Readiris™ Pro 12

ユーザガイド

目次

著作権 1

第 1 章　Readirisについて 3

時間の節約、再タイプの必要なし 3

Readirisシリーズ 6

第 2 章　Readirisのインストール 9

システム動作環境 9

ソフトウェアのインストール 10

ソフトウェアのアンインストール 11

ソフトウェアの登録 12

製品サポート 12

第 3 章　はじめに 13

Readirisの実行 13

OCRウィザードの利用 13

ユーザインターフェース 15

ユーザインターフェース言語の変更 18

第 4 章　Readiris SmartTasks 19

Readiris SmartTasksの利用 19

第 5 章　文書のスキャン 23

オプションの選択 23

画像ファイルを開く 24

用紙文書のスキャン 26

第 6 章　スキャンした文書の調整 33

第 7 章 文書を画像ファイルとして保存 39

第 8 章　文書のウィンドウイング 41

文書を自動的にウィンドウイング 41

文書を手動でウィンドウイング 43

ウィンドウイング テンプレートの使用 47

第 9 章　文書を認識 51

はじめに 51

文書の言語を選択 52

文書の特徴を定義 54

インタラクティブ学習の利用 56

フォント辞書の使用 59

第 10 章 文書のフォーマットと保存 61

文書のフォーマット 61

テキスト文書のフォーマット 63

表からなる文書のフォーマット 67

PDF文書の作成 71

PDFオプションの選択 72

PDF文書をiHQCで圧縮 74

PDF文書の目的変更 75

XPS文書の作成 76

XPSオプションの選択 78

XPS文書をiHQCで圧縮 79

グラフィック オプションの選択 80

第 11 章　設定の保存と読み込み 83

第 12 章　複数ページの文書を認識 85

複数の画像ファイルを開いて認識 85

複数ページの文書をスキャンして認識 87

複数ページの文書を編集 88

第 13 章　手書きテキストの認識 91

第 14 章　バーコードの認識 95

索引 97

著作権

ReadirisPro12-dgi-090223-01

Copyrights ©1987-2009 I.R.I.S. All Rights Reserved.

I.R.I.S. は、Readirisソフトウェア、オンライン ヘルプ
システムおよび本出版物の著作権を有します。

本資料に記載されている情報は、I.R.I.S.
の財産です。コンテンツは、予告なく変更されることがあり、I.R.I.S.はコンテンツについて公約を表明しません。本資料で説明されているソフトウェアは、本製品のご利用条件が記載されているライセンス契約に基づいて提供されます。
本ソフトウェアは、契約条件に従ってのみ使用またはコピーを行うことができます。
本出版物のいかなる部分も、I.R.I.S.の書面による事前の許可を得ずに検索システムに複製、転送、保管したり、別の言語に翻訳したりすることはできません。

本ユーザガイドは、実演を行うために架空の名前を使用しています。実在の人物、企業または組織が言及される場合、まったくの偶然によるものです。

商標

Readirisロゴ、およびReadirisは、Image Recognition Integrated Systems S.A.の商標です。
I.R.I.S. によるOCR、ICR、およびバーコードテクノロジー
I.R.I.S. によるAutoFormatと言語テクノロジー
I.R.I.S. によるBCR、およびフィールド分析テクノロジー
I.R.I.S. による iHQC圧縮テクノロジー
Apacheにより開発されたXMLパーサ。この製品はApache Software Foundationにより開発されたソフトウェアを含みます。

本ユーザガイドで参照しているその他の全製品は、各企業の商標または登録商標であり、各企業に帰属します。

第 1 章
READIRISについて

時間の節約、再タイプの必要なし

Readirisをご利用く ださり、ありがとうございます。
このソフトウェアは、皆様のテキスト、表、グラフィック、バーコード、手書きテキストを再キャプチャするのに、必ず大きな助けとなります。

コンピュータは効率的ではありますが、まず最初に情報を入力しなければなりません。
今までに15ページものレポートや数値がいっぱいの表を再タイプしたことがあるお方は、どれほど手間がかかり、時間がかかるかおわかりだと思います。
この最新術を駆使したOCRパッケージを利用すれば、用紙文書やスキャンした画像ファイルを自動的にテキスト検索、編集が可能な文書に変換し、アーカイブ、共有することもできます。
最高速が得られる認識モードと最適のOCR精度を保証する認識モードの2種類をご利用いただけます。

印刷、またはタイプされた文書をスキャンし、Readirisで認識したゾーンを指定するか、システムで自動検出し、文字認識を実行し、文書をワードプロセッサにエクスポートします。
多数のページからなる文書でも、初めから終わりまで簡単な操作で処理できます。
何時間もタイプに時間を費やさなくても、Readirisは、用紙文書を編集可能なコンピュータファイルにたった数クリックで変換できます。
手動入力と比較し、最高40倍の速さです。

ウィザードが、Readiris™を操作するのに必要な設定を順次ご説明しますから、手早く簡単に結果が得られます。

SmartTasksをご使用になると、さらにプロセスが速まります。
読み取った結果を直接ワードプロセッサやスプレッドシートに送信したり、PDFやXPSファイルなどにアーカイブできます。
ファックスを認識し、PDF文書に変換するには、Windows Explorerから画像ファイルをReadirisアプリケーション
ウィンドウにドラッグします。または、画像をReadirisに直ちにコンテキストメニューから送信します。

Readirisは、表データを認識し、スプレッドシート
ソフトウェアのワークシート、またはワードプロセッサ内の表オブジェクトとして再作成します。数値データは、即座に次の処理に利用できるようになります。

Readirisは、最新鋭の認識テクノロジーを基にしています。
フォントに依存しないテキスト認識は、自習テクニックで補完されています。システムは、文脈や言語分析を通じ、新しい文字や言葉を学ぶことができます。
つまり、認識システムのOCR精度が、使用を重ねるにつれ向上することになります。

Readirisは、最高で128ヶ国語の言語をサポートします。
中欧、バルト海諸国、キリル文字の言語、およびギリシャ語やトルコ語を含むアメリカとヨーロッパの言語すべてに対応しています。
オプションとして、Readirisはアラビア語やペルシア語（ファルシ）、ヘブライ語、アジア4カ国の言語
(日本語、韓国語、簡体字中国語、および繁体字中国語)も読むことができます。Readirisは、混合アルファベットにも対応できます。
ソフトウェアは、ギリシャ語やキリル文字の言語、アラビア語、ヘブライ語、アジア諸語の文書に出てくる「西洋文字」の言葉を検知します。別の言語に転記することができない固有名詞、ブランド名などの多くが、西洋の文字や記号で表記されていることがあります。

Readirisは、認識の後ではなく、認識中に言語学を駆使します。
そのためReadirisは、品質の劣る文書、ファックス、ドットマトリックスの印刷物などを含めあらゆる種類の文書を最高の精度で認識します。
スキャンやコピーの仕方が悪かったために、フォントが薄すぎたり、濃すぎたりする文書でも、スムーズに対処します。

ドットマトリックス記号など切れ切れになった文字を組み立てなおし、一緒になった文字も解消します。

その他にも、Readiris には(オプションの)ユーザ確認機能があります。
このユーザ確認機能(「インタラクティブ学習」)をオンにすると、認識システムがはっきりわからない文字にフラグを立てるだけではなく、システムの精度を向上させることができます。
あなたが確認したすべてのソリューションを記憶し、システムのスピードと確実性を向上し、使用を重ねるにつれ、システムがさらにインテリジェントを増します。
このパワフルな学習ツールは、数学記号や絵記号などの特殊文字をReadirisに認識させる訓練となり、ゆがんだフォントを処理させることもできます。

さらに生産性を高めるために、Readirisはテキストを認識するだけではなく、あなたに代わってフォーマットすることもできます。
様々なレベルのフォーマットをご利用いただけます。
「自動フォーマット」をご利用になると、Readirisはスキャンした文書のファックスコピーを再作成し、元の文書の文字、段落、ページの書式を維持します。
認識全体において、類似の文字が使用され、ソース文書で使用された文字サイズ、タイプスタイルが維持されます。
段組、テキストブロック、グラフィックの配置は、元の文書に従います。
Readirisは、スキャンしたページの背景写真ですら認識した文書に含めることができます。
Readirisは、グレースケールとカラーのスキャンを楽々とサポートしますから、ラインアート、白黒の写真、カラーのイラストなどどんなグラフィックでも再キャプチャーできます。
文書に表がある場合は、Readirisは実際のセルに認識し、元の表のセルの枠を再作成します。

つまり、Readiris
はあなたの文書の真のコピーを作り上げることができ、スキャンした画像ではなく、編集可能でコンパクトなテキストファイルにするのです。

スキャンしたページにあるバーコードを読み取ることもできます。また、間隔を十分にとった「ブロック体」の手書きのテキストも読み取ることができます。

Readirisは、Twain順守であり、フラットベッドやシートフィードのスキャナ、「一体型」デバイス、「MFP」(多機能周辺機器)、デジタルカメラをサポートします。スキャナに文書フィーダがない場合、インターバルスキャンで複数ページの文書を楽にスキャンできます。

READIRISシリーズ

下表は、ご利用いただけるバージョンの概要です。

Readiris Home 12 制限付き機能 認識言語25ヶ国語 PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート PDF画像-テキスト、DOCX、ODT、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFF出力を生成。

Readiris Pro 12 基本機能 128 ヶ国語の 認識言語 PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート 4種類のPDFファイル、PDF-iHQC (レベルI)、4種類のXPSファイル、XPS-	**Readiris Corporate 12** 基本機能 128 ヶ国語の 認識言語 PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート 4種類のPDFファイル、PDF-iHQC (レベルI～III)、4種類のXPSファイル、

iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成	XPS-iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成 大量認識 自動処理 文書インデックス作成 名刺認識

Readiris Pro 12 Asian	**Readiris Corporate 12 Asian**
基本機能 128 ヶ国語の 認識言語 PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート 4種類のPDFファイル、PDF-iHQC (レベルI)、4種類のXPSファイル、XPS-iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成 繁体字、簡体字の中国語認識 日本語認識 韓国語認識	基本機能 128 ヶ国語の 認識言語 PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート 4種類のPDFファイル、PDF-iHQC (レベルI～III)、4種類のXPSファイル、XPS-iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成 繁体字、簡体字の中国語認識 日本語認識 韓国語認識 大量認識 自動処理 文書インデックス作成 名刺認識

Readiris Pro 12 Middle-East*	**Readiris Corporate 12 Middle-East***
基本機能	基本機能
128 ヶ国語の 認識言語	128 ヶ国語の 認識言語
PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート	PDF、DCX、DJV、DJVU、JPG、JPEG、J2C、J2K、JP2、PNG、TIF、TIFF、BMP、PCX画像をサポート
4種類のPDFファイル、PDF-iHQC (レベルI)、4種類のXPSファイル、XPS-iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成	4種類のPDFファイル、PDF-iHQC (レベルI～III)、4種類のXPSファイル、XPS-iHQC (レベルI)、DOCX、ODT、XLS、WordML、SpreadsheetML、RTF、HTM、XML、TXT、TIFFなどを生成
アラビア語、ペルシア語(ファルシ)認識	アラビア語、ペルシア語(ファルシ)認識
ヘブライ語認識	ヘブライ語認識
	大量認識
	自動処理
	文書インデックス作成
	名刺認識

(*) Macバージョンはありません。

第 2 章
READIRISのインストール

システム動作環境

Readirisを使用するために必要な**最小システム構成**

- 486ベースIntel PC、または互換機。
  PentiumベースPCを推奨。
- 256 MB RAM。
- 120 MBのハードディスク空き容量。

  (サンプルファイルをインストールしない場合は、105MBのハードディスク空き容量で十分です)

- Windows Vista、Windows XP または Windows 2000オペレーティングシステムに対応。

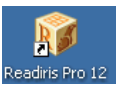
スキャナドライブの一部は、Windows
最新バージョンで稼動しない場合がありますのでご留意ください。
サポートされているプラットフォームに関しては、スキャナに同梱されていた文書をご参照ください。

ソフトウェアのインストール

Readirisのインストールの仕方:

- Windowsに管理者としてログオンするか、ソフトウェアをインストールするのに必要な管理者権限が必要となります。
- スキャナをPCに接続し、該当するソフトウェアをインストールします。
  スキャナをテストします。
  問題が生じた場合は、スキャナの製造会社にご連絡ください。
- CD-ROMドライブに、Readiris CD-ROMを挿入し、画面上の指示に従って、ソフトウェアをインストールします。
- [**Readiris**]をクリックして、インストールを開始します。(追加のソフトウェア製品、 Copernic Desktop Search Home EditionとCardiris 4 LEもお求めいただけます。)
- インストール言語を選択して[**OK**]をクリックします。
- ライセンス契約の規約を受諾します。
- **完全インストール**と**カスタムインストール**から選択できます。
  必要なオプションを選択して、毎回[**次へ**]をクリックして次に進みます。
  語彙とサンプル画像、および電子ユーザガイド、オンラインヘルプが、デフォルトでインストールされます。
- [**終了**]をクリックし、インストールを完了します。
  インストールプログラムにより、サブメニューの「**I.R.I.S.**

