Readiris™ Pro 12

Guia do usuário

Conteъdo

Direitos autorais 1

Capítulo 1 Apresentação do Readiris 3

Poupe tempo, não é preciso voltar a digitar 3

A série Readiris..... 6

Capítulo 2 Instalação do Readiris 9

Requisitos de sistema 9

Instalação do software..... 9

Desinstalação do software..... 11

Registro do software 11

Suporte ao produto..... 12

Capítulo 3 Introdução..... 13

Execução do Readiris..... 13

Usando o Assistente OCR..... 13

Interface do usuário 15

Alteração do idioma da interface do usuário..... 18

Capítulo 4 SmartTasks do Readiris..... 19

Capítulo 5 Digitalização de documentos..... 23

Seleção das opções..... 23

Abertura de arquivos de imagens..... 24

Digitalização de documentos impressos em papel 26

Capítulo 6 Ajuste de documentos digitalizados 31

Capítulo 7 Salvamento de documentos como arquivos de imagens 37

Capítulo 8 Janelas de documentos 39

Divisão automática de documentos em janelas 39

Divisão manual de documentos em janelas..... 41

Uso de modelos de janela..... 45

Capítulo 9 Reconhecimento de documentos..... 49

Introdução 49

Seleção do idioma do documento 50

Definição das características do documento 53

Uso do aprendizado interativo 55

Uso de dicionários de fontes 57

Capítulo 10 Formatação e salvamento de documentos 59

Formatação de documentos..... 59

Formatação de documentos de texto 61

Formatação de documentos baseados em tabela 65

Criação de documentos PDF..... 69

Seleção das opções PDF 70

Compactação iHQC de documentos PDF 71

Redirecionamento de documentos PDF 73

Criação de documentos XPS 74

Seleção das opções XPS 75

Compactação iHQC de documentos XPS 76

Seleção das opções de gráficos 77

Capítulo 11 Salvamento e carregamento de configurações... 81

Capítulo 12 Reconhecimento de documentos de várias páginas 83

Abertura e reconhecimento de vários arquivos de imagens 83

Digitalização e reconhecimento de documentos de várias páginas 85

Edição de documentos de várias páginas 86

Capítulo 13 Reconhecimento de texto manuscrito..... 89

Capítulo 14 Reconhecimento de códigos de barras 93

Hndice 95

Direitos autorais

ReadirisPro12-dgi-091102-02

Copyrights © 1987-2009 I.R.I.S. Todos os direitos reservados.

A I.R.I.S. detém os direitos autorais do software Readiris, do sistema de ajuda on-line e desta publicação.

As informações contidas neste documento são de propriedade da I.R.I.S. 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 contrato de licença que estabelece os termos de uso deste produto. O software pode ser usado ou copiado apenas de acordo com os termos desse contrato. Nenhuma parte desta publicação pode ser reproduzida, transmitida, armazenada em sistemas de recuperação ou traduzida para outro idioma sem o consentimento prévio, por escrito, da I.R.I.S.

Este guia do usuário utiliza nomes fictícios para fins de demonstração; referências a pessoas, empresas ou organizações reais são meras coincidências.

Marcas comerciais

Readiris e o logotipo Readiris são marcas comerciais da Image Recognition Integrated Systems S.A.
OCR, ICR e a tecnologia de código de barras da I.R.I.S.
Tecnologia AutoFormat e Lingüística da I.R.I.S.
Tecnologia de análise de campo e BCR da I.R.I.S.
Tecnologia de compactação iHQC da I.R.I.S.
Analisador de XML desenvolvido pela Apache. Este produto inclui software desenvolvido pela Apache Software Foundation.

Todos os demais produtos mencionados neste guia do usuário são marcas comerciais ou registradas de seus respectivos proprietários.

CAPÍTULO 1
APRESENTAÇÃO DO READIRIS

POUPE TEMPO, NÃO É PRECISO VOLTAR A DIGITAR

Parabéns por ter adquirido o Readiris. Este pacote irá, sem dúvida alguma, ser uma grande ajuda para a recuperação dos seus textos, tabelas, gráficos, códigos de barras e texto manuscrito.

Para usufruir de todo o rendimento dos computadores, primeiro você terá de digitar as suas informações. Se alguma vez você já teve de voltar a digitar um relatório de 15 páginas ou uma grande tabela com números, sabe o quanto isso pode ser demorado e enfadonho. Utilize este pacote de OCR (Optical Character Recognition) de última geração para automaticamente converter documentos impressos em papel ou arquivos de imagens digitalizadas em documentos de texto editáveis e pesquisáveis, que podem ser arquivados e compartilhados. Estão disponíveis dois modos de reconhecimento: um assegura velocidade máxima e o outro garante precisão ideal do OCR.

Digitalize um documento impresso ou digitado, indique as zonas que deseja reconhecer com o Readiris, ou faça com que o sistema as detecte para você, execute o reconhecimento de caracteres e exporte o documento para um processador de texto. Documentos com várias páginas são processados do começo ao fim de uma só vez. Alguns cliques no mouse irão lhe poupar longas horas de trabalho, uma vez que o Readiris converte os documentos impressos em papel em

arquivos eletrônicos editáveis: é até 40 vezes mais rápido do que a redigitação manual.

O assistente fornece orientação bastante prática durante as configurações necessárias para operar o Readiris, permitindo que você obtenha resultados de maneira rápida e fácil. Ou utilize as SmartTasks para agilizar ainda mais o processo. Você pode enviar os resultados da leitura diretamente para um processador de texto ou uma planilha, arquivá-los como arquivos PDF ou XPS etc. Para reconhecer faxes e converter documentos PDF, arraste os respectivos arquivos de imagem do Windows Explorer para a janela do aplicativo Readiris. Ou, envie uma imagem imediatamente para o Readiris através do menu de contexto.

O Readiris reconhece dados tabulares e os recria como planilhas em um software de planilha ou como objetos de tabela em um processador de texto; os dados numéricos ficam imediatamente prontos para serem processados novamente.

O Readiris se baseia nas mais avançadas tecnologias de reconhecimento. O reconhecimento de texto independente da fonte é complementado por técnicas de auto-aprendizado. O sistema aprende novos caracteres e palavras por meio de análise contextual e lingüística. Isso significa que a precisão do OCR do sistema de reconhecimento melhorará ao longo do tempo.

O Readiris oferece suporte a até 128 idiomas: existe suporte para todos os idiomas americanos e europeus, entre eles os idiomas centro-europeus, cirílicos e bálticos, bem como o grego e o turco. Como opção, o Readiris pode ler documentos em árabe, farsi e hebraico, além de quatro idiomas asiáticos: japonês, chinês simplificado, chinês tradicional e coreano. O Readiris lida até mesmo com alfabetos mistos: o software detecta as palavras “ocidentais” que aparecem em documentos gregos, cirílicos, hebraicos e asiáticos - muitos nomes próprios impossíveis de transcrever, nomes de marca etc. são escritos com os símbolos ocidentais.

O Readiris utiliza a lingüística *durante* a fase de reconhecimento, e não depois. Como resultado, o Readiris reconhece documentos de todos os tipos com grande precisão, incluindo documentos de má qualidade, faxes e impressões matriciais. Ele lida perfeitamente com documentos digitalizados e copiados de forma insatisfatória que tenham fontes demasiado claras ou demasiado escuras. Os caracteres unidos são resolvidos, enquanto formas fragmentadas, como os símbolos matriciais, são recompostas.

Além disso, o Readiris possui uma função de verificação do usuário (opcional). Quando ativada, essa função ("Aprendizado interativo") não apenas sinaliza os caracteres com relação aos quais o sistema de reconhecimento não tem certeza, mas também permite aumentar a precisão do sistema. Todas as soluções confirmadas são memorizadas, aumentando a velocidade e a confiança do sistema e tornando-o mais inteligente à medida que você avança. Essa poderosa ferramenta de aprendizado permite que você treine o Readiris em caracteres especiais, como símbolos matemáticos e dingbats, além de permitir que ele lide com fontes deformadas.

Para aumentar ainda mais a sua produtividade, o Readiris não apenas reconhece os seus textos, mas também pode formatá-los para você. Estão disponíveis vários níveis de formatação. Quando você utiliza a “autoformatação”, o Readiris recria uma cópia exata do documento digitalizado: a formatação da página, das palavras e dos parágrafos do documento original é mantida. São utilizados tipos de impressão semelhantes. Os tamanhos de ponto e os estilos de tipo, tal como foram utilizados no documento de origem, são mantidos ao longo do processo de reconhecimento. O posicionamento de colunas, blocos de texto e gráficos segue o dos documentos originais. O Readiris pode até mesmo incluir a fotografia de fundo de uma página digitalizada no documento reconhecido. E, como o Readiris pode fazer a digitalização em tons de cinza e em cores sem problemas, você pode recapturar qualquer gráfico, seja ele um desenho, uma fotografia em preto-e-branco ou uma ilustração colorida. Quando um documento inclui tabelas, o Readiris as

reorganiza em células reais e recria as bordas das células das tabelas originais.

Em outras palavras, o Readiris permite que você arquive uma cópia real dos seus documentos, sendo essa cópia constituída por arquivos de texto compactos e editáveis em vez de imagens digitalizadas.

Os códigos de barras que ocorrem em uma página digitalizada também podem ser lidos, e o mesmo se aplica a texto manuscrito, contanto que você escreva em "letras maiúsculas" bem espaçadas.

O Readiris é compatível com Twain e oferece suporte a uma ampla variedade de digitalizadores de mesa e de produção, dispositivos "all-in-one" ou "MFPs" ("Multifunctional Peripherals") e câmeras digitais. A digitalização com intervalo permite que você digitalize documentos de várias páginas de maneira eficiente quando o seu digitalizador não estiver equipado com um alimentador de documentos.

A SÉRIE READIRIS

A tabela abaixo apresenta uma visão geral das versões disponíveis:

Readiris Home 12

Recursos limitados

25 idiomas de reconhecimento

Oferece suporte para imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX

Gera arquivos PDF Imagem-Texto, DOCX, ODT, WordML, SpreadsheetM, RTF, HTM, XML, TXT, e TIFF

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

Readiris Pro 12 Asian	**Readiris Corporate 12 Asian**
Recursos básicos	Recursos básicos
128 idiomas de reconhecimento	128 idiomas de reconhecimento
Oferece suporte para imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX	Oferece suporte para imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX
Gera quatro tipos de arquivos PDF, PDF-iHQC (nível I), quatro tipos de arquivos XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF etc.	Gera quatro tipos de arquivos PDF, PDF-iHQC (nível I-III), PDF/A, quatro tipos de arquivos XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF etc.
Reconhecimento de texto em chinês	Reconhecimento de texto em chinês

tradicional e simplificado Reconhecimento de texto em japonês Reconhecimento de texto em coreano	tradicional e simplificado Reconhecimento de texto em japonês Reconhecimento de texto em coreano Reconhecimento de grandes volumes Processamento automatizado Indexação de documentos Reconhecimento de cartões de visita

Readiris Pro 12 Middle-East*	**Readiris Corporate 12 Middle-East***
Recursos básicos 128 idiomas de reconhecimento Oferece suporte para imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Gera quatro tipos de arquivos PDF, PDF-iHQC (nível I), quatro tipos de arquivos XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF etc. Reconhecimento de texto em árabe & farsi Reconhecimento de texto em hebraico	Recursos básicos 128 idiomas de reconhecimento Oferece suporte para imagens PDF, DCX, DJV, DJVU, JPG, JPEG, J2C, J2K, JP2, PNG, TIF, TIFF, BMP, PCX Gera quatro tipos de arquivos PDF, PDF-iHQC (nível I-III), PDF/A, quatro tipos de arquivos XPS, XPS-iHQC (nível I), DOCX, ODT, XLS, WordML, SpreadsheetML, RTF, HTM, XML, TXT, TIFF etc. Reconhecimento de texto em árabe & farsi Reconhecimento de texto em hebraico Reconhecimento de grandes volumes Processamento automatizado Indexação de documentos Reconhecimento de cartões de visita

(*) Nenhuma versão para Mac disponível.

CAPÍTULO 2
INSTALAÇÃO DO READIRIS

REQUISITOS DE SISTEMA

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

- Um PC com processador Intel 486 ou compatível. Recomenda-se um PC com processador Pentium.
- 256 MB de memória RAM.
- 120 MB de espaço livre no disco.

  (105 MB de espaço em disco serão suficientes se você não instalar os exemplos de arquivos.)

- Sistema operacional Windows Vista, Windows XP ou Windows 2000.

Alguns drivers de digitalizador podem não funcionar com as últimas versões do Windows. Consulte a documentação fornecida com o digitalizador para saber quais plataformas são compatíveis.

INSTALAÇÃO DO SOFTWARE

Como instalar o Readiris:

- Faça logon como administrador no Windows ou certifique-se de ter os direitos de administração necessários.

- Conecte o digitalizador ao PC e instale o software correspondente.
  Teste o digitalizador. Em caso de problemas, entre em contato com o fabricante do digitalizador.

- Insira o CD-ROM do Readiris na unidade de CD-ROM e siga as instruções na tela para instalar o software.

- Clique em **Readiris** para iniciar a instalação (produtos de software adicionais são oferecidos: Copernic Desktop Search Home Edition e Cardiris 4 LE).

- Selecione o idioma da instalação e clique em **OK**.

- Aceite os termos do contrato de licença.

