

58カ国語対応活字OCRソフトウェア

e.Typist ^{OCR} v15.0
イータイピスト

かんたん操作マニュアル



MEDIA DRIVE CORPORATION
メディアドライブ株式会社



はじめに.....	3
1. e.Typist を起動してみましよう.....	5
2. 原稿を読み込んでみましよう.....	7
3. 画像の状態を確認しまししよう.....	11
4. 認識条件を指定しまししよう.....	13
5. 認識したい範囲を指定しまししよう.....	14
6. 文字認識をしまししよう.....	19
7. 認識結果を修正しまししよう.....	20
8. データを他のソフトに転送しまししよう.....	22
9. 作業を終わりにしまししよう.....	25
10. おわりに.....	26

はじめに

e.Typist は、「OCR のソフトウェア」です。

OCR とは、画像内の文字をテキスト変換する機能です。

たとえば、あなたが今、興味を持っている趣味について、友人に紹介するため、パソコンで文章を作っているとしましょう。

ある日、新聞を読んでいると、その趣味について興味深い記事が掲載されていました。あなたはその新聞記事を「参考文章」として、自分が作っている文章内に引用したい。その場合、どうしますか？

新聞を見ながら、一文字ずつ、キーボードから入力することになるでしょう。

OCR とは、この操作を自動で行ってくれる機能です。

新聞をスキャンしたあと、自動でキーボードで打った文字に変換してくれます。

ほかにも、いろいろな使い方ができます。

- ・毎日届く大量の注文書の注文番号だけを Excel に転送したい
- ・資料の文献を一冊取り込んでおき、必要な単語を検索できるようにしたい
- ・英語の文章を取り込んで翻訳したい

上記のような場合にも、OCR は役立ちます。

OCR ソフトを使って、手持ちの文献や資料を活用してください。



● サンプル原稿

【認識サンプル A(Word 出力用)】

21 世紀となり、既在世界は今大きく揺れ動いています。

地球規模の新時代を模索し、政治・経済・文化・思想・科学をはじめとする、あらゆる分野・領域において、既成のパラダイムの組み替えが行われています。こうしたイノベーションの波は、私たちの関わるテクノロジーの分野を呑み込み、従来の技術領域にとらわれない、いわば「マルチテクノロジー新時代」の幕開けを告げました。

情報システムは分散化、パーソナル化へ、情報処理は知識・経験を取り込んだ AI (人工知能)化へ、そして情報自体は、文字・図形・音響・映像を融合した

マルチメディア化へと、私たちを取り巻く環境は大きな変貌を遂げようとしています。

また、その一方で、これまであまりにも機能を優先、その先進性の向上を競いあい、ともすればなおざりにされてきたテクノロジーと人の関わりについても、もう一度、真剣に考え直され始めてきました。すべてのテクノロジーに人が介在するのだと、いまさらながらに気づきはじめてたのです。

メディアドライブは、このような時代の潮流のなかで、高度情報化社会の技術トレンドを先取りする情報システムのシンクタンクとして、また自由な環境と発想でさまざまな技術分野に挑戦するベンチャー型企業として、今までにない、あるいは一歩進んだ情報関連機器の開発と、そのために必要な基盤技術の研究を推進し続けています。

しかも、そのねらいとするところは、誰もが手軽に使いこなせる、新しいヒューマンインターフェイスを装備した「限りなく人間の感性に近い」システムの開発なのです。



パッケージ関連商品	ソリューション関連商品	マルチメディア関連商品
WinReaderPRO	FormOCR	CrossMediatorforVideo
MacReaderPRO	WinReaderHandS	CrossMediatorPRO
e.Typist	FAXOCR	コマーシャル検出・分類
ExcelOCR/WordOCR	免許証 OCR	マルチメディア検索ライブラリ
やさしく PDFOCR	投票用紙認識分類ソフト	
やさしくファイリング	文字認識ライブラリ	
やさしく名刺ファイリング PRO	活字文字認識ライブラリ	
やさしく名刺ファイリング	手書文字認識ライブラリ	
やさしく名刺ファイリング forOutlook	名刺認識ライブラリ	
やさしくはがきファイリング		
やさしく PDF へ文字入力		
やさしく PDF へ文字入力 PRO		
やさしくビデオファイリング		

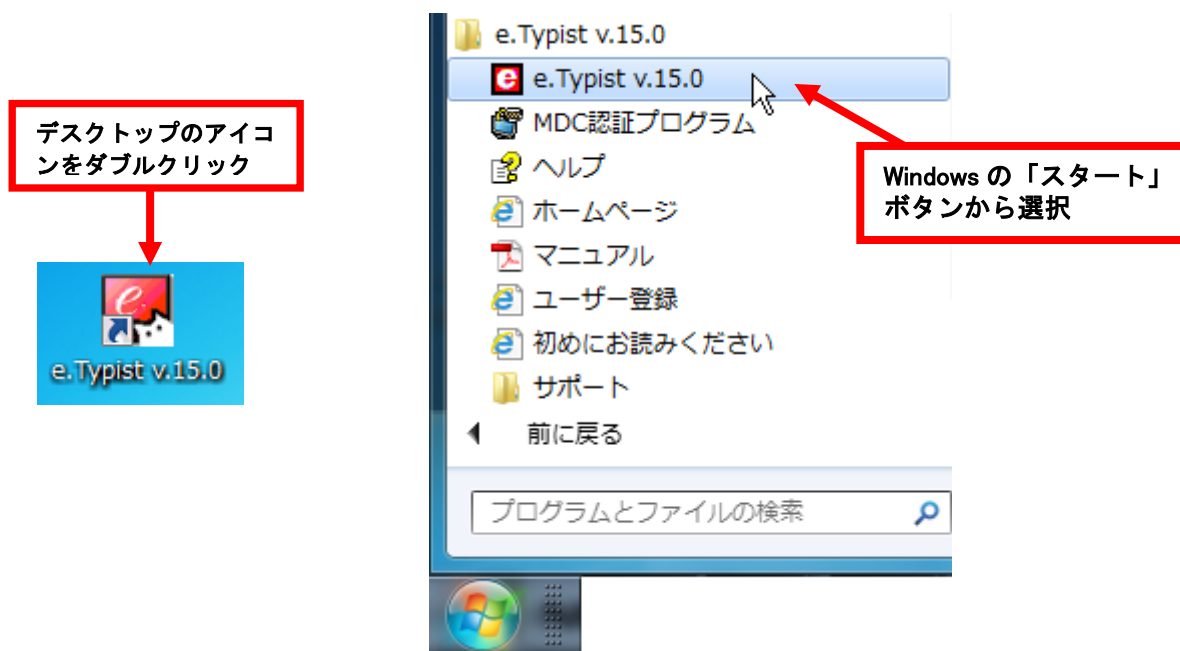
1. e.Typist を起動してみましょう

まずは e.Typist を起動してみましょう。

●Windows7 の場合には…

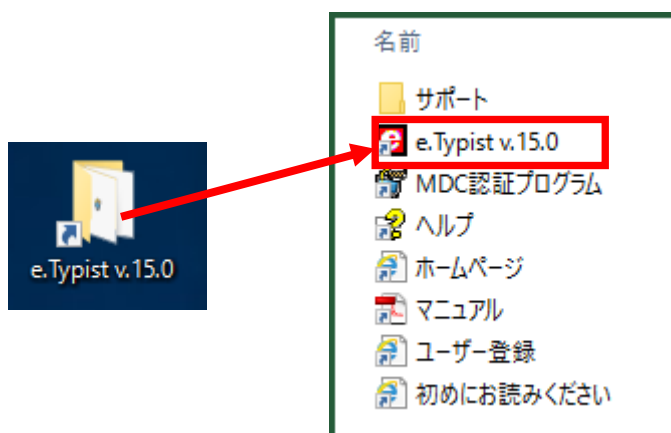
デスクトップの「e.Typist v.15.0」アイコンをダブルクリックします。

または、Windows の「スタート」ボタンから「(すべての)プログラム」→「e.Typist v.15.0」のフォルダを開き、「e.Typist v.15.0」を選択します。