Applications - Readiris」がWindows[**プログラム**]メニューに自動的に作成されます。
インストールプログラムはまた、Windowsデスクトップに Readiris
アプリケーションのショートカットを作成します。



- インストールプロセスを繰り返し、CD-ROMのその他のソフトウェアをインストールします。

ソフトウェアのアンインストール

Readirisを正しくアンインストール方法は1つしかありません。
Windows(アン)インストール ウィザードをご使用ください。
プログラムファイルを手動で削除して、Readiris、またはそのソフトウェアモジュールをアンインストール*しない*でください。

Readirisをアンインストールする仕方:

- アプリケーションを閉じます。
- [**スタート**]メニューから[**コントロールパネル**]をクリックします。
- [**プログラム**]アイコンから[**プログラムのアンインストール**]をクリックします。
- リストの中から[**Readiris**]を選択し、[**アンインストール**]をクリックします。
- 画面上の指示に従います。

ソフトウェアの登録

Readiris ライセンスをご登録ください。登録されますと、

- 今後の製品開発情報やI,R.I.S.
  の関連製品に関するご案内を差し上げます。
- 無料で製品サポートが受けられます。
- I.R.I.S. 製品が特価で提供されます。

登録の仕方:

[登録]メニューの「登録ウィザード」を利用します。
画面上の手順に従います。

製品サポート

製品の登録が済みましたら、I.R.I.S.からソフトウェアのすべての基本機能について製品サポートが受けられます。I.R.I.S.へのお問い合わせ:

ヨーロッパ:
support@irislink.com
Tel: 32 10 45 13 64

米国:
support@irisusa.com
Tel.: 1 800 447 4744

アジア太平洋:
support@irislink.com
Tel.: 852 22646133

第 3 章
はじめに

READIRISの実行

Readirisを実行する仕方:

- Windows
  [スタート]メニューから「**Readiris**」を起動するか、デスクトップ上のショートカットをダブルクリックします。

- スタートアップ画面をクリックして「Readiris」を起動します。

  OCR ウィザードが自動的に開きます。

OCRウィザードの利用

OCRードで、Readirisを操作するのに必要な設定すべてを素早く 定義することができます。

Readirisを起動する場合は、スタートアップ画面をクリックしてOCRウィザードを開始します。

ステップ1
画像ソースを選択します。
スキャナで画像をキャプチャできますし、**画像ファイル**を開く こともできます。
ご希望の**回転**、**傾き補正**のオプションを選択します。

詳細に関しては、「**オプションの選択**」をご参照ください。

Readirisに馴れるには、ソフトウェアに搭載されているサンプル画像を利用します。

画像サンプルは、Readiris CD-ROMとReadirisインストールフォルダのサブフォルダ「**サンプル**」にあります。

[**次へ**]をクリックして次のステップに進みます。

ステップ2
スキャナを選択した場合は、[**変更**]をクリックしてスキャナ設定を選択します。

スキャナ設定の詳細に関しては、「**用紙文書のスキャン**」をご参照ください。

[**OK**]をクリックして、設定を保存します。
[**次へ**]をクリックして次のステップに進みます。

ステップ3
[**変更**]をクリックして文書言語を変更します。
文書言語は、デフォルトではアメリカ英語に設定されています。
リストから必要な言語、または言語の組み合わせを選択し、[**OK**]をクリックします。
スライダを使用して必要な**速度**/**精度**設定を定めます。

詳細に関しては、「**文書の言語を選択**」をご参照ください。

[**OK**]をクリックして、設定を保存します。
[**次へ**]をクリックして次のステップに進みます。

ステップ4
[**変更**]をクリックして出力フォーマット、ターゲット
アプリケーションを変更します。デフォルトのターゲット
アプリケーションはMicrosoft Wordです。
[**送信先**]、または[**外部ファイル**]リストから必要な出力フォーマット、または
アプリケーションを選択します。
各種タブをクリックし、ご希望のオプションを選択します。
選択したフォーマット/アプリケーションに利用できないオプションは、淡色
表示になっています。

詳細に関しては、「**文書のフォーマット**」をご参照ください。

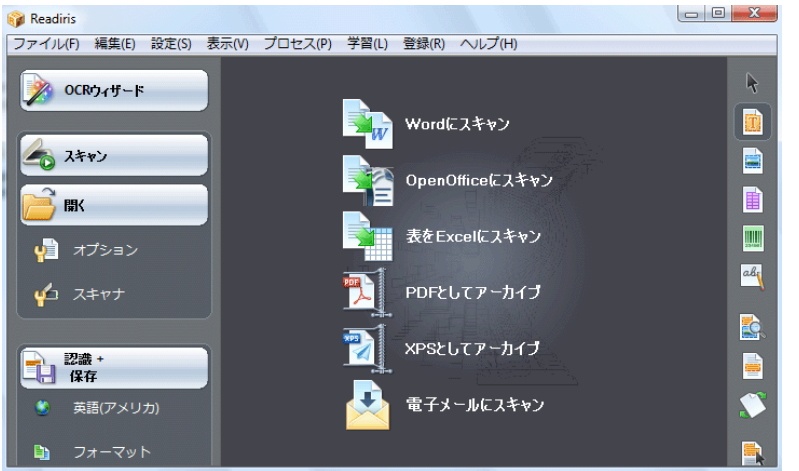
[**OK**]をクリックして、設定を保存します。
[**次へ**]をクリックして次のステップに進みます。

ステップ5
[**移動**]をクリックして文書を開いて/スキャンして認識します。

ユーザインターフェース

Readiris
インターフェースを試すには、**Readiris**スタートアップ画面をクリックし、
OCRウィザードが起動したら[**キャンセル**]をクリックします。

空白のReadirisが表示されます。



Readiris インターフェースは次から構成されています。

- **SmartTasks** (中央)

  SmartTasksは、もっとも頻繁に使用するReadiris
  機能をボタン1つで使用できるようにする所定のコマンドです。

  SmartTasksを利用し、スキャン、認識、文書をターゲット
  アプリケーションに送信したり、ご希望のフォーマットに出力します。

  SmartTasksは、デフォルト設定を適用しますが、右クリックして、特定
  のニーズに合わせて簡単に構成することもできます。

- メインツールバー(左側のツールバー)

  メインツールバーのコマンドとオプションを使用し、文書を手動でスキャ
  ン、認識します。
  その手順は、**OCRウィザード**で指示されます。

- 画像ツールバー(右側のツールバー)

画像ツールバーボタンを利用し、Readiris
インターフェースの文書を編集します。
各ボタンをポイントすると、ツールチップが表示されます。

Readirisで文書が開いたり、スキャンされると、インターフェースに3つのメインゾーンが追加されます。

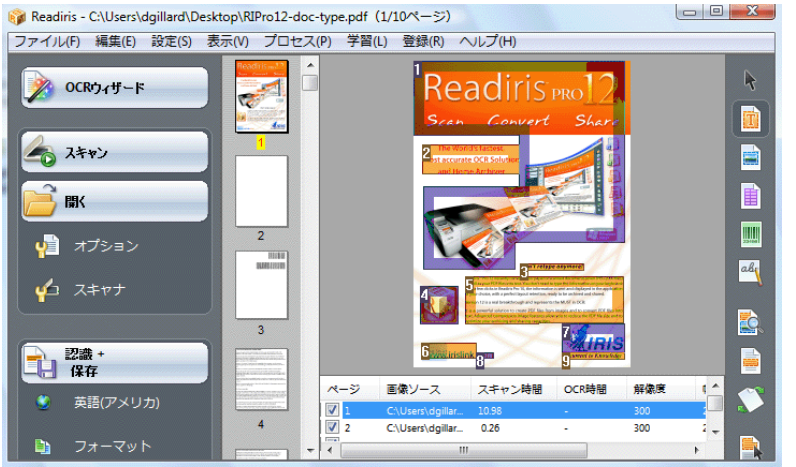
- ページツールバー(メインツールバーの右側)

  ページツールバーは、ページサムネイルを表示し、これをポイントすると設定情報が表示されます。

- 画像ウィンドウ (中央)

- 文書パネル(一番下)

  文書パネルは、Readirisで開いている文書に関する統計情報を表示します。スキャン時間、OCR時間、解像度、文書の幅と高さなどの情報です。



ユーザインターフェース言語の変更

Readirisのユーザインターフェースは、様々な言語でご利用いただけます。

ユーザインターフェース言語の変更の仕方:

- [設定]メニューから[ユーザインターフェース言語]をクリックします。
- 「言語」リストから必要な言語を選択し、[**OK**]をクリックして確認します。

注記:
間違った言語を選択した場合は、**Ctrl+U**をクリックします。「言語」ダイアログボックスが表示され、リストから別の言語を選択することができます。

第 4 章
READIRIS SMARTTASKS

READIRIS SMARTTASKSの利用

Readirisを起動するとき、Readirisスタートアップ画面をクリックし、OCRウィザードが起動したら[**キャンセル**]をクリックします。 Readiris SmartTasksが表示されます。

SmartTasksは、もっとも頻繁に使用するReadiris機能をボタン1つで使用できるようにする所定のコマンドです。

各種SmartTasksボタンで下記が行えます。

1. 文書をスキャンし、認識して、直接**Word**に送信してテキスト処理。

Microsoft Wordが、デフォルトのターゲット
アプリケーションです。
その他の利用可能なアプリケーションに関する詳細は、「**テキスト文書のフォーマット**」をご参照ください。

2.
文書をスキャンし、認識して、直接OpenOfficeに送信してテキスト処理。

OpenOffice.org Writerが、デフォルトのターゲット
アプリケーションです。
その他の利用可能なアプリケーションに関する詳細は、「**テキスト文書のフォーマット**」をご参照ください。

3. 表をスキャンし、認識して、直接Excelや他のスプレッドシートに送信。

Microsoft Excelが、デフォルトのターゲット
アプリケーションです。
詳細に関しては、「**表からなる文書のフォーマット**」をご参照ください。

4. 文書をスキャンし、認識して、PDFファイルとしてアーカイブ。

Adobe Acrobat PDF 画像-
テキストがデフォルトの出力フォーマットです。
その他の利用可能なフォーマットに関する詳細は、「**PDF文書の作成**」をご参照ください。

5. 文書をスキャンし、認識して、XPSファイルとしてアーカイブ。

XPS 画像-テキストがデフォルトの出力フォーマットです。
その他の利用可能なフォーマットに関する詳細は、「**XPS文書の作成**」をご参照ください。

6. 文書をスキャンし、認識して、直接電子メール。

デフォルトでは文書は、デフォルトの電子メールアプリケーションを介してPDF 画像-テキストとして送信されます。
その他の利用可能なフォーマットに関する詳細は、「**文書のフォーマット**」をご参照ください。

Readiris
を初めてお使いになるときには、SmartTasksを構成しなければなりません。

SmartTasksの構成の仕方:

- ご希望のSmartTasksを右クリックします。
- 画像ソースとして[**スキャナ**]、または[**画像ファイル**]を選択します。
    - [**スキャナ**]を選択した場合は、SmartTaskをクリックすると直ちにReadiris はスキャナを開始します。スキャンされた文書はインターフェースに表示され、処理、保存されます。

      SmartTasksが作動するためには、スキャナを正しく構成しなければなりません。

      スキャナの構成の仕方:

      - メインツールバーの[**スキャナ**]をクリックします。
      - [**スキャナモデル**]をクリックし、リストからお使いのスキャナを選択します。

        お使いのスキャナがリストに掲載されていない場合は、「**Twainその他のモデル**」を選択します。
      - 該当する場合は、[構成]をクリックし、Twainソースを選択します。
      - [**OK**]をクリックして、設定を保存します。

スキャナ設定と用紙文書のスキャンに関する詳細は、「**用紙文書のスキャン**」をご参照ください。

- [**画像ファイル**]を選択し、SmartTaskをクリックすると、Readiris
  が「**入力**」ダイアログボックスを表示しますので、処理したい画像ファイルをそこで選択できます。

  画像ファイルを開く詳細に関しては、「**画像ファイルを開く**」をご参照ください。

- [**構成**]をクリックし、出力**フォーマット**とその**オプション**を変更します。

  利用できる出力フォーマットとオプションは、選択したSmartTaskにより異なることにご留意ください。

  その他の利用可能なフォーマットとオプションに関する詳細は、「**文書のフォーマット**」をご参照ください。

- SmartTasksの構成が終了したら、メインツールバーのボタンを利用し、**言語**設定と画像の改善**オプション**、およびまだ必要な場合は、**スキャナの設定**を特定します。

