

2015年12月4日



パーソナルメディア株式会社

〒141-0031 東京都品川区西五反田 1-29-1 コイズミビル

TEL.03-5759-8305 FAX.03-5759-8306

E-mail: sales@personal-media.co.jp

http://www.chokanji.com/ckk/

## 見分けにくい異体字をラクラク入力できる漢字入力ツール 「超漢字検索 異体字パーツナビ」を新発売

ソフトウェアメーカーのパーソナルメディア株式会社(代表取締役:松為彰、本社:東京、電話:03-5759-8305、資本金 1,000 万円)は、見分けにくい異体字もラクラク区別して入力できる「超漢字検索 異体字パーツナビ」の開発に成功し、自治体や官公庁を含む法人のお客様向けのカスタム製品として、2016年2月1日より出荷を開始します。本製品は43,000文字以上の異体字に対応しており、電子行政サービスや各種の業務システムにおいて人名用漢字を入力する際の便利なツールとしてご利用いただけます。また、12月9日(水)～11日(金)に東京ミッドタウンで開催される「2015 TRON Symposium (TRONSHOW)」のパーソナルメディアのブース(ブース番号:B-5)にて本製品のデモンストレーションを行います。

人名用漢字には多くの異体字があり、名簿の作成や電子政府の端末操作などの場面では、これらの異体字を効率よく識別、同定して入力する必要があります。ところが、異体字の中には「渡邊」の「邊」のように100文字以上の異体字が存在するものもあり(図1)、それらの異体字間の字形の違いが分かりにくいために、文字の識別や同定、入力が容易でない場合があります。「超漢字検索 異体字パーツナビ」は、このような異体字の検索や入力を効率よく行うために、当社が新しく開発したツールです。

「超漢字検索 異体字パーツナビ」では、漢字を構成する偏(へん)、旁(つくり)などの部首や部品のうち、異体字間で差異のある部品を絞込み検索キーとして表示します。それらの絞込み検索キーの中から、入力したい異体字の字形に一致するものをユーザが指定することにより、異体字の絞込み検索が行われ、目的の異体字をすばやく見つけ出すことができます。たとえば、「葛」の異体字の場合、3画あるいは4画の「くさかんむり」を含む異体字があり、また下部に「ヒ」「メ」「入」を含む異体字があります。これらの部品のいずれかを絞込み検索キーとして指定することにより、多くの「葛」の異体字の中から、入力したい異体字を容易に特定できます(図2)。

また、本製品は読めない漢字でもすばやく検索して入力できる総合文字検索ツール「超漢字検索 Web版(\*1)」の機能を含んでおり、漢字の読みや画数はもちろん、部首や一部の部品、それらの組み合わせなどを用いて、分かりやすく直感的な操作により目的の漢字を検索できます。さらに、情報処理推進機構(IPA)が整備した「IPAmj 明朝フォント(\*2)」と Unicode IVS/IVD(\*3) の “Mojji\_Joho” コレクションに対応しており、検索した異体字を IPAmj 明朝フォントで表示したり、他の Web アプリケーションに貼り込んだり、検索した異体字の MJ 番号(MJ 文字図形名)を確認することが可能です。

今秋から始まったマイナンバー制度のメリットを活用していくには、人名などの固有名詞を含むパーソナル情報の電子化やオンライン化が不可欠です。人名用漢字として必須の異体字を効率よく入力できる「超漢字検索 異体字パーツナビ」は、そのための有効なツールとなります。パーソナルメディアでは、本製品や既存の超漢字検索関連製品に加えて、お客様の外字を含めて検索するカスタマイズ版の超漢字検索や、外字の同定、整理、変換などを行うツールの開発や販売を通じて、お客様のご要望に応じた多様な漢字ソリューションを提供していきます。

