Readiris™ Pro 12

Manual de utilizador

Índice

Direitos de autor 1

Capítulo 1 Introduzir o Readiris 3

Poupe tempo, não dactilografe mais 3

A série Readiris 6

Capítulo 2 Instalar o Readiris 11

Requisitos do sistema 11

Instalação do software 11

Desinstalar o software 13

Registo do software 13

Assistência do produto 14

Capítulo 3 Iniciar 15

Executar o Readiris 15

Utilizar o assistente de OCR 15

Interface do utilizador 17

Alterar o idioma da interface do utilizador 20

Capítulo 4 Readiris SmartTasks 21

Utilizar as Readiris SmartTasks 21

Capítulo 5 Digitalizar documentos 25

Seleccionar as opções 25

Abrir ficheiros de imagem 26

Digitalizar documentos em papel 28

Capítulo 6 Ajustar documentos digitalizados 33

Capítulo 7 Guardar documentos como ficheiros de imagem37

Capítulo 8 Dispor documentos por janelas 39

Dispor documentos por janelas automaticamente 39

Dispor documentos por janelas manualmente 41

Utilizar modelos de disposição de janelas..... 46

Capítulo 9 Reconhecer documentos..... 49

Introdução 49

Seleccionar o idioma do documento 50

Definir as características do documento..... 53

Utilizar aprendizagem interactiva 54

Utilizar dicionários de tipos de letra 57

Capítulo 10 Formatar e guardar documentos..... 59

Formatar documentos..... 59

Formatar documentos de texto 61

Formatar documentos baseados em tabelas..... 65

Criar documentos PDF..... 69

Seleccionar as opções PDF 70

Documentos PDF comprimidos por iHQC 71

Converter documentos PDF 73

Criar documentos XPS 74

Seleccionar as opções XPS 75

Documentos XPS comprimidos por iHQC 76

Seleccionar as opções dos gráficos 77

Capítulo 11 Guardar e carregar as definições 81

Capítulo 12 Reconhecer documentos multi-páginas..... 83

Abrir e reconhecer múltiplos ficheiros de imagem 83

Digitalizar e reconhecer documentos multi-páginas 85

Editar documentos multi-páginas..... 86

Capítulo 13 Reconhecer texto escrito em letra de imprensa 89

Regras de texto escrito em letra de imprensa 91

Capítulo 14 Reconhecer códigos de barras 93

Introduzir leitura de códigos de barras 93

Índice alfabético 95

Direitos de autor

ReadirisPro12-dgi-090217-01

Direitos de autor © 1987-2009 I.R.I.S. Todos os direitos reservados.

A I.R.I.S. possui os direitos de autor para o software Readiris software, para o sistema de ajuda on-line e para esta publicação.

A informação contida neste documento é propriedade da I.R.I.S. O seu conteúdo está sujeito a alterações sem aviso prévio e não representa um compromisso por parte da I.R.I.S. O software descrito neste documento é fornecido sob um acordo de licença que declara os termos de utilização deste produto. O software pode ser utilizado ou copiado apenas em conformidade com os termos desse acordo. Esta publicação não pode ser reproduzida, transmitida, guardada num sistema de recuperação ou traduzida para outro idioma sem autorização escrita prévia da I.R.I.S.

Este manual de utilizador utiliza nomes fictícios para fins de demonstração, referências a pessoas, empresas ou organizações reais são estritamente coincidentes.

Marcas comerciais

O logótipo Readiris e o Readiris são marcas comerciais da Image Recognition Integrated Systems S.A.
OCR, ICR e de código de barras da I.R.I.S.
Tecnologia AutoFormat e linguística da I.R.I.S.Tecnologia BCR e análise de campo da I.R.I.S.
Tecnologia BCR e análise de campo da I.R.I.S.
Tecnologia de compressão iHQC da I.R.I.S.
Reconhecedor XML desenvolvido pela Apache. Este produto inclui o software desenvolvido pela Apache Software Foundation.

Todos os outros produtos mencionados neste manual de utilizador são marcas comerciais ou marcas comerciais registadas dos seus respectivos proprietários.

CAPÍTULO 1
INTRODUZIR O READIRIS

POUPE TEMPO, NÃO DACTILOGRAFE MAIS

Parabéns por ter adquirido o Readiris. Este pacote de software será indubitavelmente um grande ajuda para captar novamente os textos, tabelas e gráficas, códigos de barras e texto escrito em letra de imprensa.

Apesar da eficiência dos computadores, necessita primeiro de introduzir as suas informações. Se alguma vez dactilografou um relatório de 15 páginas ou uma tabela grande de números, sabe o quão tédio e moroso pode ser. Utilize este pacote de OCR topo de gama para converter automaticamente documentos em papel ou ficheiros de imagem digitalizados em documentos pesquisáveis e editáveis que podem ser arquivados e partilhados. Estão disponíveis dois modos de reconhecimento: um garante velocidade máxima, o outro garante uma precisão de OCR ideal.

Digitalize um documento impresso ou dactilografado, indique as zonas que pretende reconhecer com o Readiris - ou coloque o sistema a detectá-las por si - execute o reconhecimento de caracteres e exporte o documento para o processador de texto. Os documentos compostos por muitas páginas são processados de início ao fim num único passo. Alguns cliques do rato é melhor que longas horas de trabalho enquanto o Readiris converte os seus documentos em papel em ficheiros de computador editáveis: é 40 vezes mais rápido que dactilografar manualmente.

O assistente guia-o suavemente pelas definições necessárias para operar o Readiris, permitindo obter resultados rápidos e fáceis. Ou utilize as SmartTasks para acelerar ainda mais o processo. É possível enviar os resultados de leitura directamente para o seu processador de texto ou folha de cálculo, arquive-os como ficheiros PDF ou XPS, etc. Para reconhecer faxes e converter documentos PDF, arraste os seus ficheiros de imagem do Windows Explorer para a janela de aplicação do Readiris.Ou envie imediatamente uma imagem para o Readiris através do menu de contexto.

O Readiris também reconhece dados tabulares e recria-os como folhas de cálculo no seu software de folhas de cálculo ou como objectos de tabela dentro do processador de texto, os dados numéricos estão imediatamente prontos para um processamento posterior.

O Readiris baseia-se nas tecnologias de reconhecimento mais avançadas. O reconhecimento de texto com tipo de letra independente é completado por técnicas auto-aprendizagem. O sistema consegue aprender novos caracteres e palavras através de análise contextual e linguística. Isto significa que a precisão de OCR do sistema de reconhecimento irá melhorar à medida que avança.

O Readiris suporta até 128 idiomas: são suportados todos os idiomas americanos e europeus, incluindo os idiomas Centro-europeus, Bálticos e Cirílicos assim como o Grego e o Turco. Opcionalmente, o Readiris consegue ler documentos em Árabe, Persa e Hebraico e quatro idiomas Asiáticos - Japonês, Chinês simplificado e tradicional e Coreano. O Readiris suporta até alfabetos mistos: o software detecta palavras "Ocidentais" que ocorrem em documentos em Grego, Cirílico, Árabe, Hebraico e Asiático - muitos nomes próprios não convertíveis, nomes de marcas etc. são escritos utilizando os símbolos Ocidentais.

O Readiris utiliza linguística *durante* a fase de reconhecimento, não após. Como resultado, o Readiris reconhece todos os tipos de documentos com uma excelente precisão, incluindo documentos de

baixa qualidade, faxes e impressões com matriz de pontos. Suporta na perfeição documentos mal digitalizados e copiados com tipos de letras demasiado claros ou escuros. Os caracteres unidos são separados enquanto que os caracteres fragmentados, como símbolos com matriz de pontos, são recompostos.

Além disso, o Readiris tem uma função de verificação de utilizador (opcional). Quando activada, a função de verificação de utilizador ("Aprendizagem interactiva") não só assinala os caracteres que o sistema de reconhecimento não tem a certeza mas também permite aumentar a precisão do sistema. Todas as soluções que confirmar são memorizadas, aumentando a velocidade e a confiança do sistema e tornando o sistema mais inteligente à medida que avança. Esta ferramenta de aprendizagem avançada também permite treinar o Readiris nos caracteres especiais como símbolos matemáticos e dingbats e suportar tipos de letra distorcidos.

Para aumentar ainda mais a produtividade, o Readiris não só reconhece os textos, como pode formatá-los por si. Estão disponíveis vários níveis de formatação. Quando utiliza a "formatação automática", o Readiris recria uma cópia fac-simile do documento digitalizado: a formatação das palavras, parágrafos e páginas do documento original é mantida. São utilizados tipos de letra semelhantes, os tamanhos e estilos utilizados no documento original são mantidos ao longo do reconhecimento. A colocação de colunas, blocos de texto e gráficos segue os documentos originais. O Readiris pode até incluir a fotografia de fundo de uma página digitalizada no documento reconhecido. E visto que o Readiris suporta facilmente digitalizações com escala de cinzentos e a cores, é possível captar quaisquer gráficos - sejam eles fotografias a preto e branco ou ilustrações a cores. Quando um documento contém tabelas, o Readiris reconhece-as em células reais e recria as margens das células das tabelas originais.

Por outras palavras, o Readiris permite arquivar uma cópia verdadeira dos seus documentos, sendo ficheiros de texto editáveis e compactos em vez de imagens digitalizadas.

Os códigos de barras que ocorrem numa página digitalizada também podem ser lidos e o mesmo aplica-se para o texto escrito em letra de imprensa, desde que escreva "letras maiúsculas" bem espaçadas.

O Readiris é compatível com Twain e suporta uma vasta gama de scanners e dispositivos "tudo em um" ou "MFPs" ("periféricos multifuncionais") e câmaras digitais. A digitalização de intervalo permite digitalizar documentos multi-páginas de forma eficiente quando o scanner não está equipado com um alimentador de documentos.

A SÉRIE READIRIS

A tabela abaixo fornece uma vista geral das versões disponíveis:

Readiris Home 12
Funcionalidades limitadas
25 idiomas de reconhecimento
Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Gera ficheiros Imagem-Texto PDF, DOCX, ODT, WordML, SpreadsheetML. RTF, HTM, XML, TXT, TIFF

Readiris Pro 12	**Readiris Corporate 12**
Funcionalidades básicas	Funcionalidades básicas
128 idiomas de reconhecimento	128 idiomas de reconhecimento
Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX	Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível I), quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc.	Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível II-III, PDF/A, quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc.
	Reconhecimento de grandes quantidades
	Processamento automatizado
	Indexação de documentos
	Reconhecimento de cartões de visita

Readiris Pro 12 Asian	**Readiris Corporate 12 Asian**
Funcionalidades básicas	Funcionalidades básicas
128 idiomas de reconhecimento	128 idiomas de reconhecimento
Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX	Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível I), quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc.	Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível II-III, PDF/A, quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc.

Reconhecimento de Chinês tradicional e simplificado Reconhecimento de Japonês Reconhecimento de Coreano	Reconhecimento de Chinês tradicional e simplificado Reconhecimento de Japonês Reconhecimento de Coreano Reconhecimento de grandes quantidades Processamento automatizado Indexação de documentos Reconhecimento de cartões de visita

Readiris Pro 12 Middle-East*	**Readiris Corporate 12 Middle-East***
Funcionalidades básicas 128 idiomas de reconhecimento Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível I), quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc. Reconhecimento de Árabe & Persa Reconhecimento de Hebraico	Funcionalidades básicas 128 idiomas de reconhecimento Suporta imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Gera quatro tipos de ficheiros PDF, PDF-iHQC (nível II-III, PDF/A, quatro tipos de ficheiros XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF, etc. Reconhecimento de Árabe & Persa Reconhecimento de Hebraico Reconhecimento de grandes quantidades Processamento automatizado

	Indexação de documentos Reconhecimento de cartões de visita

(*) Nenhuma versão para Mac disponível.

CAPÍTULO 2
INSTALAR O READIRIS

REQUISITOS DO SISTEMA

Esta é a **configuração mínima do sistema** necessária para utilizar o Readiris:

- um PC Intel 486 ou compatível. Um PC Pentium é recomendado.
- 256 MB de memória RAM.
- 120 MB de espaço livre no disco.

  (bastam 105 MB de espaço livre no disco quando não instala os ficheiros de amostra)

- os sistemas operativos Windows Vista, Windows XP ou Windows 2000.

Atente que algumas unidades de scanner podem não funcionar na versão mais recentes do Windows. Consulte a documentação fornecida com o seu scanner para descobrir quais as plataformas suportadas.

INSTALAÇÃO DO SOFTWARE

Como instalar o Readiris:

- Inicie sessão no Windows como administrador ou certifique-se de que tem os direitos de administrador necessários para instalar o software.
- Ligue o seu scanner ao PC e instale o software correspondente. Teste o seu scanner. Se ocorrer algum problema, contacte o fabricante do scanner.
- Introduza o CD-ROM do Readiris na unidade de CD-ROM e siga as instruções no ecrã para instalar o software.
- Clique em **Readiris** para iniciar a instalação (são oferecidos produtos de software adicionais: Copernic Desktop Search Home Edition e Cardiris 4 LE).
- Seleccione o idioma de instalação e clique em **OK**.
- Aceite os termos do acordo de licença.
- É oferecida uma instalação **completa** e **personalizada**. Seleccione as opções necessárias e clique em **Seguinte** sempre que estiver pronto para passar para o ecrã seguinte.
  Todos os léxicos e imagens de amostra serão instalados por defeito, assim como um manual de utilizador electrónico e um ajuda on-line.
- Clique em **Concluir** para concluir a instalação.
  O sub-menu **Aplicações I.R.I.S. - Readiris** no menu **Programas** do Windows é criado automaticamente pelo programa de instalação.
  O programa de instalação também cria um atalho para a aplicação Readiris no ambiente de trabalho do Windows.

- Repita o processo de instalação para instalar qualquer software adicional a partir do CD-ROM.

