
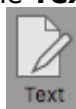
Expanding or reducing zones

- Click inside a zone.

  Notice the markers on the sides and in the corners of the zone.

- Click on a marker, hold the mouse button and drag the frame to the required size.

Moving zones

- Click inside a zone you want to move.
  The entire zone is now marked in its corresponding color.

- Hold the mouse button, and drag the frame to the position of your choice.

Deleting zones

- Click the zone you want to delete.

- Or Cmd-click to select multiple zones.

  The entire zone(s) are now marked in their corresponding color.

- Then press the Backspace button on your keyboard or right-click and choose **Delete zone(s)**.

Deleting all zones

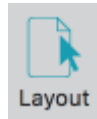- On the **Edit** menu, click **Delete All Zones**.

Deleting small zones

- On the **Edit** menu, click **Delete Small Zones**.

 All zones that are smaller than 50 pixels will be removed.

Manually drawing zones

- Click the **Layout** icon to display the Layout toolbar.

- Click the zone type you want to create. E.g. Text.
- Click and hold the mouse button and draw a frame around the section to want to recognize as text.

  Note that barcode zones must always be drawn manually. They are not created automatically by the **Page Analysis** feature.

Changing the zone type

- Ctrl-click a zone, and then select the required zone type from the Context menu.
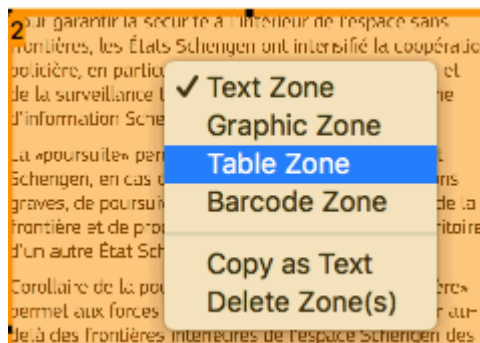
OR

- Click the **Layout** icon to display the Layout toolbar.
- Select the zone you want to change.
- Click the required zone type icon in the Layout toolbar. E.g. Graphic.
- Then click the required zone type icon again.

  💡 **Tip**: notice the tooltips at the bottom of the interface. They describe which actions you can do.

Changing the sort order of zones

- Click the **Layout** icon to display the Layout toolbar.
- Click the **Sort** icon.

  The numbers now disappear from the zones.

- Now click the recognition zones one by one, in the order you want to sort them in.

  Zones you do not click, will be excluded from recognition.

- When you have click all the zones on the page, Readiris automatically quits the Sort mode.
  You can also quit the Sort mode by clicking the **Sort** icon again. You will be prompted whether or not to keep the new sort order.

Connecting zones

Suppose you have two text zones that belong together, and you want to turn them into one single text zone. In that case:

- Click the **Layout** icon to display the Layout toolbar.

- Click the **Text** icon.

- Then draw a text zone that overlaps the two zones to connect them.
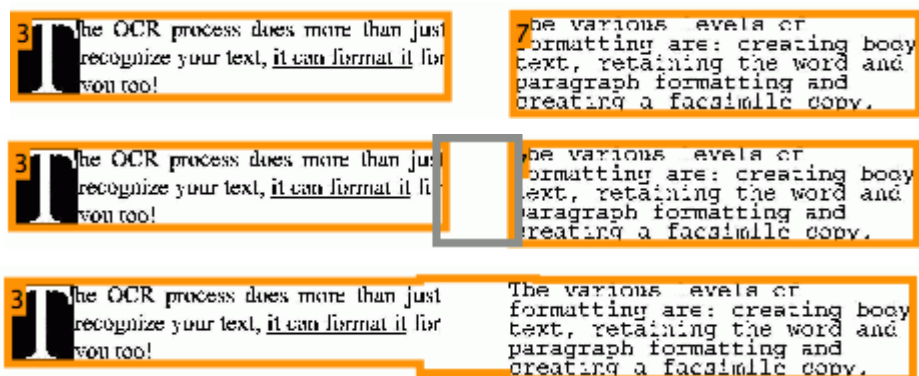
- The two zones are merged into one.

    Note that only zones of the same zone type can be connected.

OR

- Close the Layout toolbar.

- Click a zone to select it. The entire zone is now marked in its corresponding color and the markers on the borders are displayed.

- Click one of the markers and expand a side of the zone until it overlaps with another zone of the same type.
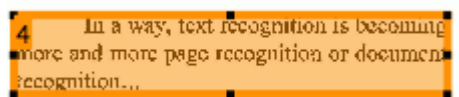
- The two zones are merged into one.

Recognizing a particular area

In case your document only contains content in a certain area of the page, you can use the **Frame** option to exclude the remainder of the page from the recognition.

- Click the **Layout** icon to display the Layout toolbar.

- Click the **Frame** icon.

- Then draw a frame around the section that contains the information.

    Everything outside the frame will be ignored.

    You will also be prompted whether you want to ignore the same area on all the pages of the current document.

    Note: the command **Frame Area to Analyze** on the **Page** menu carries out the same operation.
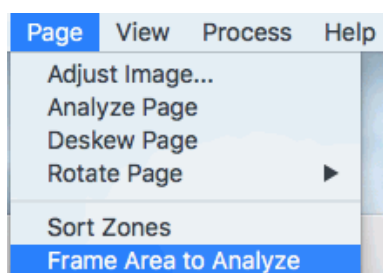
Re-doing the Page Analysis

If you are not satisfied with the changes you made, note that you can always redo the automatic page analysis.

- Go to Menu **Page** and select **Analyze Page**.

Using zoning templates

When you want to scan many similar looking documents, it might be useful to apply **zoning templates**, instead of editing the recognition zones each time. In a zoning template you indicate what your documents look like and where the text, picture or tables zones are located. This way, Readiris only looks in the indicated areas, saving you a considerable amount of time.

To create a zoning template:

- Scan the first document.

- Edit the zones to your liking.

    See Editing the Recognition Zones for more information.

- On the **Page** menu, click **Save Layout**.
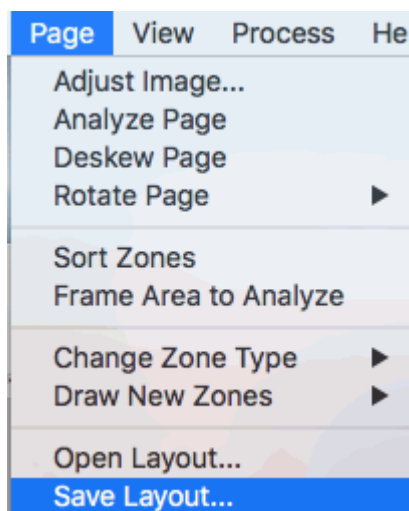
To use a zoning template:

- First, deactivate Page Analysis on the **Readiris** menu > **Preferences** > **Page analysis**.
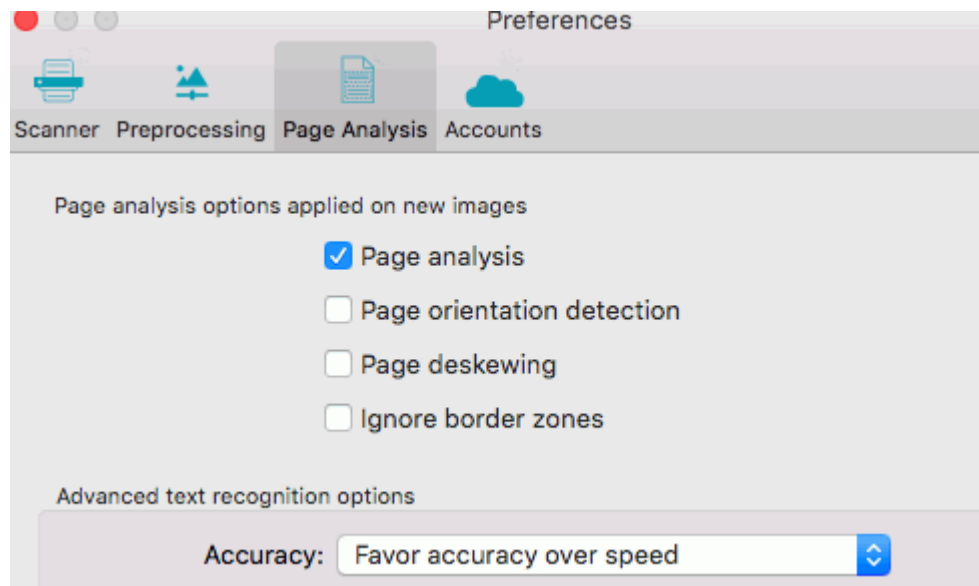
- On the **Page** menu, click **Open Layout**.

- Open the layout file you have saved.

- Choose whether to use the Layout **for the current page and new pages** or **for the current page only**.

- Then, scan or open your documents in Readiris.

To stop using the template, re-activate Page Analysis again.

Manually copying the zoning layout.

It is also possible to manually copy the zoning layout from one page to another:

- Cmd-click the zones you want to copy to another page.

  💡**Tip**: make sure the Layout toolbar is closed. Otherwise, Cmd-click won't work.

- Then press **Cmd+C** on the keyboard.

- Go to the page to which you want to copy the zoning layout.

- Click inside the image and press **Cmd+V** on the keyboard.

Using the Text Corrector

The Text Corrector allows you to correct the recognition results. Note that the Text Corrector is meant for **text corrections**. It is not a Text Editor in which you can change the layout of the document, delete entire paragraphs, merge content, etc.

To access the Text Corrector:

- Click the **Text** icon in **Image and Layout Editing** toolbar.

   Note: the **Text** icon is only available if you have at least one document open in Readiris.
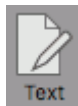
How does the Text Corrector work?

The Text Corrector displays the recognition results of the currently selected page. The original image is shown in the left-hand column and the recognized text in the right-hand column.

> **Important note:** if a page does not contain any Recognition Zones, there will be no text results to display.

To switch to horizontal view, click **H**. The recognized text is now displayed above the original image.

By default the recognized text is shown in the original layout. If you want to display only flowing text without images, select **Flowing text** in the drop-down list.

The suspected words are marked in yellow. The spelling corrector underlines the words that are detected as incorrect.

Correcting Text

- Click inside the recognized text.
  An orange frame in the original image shows where you have placed the cursor.
  The word in which the cursor is located or the text selection is highlighted in the original image

- Type in the corrections inside the recognition results.
  When a correction has been done the yellow highlighting is removed from the word.

  Tip: Ctrl-click a word to display spelling suggestions in the Context menu.

- Press Backspace on the keyboard to delete any unwanted results.

- You can also delete entire words or entire lines by selecting them with the mouse cursor and by pressing Backspace.

  Note that you can delete one line at a time.

  Also note that it is not possible to make a selection of non-contiguous characters or words.

- The standard **Copy**, **Cut** and **Paste** commands can also be used via the **Edit** menu, the Context menu or the Keyboard shortcuts.

- To **undo** corrections, press **Cmd+Z** on they keyboard. Or use the **Undo** option in the **Edit** menu.

  Note that the **Undo** operation is not limited to text corrections. It also undoes changes to the current page, page reordering, etc.

  Also note that page deletion cannot be undone.

- To switch to the next suspect word, click the corresponding button or press Tab on the keyboard.

- To move to different pages, click the thumbnails in the Pages panel.

- When you are done correcting the results, click **X** to close the Text Corrector.

Notes

- Any changes you do will be taken into account in the exported document. The changes will not be displayed in the original image, but the text behind the original images does contain the changes.

- It is not necessary to close the Text Corrector before you Export your documents.

- When the Text Corrector is open it is not possible to modify the Zone Types on the original image.

Changing the Zoom

- To zoom in or out on the original image use the corresponding zoom commands.

- To zoom in on the recognized text, use the slider.

- To change the zoom of the recognized text and the original image simultaneously, click the attachment icon, then move the slider.

Section 9: Adding Annotations for PDF export

Using Annotations

The 'Annotate' Option allows you to add comments of various types on your document.

⬦ Warning:
Those annotations are supported and included into the final document when exporting to a **PDF format only**. Optimal results are rendered with the PDF viewer Acrobat Reader. Good results with any other PDF viewer are NOT guaranteed.

To access the Annotation Mode

- Click the **Annotate** icon in Image and Layout Editing toolbar.

OR

- Go to the **Page Menu** and click on **PDF Annotations...**
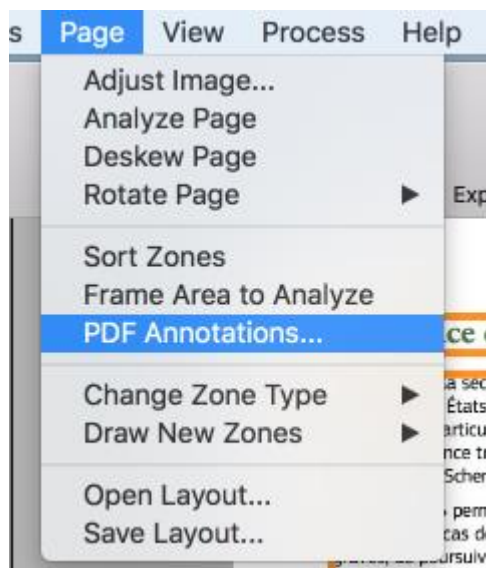
- The right hand column opens on an edition pane to create, edit and manage annotations.
    - The list of annotations. They are displayed per page, showing type and text related. The list allows the selection of one annotation at a time.
    - The **Remove button** deletes the selected annotation.
    - A drop-down list allows to select the **type of annotation** you want to add.
    - Below the type's list, a frame containing the **edition options**. The options depend on the annotation type.
    - A **Help button** that opens the HTML documentation at the relevant chapter and section.
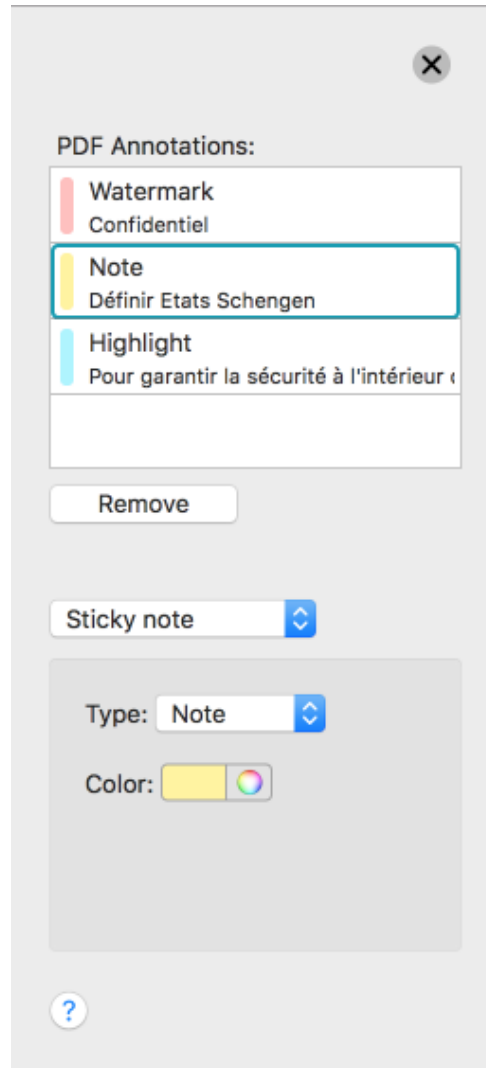
When the **annotation mode** is active :

- the selection is set on the last annotation created on the page. If the page has no annotation, then the selection is set on the last type used.