- São oferecidas uma instalação **completa** e uma instalação **personalizada**. Selecione as opções necessárias e clique em **Próximo** sempre que estiver pronto para ir para a próxima tela. Todos os léxicos e exemplos de imagens serão instalados por padrão, bem como um guia do usuário eletrônico e uma ajuda on-line.

- Clique em **Concluir** para concluir a instalação.
  O programa de instalação cria o submenu **Aplicativos I.R.I.S. - Readiris** automaticamente no menu **Programas** do Windows.
  O programa de instalação também cria um atalho para o aplicativo Readiris na área de trabalho do Windows.

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

DESINSTALAÇÃO DO SOFTWARE

Existe apenas uma forma correta de desinstalar o Readiris: com o assistente de (des)instalação do Windows. É altamente recomendável *não* desinstalar o Readiris nem qualquer um dos seus módulo de software manualmente apagando os arquivos de programa.

Para desinstalar o Readiris:

- Feche o aplicativo.
- No menu **Iniciar**, clique em **Painel de Controle**.
- No ícone **Programas**, clique em **Desinstalar um programa**.
- Selecione **Readiris** na lista e clique no botão **Desinstalar**.
- Siga as instruções na tela.

REGISTRO DO SOFTWARE

Lembre-se de registrar a licença do Readiris. Ao fazer o registro, você:

- Será informado sobre futuros desenvolvimentos do produto e sobre produtos I.R.I.S. relacionados.
- Terá direito a suporte gratuito.
- Terá direito a ofertas especiais em produtos I.R.I.S.

Para fazer o registro:

Use o **Assistente de registro** no menu **Registrar**. Siga as instruções na tela.

SUPORTE AO PRODUTO

Depois de registrar o seu produto, você terá direito a suporte ao produto da I.R.I.S. para todas as funcionalidades básicas do software. Entre em contato com a I.R.I.S. em:

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

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

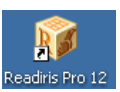
Ásia-Pacífico:
support@irislink.com
Telefone: +852 22646133

CAPÍTULO 3
INTRODUÇÃO

EXECUÇÃO DO READIRIS

Para executar o Readiris:

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



- Clique em qualquer lugar da tela de inicialização para iniciar o Readiris.

O Assistente OCR abrirá automaticamente.

USANDO O ASSISTENTE OCR

O Assistente OCR permite que você defina, rapidamente, todas as configurações necessárias para operar o Readiris.

Ao iniciar o Readiris, clique em qualquer lugar na tela de inicialização para iniciar o Assistente OCR.

Etapa 1

Selecione a **fonte da imagem**.
Você pode capturar imagens com o **digitalizador** ou abrir **arquivos de imagens**.
Selecione as opções de **rotação** e **alinhamento** que você deseja usar.

> Para obter mais informações, consulte a seção **Seleção das opções**.
>
> Para familiarizar-se com o Readiris, use os exemplos de imagens fornecidos com o software. Eles podem ser encontrados no CD-ROM do Readiris e na subpasta **Samples** da pasta de instalação do Readiris.

Clique em **Próximo** para ir para a etapa seguinte.

Etapa 2

Caso você tenha selecionado um digitalizador, clique no botão **Alterar** para selecionar as configurações do digitalizador.

> Para obter mais informações sobre as configurações do digitalizador, consulte a seção **Digitalização de documentos impressos em papel**.

Clique em **OK** para salvar as configurações.
Clique em **Próximo** para ir para a etapa seguinte.

Etapa 3

Clique no botão **Alterar** para alterar o idioma do documento. Por padrão, o idioma do documento está definido como Inglês americano.
Selecione o idioma necessário ou a combinação de idiomas e os idiomas secundários na lista e clique em **OK**.
Use o controle deslizante para definir as configurações necessárias de **Velocidade-Precisão**.

> Para obter mais informações, consulte a seção **Seleção do idioma do documento**.

Clique em **OK** para salvar as configurações.
Clique em **Próximo** para ir para a etapa seguinte.

Etapa 4

Clique no botão **Alterar** para alterar o formato de saída ou o aplicativo de destino. O aplicativo de destino padrão é o Microsoft Word.
Selecione o aplicativo ou o formato de saída necessário na lista **Enviar para** ou **Arquivo externo**.
Clique nas várias guias e selecione as opções desejadas. As opções que não estão disponíveis para o aplicativo ou o formato escolhido aparecem esmaecidas.

> Para obter mais informações, consulte o capítulo **Formatação e salvamento de documentos**.

Clique em **OK** para salvar as configurações.
Clique em **Próximo** para ir para a etapa seguinte.

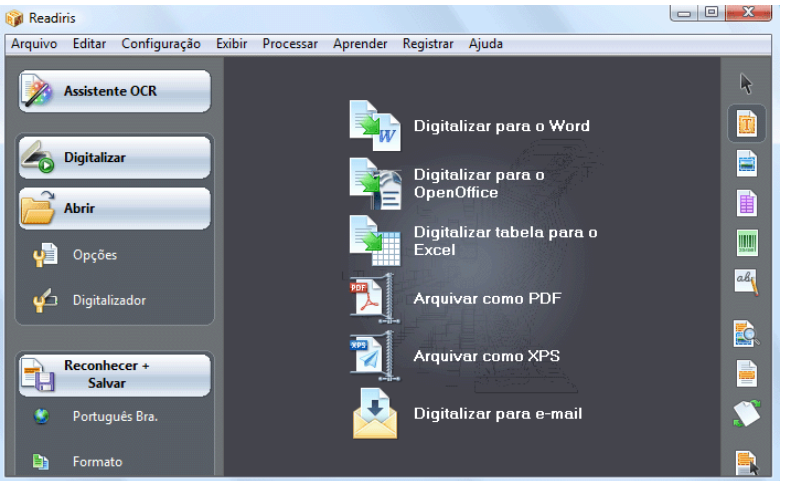
Etapa 5

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

INTERFACE DO USUÁRIO

Para explorar a interface do Readiris, clique em qualquer lugar na tela de inicialização do **Readiris** e, em seguida, clique em **Cancelar** quando o Assistente OCR for iniciado.

A interface vazia do Readiris será exibida.



A interface do Readiris é composta de:

- **SmartTasks** (no meio)

  SmartTasks são comandos predefinidos que permitem que você utilize as funções mais freqüentes do Readiris mediante o toque de um botão.

  Use as SmartTasks para digitalizar, reconhecer e enviar os documentos para o aplicativo de destino ou o formato de saída desejado.

  As SmartTasks aplicam configurações padrão, mas podem ser configuradas facilmente, clicando-se com o botão direito do mouse, para atender a necessidades mais específicas.

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

  Use os comandos e as opções da barra de ferramentas principal para digitalizar e reconhecer documentos manualmente.
  O **Assistente OCR** mostra a ordem em que você deve fazer isso.

- **Barra de ferramentas de imagem** (barra de ferramentas direita)

  Use os botões da barra de ferramentas de imagem para editar documentos na interface do Readiris.
  Aponte para os diferentes botões a fim de exibir suas dicas de ferramenta.

  Quando um documento for aberto ou digitalizado no Readiris, três zonas principais serão adicionadas à interface:

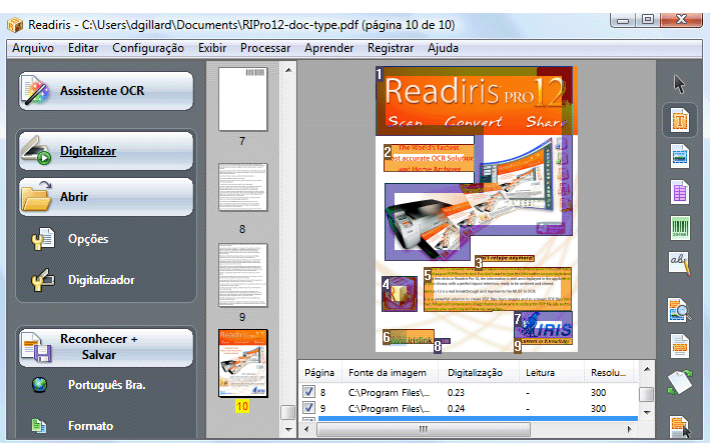
- **Barra de ferramentas de página** (à direita da barra de ferramentas principal)

  A barra de ferramentas de página exibe as miniaturas das páginas, que fornecem informações sobre as configurações quando você aponta para elas.

- **Janela de imagem** (no meio)

- **Painel de documentos** (na parte inferior)

  O painel de documentos exibe informações estatísticas sobre os documentos que estão abertos no Readiris, como a hora da digitalização e do reconhecimento, a resolução, a largura e a altura dos documentos etc.



ALTERAÇÃO DO IDIOMA DA INTERFACE DO USUÁRIO

A interface do usuário do Readiris está disponível em uma ampla variedade de idiomas.

Para alterar o idioma da interface do usuário:

- No menu **Configurações**, clique em **Idioma de interface do usuário**.
- Na lista **Idioma**, selecione o idioma necessário e clique em **OK** para confirmar.

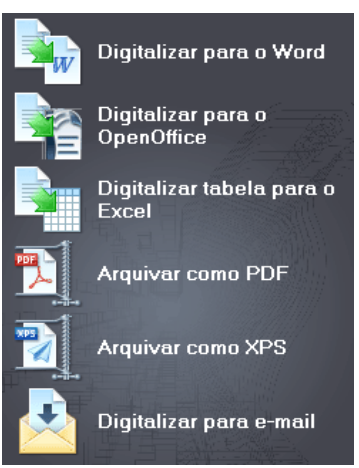
Nota: Se você selecionar um idioma incorreto, clique em **Ctrl+U**. A caixa de diálogo **Idioma** será aberta e você poderá selecionar outro idioma na lista.

CAPÍTULO 4
SMARTTASKS DO READIRIS

USO DAS SMARTTASKS DO READIRIS

Ao iniciar o Readiris, clique em qualquer lugar na tela de inicialização do Readiris e, em seguida, clique em **Cancelar** quando o Assistente OCR for iniciado. As SmartTasks do Readiris serão exibidas.

SmartTasks são comandos predefinidos que permitem que você utilize as funções mais freqüentes do Readiris mediante o toque de um botão.



Os vários botões SmartTask permitem que você:

1. Digitalize e reconheça documentos e os envie diretamente para o **Word** para processamento de texto.

   O Microsoft Word é o aplicativo de destino padrão. Consulte a seção **Formatação de documentos de texto** para saber mais sobre os outros aplicativos disponíveis.

2. Digitalize e reconheça documentos e os envie diretamente para o **OpenOffice** para processamento de texto.

   O OpenOffice.org Writer é o aplicativo de destino padrão. Consulte a seção **Formatação de documentos de texto** para saber mais sobre os outros aplicativos disponíveis.

3. Digitalize e reconheça tabelas e as envie diretamente para o **Excel** e outras planilhas.

   O Microsoft Excel é o aplicativo de destino padrão. Consulte a seção **Formatação de documentos baseados em tabela** para saber mais.

4. Digitalize e reconheça documentos e os **arquive** como arquivos **PDF**.

   Adobe Acrobat PDF Texto-Imagem é o formato de saída padrão. Consulte a seção **Criação de documentos PDF** para saber mais sobre os outros formatos disponíveis.

5. Digitalize e reconheça documentos e os **arquive** como arquivos **XPS**.

   XPS Imagem-Texto é o formato de saída padrão. Consulte a seção **Criação de documentos XPS** para saber mais sobre os outros formatos disponíveis.

6. Digitalize e reconheça documentos e os envie diretamente por **e-mail**

   Por padrão, os documentos serão enviados como PDF Imagem-Texto por meio do aplicativo de e-mail padrão. Consulte a seção **Formatação de documentos** para saber mais sobre os outros formatos disponíveis.

Ao utilizar o Readiris pela primeira vez, você precisará configurar as SmartTasks.

Para configurar as SmartTasks:

- Clique com o botão direito do mouse na SmartTask que você deseja usar.
- Selecione **Digitalizador** ou **Arquivos de imagens** como a fonte da imagem.
    - Quando **Digitalizador** for selecionado, o Readiris iniciará o digitalizador assim que você clicar na SmartTask. Os documentos digitalizados serão exibidos na interface, processados e salvos.

      O digitalizador precisa estar configurado corretamente para que a SmartTasks funcione.

      Para fazer isso:

      - Clique no botão **Digitalizador** na barra de ferramentas principal.
      - Clique em **Tipo** e selecione seu digitalizador na lista.

        Se o seu digitalizador não estiver na lista, selecione Outros modelos Twain.

      - Clique em **Configurar**, se aplicável, para selecionar a fonte Twain.
      - Em seguida, clique em **OK** para salvar as configurações.

Para obter mais informações sobre as configurações do digitalizador e sobre a digitalização de documentos impressos em papel, consulte a seção **Digitalização de documentos impressos em papel**.

- Quando você selecionar **Arquivos de imagem** e clicar na SmartTask, o Readiris abrirá a caixa de diálogo **Entrada**, na qual você poderá selecionar os arquivos de imagens que deseja processar.

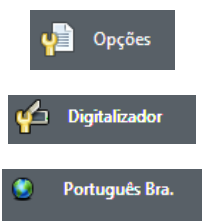
  Para obter mais informações sobre como abrir arquivos de imagens, consulte a seção **Abertura de arquivos de imagens**.

