

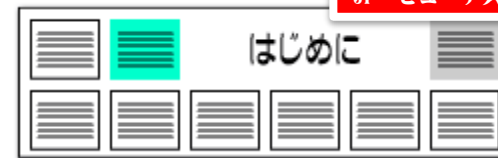
# 【 準備しましょう。】

---

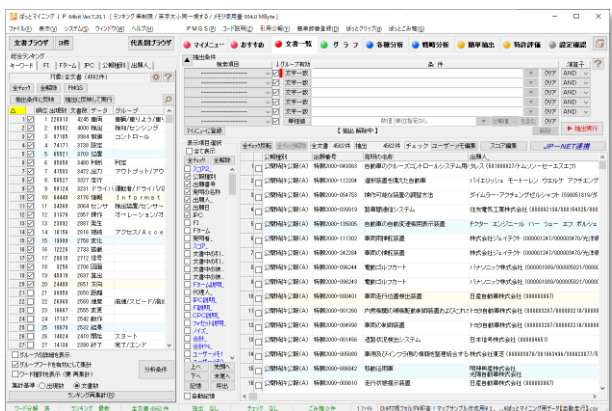
- ☑ ぱっとマイニング JPの準備



ちょっと待って！ **使う前の設定**はこれだけ！



▶ システム設定って何？



メインウィンドウ

メインウィンドウからの環境設定メニューには、ソフトの動作のための設定が集まっています。

文字の色や、形態素分析を利用するかどうか、追加単語辞書のメンテナンスなどです。

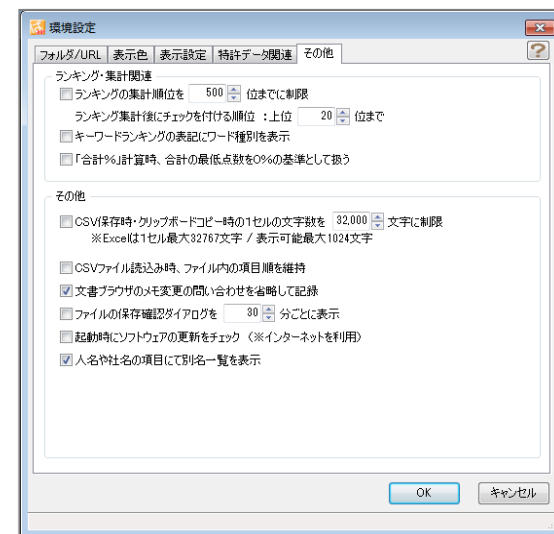
すぐソフトを使いたい場合には、**何も変更する必要はありません。**

▶ メインウィンドウの『システム』から



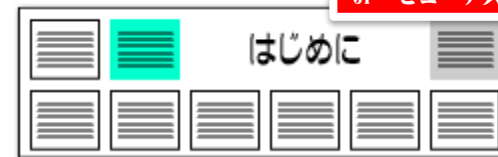
気になる場合には、メインウィンドウの [システム] の [環境設定] をクリックし、どのような設定があるのか、見ておくとよいでしょう。

▶ こんな画面です。



それぞれの設定は、別途ご説明します。

# 😊 それでも**気になる設定**はコレ！ 1



## ▶ その他の設定 1 ワード分解の だけは気をつけよう！

戦略分析 簡単抽出 特許評価 設定確認

ワード分解設定

B. ワード分解時の処理

※ ファイルを読み込み後に行う「ワード分解処理」に対応するものです。  
設定を変更すると、再度ワード分解を行う必要があります。

ワード分解を使用

辞書を使用せずにワード分解を行う

未知ワードの先頭が除去用ワードに該当すれば分解

未知ワードの末尾が除去用ワードに該当すれば分解

ワード分解の完了時に総合ランキングを自動集計

英字・ひらがな・カタカナの大文字と小文字を同一視  
(※ 数値 & 単位検索には影響しません)