DESINSTALAR O SOFTWARE

Existe apenas uma forma correcta de desinstalar o Readiris: através do assistente de (des)instalar do Windows. É altamente recomendado *não* desinstalar o Readiris ou qualquer um dos seus módulos de software ao eliminar manualmente os ficheiros de programas.

Para desinstalar o Readiris:

- Feche a aplicação.
- No menu **Iniciar**, clique em **Painel de controlo**.
- No ícone **Programas**, clique em **Desinstalar um programa**.
- Seleccione **Readiris** na lista e clique no botão **Desinstalar**.
- Siga as instruções no ecrã.

REGISTO DO SOFTWARE

Lembre-se de registar a sua licença de Readiris. Ao fazer o registo, irá:

- ser informado sobre futuros desenvolvimentos de produtos e produtos I.R.I.S. relacionados;
- receber assistência grátis do produto;
- ter direito a ofertas especiais sobre os produtos I.R.I.S.

Para registar:

Utilize o **Assistente de registo** no menu **Registar**. Siga as instruções no ecrã.

ASSISTÊNCIA DO PRODUTO

Assim que registar o produto, tem direito à assistência do produto por parte da I.R.I.S. em todas as funcionalidades de softwarebásicas. Contacte a I.R.I.S. pelo:

Europa:
support@irislink.com
Tel:+32 10 45 13 64

EUA:
support@irisusa.com
Tel.:+1 800 447 4744

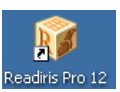
Ásia-Pacífico:
support@irislink.com
Tel.: +852 22646133

CAPÍTULO 3
INICIAR

EXECUTAR O READIRIS

Para executar o Readiris:

- Inicie o **Readiris** a partir do menu **Iniciar** do Windows ou clique duas vezes no atalho no seu ambiente de trabalho.



- Clique em qualquer parte do ecrã de arranque para lançar o Readiris.

O Assistente de OCR abre automaticamente.

UTILIZAR O ASSISTENTE DE OCR

O Assistente de OCR permite definir rapidamente todas as definições necessárias para operar o Readiris.

Quando inicia o Readiris, clique em qualquer parte do ecrã de arranque para iniciar o Assistente de OCR.

Fase 1

Seleccione a **fonte de imagem**.
É possível captar imagens utilizando o **scanner** ou abrindo **ficheiros de imagem**.
Seleccione as opções de **rotação** e de **remoção da inclinação da página** que pretende utilizar.

> Para obter mais informações, consulte a secção **Seleccionar as opções**.
>
> Para se familiarizar com o Readiris, utilize as imagens de amostra fornecidas com o software.
>
> Podem ser encontradas no CD-ROM do CD-ROM e na sub-pastas **Amostras** na pasta de instalação do Readiris.

Clique em **Seguinte** para avançar para a fase seguinte.

Fase 2

Caso seleccione um scanner, clique no botão **Alterar** para seleccionar as definições do scanner.

> Para obter mais informações sobre as definições do scanner, consulte a secção **Digitalizar documentos em papel**.

Clique em **OK** para guardar as definições.
Clique em **Seguinte** para avançar para a fase seguinte.

Fase 3

Clique no botão **Alterar** para alterar o idioma do documento. O idioma do documento está predefinido para Inglês americano.
Seleccione o idioma ou combinação de idiomas necessários a partir da lista e clique em **OK**.
Utilize o cursor para definir as definições de **Velocidade-Precisão**.

> Para obter mais informações, consulte a secção **Seleccionar o idioma do documento**.

Clique em **OK** para guardar as definições.
Clique em **Seguinte** para avançar para a fase seguinte.

Fase 4

Clique no botão **Alterar** para alterar o formato de exportação ou a aplicação pretendida. A aplicação pretendida predefinida é o Microsoft Word.
Seleccione o formato ou aplicação de exportação necessários a partir da lista **Enviar para** ou **Ficheiro externo**.
Clique nos vários separadores e seleccione as opções à sua escolha. As opções que estão indisponíveis para o formato/aplicação escolhidos surgem desactivadas.

> Para obter mais informações, consulte a secção **Formatar documentos**.

Clique em **OK** para guardar as definições.
Clique em **Seguinte** para avançar para a fase seguinte.

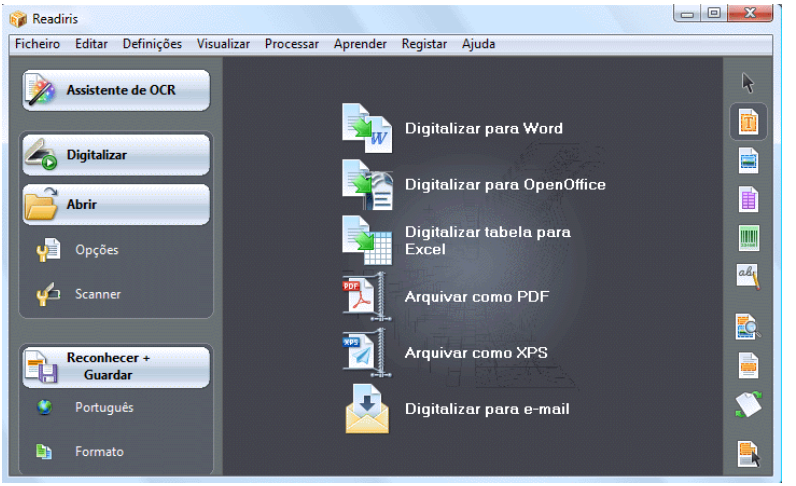
Fase 5

Clique em **AVANÇAR** para abrir/digitalizar e reconhecer o documento.

INTERFACE DO UTILIZADOR

Para explorar a interface do Readiris, clique em qualquer parte do ecrã de arranque do **Readiris** e clique em **Cancelar** quando o assistente de OCR é lançado.

Será apresentada a interface Readiris vazia.



A interface do Readiris é composta por:

- **SmartTasks** (ao centro)

  As SmartTasks são comandos predefinidos que permitem utilizar as funções Readiris mais frequentes através de um botão.

  Utilize as SmartTasks para digitalizar, reconhecer e enviar os documentos para a aplicação pretendida ou formato de exportação à sua escolha.

  As SmartTasks aplicam predefinições mas podem ser configuradas facilmente ao clicar com o botão direito para se adequarem a necessidades mais particulares.

- a **barra de ferramentas principal** (barra de ferramentas esquerda)

  Utilize os comandos da barra de ferramenta principal e opções para digitalizar e reconhecer documentos manualmente.
  A ordem pela qual é aconselhado a fazê-lo é fornecida no **Assistente de OCR**.

- a **barra de ferramentas da imagem** (barra de ferramentas direita)

  Utilize os botões da barra de ferramentas da imagem para editar documentos na interface do Readiris.
  Seleccione os diferentes botões para apresentar as suas descrições.

Quando um documento tiver sido aberto ou digitalizado no Readiris, são adicionadas três zonas principais à interface:

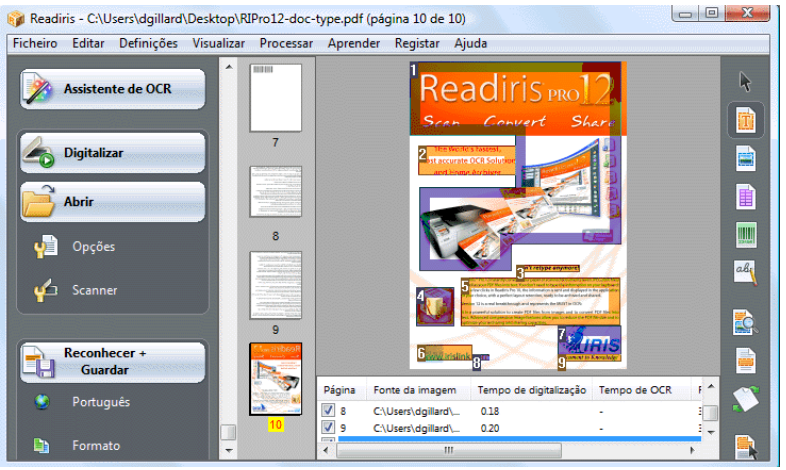
- a **barra de ferramentas da página** (à direita da barra de ferramentas)

  A barra de ferramentas da página apresenta as miniaturas da página, que fornecem informações sobre as definições quando seleccionadas.

- a **janela da imagem** (ao centro)

- o **painel de documentos** (em baixo)

  O painel de documentos apresenta informações estatísticas sobre os documentos que estão abertos no Readiris, como o tempo de digitalização e de OCR, a resolução, largura e altura dos documentos etc.



ALTERAR O IDIOMA DA INTERFACE DO UTILIZADOR

A interface do utilizador do Readiris está disponível numa grande variedade de idiomas.

Para alterar o idioma da interface do utilizador:

- No menu **Definições**, clique em **Idioma da interface do utilizador**.
- Na lista **Idioma**, seleccione o idioma necessário, de seguida, clique em **OK**para confirmar.

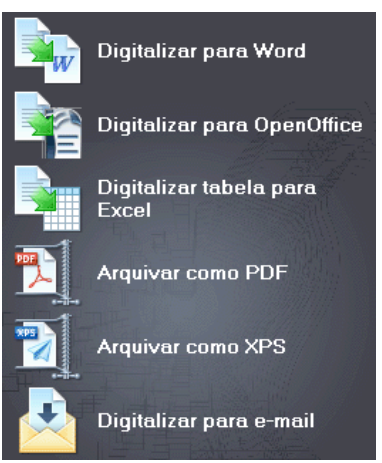
Nota: Se seleccionou um idioma incorrecto, clique em **Ctrl+U**. A caixa de diálogo **Idioma** irá abrir e poderá seleccionar outro idioma na lista.

CAPÍTULO 4
READIRIS SMARTTASKS

UTILIZAR AS READIRIS SMARTTASKS

Quando iniciar o Readiris, clique em qualquer parte do ecrã de arranque do Readiris e clique em **Cancelar** quando o assistente de OCR é lançado. Serão apresentadas as Readiris SmartTask.

As SmartTasks são comandos predefinidos que permitem utilizar as funções Readiris mais frequentes através de um botão.



Os vários botões das Readiris SmartTask permitem:

1. digitalizar e reconhecer documentos e enviá-los directamente para o **Word** para processamento de texto;

O Microsoft Word é a aplicação pretendida predefinida. Consulte a secção **Formatar documentos de texto** para saber mais sobre as outras aplicações disponíveis.

2. digitalizar e reconhecer documentos e enviá-los directamente para o **OpenOffice** para processamento de texto;

O OpenOffice.org Writer é a aplicação pretendida predefinida. Consulte a secção **Formatar documentos de texto** para saber mais sobre as outras aplicações disponíveis.

3. digitalizar e reconhecer tabelas e enviá-las directamente para o **Excel** e para outras folhas de cálculo;

O Microsoft Excel é a aplicação pretendida predefinida. Consulte a secção **Formatar documentos baseados em tabelas** para saber mais.

4. digitalizar e reconhecer documentos e **arquivá-los** como ficheiros **PDF**;

Imagem-Texto Adobe Acrobat PDF é o formato de exportação predefinido. Consulte a secção **Criar documento PDF** para saber mais sobre os outros formatos disponíveis.

5. digitalizar e reconhecer documentos e **arquivá-los** como ficheiros **XPS**;

Imagem-Texto XPS é o formato de exportação predefinido. Consulte a secção **Criar documento XPS** para saber mais sobre os outros formatos disponíveis.

6. digitalizar e reconhecer documentos e enviá-los directamente por **e-mail**.

Os documentos serão enviados como Imagem-Texto PDF por defeito através da aplicação de e-mail predefinida. Consulte a secção **Formatar documentos** para saber mais sobre os outros formatos disponíveis.

Quando utilizar o Readiris pela primeira vez, tem de configurar as SmartTasks.

Para configurar as SmartTasks:

- Clique com o botão direito na SmartTask que pretende utilizar.
- Seleccione **Scanner** ou **Ficheiros de imagem** como fonte da imagem.
    - Quando selecciona **Scanner**, o Readiris irá iniciar o seu scanner assim que clicar na SmartTask. O(s) documento(s) digitalizado(s) será(ão) apresentado(s) na interface, processado(s) e guardado(s).

      O seu scanner tem de ser configurado correctamente para que as SmartTasks funcionem.

      Para o fazer:

      - Clique no botão **Scanner** na barra de ferramentas principal.
      - Clique em **Modelo de scanner** e seleccione o seu scanner a partir da lista.

        Se o seu scanner não estiver na lista, seleccione **Outros modelos Twain**.

      - Clique em **Configurar** se aplicável, para seleccionar a fonte Twain.
      - De seguida, clique em **OK** para guardar as definições.

        Para obter mais informações sobre as definições do scanner e sobre digitalizar documentos em papel, consulte a secção **Digitalizar documentos em papel**.

- Quando selecciona **Ficheiros de imagem** e clique na SmartTask, o Readiris abre a caixa de diálogo de **Entrada** na qual é possível seleccionar os ficheiros de imagem que pretende processar.

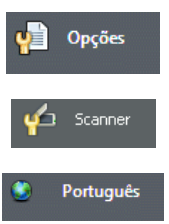
  Para obter mais informações sobre abrir ficheiros de imagem, consulte a secção **Abrir ficheiros de imagem**.

- Clique em **Configurar** para alterar o **formato** de exportação e as suas **opções**.

  Atente que os formatos e opções de exportação disponíveis dependem da SmartTask seleccionada.

  Consulte a secção **Formatar documentos** para saber mais sobre os formatos e opções disponíveis.

- Quando concluir a formatação das SmartTasks, utilize os botões na barra de ferramentas principal para especificar as definições do **idioma** e as opções demelhoramento de imagem e se ainda for necessário as **definições do scanner**.



  Para obter mais informações sobre as definições mencionadas acima, consulte as secções **Seleccionar as opções**,**Digitalizar documentos em papel** e **Seleccionar o idioma do documento**.

