## ●年間保守サービスの内容

### 1. トラブル対応

(1) トラブルコール受付

ソフトウェアの操作・機能に関するご質問、また不具合に関する内容の受付を E-Mail もしくは、FAX にて行います。このとき、契約記載番号(お客様 ID)、ご質問の内容、または、不具合の切り分けに必要な情報 ( 現象、システム環境設定など ) をご提供いただきます。

(2) ご質問と回答

E-Mail もしくは、FAX でお受けした操作・機能に関するご質問、または、不具合に関するお問い合わせに対し、回答および技術情報の提供を行います。

(3) ソフトウェアトラブルの切り分け

E-Mail もしくは、FAX でお受けしたソフトウェアトラブルに関するご質問、または、不具合に関するお問い合わせに対し、不具合の状況の把握を行います。

(4) 回答または、支援業務の報告

上記のサービスは、お客様のお問い合せ発生日より弊社営業日 5 日以内に E-Mail もしくは、FAX にて回答するものとし、未解決の案件については、改訂版の作成と提供などの回答、または、支援業務の報告を継続して行うものとします。

### 2. バージョンアップ版の提供

対応文章がバージョンアップされ対応が必要な場合、保守期間内であれば当社ライブラリのバージョンアップ版の提供を行います。

### 3. 保守の対象とならない場合

文書ファイルのフォーマットなどが、暗号化等の対応により情報が隠蔽され、公知の技術でも解析不能な場合には保守ができません。また、一部分の対処が他の部分の不具合に起因し、トレードオフを生ずる場合も同様です。

### 4. OEM 年間保守契約

本ライブラリを技術 OEM として組み込んで利用していただいているソフトウェアメーカ様には、OEM 年間保守契約を締結していただいております。これは通常の保守に加えて、当社がご提供したバージョンアップ版 ( ライブラリ ) をお客様に再配布できる権利を含めているものです。

**保守サービス**　窓　　口：株式会社データ変換研究所　技術サポート担当
受付時間：平日(月～金)の 10 時より 17 時まで。土日祝日及び年末年始を除きます。

## 変換道

Dehenken では、各種フォーマットの文書ファイルから安全にテキストデータを取り出す技術を大切にしています。さまざまな形式のフォーマットから、精度良く・安全に取り出す継続的品質・性能向上活動を「変換道」と呼んでいます。

### ●破損ファイル対象時の安全性(Broken Files Safeness)

文書ファイルやテキストファイルは、保存時や複製時にデータの破損事故が発生する場合があります。本ソフトウェアが破損したファイルであっても、安全に抽出を継続したり、中断したりできることを確認するために、意図的に破壊したファイルで網羅的に動作確認を実施しています。当社はこのような破損ファイルによる安全確認を Broken Files Safeness と呼んでいます。

### ●国際化対応の考え方

Dehenken では国際化対応をするために、内部コードにはユニコードを採用しています。特殊文字(サロゲートペアの文字等)の処理を高速化するために、内部コードとして Dehenken UCS コード(略して DUCS )を定義して処理し、その情報を公開しています。MS-Office や PDF など国際的に流通しているソフトウェアについては、英語版、中国語版、韓国語版等の OS の対応版を入手して、各国言語の OS 上の対象バージョンで動作検証用ファイルを作成しています。

### ●サードベンダ製の出力ファイルの考え方

文書ファイルが 1 メーカの仕様により開発作成している場合もありますが、フォーマットを公開してサードベンダによる文書の作成を許可している PDF / MS-Office 等の場合もあります。Dehenken では、フォーマットの仕様を作成したメーカのものを中心に行っていますが、サードベンダのソフトウェアによるものもエンドユーザに行き渡る可能性もありますので、サードベンダ製の出力ファイルでも動作検証とサポートの範囲と考えています。

### ●新規文書ファイルへの対応の考え方

各社の文書ファイルを作成するソフトウェアが新バージョンにバージョンアップするときがあります。バージョンアップ版が市場に出た時は、市販の流通経路にてバージョンアップ版を買い求め、できる限り早急に対応確認をしています。また、現行サポートしていないファイル形式が登場した場合には、市場での利用状況・認知状況をリサーチしながら、対応要望の十分高いものについては対応作業を進めます。また、機能の追加が必要な場合には、詳細をお聞きした上で調査し、対応ができるものについては柔軟に対応しています。

Dehenken

株式会社 データ変換研究所 Dehenken Limited

本　　社 〒600-8433 京都市下京区高辻通室町西入繁昌町 300-1 カノン室町四条 308
TEL 075-366-4645 FAX 075-366-4649
横浜営業所 〒231-0048 神奈川県横浜市中区蓬莱町 2-4-7 澤田聖徳ビル 204
URL : https://www.dehenken.co.jp/ E-Mail : info_ml@dehenken.co.jp
EST' D 1999 Dehenken Limited © Copyright Dehenken 2025. All rights reserved.