  上述の設定に関する詳細は、「**オプションの選択**」、「**用紙文書のスキャン**」、「**文書の言語を選択**」をご参照ください。

- これで、SmartTasksをクリックして使用できます。

Readirisは自動的に全認識プロセスを実行します。

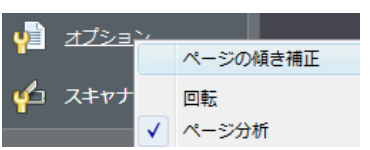
第 5 章
文書のスキャン

オプションの選択

用紙文書をスキャンしたり、画像ファイルを開く前に、画像の改善オプション数種を選択することができます。
これらのオプションを有効にすると、文書を開いたり、スキャンしているときに適用されます。

操作

- メインツールバーの[**オプション**]をクリックし、改善オプション数種を選択します。



  - [**ページの傾き補正**]をクリックして、斜めにスキャンしたページをまっすぐにします。
  - [**回転**]をポイントし、Readirisで**自動的**にページを回転するか、**90°左**、**90°右**、**180°**回転するかを決定します。

    これ2つのオプションは、スキャン処理を若干遅くしますのでご注意ください。 必要な場合にのみ選択してください。

  - デフォルトでは、[**ページ分析**]が有効になっています。

このため、スキャンした画像、開いた画像は、自動的にウィンドウで分割されます。

画像ツールバーのウィンドウイングを利用して、ページ分析結果を変更したり、手動で文書をウィンドウイングできます。
詳細に関しては、「**文書を手動でウィンドウイング**」をご参照ください。

- 設定すべて(**スキャナ設定**、**オプション**)を特定したら、[**スキャン**]をクリックして文書をスキャンするか、[**開く**]ボタンをクリックして画像ファイルを開きます。

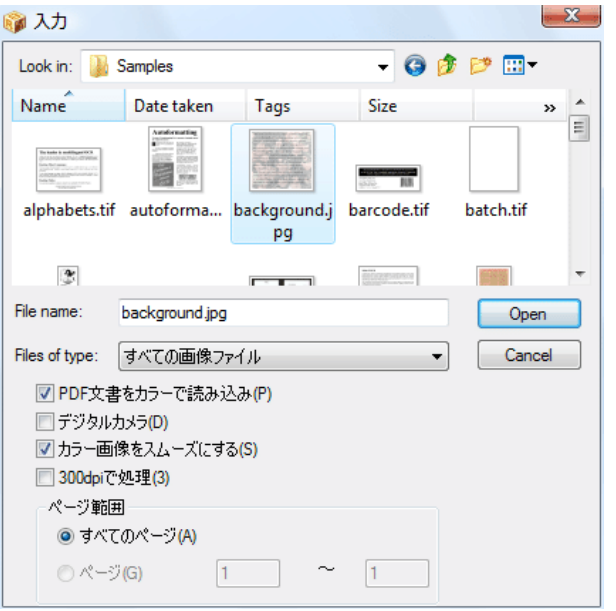
上述のオプションは、[**設定**]メニューからも利用できます。

画像ファイルを開く

Readirisでは、ご自分のスキャナでスキャンした用紙文書、様々なフォーマットのすでに現存する画像ファイルのいずれをも処理することができます。

既存の画像ファイルの開き方:

- [**開く**]ボタンをクリックして画像を検索します。



ヒント:
画像ファイルをReadiris画像ウィンドウに**ドラッグ**して開くこともできます。

ヒント:
開きたい画像ファイルを右クリックし、[**プログラムから開く**]をポイントし、[**IOCR application**]をクリックします。
Readirisソフトウェアが開き、画像を表示します。

ヒント:
複数ページの画像ファイル(TIFF画像やDCXファックス)やPDF文書を読み込む場合、ページ範囲を特定することができます。(文書の特定の章だけが必要な場合など)
ページ範囲を指定するには、メインツールバーの[**開く**]をクリックします。
[**ページ範囲**]で[**ページ**]を選択し、読み込みたいページを指定します。
「**複数の画像ファイルを開いて認識**」もご参照ください。

ヒント:
読み込みプロセスを速めるには、[**開く**]ボタンをクリックし、PDF文書の処理をする場合は[**PDF文書をカラーで読み込み**]をオフにします。

- Readirisがサポートするグラフィック フォーマット:

グラフィックのフォーマット	ファイル拡張子
Adobe Acrobat PDF	*.pdf
DCXファックス	*.dcx
DjVu画像	*.djv、*.djvu
JPEG画像	*.jpg、*.jpeg
JPEG 2000画像	*.j2c、*.j2k、*.jp2
Portable Network Graphics	*.png
(圧縮した複数ページ)TIFF画像 非圧縮、LZW、PackBits、Group 3、Group 4およびJPEG圧縮	*.tif、*.tiff
Windowsビットマップ	*.bmp
ZSoftペイントブラシ画像	*.pcx

- ご希望の画像ファイルを選択し、[**開く**]をクリックします。

注記:
[**入力**]ダイアログボックスのオプションも文書のスキャンに適用されます。これについては「**用紙文書のスキャン**」をご参照ください。

用紙文書のスキャン

Readirisでは、ご自分のスキャナでスキャンした用紙文書、様々なフォーマットのすでに現存する画像ファイルのいずれをも処理することができます。

用紙文書を処理する場合、Readirisは、あなたが[**スキャン**]をクリックすると同時にお使いのスキャナを開始し、スキャンした文書をインターフェースに表示します。

文書のスキャンの仕方:

- [**スキャナ**]をクリックし、スキャナ設定を設定します。

[**スキャナ**]ダイアログボックスのオプションの一部は、[**開く**]ダイアログボックスにもあります。

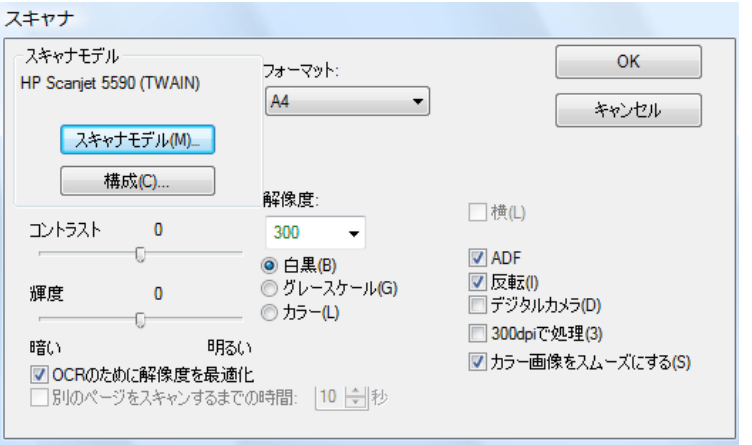
- 正しいスキャナモデルを選択します。

お使いのスキャナがリストに掲載されていない場合は、「**Twain その他のモデル**」を選択し、[**OK**]をクリックします。

- Readirisは、フラットベッドやシートフィードのスキャナほぼ全種、「一体型」デバイス、(「MFP」-多機能周辺機器)、デジタルカメラ、現在利用可能なスキャナ標準をサポートします。 Readirisは、Twain順守です。

該当する場合にのみ、[**構成**]は利用できます。
[構成]をクリックし、Twain ソースを選択します。

- スキャナ設定を選択します。



フォーマットと解像度

Readirisは、様々な用紙フォーマットと解像度をサポートします。

スキャン解像度300 dpiをご使用になるようお勧めします。

、アジア諸語のテキスト、非常に小さな印字を認識する場合は、解像度400 dpiを使用します。

カラーモード

Readirisは、文書をスキャンし、**カラー**、**白黒**、**グレースケール**の画像ファイルを開く ことができます。

コントラスト - 輝度

最適のスキャン結果を得られるように、スライダを使用して適切な輝度とコントラストの設定を定めます。

OCRのために解像度を最適化

[**OCRのために解像度を最適化**]を選択し、詳細すぎる(600 dpi以上)スキャン画像の解像度を補正します。
Readirisは、解像度を下げます。

このオプションは、詳細が少なすぎるスキャン画像の解像度を向上させることは一切ありません。

複数ページの文書のスキャン

複数ページの文書をスキャンし、文書フィーダのあるスキャナを利用している場合は、[**ADF**](自動文書フィーダ)オプションを選択します。
スキャナのフィーダにスキャンしたいページを入れ、スキャンを開始します。

フラットベッドのスキャナを使用する場合は、[**インターバルスキャン**]を選択します。
[**別のページをスキャンするまでの時間**]を選択し、何秒後にReadirisに別のページをスキャンさせるかを指定します。

詳細に関しては、「**複数ページの文書をスキャンして認識**」をご参照ください。

重要:
選択したスキャナに利用できないオプションは、淡色表示になっています。

横長の画像をスキャン

横長の画像をスキャンする場合は、[**横**]オプションを選択します。

自動露出

スキャナによっては、[**自動露出**]オプションがデフォルトで選択されている場合もあります。
このオプションは、コントラストと輝度の設定を自動的に調節します。

反転画像をスキャン

反転した白黒画像をスキャンする場合は、[**反転**]オプションを選択します。

デジタルカメラの使用

カメラをスキャンソースとして使用する場合は、[**デジタルカメラ**]を選択します。

Readirisは、特殊な認識ルーティンを活用して、デジタルカメラの画像を処理します。

スキャンソースとしてデジタルカメラを使用する場合のヒント:

- 白い文書を撮影してカメラを較正します。
- 必ず最高の解像度を選択します。
- アップで撮影する場合は、カメラのマクロモードを有効にします。
- デジタルズームではなく、光学ズームのみを使用します。
- 文書の真上でカメラを構えます。
  文書を斜めに撮影しないようにします。
- 安定した画像が取れるようにします。
  必要な場合は、三脚を使用します。
- 光沢のある用紙を撮影する場合は、フラッシュをオフにします。
- 圧縮したカメラ画像を開かないようにします。
- 環境(日光、照明、ネオン)に合わせてReadirisの輝度とコントラストの設定を調整します。
- カラーモードとしてカラー、またはグレースケールを選択します。

300dpiで処理

正しくなかったり、不明の解像度の画像を処理する場合は、[**300dpiで処理**]オプションを選択します。画像は、300 dpiの画像であるかのように処理されます。

デジタルカメラの画像は、ほとんどいつも不明です。

カラー画像をスムーズにする

スキャナによっては、カラーやグレースケールの画像をうまく 認識するには、画像をスムーズにすることが必要なため、このオプションがデフォルトで選択されています。

- 設定すべて(**スキャナ設定**、**オプション**)を特定したら、[**スキャン**]をクリックして文書をスキャンします。

  注記: 線の傾きに注意します。
  線の傾きが0.5°以上の場合は、OCRエラーが発生しやすくなります。

第 6 章
スキャンした文書の調整

非常に薄かったり濃いグレースケールやカラーの画像を開いたり、スキャンする場合、納得のいく OCR結果を得るためには、認識を行う前にこれらの画像を調整しなければならない場合があります。

画像の調整の仕方:

- カラー-グレースケール文書を開くかスキャンします。

  スキャナ設定が正しいことを確認します。

  重要:
  スキャナ設定と**調節設定**は同じに見えますが、両設定はスキャン/認識プロセスの別々の段階で適用されます。

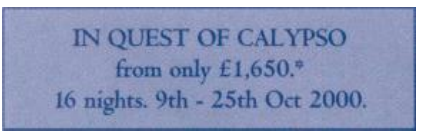
- [**プロセス**]メニューから[**画像を調節**]をクリックします。

  Readiris
  は、インテリジェントな**2値化**ルーティンを活用し、カラー-グレースケール画像を白黒画像に変換し、その画像でOCRを実行します。

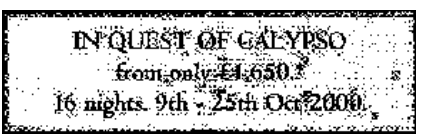
o 画像を均質にするには、[**カラー画像をスムーズにする**]を選択します。

  このオプションは、強度の違いをスムーズにすることにより、グレースケールとカラー画像をより一層均質にします。
  その結果、前景(テキスト)と背景(アートワーク)のコントラストをより鮮明にします。

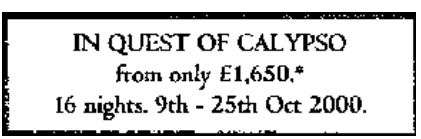
注記:
カラー背景からテキストを分離するには、スムージングしか方法がない場合も時々あります。



(元の画像)



(2値化した白黒画像)



(スムーズにした画像)