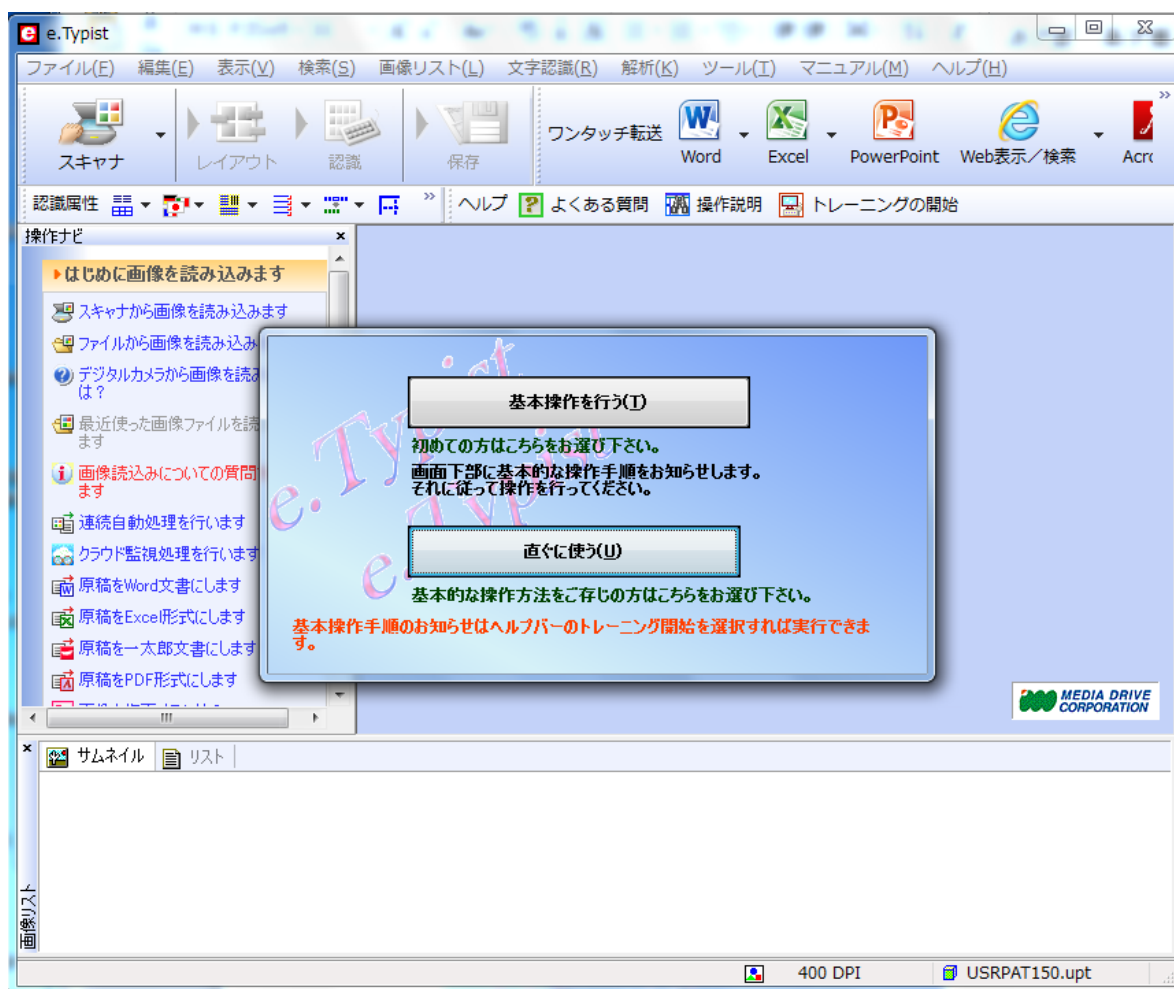


●Windows8 以降の場合には…

デスクトップに作られた「e.Typist v.15.0」フォルダを開き、「e.Typist v.15.0」をダブルクリックしてください。

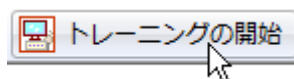


e.Typist v.15.0 が起動します。
ここでは「直ぐに使う」を選択してください。



●「基本操作を行う」について…

簡易マニュアルの操作が一通り終了したら、「基本操作を行う」を使ってみてください。
いろいろなパターンの画像を使い、1ステップずつ、操作を確認することが出来ます。
なお、2回目以降の起動時には、上記の「基本操作を行う/すぐに使う」画面は表示されませんが、ツールバーの「トレーニングを開始」ボタンを押すと「基本操作を行う」の機能が使えます。



2. 原稿を読み込んでみましょう

OCRしたい原稿を読み込みます。

まずは、この簡易マニュアル 4 ページにある原稿を使って試すこともお勧めです。

● スキャナから読み込む


新聞や書類など、「紙」になっている原稿は、スキャナから取り込みます。

1. まずは、使用するスキャナを設定します。

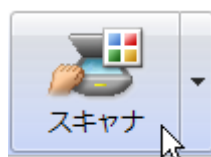
「ファイル」メニューから「スキャナの選択」を選択します。

「スキャナドライバの選択」ダイアログが表示されるので、ご利用のスキャナのドライバを選択し、「OK」ボタンを押します。



 このダイアログにスキャナ名が表示されない場合は、スキャナ用のドライバをインストールしてください。スキャナ用ドライバは、スキャナごとに用意されています。スキャナ付属のCDからインストール、または、スキャナメーカーのホームページからダウンロードしてインストールしてください。

2. 「スキャナ」ボタンをクリックします。



3. スキャナからの取り込みパネルが表示されます。

表示されてくる画面は、スキャナごとに異なりますが、大事な設定箇所は以下です。

・「解像度」(DPI)