- Por fim, clique na SmartTask para utilizá-la.

  O Readiris irá executar automaticamente o processo de reconhecimento inteiro.

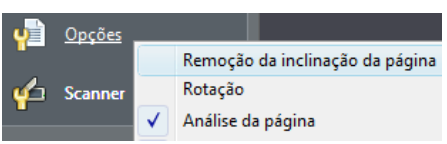
CAPÍTULO 5
DIGITALIZAR DOCUMENTOS

SELECCIONAR AS OPÇÕES

Antes de digitalizar documentos em papel ou abrir ficheiros de imagem, é possível seleccionar várias opções de melhoramento de imagem. Quando activadas, estas opções serão aplicadas durante a abertura e digitalização de documentos.

Operação

- Clique no botão **Opções** na barra de ferramentas principal para seleccionar várias opções de melhoramento de imagem.



- Clique em **Remoção da inclinação da página** para alinhar as páginas digitalizadas de forma inclinada.
- Seleccione **Rotação** e determine se pretende que o Readiris para rodar as páginas **automaticamente** ou **90° para a esquerda**, **90° para a direita** ou **180°**.

  Atente que estas duas opções abrandam um pouco o processo de digitalização. Seleccione-as apenas quando necessário.

- **Análise de páginas** está activada por defeito.

Desta forma, as imagens digitalizadas ou abertas serão automaticamente divididas em janelas.

Pode também utilizar as ferramentas de disposição por janelas na barra de ferramentas da imagem para modificar os resultados de análise da página ou para dispor manualmente os documentos por janelas. Para obter mais informações, consulte a secção **Dispor manualmente documentos por janelas**.

- Após definir todas as definições (**Definições do scanner**, **Opções**), clique no botão **Digitalizar** ou **Abrir** para digitalizar documentos ou abrir ficheiros de imagem.

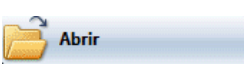
  Atente que as opções mencionadas acima estão também disponíveis no menu **Definições**.

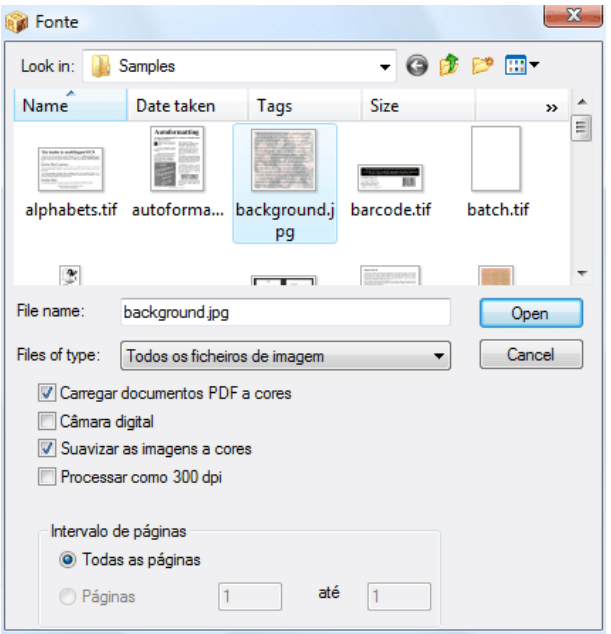
ABRIR FICHEIROS DE IMAGEM

Com o Readiris é possível processar documentos em papel que digitaliza no scanner ou processar ficheiros de imagem já existentes de vários formatos.

Para abrir ficheiros de imagem existentes:

- Clique no botão **Abrir** paraprocurar ficheiros de imagem.





Conselho: também pode **arrastar** ficheiros de imagem para a janela de imagem do Readiris para os abrir.

Conselho: Clique com o botão directo em qualquer ficheiro de imagem que pretende abrir, seleccione **Abrir com** e clique em **aplicação IOCR**. O software Readiris irá abrir e apresentar a imagem.

Conselho: quando carregar ficheiros de imagem multi-páginas (imagens TIFF e faxes DCX) e documentos PDF, é possível definir o intervalo de páginas (por exemplo, caso necessite apenas de um certo capítulo de um documento). Para o fazer, clique em **Abrir** na barra de ferramentas principal. Na área **Intervalo de páginas**, seleccione **Páginas** e introduza quais as páginas que pretende carregar. Ver também **Abrir e reconhecer múltiplos ficheiros de imagem**

Conselho:para acelerar o processo de carregamento, clique em **Abrir** e anule a selecção de **Carregar documentos PDF a cores** quando processar documentos PDF.

- O Readiris suporta os formatos de gráficos seguintes:

Formato de gráfico	**Extensão do ficheiro**
Adobe Acrobat PDF	*.pdf
DCX fax	*.dcx
Imagens DjVu	*.djv, *.djvu
Imagens JPEG	*.jpg, *.jpeg
Imagens JPEG 2000	*.j2c, *.j2k, *.jp2
Gráficos de rede portáteis	*.png
(Comprimidas e multi-páginas) imagens TIFF não comprimidas e LZW, PackBits, Grupo 3, Grupo 4 e JPEG comprimidos	*.tif, *.tiff
Mapa de bits do Windows	*.bmp
Imagens ZSoft Paintbrush	*.pcx

- Seleccione o ficheiro de imagem à sua escolha e clique em **Abrir**.

Nota: as opções da caixa de diálogo **Entrada** também se aplicam a digitalizações de documentos e estão discutidas na secção **Digitalizar documentos em papel**.

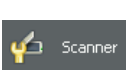
DIGITALIZAR DOCUMENTOS EM PAPEL

Com o Readiris é possível processar documentos em papel que digitaliza no scanner ou processar ficheiros de imagem já existentes de vários formatos.

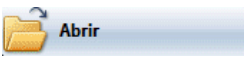
Quando processa documentos em papel, o Readiris irá iniciar o scanner assim que clicar no botão **Digitalizar** e apresentar o documento digitalizado na interface.

Para digitalizar documentos:

- Clique no botão **Scanner** para seleccionar as definições do scanner.



Atente que muitas das opções na caixa de diálogo **Scanner** também estão disponíveis na caixa de diálogo **Abrir**.



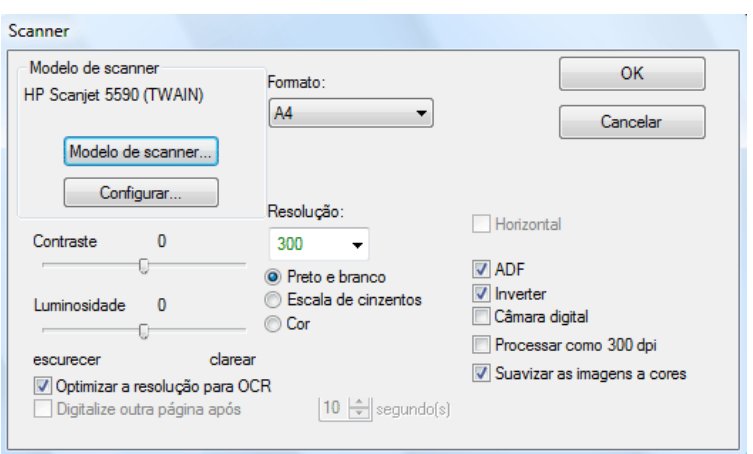
- Seleccione o modelo de scanner correcto.

  Se o seu scanner não estiver na lista, seleccione **Outros modelos Twain** e clique em **OK**.

- O Readiris suporta quase todos os scanners de mesa e de alimentação de papel, "dispositivos" tudo em um ("MFP" - "Periférico multifuncional"), câmaras digitais e scanners standard actualmente disponíveis. O Readiris é compatível com Twain.

  Atente que o botão **Configurar** só está disponível quando aplicável. Clique no botão para seleccionar a fonte Twain.

- Seleccione as definições do scanner.



Formato e resolução

O Readiris suporta uma vasta gama de formatos de papel e resoluções.

Atente que é recomendado utilizar uma resolução de digitalização de 300 ppp.

Utilize uma resolução de 400 ppp quando reconhecer texto asiático ou letras muito pequenas.

Modo a cores

O Readiris pode digitalizar documentos e abrir ficheiros de imagem a **cores**, **preto e branco** e **escala de cinzentos**.

Contraste - Luminosidade

Utilize o cursor para determinar as definições de luminosidade e de contraste adequadas, de modo a obter um resultado de digitalização ideal.

Optimizar a resolução para OCR

Seleccione **Optimizar a resolução para OCR** para corrigir a resolução de imagens digitalizadas com demasiado detalhe (superior a 600 ppp). O Readiris irá reduzir a resolução.

> Atente que esta opção nunca aumenta a resolução de imagens digitalizadas com detalhe insuficiente.

Digitalizar documentos multi-páginas

Quando digitalizar documentos multi-páginas e utilizar um scanner equipado com um alimentador de documentos, seleccione a opção **ADF** (alimentador de documentos automático). Coloque as páginas que pretende digitalizar no alimentador e comece a digitalizar.

Ou utilize **digitalização de intervalos** quando utilizar um scanner de mesa**:** seleccione a opção **Digitalizar outra página depois** e indique após quantos segundos pretende que o Readiris digitalize outra página.

Para obter mais informações, consulte a secção **Digitalizar e reconhecer documentos multi-páginas**.

Importante: quaisquer opções que estão indisponíveis para o scanner seleccionado surgem desactivadas.

Digitalizar imagens horizontais

Seleccione a opção **Horizontal** quando digitalizar imagens na horizontal.

Auto-exposição

Com alguns scanners a opção **Auto-exposição** é seleccionada por defeito. Esta opção ajusta as definições de contraste e luminosidade automaticamente.

Digitalizar imagens invertidas

Seleccione a opção **Inverter** quando digitalizar imagens a preto e branco invertidas.

Utilizar uma câmara digital

Seleccione **Câmara digital** quando utilizar uma câmara como fonte de digitalização. O Readiris utiliza rotinas de reconhecimento especiais para processar imagens de câmaras digitais.

Conselhos para utilizar uma câmara digital como fonte de digitalização:

- Calibre a câmara ao fotografar um documento branco.
- Seleccione sempre a resolução de imagem mais elevada.

- Active o modo macro da câmara para fotografar grandes planos.
- Utilize apenas zoom óptico e não zoom digital.
- Segure a câmara directamente sobre o documento. Evite fotografar o documento de forma inclinada.
- Produza imagens estáveis. Se necessário, utilize um tripé.
- Desactive o flash quando captar papel brilhante.
- Evite abrir imagens de câmara comprimidas.
- Adapte as definições de luminosidade e contraste do Readiris ao ambiente (luz do dia, luz de lâmpadas, luz néon).
- Seleccione cor ou escala de cinzentos como modo de cor.

Processar como 300 ppp

Seleccione **Processar como 300 ppp** quando processar imagens com uma resolução incorrecta ou desconhecida. As imagens serão processadas como se tivessem uma resolução de 300 ppp.

A resolução das imagens de câmaras digitais é quase sempre desconhecida.

Suavizar as imagens a cores

Esta opção é seleccionada por defeito visto que a suavização das imagens é necessária com alguns scanners para reconhecer as imagens a cores ou com escala de cinzentos com sucesso.

- Após definir todas as definições (**Definições do scanner**, **Opções**), clique em **Digitalizar** para digitalizar documentos.

Nota: preste atenção à inclinação da linha. Uma inclinação de linha superior a 0.5° aumenta o risco de erros de OCR.

CAPÍTULO 6
AJUSTAR DOCUMENTOS DIGITALIZADOS

Quando abrir ou digitalizar imagens a cores e escala de cinzentos extremamente claras ou escuras, pode ser necessário ajustar essas imagens antes de executar o reconhecimento, de modo a obter resultados de OCR satisfatórios.

Para ajustar imagens:

- Abra ou digitalize um documento com escala de cinzentos a cores.

  Certifique-se de que as **definições do scanner** estão correctas.

  Importante: as **definições do scanner** e as **definições de ajuste** parecem ser as mesmas mas atente que ambos os conjuntos de definições são aplicados em fases diferentes no processo de digitalização/reconhecimento.

- No menu **Processar**, clique em **Ajustar imagem**.

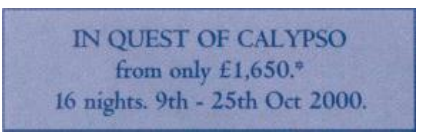
  O Readiris usa rotinas de **binarização** inteligente para converter imagens com escala de cinzentos a cores para imagens a preto e branco, que são utilizadas para executar OCR em.

o Seleccione **Suavizar a imagem a cores** para uniformizar a imagem.

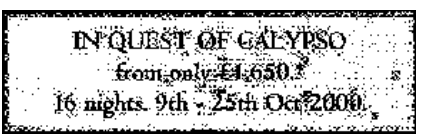
  Esta opção torna as imagens a cores e com escala de cinzentos mais homogéneas ao suavizar as diferenças na intensidade. Como

resultados, é criado um contraste mais forte entre o primeiro plano (texto) e o fundo (imagens).

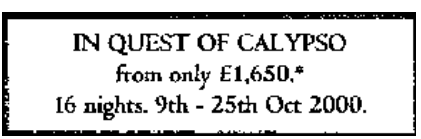
Nota: por vezes, suavizar, é a única forma de separar o texto de um fundo colorido.



(Imagem original)



(Imagem a preto e branco binarizada)



(Imagem suavizada)

- Utilize o cursor para aumentar ou diminuir a **Luminosidade**.

As definições de **Luminosidade** determinam a luminosidade geral da imagem. Utilize estas definições para escurecer ou clarear a imagem quando o texto está ilegível.

Exemplo 1: clareie uma imagem escura para eliminar o fundo da página.