- スライダを利用して[**輝度**]を調節します。

[**輝度**]設定は、画像の全体的な明るさを定めます。
テキストが判読しにくい場合は、この設定を利用して画像を暗くしたり、明るくしたりします。

例1: 暗い画像を明るくし、ページの背景を削除します。

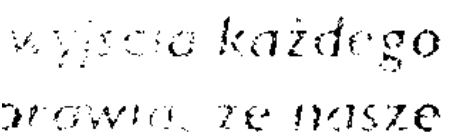
(カラー画像)

(2値化画像。
デフォルトの2値化設定により、真っ黒な画像になりました)

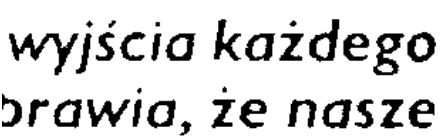
(画像を明るくすることにより、満足できる認識結果になります
)

例2:
テキストが薄すぎて、2値化画像に表示されない場合は、画像を暗くします。

(カラー画像)



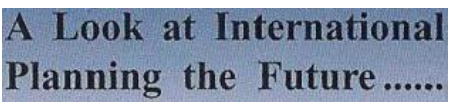
(2値化画像。
デフォルトの輝度設定により、切れ切れの文字にになりました)



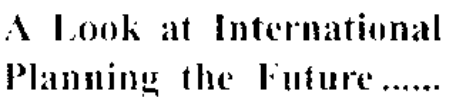
(画像を暗くすることにより、満足できる認識結果になります)

- スライダを利用して[**コントラスト**]を調節します。

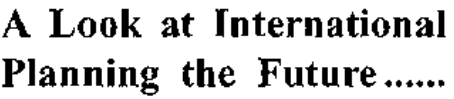
[**コントラスト**]設定は、画像の暗い部分と明るい部分のコントラストを定めます。
この設定を利用し、文字の形状がカラー背景に目立つようにします。



(カラー画像)



(デフォルトのコントラスト設定により、文字が乱れてしまいました)



(コントラスト設定を上げることにより、満足できる認識結果になります)

- スライダを利用して[**汚れの補正**]オプションを調節します。

汚れの補正は、白黒画像から小さな斑点を取り除きます。

- [**適用**]をクリックして、結果をプレビューします。
- 結果に満足できる場合は、[**OK**]をクリックします。
  満足できない場合は、設定を再度変更します。

- [**認識+保存**]をクリックして、文書を認識します。

第 7 章
文書を画像ファイルとして保存

スキャンした用紙文書は、すぐにOCRを実行する必要はありません。
画像ファイルとして保存できます。

画像ファイルとして保存する仕方:

- 文書をスキャンします。
- **[ファイル]**メニューの**[ページ全体を画像として保存]**、または[全てのページを画像として保存]をクリックします。
- 後から保存した画像ファイルを開き、認識を実行します。

グラフィックのみ保存

文書のテキストは保存しないで、グラフィック
ウィンドウだけを保存することもできます。

グラフィックのみ保存する仕方:

- 文書をスキャン、または開きます。
- [ファイル]メニューから[グラフィックを保存]をクリックします。
- 文書のグラフィックすべてが単一のファイルに保存されます。

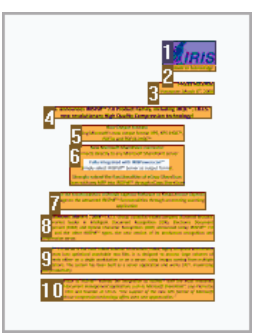
第 8 章
文書のウィンドウイング

文書を自動的にウィンドウイング

文書をスキャン、または開く とき、Readirisは自動的に[ページ分析]を適用し、文書を別々のウィンドウに分割します。

デフォルトでは、[**ページ分析**]が選択されています。
自動ページ分析を実行したくない場合は、[**オプション**]をクリックし、[**ページ分析**]の選択を解除します。

ページ分析の結果は、自動ページ分析後に手動で変更できます。
詳細に関しては、「**文書を手動でウィンドウイング**」をご参照ください。



ウィンドウのタイプ

Readirisは、ウィンドウタイプ5種を使用しています。
テキストブロック、**グラフィックゾーン**、**表**、**バーコードゾーン**、**手書きゾーン**です。

ページ分析は、自動的にテキスト、グラフィック、表ゾーンを検出します。
バーコードゾーンと 手書きゾーンは、手動で描かなければなりません。

> 詳細に関しては、「**文書を手動でウィンドウイング**」をご参照ください。

各ウィンドウタイプには、それぞれのカラーコードが割り当てられています。
テキストウィンドウはオレンジ色、グラフィックウィンドウは紫色、表ウィンドウはピンクです。バーコードゾーンは緑色、手書きゾーンは青色です。

ウィンドウは上下、左右にソートされています。
番号は、ウィンドウのソート順です。
ただし、ソート順とウィンドウタイプは、変更することができます。
詳細に関しては、「 **文書を手動でウィンドウイング**」をご参照ください。

ページ枠のテキストを無視

スキャナが実際の画像の周りに黒枠を生成した場合は、ページ分析はノイズしかないゾーンを検出しがちです。

これを避けるには、[**設定**]メニューの[**枠上のウィンドウを検出しない**]をクリックし、文書を再度スキャンします。

外部ゾーンを無視

自動的に文書をウィンドウイングする代わりに、[**外部ゾーンを無視**]を使用することもできます。
この機能は、ページ上の特定の部分1箇所だけをOCRするときに便利です。

イメージツールバーの[**外部ゾーンを無視**]を選択します。

Readirisに認識させたい部分の周りにフレームを描きます。

ページ上で選択した部分が、自動的に分析されます。

文書のすべてのページで、ページ分析から同じ外部ゾーンを除外したいか尋ねられます。

文書を手動でウィンドウイング

ページ分析による文書を自動でウィンドウイングするほかにも、Readirisは文書を手動でウィンドウイングできます。

手動ウィンドウイングは、自動ページ分析結果を修正しなければならないときに便利です。ウィンドウイング テンプレートを作成するときにも便利です。

ウィンドウイング
テンプレートの詳細に関しては、「**ウィンドウイング テンプレートの使用**」をご参照ください。

バーコードゾーンと手書きゾーンは、いつも手動で描かなければなりませんのでご留意ください。

操作

- 文書を手動でウィンドウイングするには、[オプション]をクリックし、[ページ分析]の選択を解除します。
- [スキャン]、[開く]をクリックし、文書を開いたりスキャンします。

- イメージツールバーで正しい**ウィンドウタイプ**をクリックして選択します。

  Readirisは、ウィンドウタイプ5種を使用しています。
  テキストブロック、**グラフィックゾーン**、**表**、**バーコードゾーン**、**手書きゾーン**です。

- ウィンドウイングしたいテキストブロック、グラフィック、表、バーコード、手書きソーンの周りにフレームを描きます。

  バーコードと手書きソーンの認識に関する詳細は、「**バーコードの認識**」と「**手書きテキストの認識**」をご参照ください。

- 文書のウィンドウイングが終了したら、[**認識＋保存**]をクリックしてOCRを実行します。

ウィンドウの並べ替え

ウィンドウのソート順を変更するには、画像ツールバーの[**ソート**]をクリックし、ウィンドウを1つずつ必要な順にクリックします。

終了したら、[**認識＋保存**]をクリックしてOCRを実行します。

クリックしなかったウィンドウは、認識から除外されます。

多角形を描く

文書の手動ウィンドウイングは、長方形に限られているわけではありません。長方形を結合して多角形を作成することもできます。
同じタイプのウィンドウが交差するときにはいつでも、自動的に多角形になります。

自動ページ分析

現在のページが、手動でウィンドウイングするには複雑すぎる場合は、画像ツールバーの[**ページの分析**]をクリックしてページを自動的にウィンドウイングします。

バーコードゾーンと手書きゾーンは、いつも手動で描かなければなりませんのでご留意ください。

ウィンドウタイプの変更

ウィンドウタイプを変更する場合は、[**ウィンドウ**]、[**タイプ**]の順にポイントし、必要なウィンドウタイプをクリックします。

複数のウィンドウのウィンドウタイプを同時に変更することもできます。

- 画像ツールバーの**ポインタ**ボタンをクリックします。

- **Shift**を押しながら、複数のウィンドウを選択します。
- 選択したウィンドウのいずれかを右クリックし、[**ウィンドウ**]、[**タイプ**]の順にポイントし、必要なウィンドウタイプをクリックします。

ウィンドウサイズの修正

修正したいウィンドウをクリックします。

(ウィンドウの側面、隅にある)マーカーの上にマウスポインタを置きます。

マーカーをクリックしてマウスをドラッグし、ウィンドウサイズを修正します。

ウィンドウの移動

移動したいウィンドウを選択します。

ウィンドウの内側をクリックし、マウスをドラッグしてウィンドウの位置を変更します。

特定ウィンドウの認識

認識したいページを右クリックし、[**テキストとしてコピー**]を選択します。

結果は、本文としてクリップボードに送信されます。
これは、手書きテキストにも利用できます。

グラフィック ウィンドウとバーコード
ウィンドウもクリップボードにコピーできます。

ウィンドウの削除

削除したいウィンドウを選択するか、[**編集**]メニューの[**すべて選択**]をクリックし、すべてのウィンドウを選択します。

[**編集**]メニューの[**切り取り**]、[**クリア**]を選択し、ウィンドウを切り取ったり、削除します。

または、

選択したウィンドウを右クリックし、[**ウィンドウ**]をポイントしてから[**削除**]をクリックします。

小さいウィンドウの削除

たとえばファックスなどの文書によっては、ページに斑点が「散在」していることがよくあり、Readiris
がテキストを含んでいない余計なウィンドウを作ってしまうことになります。

小さいウィンドウすべてを消去するには、[**編集**]メニューの[**小さいウィンドウを削除**]をクリックします。

このオプションは、1.27cm未満のウィンドウすべてを消去し、残りのゾーンを再ソートします。

ウィンドウイング テンプレートの使用

類似のページレイアウトの文書を多数OCRする場合は、自動ページ分析の代わりに**ウィンドウイング テンプレート**を使用すると便利です。
これを使用すると、同じウィンドウ構造がすべてのスキャンした、または開いている文書に適用されますから、処理にかかる時間が短縮されます。

操作

- イメージツールバーのボタンを使用して、文書の最初のページを手動でウィンドウイングします。

  詳細に関しては、「**文書を手動でウィンドウイング**」をご参照ください。

- [ファイル]メニューから[レイアウトの保存]をクリックします。

- メインツールバーの[**開く**]、[**スキャン**]をクリックして、文書の他のページを開くかスキャンします。
- [**ファイル**]メニューから[**レイアウトの読み込み**]をクリックします。
- 保存したレイアウトファイルを選択します。
- レイアウトを開いているページ、またはスキャンしたページ**すべて**に適用する場合は、「**レイアウトファイル**」ダイアログボックスの[**すべてのページにこのレイアウトを適用**]を選択します。
- [**開く**]をクリックしてレイアウトファイルを読み込みます。

  Readirisに文書を追加するときは、デフォルトでページ分析が有効になっているためレイアウトファイルを再度読み込まなければなりません。

外部ゾーンを無視

ウィンドウイング
テンプレートの代わりに、[**外部ゾーンを無視**オプションを使用できます。このオプションでは、ページ上でOCRが必要な場所を1箇所特定することができます。OCR部分以外のデータはすべて、認識から除外されます。

操作

- イメージツールバーの[**外部ゾーンを無視**]を選択します。

- Readirisに認識させたい部分の周りにフレームを描きます。

  ページ上で選択した部分が、自動的に分析されます。

文書のすべてのページで同じ外部ゾーンを無視するかどうか尋ねられます。

この機能をキャンセルする場合は、画像ツールバーの[**ページの分析**]をクリックし、**ページ分析**を再実行します。

- [**認識+保存**]をクリックして、OCRを実行します。

第 9 章
文書を認識

はじめに

文書を 認識するために、Readirisは認識段階で言語学を駆使しています。
そのためReadirisは、あらゆる文書のテキスト、表、グラフィック、バーコード、手書きのテキストを認識できます。
Readirisは、複雑な段組みの文書や品質の悪い文書、ファックス、ドットマトリックスの印刷物、スキャンやコピーの仕方が悪かったために、フォントが薄かったり濃かったりする文書などにも対応できます。

Readirisは、128ヶ国語の言語に対応しています。
中欧、バルト海諸国、キリル文字の言語、およびギリシャ語やトルコ語を含むアメリカとヨーロッパの言語すべて対応しています。
オプションとして、Readirisはアラビア語やペルシア語（ファルシ）、ヘブライ語、アジア4カ国の言語
(日本語、韓国語、簡体字中国語、および繁体字中国語)も読むことができます。Readirisは、混合アルファベットにも対応できます。
ソフトウェアは、ギリシャ語やキリル文字の言語、アラビア語、ヘブライ語、アジア諸語の文書に出てくる「西洋文字」の言葉を検知します。別の言語に転記することができない固有名詞、ブランド名などの多くが、西洋の文字や記号で表記されていることがあります。