「300」/「400」/「600」を選択します。

原稿の文字が小さい場合は「600」で読み込んでください。

・「色」

原稿にあわせて選びます。

おおまかに、カラー/グレースケール/モノクロが用意されています。

・「濃度」

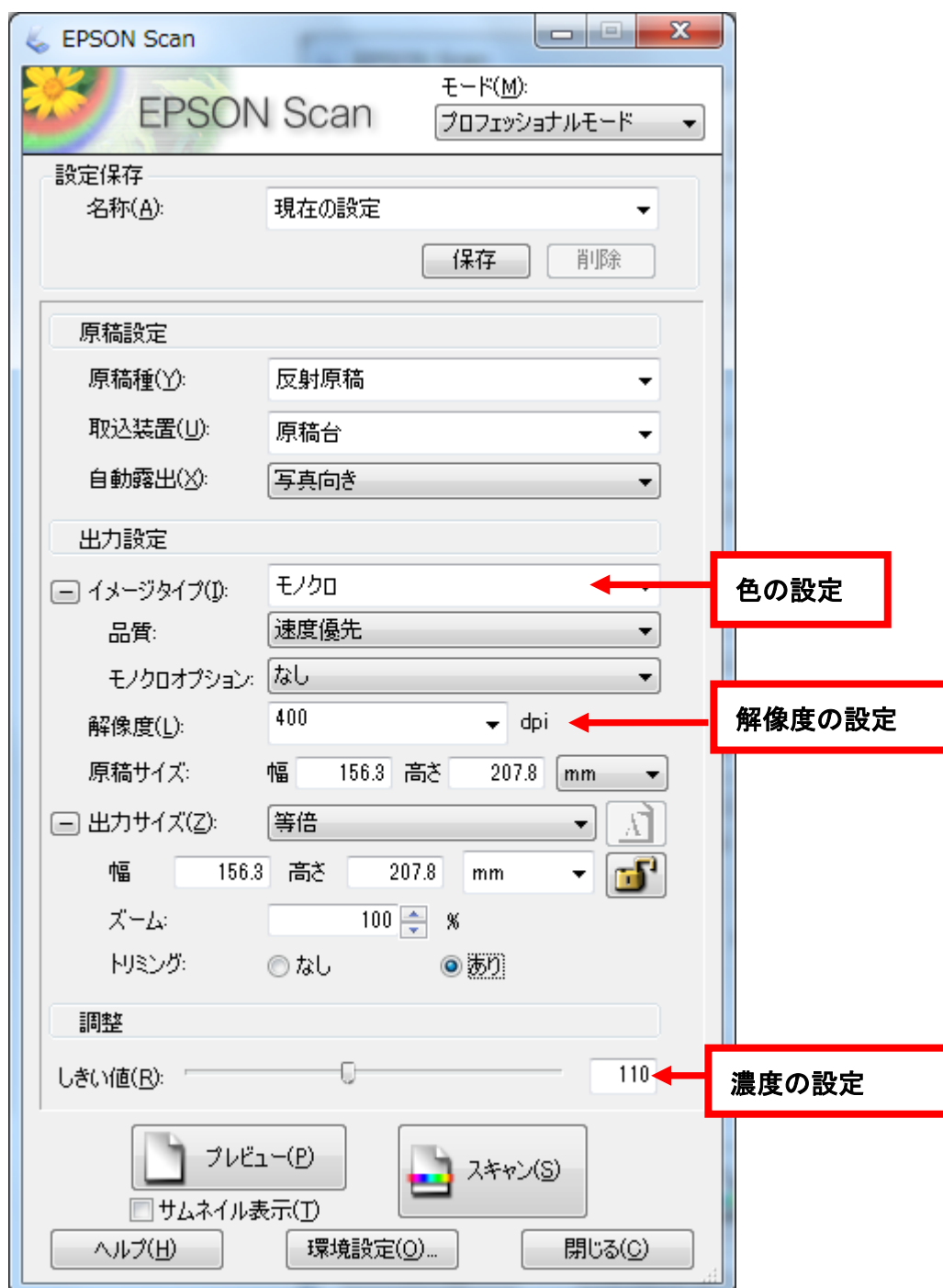
文字がかすれたり、つぶれない、また裏写りしないように濃度を調整してください。

・「原稿のセット方法」

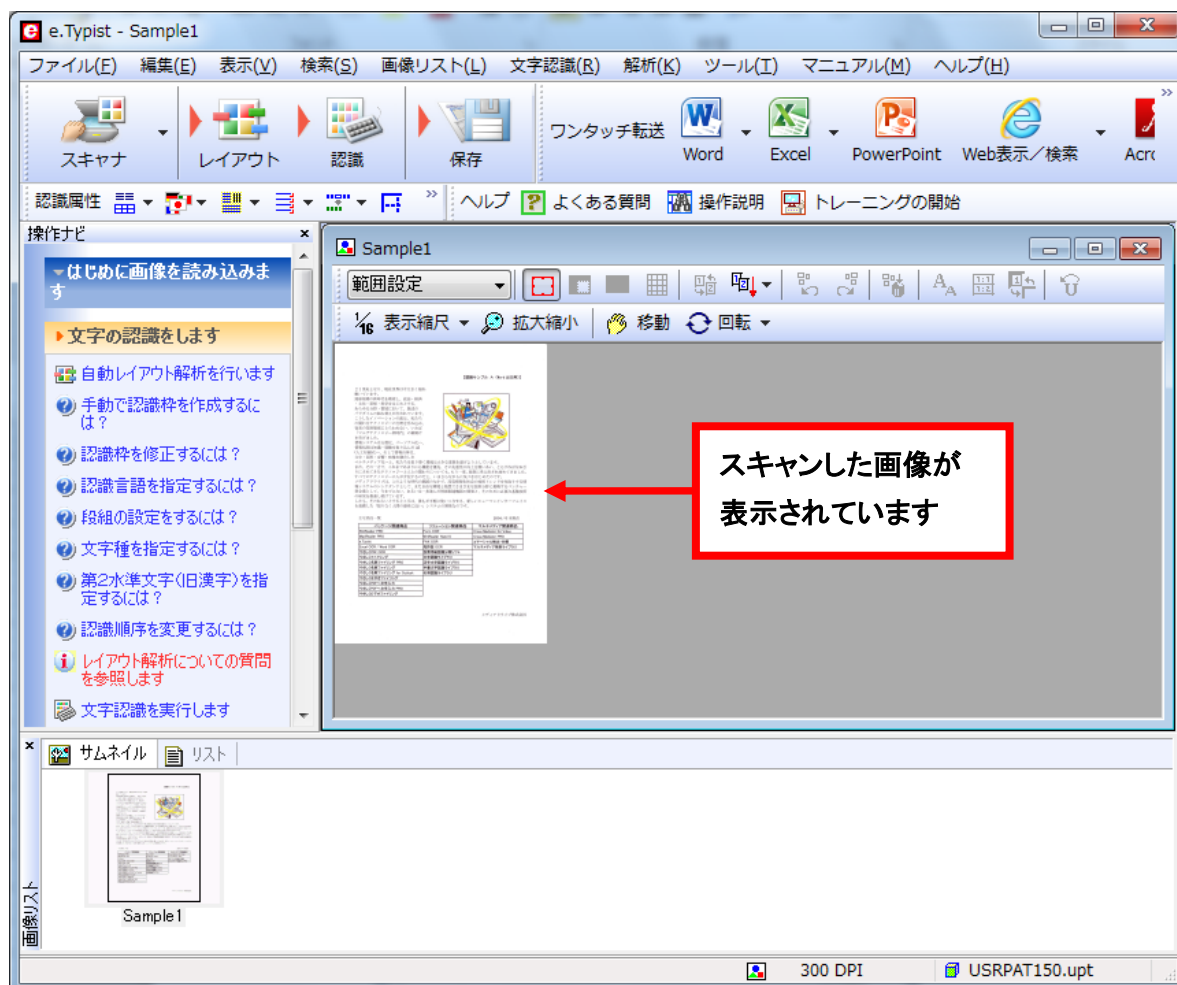
原稿台に原稿を置くときには、できるだけ、傾かないようにセットします。


傾きは、正しい認識をするために、大きく影響します。

<EPSON スキャナの例>

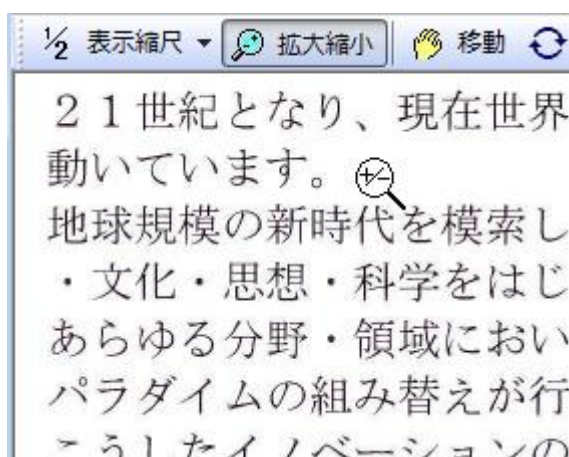


4. 「スキャン」ボタンを押すと、画像がスキャンされます。
 スキャナが止まったら、「閉じる」ボタンでスキャンパネルを閉じます。
5. スキャンした画像が e.Typist 上に表示されます。

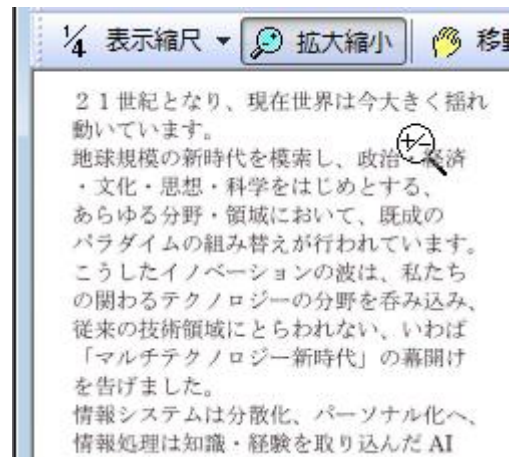


💡 画像が小さい場合には、「拡大縮小」ボタン  を使って大きく表示します。「拡大縮小」ボタンを押し、画像上でマウス左クリックすると拡大、右クリックすると縮小表示します。現在の表示縮尺は、拡大縮小ボタンの隣に表示されます。