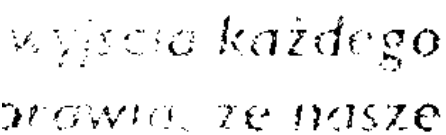
(Imagem a cores)

(Imagem binarizada. As definições de binarização predefinidas produzem uma imagem negra)

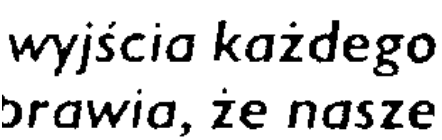
(A imagem clareada produz resultados de reconhecimento satisfatórios)

Exemplo 2: escureça uma imagem quando o texto for tão claro que não aparece na imagem binarizada.

(Imagem a cores)



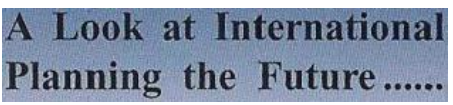
(Imagem binarizada. As definições de luminosidade predefinidas produzem caracteres fragmentados)



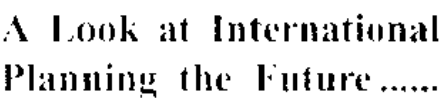
((A imagem escurecida produz resultados de reconhecimento satisfatórios)

- Utilize o cursor para aumentar ou diminuir o **Contraste**.

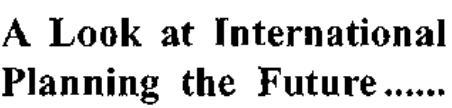
As definições de **Contraste** determinam o contraste entre as zonas mais escuras e mais claras de uma imagem. Utilize estas definições para realçar as formas dos caracteres num fundo colorido.



(Imagem a cores)



(As definições de contraste predefinidas produzem caracteres quebrados)



(As definições de contraste aumentadas produzem resultados de reconhecimento satisfatórios)

- Utilize o cursor para aumentar ou diminuir as opções de **Remoção de pontos**.

Remover pontos pequenos de imagens a preto e branco.

- Clique em **Aplicar** para pré-visualizar os resultados.
- Se os resultados forem satisfatórios, clique em **OK**. Caso contrário, altere novamente as definições.
- Clique em **Reconhecer + Guardar** para reconhecer o documento.

CAPÍTULO 7
GUARDAR DOCUMENTOS COMO FICHEIROS DE IMAGEM

Os documentos em papel que digitaliza não necessitam de ser reconhecidos imediatamente. Podem ser guardados como ficheiros de imagem.

Para o fazer:

- Digitalize o documento.
- No menu **Ficheiro**, clique nos comandos **Guardar página inteira como Imagem** ou **Guardar todas as páginas como Imagem**.
- De seguida, abra o ficheiro de imagem guardado e execute o reconhecimento.

Guardar gráficos apenas

Pode também escolher guardar as janelas de gráficos sem o texto do documento.

Para o fazer:

- Digitalize ou abra o documento.
- No menu **Ficheiro**, clique em **Guardar gráficos**.
- Todos os gráficos do documento serão guardados num único ficheiro.

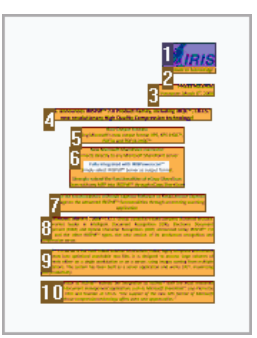
CAPÍTULO 8
DISPOR DOCUMENTOS POR JANELAS

DISPOR DOCUMENTOS POR JANELAS AUTOMATICAMENTE

Quando digitalizar ou abrir documentos, o Readiris irá aplicar automaticamente a **Análise de páginas** para dividir os documentos em janelas diferentes.

A opção **Análise de páginas** está activada por defeito. Clique no botão **Opções** e desactive a **Análise de páginas** caso pretenda evitar a análise de páginas automática.

Os resultados da análise de páginas podem ser modificados manualmente após a análise de páginas automática. Para obter mais informações, consulte a secção **Dispor manualmente documentos por janelas**.



Tipos de janelas

O Readiris utiliza tipos de janelas cinco: **blocos de texto**, **zonas de gráficos** e **tabelas**, **zonas de códigos de barras** e **zonas escritas em letra de imprensa**.

A análise de páginas detecta zonas de texto, gráficos e tabelas. As zonas de códigos de barras e as zonas escritas em letra de imprensa necessitam de ser desenhadas manualmente.

> Para obter mais informações, consulte a secção **Dispor manualmente documentos por janelas**.

Cada tipo de janela tem o seu próprio código de cor: as janelas de texto são cor-de-laranja, os gráficos são roxos e as janelas de tabelas cor-de-rosa.As zonas de códigos de barras são verdes e as zonas escritas em letra de imprensa são azuis.

As janelas são ordenadas de cima para baixo e da esquerda para a direita. Os números indicam a ordem das janelas. No entanto, a ordem e os tipos de janelas podem ser alterados. Para obter mais informações, consulte a secção **Dispor manualmente documentos por janelas**.

Ignore as zonas de texto nas margens das páginas

Quando o scanner gera margens pretas em torno da imagem, a análise das páginas tende a encontrar zonas em que só existe ruído.

Para evitar isto, clique em **Não detectar janelas nas margens** no menu **Definições** e digitalize novamente o documento.

Ignorar zona exterior

Como alternativa à disposição automática por janelas, a função **Ignorar zona exterior** pode ser utilizada. Esta função é útil quando apenas uma área em particular nas páginas do documento necessita de ser reconhecida.

Seleccione **Ignorar zona exterior** ao clicar no botão correspondente na barra de ferramentas da imagem.

Desenhe uma moldura em torno da parte da página que pretende que o Readiris reconheça.

> A parte da página que seleccionar será analisada automaticamente.
>
> Ser-lhe-á solicitado se pretende excluir a mesma zona exterior da análise de páginas em todas as páginas do documento.

DISPOR DOCUMENTOS POR JANELAS MANUALMENTE

Além de dispor documentos por janelas automaticamente através da **Análise de páginas**, o Readiris permite dispor os documentos por janelas manualmente.

A disposição por janelas manual é útil quando tiver de modificar os resultados da análise de páginas automática. Também é útil para utilizar quando criar modelos de disposição por janelas.

> Para obter mais informações sobre os modelos de disposição por janelas, consulte a secção **Utilizar modelos de disposição por janelas**.
>
> Atente que as zonas de códigos de barras e as zonas escritas em letra de imprensa necessitam sempre de ser dispostas por janelas manualmente.

Operação

- De modo a dispor um documentos por janelas manualmente, clique primeiro no botão **Opções** e anule a selecção de **Análise de páginas**.

- Abra ou digitalize o documento ao clicar no botão **Digitalizar** ou **Abrir**.

- Seleccione o **tipo de janela** correcto ao clicar no tipo de janela correspondente na barra de ferramentas da imagem.

  O Readiris utiliza cinco tipos de janelas: **blocos de texto**, **zonas de gráficos** , **tabelas**, **zonas de códigos de barras** e **zonas escritas em letra de imprensa**.

- Desenhe uma moldura em torno dos blocos de texto, gráficos, tabelas, códigos de barras e zonas escritas em letra de imprensa que pretende dispor por janelas.

  Para obter mais informações sobre o reconhecimento de códigos de barras e texto escrito em letra de imprensa, consulte as secções **Reconhecer códigos de barras** e **Reconhecer texto escrito em letra de imprensa**, respectivamente.

- Quando concluir a disposição do documento por janelas, clique no botão **Reconhecer + Guardar** para executar o OCR.

Ordenar janelas

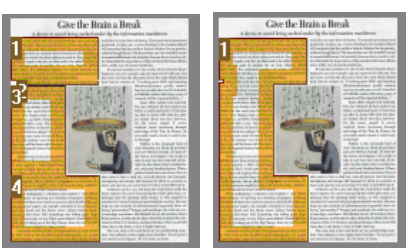
Para alterar a ordem das janelas, clique no botão **Ordenar** na barra de ferramentas da imagem e clique nas janelas uma a uma na ordem necessária.

Quando concluir, clique no botão **Reconhecer + Guardar** para executar o OCR.

As janelas em que não clica, serão excluídas do reconhecimento.

Desenhar polígonos

Dispor os documentos por janelas manualmente não está limitado a formas rectangulares. É possível criar janelas poligonais ao unir janelas rectangulares. Sempre que duas janelas do mesmo tipo se cruzam, elas tornam-se automaticamente num polígono.



Análise de páginas automática

Caso a página actual seja demasiado complexa para dispor manualmente por janelas, clique no botão **Análise de páginas** na barra de ferramentas da imagem para dispor a página por janelas automaticamente.

Atente que as zonas de códigos de barras e as zonas escritas em letra de imprensa necessitam sempre de ser desenhadas manualmente.

Alterar o tipo de janela

Para alterar o tipo de janela de uma janela, clique com o botão direito na janela, seleccione **Janela**, de seguida seleccione **Tipo** e depois clique no tipo de janela necessário.

É também possível alterar o tipo de janela de várias janelas em simultâneo:

- Clique no botão do **ponteiro** na barra de ferramentas da imagem.

- Mantenha premida a tecla **Shift** enquanto selecciona múltiplas janelas.
- Clique com o botão direito nas janelas seleccionadas, seleccione **Janela**, de seguida seleccione **Tipo** e depois clique no tipo de janela necessário.

Modificar o tamanho da janela

Clique na janela que pretende modificar.

Coloque o cursor do rato sobre um marcador (nos lados e nos cantos da janela).

Clique no marcador e arraste o rato para modificar o tamanho da janela.

Mover janelas

Seleccione a janela que pretende mover.

Clique dentro da janela e arraste o rato para modificar a posição da janela.

Reconhecer uma janela em particular

Clique com o botão direito na janela que pretende reconhecer e seleccione **Copiar como texto**.

> Os resultados são enviados para a área de transferência como corpo de texto. Isto também funciona para texto escrito em letra de imprensa.
>
> As janelas de gráficos e as janelas de códigos de barras também pode ser copiadas para a área de transferência.

Eliminar janelas

Seleccione a(s) janela(s) que pretende eliminar ou clique no comando **Seleccionar tudo** no menu **Editar** para seleccionar todas as janelas.

Seleccione os comandos **Cortar** ou **Apagar** no menu **Editar** para cortar ou eliminar as janelas.

ou

Clique com o botão direito nas janelas seleccionadas, seleccione **Janela**, de seguida, clique em **Eliminar**.

Eliminar janelas pequenas

Alguns documentos, faxes por exemplo, têm frequentemente pontos "espalhados" nas páginas, fazendo com que o Readiris crie janelas supérfluas que não contêm texto.

Para apagar todas as janelas pequenas, clique em **Eliminar janelas pequenas** no menu **Editar**.

Esta opção apaga todas as janelas mais pequenas que 0,5" e ordena novamente as restantes zonas.

UTILIZAR MODELOS DE DISPOSIÇÃO DE JANELAS

Quando reconhecer vários documentos com um esquema de página semelhante, convém utilizar os **modelos de disposição por janelas** em vez de análise de páginas automática. Dessa forma, é aplicada a mesma estrutura de janelas a todos os documentos digitalizados ou abertos, o que acelera o processo.

Operação

- Disponha manualmente a primeira página do documento por janelas utilizando os botões da barra de ferramentas da imagem.

  Para obter mais informações, consulte a secção **Dispor manualmente documentos por janelas**.

- No menu **Ficheiro**, clique no comando **Guardar esquema**.
- Abra ou digitalize as outras páginas do documento ao clicar no botão **Abrir** ou **Digitalizar** na barra de ferramentas principal.
- No menu **Ficheiro**, clique no comando **Carregar esquema**.
- Seleccione o ficheiro de esquema que guardou.
- Para aplicar o esquema a **todas** as páginas abertas ou digitalizadas, seleccione **Aplicar esquema a todas as páginas** na caixa de diálogo **Ficheiro de esquema**.
- Clique em**Abrir** para carregar o ficheiro de esquema.

  Atente que quando adiciona um documento ao Readiris, o ficheiro de esquema tem de ser novamente carregado visto que a análise de páginas está activada por defeito.

Ignorar zona exterior

Como alternativa aos modelos de disposição por janelas, é possível utilizar a opção **Ignorar zona exterior.** Dessa forma, é possível

definir uma área em particular na página que necessita de ser reconhecida. Quaisquer dados fora da área de OCR serão excluídos do reconhecimento.

Operação

- Seleccione **Ignorar zona exterior** ao clicar no botão correspondente na barra de ferramentas da imagem.

- Desenhe uma moldura em torno da parte da página que pretende que o Readiris reconheça.

  A parte da página que seleccionar será analisada automaticamente.

  Ser-lhe-á solicitado se pretende ignorar a mesma zona exterior para todas as páginas do documento.

  Para cancelar esta função, execute novamente **Análise de páginas** ao clicar no botão **Análise de páginas** na barra de ferramentas da imagem.

- Clique em **Reconhecer + Guardar** para executar o OCR.

CAPÍTULO 9
RECONHECER DOCUMENTOS

INTRODUÇÃO

Para reconhecer documentos, o Readiris aplica linguística durante a fase de reconhecimento. Como resultado, o Readiris reconhece texto, tabelas e gráficos, códigos de barras e texto escrito em letra de imprensa em todos os tipos de documentos. O Readiris suporta até documentos com colunas complexos, documentos de baixa qualidade, faxes, impressões com matriz de pontos, documentos mal digitalizados e copiados com tipos de letras demasiado claras ou escuras, etc.

O Readiris suporta 128 **idiomas**: são suportados todos os idiomas americanos e europeus, incluindo os idiomas Centro-europeus, Bálticos e Cirílicos assim como o Grego e o Turco. Opcionalmente, o Readiris consegue ler documentos em Árabe, Persa e Hebraico e quatro idiomas Asiáticos - Japonês, Chinês simplificado e tradicional e Coreano. O Readiris suporta até alfabetos mistos: o software detecta palavras "Ocidentais" que ocorrem em documentos em Grego, Cirílico, Árabe, Hebraico e Asiático - muitos nomes próprios não convertíveis, nomes de marcas etc. são escritos utilizando os símbolos Ocidentais.