- the zone boundaries on the page are no more visible (in the image view).

Note: the Annotation mode is enabled even when no document is imported in Readiris, because an annotation of type **watermark** can be defined at any time.

Annotation Types

The supported annotation types are:

- Watermark
- Free text
- Sticky note
- Audio note
- Highlight, underline or strikethrough Text
- File Attachment
- Stamp

When saved, all annotations are associated with a creation date and the username of their author.

When a page is modified (rotation, text edition,...), the annotations are kept and repositioned or rotated if needed.

To remove an annotation on a page

From the **Annotation Pane**:

1. In the PDF Annotation's list, select the Annotation type you want to remove
2. Click on the **Remove button**.

From the **Image View**:

1. Right-click on the Annotation. A **Context menu** opens.
2. Click on the **Remove option**.

To quit the Annotation Mode

- Click on the cross button ( ) in the upper right corner of the annotation pane.

To export as PDF

Please refer to the specific options for [PDF](#) output format.

> ⬥**Warning**:
> When selecting another format than PDF for export, a warning pop-up window is displayed: Annotations are not supported by other formats than PDF.
> The option 'Do not show this message again' can be selected.
>
>
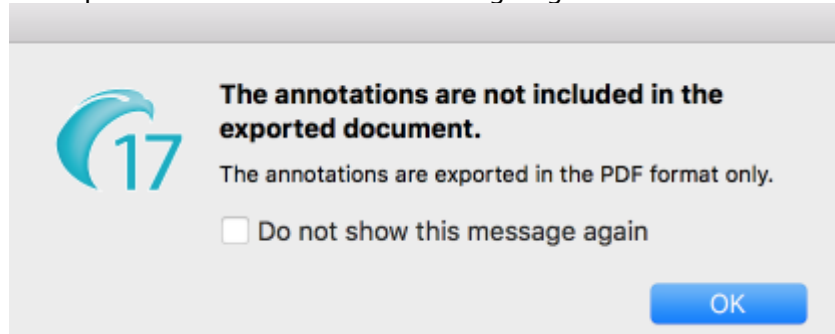
Annotation Types

Watermark

The **watermark annotation** is a text drawn across a document's page.

To add a Watermark:

In the **Annotation pane**:

1. Select **Watermark** in the drop-down list of Annotation types
2. Type in the line of Text you want
3. Use the **edition options**

Edition options:

- Text **definition**
- Text **orientation**: horizontal or diagonal
- Text **color** (color palette)
- Text **opacity** (transparency scale)
- Text **size** (scale from small to large)

Watermark Edition Interface
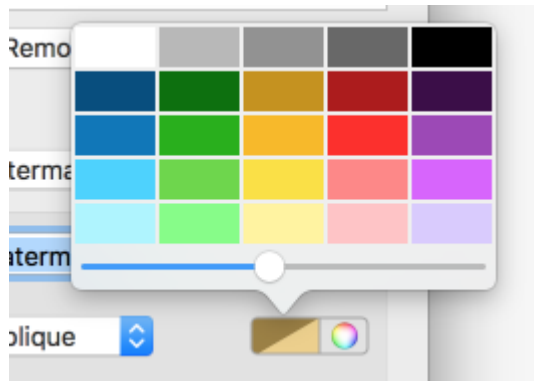
Quick selection of color and opacity
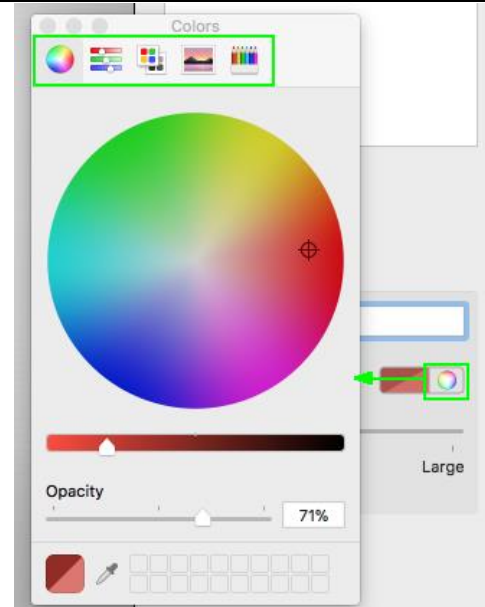
Advanced selection of color and opacity

Mac offers a menu of options for defining the colors.

Default features of a Watermark:

- The text is a single line

- The text is centered horizontally and vertically on the page

- The text font style is predefined

- The watermark is added to all the pages of the document. As an **option** in the Pages Panel, you can hide the watermark from the PDF Export on the pages you have selected. Right-click to open a **Context menu**, then select the option '**Hide Watermark in PDF Export**'.

Notes:

- When exported into a PDF file, the watermark is permanent. It cannot be hidden when displayed or printed.

- The watermark attributes are not reset when a new document is started and it is possible to set or edit them even if there is no page imported in Readiris.

- The watermark is also added to documents created in batch, in watched folder mode or by conversion with the right mouse click in file explorer (**Finder**).

Free text

The **free text annotation** is made of one or several text lines added on a document's page.

To add a Free text:
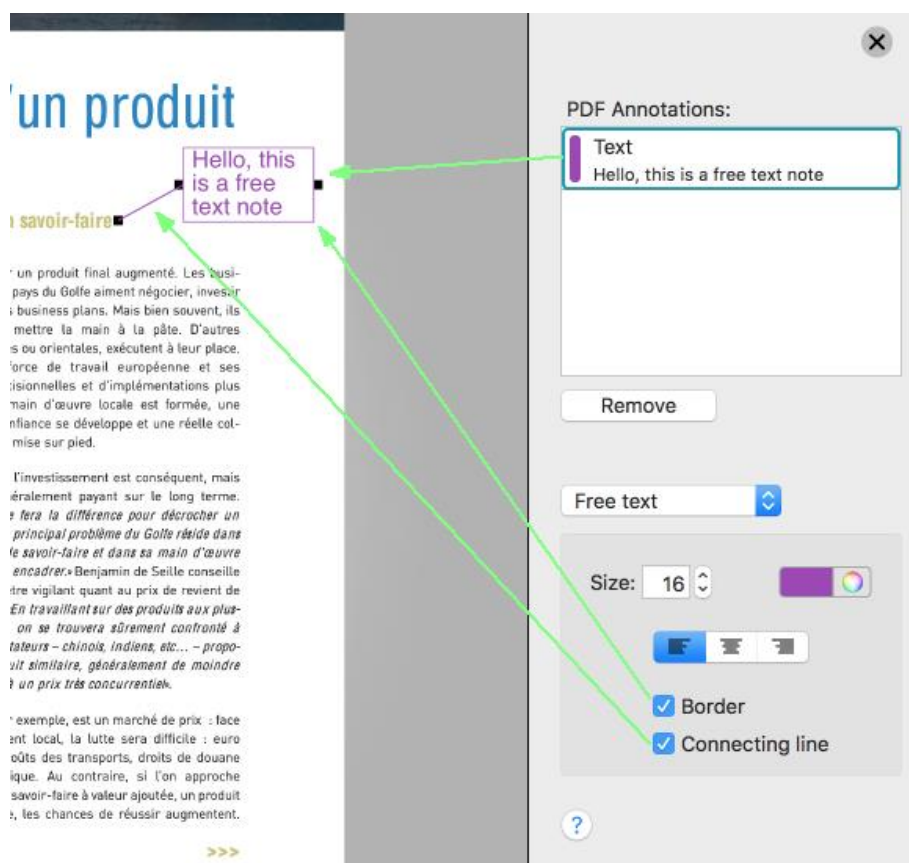
In the **Annotation pane**:

1. Select **Free text** in the drop-down list of Annotation types

2. Click on the page where you want the free text to be displayed. The text is laid out in a rectangular box. Click again inside the box to make the keyboard blinking cursor appear.

3. Type in your text. The box height automatically adapts to fit the required space. You can adapt the box width when the cursor shows a **horizontal resize icon** ⊹

4. You can move the position of your free text when it is selected and the **move cursor** appears

Edition options:

- Text **Font size**

- Text **color** (and border and connecting line if selected): **quick color selection** or **advanced color selection**
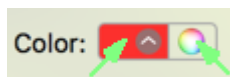
- Text **alignment**: left, center or right

- Text **border**

- **Connecting line**. When **Connecting line** is selected, a small 'tail' is attached to the text box. The end point of that line can be moved where you want on the page (thanks to the **move cursor**). When the end point is moved, the start point location is updated automatically to prevent from the connecting line to be drawn over the annotation text.

Default features of a Free text:

- The text is drawn on a semi-transparent white background. That way it is easier to read without masking completely the content behind.
- The text font style predefined.
- The connecting line has the same width as the border line.

Sticky Note

The **sticky note annotation** appears as an icon on a document's page. When single-clicking on the icon, a pop-up window with the text content is shown.

**To add a Sticky Note:**

In the **Annotation pane**:

1. Select **Sticky note** in the drop-down list of Annotation types
2. Select a sticky note type in the drop-down list of types
3. Select a color for the sticky note
4. Click on the page where you want the sticky note to be shown. The annotation is created. The text content is still empty
5. Type your text in the text box attached to the icon
6. You can move the position of the note when it is selected and the **move cursor** appears

Edition options:

- Icon **Color**: **quick color selection** or **advanced color selection**
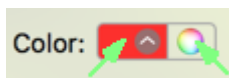
- Various **sticky note types**:

  o **Comment**: with a speech bubble icon

  o **Key**: with a star icon

  o **Note**: with an information symbol icon

  o **Help**: with a help icon

  o **Insert**: with a plus symbol icon

  o **Question**: with a question mark symbol

  o **Remark** : with an exclamation mark symbol

- **Text Content**: the text is edited in the pop-up window once the icon is clicked. The size of the pop-up window can be adapted when the cursor shows a resize arrow ↙ from one of its lower corner.
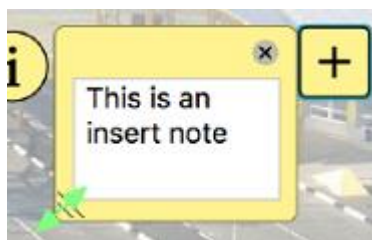
Note: The text color or text font of a Sticky note cannot be edited. It is not supported by the PDF format.

Audio Note

An **Audio annotation** is a short sound recording embedded in the file.

It appears as a **loud speaker icon** on the document's page .

To add an Audio record:

In the **Annotation pane**:

1. Select **Audio note** in the drop-down list of Annotation types
2. Click on the page where you want the icon of the audio note to be placed
3. Record your audio annotation by using the recording buttons
4. You can move the position of your audio comment when it is selected and the **move cursor** appears

Edition options:

- Icon **color**: **quick color selection** or **advanced color selection**
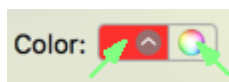
- **Recording buttons**:
  - **Record**: Click **Record** for the system to start the recording. The label says 'Recording...' and you can see the time elapsing.
  - **Stop**: Click **Stop** to stop the recording or the playback.
  - **Pause**: Click **Pause** to suspend the recording or the playback. Click *record*, *play* or *pause* button to resume the action.
  - **Play**: Click **Play** to play the recording. The label says 'Playing...' and you can see the duration time of your recording elapsing.
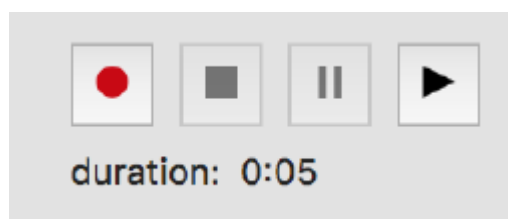
Recording buttons only
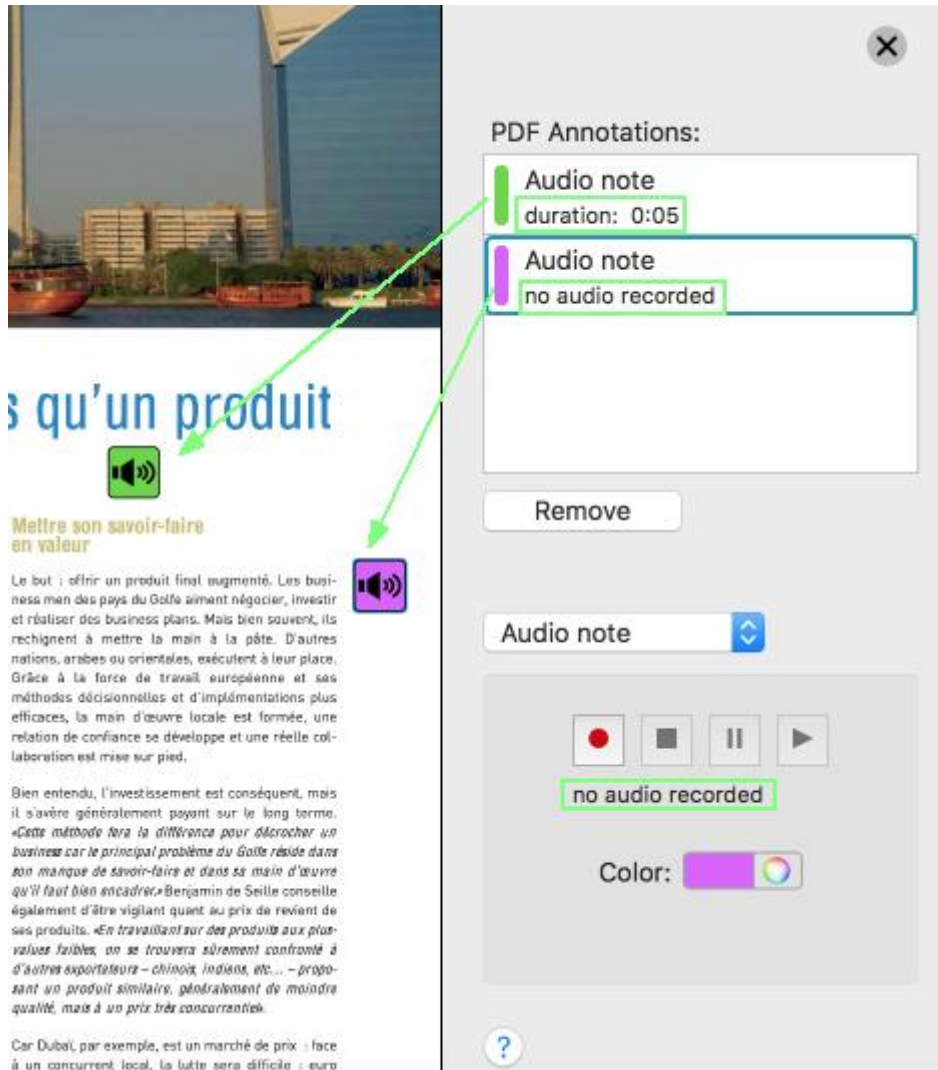
Full interface for Audio annotation