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

  As opções e os formatos de saída disponíveis dependem da SmartTask selecionada.

  Consulte a seção **Formatação de documentos** para saber mais sobre os formatos e as opções disponíveis.

- Quando terminar de configurar as SmartTasks, use os botões da barra de ferramentas principal para especificar as configurações de **idioma** e as **opções** de aperfeiçoamento de imagens e, se ainda for necessário, as **configurações do digitalizador**.



  Para obter mais informações sobre as supramencionadas configurações, consulte as seções **Seleção das opções**, **Digitalização de documentos impressos em papel** e **Seleção do idioma do documento**.

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

  O Readiris percorrerá todo o processo de reconhecimento automaticamente.

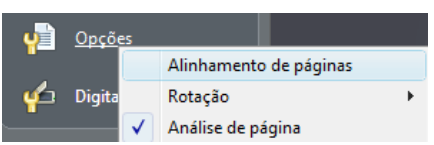
CAPÍTULO 5
DIGITALIZAÇÃO DE DOCUMENTOS

SELEÇÃO DAS OPÇÕES

Antes de digitalizar documentos impressos em papel ou abrir arquivos de imagens, você poderá selecionar várias opções de aperfeiçoamento de imagens. Quando ativadas, essas opções serão aplicadas durante a abertura e a digitalização de documentos.

Operação

- Clique no botão **Opções** na barra de ferramentas principal para selecionar várias opções de aperfeiçoamento de imagens.



  - Clique em **Alinhamento de páginas** para endireitar as páginas que foram digitalizadas em ângulo.
  - Aponte para **Rotação** e determine se você deseja que o Readiris gire as páginas **automaticamente** ou **90° para a esquerda**, **90° para a direita** ou **180°**.

    Essas duas opções desaceleram um pouco o processo de digitalização. Só as selecione quando for necessário.
  - A opção **Análise de página** é ativada por padrão.

Dessa maneira, as imagens digitalizadas ou abertas serão divididas em janelas automaticamente.

Você também pode usar as ferramentas de janela, na barra de ferramentas de imagem, para modificar os resultados da análise de página ou para dividir os documentos em janelas manualmente. Para obter mais informações, consulte o capítulo **Janelas de documentos**.

- Quando você terminar de definir todas as configurações (**Configurações do digitalizador**, **Opções**), clique no botão **Digitalizar** ou **Abrir** para digitalizar os documentos ou abrir os arquivos de imagens.

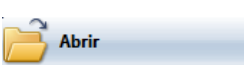
As opções supramencionadas também estão disponíveis no menu **Configurações**.

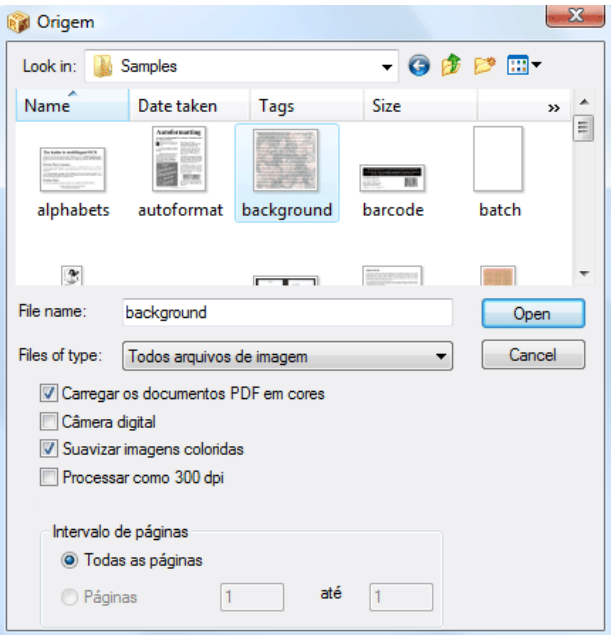
ABERTURA DE ARQUIVOS DE IMAGENS

Com o Readiris, você poderá processar documentos impressos em papel que digitalizar com um digitalizador ou processar arquivos de imagens já existentes de vários formatos.

Para abrir arquivos de imagens existentes:

- Clique no botão **Abrir** para procurar arquivos de imagens.





Dica: Você também pode **arrastar** arquivos de imagens até a janela de imagem do Readiris para abri-los.

Dica: Clique com o botão direito do mouse em qualquer arquivo de imagem que você deseja abrir, aponte para **Abrir com** e clique em **Aplicativo IOCR**. O software Readiris será aberto e exibirá a imagem.

Dica: Ao carregar documentos PDF e arquivos de imagens de várias páginas (imagens TIFF e fax DCX), você poderá definir o intervalo de páginas (caso, por exemplo, precise apenas de um determinado capítulo de um documento). Para fazer isso, clique em **Abrir** na barra de ferramentas principal. Na área **Intervalo de páginas**, selecione **Páginas** e insira quais páginas deseja carregar. Consulte também **Abertura e reconhecimento de vários arquivos de imagens**.

Dica: Para agilizar o processo de carregamento, clique no botão **Abrir** e desmarque **Carregar os documentos PDF em cores** ao processar documentos PDF.

- O Readiris aceita os seguintes formatos gráficos:

Formato gráfico	Extensão do arquivo
Adobe Acrobat PDF	*.pdf
Fax DCX	*.dcx
Imagens DjVu	*.djv, *.djvu
Imagens JPEG	*.jpg, *.jpeg
Imagens JPEG 2000	*.j2c, *.j2k, *.jp2
PNG (Portable Network Graphics)	*.png
Imagens TIFF (compactadas e de várias páginas) descompactadas e LZW, PackBits, Grupo 3, Grupo 4	*.tif, *.tiff
Bitmap do Windows	*.bmp
Imagens ZSoft Paintbrush	*.pcx

- Selecione o arquivo de imagem desejado e clique em **Abrir**.

Nota: As opções da caixa de diálogo **Entrada** também se aplicam à digitalização de documentos e são analisadas na seção **Digitalização de documentos impressos em papel**.

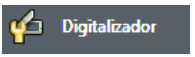
DIGITALIZAÇÃO DE DOCUMENTOS IMPRESSOS EM PAPEL

Com o Readiris, você poderá processar documentos impressos em papel que digitalizar com um digitalizador ou processar arquivos de imagens já existentes de vários formatos.

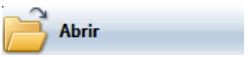
Ao processar documentos impressos em papel, assim que você clicar no botão **Digitalizar**, o Readiris iniciará o digitalizador e exibirá o documento digitalizado na interface.

Para digitalizar documentos:

- Clique no botão **Digitalizador** para definir as configurações do digitalizador.



Várias das opções da caixa de diálogo **Digitalizador** também estão disponíveis na caixa de diálogo **Abrir**.



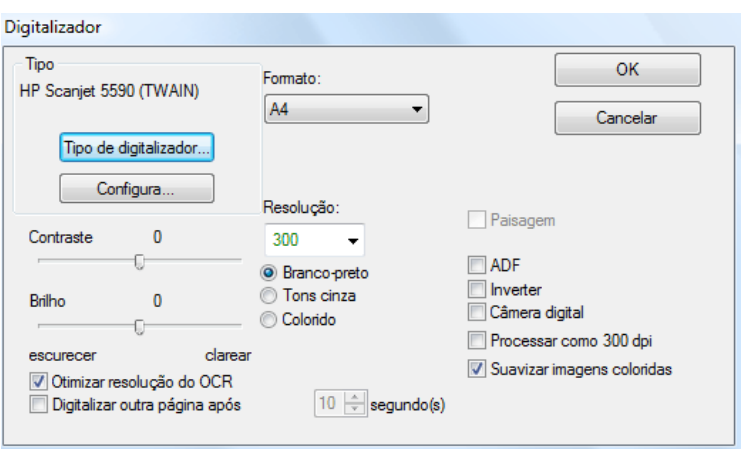
- Selecione o modelo correto de digitalizador.

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

- O Readiris é compatível com quase todos os digitalizadores de mesa e de produção, dispositivos "all-in-one" ("MFP" - "Multifunctional Peripheral"), câmeras digitais e padrões de digitalizador disponíveis no momento. O Readiris é compatível com o padrão Twain.

  O botão **Configurar** só estará disponível quando aplicável. Clique nele para selecionar a fonte Twain.

- Selecione as configurações do digitalizador.



Formato e resolução

O Readiris aceita uma ampla variedade de resoluções e formatos de papel.

É recomendável usar uma resolução de digitalização de 300 dpi.

Use uma resolução de 400 dpi ao reconhecer , texto asiático ou letras miúdas.

Modo de cor

O Readiris pode digitalizar documentos e abrir arquivos de imagens em **cores**, **preto-e-branco** e **tons de cinza**.

Contraste - Brilho

Use o controle deslizante para determinar as configurações adequadas de brilho e de contraste, de forma a obter o resultado ideal da digitalização.

Otimizar a resolução do OCR

Selecione **Otimizar resolução do OCR** para corrigir a resolução das imagens digitalizadas com muitos detalhes (acima de 600 dpi). O Readiris reduzirá a resolução.

Essa opção nunca aumenta a resolução das imagens digitalizadas com poucos detalhes.

Digitalizar documentos de várias páginas

Ao digitalizar documentos de várias páginas usando um digitalizador equipado com um alimentador de documentos, selecione a opção **ADF** (Automatic Document Feeder, alimentador automático de documentos). Coloque as páginas que você deseja digitalizar no alimentador e inicie a digitalização.

Ou use a **digitalização com intervalo** quando utilizar um digitalizador de mesa**:** selecione a opção **Digitalizar outra**

página após e indique depois de quantos segundos o Readiris deverá digitalizar outra página.

Para obter mais informações, consulte a seção **Digitalização e reconhecimento de documentos de várias páginas**.

Importante: Quaisquer opções que não estiverem disponíveis para o digitalizador selecionado aparecerão esmaecidas.

Digitalizar imagens na orientação paisagem

Selecione a opção **Paisagem** ao digitalizar imagens na orientação paisagem.

Exposição automática

Com alguns digitalizadores, a opção **Exposição automática** é selecionada por padrão. Essa opção ajusta as configurações de contraste e de brilho automaticamente.

Digitalizar imagens invertidas

Selecione a opção **Inverter** ao digitalizar imagens em preto-e-branco invertidas.

Usar uma câmera digital

Selecione **Câmera digital** quando você estiver usando uma câmera como fonte da digitalização. O Readiris utiliza rotinas especiais de reconhecimento para processar imagens de câmera digital.

Dicas para usar uma câmera digital como fonte da digitalização:

- Calibre a câmera fotografando um documento branco.
- Sempre selecione a maior resolução de imagem.
- Ative o modo macro da câmera para tirar closes.
- Use apenas o zoom óptico, não use o zoom digital.

- Segure a câmera diretamente acima do documento. Evite fotografar o documento em ângulo.
- Produza imagens estáveis. Use um tripé se necessário.
- Desative o flash ao capturar papel de qualidade fotográfica.
- Evite abrir imagens de câmera compactadas.
- Adapte as configurações de contraste e de brilho do Readiris ao ambiente (luz do dia, luz de lâmpada, luz neon).
- Selecione cor ou tons de cinza como o modo de cor.

Processar como 300 dpi

Selecione **Processar como 300 dpi** ao processar imagens com uma resolução incorreta ou desconhecida. As imagens serão processadas como se tivessem uma resolução de 300 dpi.

A resolução de imagens de câmera digital é quase sempre desconhecida.

Suavizar imagens coloridas

A opção Suavizar imagens coloridas é selecionada por padrão, uma vez que a suavização de imagens é necessária para que alguns digitalizadores reconheçam imagens coloridas e em tons de cinza com êxito.

- Quando você terminar de definir todas as configurações (**Configurações do digitalizador**, **Opções**), clique em **Digitalizar** para digitalizar os documentos.

Nota: Preste atenção à inclinação. Uma inclinação superior a 0,5° aumenta o risco de erros do reconhecimento.

CAPÍTULO 6
AJUSTE
DE DOCUMENTOS DIGITALIZADOS

Ao abrir ou digitalizar imagens coloridas ou em tons de cinza extremamente claras ou extremamente escuras, talvez seja necessário ajustá-las antes de executar o reconhecimento, para obter resultados satisfatórios do OCR.

Para ajustar imagens:

- Abra ou digitalize um documento colorido ou em tons de cinza.

  Verifique se as **configurações do digitalizador** estão corretas.

  Importante: As **configurações do digitalizador** e as **configurações do ajuste** parecem ser idênticas, mas observe que os dois conjuntos de configurações são aplicados em estágios diferentes do processo de digitalização/reconhecimento.

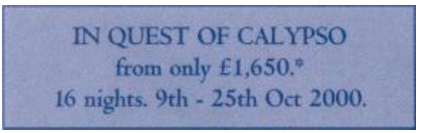
- No menu **Processar**, clique em **Ajustar imagem**.

  O Readiris utiliza rotinas inteligentes de **binarização** para converter imagens coloridas e em tons cinza em imagens em preto-e-branco, que são usadas para o reconhecimento.

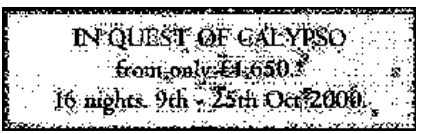
- Selecione **Suavizar imagem colorida** para nivelar a imagem.