漢字・カタカナ・英文字の間での分解

1. 漢字・カタカナ・英文字の前後で分解  
例: 「ICカード端末」を「IC」「カード」「端末」として分解する。

2. 英文字の前後のみ分解 (漢字・カタカナの連続をひとつのワードと見なす)  
例: 「ICカード端末」を「IC」「カード端末」として分解する。

除去用ワード辞書設定

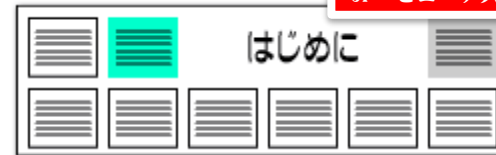
### 【注意】 「ワード分解を使用する」のチェックの意味

チェックを外すと、ワード分解を利用している機能は使えなくなります。ワード分解していないと、文章の言葉をバラバラにしないので、「言葉 = ワードを基にした分析」、たとえばワードのランキング機能や、ワードを利用したグラフ表示などができなくなります。

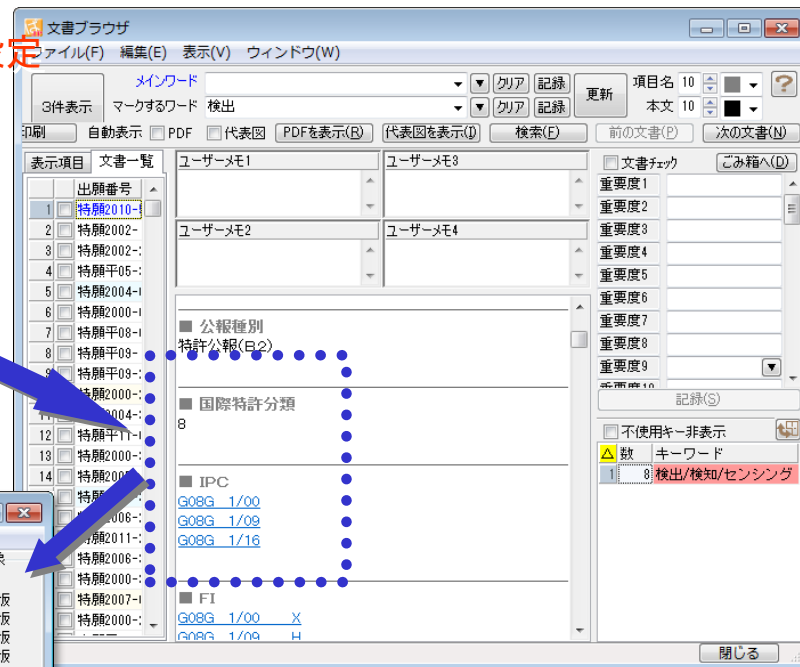
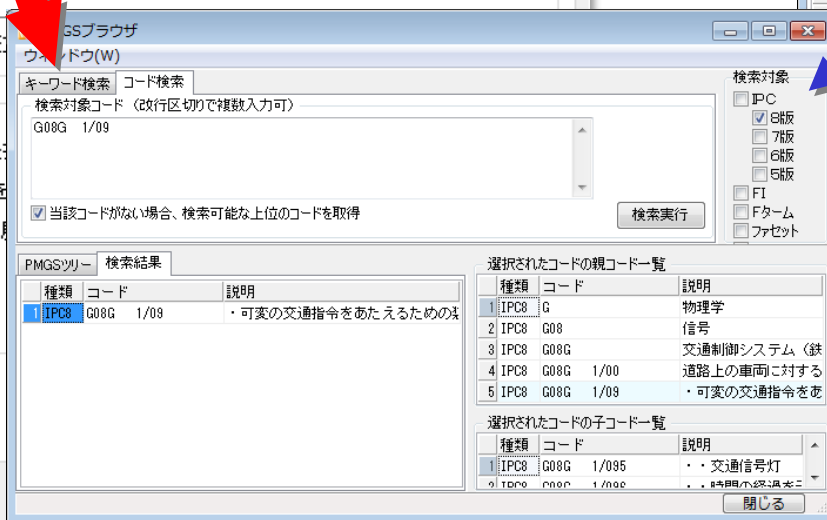
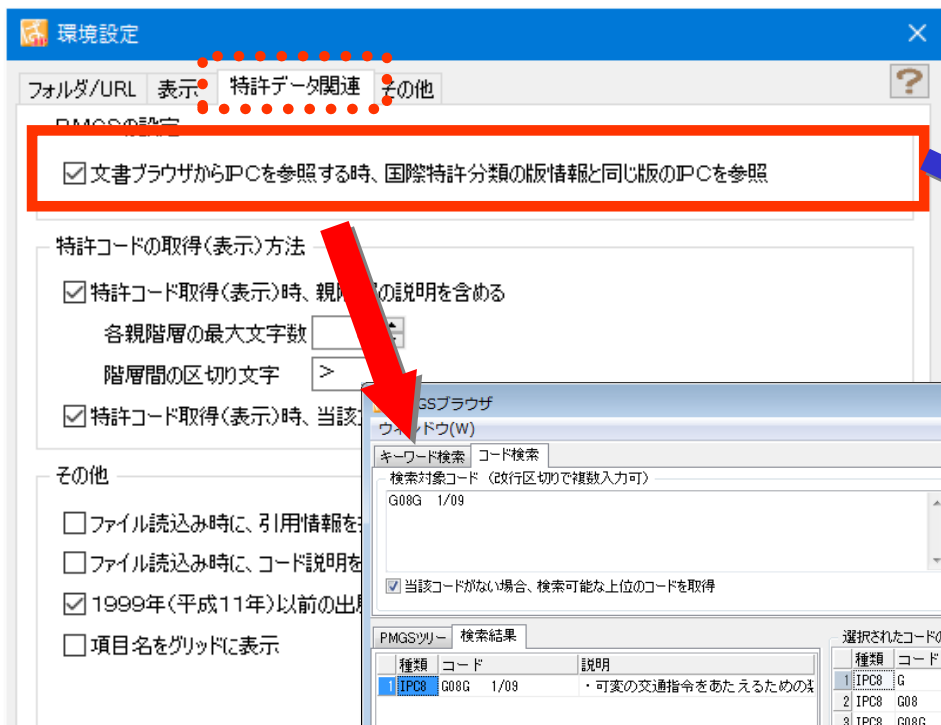
では、なぜ「ワード分解を利用しない」というモードがあるのでしょうか。「ワード分解」= 言葉をバラバラにするには、コンピュータに大変な負荷が一時的にかかります。ファイルの読み込み時に、読み込むファイルが大きいほど、この時間が増大します。