Default features of an Audio record:

- **Recording buttons** are enabled only when relevant.
- A **confirmation box** is displayed before overwriting an existing record.
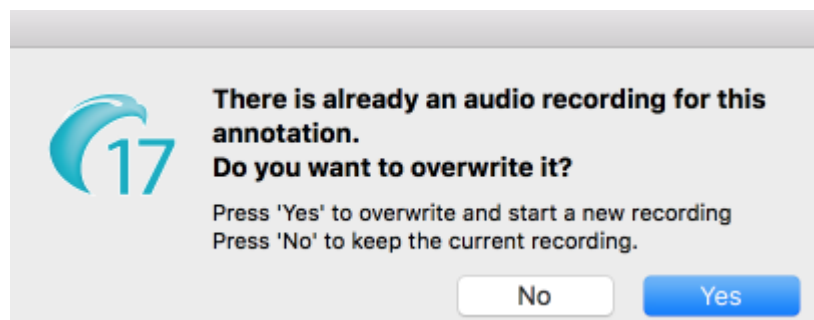
- The internal microphone (built-in) of your Mac is used to register the audio notes. Please refer to the **System preferences/Sound** if you wish to change the device for sound input. The **MAC help** can provide you more information.
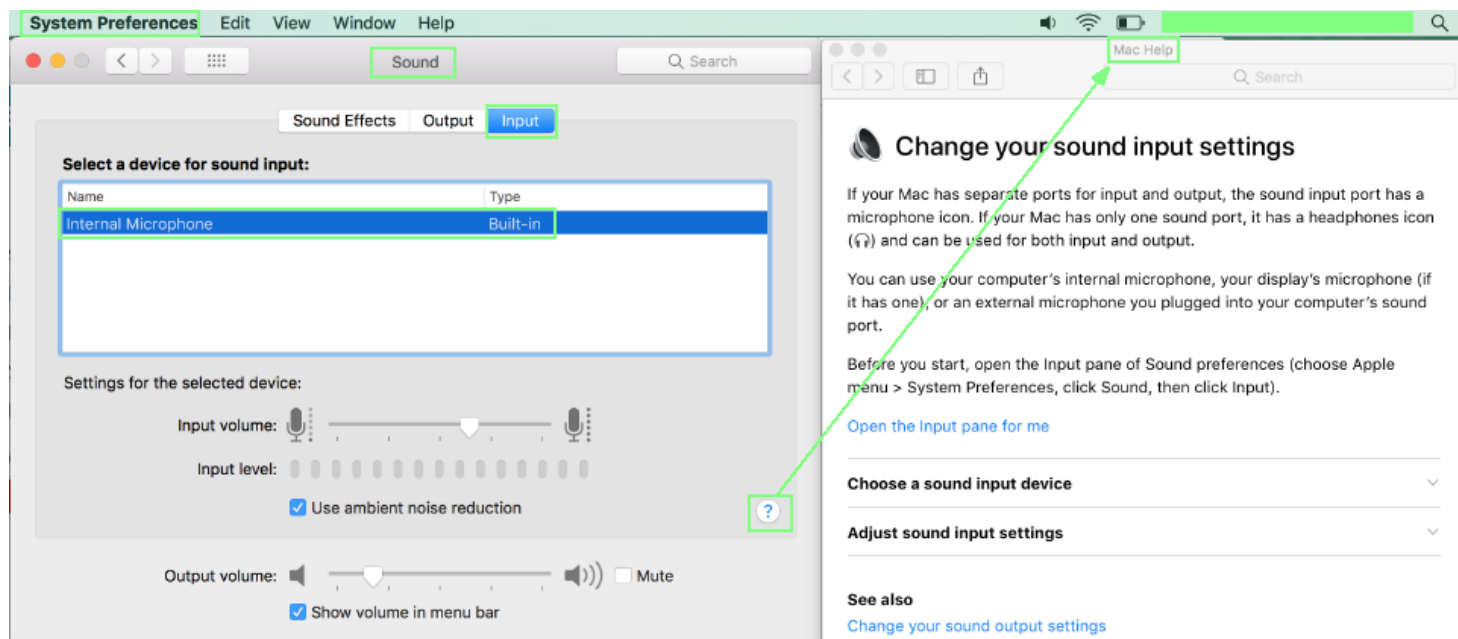
To play an Audio annotation in the PDF export (Acrobat Reader):

1. Click on the loud speaker icon. The **audio record** is played on the default audio output of your device.

Highlight, Underline or Strikethrough Text

These **text annotations** consist in adding a graphical effect on existing text.

To highlight, underline or strikethrough some existing Text:

In the **Annotation pane**:

1. Select one of the three types in the drop-down list of Annotation types
2. Create your highlight, underline or strikethrough by selecting the text into the image view
3. The Annotation is created as soon as the selection is done

Edition options:

- The **color** of the annotation: **quick color selection** or **advanced color selection**
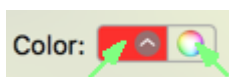
- Add a **Comment**: in some PDF viewers, this comment is shown when the mouse cursor runs over the marked text.

Note: For one same text selection, the type cannot be changed/edited after saving. It has to be removed and recreated.

File Attachment

A **File attachment annotation** embeds a file at a selected location in your document, so that after the exportation as PDF, the reader can open it for viewing. By attaching files as a comment, you can reference longer documents that cannot easily be pasted into a sticky note or text box. To view an attachment, the reader must have an application installed that can open the attachment.

To add a File Attachment:

In the **Annotation pane**:

1. Select **File attachment** in the drop-down list of Annotation types

2. Click on the page where you want the icon of an attached file to be displayed. The icon shows a paper clip

3. The file browser window opens. Select the file to attach.

4. The annotation is created. The name of the file attachment is displayed above the color palette.

5. You can move the position of the icon when it is selected and the **move cursor** appears

Edition options:

- Icon **Color**: **quick color selection** or **advanced color selection**
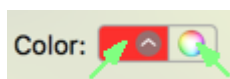
- Add a **Comment**: in some PDF viewers, this comment is shown when the mouse cursor runs over the icon.

Stamp

You apply a **stamp** to an e-document in much the same way you apply a rubber stamp to a paper document. You can choose from a list of predefined stamps, or you can create your own stamps.

To add a Stamp:

In the **Annotation pane**:

1. Select **Stamp** in the drop-down list of Annotation types. The most recently used stamp is selected.
2. Select a stamp from the list by clicking on the area showing the currently selected stamp.
3. On the page, the cursor takes the shape of a rubber stamp. Click where you want the stamp to be displayed.
4. You can move the position of your stamp when it is selected and the **move cursor** appears .

Edition options:

In the **Annotation pane**:

- Stamp **size**: slider from small to large
- Stamp **orientation**: you can choose among a list of predefined orientations: horizontal, 30 degrees, 45 degrees, 60 degrees or vertical

In the **Configuration window**:

1. Click on the button **Edit....** The **Configuration window** opens.

Configuration window

2. Select a stamp from the list on the left and choose your settings on the right:

- **First line of text**: free text

- **Second line of text**: free text. By default, some predefined stamps show the current date on the second line. The latter can be edited.

  TIP: to add the current date or current time to your stamp, you can type in those predefined variables:

  1. **$DATE**, **$DATE_LONG** and **$DATE_SHORT**: are replaced by the system date according to the date formats defined in the **System Preferences**
  o **$TIME**, **$TIME_LONG** and **$TIME_SHORT**: are replaced by the system time according to the time formats defined in the **System Preferences**

- **Font size of second line**: a slider controls the font size of the text on second line: it can be smaller or equal to the font size of the first line, but not larger.

- **Color**: **quick color selection** or **advanced color selection**
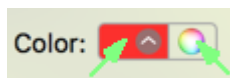
- **Border type**: 5 possibilities

  o No border

  o Rectangle

  o Rectangle with rounded corners

  o Arrow (to the left)

  o Arrow (to the right)

  Note: for the two Arrow types,the color choice will affect the stamp background, the text color being selected automatically.

- **Border width** (disabled if border type is 'no border'): from 0,1 to 20 pt.

Create a custom stamp or delete a stamp from the list:

A custom stamp can be defined at any time from the **Settings Menu** > **PDF Annotation Stamps...**

1. The **Configuration window** opens.

2. Click on the ➕ button at the bottom of the list to create a new custom stamp. Edit.

3. Or, select an existing stamp and click on the ➖ button to remove it from the list.

Default features of a Stamp:

- The stamps list is saved and remains available after Readiris is quit and restarted. The list is reset when [Restore Factory Settings](#) is invoked.

- Editing or deleting a stamp in the predefined list does not affect the annotations of the same type already in use.

- When a new stamp has been created, the font size of its second line and its border attributes are used as defaults for the next created stamp.

Section 10: Configure the Export

Three default **Export** buttons are available in the top toolbar to export your documents as **PDF**, **DOCX** and **XLSX.**

See below to change your favorite output formats.

To export documents directly:

- Click one of the **Export** buttons.
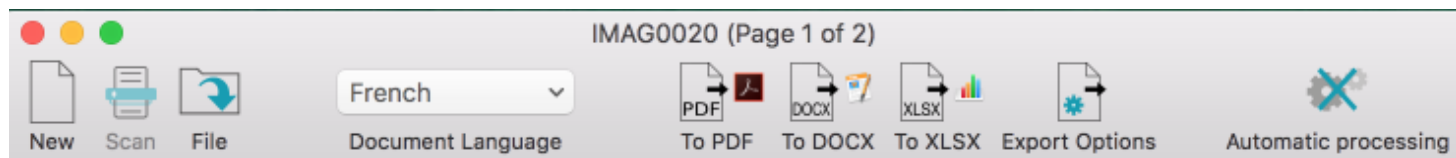
- Enter a name for the documents and choose the location where to save it.
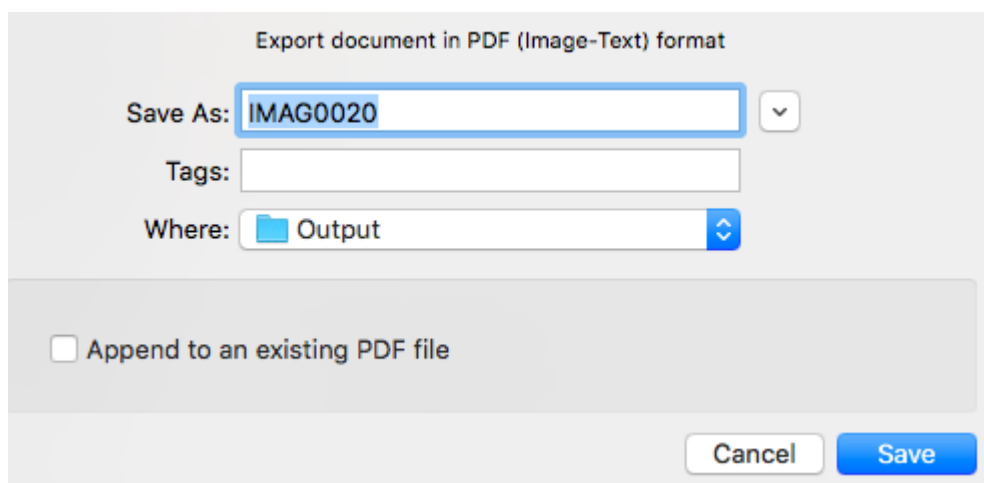
- When you select **Export as PDF** you have the possibility to add the current document to an existing PDF file.

    - To do so, select **Append to an existing PDF file**.

    - Then select the PDF file to which you want to add the current document.

- Click **Save**. The processed documents will be opened in the default application for the selected format.
  E.g. Preview for PDF, Pages for Word, etc.

To modify the default settings of an Export button:

- Click and hold an **Export** button for a few seconds until the **Options** menu appears.

- Move the cursor to the **Option** menu to open it.
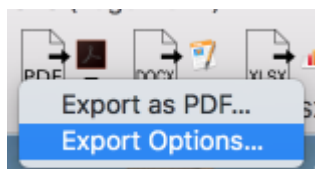
- Here you can change all related options and change the destination.

To select a different output format than the ones offered by the default Export buttons:

- Click **Export Options**.
  In this menu all available output formats are displayed.

- Drag the required Output Format from the **Others** list to the **Favorites** list. You can arrange the order of your favorites by dragging and dropping them where you want.
  💡 **TIP**: in the **Others** list, clicking on the [↑] button will set the format directly as your first favorite.