O Readiris baseia-se nas tecnologias de reconhecimento mais avançadas. O reconhecimento de texto com tipo de letra independente é completado por técnicas auto-aprendizagem. O sistema consegue aprender novos caracteres e palavras através de análise contextual e linguística. Isto significa que a precisão de

OCR do sistema de reconhecimento irá melhorar à medida que avança.

Além disso, o Readiris tem uma função de verificação de utilizador (opcional). Quando activada, a função de verificação de utilizador (**Aprendizagem interactiva**) não só assinala os caracteres que o sistema de reconhecimento não tem a certeza mas também permite aumentar a precisão do sistema. Todas as soluções que confirmar são memorizadas temporariamente durante o reconhecimento, aumentando a velocidade e a confiança do sistema e tornando o sistema mais inteligente à medida que avança. Esta ferramenta de aprendizagem avançada também permite treinar o Readiris nos caracteres especiais como símbolos matemáticos e dingbats e suportar tipos de letra distorcidos.

Os resultados de aprendizagem interactivos podem também ser guardados permanentemente em **dicionários de tipos de letra** para utilização futura.

SELECCIONAR O IDIOMA DO DOCUMENTO

O Readiris oferece OCR em 128 idiomas. O Readiris suporta todos os idiomas americanos e europeus, incluindo os idiomas Centro-europeus, Cirílicos e Bálticos assim como o Grego e o Turco.

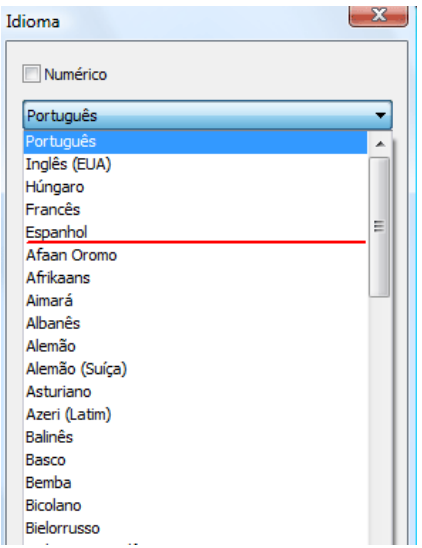
O Readiris Pro Asiático e o Readiris Corporate Asiático reconhece adicionalmente documentos em Japonês, Chinês simplificado, Chinês tradicional e Coreano.

O Readiris Pro Médio-Oriente e o Readiris Corporate Médio-Oriente reconhece adicionalmente documentos em Árabe, Persa e Hebraico.

De modo ao Readiris reconhecer um documento, o idioma do documento tem de ser especificado.

Para o fazer:

Clique no botão **globo** na barra de ferramentas principal e seleccione o idioma à sua escolha na lista pendente.



Os 5 idiomas mais seleccionados recentemente são deslocados para o topo da lista de idiomas.

Importante: seleccione o idioma do documento *antes*de executar a análise da página quando lida com documentos asiáticos, hebraicos e arábicos. São utilizadas rotinas de análise de página específicas para estes documentos.

O reconhecimento pode também estar limitado a um carácter **numérico** definido para reconhecer idealmente tabelas e valores. Então, o Readiris só reconhece os numerais 0-9 e a série seguinte de símbolos:

+ sinal mais	* asterisco	/ barra	% sinal de percentagem	, vírgula
. ponto	(abrir parêntesis	) fechar parêntesis	- hífen	= sinal de igual
$ sinal de dólar	£ sinal de libra	€ sinal de euro	¥ sinal de iene	

Para activar o modo numérico, seleccione **Numérico** na caixa de diálogo **Idioma**.

Reconhecer documentos com idiomas mistos

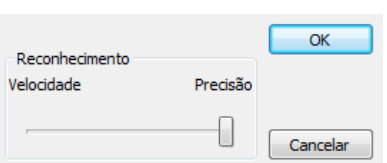
O Readiris também permite activar conjuntos de caracteres mistos. Dessa forma o Readiris muda automaticamente de idioma a meio de uma frase e reconhece palavras em Inglês (nomes próprios etc.) que ocorrem em idiomas "exóticos".

Clique no botão **globo** na barra de ferramentas principal e seleccione a combinação de idiomas necessária na lista pendente de idiomas.

Nota: quando processar documentos asiáticos ou hebraicos, os conjuntos de caracteres mistos são usados automaticamente.

Velocidade-Precisão

Seleccione a relação correcta entre **velocidade de OCR** e **precisão de OCR**.



Esta relação está disponível para os alfabetos latino, cirílico e grego.

Conselho: favoreça a precisão sobre a velocidade quando a qualidade de imagem for bastante fraca.

DEFINIR AS CARACTERÍSTICAS DO DOCUMENTO

Junto do idioma do documento, outras características do documento como o **Tipo de letra** e **Espaço entre caracteres** desempenham uma função importante no processo de reconhecimento.

Tipo de letra

O Readiris distingue entre documentos impressos "normais" e com matriz de pontos. Os símbolos da matriz de pontos (do tipo de 9 pinos) são feitos de pontos isolados, separados.

São necessárias técnicas de segmentação e reconhecimento especiais para reconhecer documentos com matriz de pontos e necessitam de ser activadas.

Far out in the uncharted back

Para seleccionar o tipo de letra:

- No menu **Definições**, seleccione **Tipo de letra**.
- O tipo de letra está predefinido para **Automático**.

  Dessa forma, o Readiris reconhece matriz de pontos de "25 pinos" ou "NLQ" (Qualidade intermédia) ou outra impressão "normal".

- Para reconhecer apenas documentos impressos com matriz de pontos, clique em **Matriz de pontos**.

  O Readiris irá reconhecer documentos impressos com matriz de pontos denominados ""draft"" ou de "9 pinos".

Espaço entre caracteres

O espaço entre caracteres é o número de caracteres por polegadas num determinado tipo de letra. O espaço de caracteres pode ser **fixo**, neste caso todos os caracteres têm a mesma largura ou **proporcional**, neste caso os caracteres têm um largura diferente.

Para seleccionar o espaço entre caracteres:

- No menu **Definições**, seleccione **Espaço entre caracteres**.
- O espaço entre caracteres está predefinido para **Automático**.
- Clique em **Fixo** se todos os caracteres do tipo de letra tiverem a mesma largura. Isto é frequente nos documentos escritos à máquina antigos.
- Clique em **Proporcional** se os caracteres do tipo de letra tiverem uma largura diferente. Praticamente todos os tipos de letra em jornais, revistas e livros são proporcionais.

Importante: estas características de documentos não se aplicam a documentos asiáticos, hebraicos ou arábicos.

UTILIZAR APRENDIZAGEM INTERACTIVA

O Readiris oferece uma função de aprendizagem interactiva. Através de **Aprendizagem interactiva** é possível treinar o sistema de reconhecimento acerca dos tipos de letra e formas de caracteres e corrigir os resultados de OCR, se necessário. Durante a aprendizagem interactiva, quaisquer caracteres que o sistema de

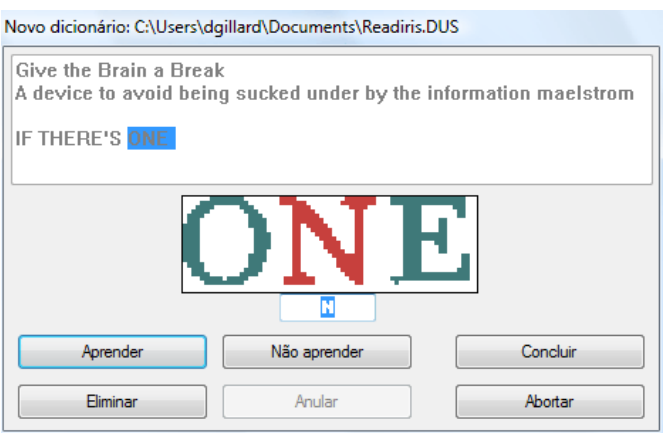
reconhecimento não tem a certeza são apresentados numa janela de pré-visualização, em combinação com a palavra semelhante e a solução proposta.

A Aprendizagem interactiva pode melhorar substancialmente a precisão do sistema de reconhecimento e é particularmente útil quando reconhecer formas distorcidas, deformadas. A Aprendizagem interactiva pode também ser utilizada para treinar o Readiris acerca dos símbolos especiais que não consegue reconhecer inicialmente, como símbolos matemáticos e científicos e dingbats.

Para activar a aprendizagem interactiva:

- No menu **Aprender** clique em **Aprendizagem interactiva**.
- Determine as definições necessárias e clique no botão **Reconhecer + Guardar** para executar o documento.
- No final do reconhecimento, o Readiris entra na fase de aprendizagem interactiva.

São apresentados os caracteres que o sistema de reconhecimento não tem a certeza.



Se os resultados estiverem correctos:

- Clique no botão **Aprender** para guardar o resultado como certo.

  Os resultados de aprendizagem são guardados temporariamente na memória do computador, durante o reconhecimento. O Readiris não irá apresentar os caracteres aprendidos quando reconhecer o resto do documento.

  Quando é um documento novo é reconhecido, os resultados de aprendizagem são apagados.

  Para guardar permanentemente os resultados de aprendizagem , utilize um dicionário de tipos de letra. Para obter mais informações, consulte a secção **Utilizar dicionários de tipos de letra**.

- Clique em **Concluir** para guardar todas as soluções que o software oferece.

Se os resultados estiverem incorrectos:

- Introduza os caracteres correctos e clique no botão **Aprender**.

ou

- Clique em **Não aprender** para guardar o resultado como incerto.

  Utilize este comando para caracteres danificados que podem ser confundidos com outros caracteres quando aprendidos, por exemplo, o número 1 e a letra I, que têm uma forma idêntica em muitos tipos de letra.

- Clique em **Eliminar** para eliminar caracteres da exportação.

  Utilize este botão para evitar que ruído do documento surja no ficheiro de exportação.

- Clique em **Anular** para corrigir erros.

  O Readiris mantém registo das últimas 32 operações.

- Clique em **Abortar** para abortar a aprendizagem interactiva.

Todos os resultados de aprendizagem serão eliminados. Da próxima vez que clicar em **Reconhecer + Guardar**, a aprendizagem interactiva irá iniciar novamente.

UTILIZAR DICIONÁRIOS DE TIPOS DE LETRA

Quando digitalizar muitos documentos do mesmo tipo, qualidade do tipo de letra e da impressão, pode não pretender repetir sempre o processo de aprendizagem. Por isso, convém utilizar dicionários de tipos de letras. Os dicionários de tipos de letras contêm informações sobre os tipos de letra aprendidos durante a **aprendizagem interactiva** e pode aumentar substancialmente os resultados de reconhecimento.

Atente que os dicionários de tipos de letra estão limitados a 500 formas. É aconselhável criar dicionários separados para aplicações específicas.

Para criar um dicionário de tipos de letra novo:

- No menu **Aprender**, clique no comando **Novo dicionário de tipos de letra**.

  O Readiris irá abrir o dicionário **Readiris.dus** por defeito. Altere o nome de ficheiro e clique em **Guardar** para guardá-lo como um dicionário novo.

- Clique em **Aprendizagem interactiva** no menu **Aprender** para a activar.

- Clique em **Reconhecer + Guardar** para reconhecer o documento.

- O Readiris introduz a fase de aprendizagem interactiva. Utilize os botões da caixa de diálogo para guardar caracteres no dicionário de tipos de letra.

Para utilizar um dicionário de tipos de letra existente:

- No menu **Aprender** clique em **Dicionário de tipos de letra**.
- Seleccione o dicionário que pretende utilizar e clique em **Abrir**.
- No menu **Aprender** clique em **Anexar dicionário de tipos de letra** ou **Ler dicionário de tipos de letra**.

  Quando seleccionar **Anexar dicionário de tipos de letra**, certifique-se de que activa a **Aprendizagem interactiva**. O Readiris irá reconhecer as formas dos caracteres guardadas no dicionário e utilize a aprendizagem interactiva, permitindo guardar novas informações no dicionário.

  Quando seleccionar **Ler dicionário de tipos de letra**, o Readiris irá reconhecer as formas de caracteres guardadas no dicionário mas não irá adicionar novo conteúdo, mesmo que a **Aprendizagem interactiva** esteja activada. Atente que ainda é útil utilizar a **Aprendizagem interactiva** para verificar e se necessário, corrija os resultados de reconhecimento que não estão guardados no dicionário de tipos de letra.

  Atenção: não clique em **Dicionário de tipos de letra** no menu **Aprender** e abra um dicionário existente enquanto o modo de dicionário **Novo dicionário** está activado. Caso contrário, os conteúdos do dicionário de tipos de letra existente serão apagados.

- Clique em **Reconhecer + Guardar** para reconhecer o documento.

CAPÍTULO 10
FORMATAR E GUARDAR DOCUMENTOS

FORMATAR DOCUMENTOS

Os documentos que reconhece no Readiris podem ser guardados em vários formatos de exportação. O Readiris guarda os resultados de OCR como ficheiros PDF Adobe Acrobat, ficheiros Microsoft XPS, ficheiros de texto Word, WordML, RTF e OpenDocument, ficheiros HTML e XML, folhas de cálculo SpreadsheetML e ficheiros de texto Ansi e Unicode.

Além disso, o Readiris pode exportar resultados directamente para essas aplicações como Microsoft Word e Excel, Adobe Reader, Microsoft XPS Viewer, os principais navegadores de Internet e software de e-mail etc.

> Nota: Os ficheiros RTF criados pelo Readiris também podem ser abertos nas aplicações com base na Internet AjaxWrite, ThinkFree, Zoho Writer e Google Writely (que abre documentos RTF e HTML). Quando utilizar uma destas aplicações, certifique-se de que selecciona a opção de esquema **Reter formatação de palavra e texto**.