  Essa opção torna imagens coloridas e em tons cinza mais homogêneas ao atenuar as diferenças de intensidade. Como resultado, é criado um contraste mais intenso entre o primeiro plano (texto) e o fundo (arte).

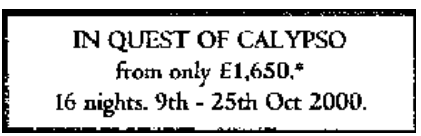
  Nota: Às vezes, a suavização é a única maneira de separar o texto de um fundo colorido.



(Imagem original)



(Imagem em preto-e-branco binarizada)



(Imagem suavizada)

- Use o controle deslizante para aumentar ou reduzir o **Brilho**.

  As configurações de **Brilho** determinam o brilho geral da imagem. Use essas configurações para escurecer ou clarear a imagem quando o texto estiver ilegível.

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

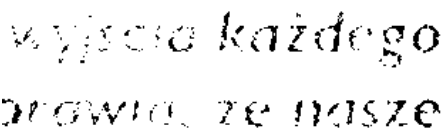
(Imagem colorida)

(Imagem binarizada. As configurações de binarização padrão produzem uma imagem preta.)

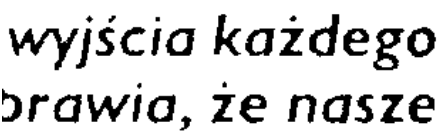
(A imagem clareada produz resultados satisfatórios do reconhecimento.)

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

(Imagem colorida)



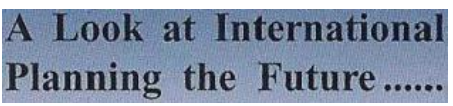
(Imagem binarizada. As configurações de brilho padrão produzem caracteres fragmentados.)



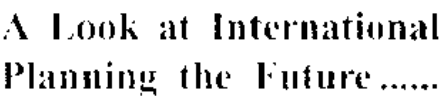
(A imagem escurecida produz resultados satisfatórios do reconhecimento.)

- Use o controle deslizante para aumentar ou reduzir o **Contraste**.

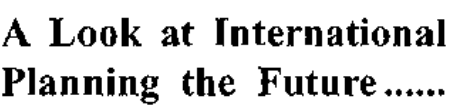
  As configurações de **Contraste** determinam o contraste entre as zonas mais escuras e mais claras de uma imagem. Use essas configurações para realçar as formas dos caracteres em um fundo colorido.



(Imagem colorida)



(As configurações de contraste padrão produzem caracteres interrompidos.)



(As configurações de maior contraste produzem resultados satisfatórios do reconhecimento.)

- Use o controle deslizante para aumentar ou reduzir as opções **Limpar**.

  A limpeza remove os pontinhos de imagens em preto-e-branco.

- Clique em **Aplicar** para visualizar os resultados.
- Se os resultados estiverem satisfatórios, clique em **OK**. Caso contrário, altere as configurações novamente.

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

CAPÍTULO 7
SALVAMENTO DE DOCUMENTOS COMO ARQUIVOS DE IMAGENS

Não é preciso reconhecer imediatamente os documentos impressos em papel que você digitalizar. Eles podem ser salvos como arquivos de imagens.

Para fazer isso:

- Digitalize o documento.
- No menu **Arquivo**, clique no comando **Salvar toda a página como imagem** ou **Salvar todas as páginas como imagem**.
- Mais tarde, abra o arquivo de imagem salvo e execute o reconhecimento.

Salvar somente gráficos

Você também pode optar por salvar as janelas de gráficos sem o texto do documento.

Para fazer isso:

- Digitalize ou abra o documento.
- No menu **Arquivo**, clique em **Salvar gráficos**.
- Todos os gráficos do documento serão salvos em um único arquivo.

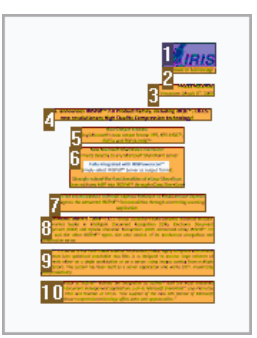
CAPÍTULO 8
JANELAS DE DOCUMENTOS

DIVISÃO AUTOMÁTICA DE DOCUMENTOS EM JANELAS

Ao digitalizar ou abrir documentos, o Readiris aplicará a **Análise de página** automaticamente para dividir os documentos em diversas janelas.

A opção **Análise de página** é selecionada por padrão. Clique no botão **Opções** e desative **Análise de página** caso queira evitar a análise automática de página.

Os resultados da análise de página poderão ser modificados manualmente depois da análise automática de página. Para obter mais informações, consulte a seção **Divisão manual de documentos em janelas**.



Tipos de janela

O Readiris utiliza cinco tipos de janela: **blocos de texto**, **zonas gráficas**, **tabelas**, **zonas de código de barras** e **zonas manuscritas**.

A análise de página detecta zonas gráficas, de texto e de tabela automaticamente. As zonas de código de barras e as zonas manuscritas precisam ser desenhadas manualmente.

> Para obter mais informações, consulte a seção **Divisão manual de documentos em janelas**.

Cada tipo de janela tem seu próprio código de cor: as janelas de texto são laranjas, os gráficos são roxos e as janelas de tabela são rosas. As zonas de código de barras são verdes e as zonas manuscritas são azuis.

As janelas são ordenadas de cima para baixo, da esquerda para a direita. Números indicam a ordenação das janelas. No entanto, a ordenação e os tipos de janela podem ser alterados. Para obter mais informações, consulte a seção **Divisão manual de documentos em janelas**.

Ignorar zonas de texto nas bordas da página

Quando o digitalizador gera bordas pretas ao redor da imagem real, a análise de página costuma localizar zonas onde há apenas ruído.

Para evitar isso, clique em **Não detectar janelas nas bordas** no menu **Configurações** e digitalize o documento novamente.

Ignorar a zona exterior

Como uma alternativa à divisão automática de documentos em janelas, pode-se utilizar a função **Ignorar a zona exterior**. Essa função é útil quando é preciso reconhecer somente uma área específica nas páginas do documento.

Selecione **Ignorar a zona exterior** clicando no botão correspondente na barra de ferramentas de imagem.

Desenhe um quadro ao redor da parte da página que o Readiris deverá reconhecer.

> A parte da página selecionada será analisada automaticamente.
>
> Você será solicitado a informar se deseja excluir a mesma zona externa da análise de página de cada página do documento.

DIVISÃO MANUAL DE DOCUMENTOS EM JANELAS

Além da divisão automática de documentos em janelas por meio da análise de página, o Readiris permite dividir os documentos em janelas manualmente.

A divisão manual de documentos em janelas é útil quando é preciso modificar os resultados da análise automática de página. Ela também é útil durante a criação de modelos de janela.

> Para obter mais informações sobre modelos de janela, consulte a seção **Uso de modelos de janela**.
>
> As zonas de texto manuscrito e de código de barras sempre precisam ser divididas em janelas manualmente.

Operação

- Para dividir um documento em janelas manualmente, primeiro clique no botão **Opções** e desmarque **Análise de página**.
- Abra ou digitalize o documento clicando no botão **Digitalizar** ou **Abrir**.

- Selecione o **tipo de janela** correto, clicando no botão de tipo de janela correspondente na barra de ferramentas de imagem.

  O Readiris utiliza cinco tipos de janela: **blocos de texto**, **zonas gráficas**, **tabelas**, **zonas de código de barras** e **zonas manuscritas**.

- Desenhe um quadro ao redor dos blocos de texto, dos gráficos, das tabelas e das zonas manuscritas que você deseja dividir em janelas.

  Para obter mais informações sobre o reconhecimento de códigos de barras e manuscritos, consulte as seções **Reconhecimento de códigos de barras** e **Reconhecimento de texto manuscrito**, respectivamente.

- Quando terminar de dividir o documento em janelas, clique no botão **Reconhecer + Salvar** para executar o reconhecimento.

Ordenar janelas

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

Quando terminar, clique no botão **Reconhecer + Salvar** para executar o reconhecimento.

> As janelas nas quais você não clicar serão excluídas do reconhecimento.

Desenhar polígonos

A divisão manual de documentos em janelas não se limita a formas retangulares. Você pode criar janelas poligonais juntando as retangulares. Sempre que duas janelas do mesmo tipo formarem uma interseção, elas se tornarão um polígono automaticamente.

Análise automática de página

Caso a página atual seja muito complexa para a divisão manual em janelas, clique no botão **Analisar página** na barra de ferramentas de imagem para dividir a página em janelas automaticamente.

> As zonas manuscritas e de código de barras sempre precisam ser desenhadas manualmente.

Alterar o tipo de janela

Para alterar o tipo de janela de uma janela, clique com o botão direito do mouse na janela, aponte para **Janela**, em seguida, para **Tipo** e clique no tipo de janela necessário.

Você também pode alterar o tipo de janela de várias janelas simultaneamente:

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

- Mantenha pressionada a tecla **Shift** enquanto seleciona várias janelas.
- Clique com o botão direito do mouse em qualquer uma das janelas selecionadas, aponte para **Janela**, em seguida, para **Tipo** e clique no tipo de janela necessário.

Modificar o tamanho da janela

Clique na janela que você deseja modificar.

Coloque o ponteiro do mouse sobre um marcador (nas laterais e nos cantos da janela).

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

Mover janelas

Selecione a janela que você deseja mover.

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

Reconhecer uma janela específica

Clique com o botão direito do mouse na janela que você deseja reconhecer e selecione **Copiar como texto**.

> Os resultados são enviados para a área de transferência como texto de corpo. Isso também funciona com texto manuscrito.

Também é possível copiar janelas de gráficos e de códigos de barras para a área de transferência.

Excluir janelas

Selecione as janelas que você deseja excluir ou clique no comando **Selecionar tudo**, no menu **Editar**, para selecionar todas as janelas.

Selecione os comandos **Recortar** ou **Apagar**, no menu **Editar**, para recortar ou excluir as janelas.

ou

Clique com o botão direito do mouse nas janelas selecionas, aponte para **Janela** e clique em **Excluir**.

Excluir janelas pequenas

Alguns documentos, faxes por exemplo, normalmente têm pontos "desgarrados" nas páginas, o que faz com que o Readiris crie janelas supérfluas que não contêm texto.

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

Essa opção apaga todas as janelas com menos de 1,3 cm e reordena as demais zonas.

USO DE MODELOS DE JANELA

Durante o reconhecimento de muitos documentos com um gabarito de página semelhante, talvez seja útil utilizar **modelos de janela** em vez da análise automática de página. Dessa maneira, a mesma estrutura de janela será aplicada a todos os documentos digitalizados ou abertos, o que agiliza o processo.

Operação

- Divida a primeira página do documento em janelas manualmente, usando os botões da barra de ferramentas de imagem.

  Para obter mais informações, consulte a seção **Divisão manual de documentos em janelas**.

- No menu **Arquivo**, clique no comando **Salvar gabarito**.
- Abra ou digitalize as outras páginas do documento clicando no botão **Abrir** ou **Digitalizar** na barra de ferramentas principal.
- No menu **Arquivo**, clique no comando **Carregar gabarito**.
- Selecione o arquivo de gabarito salvo.
- Para aplicar o gabarito a **todas** as páginas abertas ou digitalizadas, selecione **Aplicar gabarito em todas as páginas** na caixa de diálogo **Arquivo de gabarito**.
- Clique em **Abrir** para carregar o arquivo de gabarito.

  Quando você adicionar um documento ao Readiris, o arquivo de gabarito precisará ser carregado novamente, uma vez que a análise de página é ativada por padrão.

Ignorar a zona exterior

Como uma alternativa aos modelos de janela, você pode utilizar a opção **Ignorar a zona exterior.** Dessa maneira, você poderá definir uma área específica da página que precisa ser reconhecida. Qualquer dado fora da área OCR será excluído do reconhecimento.

Operação

- Selecione **Ignorar a zona exterior** clicando no botão correspondente na barra de ferramentas de imagem.

- Desenhe um quadro ao redor da parte da página que o Readiris deverá reconhecer.

  A parte da página selecionada será analisada automaticamente.

  Você será solicitado a informar se deseja ignorar a mesma zona exterior em todas as páginas do documento.

  Para cancelar essa função, execute a **Análise de página** novamente, clicando no botão **Analisar página** na barra de ferramentas de imagem.

- Clique em **Reconhecer + Salvar** para executar o reconhecimento.

CAPÍTULO 9
RECONHECIMENTO DE DOCUMENTOS

INTRODUÇÃO

Para reconhecer documentos, o Readiris aplica lingüística durante a fase de reconhecimento. Como resultado, o Readiris reconhece texto, tabelas, gráficos, códigos de barras e texto manuscrito em todos os tipos de documentos. O Readiris lida até mesmo com documentos complexos em forma de colunas, documentos de má qualidade, faxes, impressões matriciais, documentos digitalizados e copiados de forma insatisfatória que contenham fontes demasiado claras ou demasiado escuras etc.

O Readiris oferece suporte a 128 **idiomas**: existe suporte para todos os idiomas americanos e europeus, entre eles os idiomas centro-europeus, cirílicos e bálticos, bem como o grego e o turco. Como opção, o Readiris pode ler documentos em árabe, farsi e hebraico, além de quatro idiomas asiáticos: japonês, chinês simplificado, chinês tradicional e coreano. O Readiris lida até mesmo com alfabetos mistos: o software detecta as palavras "ocidentais" que aparecem em documentos gregos, cirílicos, hebraicos e asiáticos - muitos nomes próprios impossíveis de transcrever, nomes de marca etc. são escritos com os símbolos ocidentais.