- (※1) 「超漢字検索 異体字パーツナビ」で扱える文字は「超漢字検索 Web 版」と共通であり、その総数は 186,487 字です。ただし、この文字数は、複数の文字コード規格に含まれる同一字形の文字を重複して数えています。同一字形の文字の重複を除いた場合の総文字数は、約 10 万字です(同一字形の解釈によりこの数は変動します)。この中には、大漢和辞典に収録されている約 5 万字のほか、住基ネット用の文字を含む多数の人名用異体字、変体仮名、中国簡体字や中国伝統字、ハングル、世界各国の文字や記号などが含まれます。
- (※2) IPAmj 明朝フォントは、経済産業省の委託事業である「文字情報基盤構築に関する研究開発事業」の成果として、IPA から無償で公開されている漢字の文字フォントです。「戸籍統一文字」の 55,270 字や「住民基本台帳ネットワークシステム統一文字」の 19,563 字など、約 6 万字の漢字を含んでいます。
- (※3) Unicode IVS/IVD とは、Unicode において漢字の異体字をコード化して扱うための規格です。異体字を示す字形の一覧が IVD(Ideographic Variation Database: 漢字字形データベース)として登録、公開されており、テキストデータ中の異体字の区別を IVS(Ideographic Variation Sequence: 異体字シーケンス)によって指定します。IVD には、文字情報基盤に対応した “MojiJoho” などのコレクションが登録されています。
- ※ 本製品には、東京都中小企業振興公社の実施する平成 26 年度新製品・新技術開発助成事業の採択案件「異体字入力用の漢字検索ライブラリの開発」の研究成果の一部を利用しています。

## 補足資料

### ■「超漢字検索 異体字パーツナビ」の特長

- 人名などに使用される多数の異体字の中から、入力すべき目的の漢字を効率よく検索して確定し、Web アプリケーション等の画面に入力できます。名簿の作成や電子政府の端末操作などの場面において、作業負荷の軽減や作業時間の短縮、誤りの低減が可能となります。
  - 漢字を構成する偏(へん)、旁(つくり)などの部首や部品のうち、異体字間で差異のある部品のみを絞込み検索キーとして表示します。ユーザは、それらの絞込み検索キーを指定することによって、最小限の確認操作で目的の異体字を特定できます。
  - 「邊」や「辺」の異体字 122 文字、「荒」の異体字 44 文字、「齋」の異体字 41 文字などを含めて、43,000 文字以上の異体字に対応しています。
  - 異体字の検索機能は WebAPI(※4)として提供されますので、住民票管理、戸籍管理、会員名簿のデータベース管理など、各種の業務用 Web アプリケーションの中に組み込んで利用することができます。
  - 情報処理推進機構(IPA)が整備した「IPAmj 明朝フォント」と Unicode IVS/IVD の “MojiJoho” コレクションに対応しており、文字情報基盤の成果を活用したシステムを構築できます。
  - 異体字の検索機能に加えて、累計 10 万人以上のユーザを持つ総合文字検索ツール「超漢字検索」の漢字検索機能が利用できます。漢字の読みや画数はもちろん、部首や一部の部品、それらの組み合わせなどを用いた、直感的で分かりやすい操作による漢字の検索が可能です。
- (※4) ウェブサイトや Web アプリケーションで使用する機能を提供するためのプログラミングインタフェースです。API は Application Programming Interface の略です。

## ■「超漢字検索 異体字パーツナビ」の異体字検索機能

「超漢字検索 異体字パーツナビ」では、漢字を構成する偏(へん)、旁(つくり)などの部首や部品のうち、異体字間で差異のある部品を絞込み検索キーとして表示します。絞込み検索キーは、異体字間で差異のある部品に限られており、すべての異体字に共通して含まれる部品は絞込み検索キーとして表示されません。そのため、「超漢字検索 異体字パーツナビ」の表示する絞込み検索キー自体が、異体字の字形の違いを確認する際に注目すべき部分を表わしています。

たとえば、「葛」は「日」という部品を含んでいますが、「葛」のすべての異体字が共通して「日」を含んでいるため、「日」は異体字の絞込み検索キーとしては表示されません。「葛」の異体字の検索では、3画と4画の「くさかんむり」や下部の「ヒ」「メ」「入」などの部品が絞込み検索キーとして表示されます。これらの部品は、「葛」の異体字間で字形の異なる部分であり、異体字を見分ける際に注目すべきポイントとなっています。

また、絞込み検索キー、すなわち異体字間で差異のある部品は、候補となる異体字が絞り込まれるにつれて減少し、残った異体字を区別するために不必要な情報は表示されません。このため、ユーザは最小限の確認操作によって目的の異体字を特定することが可能です。