Readirisは、最新鋭の認識テクノロジーを基にしています。
フォントに依存しないテキスト認識は、自習テクニックで補完されています。システムは、文脈や言語分析を通じ、新しい文字や単語を学ぶことができます。
つまり、認識システムのOCR精度が、使用を重ねるにつれ向上することになります。

その他にも、Readiris には(オプションの)ユーザ確認機能があります。
このユーザ確認機能(**インタラクティブ学習**)をオンにすると、認識システムがはっきりわからない文字にフラグを立てるだけではなく、システムの精度を向上させることができます。
あなたが確認したすべてのソリューションを認識中に一時的に記憶し、システムのスピードと自信を向上し、使用を重ねるにつれ、システムがさらにインテリジェントを増します。
このパワフルな学習ツールは、数学記号や絵記号などの特殊文字をReadirisに認識させる訓練となり、ゆがんだフォントを処理させることもできます。

インタラクティブ学習の結果は、**フォント辞書**に恒久的に保存し、将来の使用に備えることができます。

文書の言語を選択

Readirisは、128ヶ国語でOCRを実行できます。
Readirisは、中欧、バルト海諸国、キリル文字の言語、およびギリシャ語やトルコ語を含むアメリカとヨーロッパの言語すべてに対応しています。

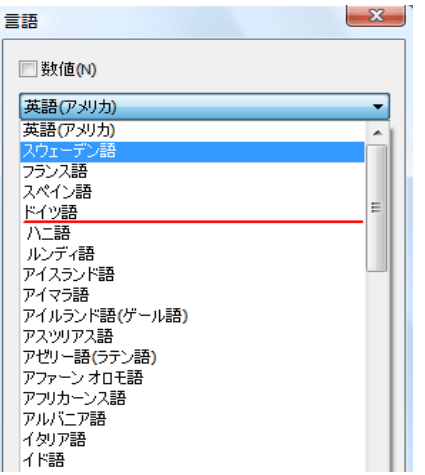
Readiris Pro AsianとReadiris Corporate Asianは、さらに日本語、簡体字中国語、繁体字中国語および韓国語の文書を認識します。

Readiris Pro Middle-EastとReadiris Corporate Middle-Eastは、さらにアラビア語やペルシア語(ファルシ)、ヘブライ語の文書を認識します。

Readirisが文書を認識するためには、文書の言語を指定しなければなりません。

文書の言語を指定する仕方:

メインツールバーの[**グローブ**]をクリックし、ドロップダウンリストからご希望の言語を選択します。



最近選択した言語の5つが言語リストの上位に移動します。

重要:
アジア諸語、ヘブライ語、アラビア語の文書を扱う場合は、ページ分析を実行する*前*に文書の言語を選択します。
特殊なページ分析ルーティンが、これらの文書に使用されます。

認識を**数値**に限定し、表やデータを最適に認識することもできます。 Readirisは、数字0～9と下記の記号のみを認識します。

+ プラス記号	* 星印	/ スラッシュ	% パーセント記号	, コンマ
. ピリオド	(左丸括弧s	) 右丸括弧s	- ハイフン	= イコール記号
$ ドル記号	£ ポンド記号	€ ユーロ記号	¥ 円記号	

数値モードをアクティベートするには、[**言語**]ダイアログボックスの[**数値**]を選択します。

混合言語の文書を認識

Readirisは、混合文字セットも有効にできます。
これを行うことにより、Readirisは文章の途中で自動的に言語を切り替え、「非インドヨーロッパ」言語の中に出てくる英語の単語(固有名詞など)を認識できます。

メインツールバーの[**グローブ**]をクリックし、ドロップダウンリストから必要な言語の組み合わせを選択します。

注記
アジア諸語、ヘブライ語の文書を処理する場合は、混合文字セットは、自動的に使用されます。

速度と精度

OCR速度とOCR精度の適切な均衡を図ります。

ラテン、キリル文字、ギリシャ語のアルファベットに、このバランスが利用できます。

ヒント: 画像の品質がよくない場合は、精度を重視します。

文書の特徴を定義

認識プロセスでは、文書の言語に続き、**フォントの種類**や**文字ピッチ**などその他の文書の特徴が大切な役割を果たします。

フォントの種類

Readirisは、「 標準」 と「 ドットマトリックス」 で印刷した文書を識別します。
ドットマトリックス記号(タイプ9ピン)は、個別の点から形成されています。

ドットマトリックス文書を認識するには、特殊な分割と 認識テクニックが必要であり 、それをアクティ ベートしなければなりません。

Far out in the uncharted back

フォントの種類の選択の仕方:

- [設定]メニューから[フォントの種類]をポイントします。
- フォントの種類は、デフォルトでは［自動］に設定されています。

  自動の場合は、Readiris は「25ピン」、「NLQ」(Near Letter Quality) ドットマトリックス、または他の「標準」印刷かを認識します。

- ドットマトリックス文書だけを認識する場合は、［ドットマトリックス］をクリックします。

  Readirisは、いわゆる「ドラフト」、または「9ピン」ドットマトリックス印字の文書を認識します。

文字ピッチ

文字ピッチとは、印字面1インチ当り の文字数こと です。
文字ピッチは、すべての文字幅が同一の[等幅、または文字によって幅が異なる[プロポーショ ナル]に設定できます。

文字ピッチの選択の仕方:

- [設定]メニューから[文字ピッチ]をポイントします。

- 文字ピッチは、デフォルトでは［**自動**］に設定されています。
- 印字面の文字すべてを同じ幅にそろえる場合は、[**等幅**]をクリックします。
  旧式のタイプライターの文書は、よく等幅フォントになっています。
- 印字面の文字の幅を同一にしない場合は、[**プロポーショナル**]をクリックします。
  実質的には、新聞、雑誌、書籍のフォントすべてが、プロポーショナルになっています。

重要:
これらの文書の特徴は、アジア諸語、ヘブライ語、アラビア語の文書には適用されません。

インタラクティブ学習の利用

Readirisは、インタラクティブ学習機能を搭載しています。
インタラクティブ学習により、フォント、文字の形状に関し認識システムを訓練し、必要に応じてOCR結果を修正できます。
インタラクティブ学習の際に、認識システムがはっきりとは認識できない文字は、その語源と解決法の案とともにプレビューウインドウに表示されます。

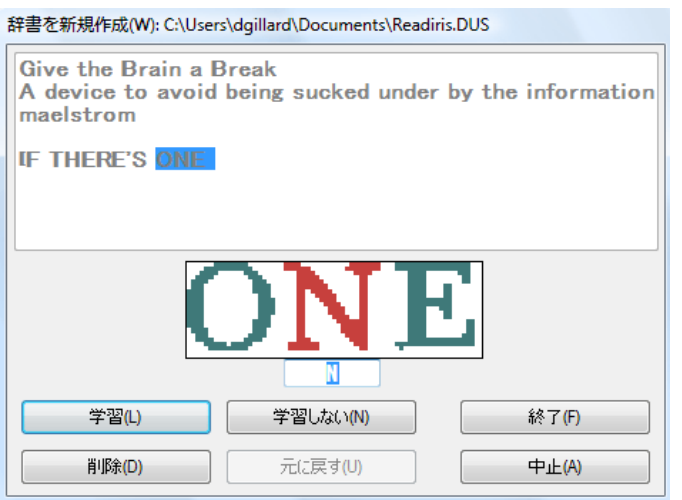
インタラクティブ学習は、認識システムの精度を大幅に強化し、特に歪んだり、損傷した用紙を認識する場合に便利です。
インタラクティブ学習は、Readirisが最初は認識できない特殊記号に関する

訓練にも使用できます。その例として、数学記号、科学記号、絵記号などが挙げられます。

インタラクティブ学習を有効にする仕方:

- [学習]メニューから[インタラクティブ学習]をクリックします。
- 必要な設定を定め、[認識+保存]をクリックして、文書を認識します。
- 認識の最後に、Readirisはインタラクティブ学習フェーズに入ります。

認識システムがはっきりとは認識できない文字は、表示されます。



結果が正しい場合:

- [学習]をクリックして、結果を確かなものとして保存します。

学習結果は、認識の間、一時的にコンピュータメモリに保存されます。
Readirisは、残りの文書をOCRしているときに、学習した文字をそれ以降は表示しません。

新規文書がOCRされたときに、学習結果は消去されます。

学習結果を恒久的に保存するには、フォント辞書を使用します。詳細に関しては、「**フォント辞書の使用**」をご参照ください。

- [終了]をクリックして、ソフトウェアが提供したソリューションすべてを保存します。

結果が正しくない場合:

- 正しい文字を入力し、[学習]をクリックします。

または、

- [学習しない]をクリックして、結果を不確かとして保存します。

  学習した場合、他の文字と混同される可能性がある損傷した文字にこのコマンドを使用します。例えば、数字の1と小文字の l は、多くのフォントで同一です。

- [削除]をクリックして、出力から文字を削除します。

  文書ノイズが出力ファイルに表示されるのを防止するのにこのボタンを使用します。

- [元に戻す]をクリックして、間違いを訂正します。

  Readirisは、前回32個の操作を記憶しています。

- [中止]をクリックしてインタラクティブ学習を中止します。

すべての学習結果が削除されます。
次回[認識+保存]をクリックすると、インタラクティブ学習は再開されます。

フォント辞書の使用

種類、フォントの品質、印刷の品質が同じ多数の書類をスキャンする場合、毎回学習プロセスを繰り返すのを避けたい場合があります。
そのため、フォント辞書の使用が便利です。
フォント辞書は、インタラクティブ学習の際に学んだフォント情報を含有しており、認識結果を大幅に増大させることができます。

フォント辞書は、最高500個の形状に限られていますのでご注意ください。
特定のアプリケーション用に、別個の辞書を作成するようお勧めします。

新規フォント辞書の作成の仕方:

- [学習]メニューから[新規フォント辞書]をクリックします。

  デフォルトでは、Readirisは辞書**Readiris.dus**を開きます。
  ファイル名を変更して[保存]をクリックし、新規辞書として保存します。

- [学習]メニューの[インタラクティブ学習]をクリックしてアクティベートします。

- [認識+保存]をクリックして、文書を認識します。

- Readirisは、インタラクティブ学習フェーズに入ります。ダイアログボックスのボタンを使用して、フォント辞書に文字を保存します。

既存のフォント辞書の使い方:

- [**学習**]メニューから[**フォント辞書**]をクリックします。
- ご希望の辞書を選択し、[**開く**]をクリックします。
- [**学習**]メニューから[**フォント辞書に追加**]、または[**フォント辞書を閲覧**]をクリックします。

  [**フォント辞書に追加**]を選択する場合は、必ず[**インタラクティブ学習**]をオンにします。
  Readirisは、辞書に保存されている文字の形を認識してインタラクティブ学習に使用し、新しい情報を辞書に保存することができます。

  [**フォント辞書を閲覧**]を選択した場合は、Readirisは辞書に保存されている文字の形を認識しますが、[**インタラクティブ学習**]がオンになっていても新しいコンテンツは辞書に追加されません。
  しかし[**インタラクティブ学習**]を選択すると、必要に応じてフォント辞書に保存されていない認識結果を修正しますから便利です。

  注意:
  [**新規フォント辞書**]がオンになっているときは、[**学習**]メニューの[**フォント辞書**]をクリックして既存の辞書を開かないでください。
  これを行うと、既存のフォント辞書のコンテンツが消去されてしまいます。

- [**認識+保存**]をクリックして、文書を認識します。

第 10 章
文書のフォーマットと保存

文書のフォーマット

ReadirisでOCRする文書は、様々な出力フォーマットで保存できます。Readirisは、OCR結果をAdobe Acrobat PDFファイル、Microsoft XPSファイル、Word、RTFとOpenDocumentテキストファイル、HTMLとXMLファイル、WordMLとSpreadsheetMLワークシート、ANSIとUnicodeテキストファイルで保存します。

この他にも、Readirisは結果をMicrosoft Word、Excel、Adobe Reader、Microsoft
XPSビューア、代表的なWebブラウザ、電子メールソフトウェアなどのターゲット アプリケーションに直接エクスポートできます。

> 注記:
> Readirisによって作成されたRTFファイルは、ウェブベースの事務アプリケーションであるAjaxWrite、ThinkFree、Zoho Writer、Google Writelyでも開くことができます。(Google WritelyはRTF、HTML文書両方を開くことができます)
> これらのアプリケーションを使用する場合は、必ずレイアウトオプションの[文字と段落の書式を維持]を選択します。