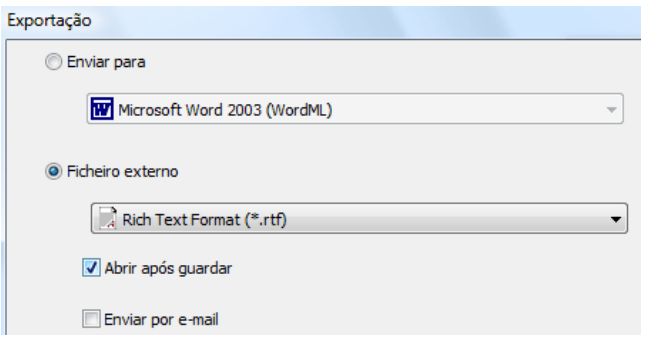
Operação

- Clique no botão **Formato** na barra de ferramentas principal para seleccionar o modo de exportação.

  O Readiris:

- envia documentos para uma aplicação, que irá abrir automaticamente, ou;
- guarda documentos como um ficheiro externo.

A opção **Enviar por e-mail** cria uma nova mensagem de e-mail e introduz o documento reconhecido como anexo do e-mail.



- Clique nos diferentes separadores para seleccionar as definições que pretende aplicar.

  As definições que estão indisponíveis para o formato de exportação seleccionado surgem desactivadas.

  Os formatos de ficheiro de exportação assim como as suas opções são discutidos nas secções abaixo.

- Clique em **Reconhecer + Guardar** para executar o reconhecimento e guardar os documentos.

  Os resultados de OCR podem ser exportados várias vezes sem repetir o reconhecimento. Clique novamente no botão **Formato** e altere o formato de texto e as opções de formatação. De seguida, clique novamente em **Reconhecer + Guardar**.

  Para procurar e ordenar razões, o Readiris permite definir **propriedades do documento** de exportação PDF, XPS, Word, RTF, WordML, SpreadsheetML e HTML. Para definir as propriedades de documento de um documento, clique em **Propriedades de**

documento no menu **Ficheiro**. Atente que as opções das propriedades de documento também acessíveis na caixa de diálogo **Ficheiro de exportação**, que abre quando clica em **Reconhecer + Guardar**.

Atente que, quando guardar um documento multi-páginas como **ficheiro externo**, é possível criar um ficheiro de exportação em separado no Readiris ou guardar todas as páginas que pertencem ao mesmo documento para um único ficheiro de exportação. Basta clicar nas opções correspondentes na caixa de diálogo **Ficheiro de exportação**: **Criar um ficheiro por página** e **Criar um ficheiro por documento**, respectivamente.

Atente que, no entanto, as opções **Criar um ficheiro por página** e **Criar um ficheiro por documento** só estão disponíveis quando guardar documentos como ficheiro externo e não quando abrir documentos na aplicação pretendida.

FORMATAR DOCUMENTOS DE TEXTO

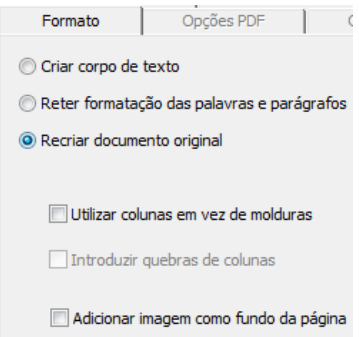
Com o Readiris é possível gerar vários tipos de formatos de exportação baseados em texto. O Readiris oferece exportação de texto Word,WordML, RTF, txt e OpenDocument.

Para gerar ficheiros de exportação baseados em texto:

- Clique no botão **Formato** na barra de ferramentas principal e seleccione o formato de exportação à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**.
- Dependendo do formato de texto que seleccionou, estão disponíveis várias opções de formatação.

  Quaisquer opções que estão indisponíveis para o formato de texto seleccionado surgem desactivadas.

Opções de esquema



- A opção **Criar corpo de texto** evita a formatação de texto por parte do Readiris.

  O Readiris gera um texto corrido, contínuo.

- A opção **Reter formatação das palavras e parágrafo** ocupa uma posição intermédia entre corpo de texto e formatação automática.

  O tipo, tamanho e estilo de letra são mantidos ao longo do reconhecimento.

  Os separadores e o alinhamento de cada bloco são recriados.

  Os blocos de texto e as colunas não são recriados; os parágrafos simplesmente seguem-se uns aos outros.

  As tabelas são novamente captadas correctamente.

- A opção **Recriar documento original** recria uma cópia fac-simile do documento original.

  O Readiris gera uma cópia autêntica do documento original, deixa de ser uma imagem digitalizada.

  O Readiris também recria quaisquer hiperligações para endereços de correio electrónico e web sites.

- A opção **Utilizar colunas em vez de molduras** cria documentos com colunas.

  Os textos com colunas são mais fáceis de editar que os documentos com várias molduras: o texto flui naturalmente de uma coluna para a próxima.

  Nota: quando o sistema não consegue detectar colunas no documento original, este modo de formatação utiliza molduras como uma posição de reserva.

- A opção **Introduzir quebras de colunas** introduz uma quebra de coluna no final de cada coluna.

  Qualquer texto que edita, adiciona ou retira, permanece dentro da coluna; nenhum texto flui automaticamente através de uma quebra de coluna.

  Conselho: desactive esta opção quando colocar o corpo de texto em colunas. Irá garantir o fluxo natural do texto de uma coluna para a outra,

- A opção **Adicionar imagem como fundo da página** coloca a imagem digitalizada como fundo da página sob o texto reconhecido.

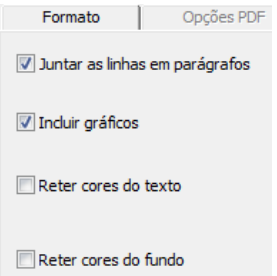
  No entanto, esta opção aumenta substancialmente o tamanho de ficheiro dos ficheiros de exportação.

  O formato **Texto-Imagem PDF** fornece o mesmo resultado para ficheiros PDF.

  A opção **Reter cores do fundo** no separador **Opções** fornece uma alternativa menos drástica, mais compacta.

Opções gerais

- Clique no separador **Opções** para seleccionar as opções gerais:



- A opção **Juntar as linhas em parágrafos** permite detecção automática de parágrafos.

  O Readiris molda o texto reconhecido até iniciar um novo parágrafo e "volta a unir" as palavras com hífen no final de uma linha.

- A opção **Incluir gráficos** inclui os gráficos em ficheiros com formatação automática.

  Isto é essencial para criar uma cópia real de um documento.

  Utilize as opções de gráficos no separador **Gráficos** para determinar o modo de calor e resolução dos gráficos guardados dentro dos ficheiros de exportação.

- A opção **Reter cores do texto** mantém as cores originais do texto ao longo do reconhecimento.

- A opção **Reter cores do fundo** mantém as cores do fundo da página ao longo do reconhecimento.

  Uma cor de fundo uniforme é criada por parágrafo no ficheiro de exportação.

Tamanhos do papel

Dependendo do formato que seleccionou, é possível seleccionar os tamanhos de papel preferidos:

- Clique no separador **Tamanho do papel** e utilize os botões direccionais para aplicar e excluir tamanhos de papel.

- O Readiris irá analisar os tamanhos de papel activos pela ordem indicada e irá utilizar o primeiro tamanho de papel que é suficientemente grande para suportar o documento digitalizado.

FORMATAR DOCUMENTOS BASEADOS EM TABELAS

Com o Readiris é possível exportar dados tabulares para folhas de cálculo, processadores de texto e navegadores de Internet: as tabelas são reconstruídas célula por célula em folhas de cálculo e introduzidas como objectos de tabela em ficheiros de processador de texto.

O Readiris reconhece tabelas com grelhas e tabelas sem grelhas.

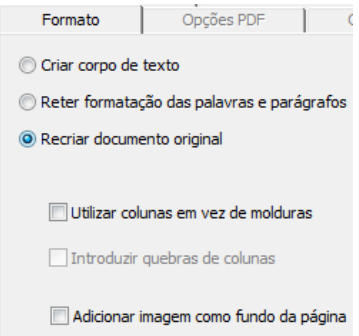
Performance test optical media	
CD-ROM Digital Versatile Disk	Average time (n
CD-ROM 32x speed	60
DVD	58
Tested on 333 MHz Pentium II	

123.985	69	31
987.745	129	2
19	149	91
287.410	49	52

(com grelhas) (sem grelhas)

Para gerar documentos baseados em tabelas:

- Clique no botão **Formato** na barra de ferramentas principal e seleccione o formato de exportação à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**.
- Seleccione a opção de esquema à sua escolha.

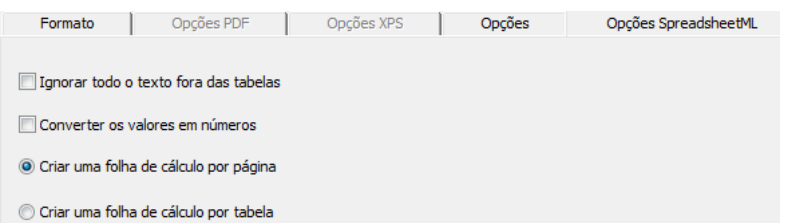


Para obter mais informações sobre as opções de formatação, consulte a secção **Formatar documentos de texto**.

Opções SpreadsheetML

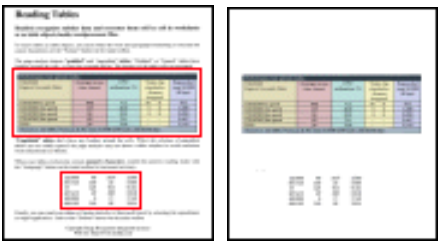
Quando seleccionar **Microsoft Excel 2002/2003** como a aplicação pretendida, estão disponíveis opções **SpreadsheetML** específicas. Clique no separador **Opções SpreadsheetML** para apresentar as opções:

Atente que a opção de esquema **Recriar documento original** fica indisponível quando este formato é seleccionado.



- A opção **Ignorar todo o texto fora das tabelas** guarda as tabelas e ignora todos os outros resultados de reconhecimento.

  Todos os dados dentro das tabelas são novamente captados; quaisquer dados fora das tabelas não o são.



- A opção **Converter os valores em números** codifica os valores reconhecidos em números.

  Como resultado, é possível executar operações aritméticas nessas células. As células de texto (em qualquer tabela) permanecem texto.

  Atente que apenas os valores dentro das tabelas são codificados como números.

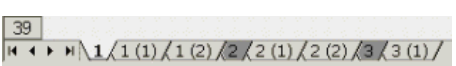
  O Excel executa exclusivamente operações matemáticas em dados codificados como números.

- A opção **Criar uma folha de cálculo por página** certifica-se de que é criada uma folha de cálculo por página digitalizada.

  Se uma página tiver tabelas e texto, serão todos colocados na mesma folha de cálculo.

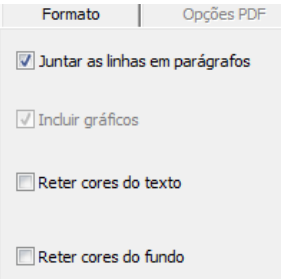
- A opção **Criar uma folha de cálculo por tabela** coloca cada tabela numa folha de cálculo em separado e inclui o texto reconhecido (fora das tabelas) noutra folha de cálculo.

  Se o documento reconhecido tiver **várias páginas**, verá essa estrutura repetida por página.



Opções gerais

Clique no separador **Opções** para seleccionar as opções gerais:



- A opção **Juntar as linhas em parágrafos** permite detecção automática de parágrafos.

  O Readiris molda o texto reconhecido até iniciar um novo parágrafo e "volta a unir" as palavras com hífen no final de uma linha.

- A opção **Reter cores do fundo** recria a cor de fundo de cada célula.

	A	B	C
1	Performance optical media		
2	CD-ROM	Average access	CPU
3	Digital Versatile Disk	time (msec)	utilization (%)
4	CD-ROM 24x speed	80	58.2
5	CD-ROM 32x speed	60	72.1
6	DVD	58	78.9
7	Tested on 333 MHz Pentium II with 64 MB RAM and 4 GB HD		

Tamanhos do papel

Dependendo do formato que seleccionou, é possível indicar os tamanhos de papel preferidos:

- Clique no separador **Tamanho do papel** e utilize os botões direccionais para aplicar e excluir tamanhos de papel.

- O Readiris irá analisar os tamanhos de papel activos pela ordem indicada e irá utilizar o primeiro tamanho de papel que é suficientemente grande para suportar o documento digitalizado.

CRIAR DOCUMENTOS PDF

O Readiris gera quatro tipos de exportação **PDF**: **Texto**,**Texto-Imagem**, **Imagem-Texto** e **Imagem**.

Para gerar exportação PDF:

Clique no botão **Formato** na barra de ferramentas principal e seleccione o tipo de PDF à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**:

Imagem PDF

Quando selecciona **Imagem** PDF, o Readiris gera documentos PDF apenas de imagem, não executa o OCR.

Imagem-Texto PDF

Quando selecciona **Imagem-Texto** PDF, o Readiris reconhece texto e cria ficheiros PDF pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está colocada na parte superior do texto.

Com este formato pode visualizar sempre o documento original (como foi digitalizado), fazer procuras e copiar-colar o texto reconhecido, que está oculto sob a imagem. Como resultado, este formato é útil para **fins de arquivamento**.

Texto PDF

Quando selecciona **Texto** PDF, o Readiris reconhece o texto e cria ficheiros PDF pesquisáveis.

A imagem da página não está contida nestes ficheiros PDF de camada única.