  See Select the Output Format for an overview.
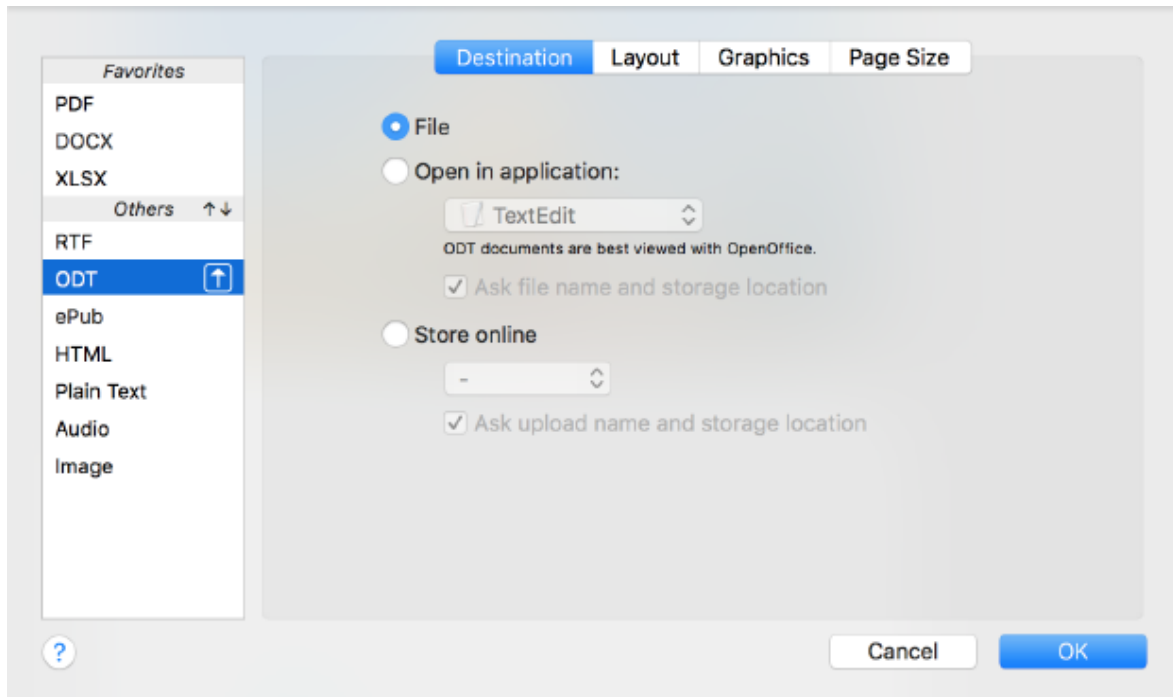
Starting a new project

When you are done exporting your documents and you want to start a new project, click **New** on the top toolbar.
This command erases your documents from Readiris.
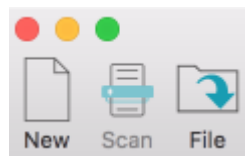
Select the Output Format

The documents you process with Readiris can be saved in a large number of output formats, ranging from plain text files to fully formatted Word processing and Spreadsheet files, and PDF files. With Readiris you can also turn Image PDFs - in which you cannot select or copy any text - into fully text-searchable PDF documents.

To access the available Output Formats:

- Click **Export Options** on the top toolbar.
- Drag the required output format from the **Others** list to the **Favorites** list.
- Click **OK** to save the settings.

In the following topics you find an overview of the supported output formats. For each of these output formats, specific layout options are available.

PDF

PDF stands for "Portable Document Format". It is a format developed by Adobe Systems Inc. to facilitate document exchange. PDF files are device-independent and resolution-independent, and can be opened on any Operating System (Windows, Mac OS, Linux, iOS, Android, etc.) PDF files are mostly used for storage purposes. Note that the content of PDF files cannot be edited (easily).
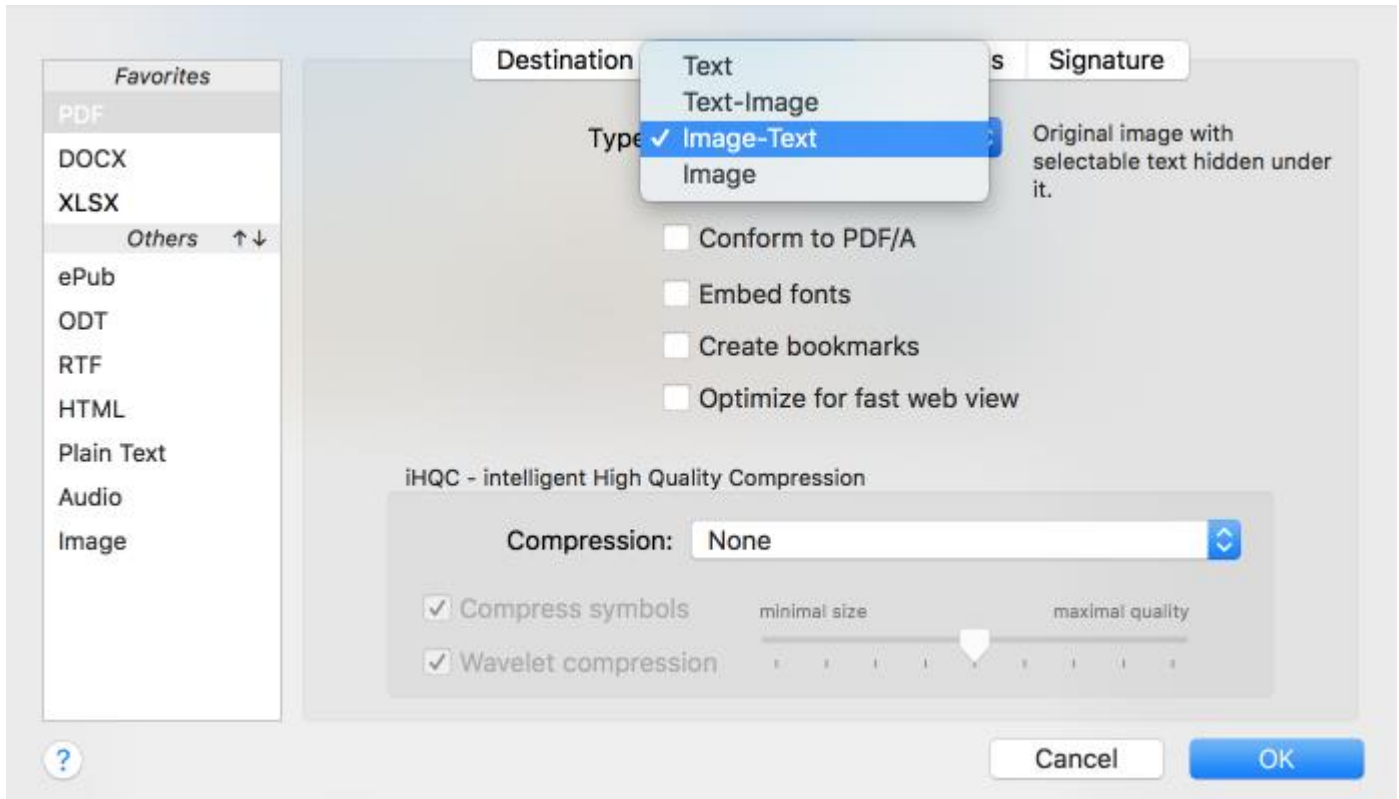
PDF Options

Type

With Readiris you can generate 4 types of PDF files:

- **PDF Image-Text**. This file type is most commonly used. It contains two layers: the recognized text, and the original image on top of the text. This way, you have both access to the recognized text and you still see the original image.

> **Note**: since the image covers the text, any recognition mistakes will not be visible.

- **PDF Text-Image**. This file type is the opposite of PDF Image-Text. It contains the original image in the background, and the recognized text on top of the image.

> **Note**: any recognition mistakes will be visible in this format.

- **PDF Text**. This file type contains the recognized text, but does not contain the original image of your document. Any images in the original document are included as graphics in the PDF file.
- **PDF Image**. When you select this file type, Readiris does not execute the text recognition on your document. Your PDF file will not be text-searchable, it only contains the image of your original document.

Conform to PDF/A *(Readiris Corporate only)*

It generates PDF files that are suited for long-term archiving.

Embed fonts

The option **Embed fonts**, embeds - or includes - the original fonts of your documents in the PDF file. This way, other users will always see the documents in their original fonts on their computer, even if they don't have the particular font you used installed on their computer.

Note that embedding fonts increases the file size of your output documents.

Create bookmarks

Bookmarks are types of links inside a PDF document, which contain representative information about the section they are linked to. Bookmarks give structure to your PDF documents. They are grouped in the **Bookmarks** panel in the **Navigation** pane.

In Readiris, the option **Create bookmarks** creates a bookmark for each text zone, image zone and table zone inside a PDF file.

Optimize for fast web view

This option restructures a PDF document for page-per-page downloading from web servers. This option is useful when accessing the documents you sent to the Cloud.

iHQC Compression

In Readiris you also have the possibility to hyper-compress your PDF documents with iHQC. iHQC stands for **intelligent High-Quality Compression**, I.R.I.S.' proprietary, efficient compression technology. iHQC is to images what MP3 is to music and what DivX is to movies.

iHQC is available for **Image-Text** and **Image PDFs**.

Select the compression type:

(Available in Readiris Pro and Corporate)

- **Good quality - best compatibility**: no use of wavelet compression (JPEG2000) and no use of symbols compression (JBIG2).

- **Good compression - best compatibility**: does the same as the previous option but with more aggressive compression factors for the background and foreground images.

(Available in Readiris Corporate only)

- **Better compression - good compatibility**: use of symbols compression (JBIG2) but no use of wavelet compression (JPEG2000).

- **Better compression - good quality**: use of symbols compression and wavelet compression with moderately aggressive compressions factors

- **Best compression**: use of symbols compression and wavelet compression with more aggressive compression factors.

- **Custom :** use of symbols compression, wavelet compression and compression factors are set freely.
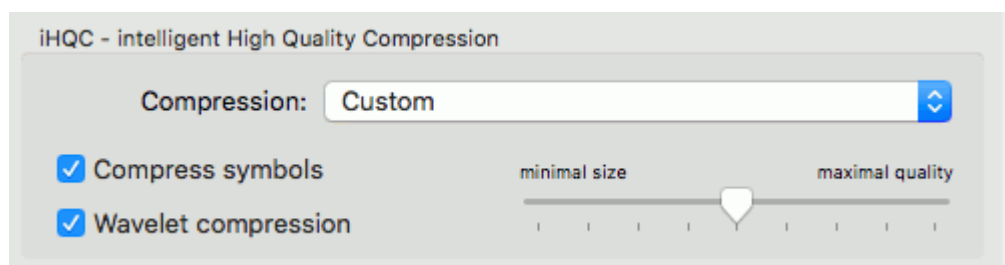
PDF Passwords

(This section applies to Readiris Corporate only)

The PDF documents you generate with Readiris can be password-protected. You can set a password to open the documents, and a password that restricts modifications to the documents.

- When you set an **open document password**, you will be prompted to enter that password when opening the PDF output.

- When you set a **permissions password**, you will only be able to perform the actions specified in the security settings. If you do want to change these settings, you must enter the permissions password.

    The Readiris security settings are similar to the standard protection features offered by Adobe Acrobat.

    Note that in Readiris the **open document password** and **permissions password** must be different.
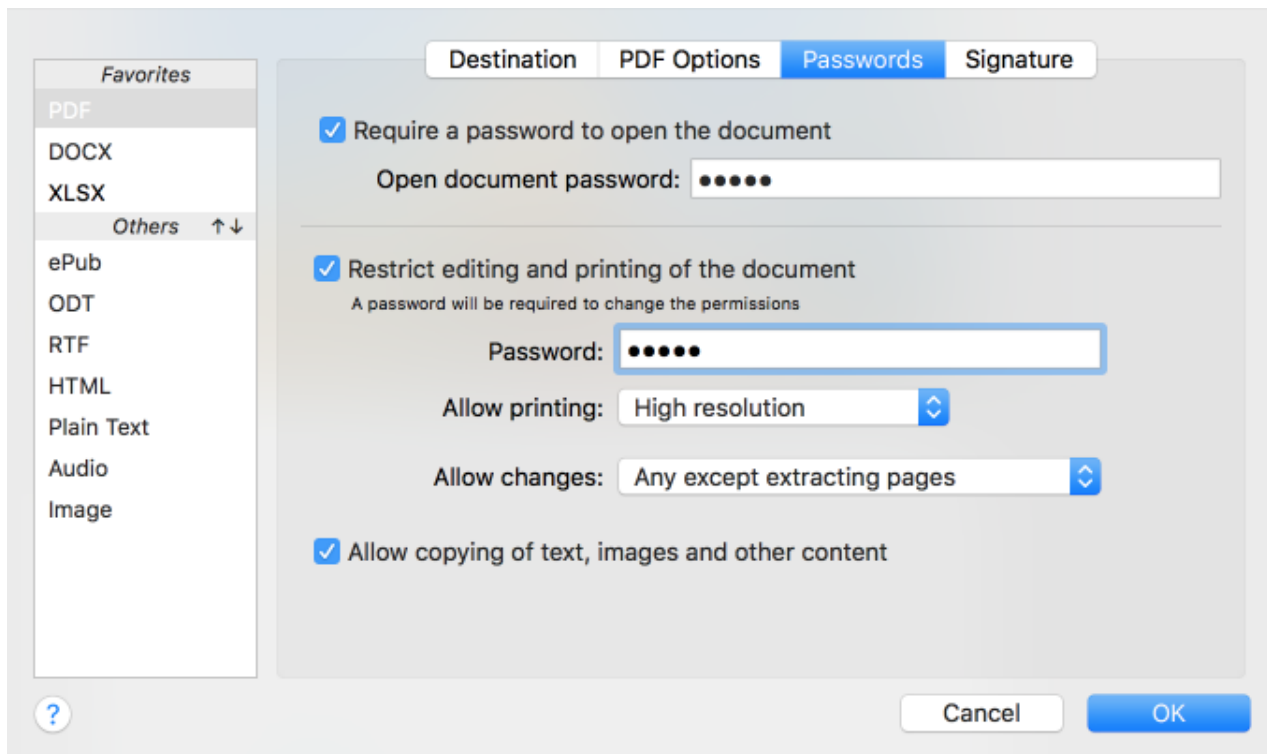
PDF Signature

(This section applies to Readiris Corporate only)

The PDF documents you generate with Readiris can be digitally signed. Digital signatures identify the person who created the PDF documents; they authenticate the identity of the author, certify a document and help prevent unwanted changes to PDF documents.

Note: Readiris does not generate digital signatures. You need to have a digital certificate at your disposal to generate signed documents with Readiris. You can create digital certificates with **Adobe Acrobat**, with the certificate assistant in the **Keychain Access**, or obtain them from companies like **VeriSign**.

To generate digitally signed documents:

- Click the **Signature** tab to access the signing options.

- Select **Sign with a digital identity** then browse for the certificate of your choice.