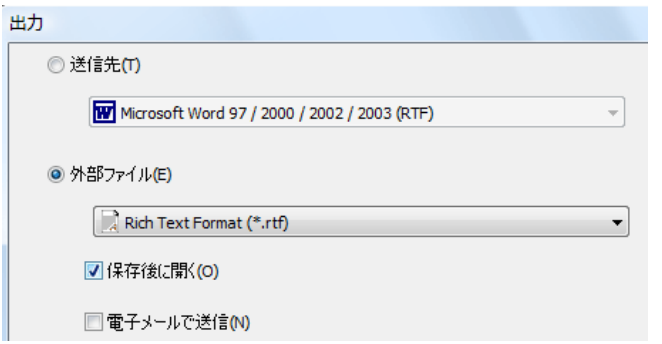
操作

- メインバーの[フォーマット]をクリックし、出力モードを選択します。

  Readirisは下記のいずれかを行います。

- 文書をアプリケーションに送信し、自動的に開く。
- 外部ファイルとして文書を保存。

[**電子メールで送信**]オプションは、新規メッセージを作成し、認識した文書を電子メール添付として挿入します。



- 各タブをクリックして、ご希望の設定を選択します。

選択した出力フォーマットに利用できない設定は、淡色表示になっています。

最も一般的に使用される出力ファイルフォーマットとそのオプションは、下記の項で説明されます。

- [**認識+保存**]をクリックして、文書を認識し、保存します。

OCR結果は、認識を繰り返さなくとも数回エクスポートできます。
[**フォーマット**]ボタンを再クリックし、テキストフォーマットとフォーマットオプションを変更します。
[**認識+保存**]を再度クリックします。

検索と並び替えのために、ReadirisではPDF、XPS、Word、RTF、WordML、SpreadsheetML、HTMLの**文書プロパティ**を定義できます。

文書のプロパティを定義するには、[**ファイル**]メニューの[**文書プロパティ**]をクリックします。 文書プロパティオプションは、[**認識+保存**]をクリックすると表示される「**出力ファイル**」ダイアログボックスからもアクセスできます。

複数ページの文書を**外部ファイル**として保存する場合、Readirisに各ページ用に別々の出力ファイルを作成することがもきますし、同じ文書に属する全ページを1つの出力ファイルに保存することもできます。
「**出力ファイル**」ダイアログボックスで、前者の場合は「**ページごとにファイルを１つ作成**」、後者の場合は「**文書ごとにファイルを１つ作成**」をクリックします。

ただし、「**ページごとにファイルを１つ作成**」と「**文書ごとにファイルを１つ作成**」オプションは、文書を外部ファイルとして保存する場合にのみ利用でき、文書をターゲットアプリケーションで開くときには利用できません。

テキスト文書のフォーマット

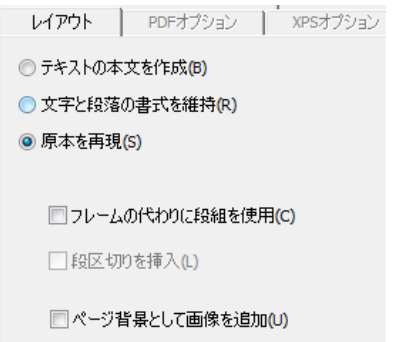
Readirisでは、数種のテキストベースの出力フォーマットを生成できます。Readirisは、Word、WordML、RTF、txt、OpenDocumentテキスト出力を提供します。

テキストベース出力ファイルの生成の仕方:

- メインツールバーの[**フォーマット**]をクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望の出力フォーマットを選択します。
- テキストフォーマットにもよりますが、数種のフォーマットオプションが利用できます。

  選択したテキストフォーマットに利用できないオプションは、淡色表示になっています。

レイアウトオプション



- [テキストの本文を作成]オプションは、Readirisによるテキストフォーマットを避けます。

  Readirisは、一続きの本文を生成します。

- [文字と段落の書式を維持]オプションは、本文と自動フォーマットの中間になります。

  フォントの種類、サイズ、字体は、認識全体で維持されます。

  タブと各ブロックの置換は再作成されます。

  テキストブロックと段は再作成されません。段落は流れるだけになります。

  表は正確に再キャプチャされます。

- [原本を再現]オプションは、元の文書のファックスコピーを再作成します。

  Readirisは、原書の真のコピーを生成し、スキャンした画像ではなくなります。

  Readirisは、電子メールアドレスやWebサイトのハイパーリンクも再作成します。

- [**フレームの代わりに段組を使用**]オプションは、段組のある文書を作成します。

  段組したテキストのほうが、複数のフレームを含む文書より編集が簡単で、 テキストは段から段へと自然に流れます。

  注記:
  システムが原書に段組を検知できない場合は、このフォーマットモードはフレームを代替として使用します。

- [**段区切りを挿入**]オプションは、各段の終わりに強制段区切りを挿入します。

  編集、追加、削除したテキストは、その段の中に残り、段区切りを超えて、自動的にテキストが流れ込むことはありません。

  ヒント:
  段組をした本文がある場合は、このオプションを無効にします。
  これにより、テキストが段から段へと自然に流れるようにできます。

- [**ページ背景として画像を追加**]オプションは、スキャンした画像を認識したテキストの下にページ背景として配置します。

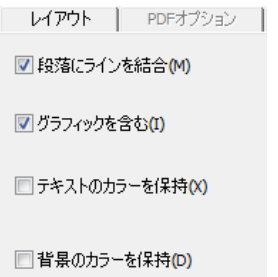
  しかし、このオプションは、出力ファイルのサイズを大幅に増大します。

  [**PDFテキスト-画像**]フォーマットは、PDFファイルと同じ結果になります。

  [**オプション**]タブの[**背景のカラーを保持**]オプションは、それほど極端ではなく、よりコンパクトな方法になります。

一般オプション

- [**オプション**]タブをクリックし、一般オプションを選択します。



- [**段落にラインを結合**]オプションは、段落自動検出を有効にします。

  Readirisは、新しい段落が始まるまで認識したテキストをワードラップし、「ハイフン」のある単語を行末で「再びくっつけ」ます。

- [**グラフィックを含む**]オプションは、自動フォーマットしたファイルのグラフィックを含みます。

  これは、文書の真のコピーを作成するのに不可欠です。

  [**グラフィック**]タブのグラフィックオプションを使用し、出力ファイルの中に保存されるグラフィックのカラーモードと解像度を定義します。

- [**テキストのカラーを保持**]オプションは、認識全体でテキストの元のカラーを維持します。

- [**背景のカラーを保持**]オプションは、認識全体でページ背景のスポットカラーを維持します。

出力ファイルの段落ごとに、均一の背景カラーが作成されます。

用紙サイズ

選択したフォーマットによりますが、希望の用紙サイズを選択できます。

- [用紙サイズ]タブをクリックし、矢印ボタンで用紙サイズを適用し、実行します。
- Readirisは、アクティブな用紙サイズを指定した順番をたどり、スキャンした文書に十分な大きさの最初の用紙サイズを使用します。

表からなる文書のフォーマット

Readirisを利用して、表データをスプレッドシート、ワードプロセッサ、Webブラウザに出力できます。
表はワークシートのセルごとに再構成され、ワードプロセッサファイルに表オブジェクトとして挿入されます。

Readirisは、罫線の有無にかかわらず、両タイプの表を認識します。

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	

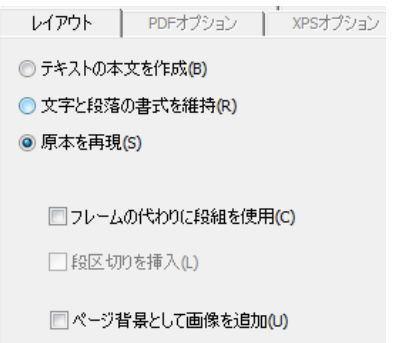
(罫線あり)

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

(罫線なし)

表からなる文書の生成の仕方:

- メインツールバーの[**フォーマット**]をクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望の出力フォーマットを選択します。
- ご希望のレイアウトオプションを選択します。

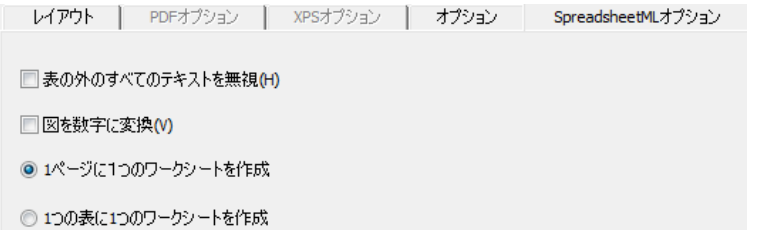


フォーマットオプションの詳細に関しては、「**テキスト文書のフォーマット**」をご参照ください。

SpreadsheetMLオプション

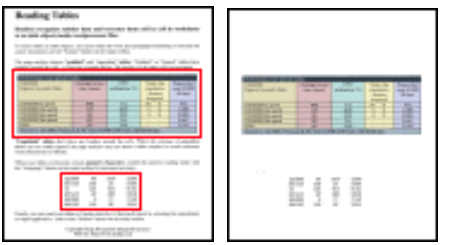
Microsoft Excel 2002/2003をターゲットアプリケーションとして選択する場合、特定のSpreadsheetMLオプションが利用できます。「 SpreadsheetML オプション」タブをクリックして、オプションを表示します。

このフォーマットを選択したときは、レイアウトオプションの[**原本を再現**]は利用できません。



- [**表の外のすべてのテキストを無視**]オプションは、表を保存し、その他すべての認識結果を無視します。

  表の中のデータはすべて回収されますが、表の外のデータは回収されません。



- [**データを数字に変換**]オプションは、認識したデータ値を数字としてエンコードします。

  その結果、これらのセルで演算操作が行えます。
  (あらゆる表の)テキストセルはテキストのままです。

  表の中のデータだけが、数字としてエンコードされることにご留意ください。

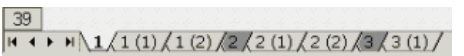
  Excelは、数字としてエンコードされたデータだけに演算操作を実行します。

- [**1ページに１つのワークシートを作成**]オプションは、スキャンしたページごとにワークシートを1つ作成します。

ページに表とテキストがある場合は、すべてが同じワークシートに配置されます。

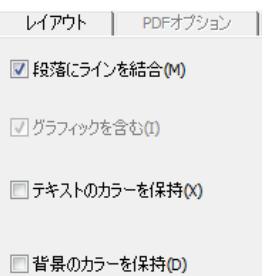
- [**1つの表に 1 つのワークシートを作成**]オプションは、別個のワークシートに各表を配置し、認識したテキスト(表の外)を別のワークシートに入れます。

認識した文書に**数ページ**がある場合は、ページごとに構造が繰り返されます。



一般オプション

[**オプション**]タブをクリックし、一般オプションを選択します。



- [**段落にラインを結合**]オプションは、段落自動検出を有効にします。

Readirisは、新しい段落が始まるまで認識したテキストをワードラップし、「ハイフン」のある単語を行末で「再びくっつけ」ます。

- [**背景のカラーを保持**]オプションは、各セルの背景のカラーを再作成します。

	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		

用紙サイズ

選択したフォーマットによりますが、希望の用紙サイズを指定することができます。

- [用紙サイズ]タブをクリックし、矢印ボタンで用紙サイズを適用し、実行します。
- Readirisは、アクティブな用紙サイズを指定した順番をたどり、スキャンした文書に十分な大きさの最初の用紙サイズを使用します。

PDF文書の作成

Readirisは、テキスト、テキスト-画像、画像
テキスト、画像の4種類のPDF出力を生成します。

PDF出力の仕方:

メインツールバーの[フォーマット]をクリックし、[送信先]、または[外部ファイル]ドロップダウンリストからご希望のPDFタイプを選択します。

PDF画像

PDF画像を選択した場合は、Readirisは画像のみのPDF文書を生成し、OCRを実行しません。

PDF画像-テキスト

PDF**画像-テキスト**を選択した場合は、Readirisはテキストを認識し、ページ画像と 認識したテキストを含んだ検索可能なPDFファイルを作成します。

ページ画像は、テキストの上に配置されます。

このフォーマットでは、(スキャンしたままの)元の文書をいつも表示することができます。また、画像の下に隠れているOCR済みテキストを検索したり、コピー&貼り付けをすることができます。
その結果、このフォーマットは**保管を目的とする場合**に便利です。