O Readiris se baseia nas mais avançadas tecnologias de reconhecimento. O reconhecimento de texto independente da fonte é complementado por técnicas de auto-aprendizado. O sistema aprende novos caracteres e palavras por meio de análise contextual

e lingüística. Isso significa que a precisão do OCR do sistema de reconhecimento melhorará ao longo do tempo.

Além disso, o Readiris possui uma função de verificação do usuário (opcional). Quando ativada, essa função (**Aprendizado interativo**) não apenas sinaliza os caracteres com relação aos quais o sistema de reconhecimento não tem certeza, mas também permite aumentar a precisão do sistema. Todas as soluções confirmadas são memorizadas temporariamente durante o reconhecimento, aumentando a velocidade e a confiança do sistema e tornando-o mais inteligente à medida que você avança. Essa poderosa ferramenta de aprendizado permite que você treine o Readiris em caracteres especiais, como símbolos matemáticos e dingbats, além de permitir que ele lide com fontes deformadas.

Os resultados do aprendizado interativo também podem ser armazenados permanentemente em **dicionários de fontes** para uso futuro.

Outra maneira de aumentar a precisão do reconhecimento é utilizar **léxicos do usuário**. Você pode criar léxicos do usuário personalizados, com terminologia específica que deseja que o Readiris reconheça.

SELEÇÃO DO IDIOMA DO DOCUMENTO

O Readiris oferece OCR em 128 idiomas. O programa aceita todos os idiomas americanos e europeus, entre eles os idiomas centro-europeus, cirílicos e bálticos, bem como o grego e o turco.

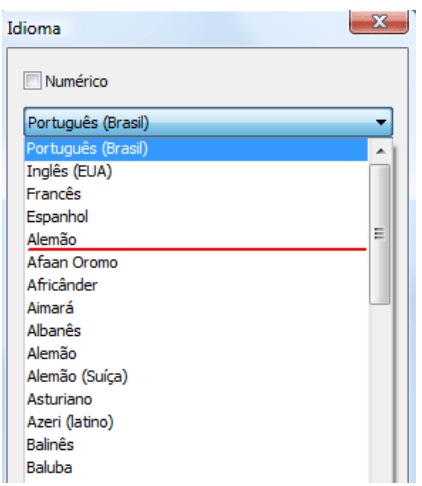
O Readiris Pro Asian o Readiris Corporate Asian também reconhecem documentos em japonês, chinês simplificado, chinês tradicional e coreano.

O Readiris Pro Middle-East e o Readiris Corporate Middle-East também reconhecem documentos em árabe, farsi e hebraico.

Para que o Readiris reconheça um documento, o idioma do documento deverá ser especificado.

Para fazer isso:

Clique no botão do **globo** na barra de ferramentas principal e selecione o idioma desejado na lista suspensa.



Os cinco idiomas selecionados mais recentemente são movidos para o início da lista de idiomas.

Importante: Quando você estiver lidando com documentos asiáticos e hebraicos, selecione o idioma do documento *antes* de executar a análise de página. São utilizadas rotinas específicas de análise de página para esses documentos.

O reconhecimento também pode se limitar a um conjunto de caracteres **numéricos** para o reconhecimento ideal de tabelas e figuras. O Readiris só reconhecerá os números 0-9 e as seguintes séries de símbolos:

+ sinal de mais	* asterisco	/ barra	% sinal de porcentagem	, vírgula
. ponto	(parêntese de abertura	) parêntese de fechamento	- hífen	= sinal de igualdade
$ símbolo de dólar	£ símbolo de libra	€ símbolo de euro	¥ símbolo de iene	

Para ativar o modo numérico, selecione **Numérico** na caixa de diálogo **Idioma**.

Reconhecer documentos com idiomas misturados

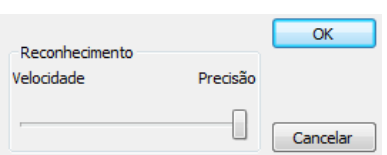
O Readiris também permite ativar conjuntos de caracteres mistos. Dessa maneira, o Readiris alterna os idiomas no meio de uma sentença automaticamente e reconhece palavras em inglês (nomes próprios etc.) que ocorrem em idiomas "exóticos".

Clique no botão do **globo** na barra de ferramentas principal e selecione a combinação de idiomas necessária na lista suspensa de idiomas.

Nota: Quando documentos asiáticos ou hebraicos são processados, conjuntos de caracteres mistos são usados automaticamente.

Velocidade - Precisão

Selecione o equilíbrio certo entre **velocidade do reconhecimento** e **precisão do reconhecimento**.



Esse equilíbrio está disponível para os alfabetos latino, cirílico e grego.

Dica: Quando a imagem for de má qualidade, dê preferência à precisão em detrimento da velocidade.

DEFINIÇÃO DAS CARACTERÍSTICAS DO DOCUMENTO

Além do idioma do documento, outras características do documento, como o **Tipo de fonte** e o **Espaçamento de caracteres** desempenham um papel importante no processo de reconhecimento.

Tipo de fonte

O Readiris faz distinção entre documentos "regulares" e documentos impressos em impressora matricial. Os símbolos matriciais (do tipo 9 pinos) são compostos de pontos separados isolados.

Para reconhecer documentos matriciais, são necessárias técnicas especiais de segmentação e de reconhecimento, que precisam ser ativadas.

Far out in the uncharted back

Para selecionar o tipo de fonte:

- No menu **Configurações**, aponte para **Tipo de fonte**.
- O tipo de fonte é definido, por padrão, como **Automático**.

Dessa maneira, o Readiris reconhece a impressão matricial de "25 pinos" ou "NLQ" (Near Letter Quality) ou outra impressão "normal".

- Para reconhecer somente documentos impressos em impressora matricial, clique em **Matricial**.

  O Readiris reconhecerá os documentos impressos em impressora matricial de "9 pinos" ou "rascunho".

Espaçamento de caracteres

O espaçamento de caracteres é o número de caracteres por polegada em um tipo de letra. O espaçamento de caracteres pode ser **fixo** (nesse caso, todos os caracteres têm a mesma largura) ou **proporcional** (nesse caso, os caracteres têm larguras diferentes).

Para selecionar o espaçamento de caracteres:

- No menu **Configurações**, aponte para **Espaçamento de caracteres**.
- O espaçamento de caracteres é definido, por padrão, como **Automático**.
- Clique em **Fixo** se todos os caracteres do tipo de letra tiverem a mesma largura. Isso normalmente ocorre em documentos antigos datilografados em máquina de escrever.
- Clique em **Proporcional** se os caracteres do tipo de letra tiverem larguras diferentes. Praticamente todas as fontes encontradas em jornais, revistas e livros são proporcionais.

Importante: Estas características de documento não se aplicam a documentos asiáticos, hebraicos e árabes.

USO DO APRENDIZADO INTERATIVO

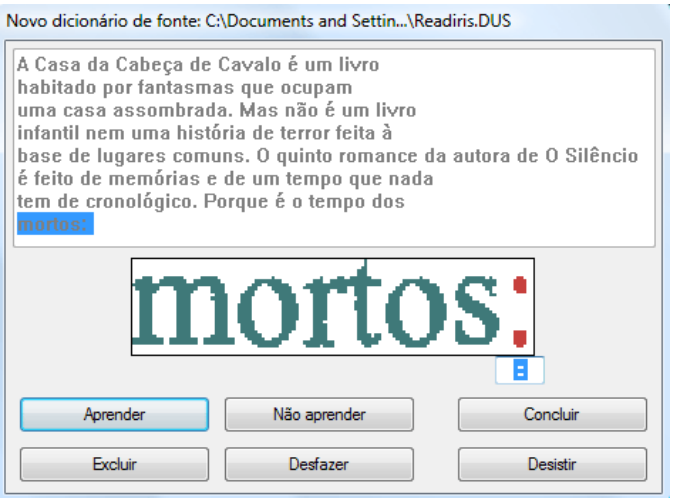
O Readiris oferece uma função de aprendizado interativo. Por meio do **Aprendizado interativo**, você pode treinar o sistema de reconhecimento quanto a fontes e formas de caracteres, bem como corrigir os resultados do reconhecimento, se necessário. Durante o aprendizado interativo, se o sistema de reconhecimento não tiver certeza quanto a qualquer caractere, este será exibido em uma janela de visualização, em combinação com sua palavra original e a solução proposta.

O aprendizado interativo pode melhorar, substancialmente, a precisão do sistema de reconhecimento, sendo particularmente útil durante o reconhecimento de formas distorcidas e alteradas. Ele também pode ser usado para treinar o Readiris quanto a símbolos especiais que ele não consegue reconhecer inicialmente, como símbolos matemáticos e científicos e dingbats.

Para ativar o aprendizado interativo:

- No menu **Aprender**, clique em **Aprendizado interativo**.
- Defina as configurações necessárias e clique no botão **Reconhecer + Salvar** para reconhecer o documento.
- No final do reconhecimento, o Readiris entra na fase de aprendizado interativo.

  Serão exibidos os caracteres com relação aos quais o sistema de reconhecimento não tiver certeza.



Se os resultados estiverem corretos:

- Clique no botão **Aprender** para salvar os resultados como seguros.

  Os resultados do aprendizado são armazenados temporariamente na memória do computador, pela duração do reconhecimento. O Readiris não exibirá mais os caracteres aprendidos ao reconhecer o resto do documento.

  Quando um novo documento for reconhecido, os resultados do aprendizado serão apagados.

  Para salvar os resultados do aprendizado permanentemente, use um dicionário de fontes. Para obter mais informações, consulte a seção **Uso de dicionários de fontes**.

- Clique em **Concluir** para salvar todas as soluções oferecidas pelo software.

Se os resultados estiverem incorretos:

- Digite os caracteres corretos e clique no botão **Aprender**.

  ou

- Clique em **Não aprender** para salvar o resultado como inseguro.

  Use esse comando para caracteres danificados que poderiam ser confundidos com outros se aprendidos. Por exemplo, o número 1 e a letra I, que têm uma forma idêntica em muitas fontes.

- Clique em **Excluir** para excluir caracteres da saída.

  Use esse botão para evitar que ruído do documento apareça no arquivo de saída.

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

  O Readiris consegue desfazer as últimas 32 operações.

- Clique em **Desistir** para desistir do aprendizado interativo.

  Todos os resultados do aprendizado serão excluídos. Na próxima vez em que você clicar em **Reconhecer + Salvar**, o aprendizado interativo será reiniciado.

USO DE DICIONÁRIOS DE FONTES

Ao digitalizar vários documentos do mesmo tipo, com a mesma qualidade de fonte e a mesma qualidade de impressão, não convém repetir o processo de aprendizado a cada vez. Portanto, é útil usar os dicionários de fontes. Os dicionários de fontes contêm informações sobre as fontes obtidas durante o **aprendizado interativo** e podem aumentar os resultados do reconhecimento de forma significativa.

Os dicionários de fontes se limitam a 500 formas. É recomendável criar dicionários separados para aplicativos específicos.

Para criar um novo dicionário de fontes:

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

O Readiris abrirá o dicionário **Readiris.dus** por padrão. Altere o nome do arquivo e clique em **Salvar** para salvá-lo como um novo dicionário.

- Clique em **Aprendizado interativo** no menu **Aprender** para ativá-lo.
- Clique em **Reconhecer + Salvar** para reconhecer o documento.
- O Readiris entra na fase de aprendizado interativo. Use os botões da caixa de diálogo para salvar caracteres no dicionário de fontes.

Para usar um dicionário de fontes existente:

- No menu **Aprender**, clique em **Dicionário de fontes**.
- Selecione o dicionário que você deseja usar e clique em **Abrir**.
- No menu **Aprender**, clique em **Adicionar ao dicionário** ou **Ler dicionário**.

Ao selecionar **Adicionar ao dicionário**, certifique-se de ativar **Aprendizado interativo**. O Readiris reconhecerá as formas dos caracteres armazenadas no dicionário e usará o aprendizado interativo, permitindo que você armazene novas informações no dicionário.

Quando você selecionar **Ler dicionário**, o Readiris reconhecerá as formas dos caracteres armazenadas no dicionário, mas não adicionará novo conteúdo ao dicionário, mesmo que o **Aprendizado interativo** esteja ativado. Ainda é útil usar o **Aprendizado interativo** para verificar e, se necessário, corrigir os resultados do reconhecimento que não foram salvos no dicionário de fontes.

Atenção: Não clique em **Dicionário de fontes** no menu **Aprender** e abra um dicionário existente enquanto o modo de dicionário **Novo dicionário** estiver ativado. Caso contrário, o conteúdo do dicionário de fontes existente será apagado.

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

CAPÍTULO 10
FORMATAÇÃO E SALVAMENTO DE DOCUMENTOS

FORMATAÇÃO DE DOCUMENTOS

Os documentos reconhecidos no Readiris podem ser salvos em vários formatos de saída. O Readiris salva os resultados do reconhecimento como arquivos PDF do Adobe Acrobat, arquivos Microsoft XPS, arquivos de texto Word, WordML, RTF e OpenDocument, arquivos HTML e XML, planilhas SpreadsheetML e arquivos de texto Ansi e Unicode.