<1/2 まで拡大>



<1/4 の表示>



● 画像ファイル・PDF ファイルから読み込む

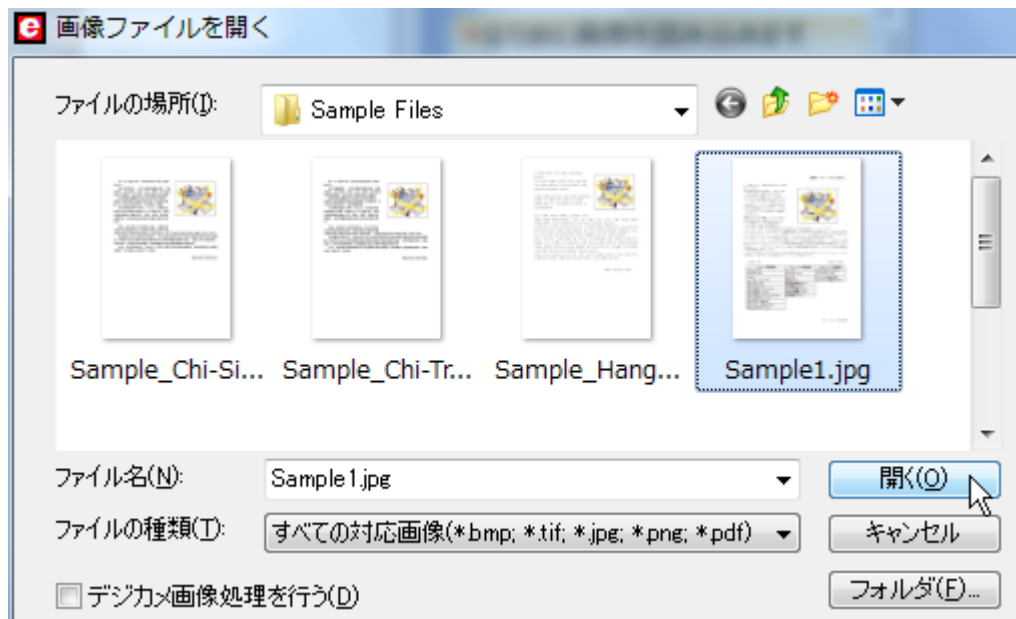
すでに PC 内に画像ファイルとして取り込まれている場合、それを直接、e.Typist に読み込むことも可能です。

たとえば、インターネット上から取り込んだ PDF ファイル、以前にスキャンして保存しておいたファイルを読む場合に使います。

1. ツールバーの「スキャナ」ボタン右側「▼」を押し、「ファイル読込」を選択します。「ファイル」ボタンに切り替わるのでこれを押します。



2. 「画像ファイルを開く」ダイアログが表示されるので、読み込みたいファイルを選択し、「開く」とすると、画像が読み込まれます。




3. 画像の状態を確認しましょう

読み込んだ画像の状態によって、うまく認識できるか、かわってきます。
どうもうまくいかない場合は、画像の状態を確認してください。

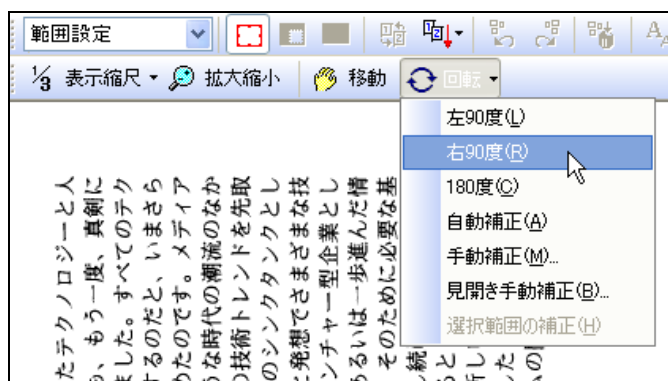
●原稿の向きは正しいですか？

原稿の向きが正しくないと、まったく認識できません。

90度横倒しになっていたり、180度逆さになっている場合は、画像の回転が必要です。

 回転・「回転」ボタンを使って正しい向きにしてください。

下図のような場合には「右90度」を選択すると正しい向きになります。

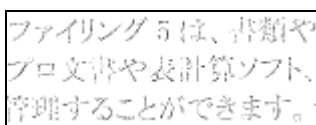


●原稿の状態は良好ですか？

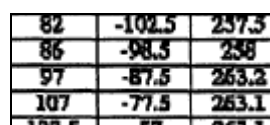
読み込んだ文字がはっきりと読める状態でないと、正しく認識できません。

下図のような場合には、濃度を調整しながら再度スキャンしてください。

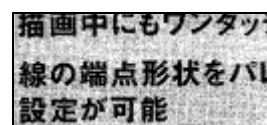
<濃度が薄い>



<濃度が濃い>

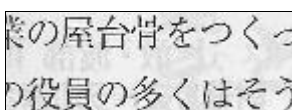


<下地の色が濃い>

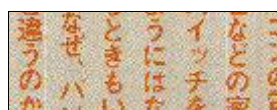


また、裏写りしていたり、下地・文字の色が濃くはつきり読めない場合も、濃度を薄くする・モノクロでスキャンするなど工夫し、再度、スキャンしてください。

<裏写り>

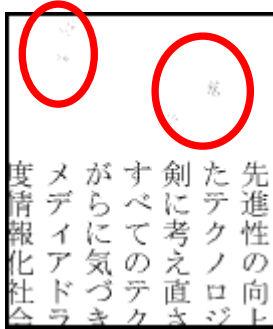


<下地と文字の色が近い>

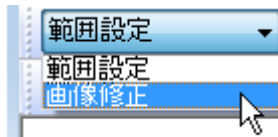


●「汚れ」はないですか？

取り込んだ画像に「ゴミ」のような汚れが入っている場合は、削除します。



1. 画像ウインドウの「範囲設定」をクリックし、「画像修正」を選択します。

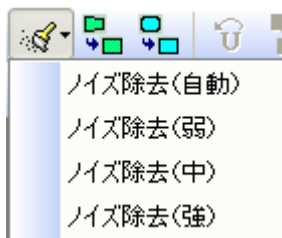


2. 画像ウインドウのツールバーのボタンが下のようになります。

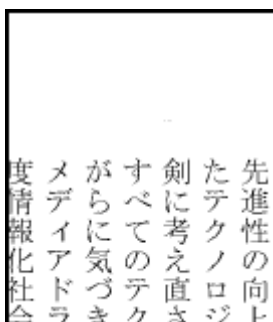


3. 「ノイズ除去」ボタンをクリックすると、除去レベルが表示されます。

除去するレベルを選択してください。



4. きれいになりました。



💡 ノイズ除去でもきれいにならない場合は、🧼「消しゴム」ボタンを使ってゴミを消してください。

また、ノイズ除去は、モノクロの画像に対してのみ、使用できます。

カラーやグレースケールの画像は、「ノイズ除去」ではなく、「消しゴム」ボタンで削除してください。

4. 認識条件を指定しましょう

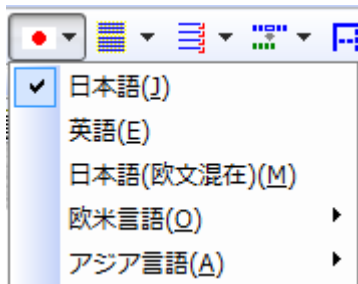
認識する前に、今回の原稿はどんな原稿か指定しておくことで、認識率がアップします。
画面左上の「認識属性」ツールバーで、認識条件の設定を行います。




●「認識言語」を指定する

原稿にあわせて、言語を設定してください。

● 「認識言語指定」ボタンをクリックすると、認識可能な言語が表示されます。



日本語内に欧米語の単語が含まれる文章が多い場合は、「日本語(欧文混在)」を選択してください。

インターネット (Internet) の普及が進み、テクノロジー (technology)

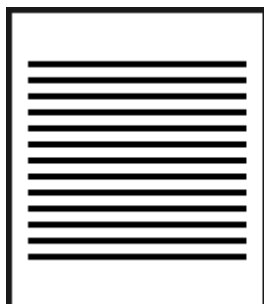
● 段組を指定する

「段組」とは、原稿のレイアウトのことです。

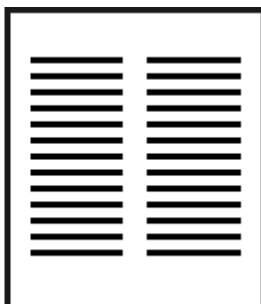
「自動」と、「横一段/横多段」、「縦一段/縦多段」があります。

簡単なレイアウトの原稿は、「自動」のままでも正しい段組として認識してくれます。

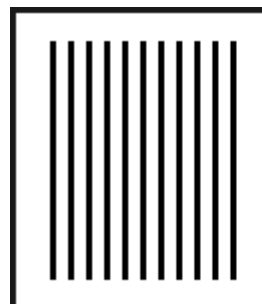
<横一段>



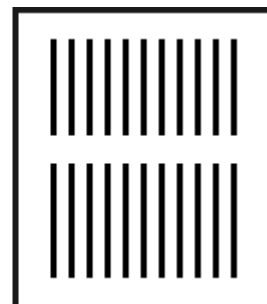
<横多段>



<縦一段>



<縦多段>



5. 認識したい範囲を指定しましょう

認識する範囲を指定します。

●自動で認識範囲を指定する

原稿全体を認識させる場合には、まず、自動で指定してみます。

「レイアウト」ボタンを押します。



画像全体に、認識枠を作成します。

認識枠は、それぞれ、色分けされます。

緑—「文章」として認識する枠です。

青—「表」として認識する枠です。

区切り位置にタブを入れるので、Excel 転送時、項目がセルで区切られます。

赤—「図」として認識する枠です。文字認識せず、そのまま出力します。

インターフェース関連商品	ソリューション関連商品	マルチメディア関連商品
WinReader PRO	Form OCR	Direct Mediator for Video
MacReader PRO	WinReader Next S	Direct Mediator PRO
sTypeit	FAX OCR	コミュニケーション特化型分類
Excel OCR / Word OCR	価格証 OCR	マルチメディア音楽ライブラリ
おしくPDF OCR	投票用紙読取りソフト	
おしくファインディング	文字認識ライブラリ	
おしく名前ファインディング PRO	漢字文字認識ライブラリ	
おしく名前ファインディング	半角文字認識ライブラリ	
おしく名前ファインディング for Outlook	名前認識ライブラリ	
おしくはがきファインディング		
おしくPDFへ文字入力		
おしくPDFへ文字入力 PRO		
おしくPDF文字入力		