Texto-Imagem PDF

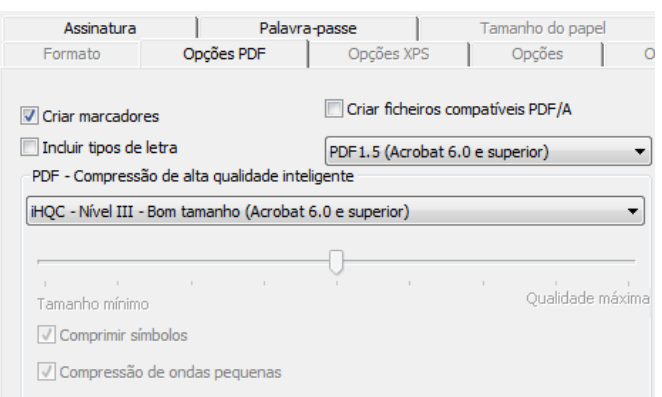
Quando selecciona **Imagem-Texto** PDF, o Readiris reconhece texto e cria documentos PDF pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está contida sob o texto.

SELECCIONAR AS OPÇÕES PDF

Para seleccionar as opções PDF:

- Clique no botão **Formato** na barra de ferramentas principal e seleccione o tipo de PDF à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**.
- Dependendo do **tipo de PDF** que seleccionou, estão disponíveis várias opções. Clique no separador **Opções de PDF** para aceder às opções:



Criar marcadores

A opção **Criar marcadores** cria marcadores para cada bloco de texto, gráfico e tabela em ficheiros PDF Adobe Acrobat.

Incluir tipos de letra

Seleccione a opção **Incluir tipos de letra** para incluir tipos de letra em ficheiros PDF Adobe Acrobat.

Incluir tipos de letra previne a substituição do tipo de letra e garante que os leitores, independentemente da configuração do computador, visualizem o texto com os tipos de letra originais.

Incluir tipos de letra aumenta ligeiramente o tamanho do ficheiro dos documentos reconhecidos.

DOCUMENTOS PDF COMPRIMIDOS POR IHQC

Além de quatro tipos de exportação PDF "normal", a Readiris oferece a exportação PDF comprimida por iHQC: Os documentos PDF dos tipos **Imagem-Texto** e **Imagem** pode ser altamente comprimidos através da iHQC sem perda da qualidade de imagem.

iHQC significa **intelligent High-Quality Compression (Compressão inteligente de alta qualidade)**, a tecnologia de compressão eficaz, propriedade da I.R.I.S.. A iHQC está para as imagens como o MP3 está para a música e o DivX para os filmes.

Para gerar exportação PDF comprimida por iHQC:

- Clique no botão **Formato** na barra de ferramentas principal e escolha entre os dois modos de exportação.
- Na lista **Enviar para** ou **Ficheiro externo**, seleccione o tipo de PDF à sua escolha: **Imagem-Texto PDF** ou **Imagem PDF**.

- No separador **Opções PDF**, seleccione o nível de compressão necessário.

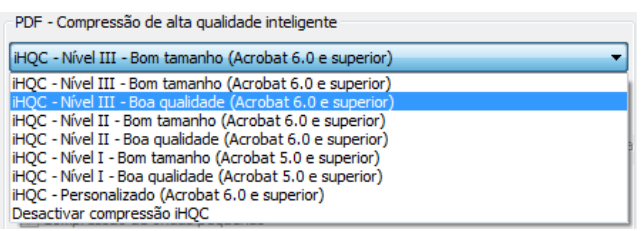
  O Readiris Pro suporta **Nível I - Tamanho bom** e **Nível I - Compressão de boa** qualidade.

  O Readiris Corporate também suporta **Nível II e III Compressão de tamanho bom** e **Boa qualidade** assim como compressão **Personalizada**.

  Na compressão de Nível II a opção **Comprimir símbolos** é automaticamente activada para comprimir o texto de forma compacta.

  Na compressão de Nível III também a opção **Compressão de ondas pequenas** é automaticamente activada para comprimir gráficos de forma compacta.

  Quando selecciona a compressão **Personalizada**, é possível activar ou desactivar estas opções de forma independente. Também é possível utilizar o cursor para definir a relação **Tamanho bom - Boa qualidade**.



- Clique em **Reconhecer + Guardar** para reconhecer o documento e guardá-lo como ficheiro PDF comprimido por iHQC.

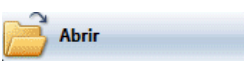
CONVERTER DOCUMENTOS PDF

Juntamente com gerar documentos PDF, o Readiris também pode converter ficheiros PDF: O Readiris converte PDFs de **imagem** em PDFs de **texto** ou outro formato de texto suportado e **desbloqueia** conteúdo PDF só de leitura.

> **Aviso**: O Readiris não abre documentos PDF protegidos por palavra-passe.

Operação

- Clique no botão **Abrir** na barra de ferramentas principal e seleccione o ficheiro PDF que pretende que o Readiris converta.



> Na área **Intervalo de páginas** da caixa de diálogo, seleccione **Páginas** e indique quais as páginas que pretende que o Readiris abra.

- Clique no botão **Abrir** na caixa de diálogo para abrir o ficheiro PDF à sua escolha.
- Clique no botão **Formato** na barra de ferramentas principal e seleccione o tipo de PDF à sua escolha.

> Para obter mais informações sobre os tipos de PDF, consulte a secção **Criar documentos PDF**.

- Clique no botão **Reconhecer + Guardar** para converter o documento.

CRIAR DOCUMENTOS XPS

O Readiris gera quatro tipos de ficheiros XPS: **Texto**, **Texto-Imagem**, **Imagem-Texto** e **Imagem**.

XPS significa XML Paper Specification (Especificação do papel XML) e é um formato de esquema fixo desenvolvido pela Microsoft.

Para gerar exportação XPS:

Clique no botão **Formato** na barra de ferramentas principal e seleccione o tipo de XPS à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**:

Imagem XPS

Quando selecciona **Imagem** XPS, o Readiris gera documentos XPS apenas de imagem, não executa o OCR.

Imagem-Texto XPS

Quando selecciona **Imagem-Texto** XPS, o Readiris reconhece texto e cria ficheiros XPS pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está colocada na parte superior do texto.

Com este formato pode visualizar sempre o documento original (como foi digitalizado), fazer procuras e copiar-colar o texto reconhecido, que está oculto sob a imagem. Como resultado, este formato é útil para **fins de arquivamento**.

Texto XPS

Quando selecciona **Texto** XPS, o Readiris reconhece o texto e cria ficheiros XPS pesquisáveis.

A imagem da página não está contida nestes ficheiros XPS de camada única.

Texto-Imagem XPS

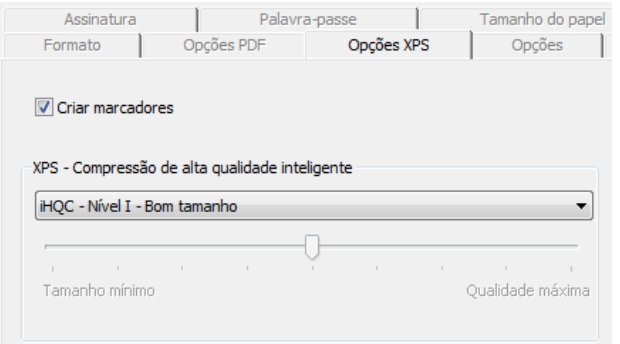
Quando selecciona **Texto-Imagem** XPS, o Readiris reconhece texto e cria documentos XPS pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está contida sob o texto.

SELECCIONAR AS OPÇÕES XPS

Para seleccionar as opções XPS:

- Clique no botão **Formato** na barra de ferramentas principal e seleccione o tipo de XPS à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**.
- Dependendo do **tipo de XPS** que seleccionou, estão disponíveis várias opções. Clique no separador **Opções de XPS** para aceder às opções:



Criar marcadores

A opção **Criar marcadores** cria marcadores para cada bloco de texto, gráfico e tabela em ficheiros XPS da Microsoft.

DOCUMENTOS XPS COMPRIMIDOS POR IHQC

Além de quatro tipos de exportação XPS "normal", a Readiris oferece a exportação XPS comprimida por iHQC. Os documentos XPS dos tipos **Imagem-Texto** e **Imagem** pode ser altamente comprimidos através da iHQC.

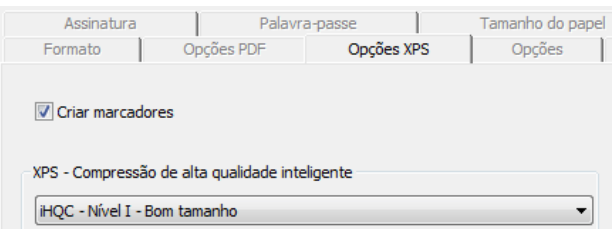
iHQC significa **intelligent High-Quality Compression (Compressão inteligente de alta qualidade)**, a tecnologia de compressão eficaz, propriedade da I.R.I.S.. A iHQC está para as imagens como o MP3 está para a música e o DivX para os filmes.

Para gerar exportação XPS comprimida por iHQC:

- Clique no botão **Formato** na barra de ferramentas principal e escolha entre os dois modos de exportação.

- Na lista **Enviar para** ou **Ficheiro externo**, seleccione o tipo de XPS à sua escolha: **Imagem-Texto XPS** ou **Imagem XPS**.
- No separador **Opções XPS**, seleccione o nível de compressão necessário.

  O Readiris suporta **Nível I - Tamanho bom** e **Nível I Compressão de boa** qualidade.

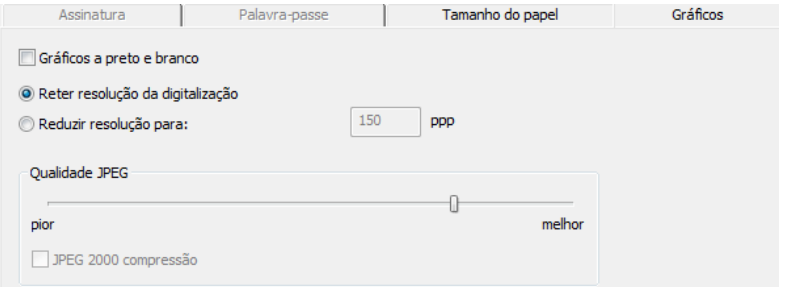


SELECCIONAR AS OPÇÕES DOS GRÁFICOS

Dependendo do formato de exportação e aplicação pretendida que seleccionar, podem estar disponíveis opções de gráficos avançadas. As opções de gráficos podem ser utilizadas para alterar a qualidade de imagem e resolução.

Para aceder às opções dos gráficos:

- Clique no botão **Formato** na barra de ferramentas principal e seleccione o formato de exportação à sua escolha na lista pendente **Enviar para** ou **Ficheiro externo**.
- Clique no separador **Gráficos** para apresentar as opções.



Modo a cores

O Readiris guarda gráficos a **cor** por defeito.

Seleccione **Preto e branco** para guardar gráficos a preto e branco.

Resolução

O Readiris **retém a resolução de digitalização** por defeito.

Pode também escolher **reduzir a resolução**.

Conselho: Quando guardar documentos como ficheiros HTML para colocar num web site, reduza a resolução para 70 ppp (resolução de ecrã).

Qualidade JPEG

Gráficos guardados dentro de documentos PDF, XPS, Word e RTF são guardados no formato JPEG.

Utilize o cursor para ajustar a qualidade JPEG.

Compressão JPEG 2000

Quando guardar ficheiros no formato PDF ou XPS , o Readiris pode aplicar a compressão JPEG 2000 às imagens a cores-escala de cinzentos guardadas dentro desses ficheiros.

JPEG 2000 é a versão mais recente e mais compacta da norma JPEG.

Seleccione a opção **Compressão JPEG 2000** para a aplicar.

CAPÍTULO 11
GUARDAR E CARREGAR AS DEFINIÇÕES

Quaisquer definições que defina no Readiris são guardadas automaticamente para uma utilização futura após fechar a aplicação. Para restaurar as predefinições, clique no comando **Restaurar predefinições** no menu **Ficheiro**.

Quando digitalizar vários grupos de documentos que necessitem todos de definições diferentes, é útil guardar **ficheiros de definições** em separado para cada grupo.

Operação

- Seleccione as definições que pretende utilizar para um determinado grupo de documentos.
- No menu **Ficheiro**, clique no comando **Guardar definições**.
- Quando digitalizar ou abrir um documento do mesmo grupo posteriormente, clique no comando **Carregar definições** no menu **Ficheiro**.
- Seleccione o ficheiro de definições correcto e clique no botão **Abrir**.
- Clique em **Reconhecer + Guardar** para reconhecer o documento, utilizando as definições correctas.

Nota: o comando **Informações** no menu **Ficheiro** proporciona uma vista geral das definições mais importantes que seleccionou.

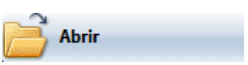
CAPÍTULO 12
RECONHECER DOCUMENTOS MULTI-PÁGINAS

ABRIR E RECONHECER MÚLTIPLOS FICHEIROS DE IMAGEM

O Readiris destina-se a processar múltiplos ficheiros de imagem de cada vez.

Para abrir múltiplos ficheiros de imagem:

- Clique em **Abrir** na barra de ferramentas principal.



- Seleccione os ficheiros de imagem que pretende abrir:
    - Seleccione o primeiro ficheiro de imagem e prima a tecla **Ctrl** à medida que selecciona imagens adicionais ou;
    - Seleccione um intervalo contínuo de ficheiros de imagem ao clicar na primeira imagem e ao premir a tecla **Shift** à medida que selecciona a última imagem.

Nota: quando abre um único ficheiro que consiste em várias páginas, por exemplo, um ficheiro TIFF ou documento PDF multi-páginas, é possível indicar o intervalo de páginas. Na área **Intervalo de páginas** da caixa de diálogo, seleccione **Páginas** e indique quais as páginas que pretende reconhecer.

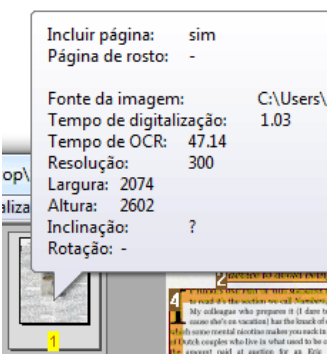
- Clique no botão **Abrir** para abrir a(s) imagem(s).