Além disso, o Readiris pode exportar os resultados diretamente para aplicativos de destino como Microsoft Word e Excel, Adobe Reader, Microsoft XPS Viewer, os principais navegadores da Web, software de e-mail etc.

> Nota: Os arquivos RTF criados pelo Readiris também podem ser abertos nos aplicativos on-line para escritório AjaxWrite, ThinkFree, Zoho Writer e Google Writely (que abre documentos RTF e HTML). Ao utilizar um desses aplicativos, certifique-se de selecionar a opção de gabarito **Manter formatação de palavras e parágrafos**.

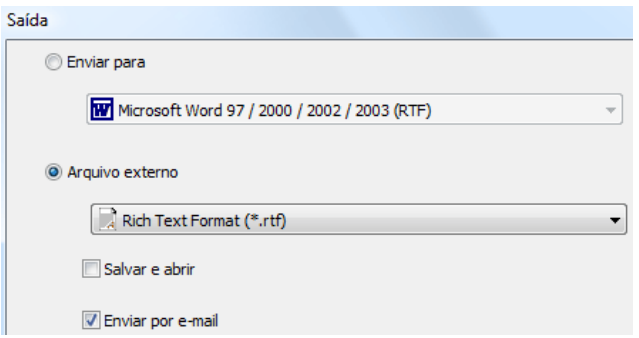
Operação

- Clique no botão **Formato** na barra de ferramentas principal para selecionar o modo de saída.

  O Readiris:

- envia os documentos para um aplicativo, que será aberto automaticamente, ou
- salva os documentos como um arquivo externo.

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



- Clique nas diferentes guias para selecionar as opções que você deseja aplicar.

As configurações que não estão disponíveis para o formato de saída selecionado aparecem esmaecidas.

Os formatos de saída mais usados, bem como suas opções, são analisados nas seções abaixo.

- Clique em **Reconhecer + Salvar** para executar o reconhecimento e salvar os documentos.

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

Por motivos de pesquisa e de ordenação, o Readiris permite definir **propriedades dos documentos** de saída PDF, XPS, Word, RTF, WordML, SpreadsheetML e HTML. Para definir as propriedades de um documento, clique em **Propriedades do documento** no menu

Arquivo. As opções de propriedades do documento também podem ser acessadas na caixa de diálogo **Arquivo de saída**, que é aberta quando você clica em **Reconhecer + Salvar**.

Ao salvar um documento de várias páginas como um **arquivo externo**, você poderá criar um arquivo de saída separado para cada página no Readiris ou salvar todas as páginas pertencentes ao mesmo documento em um único arquivo de saída. Basta clicar nas opções correspondentes na caixa de diálogo **Arquivo de saída**: **Criar um arquivo por página** e **Criar um arquivo por documento**, respectivamente.

No entanto, as opções **Criar um arquivo por página** e **Criar um arquivo por &documento** só estarão disponíveis quando os documentos forem salvos como um arquivo externo, e não quando os documentos forem abertos em um aplicativo de destino.

FORMATAÇÃO DE DOCUMENTOS DE TEXTO

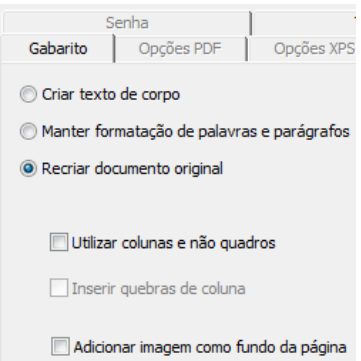
Com o Readiris, você pode gerar vários tipos de formatos de saída de texto. O Readiris oferece saída de texto Word, WordML, RTF, txt e OpenDocument.

Para gerar arquivos de saída de texto:

- Clique no botão **Formato** na barra de ferramentas principal e selecione o formato de saída desejado na lista suspensa **Enviar para** ou **Arquivo externo**.
- Dependendo do formato de texto selecionado, várias opções de formatação estão disponíveis.

  As opções que não estão disponíveis para o formato de texto selecionado aparecem esmaecidas.

Opções de gabarito



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

  O Readiris gera um texto corrido contínuo.

- A opção **Manter formatação de palavras e parágrafos** ocupa uma posição intermediária entre o texto do corpo e a autoformatação.

  O tipo, o tamanho e o estilo de fonte são mantidos durante o reconhecimento.

  As tabulações e o alinhamento de cada bloco são recriados.

  As colunas e os blocos de texto não são recriados; cada parágrafo simplesmente segue o anterior.

  As tabelas são recapturadas corretamente.

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

  O Readiris gera uma cópia verdadeira do documento original, em vez de uma imagem digitalizada.

O Readiris também recria quaisquer hiperlinks para endereços de e-mail e sites.

- A opção **Utilizar colunas e não quadros** cria documentos em formato de colunas.

  Textos em formato de colunas são mais fáceis de editar do que documentos que contêm muitos quadros: o texto flui naturalmente de uma coluna para a próxima.

  Nota: Quando o sistema não puder detectar colunas no documento original, esse modo de formatação usará quadros como uma alternativa.

- A opção **Inserir quebras de coluna** insere uma quebra de coluna forçada no fim de cada coluna.

  Qualquer texto que você editar, adicionar ou remover permanecerá dentro de sua coluna; nenhum texto jamais atravessará automaticamente uma quebra de coluna.

  Dica: Desative essa opção quando trabalhar com texto de corpo em formato de colunas. Isso garante o fluxo natural do texto de uma coluna para a seguinte.

- A opção **Adicionar imagem como fundo da página** posiciona a imagem digitalizada como o fundo da página, abaixo do texto reconhecido.

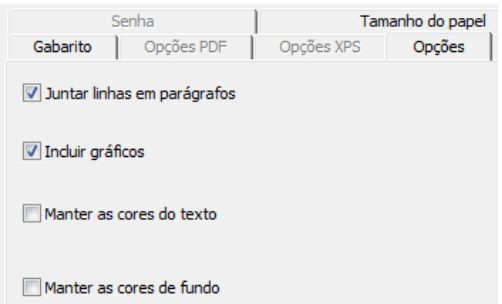
  No entanto, essa opção aumenta o tamanho dos arquivos de saída substancialmente.

  O formato **PDF Texto-Imagem** proporciona o mesmo resultado para arquivos PDF.

  A opção **Manter as cores de fundo**, na guia **Opções**, fornece uma alternativa mais compacta e menos drástica.

Opções gerais

Clique na guia **Opções** para selecionar as opções gerais:



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

  O Readiris quebra automaticamente as linhas do texto reconhecido até que um novo parágrafo seja iniciado e "junta novamente" as partes das palavras que sofreram separação silábica no final da linha.

- A opção **Incluir gráficos** inclui os gráficos em arquivos autoformatados.

  É essencial criar uma cópia verdadeira de um documento.

  Use as opções gráficas da guia **Gráficos** para determinar o modo de cor e a resolução dos gráficos armazenados nos arquivos de saída.

- A opção **Manter as cores do texto** mantém as cores originais do texto durante o reconhecimento.

- A opção **Manter as cores de fundo** mantém as cores especiais do fundo da página durante o reconhecimento.

  Uma cor de fundo uniforme é criada por parágrafo no arquivo de saída.

Tamanhos de papel

Dependendo do formato selecionado, você pode selecionar os tamanhos de papel preferidos:

- Clique na guia **Tamanho do papel** e use os botões de seta para aplicar e excluir tamanhos de papel.
- O Readiris percorre os tamanhos de papel ativos na ordem indicada e utiliza o primeiro tamanho que seja suficientemente grande para conter o documento digitalizado.

FORMATAÇÃO DE DOCUMENTOS BASEADOS EM TABELA

Com o Readiris, você pode transferir dados tabulares para planilhas, processadores de texto e navegadores da Web: as tabelas são reconstruídas, célula por célula, em planilhas e inseridas como objetos de tabela em arquivos de processador de texto.

O Readiris reconhece tabelas com e sem grade.

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	

(com grade)

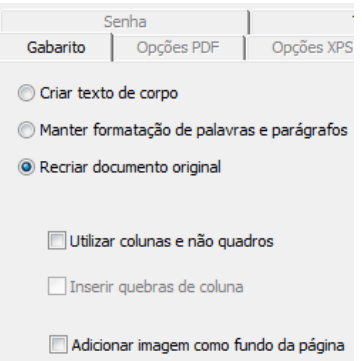
123.985	69	31
987.745	129	2
19	149	91
287.410	49	52

(sem grade)

Para gerar documentos baseados em tabela:

- Clique no botão **Formato** na barra de ferramentas principal e selecione o formato de saída desejado na lista suspensa **Enviar para** ou **Arquivo externo**.

- Selecione as opções de gabarito desejadas.

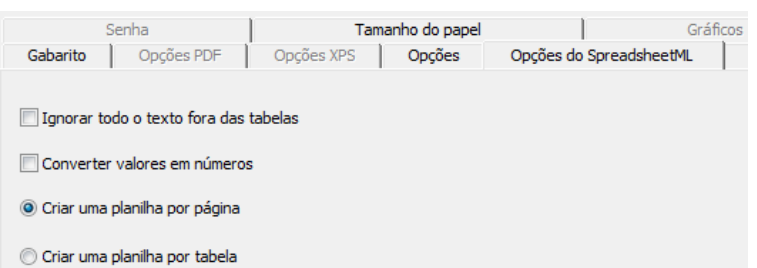


Para obter mais informações sobre as opções de formatação, consulte a seção **Formatação de documentos de texto**.

Opções do SpreadsheetML

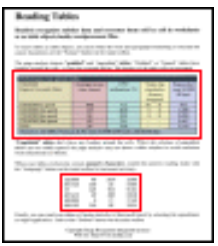
Quando o **Microsoft Excel 2002/2003** é selecionado como o aplicativo de destino, opções específicas do **SpreadsheetML** são disponibilizadas. Clique na guia **Opções do SpreadsheetML** para exibi-las:

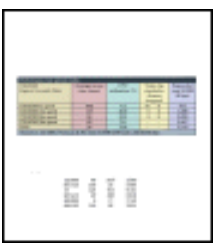
A opção de gabarito **Recriar documento original** ficará indisponível quando esse formato for selecionado.



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

  Todos os dados dentro das tabelas são recapturados; quaisquer dados fora das tabelas são ignorados.





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

  Como resultado, você pode executar operações aritméticas nessas células. As células de texto (em qualquer tabela) permanecem como texto.

  Somente os valores dentro de tabelas são codificados como números.

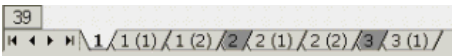
  O Excel executa operações matemáticas exclusivamente em dados codificados como números.

- A opção **Criar uma planilha por página** faz com que apenas uma planilha seja criada por página digitalizada.

  Se uma página contiver tabelas e texto, tudo será colocado na mesma planilha.

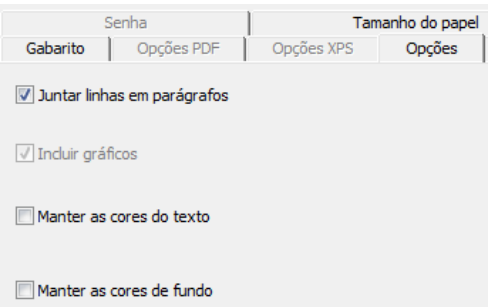
- A opção **Criar uma planilha por tabela** posiciona cada tabela em uma planilha separada e inclui o texto reconhecido (fora das tabelas) em outra planilha.

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



Opções gerais

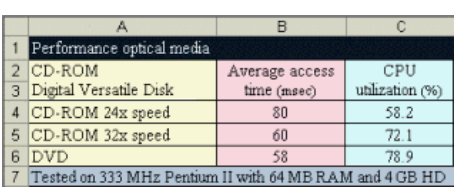
Clique na guia **Opções** para selecionar as opções gerais:



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

  O Readiris quebra automaticamente as linhas do texto reconhecido até que um novo parágrafo seja iniciado e "junta novamente" as partes das palavras que sofreram separação silábica no final da linha.

- A opção **Manter as cores de 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 de papel

Dependendo do formato selecionado, você pode indicar os tamanhos de papel preferidos:

- Clique na guia **Tamanho do papel** e use os botões de seta para aplicar e excluir tamanhos de papel.
- O Readiris percorre os tamanhos de papel ativos na ordem indicada e utiliza o primeiro tamanho que seja suficientemente grande para conter o documento digitalizado.

CRIAÇÃO DE DOCUMENTOS PDF

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

Para gerar saída PDF:

Clique no botão **Formato** na barra de ferramentas principal e selecione o tipo de PDF desejado na lista suspensa **Enviar para** ou **Arquivo externo**:

PDF Imagem

Quando você seleciona PDF **Imagem**, o Readiris gera documentos PDF somente imagem; ele não executa o OCR.

PDF Imagem-Texto

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

A imagem da página é colocada por cima do texto.

Com este formato, você sempre pode ver o documento original (como ele foi digitalizado), como também procurar e copiar-colar o texto reconhecido, que está oculto abaixo da imagem. Por isso, esse formato é útil para **fins de arquivamento**.

PDF Texto

Quando você seleciona PDF **Texto**, o Readiris reconhece o texto e cria arquivos PDF pesquisáveis.

A imagem da página não está contida nos arquivos PDF de uma única camada.