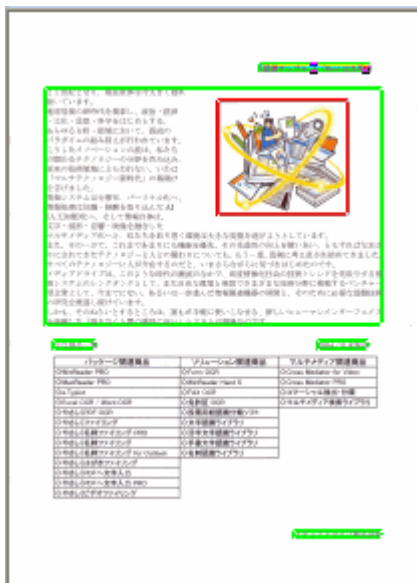
● 不要な認識枠を削除する


自動レイアウトで、不要な部分にも認識枠を作った場合、または、画像の一部だけを認識したい場合には、認識枠を削除します。

1. 削除する認識枠内をマウスでクリックします。



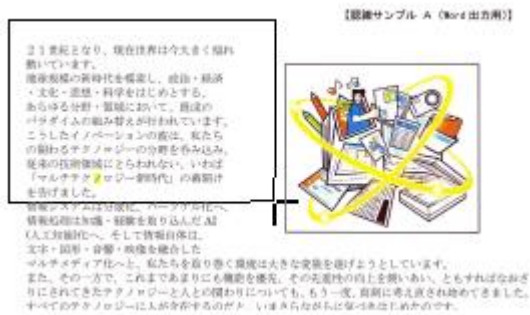
2. キーボードの[Delete]キーを押すと、認識枠が削除されます。



認識枠を一度にすべて削除する場合は、画像ウィンドウの「レイアウト取消」ボタン  を押します。確認メッセージ「全ての認識枠を削除します」を「OK」とすると、全認識枠が削除されます。

● 認識枠を手動で作成する

1. 認識したい部分を、マウスの左ボタンを押しながら、ドラッグで囲みます。



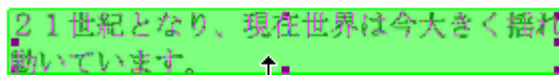
2. 囲まれた部分に、認識枠が作成されます。



● 認識枠の大きさを変更する

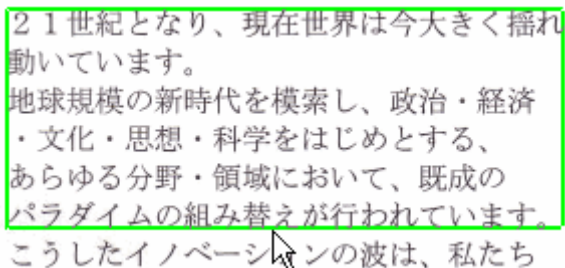
作成した認識枠の大きさを変更することができます。

1. 変更する認識枠内をマウスでクリックします。
2. 外枠にマウスカーソルを合わせます。マウスカーソルの形が変わります



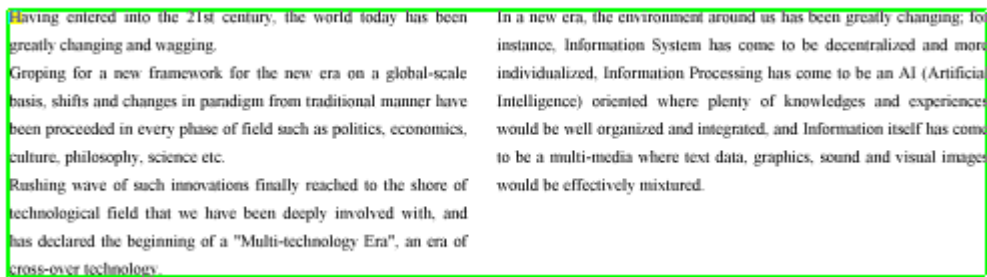
地球規模の新時代を模索し、政治・経済・文化・思想・科学をはじめとする、あらゆる分野・領域において、既成のパラダイムの組み替えが行われています。

3. 大きさを変更する方向に向かってドラッグすると、大きさを変更できます。

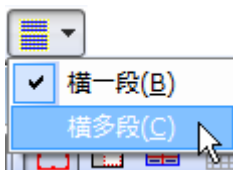


● 段組を変更する

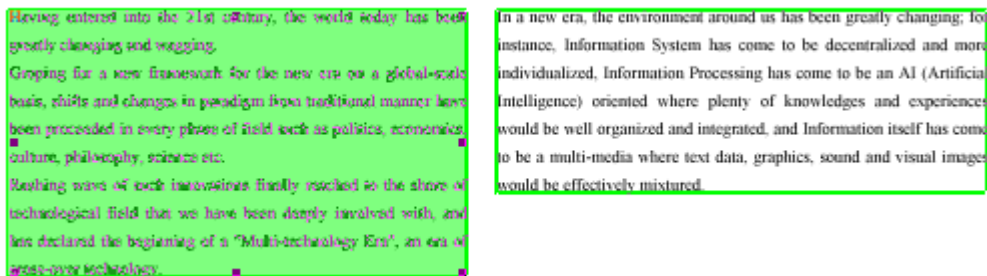
自動でレイアウトすると、多段組を「一段」に間違って判断する場合があります。特に、段落と段落の間隔が空いていない場合に起こりやすくなります。



この場合には、認識枠を選択した状態で、認識属性ツールバーの「段組」ボタンを「多段」にします。



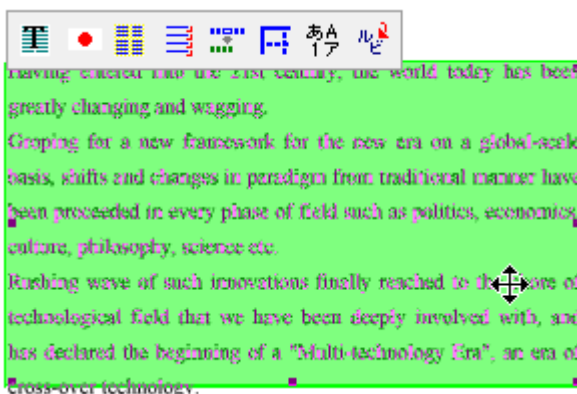
認識枠が段落ごとに自動で分割されます。



● 認識枠の設定を変更する

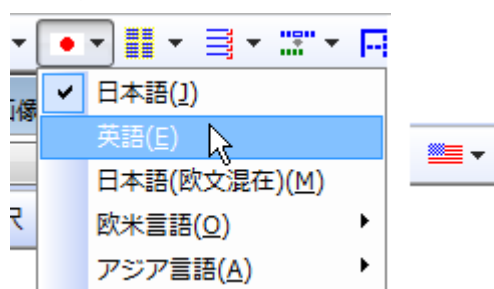
認識枠内をクリックすると、その認識枠に指定されている認識属性が表示されます。認識属性を見ると、たとえば以下は英語の文章ですが、言語が「日本語」に設定されていることがわかります。