- Click **OK**

The digital signature will be included at next export as PDF.
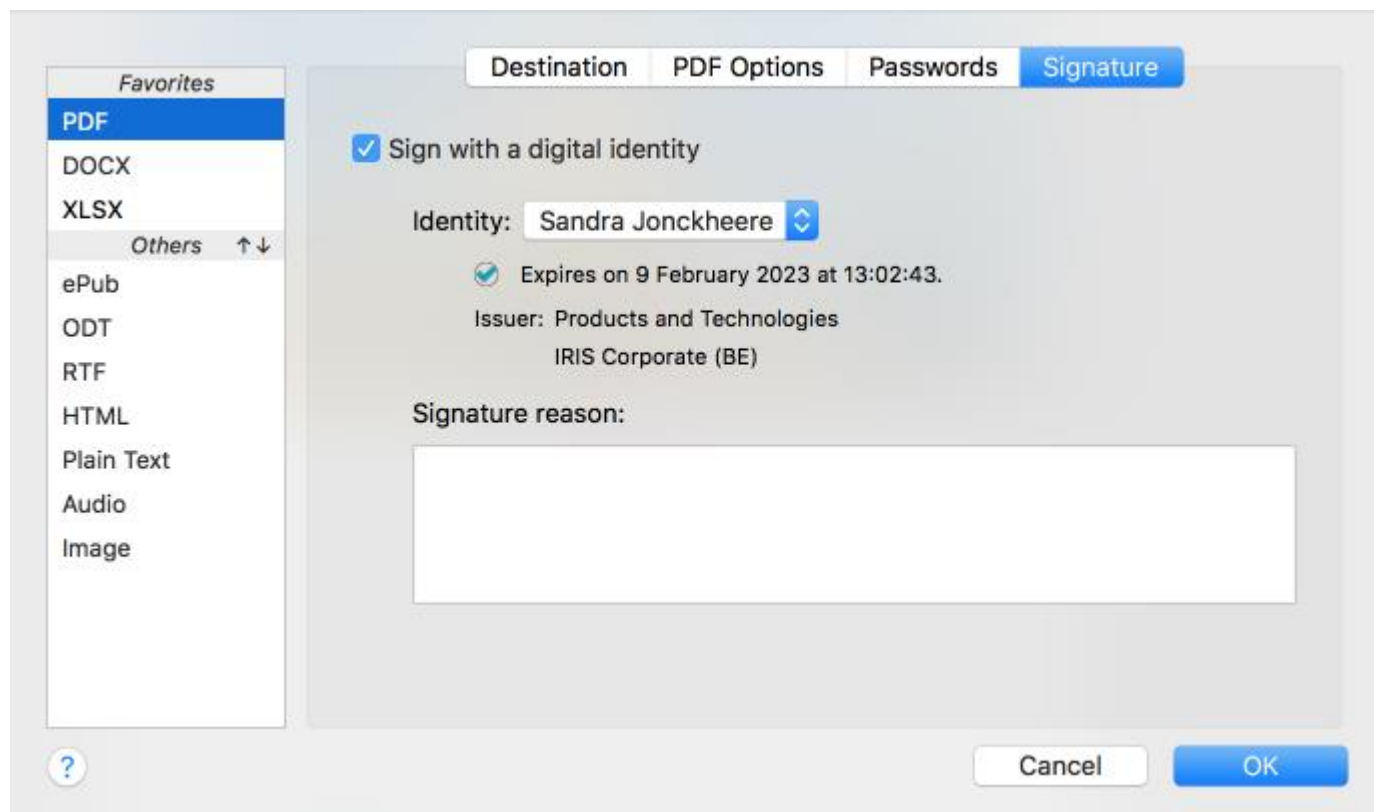
💡**Tip**: If you do not see any identity in the drop down list, make sure you have a valid certificate (digital identity) and its associated private key available in **Keychain Access**.

See the Adobe documentation for more information about certificates.

When you use a digital signature, it appears in the **Signature Panel** of Adobe Acrobat Reader DC.

DOCX, ODT, RTF

DOCX

DOCX is the standard text processor format used since Microsoft Word 2008. It is a standard format in several applications on the Microsoft Windows operating systems.

DOCX is also supported by Pages for Mac and DOCX with simplified layout is supported in TextEdit.

ODT

ODT stands for "Open DocumentText". It is an open-source file format.

ODT files can be opened with any OpenOffice-compatible text processor.

RTF

RTF stands for "Rich Text Formatting". It is a free document file format developed by Microsoft Inc. to facilitate document exchange.

Use the RTF format when you do not have the possibility to use the DOCX or ODT formats. It is recommended to use Microsoft Word to open RTF documents generated by Readiris.

Layout Options

The **Layout Options** play an important role when generating Word processor documents (.docx, .rtf, .odt). They determine for a large part what your final documents will look like.

Below you find a description of what each Layout option does and examples of real documents.

> **Note**: if any Layout option is unavailable for the output format you selected, this means the option in question is not supported.

- The option **Create body text** generates a continuous, running line of text. The result is a document without any formatting. The formatting is to be done manually, by the user, in another application.

Example of body text
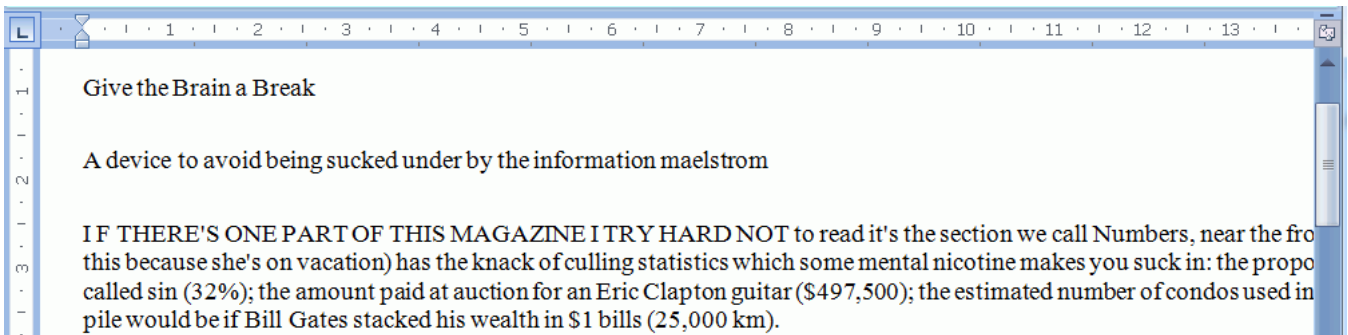
- The option **Retain word and paragraph formatting** keeps the general format structure of your scanned document.

  The font type, size and type style are maintained across the recognition process.

  The tabs and the alignment of each block are recreated.

  The text blocks and columns aren't recreated; the paragraphs just follow each other.

  Tables are recaptured correctly.

Example of word and paragraph formatting
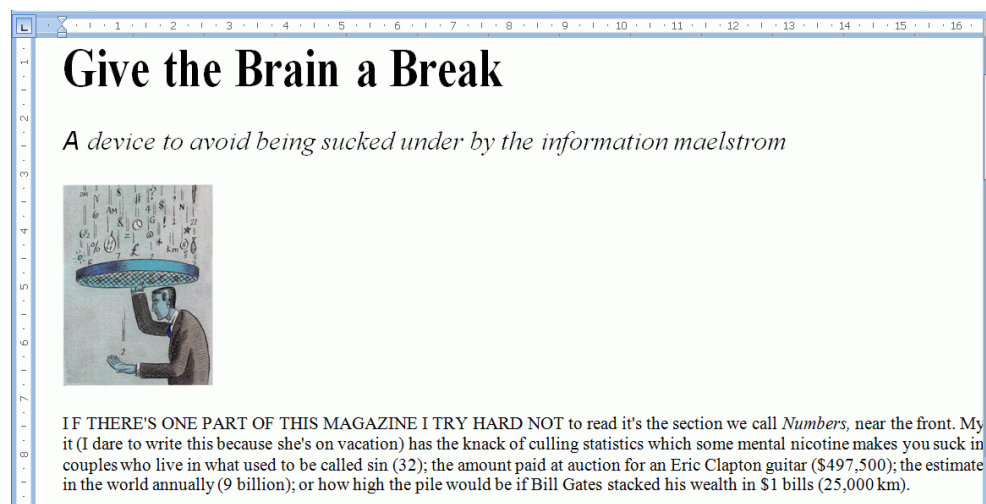
- The option **Recreate source document** tries to stay as close to the original layout as possible.

  The text blocks, tables and pictures are re-created in the same place as the original.

  The word and paragraph formatting are maintained.

  Hyperlinks are re-created too.

Example of a re-created source document

- The option **Use columns instead of frames** creates columns instead of text frames to position the information on the page.

    Columnized texts are easier to edit than documents containing multiple frames: the text flows naturally from one column to the next.

    Note: should Readiris be unable to detect columns in the source document, then frames are still used as substitute solution.

    💡 **Tip**: use this option when generating Word documents.

- The option **Insert column breaks** inserts a hard column break at the end of each column, which maintains the format of the text. Any text you edit, add or remove, remains inside its column; no text flows automatically across a column break.

> 💡 **Tip**: disable this option (**Insert column breaks**) when you have body text that contains columns. You'll ensure the natural flow of the text from one column to the next.

- The option **Add image as page background** places the scanned image as page background beneath the recognized text.

> **Note**: this option increases the file size of the output files substantially.

    The format **PDF Text-Image** modifies PDF files in the same manner.

    To avoid the file increase discussed above, select the option **Maintain colors of background**. This provides a similar but less drastic, more compact alternative.

- The option **Merge lines into paragraph**s enables automatic paragraph detection.

    Readiris wordwraps the recognized text until a new paragraph starts, and reglues hyphenated words at the end of a line.

- The option **Include graphics** includes images in your output documents.

    Clear this option in case you only need lay-outed text.

- The option **Maintain colors of text** keeps the original color of the text in your source documents.

- The option **Maintain colors of background** recreates the background color of each document.

Graphics Options

The **Graphics** Options determine how the images will be presented in the output documents, e.g. in black-and-white or in color. They also define their resolution.

> **Note**: when you select an output format or application that does not support images, then the Graphics Options will be unavailable in Readiris.

Below you find a description of what each option does.

Color type

Readiris saves graphics in their **original** depth by default.

Readiris can also save graphics in **black-and-white**, **grayscale** and **color**.

Image quality

You can choose between Low, Normal and High quality graphics.

Resolution

Readiris retains the **original** resolution by default.

You can also choose to reduce the resolution to a lower **dpi** (between 50 and 300).

Note that you cannot increase the resolution.

Page Size Options

In Readiris the page size of the documents you scan and open does not necessarily have to be the same as the page size of your output documents.

When generating DOCX, XLSX, ODT, RTF and HTML documents, you can select the preferred page sizes:
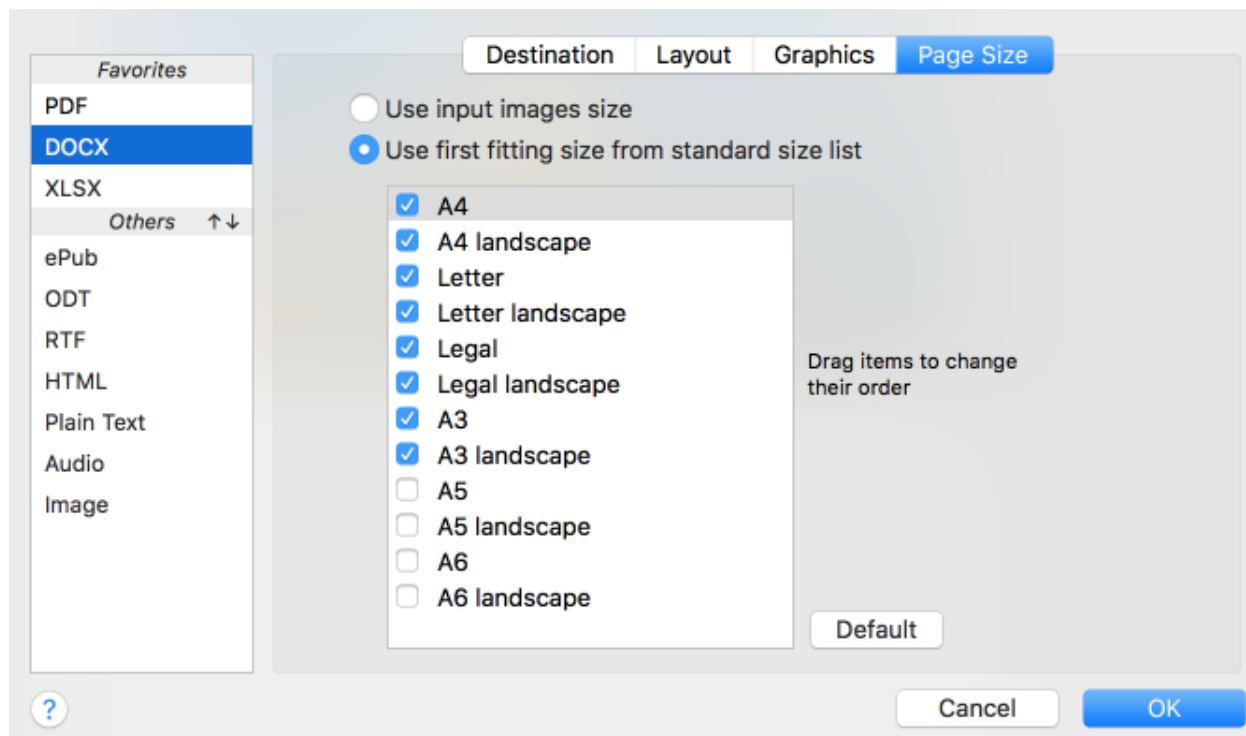
- Select **Use input images size** to generate output documents that have the exact same size as the original files.

- If you want to use a **standard page size** - in case you want to print your documents for instance - select **Use first fitting size from standard size list**.

- Then select your preferred page sizes.

- Now sort the selected page sizes in the order of your preference.
  Readiris tries to use the first page size in the list. If your document is too big for the page size, Readiris tries to second page size in the list, and so on.
  So make sure you drag your preferred page sizes to the top of the list.

- When you are done, click **OK** to save and close the settings.

    Tip: to restore the default settings, click **Default**.

ePub

ePub is a free and open e-book standard by the International Digital Publishing Forum (IDPF). EPUB is designed for reflowable content, meaning that the presentation of the content automatically adapts itself to the device on which it is opened. EPUB also supports fixed-layout content.

Note that Readiris only creates body text in ePub files. Images are not included.

Text Options

The option **Merge lines into paragraphs** is selected by default. Readiris wordwraps the recognized text until a new paragraph starts, and reglues hyphenated words at the end of a line.
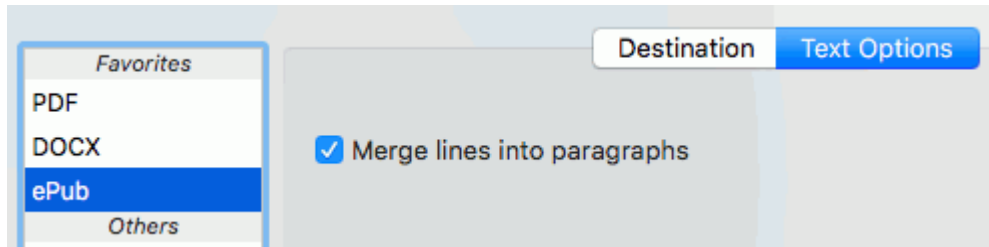
HTML

HTML stands for "Hypertext Markup Language". It is the predominant markup language for web pages. It provides a means to describe the structure and formatting of text-based information in a document. This file format can be opened in Microsoft Excel, in Web browsers such as Safari, and in Web page editors such as Adobe Dreamweaver.