PDFテキスト

PDF**テキスト**を選択した場合は、Readirisはテキストを認識し、検索可能なPDFファイルを作成します。

この単層PDFファイルには、ページ画像は含まれていません。

PDFテキスト-画像

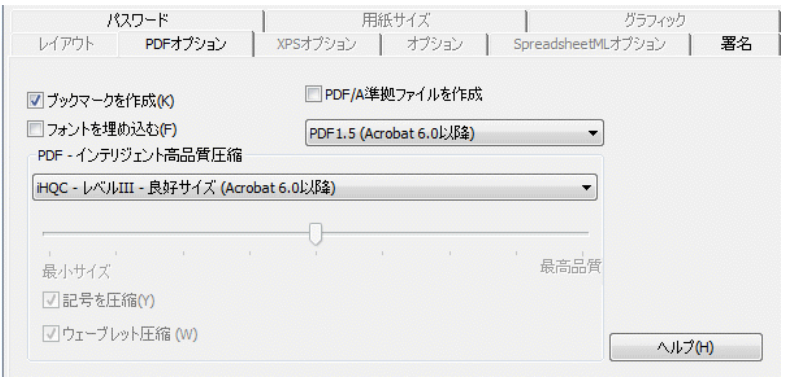
PDF**テキスト-画像**を選択した場合は、Readirisはテキストを認識し、ページ画像と 認識したテキストを含んだ検索可能なPDF文書を作成します。

ページ画像は、テキストの下にあります。

PDFオプションの選択

PDFオプションの選択の仕方:

- メインツールバーの[**フォーマット**]をクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望のPDFタイプを選択します。
- 選択した**PDFタイプ**にもよりますが、数種のオプションが利用できます。
  [**PDFオプション**]タブをクリックして、アクセスします。



ブックマークの作成

[**ブックマークの作成**]オプションは、Adobe Acrobat PDFファイルの各テキストブロック、グラフィック、表のブックマークを作成します。

フォントを埋め込む

[**フォントを埋め込む**]オプションを選択し、Adobe Acrobat PDFファイルのフォントを埋め込みます。

フォントを埋め込むと、フォント代用を防止し、読者のコンピュータ構成にかかわらず、読者はテキストを元のフォントで表示できます。

フォントの埋め込みは、認識した文書のファイルサイズを若干増大します。

PDF文書をiHQCで圧縮

「標準」PDF出力のほかにも、ReadirisはiHQC圧縮PDF出力も提供します。「**画像-テキスト**」と「**画像**」のPDF文書は、画像品質を損なうことなく、iHQC方法を利用して超圧縮できます。

iHQCは、**intelligent High-Quality Compression**の略であり、I.R.I.S.が開発し、特許を有している効率性の高い圧縮技術です。

iHQC圧縮出力の生成の仕方:

- メインツールバーの[**フォーマット**]ボタンをクリックし、2種類の出力モードから選択します。
- [**送信先**]、または[**外部ファイル**]リストからご希望のPDFタイプ、[**PDF 画像-テキスト**]、または[**PDF画像**]のいずれかを選択します。
- [**PDFオプション**]タブで、所定の圧力レベルを選択します。

  Readiris Proは、**レベルI - 良好サイズ**、**レベルI - 高品質**圧縮をサポートします。

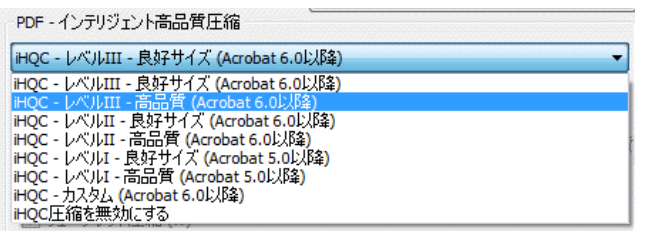
  Readiris Corporateは、**レベルII とIII 良好サイズ**、**高品質**、および**カスタム**圧縮もサポートします。

  レベルII
  圧縮では、[**記号を圧縮**]オプションが自動的に有効になり、テキストをコンパクトに圧縮します。

  レベルIII

圧縮では、[**ウェーブレット圧縮**]オプションが自動的に有効になり、グラフィックをコンパクトに圧縮します。

[**カスタム**]圧縮を選択すると、これらのオプションを個別に有効、無効にすることができます。
スライダを利用して、**良好サイズと高品質**の比率を定義することもできます。



- [**認識+保存**]をクリックして、文書を認識し、iHQC圧縮PDFファイルとして保存します。

PDF文書の目的変更

PDF文書を生成するほかにも、ReadirisはPDFファイルの目的変更を行うことが出来ます。
Readirisは、**画像**PDFを**テキスト**PDFやその他のサポートされているあらゆるテキストフォーマットに変換できますし、読取専用PDFコンテンツを**解除**できます。

> **警告:**
> Readirisは、ユーザパスワードで保護されたPDF文書は開きません。

操作

- メインツールバーの[**開く**]をクリックし、Readirisに目的を変更させたいPDFファイルを選択します。

  ダイアログボックスの[**ページ範囲**]で[**ページ**]を選択し、Readirisで開きたいページを指定します。

- ダイアログボックスの[**開く**]をクリックして、ご希望のPDFファイルを開きます。
- メインツールバーの[**フォーマット**]をクリックし、ご希望のPDFタイプを選択します。

  PDFタイプの詳細に関しては、「**PDF文書の作成**」をご参照ください。

- [**認識+保存**]をクリックして、文書の目的を変更します。

XPS文書の作成

Readirisは、**テキスト**、**テキスト-画像**、**画像**
テキスト、**画像**の4種類のXPSファイルを生成します。

XPSは、XML Paper Specificationの略であり、Microsoft社が開発した固定レイアウト書式です。

XPS出力の仕方:

メインツールバーの[**フォーマット**ボタンをクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望のXPSタイプを選択します。

XPS画像

XPS**画像**を選択した場合は、Readirisは画像のみのXPS文書を生成し、OCRを実行しません。

XPS画像-テキスト

XPS**画像-テキスト**を選択した場合は、Readirisはテキストを認識し、ページ画像と認識したテキストを含んだ検索可能なXPSファイルを作成します。

ページ画像は、テキストの上に配置されます。

このフォーマットでは、(スキャンしたままの)元の文書をいつも表示することができます。また、画像の下に隠れているOCR済みテキストを検索したり、コピー&貼り付けをすることができます。
その結果、このフォーマットは**アーカイブを目的とする場合**に便利です。

XPSテキスト

XPS**テキスト**を選択した場合は、Readirisはテキストを認識し、検索可能なXPSファイルを作成します。

この単層XPSファイルには、ページ画像は含まれていません。

XPSテキスト-画像

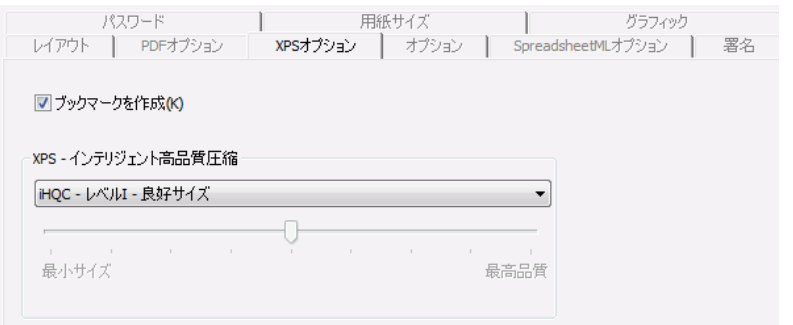
XPS**テキスト-画像**を選択した場合は、Readirisはテキストを認識し、ページ画像と認識したテキストを含んだ検索可能なXPS文書を作成します。

ページ画像は、テキストの下にあります。

XPSオプションの選択

XPSオプションの選択の仕方:

- メインツールバーの[**フォーマット**]をクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望のXPSタイプを選択します。
- 選択した**XPSタイプ**にもよりますが、数種のオプションが利用できます。
  [**XPSオプション**]タブをクリックして、アクセスします。



ブックマークの作成

[**ブックマークの作成**]オプションは、Microsoft XPSファイルの各テキストブロック、グラフィック、表のブックマークを作成します。

XPS文書をiHQCで圧縮

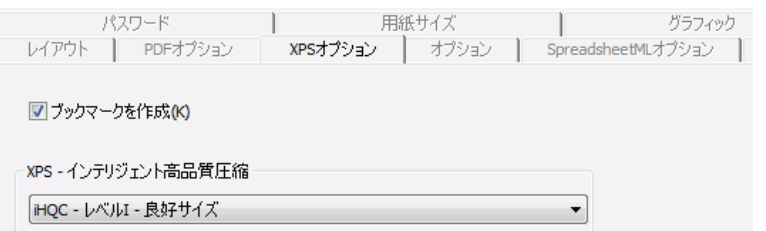
「 標準」 XPS出力のほかにも、ReadirisはiHQC圧縮XPS出力も提供します。「 画像
テキスト」と「 画像」 のXPS文書は、iHQC方法を利用して超圧縮できます。

iHQCは、**intelligent High-Quality Compression**の略であり、I.R.I.S.が開発し、特許を有している効率性の高い圧縮技術です。iHQCは、MP3が音楽、DivXが映画を圧縮しているのと同じ要領のものです。

iHQC圧縮XPS出力の生成の仕方:

- メインツールバーの[**フォーマット**]をクリックし、2種類の出力モードから選択します。
- [**送信先**]、または[**外部ファイル**]リストからご希望のXPSタイプ、[**XPS画像-テキスト**]、または[**XPS画像**]のいずれかを選択します。
- [**XPSオプション**]タブで、所定の圧力レベルを選択します。

  Readirisは、**レベルI 良好サイズ**、**レベルI 高品質**圧縮をサポートします。



グラフィック オプションの選択

選択した出力フォーマットとターゲット
アプリケーションにより、詳細なグラフィックオプションが利用できる場合があります。
グラフィックオプションを利用して、画像の品質や解像度を変更することができます。

グラフィックオプションにアクセスする仕方:

- メインツールバーの[**フォーマット**]をクリックし、[**送信先**]、または[**外部ファイル**]ドロップダウンリストからご希望の出力フォーマットを選択します。
- [**グラフィック**]タブをクリックして、オプションを表示します。

カラーモード

デフォルトでは、Readirisはグラフィックを**カラー**で保存します。

白黒でグラフィックを保存する場合は、[**白黒**]を選択します。

解像度

デフォルトでは、**Readirisはスキャン解像度**を維持します。

解像度を下げるオプションも選択できます。

ヒント:
文書をHTMLファイルとして保存し、Webサイトに掲載する場合は、解像度を70 dpi (スクリーン解像度)に下げます。

JPEGの品質

PDF、XPS、Word、RTF文書の中に保存されたグラフィックは、JPEGで保存されています。

スライダを使用して、JPEGの品質を調整します。

JPEG 2000 圧縮

PDF、XPSフォーマットでファイルを保存する場合は、Readirisは、これらのファイルの中に保存されているカラー画像、またはグレースケール画像にJPEG 2000を適用できます。

JPEG
2000は、JPEG標準の最新でよりコンパクトなバージョンです。

[**JPEG 2000圧縮2000**]オプションを選択して適用します。

第 11 章
設定の保存と読み込み

Readirisで指定した設定はすべて、アプリケーションを終了した後にも将来使用できるよう自動的に保存されます。
出荷時の設定に戻すには、[**ファイル**]メニューの[**出荷時の設定に戻す**]コマンドをクリックします。

別々の設定が必要な各種グループの文書をスキャンする場合、各グループ用に別個の**設定ファイル**を保存すると便利です。

操作

- 特定の文書グループに使用したい設定を選択します。
- [**ファイル**]メニューから[**設定を保存**]をクリックします。
- 後から同じグループの文書をスキャンしたり開く場合は、[**ファイル**]メニューの[**設定の読み込み**]をクリックします。
- 正しい設定ファイルを選択し、[**開く**]をクリックします。
- [**認識+保存**]をクリックして、正しい設定を使用して文書を認識します。

注記:
[**ファイル**]メニューの[**情報**]コマンドは、選択した設定のもっとも大切な概要を表示します。

第 12 章
複数ページの文書を認識

複数の画像ファイルを開いて認識

Readirisは、複数の画像ファイルを同時に処理できるようデザインされています。

複数の画像ファイルの開け方:

- メインツールバーの[開く]をクリックします。

- 開きたい画像ファイルを選択します。
    - 最初の画像ファイルを選択し、**Ctrl**を押さえて次の画像を選択します。
    - 続きにリストされている画像ファイルを選択する場合は、最初の画像をクリックし、**Shift**キーを押さえて最後の画像を選択します。

注記:
複数ページからなる単一のファイル(複数ページのTIFFファイル、PDFファイルなど)を開く場合は、ページ範囲を指定できます。
ダイアログボックスの[ページ範囲]で[ページ]を選択し、認識したいページを指定します。