この状態で認識すると、英語が日本語として認識されてしまいます。

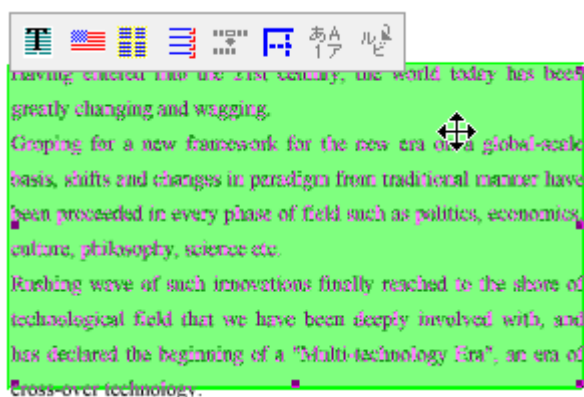



これを「英語」に直してみましょう。

枠を選択した状態で、認識属性ツールバーの言語ボタンを押し、「英語」を選択します。



言語が「英語」に変わりました。



 認識枠を作成したあとに、言語や段組を変更する場合には、必ず認識枠を選択した状態で、設定を変更してください。

認識枠が選択されていない状態では、その枠に対して変更が反映されません。

次回、新しく作った認識枠に設定が適用されてしまいます。

ご注意ください。



6. 文字認識をしましょう

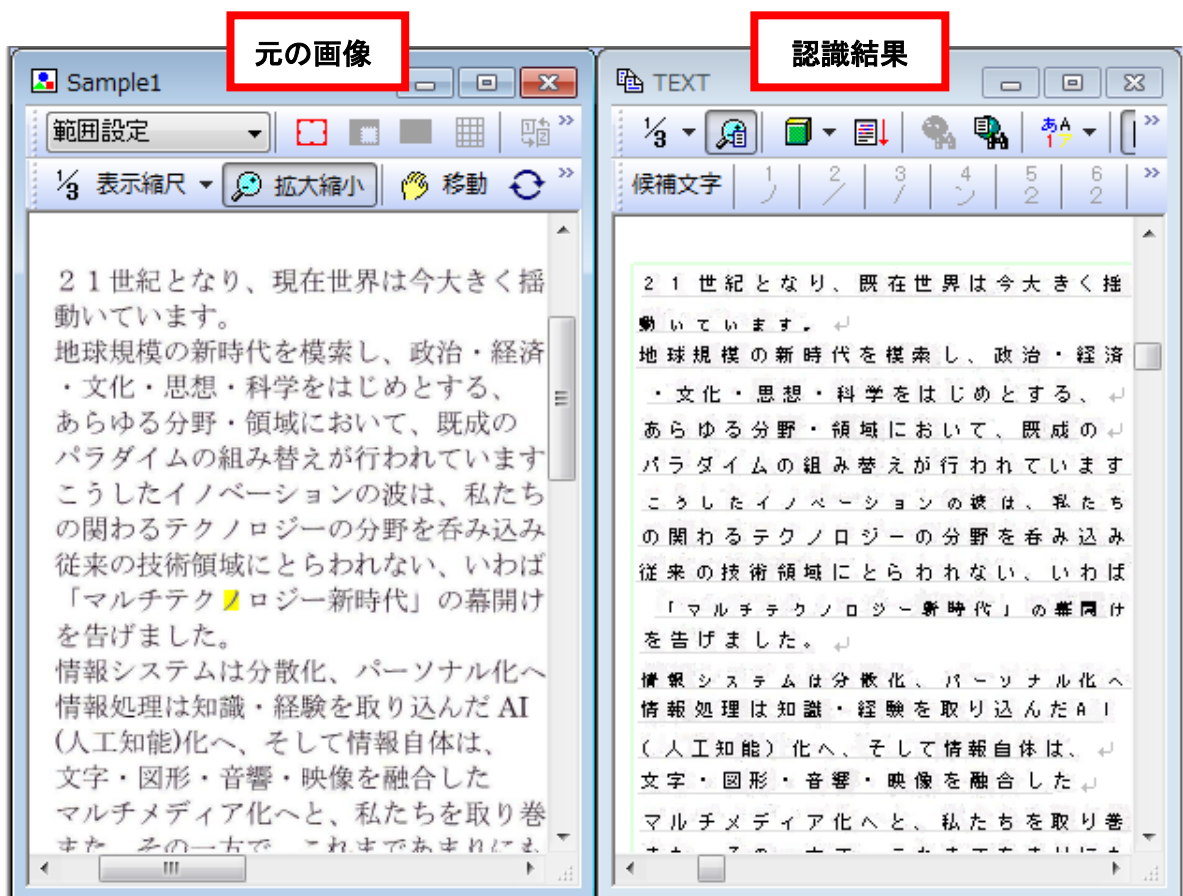
では実際に文字認識をしてみます。

「認識」ボタンをクリックします。



認識が行われて、認識結果が表示されます。

左側が「取り込んだ画像」、右側が「認識結果」です。



認識結果は元原稿のレイアウトを保った状態で表示されますが、文字を「左詰め」で表示することもできます。

「表示」メニューから「テキストレイアウト表示」を選択し、チェックをオフにします。


【認識サンプルA(Word出力用)】
21世紀となり、現在世界は今大きく揺れ
動いています。
地球規模の新時代を模索し、政治・経済

元に戻す場合は、再度、同じメニューを選択してください。

7. 認識結果を修正しましょう

認識結果が間違っていないか確認します。

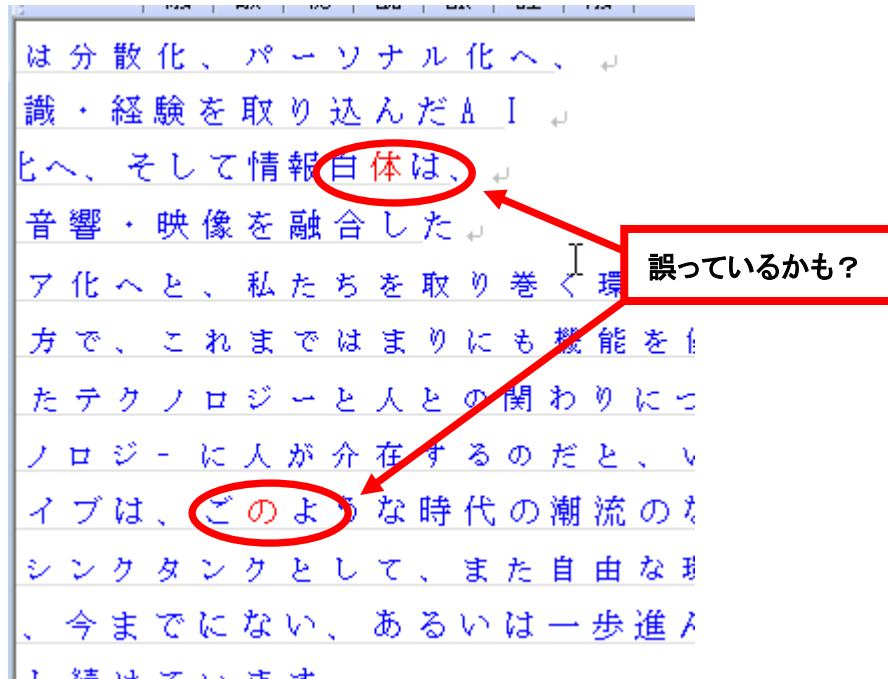
● 間違った箇所を探しやすくする

認識結果ウインドウのツールバーにある「全文解析」ボタンを押します。
認識を間違っている可能性が高い箇所が色分けで表示されます。

青い文字—正しいと判断された文字

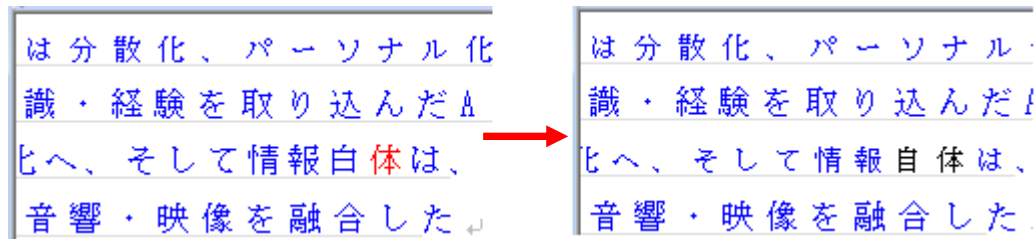
赤い文字—誤り、また前後の文字に誤りがあると判断された文字

まずは、赤い文字の辺りを中心にチェックしてください。



● キーボードから修正する

文字が間違っていたら、直接、キーボードから修正することができます。
また、不必要な文字は、Delete キーで削除することができます

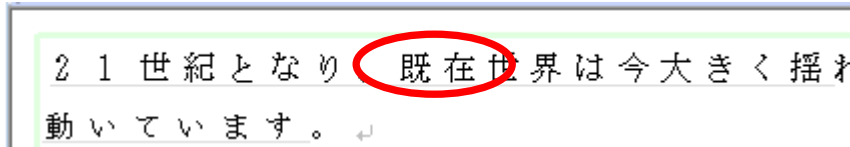


●「候補文字」で修正する

候補文字とは、認識結果として出力された文字と近い形の文字です。

間違った文字を修正する場合、候補文字の中に正しい文字があれば、クリックで選択するだけで入れ替わります。

下図の場合、「現在」の「現」を「既」と間違えています。



間違えている文字「既」にカーソルを置くと、上部ツールバーに「候補文字」が表示されます。正しい文字は「現」なので、「2 現」をクリックします。



正しい文字に変更されました。

21世紀となり、**現**在世界は今大きく揺れ動いた新時代を模索し、政治・経済・文化

- 💡 該当の文字上で右クリックしても、候補文字メニューが表示されます。ここから正しい文字を選択しても文字修正ができます。



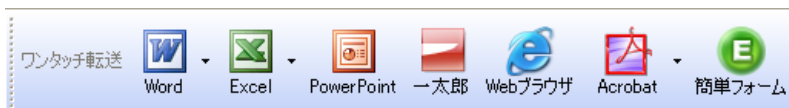
8. データを他のソフトに転送しましょう

文字認識が出来たら、他のソフトに転送することができます。使い慣れているソフトに転送し、そちらでレイアウト編集をすることもできます。

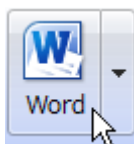
<転送できる主なソフト>

- ・マイクロソフト Office (Word/Excel/PowerPoint)
- ・Acrobat
- ・一太郎

画面上部のワンタッチ転送ツールバーには、現在、インストールされているソフトのアイコンが表示されています。転送したいソフトのアイコンをクリックすれば、文字が転送された状態でソフトが起動します。