Note: HTML is the recommended format when saving documents to Evernote.

Layout and Graphics Options

The same **Layout** and **Graphics** options are available as for <u>DOCX, ODT, RTF</u>.

XLSX

XLSX is the standard spreadsheet file format used since Microsoft Excel 2008. XLSX files are created using the Open XML standard. Each cell in an XLSX file can have a different formatting.

Layout Options

Worksheet

- The option **Create one worksheet per page** sees to it that one worksheet is created per scanned page.

  If a page contains tables and text, all is placed inside the same worksheet.

- The option **Create one worksheet per table** places each table in a separate worksheet and includes the recognized text (outside the tables) in another worksheet.

 If the document being processed contains more than one page, each page will be processed in the same manner.

  This option is useful when processing tables of different sizes and different headings.

Other options

- The option **Retain word style and text paragraphs** keeps the general format structure of your scanned document.

  The font type, size and type style are maintained across the recognition process.

  The tabs and the alignment of each block are recreated.

  The text blocks and columns aren't recreated; the paragraphs just follow each other.

  Tables are recaptured correctly.

  Pictures are not captured.

- The option **Ignore all text outside the tables** saves the tables and ignores all other recognition results.

  All data inside the tables is recaptured; any data outside the tables is not.

Example image
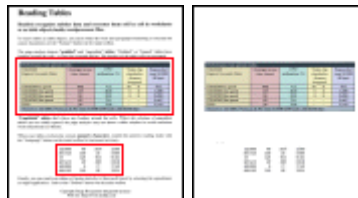
- **The option Convert figures into numbers encodes recognized figures as numbers.**

  As a result, you can execute arithmetical operations on those cells. The text cells (in any table) remain text.

  Note that only figures inside tables are encoded as numbers.

- **The option Merge lines into paragraphs enables automatic paragraph detection.**

  Readiris wordwraps the recognized text until a new paragraph starts, and reglues hyphenated words at the end of a line.

- The option **Maintain colors of text** keeps the original color of the text in your source documents.

- The option **Maintain colors of background** recreates the background color of each document.

  Note: in a spreadsheet application, this option recreates the background color of each cell.

Page Size

The same **Page Size** options are available as for [DOCX, ODT, RTF](#).

Plain Text

Plain Text generates a standard text document that contains unformatted text. Plain Text files can be opened in any text editing or word processing application.

Text Options

The option **Merge lines into paragraphs** is selected by default. Readiris wordwraps the recognized text until a new paragraph starts, and reglues hyphenated words at the end of a line.

Audio

Audio .aiff is an audio file format standard for storing sound data. Save your documents in .aiff to have the text read out to you.

Text to Speech Options

- Click the **Voice** list and select which voice should read out the text.

  Voices marked with (*) must be downloaded before use.

- Move the **Speaking rate** slider to adjust the speed.
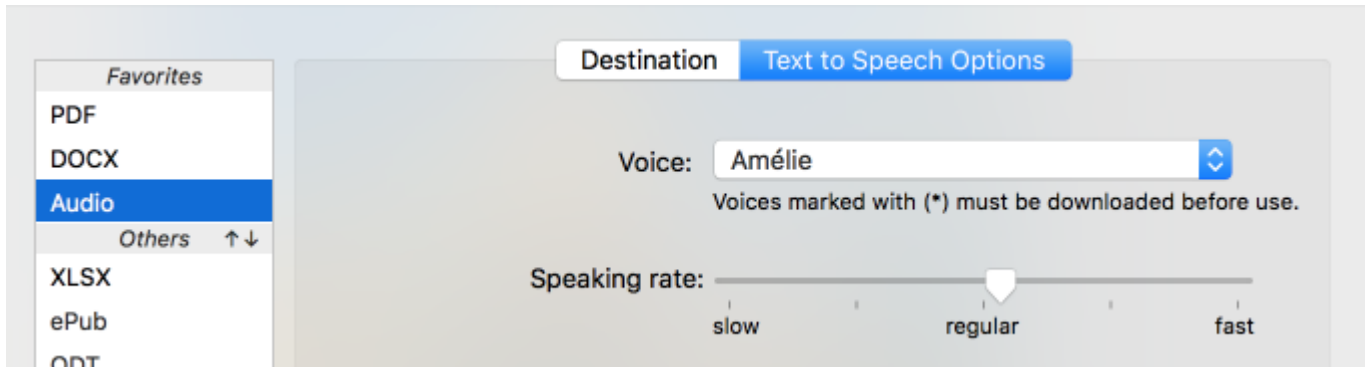
💡 **Tip:** to have the output read out to you directly, click the **Destination** tab, select **Open in application** and select a supported application.

Image

Although Readiris is a text recognition application, it can also save documents as image files.

Image Options

The following image file formats are supported as output: TIFF, PNG and JPEG.

Other ways of saving documents as image files

- Ctrl-click a Graphic zone in your documents, and select **Copy as Graphic**.
  This command copies the image to the Clipboard.

- The image can now be pasted into another application.
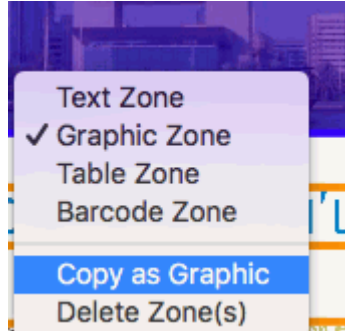
- Or drag a thumbnail from the Pages panel or Thumbnail view to the Desktop to save the entire page as JPEG image.

Select the Destination

The documents you process with Readiris can be send to different destinations. You can save them locally as files, open them in supported applications and send them to the online Storage Systems Evernote, Dropbox, Google Drive, Microsoft OneDrive and Box. Readiris Corporate also exports documents to Microsoft SharePoint.

To access the Destination Options:

- Click and hold one of the **Export** buttons for a few seconds until the **Options** menu appears.

> **Note**: if you want to use a different Output Format, click **Export Options** on the top toolbar.

- Move the cursor to the **Option** menu to open it.

- Then click the **Destination** tab.
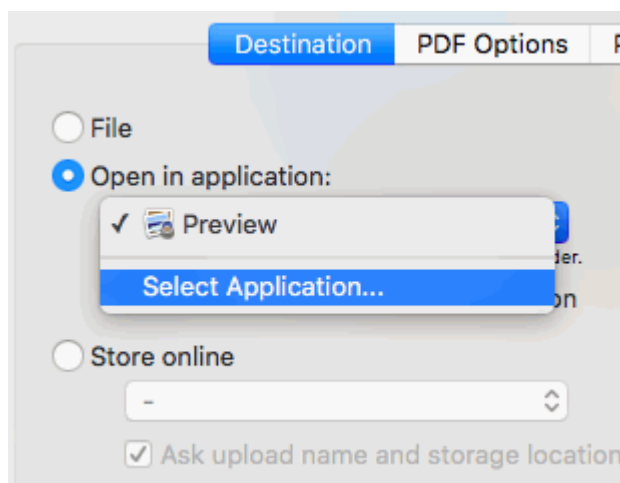
- Select **File** to save the documents locally.

  In this case the documents are simply stored on your computer and you can export and share them later.

  The pages that are opened in Readiris are saved into a single output file, except when you select an image file as output format. In that case Readiris saves them as separate image files.
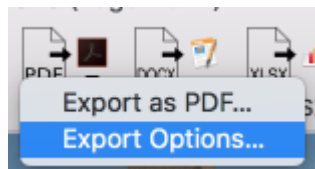
- Select **Open in application** to open the documents immediately after export.

  The default application on your Mac for the selected output format is selected. To choose a different application, click inside the list and click **Select Application**.

- Select **Store online** to send your documents to a configured online Storage System.
  Note that Internet access is required.

  - Select the required Storage System from the list.
    If there aren't any Storage Systems in the list this means you haven't configured them yet.

  - Click **Accounts** to configure a new online Storage System. See Configure an online Storage System to learn how to do so.

- When you are done, click **OK**.

Configure an online Storage System

Important: before you can configure an online Storage System in Readiris your first need to have a valid account for that Storage System.

Configuring an online Storage System can either be done via **Export Options** > **Destination** tab > **Store online** > **Accounts.**
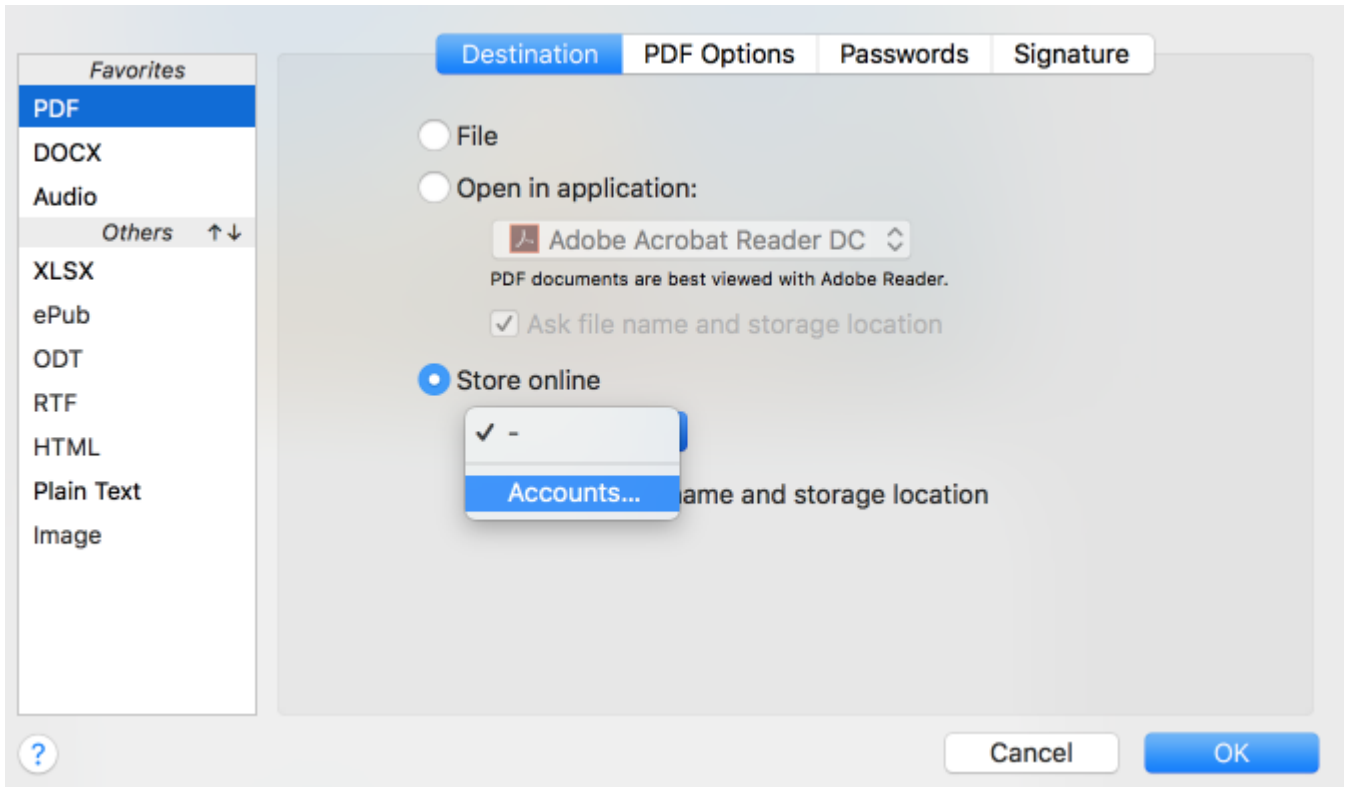
Or via the **Readiris** menu > **Preferences** > **Accounts**.

To configure an account:

- Click the plus sign at the bottom of the window.
- Select an account from the **Account type** list.
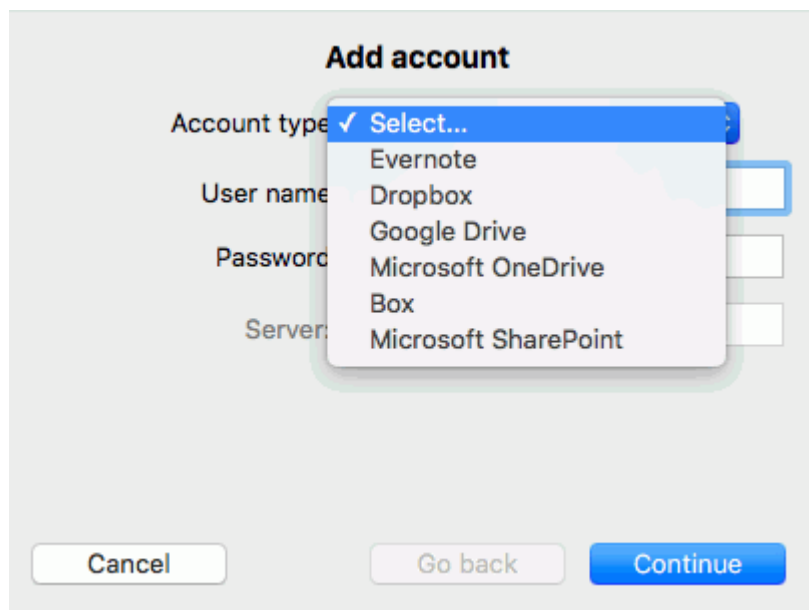
Google Drive, Dropbox, OneDrive, Evernote and Box

- Enter your **Email** and **Password**, then click **Sign in**.
- Select the option that allows Readiris to access your account.
- Select the folder to which you want to export documents, and then click **Finish**.