PDF Texto-Imagem

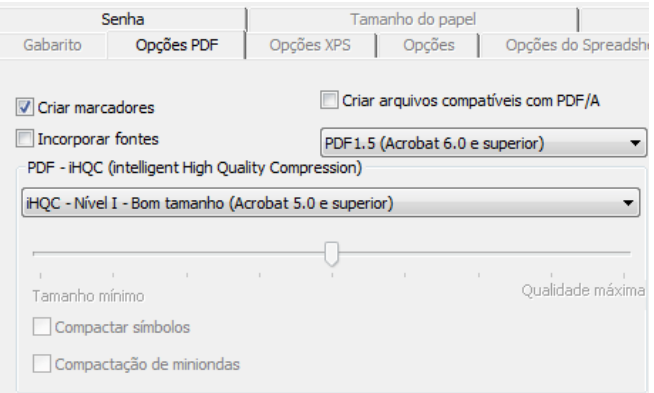
Quando você seleciona PDF **Texto-Imagem**, o Readiris reconhece o texto e cria documentos PDF pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está contida abaixo do texto.

SELEÇÃO DAS OPÇÕES PDF

Para selecionar as opções PDF:

- Clique no botão **Formato** na barra de ferramentas principal e selecione o tipo de PDF desejado na lista suspensa **Enviar para** ou **Arquivo externo**.
- Dependendo do **tipo de PDF** selecionado, várias opções estarão disponíveis. Clique na guia **Opções PDF** para acessá-las:



Criar marcadores

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

Incorporar fontes

Selecione a opção **Incorporar fontes** para incorporar fontes a arquivos PDF do Adobe Acrobat.

A incorporação de fontes evita a substituição de fontes e garante que os leitores, independentemente da configuração do computador, vejam o texto com suas fontes originais.

A incorporação de fontes aumenta um pouco o tamanho do arquivo dos documentos reconhecidos.

COMPACTAÇÃO IHQC DE DOCUMENTOS PDF

Além dos quatro tipos de saída PDF "regulares", o Readiris oferece saída PDF compactada iHQC: os documentos PDF dos tipos

Imagem-Texto e **Imagem** podem ser hipercompactados por meio de iHQC sem perda de qualidade de imagem.

iHQC significa **intelligent High-Quality Compression**, a eficiente tecnologia de compactação exclusiva da I.R.I.S. A iHQC é para imagens o que o MP3 é para músicas e o DivX é para filmes.

Para gerar saída PDF compactada iHQC:

- Clique no botão **Formato** na barra de ferramentas principal e escolha um dos dois modos de saída.
- Na lista **Enviar para** ou **Arquivo externo**, selecione o tipo de PDF desejado: **PDF Imagem-Texto** ou **PDF Imagem**.
- Na guia **Opções PDF**, selecione o nível de compactação necessário.

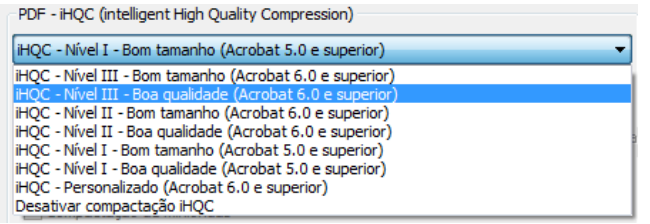
  O Readiris Pro aceita a compactação **Nível I - Bom tamanho** e **Nível I - Boa qualidade**.

  O Readiris Corporate também aceita a compactação **Níveis II e III - Bom tamanho** e **Boa qualidade**, bem como a compactação **Personalizado**.

  Na compactação Nível II, a opção **Compactar símbolos** é ativada automaticamente para compactar o texto de forma condensada.

  Na compactação Nível III, a opção **Compactação de miniondas** é ativada automaticamente para compactar gráficos de forma condensada.

  Ao selecionar a compactação **Personalizado**, você poderá ativar ou desativar essas opções independentemente uma da outra. Você também poderá usar o controle deslizante para definir a relação **Bom tamanho - Boa qualidade**.



- Clique em **Reconhecer + Salvar** para reconhecer o documento e salvá-lo como um arquivo PDF compactado iHQC.

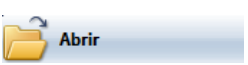
REDIRECIONAMENTO DE DOCUMENTOS PDF

Além de gerar documentos PDF, o Readiris também pode redirecionar arquivos PDF: o Readiris converte PDFs de **imagem** em PDFs de **texto** ou em outro formato de texto compatível e **bloqueia** conteúdo PDF somente leitura.

> **Aviso**: O Readiris não abre documentos PDF protegidos por senha pelo usuário.

Operação

- Clique no botão **Abrir** na barra de ferramentas principal e selecione o arquivo PDF que o Readiris deverá redirecionar.



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

- Clique no botão **Abrir** na caixa de diálogo para abrir o arquivo PDF desejado.

- Clique no botão **Formato** na barra de ferramentas principal e selecione o tipo de PDF desejado.

Para obter mais informações sobre os tipos de PDF, consulte a seção **Criação de documentos PDF**.

- Clique no botão **Reconhecer + Salvar** para redirecionar o documento.

CRIAÇÃO DE DOCUMENTOS XPS

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

XPS significa XML Paper Specification e é um formato de gabarito fixo desenvolvido pela Microsoft.

Para gerar saída XPS:

Clique no botão **Formato** na barra de ferramentas principal e selecione o tipo de XPS desejado na lista suspensa **Enviar para** ou **Arquivo externo**:

XPS Imagem

Quando você seleciona XPS **Imagem**, o Readiris gera documentos XPS somente imagem; ele não executa o OCR.

XPS Imagem-Texto

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

A imagem da página é colocada por cima do texto.

Com este formato, você sempre pode ver o documento original (como ele foi digitalizado), como também procurar e copiar-colar o

texto reconhecido, que está oculto abaixo da imagem. Por isso, esse formato é útil para **fins de arquivamento**.

XPS Texto

Quando você seleciona XPS **Texto**, o Readiris reconhece o texto e cria arquivos XPS pesquisáveis.

A imagem da página não está contida nos arquivos XPS de uma única camada.

XPS Texto-Imagem

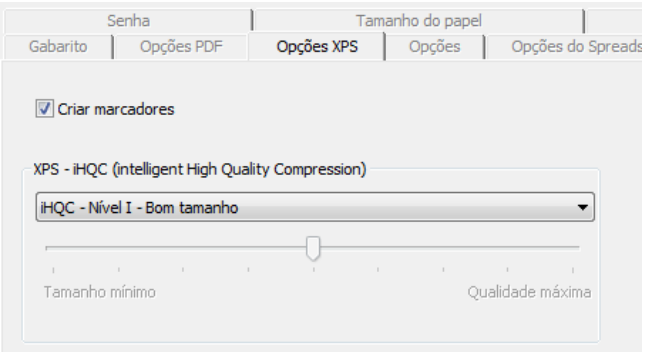
Quando você seleciona XPS **Texto-Imagem**, o Readiris reconhece o texto e cria documentos XPS pesquisáveis que contêm a imagem da página e o texto reconhecido.

A imagem da página está contida abaixo do texto.

SELEÇÃO DAS OPÇÕES XPS

Para selecionar as opções XPS:

- Clique no botão **Formato** na barra de ferramentas principal e selecione o tipo de XPS desejado na lista suspensa **Enviar para** ou **Arquivo externo**.
- Dependendo do **tipo de XPS** selecionado, várias opções estarão disponíveis. Clique na guia **Opções XPS** para acessá-las:



Criar marcadores

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

COMPACTAÇÃO IHQC DE DOCUMENTOS XPS

Além dos quatro tipos de saída XPS "regulares", o Readiris oferece saída XPS compactada iHQC. Os documentos XPS dos tipos **Imagem-Texto** e **Imagem** podem ser hipercompactados por meio de iHQC.

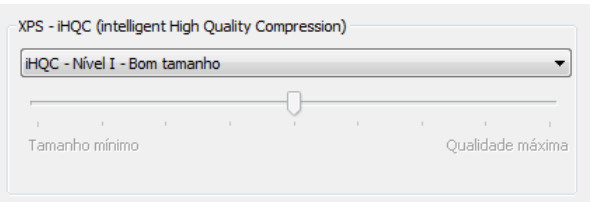
iHQC significa **intelligent High-Quality Compression**, a eficiente tecnologia de compactação exclusiva da I.R.I.S. A iHQC é para imagens o que o MP3 é para músicas e o DivX é para filmes.

Para gerar saída XPS compactada iHQC:

- Clique no botão **Formato** na barra de ferramentas principal e escolha um dos dois modos de saída.
- Na lista **Enviar para** ou **Arquivo externo**, selecione o tipo de XPS desejado: **XPS Imagem-Texto** ou **XPS Imagem**.

- Na guia **Opções XPS**, selecione o nível de compactação necessário.

  O Readiris aceita a compactação **Nível I - Bom tamanho** e **Nível I - Boa qualidade**.

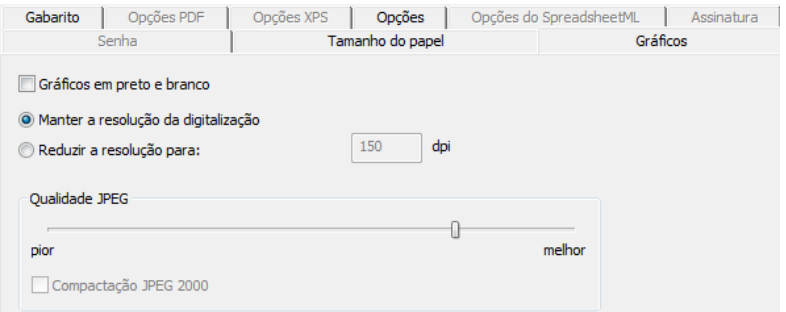


SELEÇÃO DAS OPÇÕES DE GRÁFICOS

Dependendo do formato de saída e do aplicativo de destino selecionados, opções avançadas de gráficos podem estar disponíveis. As opções de gráficos podem ser usadas para alterar a resolução e a qualidade da imagem.

Para acessar as opções de gráficos:

- Clique no botão **Formato** na barra de ferramentas principal e selecione o formato de saída desejado na lista suspensa **Enviar para** ou **Arquivo externo**.
- Clique na guia **Gráficos** para as opções.



Modo de cor

Por padrão, o Readiris salva gráficos em **cores**.

Selecione **Preto-e-branco** para salvar os gráficos em preto-e-branco.

Resolução

Por padrão, o Readiris **mantém a resolução da digitalização**.

Você pode optar por **reduzir a resolução**.

Dica: Ao salvar documentos como arquivos HTML para publicação em um site, reduza a resolução para 70 dpi (resolução da tela).

Qualidade JPEG

Os gráficos armazenados em documentos PDF, XPS, Word e RTF são salvos no formato JPEG.

Use o controle deslizante para ajustar a qualidade JPEG.

Compactação JPEG 2000

Ao salvar arquivos no formato PDF ou XPS, o Readiris pode aplicar a compactação JPEG 2000 às imagens em cores ou em tons de cinza armazenadas nesses arquivos.

JPEG 2000 é a última versão mais compacta do padrão JPEG.

Selecione a opção **Compactação JPEG 2000** para aplicá-la.

CAPÍTULO 11
SALVAMENTO E CARREGAMENTO DE CONFIGURAÇÕES

SALVAMENTO E CARREGAMENTO DE CONFIGURAÇÕES

Quaisquer configurações especificadas no Readiris serão salvas automaticamente para uso futuro depois que o aplicativo for fechado. Para restaurar as configurações de fábrica, clique no comando **Restaurar configuração de fábrica** no menu **Arquivo**.

Ao digitalizar vários grupos de documentos que exigem configurações diferentes, é útil salvar **arquivos de configurações** separados para cada grupo.

Operação

- Selecione as configurações que você deseja utilizar para um determinado grupo de documentos.
- No menu **Arquivo**, clique no comando **Salvar configuração**.
- Ao digitalizar ou abrir um documento do mesmo grupo posteriormente, clique no comando **Carregar configuração** no menu **Arquivo**.
- Selecione o arquivo de configurações correto e clique no botão **Abrir**.

- Clique em **Reconhecer + Salvar** para reconhecer o documento, usando as configurações corretas.

  Nota: O comando **Informações**, no menu **Arquivo**, apresenta uma visão geral das configurações mais importantes selecionadas.

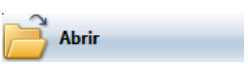
CAPÍTULO 12
RECONHECIMENTO DE DOCUMENTOS DE VÁRIAS PÁGINAS

ABERTURA E RECONHECIMENTO DE VÁRIOS ARQUIVOS DE IMAGENS

O Readiris foi concebido para processar vários arquivos de imagens ao mesmo tempo.

Para abrir vários arquivos de imagens:

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



- Selecione os arquivos de imagens que você deseja abrir:
    - Selecione o primeiro arquivo de imagem e mantenha pressionada a tecla **Ctrl** enquanto seleciona outras imagens ou
    - Selecione um intervalo contínuo de arquivos clicando na primeira imagem e mantendo pressionada a tecla **Shift** enquanto seleciona a última imagem.

Nota: Ao abrir um único arquivo que consiste em várias páginas, por exemplo, um documento PDF ou um arquivo TIFF de várias páginas, você poderá indicar o intervalo de páginas. Na área **Intervalo de páginas** da caixa de diálogo, selecione **Páginas** e indique quais páginas deseja reconhecer.