ここでは「Word」を選択してみます。



「Wordに転送」ダイアログが表示されます。

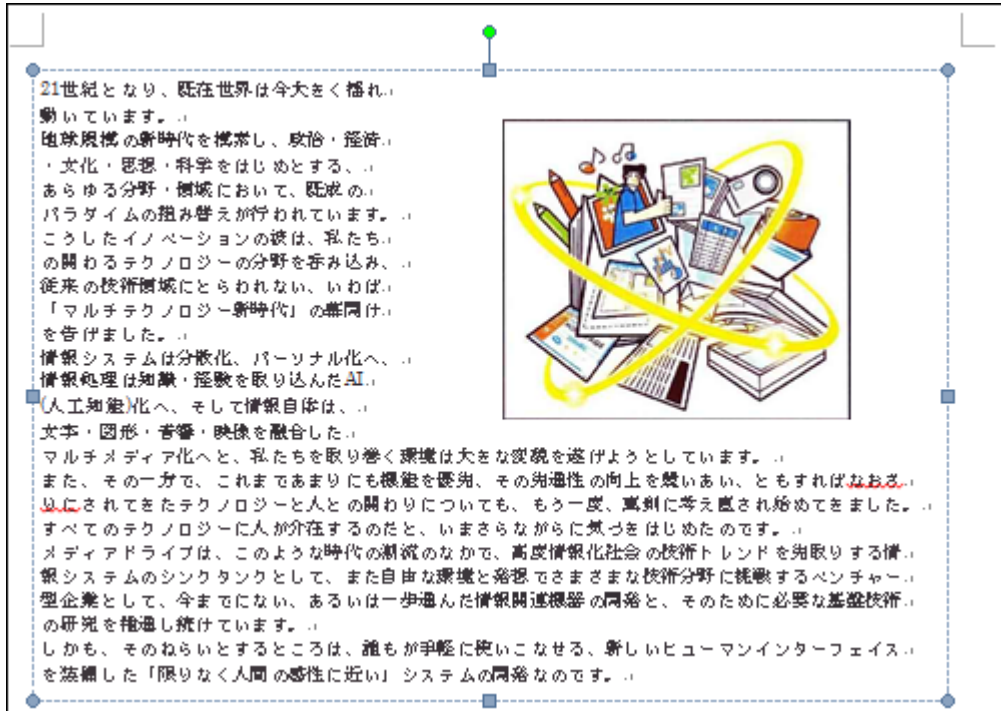


●転送モード

転送するソフトによって、いくつか転送モードが用意されています。
最初はすべて転送し、使いやすい転送方法を見つけてください。

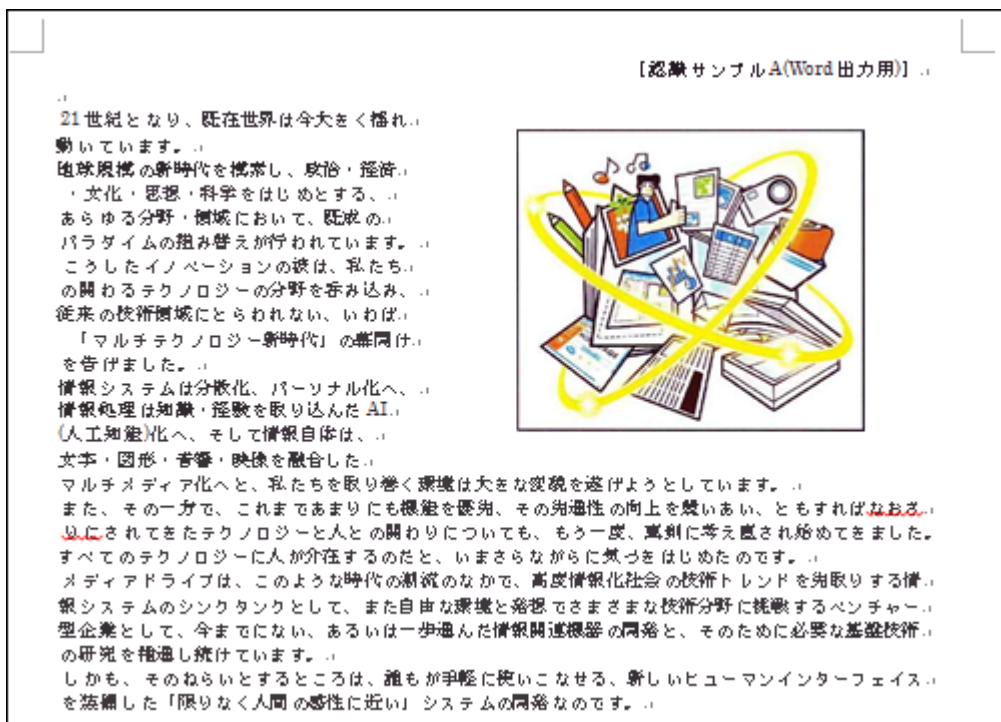
<レイアウト転送>

元原稿のレイアウトをできるだけ再現して転送します。
文章は、テキストボックスに入った状態で転送します。



<段組転送>

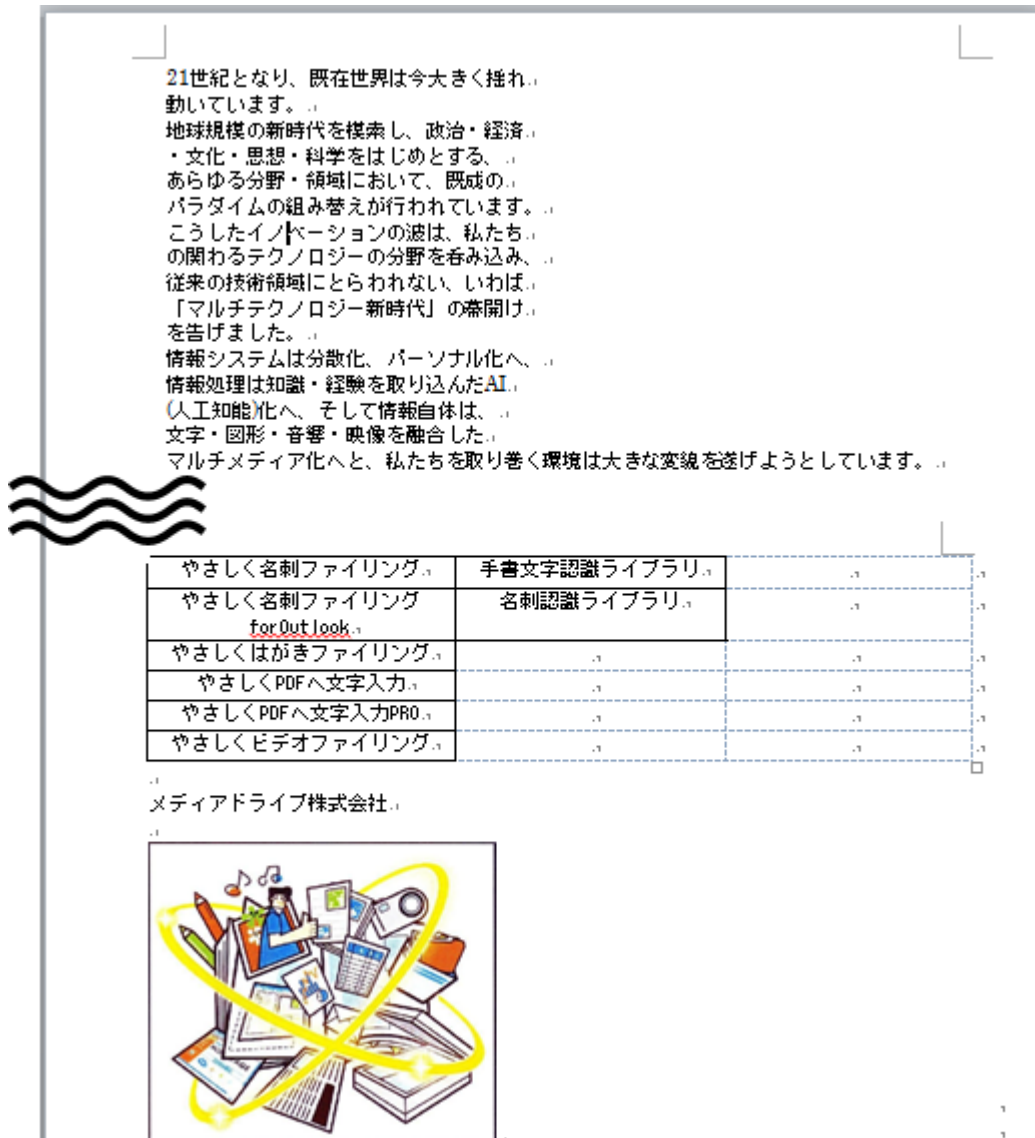
元原稿のレイアウトをできるだけ再現します。文章は、テキストボックスに入らない状態で転送します。



<テキスト転送>

段組は無視し、左詰めで転送します。


レイアウトは、転送先ソフトで作成してください。



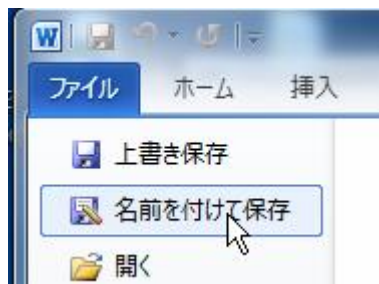
The screenshot shows a document with a text block on the left and a table on the right. The text block contains several paragraphs of Japanese text. The table has five rows and two columns. The first row contains 'やさしく名刺ファイリング' and '手書文字認識ライブラリ'. The second row contains 'やさしく名刺ファイリング for Outlook' and '名刺認識ライブラリ'. The third row contains 'やさしくはがきファイリング'. The fourth row contains 'やさしくPDFへ文字入力'. The fifth row contains 'やさしくPDFへ文字入力PRO'. The sixth row contains 'やさしくビデオファイリング'. Below the table is the text 'メディアドライブ株式会社' and an illustration of various office supplies and technology.