ワードを利用した分析やグラフ描画の必要のない業務、「査読業務」「文書のソート・抽出業務」に【ばっとマイニングJP】を利用する場合、このチェックをはずしておいたほうが、快適に利用できるということになります！

# それでも気になる設定はコレ！ 2



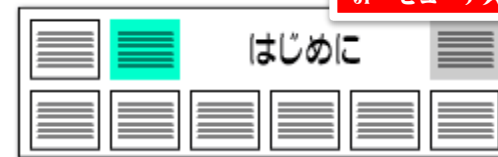
## ▶ 特許データ関連の設定 1 PMGS (パテントマップガイド) の設定



文書ブラウザのIPCコードをクリックすると、PMGSブラウザーが起動します。

このとき、上図の例であれば、IPC8版のコードだけを表示するか、それとも全版の該当コードを表示するかを、環境設定で指定しておくことができます。

# 😊 それでも気になる設定はコレ！3



## ▶ 特許データ関連の設定2 特許コード・公報番号の設定

環境設定

フォルダ/URL 表示色 表示設定 **特許データ関連** その他

PMGSの設定

文書ブラウザからIPCを参照する時、国際特許分類の版情報と同じ版のIPCを参照

特許コードの取得(表示)方法

特許コード取得(表示)時、親階層の説明を含める

各親階層の最大文字数

階層間の区切り文字

特許コード取得(表示)時、当該コードの情報がない場合は【該当コードなし】と表示

その他

ファイル読み込み時に、引用情報を抽出

先行再公表を再公表として扱う(※要再起動)