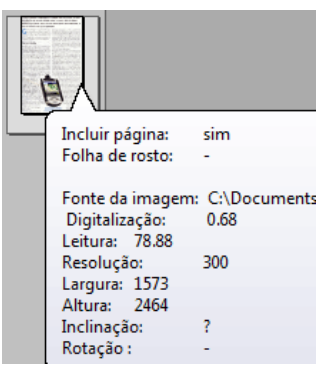
- Clique no botão **Abrir** para abrir as imagens.

  Você também pode arrastar e soltar arquivos de imagens do Windows Explorer para a janela de imagem do Readiris para abri-los.

- A barra de ferramentas de página exibirá os arquivos de imagens abertos.

  Dica: Mantenha o cursor do mouse sobre as miniaturas das páginas para exibir as informações de configuração de cada página.

  A barra de ferramentas de página pode ser usada para editar documentos de várias páginas. Para obter mais informações, consulte a seção **Edição de documentos de várias páginas**.



- Determine as configurações do reconhecimento e clique em **Reconhecer + Salvar** para executar o reconhecimento.

- Caso deseje abrir ou digitalizar outras imagens para o documento atual, clique no botão **Digitalizar** ou **Abrir** na barra de ferramentas principal.

  Será perguntado se você deseja excluir o documento atual ou não. Clique em **sim** para excluir o documento atual e iniciar um novo documento ou clique em **não** para adicionar outras digitalizações ao documento atual.

DIGITALIZAÇÃO E RECONHECIMENTO DE DOCUMENTOS DE VÁRIAS PÁGINAS

O Readiris foi concebido para processar documentos de várias páginas.

O Readiris Pro processa documentos de até 50 páginas.

O Readiris Corporate processa documentos com um número ilimitado de páginas.

Para digitalizar documentos de várias páginas no Readiris, você pode usar a função **alimentador automático de documentos**, ao utilizar um digitalizador de produção, ou usar a função **digitalização com intervalo**, ao utilizar um digitalizador de mesa.

Digitalizar documentos de várias páginas usando um alimentador de documentos (digitalizador de produção)

- Clique no botão **Digitalizador** na barra de ferramentas principal e selecione a opção **ADF** (Automatic Document Feeder, alimentador automático de documentos).
- Coloque as páginas no alimentador de documentos do digitalizador e clique em **Digitalizar** para iniciar a digitalização.
- Clique em **Reconhecer + Salvar** para reconhecer os documentos.

Digitalizar documentos de várias páginas usando a digitalização com intervalo (digitalizador de mesa)

- Clique no botão **Digitalizador** na barra de ferramentas principal.

- Selecione **Digitalizar outra página após** e indique o intervalo de tempo usando os botões de seta.

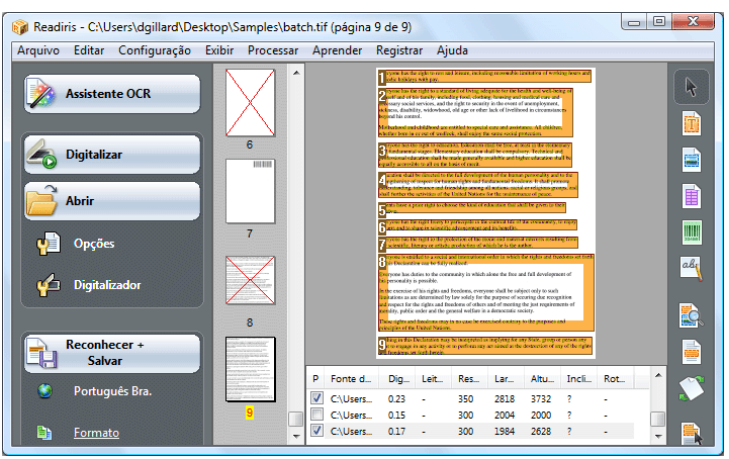
  O digitalizador digitalizará outra página automaticamente após o número indicado de segundos, sem que você precise clicar no botão **Digitalizar** a cada vez.

  Para terminar a digitalização automática, clique em **Desistir** na caixa de diálogo de digitalização com intervalo ou pressione **ESC** no teclado.

  Para congelar o intervalo de digitalização, clique em **Pausar** na caixa de diálogo de digitalização com intervalo ou pressione a barra de espaço no teclado. Clique em **Continuar** quando estiver pronto para prosseguir.

EDIÇÃO DE DOCUMENTOS DE VÁRIAS PÁGINAS

Quando vários documentos são abertos ou digitalizados no Readiris, a barra de ferramentas de página exibe suas miniaturas. As miniaturas na barra de ferramentas de página podem ser usadas para editar documentos de várias páginas.



Para mover uma página em um documento:

- Clique com o botão direito do mouse na página que deseja mover e clique em **Selecionar página**.
- Arraste a página até a posição correta.
- Ou clique com o botão direito do mouse em uma página e clique em **Mover página para cima** ou **Mover página para baixo**.

Para excluir uma página:

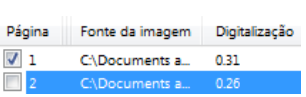
- Clique com o botão direito do mouse na página que deseja excluir e clique em **Apagar página**.
- Ou selecione a página e pressione no botão **Delete** no teclado.

Para excluir uma página do reconhecimento:

- Clique com o botão direito do mouse na página que deseja excluir e clique em **Excluir página**.
- Ou limpe a caixa correspondente ao número da página no painel de documentos.

As páginas excluídas são eliminadas da barra de ferramentas de página.

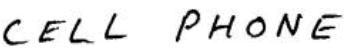
As páginas excluídas serão ignoradas quando você imprimir as imagens digitalizadas e quando salvar as digitalizações em arquivos de imagens de várias páginas.



Dica: Os comandos **Incluir todas as páginas** e **Excluir todas as páginas**, do menu **Editar**, aplicam-se a todas as páginas simultaneamente.

CAPÍTULO 13
RECONHECIMENTO DE TEXTO MANUSCRITO

Além de códigos de barras, gráficos, tabelas e texto digitado, o Readiris reconhece texto **manuscrito**. O manuscrito consiste em letras maiúsculas separadas.



É necessário software de reconhecimento inteligente de caracteres (ICR - Intelligent Character Recognition) altamente especializado para reconhecer caracteres manuscritos.

Para reconhecer manuscrito:

- Clique no botão de manuscrito na barra de ferramentas de imagem.

- Desenhe um quadro ao redor do texto manuscrito.
- Clique em **Reconhecer + Salvar** na barra de ferramentas principal.

Todo o documento, inclusive o texto manuscrito, será reconhecido.

Nota: Clique com o botão direito do mouse na zona manuscrita e clique em **Copiar como texto** para reconhecer somente a zona manuscrita e enviá-la para a área de transferência.

Símbolos reconhecidos

O reconhecimento de manuscrito limita-se ao alfabeto latino e aceita números (0 a 9), letras maiúsculas (A a Z) e estes sinais de pontuação: vírgula, ponto, sinal de mais e hífen.

Acentos, tremas e outros caracteres especiais não são aceitos.

Notas

- O Readiris suporta **manuscrito**, não **escrita à mão**.

  Para obter mais informações, consulte a seção **Regras de manuscrito**.

- Caracteres maiúsculos são substituídos por caracteres minúsculos após o reconhecimento, a menos que ocorram no início de uma sentença.

- As características do documento (idioma, tipo de fonte e espaçamento de caracteres) não se aplicam ao manuscrito.

- O aprendizado interativo também não se aplica. A tecnologia ICR se baseia em mais de um milhão de exemplos de escrita.

REGRAS DE MANUSCRITO

Várias regras devem ser levadas em consideração para que o Readiris reconheça manuscrito:

- Escreva caracteres regulares bem espaçados.

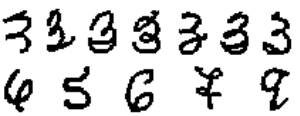
Observe como os caracteres A, G e Q são escritos.

- Use uma caneta esferográfica bem grossa.

  Canetas pretas produzem resultados melhores do que as azuis.

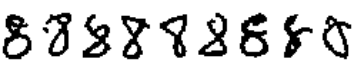
  Não use lápis.

- Não estilize demais.



  Caracteres excessivamente estilizados aumentam o risco de erros de reconhecimento.

- Não abra voltas que deveriam estar fechadas; não feche voltas que deveriam estar abertas.



- Evite caracteres interrompidos.

- Evite tornar a traçar.

  Isso reduz a qualidade da imagem e a nitidez dos símbolos manuscritos.

  Os caracteres inteiramente tachados não são reconhecidos.

- Escreva números 1 corretamente.

  O número 1 pode ser escrito no estilo "anglicizado" ou "europeu".

O número 1 pode ser sublinhado ou não.

A barra de sublinhado horizontal não precisa encostar no resto da forma da fonte.

Dica: Quando não conseguir obter resultados ideais, use o **formulário de escrita** da I.R.I.S. e adapte seu estilo de escrita. O formulário de escrita em branco da I.R.I.S. funciona como um modelo de página inteira no qual as letras maiúsculas são preenchidas corretamente e com o tamanho certo. O formulário pode ser encontrado no CD-ROM do Readiris e na pasta de instalação do Readiris.

CAPÍTULO 14
RECONHECIMENTO DE CÓDIGOS DE BARRAS

INTRODUÇÃO À LEITURA DE CÓDIGOS DE BARRAS

Além do reconhecimento óptico de caracteres de 128 idiomas, o Readiris também oferece a leitura de códigos de barras.

Há suporte para todas as simbologias conhecidas de código de barras: Codabar, Código 128, Código 39, Código 39 estendido, Código 39 HIBC, Código 93, Datalogic 2 de 5, Código Distinto 2 de 5, EAN-13, EAN-8, 2 de 5 intercalado, MSI farmacêutico, MSI-Plessey, Código de patch Kodak, PDF-417, PostNet, UCC-128, UPC-A e UPC-E.

Para que o Readiris execute o reconhecimento, são necessários códigos de barras impressos a laser e a jato de tinta. Códigos de barras impressos em impressoras matriciais não são aceitos, uma vez que não produzem contraste suficiente e sua resolução é, quase sempre, limitada a 60 dpi.

Leitura manual de códigos de barras

- Determine quais códigos de barras o Readiris deverá reconhecer:
- No menu **Configurações**, clique em **Códigos de barras**.
- Selecione as simbologias que o Readiris deverá reconhecer.
- Determine se o Readiris deverá verificar ou remover os caracteres-dígitos.
- Clique no botão de código de barras na barra de ferramentas de imagem e desenhe um quadro ao redor das zonas de códigos de barras no documento.

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

Todo o documento, inclusive o conteúdo dos códigos de barras, será reconhecido.

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

HNDICE

A

ADF 85

ajuste de documentos digitalizados..... 31

alimentador automático de documentos..... 85

alinhamento 23

alinhamento de páginas..... 23

alteração da interface do usuário 18

análise de página..... 23

aprendizado interativo..... 55

arquivo de configurações 81

arquivos de gabarito..... 45

Assistente..... 13

B

barra de ferramentas de imagem 17

barra de ferramentas principal 16

brilho..... 28, 34

C

câmera digital 29

características do documento 53

carregamento de configurações 81

códigos de barras 93

compactação JPEG 2000 79

configurações de fábrica.....81

configurações do digitalizador 27

contraste 28, 34

cor de fundo.....63

cor de fundo de células de tabela 68

D

desinstalação do Readiris 11

dicionários de fontes.....57

digitalização com intervalo...85

divisão automática em janelas 39

divisão manual em janelas ... 41

documentos asiáticos 4, 7, 50

documentos de planilha 67

documentos de texto 61

documentos de várias páginas 83, 85

documentos em árabe....4, 7, 51

documentos em hebraico... 4, 7, 51

E

edição de documentos de várias páginas 86

edição para a Ásia 4, 7, 50

edição para o Oriente Médio. 4, 7, 51

envio de documentos por e-mail 60

espaçamento de caracteres ... 54

F

formatos de imagem compatíveis 26

formatos de saída 59

I

idioma 50

idioma da interface do usuário 18

idiomas mistos.....52

imagem colorida.....28, 31

imagem em preto-e-branco...28

imagem em tons de cinza 28

imagens invertidas.....29

inclinação 30

instalação.....9

interface do usuário 15

L

limpeza 34

M

manuscrito 89, 90

matricial.....54

modelos de janela 45

modo de cor.....28

N

numérico.....51

O

opções.....23

opções de gabarito 62

opções de gráficos 77

opções PDF 70

opções XPS 75

P

páginas 86

excluir 87

mover 87

selecionar 87

painel de documentos..... 17

planilhas..... 67

precisão x velocidade..... 52

propriedades do documento . 59

R

recriação de documentos originais 62

redirecionamento de documentos PDF..... 73

registro 11

requisitos de sistema 9

resolução 27

restauração de configurações de fábrica 81

rotação 23

S

saída Excel 59

saída HTML..... 59

saída OpenDocument 59

saída PDF 59, 69

saída PDF iHQC.....71

saída RTF 59

saída SpreadsheetML 59

saída Unicode 59

saída Word 59

saída WordML.....59

saída XML.....59

saída XPS 59, 74

saída XPS iHQC.....76

salvamento de configurações 81

SmartTasks 16, 19

suavização de imagens coloridas 30, 31

suporte ao produto 12

T

tabelas.....65

tamanhos de papel 65

tipo de fonte.....53

V

velocidade x precisão 52