やさしく名刺ファイリング	手書文字認識ライブラリ
やさしく名刺ファイリング for Outlook	名刺認識ライブラリ
やさしくはがきファイリング	
やさしくPDFへ文字入力	
やさしくPDFへ文字入力PRO	
やさしくビデオファイリング	

メディアドライブ株式会社



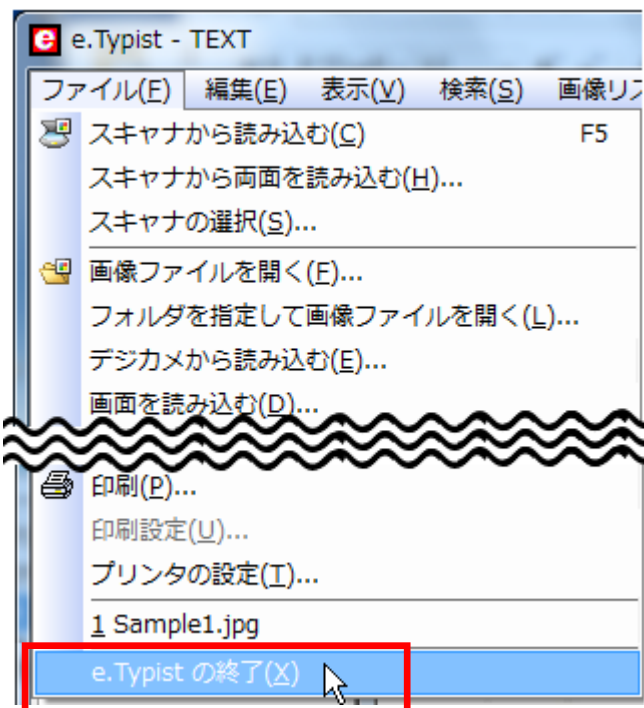
転送したデータは、普段、Word で入力したデータと同じ状態になっています。
Word で体裁を整えた後は、ファイルを保存 (docx 形式で保存) してください。



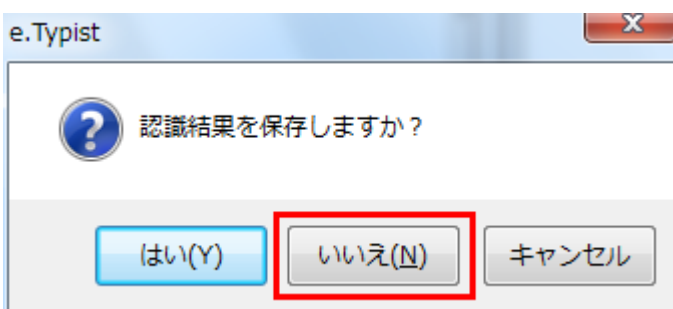
9. 作業を終わりにしましょう

データの転送・転送先の保存ができたなら、e.Typist も終わりにしましょう。

e.Typist の「ファイル」メニューから、「e.Typist の終了」を選択します。



認識結果を転送した場合、「保存しますか?」と確認メッセージが表示されます。すでに Word など別ソフトに転送し、保存されているので、e.Typist に残っている認識結果を削除しても問題ありません。



これで e.Typist が終了します。



10. おわりに

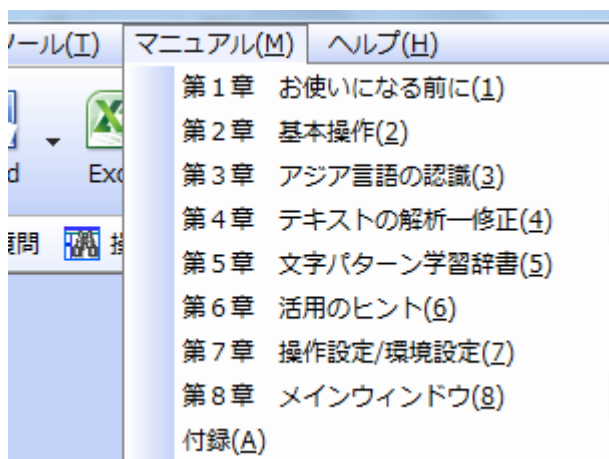
以上で「かんたん操作マニュアル」は終了です。

e.Typist がどんなソフトか、ご理解いただけましたでしょうか。

e.Typist には、ここで説明した以外にも、まだ便利な機能が用意されています。

「環境設定」や「操作設定」の項目をチェックしたり、メニューから表示を切り替えてみるなど、自分にあった使い方を探してみてください。

また、「ここも自動でできないのかな?」「いつも同じ文字を間違ってしまう、何か修正方法がないかな?」「表は認識できないの?」などの疑問があれば、e.Typist の「マニュアル」メニューから、詳細なマニュアルをご覧ください。



「こんな機能があったらいいな」など要望、この簡易マニュアルへのご意見がありましたら、メールにてご連絡ください。



メール … online-support@mediadrive.co.jp

それでは、e.Typist を使って、ますます便利に PC をご活用ください。