たとえば、「葛」の異体字を3画の「くさかんむり」で絞り込んだ場合、残った「葛」の異体字はすべて3画の「くさかんむり」を共通に含んでいますので、「くさかんむり」の部分にはもう字形の違いがありません。そのため、3画の「くさかんむり」で絞り込んだ後は、絞込み検索キーから3画と4画の「くさかんむり」が消え、それ以外の部分で字形の違いを表わす「ヒ」「メ」「入」などの部品のみが絞込み検索キーとして残ります。ユーザは、これらの残った部品に対応した字形の違いを確認することによって、目的の異体字を効率よく特定していくことができます。

「邊」の場合は、異体字が122文字あるため、この中から目視だけで目的の異体字を見つけ出すのは大変ですが、絞込み検索キーによる異体字検索機能を使えば、「邊」の異体字の入力もスムーズに行えます。たとえば、「邊」の異体字122文字(図3)に対して、「方」で絞込み検索を行うと、候補の数は35文字となります(図4)。ここからさらに2点しんじょうで絞込み検索を行うと、残った候補の数は17文字となります(図5)。この中からであれば、目的の異体字を目視で選ぶことも比較的容易ですし、さらに別の部品を使って絞込み検索を続けることもできます。「邊」の異体字の中には、1点しんじょうや2点しんじょうを含むもの、「口」を含むもの、「方」を含むものなどがありますので、これらの部品を使って絞込み検索を行い、候補となる異体字の数を大幅に減らしていくことによって、目的の異体字を効率よく見つけ出すことができます。

## ■参考情報、リンク集

文字情報技術促進協議会 (旧: IVS 技術促進協議会)

<http://citpc.jp/>

IPA 文字情報基盤整備事業と IPAmj 明朝フォント

<http://mojikiban.ipa.go.jp/>

超漢字検索 Web 版

[http://www.chokanji.com/ckk/press\\_ckkweb.html](http://www.chokanji.com/ckk/press_ckkweb.html)

超漢字検索 文字情報基盤対応版

<http://www.chokanji.com/ckk/ckkmj.html>

超漢字検索ウェブサイト(Windows 版、Linux 版、iOS 版、Android 版)

<http://www.chokanji.com/ckk/>

パーソナルメディア株式会社

<http://www.personal-media.co.jp/>

## ■ 図表

図 1. 「超漢字検索 異体字パーツナビ」の画面例



図 3. 「邊」の異体字(122 文字)

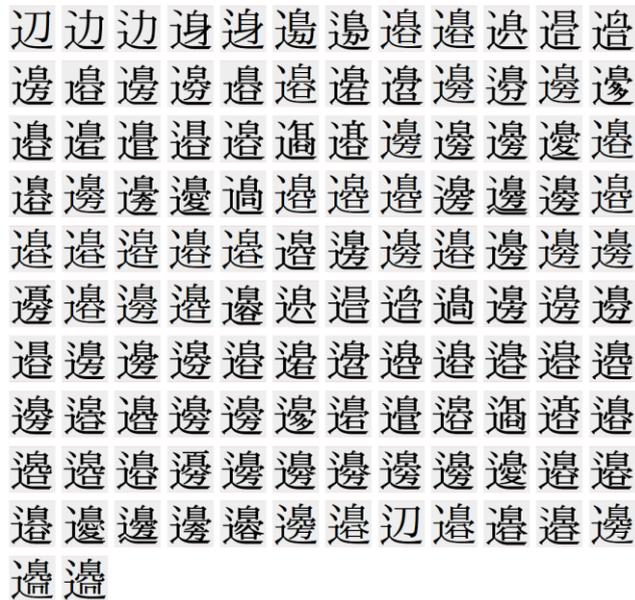


図 2. 「葛」の異体字の差異を示す絞込み検索キーと絞込み結果



上:「葛」の異体字を「ヒ」で絞り込んだ場合

下:「葛」の異体字を 4 画の草かんむりで絞り込んだ場合

図 4. 「邊」の異体字で「方」を含むもの(35 字)



図 5. 「邊」の異体字で「方」と 2 点しんによるもの(17 字)



- TRON は“The Real-time Operating system Nucleus”の略称です。
- 超漢字はパーソナルメディア株式会社の商標です。他の商品名などは、各社の商標または登録商標です。
- 本資料に記載された製品の仕様、外観イメージ、価格などは、発表日現在のものです。最終的に販売される製品では、変更されることがありますので、あらかじめご了承ください。ご購入の際は、最新情報をご確認ください。