ファイル読み込み時に、コード説明を取得

1999年(平成11年)以前の出願・公報番号を和暦で表示

項目名をグリッドに表示

JP-NET / NewCSS インストールフォルダの場所

JP-NETがインストールされています

OK キャンセル

クリック

『特許コード取得(表示)時、親階層の説明を含める』…チェックを入れると、特許分類コードの説明表示に、親階層の説明を含めることができます。説明表示をシンプルにしたい場合はチェックを外します。

各親階層の最大文字数、階層間の区切り文字を指定することができます。

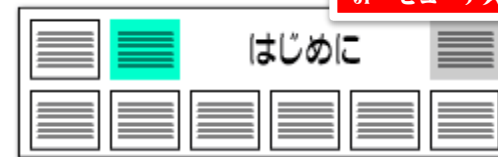
『特許コード取得(表示)時、当該コードがない場合は【当該コードなし】と表示』…当該コードの説明がない場合に、空白表示にするか、【当該コード無し】と表示するかを選択することができます。

“先行再公表”(再公表特許発行前のPCT出願:JP-NETの独自種別)の公報を、後に発行される“再公表”と同様に扱う設定とするチェックボックスです。

“再公表”公報が発行されると、同じ出願内容にもかかわらず“先行再公表”公報とは別の公報としてぱっとマイニングJPで認識されるという事例があったため、“先行再公表”の公報も“再公表”公報とみなし、別々のものとして扱わないように設定できる機能を追加しました。

ぱっとマイニングJPでは、データ読み込み時に、出願番号、公報番号をすべて西暦表示に統一します。JP-NETデータのままだの出願番号、公報番号にしたい場合は、ここにチェックを入れることで、平成11年以前の出願番号、公報番号を和暦表示に変更します。

# それでも気になる設定はコレ！4



## ▶ ワード分解設定

## ワード分解やファイル読み込み時の設定を容易に確認できる

マイメニュー ● おすすめ ● 文書一覧 ● グラフ ● 各種分析 ● 戦略分析 ● 簡単抽出 ● 特許評価 ● 設定確認

● 設定一覧 ● キーワードランキング分析条件 ● 辞書一覧 ● **ワード分解設定**

**A. 文書データの読み込み時の処理**  
 ※ ファイルの「読み込み時の処理」に対応するものです。  
 「読み込み済みのデータ」は影響を受けないので、  
 変更後に再度読み込みを行う必要があります。  
 ※ 「読み込み時に整形されるため、「整形前の状態」に戻すことはできません。

スペースの取扱い  
 そのまま扱い、変換処理を行わない  
 スペースを全て除外  
 連続するスペースを一つに纏める(推奨)

長音符とマイナス(ハイフン・ダッシュ)の変換  
 そのまま扱い、変換処理を行わない  
 すべて長音符に変換  
 すべてマイナスに変換  
 仮名文字の後は長音符、その他はマイナスに変換

読み込み時に除去する文字 (改行で複数指定可)

**B. ワード分解時の処理**  
 ※ ファイルを読み込み後に行う「ワード分解処理」に対応する設定を変更すると、再度ワード分解を行う必要があります。

ワード分解を使用  
 辞書を使用せずにワード分解を行う(特殊用途)  
 未知ワードの先頭が除去用ワードに該当すれば分解  
 未知ワードの末尾が除去用ワードに該当すれば分解  
 ワード分解の完了時に総合ランキングを自動集計  
 英字・ひらがな・カタカナの大文字と小文字を同一視  
 (※ 数値も単位検索には影響しません)

漢字・カタカナ・英文字の間での分解  
 1. 漢字・カタカナ・英文字の前後で分解  
 例: 「ICカード端末」を「IC」「カード」「端末」として分解する。  
 2. 英文字の前後のみ分解(漢字・カタカナの連続をひとつのワードと見なす)  
 例: 「ICカード端末」を「IC」「カード端末」として分解する。  
 3. 分解しない(漢字・カタカナ・英文字の連続をひとつのワードと見なす)  
 例: 「ICカード端末」をそのままひとまとまりとする。