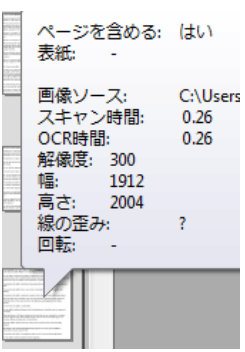
- [**開く**]をクリックして画像を開きます。

  Windows Explorerから**Readiris**画像ウィンドウに画像ファイルをドラッグ・アンド・ドロップして開くこともできます。

- ページツールバーが、開いた画像ファイルを表示します。

  ヒント:
  ページのサムネイルの上にマウスカーサーを置き、ページごとの設定情報を表示することができます。

  ページツールバーは、複数ページの文書編集にも使用できます。
  詳細に関しては、「**複数ページの文書を編集**」をご参照ください。



- 認識設定を定め、[**認識+保存**]をクリックして、認識を実行します。

- 現在の文書に別の画像を開いたり、スキャンしたい場合は、メインツールバーの[**スキャン**]、または[**開く**]ボタンをクリックします。

  現在の文書を削除したいかどうか尋ねられます。
  現在の文書を削除して新規文書を開始する場合は[**はい**]、現在

の文書に別のスキャンを追加する場合は[**いいえ**]をクリックします。

複数ページの文書をスキャンして認識

Readirisは、複数ページの文書を処理できるようデザインされています。

Readiris HomeとReadiris Proは、50ページまで処理できます。

Readiris Corporateは、処理できるページ数に制限はありません。

Readirisに複数ページをスキャンするには、シートフィードスキャナの場合は[**自動文書フィーダ**]、フラットベッドスキャナの場合は[**インターバル スキャン**]を使用します。

文書フィーダ(シートフィード スキャナ)で複数ページの文書をスキャン

- メインツールバーの[**スキャナ**]をクリックし、[**ADF**](自動文書フィーダ)を選択します。
- スキャナの文書フィーダにページを入れ、[**スキャン**]をクリックしてスキャンを開始します。
- [**認識+保存**]をクリックして、文書を認識します。

インターバル スキャン(フラットベッド スキャナ)で複数ページの文書をスキャン

- メインツールバーの[**スキャナ**]ボタンをクリックします。
- [**別のページをスキャンするまでの時間:**]を選択し、矢印ボタンを使用して間隔(秒)を指定します。

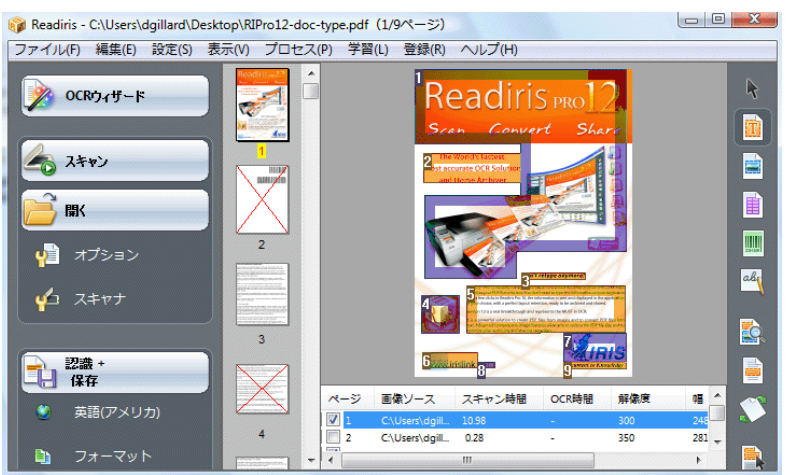
毎回[**スキャン**]ボタンをクリックしなくても、指定された秒後に、スキャナは自動的に別のページをスキャンします。

自動スキャンを終了するには、スキャンダイアログボックスの[**中止**]をクリックするか、キーボードの[**ESC**]を押します。

自動スキャンを一時停止するには、インターバルスキャンダイアログボックスの[**一時停止**]をクリックするか、キーボードのスペースバーを押します。
続行する準備ができたら、[**再開**]をクリックします。

複数ページの文書を編集

Readirisで複数の文書が開いたりスキャンされた場合、ページツールバーが各文書のサムネイルを表示します。
ページツールバーのサムネイルは、複数ページの文書編集にも使用できます。



文書内でページを移動

- 移動したいページを右クリックし、[**ページを選択**]をクリックします。

- ページを正しいポジションにドラッグします。
- または、ページを右クリックし、[**ページを上に移動**]、[**ページを下に移動**]をクリックします。

ページの削除

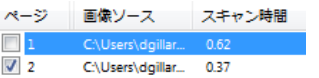
- 削除したいページを右クリックし、[**ページを削除**]をクリックします。
- または、ページを選択して、キーボードの[**Delete**]ボタンを押します。

認識からページを除外:

- 除外したいページを右クリックし、[ページを除外]をクリックします。
- または、文書パネルのページ番号ボックスをクリアします。

除外されたページは、ページツールバー上で×印がつきます。

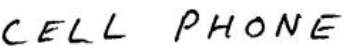
スキャンした画像を印刷したり、スキャンを複数の画像ファイルに保存する場合、除外されたページは無視されます。



ヒント:
[**編集**]メニューの[**すべてのページを含める**]と[**すべてのページを除外**]コマンドは、すべてのページに同時に適用されます。

第 13 章
手書きテキストの認識

タイプされたテキスト、表、グラフィック、バーコードのほかにも、
Readirisは**手書き**テキストを認識します。
手書きとは、1字1字別になったブロック体の文字のことです。



非常に特殊なICRソフトウェア(インテリジェント文字認識)でないと、手書きの文字は認識できません。

手書きの認識の仕方:

- メインツールバーの[手書き]をクリックします。

- 手書きテキストの周りにフレームを描きます。
- メインツールバーの[認識+保存]をクリックします。

手書きを含め、文書全体が認識されます。

注記:
手書きゾーンを右クリックし、[**テキストとしてコピー**]をクリックして、手書きゾーンのみを認識してクリップボードに送信します。

認識した記号

手書き認識は、ラテンアルファベットに限られ、数字(0～9)と大文字(A～Z)、句読点のコンマ、ピリオド、+サイン、ハイフンをサポートします。

アクセント、ウムラウト、その他の特殊文字はサポートされません。

注記

- Readirisは、**ブロック体**をサポートしますが、**筆記体**はサポートしません。

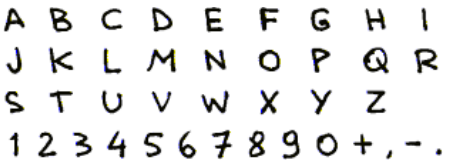
  詳細に関しては、「**手書きルール**」をご参照ください。

- 認識後、大文字は小文字に置き換えられます。ただし、文頭にある場合はそのままになります。
- 文書の言語特徴であるフォントの種類、文字ピッチなどは、手書きには適用されません。
- インタラクティブ学習も適用されません。
  ICRテクノロジーは、100万以上の筆記サンプルを基にしています。

手書きルール

Readirisが手書きを認識するためには、数点のルールを顧慮しなければなりません。

- 標準的で、間隔をあけた文字を書くこと。



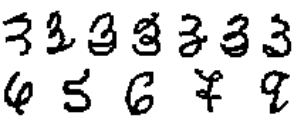
A、G、Qの文字が、どのように書かれているかにご注目ください。

- 十分に濃いボールペンを使用すること。

  黒いペンのほうが青いペンより結果がよくなります。

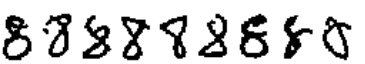
  鉛筆を使わないこと。

- あまり文字に飾りを入れないこと。



  飾りを入れすぎた書き方の文字は、OCRエラーが発生しやすくなります。

- ループの閉じなければならない部分は閉じ、閉じてはいけない部分は開けたままにすること。



- 壊れた文字を避けること。

- なぞらないこと。

なぞると画像品質が低下し、手書きした記号がはっきりしなくなります。

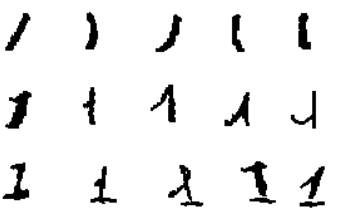
抹消された文字は、認識されません。

- 数字の1を正しく書くこと。

数字の1は、「英国式」と「ヨーロッパ風」で書くことができます。

1の下の部分に線を入れても入れなくてもかまいません。

下の横線は、縦線につながっていなくてもかまいません。



ヒント:
最適の結果が得られなかった場合は、I.R.I.S.の**筆記フォーム**を利用し、あなたの筆記スタイルを適用します。
白紙の.R.I.S.筆記フォームは、フルページのテンプレートとなり、ブロック体の文字を正しいサイズで、正確に記載できます。
フォームは、Readiris CD-ROMとReadirisインストールフォルダにあります。

第 14 章
バーコードの認識

バーコードの読み出しについて

128
ヶ国語の光学式文字認識のほかにも、Readirisはバーコードの読み出しもできます。

一般的に普及しているバーコードに対応しています。Codabar(NW-7)、Code 128、Code 39、Code 39 extended、Code 39 HIBC、Code 93、Datalogic 2 of 5、Discrete 2 of 5、EAN-13、EAN-8、Interleaved 2 of 5、MSI Pharmaceutical、MSI-Plessey、Kodak patch code、PDF-417、PostNet、UCC-128、UPC-A、UPC-Eに対応できます。

ReadirisがOCRを実行するには、レーザープリンタ、インクジェットプリンタで印刷したバーコードが必要です。
マトリックスで印刷したバーコードは、十分なコントラストがなく、解像度がほとんど60 dpiに限られているため、サポートされていません。

バーコードの手動読み出し

- Readirisに認識してほしいバーコードを特定します。

- [**設定**]メニューから[**バーコード**]をクリックします。
- Readirisに認識してほしいコードを選択します。
- Readirisが検査数字を確認、または削除するかを特定します。

- 画像ツールバーの[バーコード]をクリックし、文書内のバーコードゾーンの周りにフレームを描きます。

- メインツールバーの[**認識+保存**]をクリックします。

バーコードを含め、文書全体が認識されます。

注記:
バーコードゾーンを右クリックし、[**データとしてコピー**]をクリックして、コンテンツをクリップボードにコピーします。

索引

A

ADF 87

E

Excel出力 61

H

HTML出力 61

J

JPEG 2000 圧縮 81

O

OpenDocument出力 61

P

PDF iHQC出力..... 74

PDFオプション 72

PDF出力 61, 71

PDF文書の目的変更 75

R

Readirisのアンインストール
..... 11

RTF出力.....61

S

SmartTasks 16, 19

SpreadsheetML出力.....61

U

Unicode出力 61

W

WordML出力.....61

Word出力.....61

X

XML出力 61

XPS iHQC出力 79

XPSオプション.....78

XPS出力.....61, 76

あ

アジア言語版.....4, 7, 52

アジア言語文書 4, 7, 52

アラビア語文書 4, 8, 52

い

インストール 10

インターバル スキャン..... 87

インタラクティブ学習..... 56

う

ウィザード 13

ウィンドウイング
テンプレート 47

お

オプション 23

か

カラーモード 28

カラー画像 28, 33

カラー画像をスムーズにする
..... 31, 33

き

キャラクターピッチ..... 55

く

グラフィック オプション... 80

グレースケール画像..... 28

こ

コントラスト 28, 36

さ

サポートされている画像フォ
ーマット 26

し

システム動作環境 9

す

スキャナ設定 27

スキャンした文書の調整 33

スピードと精度 54

スプレッドシート文書 67

て

テキスト文書 63

デジタルカメラ 29

と

ドットマトリックス 55

は

バーコード 95

ふ

フォントの種類 55

フォント辞書 59

へ

ページ 89

移動..... 89

削除..... 89

除外..... 89

選択..... 89

ページ傾き補正 23

ページ分析 23, 41

ヘブライ語文書 4, 8, 52

め

メインツールバー..... 16

ゆ

ユーザー
インターフェース言語..... 18

ユーザインターフェース.... 15

ユーザインターフェースの変
更 18

れ

レイアウトオプション..... 64

レイアウトファイル..... 47

わ

ワークシート.....67

漢字

汚れの補正.....36

画像ツールバー.....16

解像度.....28

回転.....23

輝度.....28, 34

傾き補正.....23

原本の文書を再現.....64

言語.....52

混合言語.....54

自動ウィンドウイング 41

自動文書フィーダ 87

手書き.....91, 92

手動ウィンドウイング 43

出荷時の設定.....83

出荷時の設定に戻す 83

出力フォーマット 61

数字.....53

精度とスピード.....54

製品サポート 12
設定の読み込み 83
設定ファイル 83
設定を保存 83
線の傾き 31
中東言語版 4, 8, 52
電子メールで文書を送信.... 62
登録 12
背景の色 65
白黒画像 28
反転画像.....29
表.....67
表セルの背景色.....70
複数ページの文書.....85, 87
複数ページの文書を編集....88
文書の特徴.....54
文書パネル.....17
文書プロパティ 63
用紙サイズ.....67