  Atente que também pode arrastar e largar ficheiros de imagem do Windows Explorer para a janela de imagem do Readiris para os abrir.

- A barra de ferramentas da página irá apresentar os ficheiros de imagem abertos.

  Conselho: mantenha o cursor do rato sobre as miniaturas da página para apresentar as informações das definições por página.

  A barra de ferramentas da página pode ser utilizada para editar os documentos multi-páginas. Para obter mais informações, consulte a secção **Editar documentos multi-páginas**.



- Determine as definições de reconhecimento e clique em **Reconhecer + Guardar** para executar o reconhecimento.

- Caso pretenda abrir ou digitalizar imagens adicionais para o documento actual, clique no botão **Digitalizar** ou **Abrir** na barra de ferramentas principal.

  Ser-lhe-á solicitado se pretende eliminar o documento actual ou não. Clique em **sim** para eliminar o documento actual e comece um novo ou clique em **não** para adicionar as digitalizações adicionais para o documento actual.

DIGITALIZAR E RECONHECER DOCUMENTOS MULTI-PÁGINAS

O Readiris destina-se a processar documentos que consistem em várias páginas.

> Os documentos de processo do Readiris Home e Readiris Pro até 50 páginas.
>
> O Readiris Corporate processa documentos com um número ilimitado de páginas.

Para digitalizar documentos multi-páginas no Readiris, pode utilizar a função **alimentador de documentos automático** quando utilizar um scanner de alimentação de papel ou utilize a função **digitalização de intervalos** quando utilizar um scanner de mesa.

Digitalizar documentos multi-páginas com um alimentador de documentos (scanner de alimentação de papel)

- Clique no botão **Scanner** na barra de ferramentas principal e seleccione a opção **ADF** (alimentador de documentos automático)
- Coloque as páginas no alimentador de documentos do scanner e clique em **Digitalizar** para começar a digitalizar.
- Clique em **Reconhecer + Guardar** para reconhecer os documentos.

Digitalizar documentos multi-páginas com digitalização de intervalos (scanner de mesa)

- Clique no botão **Scanner** na barra de ferramentas principal.
- Seleccione **Digitalizar outra página depois** e indique o intervalo de tempo utilizando os botões direccionais.

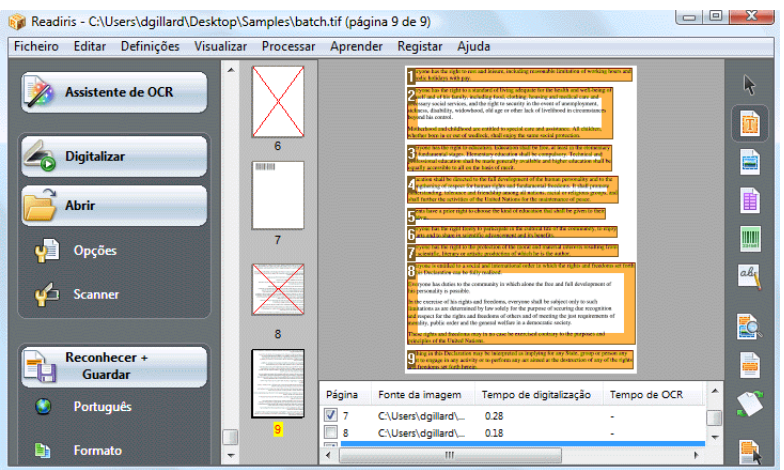
O scanner irá digitalizar automaticamente outra página depois do número indicado de segundos sem ter de clicar sempre no botão **Digitalizar**.

Clique em **Abortar** na caixa de diálogo da digitalização de intervalos para terminar a digitalização automática ou premir **ESC** no teclado.

Clique em **Pausa** na caixa de diálogo da digitalização de intervalos para pausar o intervalo de digitalização ou prima a barra de espaço no teclado. Clique em **Retomar**quando estiver pronto para continuar.

EDITAR DOCUMENTOS MULTI-PÁGINAS

Quando são abertos ou digitalizados múltiplos documentos no Readiris, a barra de ferramentas da página apresenta as miniaturas. As miniaturas na barra de ferramentas da página pode ser utilizada para editar os documentos multi-páginas.



Deslocar uma página dentro de um documento:

- Clique com o botão direito na página que pretende deslocar e clique em **Seleccionar página**.
- Arraste a página para a posição correcta.
- Ou clique com o botão direito na página e clique em **Deslocar página para cima** ou **para baixo**.

Eliminar uma página:

- Clique com o botão direito na página que pretende eliminar e clique em **Eliminar página**.
- Ou seleccione a página e prima a tecla **Delete** no seu teclado.

Excluir uma página para reconhecimento:

- Clique com o botão direito na página que pretende excluir e clique em **Excluir página**.
- Ou apague a caixa do número de página no painel de documento.

As páginas excluídas são rasuradas na barra de ferramentas da página.

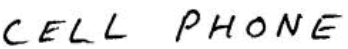
As páginas excluídas são ignoradas quando imprime as imagens digitalizadas e quando guarda as digitalizações para ficheiros de imagem multi-páginas.

Página	Fonte da imagem	Tempo de digitalização
☑ 8	C:\Users\dgillard...	0.20
☑ 9	C:\Users\dgillard...	0.21

Conselho: os comandos **Incluir todas as páginas** e **Excluir todas as páginas** no **menu Editar** aplicam-se a todas as páginas em simultâneo.

CAPÍTULO 13
RECONHECER TEXTO ESCRITO EM LETRA DE IMPRENSA

Juntamente com o texto dactilografado, tabelas, gráficos e códigos de barras, o Readiris reconhece texto **escrito em letra de imprensa**. Texto escrito em letra de imprensa consiste em letras maiúsculas separadas.



É necessário software ICR (Reconhecimento Inteligente de Caracteres) altamente especializado para reconhecer caracteres de texto escrito em letra de imprensa.

Para reconhecer texto escrito em letra de imprensa:

- Clique no botão de texto escrito em letra de imprensa na barra de ferramentas da imagem.

- Desenhe uma moldura em torno do texto escrito em letra de imprensa.
- Clique em **Reconhecer + Guardar** na barra de ferramentas principal.

O documento inteiro incluindo o texto escrito em letra de imprensa será reconhecido.

Nota: clique com o botão direito na zona de texto escrito em letra de imprensa e clique em **Copiar como dados** para reconhecer apenas a zona de texto escrito em letra de imprensa e enviá-la para a área de transferência.

Símbolos reconhecidos

O reconhecimento de texto escrito em letra de imprensa está limitado ao alfabeto latino e suporta numerais (0-9), letras maiúsculas (A-Z) e os sinais de pontuação vírgula, ponto, sinal de mais e hífen.

Os acentos, tremas e outros caracteres especiais não são suportados.

Notas

- O Readiris suporta texto escrito em letra de **imprensa**, não texto **manuscrito**.

  Para obter mais informações, consulte a secção **Regras do texto escrito em letra de imprensa**.

- Os caracteres com letras maiúsculas são substituídos por caracteres com letras minúsculas após o reconhecimento, a menos que surjam no início de uma frase.

- O idioma, tipo de letra e espaço entre caracteres característicos do documento, não se aplicam ao texto escrito em letra de imprensa.

- A aprendizagem interactiva também não se aplica. A tecnologia ICR baseia-se em mais de um milhão de amostras de escrita.

REGRAS DE TEXTO ESCRITO EM LETRA DE IMPRENSA

Devem ser consideradas várias regras para o Readiris reconhecer texto escrito em letra de imprensa:

- Escreva caracteres regulares, bem espaçados.

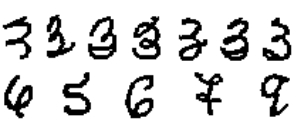
  Atente como são escritos os caracteres A, G e Q.

- Utilize uma caneta suficientemente espessa.

  As canetas pretas produzem melhores resultados que as canetas azuis.

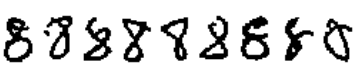
  Não utilize lápis.

- Não estilize em demasia.



  Os caracteres demasiadamente estilizados aumentam o risco de erros de OCR.

- Não abra círculos que deviam estar fechados, não feche círculos que deviam estar abertos.



- Evite caracteres separados.

- Evite retraçar

Retraçar reduz a qualidade e nitidez de imagem dos símbolos escritos em letra de imprensa.

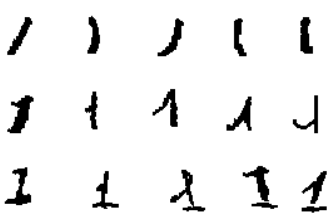
Os caracteres que estão inteiramente rasurados não serão reconhecidos.

- Escreva os uns correctamente.

O número 1 pode ser escrito no estilo "anglicano" e "europeu".

Os uns podem ser sublinhados ou não.

A barra de sublinhar horizontal não tem de tocar no resto da forma da letra.



Conselho: quando são obtidos resultados que não ideais, utilize o **formulário de escrita** da I.R.I.S. e adapte o seu estilo de escrita. O formulário de escrita da I.R.I.S. em branco serve como modelo de página completa para as letras maiúsculas que podem ser preenchidas correctamente e com o tamanho certo. O formulário pode ser encontrado no CD-ROM da Readiris e na pasta de instalação da Readiris.

CAPÍTULO 14
RECONHECER CÓDIGOS DE BARRAS

INTRODUZIR LEITURA DE CÓDIGOS DE BARRAS

Juntamente com o reconhecimento óptico de caracteres de128 idiomas, o Readiris oferece também leitura de códigos de barras.

São suportadas todas as simbologias gerais de códigos de barras: Codabar, Code 128, Code 39, Code 39 extended, Code 39 HIBC, Code 93, Datalogic 2 de 5, Discrete 2 de 5, EAN-13, EAN-8, Interleaved 2 de 5, MSI Pharmaceutical, MSI-Plessey, Kodak patch code, PDF-417, PostNet, UCC-128, UPC-A e UPC-E.

Atente que são necessários códigos de barras impressos a laser ou a jactos de tinta para que o Readiris execute o OCR. Os códigos de barras de impressão matricial não são suportados visto que não produzem contraste suficiente e a sua resolução está em grande parte limitada a 60 ppp.

Leitura manual de códigos de barras

- Determine quais os códigos de barras que pretende que o Readiris reconheça:

- No menu **Definições**, clique em **Códigos de barras**.
- Seleccione as simbologias que pretende que o Readiris reconheça.
- Determine se pretende que o Readiris verifique ou retire os dígitos de controlo.

- Clique no botão do código de barras na barra de ferramentas da imagem e desenhe uma moldura em torno das zonas dos códigos de barras no documento.

- Clique em **Reconhecer + Guardar** na barra de ferramentas principal.

O documento inteiro incluindo o conteúdo do código de barras será reconhecido.

Nota: clique com o botão direito na zona do código de barras e clique em **Copiar como dados** para copiar o seu conteúdo para a área de transferência.

ÍNDICE ALFABÉTICO

A

a retirar a inclinação..... 25

ADF 85

ajustar documentos digitalizados..... 33

alimentador de documentos automático 85

alterar a interface do utilizador 20

análise da página..... 25, 39

aprendizagem interactiva 54

assistência do produto..... 14

Assistente..... 15

B

barra de ferramentas da imagem 19

barra de ferramentas principal 18

C

câmara digital..... 31

características do documento 53

carregar definições 81

códigos de barras.....93

Compressão JPEG 2000 78

contraste 30, 36

converter documentos PDF ..73

cor de fundo.....63

cor de fundo das células das tabelas.....68

D

definições do scanner 28

desinstalar o Readiris.....13

dicionários de tipos de letra ..57

digitalização de intervalo.....85

disposição automática de janelas.....39

disposição manual de janelas 41

Documentos arábicos ...4, 8, 50

Documentos asiáticos ...4, 7, 50

documentos de texto.....61

Documentos hebraicos ..4, 8, 50

documentos multi-páginas .. 83, 85

documentos spreadsheet..... 65

E

Edição Asiática 4, 7, 50

Edição Médio-Oriente...4, 8, 50

editar documentos multi-páginas 86

enviar documentos por e-mail 60

espaço entre caracteres..... 54

Exportação Excel 59

Exportação HTML 59

Exportação OpenDocument . 59

Exportação PDF 59, 69

Exportação PDF iHQC 71

Exportação RTF 59

Exportação SpreadsheetML . 59

Exportação Unicode..... 59

Exportação Word 59

Exportação WordML 59

Exportação XML 59

Exportação XPS 59, 74

Exportação XPS iHQC 76

F

ficheiro de definições 81

ficheiros de esquema 46

folhas de cálculo 65

formatos de exportação.....59

formatos de imagem suportados 28

G

guardar definições 81

I

idioma.....50

idioma da interface do utilizador 20

idiomas mistos.....52

imagem a cores.....30, 33

imagem a preto e branco.....30

imagem em escala de cinzentos 30

imagens invertidas.....31

inclinação da linha.....32

instalação.....11

interface do utilizador.....17

L

luminosidade..... 30, 34

M

matriz de pontos..... 53

modelos de disposição de janelas 46

modo a cores 30

N

numérico 51

O

opções 25

opções de esquema..... 62

opções dos gráficos 77

Opções PDF 70

Opções XPS 75

P

páginas 86

deslocar 87

eliminar 87

excluir 87

seleccionar..... 87

painel do documento 19

precisão vs. velocidade 52

predefinições 81

propriedades dos documentos 59

R

recriar documentos originais 62

registo 13

remoção da inclinação da página 25

remoção de pontos.....36

requisitos do sistema 11

resolução 30

restaurar predefinições.....81

rotação 25

S

SmartTasks 18, 21

suavizar as imagens a cores .32, 33

T

tabelas.....65

tamanhos do papel 64

texto escrito em letra de imprensa 89, 91

tipo de letra.....53

V

velocidade vs. precisão.....52