漢字・カタカナ が連続している文字列(文字の連続)のワード分解方法  
 辞書に登録されたワードの「直前」で、分解  
 例: 「ジメチル」が未登録、「メチル」が辞書に登録済の場合  
 「ジ」「メチル」と分解する。  
 辞書に登録されたワードの「直後」で、分解  
 例: 「ハイドロゲン」が未登録、「ハイドロ」が辞書に登録済の場合  
 「ハイドロ」「ゲン」と分解する。

**C. 分析条件等の入力時の処理**  
 アンダーバーと空白、空白除去を同一視 「\_」「 |」  
 ハイフンと長音記号を同一視 「-」「ー」  
 中黒とスラッシュ記号を同一視 「・」「/」

初期値に戻す 設定読み込 設定保存 適用 キャンセル

クリックすると、「設定読み込」や「設定保存」を行なうことができます。

クリック

**A. 文書データ読み込み時の処理**

**B. ワード分解時の処理**

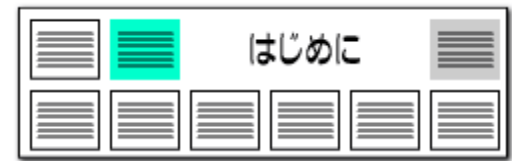
除去用ワードの設定

**C. 分析条件等の入力時の処理**

- 『ワード分解設定』は、以前のバージョンでは『環境設定』の中にありましたが、分析の根幹に関わる設定ですので、いつでも確認しやすいように、『設定確認』タブ内に移動させました。
- 公報データの中には、複数のスペースが連続していたり、長音符「ー」とハイフン「-」が混在していることがあります。「インターネット」と「インターネット」のような状態です。これらのスペースや長音符、ハイフンの扱いをここで設定することができます。
- 「ワード分解時の処理」では、チェックに合わせて、未知ワードを、ワード辞書に登録されているワードに分解します。文字数の多い薬品名などが複数の短い単語に分解される場合などがあるため、状況に応じてチェックをはずしてください。設定を変更した場合は、「適用」ボタンを押すと、チェック状態の変更が反映されます。
- 「初期値に戻す」ボタンをクリックすると、ワード分解設定を初期状態に戻します。
- 辞書を使用せずにワード分解を行なうと、文書中の分解の設定が難しい特殊なワードのかたまりを見つけることもできます。
- 除去用ワードを設定すると、「前記車両」の「前記」などの特許特有の表現を除去します。辞書不使用時の分析にも便利な機能です。
- 『設定保存』のボタンで、この画面の辞書設定を保存することが出来ます。保存したファイルは『設定読み込』のボタンで選択し、反映することが出来ます。



# それでも**気になる設定**はコレ！5



## ▶ 重要度の設定

重要度は情報を仕分けるための分類フラグです。

分析目的や社内の分類ルールなどに合わせて、自由に設定していただくことができます。

重要度の設定ファイルはCSV形式で保存されるので、社内でも共有するなどの利用方法が可能です。

ひとつの重要度の中で複数の項目を選択することができる

重要度をCSVファイルとして保存しておくことができるので、分析内容に合わせて、重要度設定を使い分けることができる！

優先順	行数	ディレク...	ファイル名	変更
<input checked="" type="checkbox"/>	50	ユーザ...	ユーザー・重要度設定(自動運転分析).csv	
<input type="checkbox"/>	50	ユーザ...	ユーザー・重要度設定.csv	
<input type="checkbox"/>	50	ユーザ...	ユーザー・重要度設定(ダウンロード分析).csv	
<input type="checkbox"/>	50	ユーザ...	ユーザー・重要度設定(ABS分析).csv	

ファイル名や保存場所は自由に設定できる

表示名や重要度の内容は自由に変更することができる

昇順／降順の並び替えができる

設定ファイルに未登録の重要度がデータ内にある場合、このボタンで抽出できる

ソート

候補追加

クリック

適用

適用して閉じる

キャンセル