Note that you can add multiple accounts of each type.
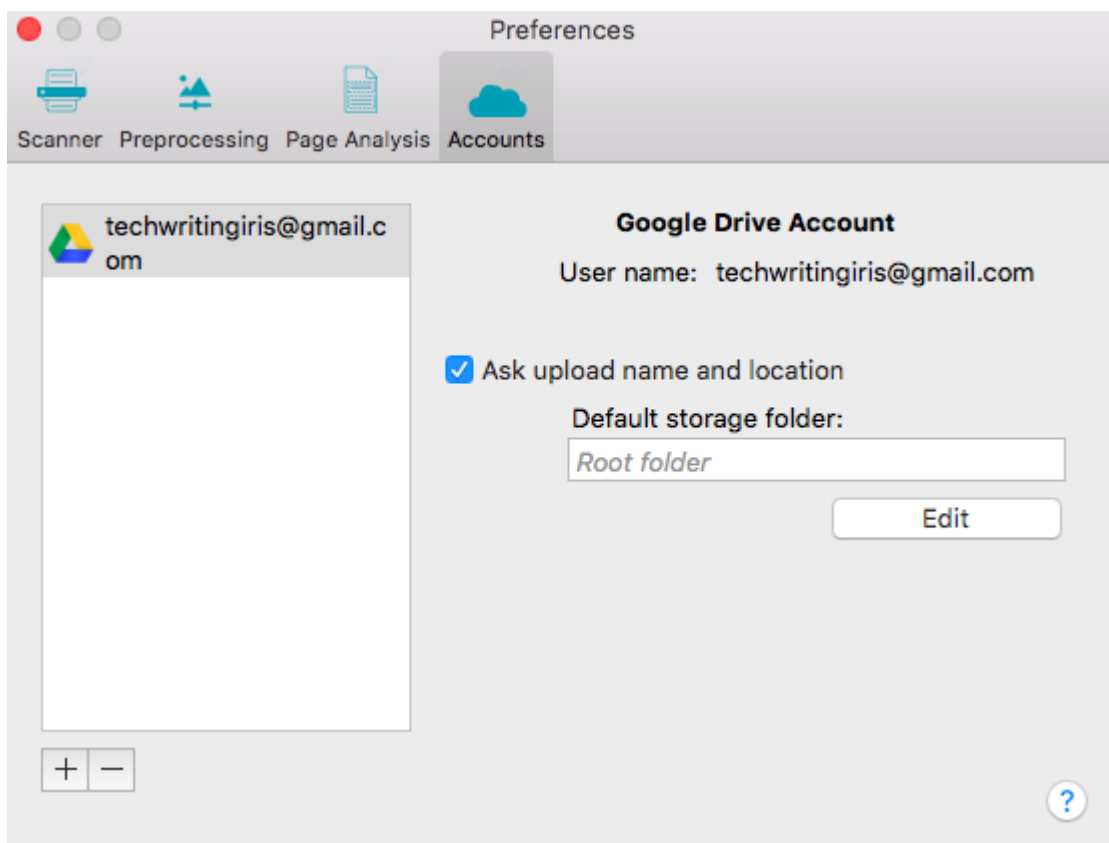
Once an account has been configured, the option **Ask upload name and location** is selected by default. If you keep this option selected, you will be prompted to enter a file name and an export location each time you export a document. To skip this step, unselect this option. Readiris will now use the default export location you configured and it will preserve the original name of your document.

To change the default export location, click **Edit** and select a different location.
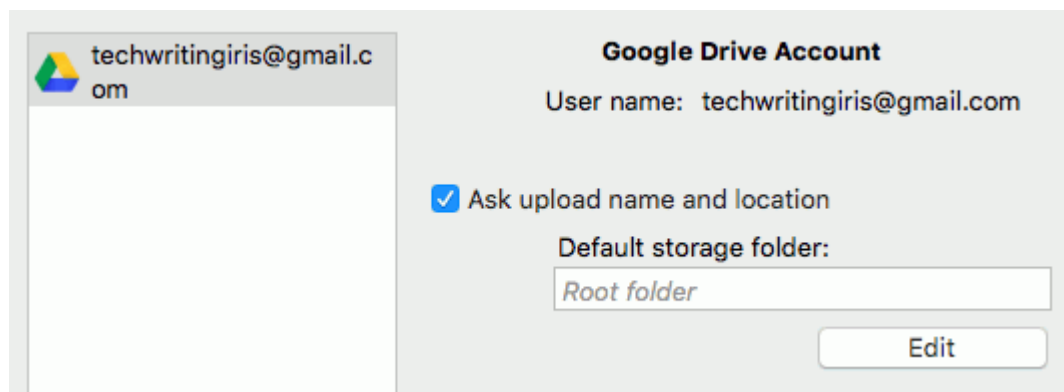
When you have configured an Evernote account, the option **Synchronize after upload** is also selected by default. If you keep this option selected, and if the Evernote application has been installed and correctly configured on your Mac, Readiris sends a synchronization command to Evernote after it has uploaded one or more documents to the Evernote servers.

Microsoft SharePoint

- Enter your **Username** and **Password**.
- Enter the server to connect to.
- Then click **Continue**.
- Select the required folder to export to.
- If necessary, select **Content Type**.

Removing an online Storage System

To remove an online Storage System from the Readiris configuration:

- Access your configured accounts as described above.
- Select the account you want to remove.
- Then click the minus sign at the bottom of the window.

Section 11: Processing Batches of Documents

Batch Processing

(This section applies to Readiris Corporate only)

When you are processing entire batches of image files you can use Readiris' **Batch Processing** folder. Batch Processing executes the recognition on all scanned images in a specific folder. Indicate to Readiris in which folder your documents are located, start the processing and all your documents will be converted to the required output format.

To start batch processing:

- On the **Process** menu, click **Batch Processing**.
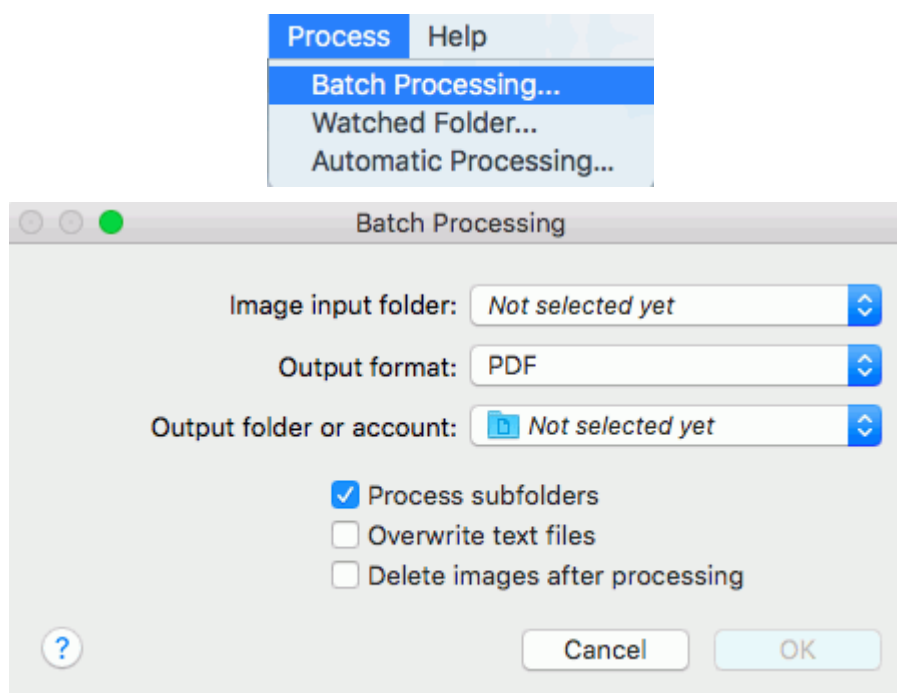
- Select an **Image input folder**.

  This is the folder from which you scan the documents.

- Select an **Output format**.

  See the topic Select the Output Format for more information.

- Select an **Output folder or account**.

  This is the folder or account where your processed documents will be stored.

  Note: to select a Cloud account, you must first configure it in **Readiris** > **Preferences** > **Accounts**. See the topic Select the Destination for more information.

> **Important note**: do not select an output folder that is also being used as Watched Folder.

Batch Processing Options

- In case you want to process the underlying subfolders of the Input folder, select **Process subfolders**.

- If you want Readiris to overwrite existing output files in the Text output folder, select **Overwrite text files**.

- If you want Readiris to delete the original images from the Input after processing, select **Delete images after processing**.

When you are done selecting the options, click OK for Readiris to start Batch processing the folder.

Readiris processes the images of all supported file formats. You cannot limit the recognition to files of a specific file format.

The recognized documents get the same file name as the original image files.

A log file is created per batch, containing the processing date and the document names and paths.

Watched Folder

(This section applies to Readiris Corporate only)

If you need to scan entire batches of documents, on a scanner equipped with a document feeder, for instance, you can use Readiris' Watched Folder feature. This way, you can scan all your documents to a specific folder that is being monitored by Readiris. Whenever Readiris spots a new document inside the Watched Folder, it processes it and sends it to an output folder.

To set up a Watched Folder:

- On the **Process** menu, click **Watched Folder**.
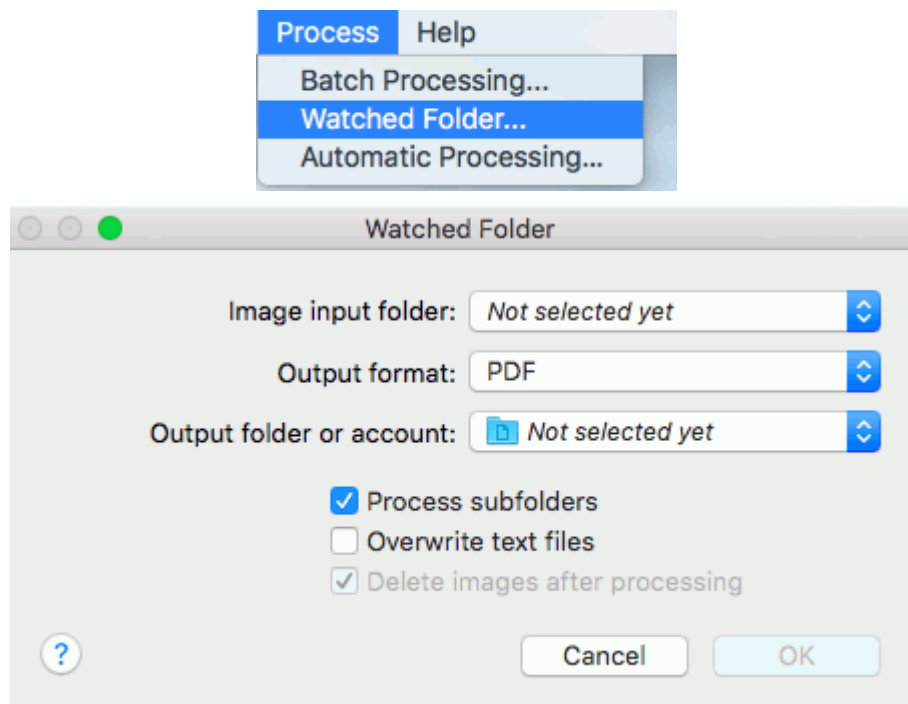
- Select an **Image input folder**.

  This is the folder from which you scan the documents.

- Select an **Output format**.

  See the topic Select the Output Format for more information.

- Select an **Output folder or account**.

  This is the folder or account where your processed documents will be stored.

  Note: to select a Cloud account, you must first configure it in **Readiris** > **Preferences** > **Accounts**. See the topic Select the Destination for more information.

Important note: do not select an output folder that is also being used as Watched Folder.

Watched Folder Options

- In case you want to process the underlying subfolders of the Watched Folder, select **Process subfolders**.

- If you want Readiris to overwrite existing output files in the Text output folder, select **Overwrite text files**.

- If you want Readiris to delete the original images from the Watched Folder after processing, select **Delete images after processing**.

When you are done selecting the options, click OK for Readiris to start monitoring the Watched Folder.

Readiris will process any images it spots inside the folder.
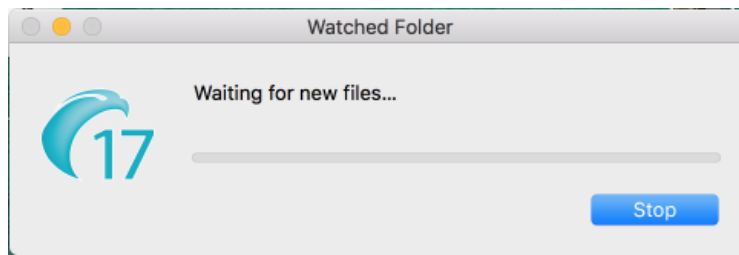
The monitoring continues even if you quit Readiris.

To add documents to the Watched Folder while it is being monitored by Readiris:

- Scan documents to the Watched Folder, or
- Move existing image files to the Watched Folder.

How to

How to configure your scanner in Readiris

Readiris supports all **Twain 1.9** compliant scanners, all **Image Capture** compliant scanners, and **IRIScan 3** and higher scanners.

Most Image Capture scanners are plug and play. Before you can use a Twain scanner, however, its drivers need to be installed on your Mac.

Twain scanners

- Connect your scanner to your Mac and power it on.

- Install your scanner's Twain driver.

    In general, drivers can be found on the CD-ROM provided with your scanner and on the website of your scanner manufacturer. They are not provided by I.R.I.S. Note that some scanner drivers may not work under the latest versions of Mac OS. See the documentation supplied with your scanner to find out which platforms are supported. If necessary, contact your scanner manufacturer.

IRIScan scanners

- Connect your *IRIScan scanner to your Mac and power it on.

- Install your scanner's Twain driver.

    *Note that you do not need to install any drivers when using an **IRISCard Anywhere, IRIScan Anywhere** or **IRIScan Book** scanner. They are not Twain driver-based scanners that can be used from within any application to scan documents. Instead, they are designed to be used on their own, disconnected from any computer. The documents you scan are stored in the scanner's internal memory (or optional SD card / USB flash drive). From there they can be loaded directly into Readiris when you connect the scanner or insert the SD card or USB flash drive into your Mac.

To check if the scanner driver has been installed successfully:

- In the **Readiris** menu, click **Preferences**.

- Click the **Scanner** tab.

- You should see your scanner listed.

If your scanner is not listed, its driver has not been installed successfully.

- Once the scanner has been configured you can start scanning, using the default scanner settings.

# How to recognize barcodes

Readiris supports most widespread 2D barcodes, both laser printed and inkjet printed.

In **Readiris Corporate**, barcodes can also be used to separate documents in a batch. Note that in this case the content of the barcodes is not converted into text, the barcodes are only used to do the separation.

Recognizing barcodes manually

- Scan or Open your document.

- Click the **Layout** icon in the Image and Layout Editing toolbar.
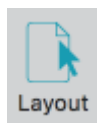
- Then click the **Barcode** icon in the menu that appears, and draw a frame around the barcode zone.

- Close the Layout toolbar and process your document. The content of the barcode will be processed as well.

Recognizing a particular barcode

In case you need to use the content of a particular barcode:

- Repeat the previous steps to draw a barcode zone.

- Ctrl-click the barcode zone and select **Copy as data**.

    The recognized content of the barcode is copied to the Clipboard. From there you can paste it into other applications.

Recognizing barcodes to separate document batches

(This section applies to Readiris Corporate only)

Barcodes can be used as separators to separate documents in a document batch. Readiris can automatically look for barcode pages and mark them as cover page, indicating the beginning of a new document.

- On the **Settings** menu click **Document Separation and Indexing**.

- Select **Detect cover pages with a barcode**.

    If necessary, indicate specific content Readiris should look for. For more information see the section How to separate document batches.

    Note: the barcode reading results can also be included in an XML index. Select the option **Generate an XML index** and check the box **Include text of cover pages in index**.

- Click **OK** to save the settings. Then process your documents.

How to obtain good recognition results

When an image has been analyzed in Readiris and its recognition accuracy is not sufficient, a warning sign is displayed both in the Page thumbnail and below the top toolbar.
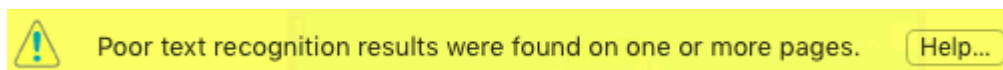
There might be various causes for unsatisfactory recognition results, since the recognition depends on a number of factors. In this topic we list the factors that influence the recognition.

- The document language

  Make sure you select the right language in the Language list. See the section Selecting the Document Language for more information.

- The quality of the scanned documents

  - If the scanned documents are too bright, or too dark, Readiris will have trouble recognizing them. Luckily there is a feature that allows you to adjust the quality of scanned documents. See the section Adjusting the Image Quality to learn how to do so.

  - If digital camera images are out of focus, or were made in bad lighting conditions (lack of brightness, reflections, weak contrast, etc.), Readiris will also have trouble recognizing them. See the sections Opening Digital Camera Images and Optimizing the Scanned Documents > Deskewing Digital Camera Images to learn how to process digital camera images correctly.

- The scanner settings

  If you are using a scanner to scan documents, the image quality is also linked to the scanner settings:

  - Make sure the scan resolution is set to 300 dpi for regular documents and to 400 dpi for small print and Asian languages.

  - Make sure documents are scanned so the text is in correct orientation. If necessary, **rotate** their images in Readiris.

  - Make sure the documents are not skewed.

  See the section Scanning Documents, Preprocessing Options and Page Analysis Options to make sure the right settings have been selected for your scanner.

Please note that an image of very poor quality won't be correctly recognized even after verification of the above points.

How to process digital camera images

See Opening Digital Camera Images and Deskewing Digital Camera Images.

How to save the contents of a particular zone

Besides saving entire documents, you can also save the contents of particular zones (available in **Layout** mode or in default mode):

- **Ctrl**-click the zone you want to recognize
- Then click **Copy as ...**
- The results are copied to the clipboard. You can now **paste** them into other applications.

Note that you can also send content directly to Evernote, if the connection is configured correctly.
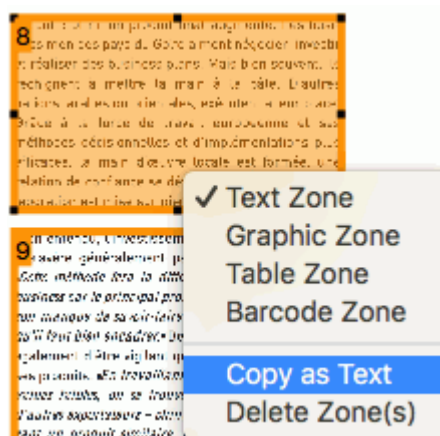
How to separate documents

(This section applies to Readiris Corporate only)

When you are scanning entire stacks of paper - on a scanner equipped with a document feeder for instance - or loading many image files into Readiris, it can be useful to indicate where one document ends and another begins. That way, Readiris can generate a separate output file for each document. This avoids your having to scan only one document at a time.

Example

Let's say you are scanning a stack of 30 pages. The first 10 pages belong to Document 1, pages 11 to 20 belong to Document 2 and pages 21 to 30 belong to Document 3. If you indicate that your stack of paper in fact contains 3 documents, Readiris will generate 3 separate output files.

How to separate paper documents

With Readiris, there are 2 ways to separate paper documents. You can do this by means of:

- **blank pages**

  In this case you simply put a blank page between two sheets of paper to indicate where the new document begins. Note that you can delete all blank pages simultaneously after recognition should this be necessary: click the command **Delete Blank Pages** on the **Process** menu to do so.

- **cover pages containing a barcode**

  In this case you can, for instance, put a barcode sticker on each first page of a document.

Note: you could also skip this step and separate the scanned images manually in Readiris. See How to separate image files below.

To access the document separation options:

- Click the **Settings** menu and click **Document Separation and Indexing**.

- Indicate which separation method you are going to use: **Detect blank pages** or **Detect cover pages with a barcode**.

  When you are using barcode pages as cover page, you can indicate specific data your barcodes should **contain** in order for Readiris to consider them to be barcode pages. Insert your company name for instance, **I.R.I.S.** in our case, in the field **containing**. Only barcodes that contain the data **'I.R.I.S.'** will be marked as cover pages and will be used to split up your document batch into separate documents. You can also add a variable part to the data, for instance the scanning date. This variable part will indicate the specific indexing data of each individual document.

- In case you want Readiris to generate an XML index file containing detailed information on the processed documents, select the option **Generate an XML index**.

  An XML index file will be created per document. The index file contains detailed information such as the detected barcode separator, the page range, the output file name and the cover page text (if selected).

  To include the text of the cover pages in the XML index, select the corresponding option. Note that these reading results are not included in the output document.

- If your cover pages also contain content you want to include in the processed output documents, select **Recognize cover pages**.

- Then click **OK** to save the settings and start processing document batches.

  The document batch will be split up and can be saved in separate output documents.

How to separate image files

- Click the **Open** button on the main toolbar and select the documents you want to open.
  The Pages panel will display the page thumbnails.

  💡 **Tip**: Use the Batch Processing or Watched folder function when scanning large volumes of documents.

- **Ctrl**-click the page where you want the new document to begin, and click **Split Here**.
  Or select the page where you want the new document to begin, and click the **Edit** menu > **Split Here**.
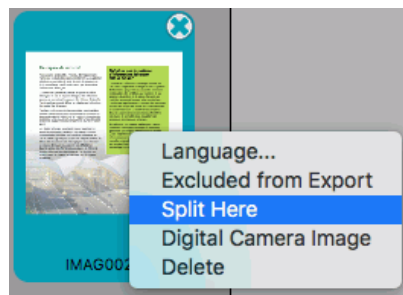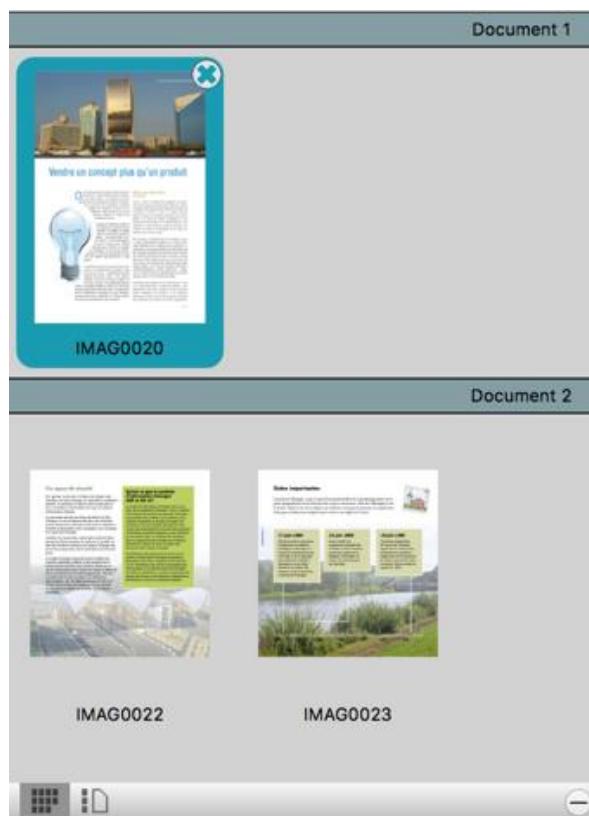

  Note that the **Split Here** command cannot be used on the first page of a document, since it is already the first page.

  Also note that it is not possible to select multiple pages and then use **Split Here**.

- The Pages panel now displays the split documents.

Note: in Thumbnail view ( ), each new document is put on a different line, to facilitate the overview.

- Both in the Pages panel and in Thumbnail view you can drag pages to different locations, and add new pages by drag-and-dropping them into Readiris.

    Note however that you cannot move a page or insert a new page before the first page of Document 2 and following. It is possible on the other hand to move a page or insert a new pages before the first page of Document 1.

- Click the one of the **Export** buttons to process the documents.

How to merge documents

To merge documents you have split:

- Ctrl-click the first page of a document, and click **Merge Documents**.
  Or select the first page of a document, and click the **Edit** menu > **Merge Documents**.

- The document you selected will be merged with the previous document.

    Note that the **Merge Documents** command does not work on the first document, since there is no preceding document to merge with.
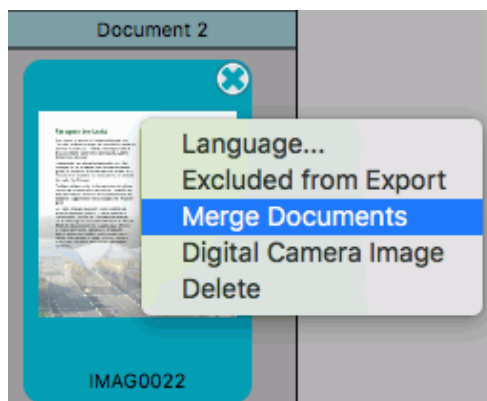
How to recognize forms

Note that Readiris is not designed to recognize complex forms. Readiris recognizes documents by converting pictures of text characters and numbers into actual text and numbers. Forms, on the other hand, contain many graphic elements such as lines, boxes, check boxes and the like, which complicate Readiris' tasks.

Use the following rule of thumb if you try to recognize forms: if you would be able to re-create the form fairly easily from scratch inside Microsoft Word or Pages for Mac, Readiris has a good chance of processing it. If it would take an extraordinary effort to re-create it in Word, then Readiris will have the same difficulty.

Example

Form Readiris would not be able to process correctly
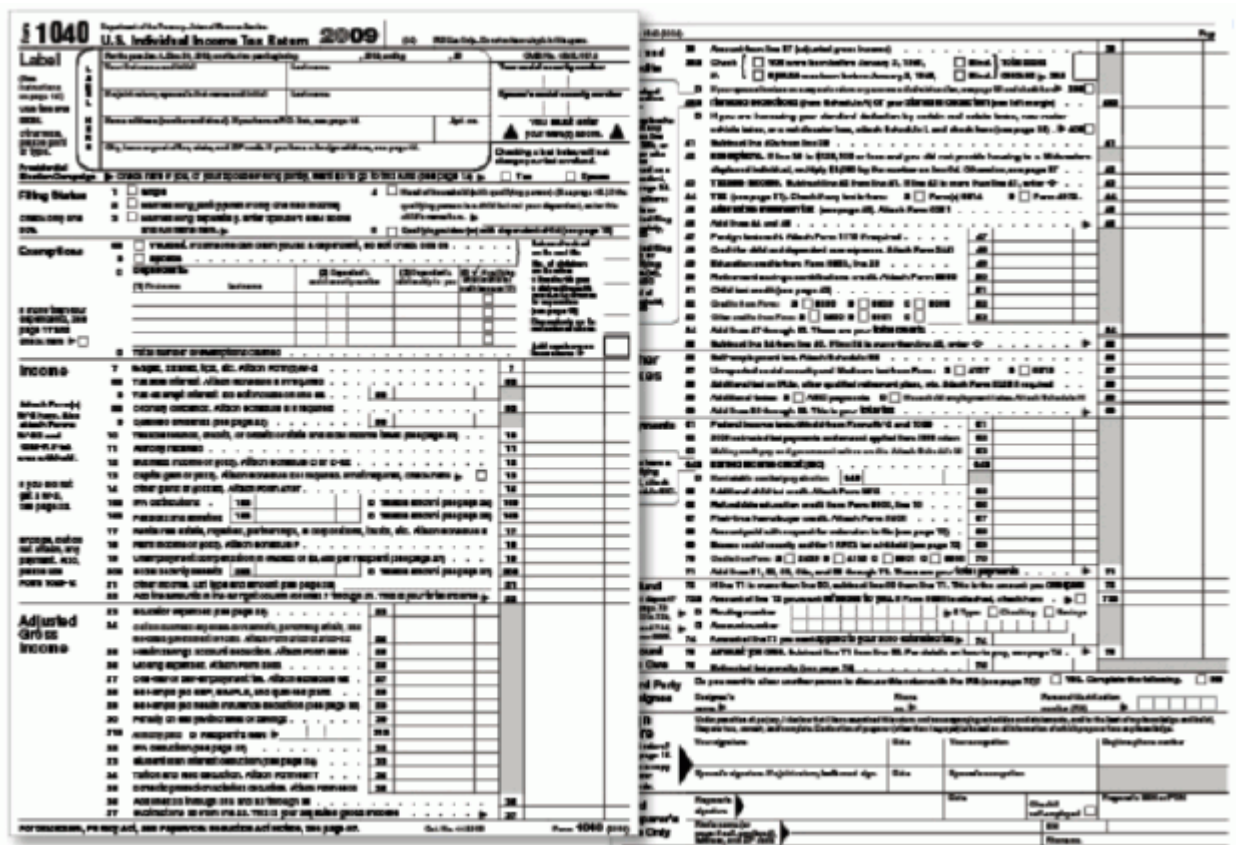
How to process low-resolution images

A resolution of **300** to **400** dpi is required to obtain optimal results in Readiris. When you try to scan or open images of a resolution lower than 200 dpi or higher than 600 dpi, Readiris will display a warning.

When you are faced with these kind of images, the best thing you can do is re-can the images in an appropriate resolution. If that's not possible, there are few tips and tricks that help you obtain better results:

Tip 1

When scanning images of an **unknown resolution** or images of which the file header displays an **incorrect resolution**, make sure to select the option **Process as 300 dpi**:

- On the **Readiris** menu, click **Preferences**.
- Click the **Preprocessing** tab.
- Select the option **Process as 300 dpi**.

When this option is selected, Readiris treats the images as if they had a 300 dpi resolution.

Tip 2 *(Readiris Corporate only)*

Make sure that favor **OCR accuracy** over speed when processing low-quality documents:

- On the **Readiris** menu, click **Preferences**.
- Click the **Page Analysis** tab.
- Make sure the option **Favor accuracy over speed** is selected.

Note: When you are processing low-resolution faxes, there is no need to take special measures. Readiris detects the typical fax resolutions 100 x 200 dpi (normal quality), 200 x 200 dpi (fine quality) and 200 x 400 dpi (superfine quality) and preprocesses these images automatically to ensure optimal results.