品質マネジメントシステム ISO 9001:2015 の認証取得

株式会社データ変換研究所は、2011 年 9 月 27 日付けで全社統一の品質マネジメントシステムとして DNV より ISO-9001:2008 の認証を取得しました。(現在は2015年版に移行)

認証の対象は「ソフトウエア商品の営業、設計・開発・製造」です。

Certificate No. : C741343
Initial certification date : 27 September, 2011

高精度・超高速テキスト抽出ライブラリ

# デ変研TFライブラリ

# TF Library Dehenken

Dehenken

## ●業界で多数の導入実績

本ライブラリは、テキスト抽出ソフトウェアとして業界での多数の導入実績があります。全文検索ソフトウェア、セキュリティ監査ソフトウェア、ケータイ向けMS-Word / PDF 添付文書中テキスト表示等に、テキストを抽出する部品(ライブラリ)としてご採用いただいています。

## ●多彩な出力のためのオプション

- プロパティ情報を出力することができます。MS-Office や PDF ファイルでは、プロパティに表示される作者、会社名といった情報を出力することができます。
- 半角カナ文字を全角カナ文字に置き換えて出力します。
- さまざまな文字コードを自動判別します。テキストの文字コードは、日本語の場合JISコード(ISO-2022-JP)、EUC-JP、Shift_JIS、UTF-8、UTF-16 などの文字コードが利用されますが、デ変研TFライブラリは自動判定(一部優先判定)することができます。

## ●開発生産性を高める使用環境の提供

本ライブラリを用いてアプリケーションを開発するOEMユーザを支援するために、開発しやすいように工夫と様々な情報提供を行っています。API使用時のサンプルソース(再利用できる使用権付与)。テキスト抽出したいファイルをオンメモリにてAPIに与え、ライブラリがファイルタイプを自動判別しテキスト抽出します。このときのファイルの判定には、拡張子(サフィックス)の情報を使いません。

## ●マルチスレッドに対応

本ライブラリは、CPU数やコア数に応じてテキスト抽出の分散処理による速度の向上ができるよう、マルチスレッドに対応しています(マルチスレッドセーフ)。

文書ファイルからの高精度・超高速テキスト抽出ライブラリ

# デ変研TFライブラリ

**デ変研 TF ライブラリは、MS-Office、PDF、一太郎等の文書ファイルから、テキストを抽出する組込型のテキストフィルタ (Text Filter) ライブラリです。フォーマットの内部のバイナリデータを解析し、プロパティ情報とテキスト情報を高精度・超高速に抽出します。**

## ●ユニコード対応とローカル文字コードの対応

本ライブラリは、日本語文字コードの２つの方式Windows-31J (CP932) とJIS X0213:2004 に対応しています。これらの文字コードの変換を指示した上で、JIS (ISO-2022-JP) / EUC-JP / Shift_JIS / UTF-8 / UTF-16 の変換を行います。本ライブラリのマルチランゲージ版はヨーロッパ圏の文字コードISO-8859 や、アジアの中国語、韓国語について対応しています。(GB2312 / GB18030 / KSC 5601 / KS X 1001 / Big5 / CP シングルバイト・ダブルバイト 等)。

## ●安全のための制限設定 (doccat.conf)

テキスト抽出時の使用リソースの限界を、あらかじめ制限設定ファイル(doccat.conf)に記述することにより、リソースの消費量を制限できます。制限できるのは、文書ファイル容量の制限(MAX_FILE_SIZE)、出力テキストの使用バッファサイズの制限(MAX_BUF_SIZE)、抽出テキストサイズ制限(MAX_TEXT_SIZE)の３つです。

## ●MS-Office 関連ファイルに関するパスワード保護等について

| | | |
|---|---|---|
| Word | 読み取りパスワードによる保護 | テキスト抽出できません |
| | 書き込みパスワードによる保護 | テキスト抽出します |
| | 読み書きパスワードによる保護 | テキスト抽出できません |
| Excel | ブック保護 | テキスト抽出します |
| | シート保護 | テキスト抽出します |
| | 読み取りパスワードによる保護 | テキスト抽出できません |
| | 書き込みパスワードによる保護 | テキスト抽出します |
| | 読み書きパスワードによる保護 | テキスト抽出できません |
| PowerPoint | 読み取りパスワードによる保護 | テキスト抽出できません |
| | 書き込みパスワードによる保護 | テキスト抽出します |
| | 読み書きパスワードによる保護 | テキスト抽出できません |
| PDF | 40bitsRC4 | テキスト抽出できます (*1) |
| | 128bitsRC4 | テキスト抽出できます (*1) |
| | 128bitsAES | テキスト抽出できます (*1) |
| | 256bitsAES | テキスト抽出できます (*1) |

PDF ファイルの暗号方式に関しては、デコードする際に暗号方式をライブラリの内部で自動判別しています。

*1 正しいユーザーパスワードを指定することが必要です。

TF Library

## ■対応 OS

**Red Hat Linux**
AS3 / ES3 / WS3 / AS4 / ES4 / WS4 / EL5 / EL6 / EL7 / EL8 / EL9 / EL10
**Windows** 2000 / XP / Vista / 7 / 8 / 8.1 / 10 / 11
**Windows Server** 2000 / 2003 / 2008 / 2008R2 / 2012 / 2012R2 / 2016 / 2019 / 2022 / 2025

## ■対応 C コンパイラ

**Windows** Microsoft Visual Studio 2008 以上
**Linux** Gnu C Compiler(gcc)
**【デ変研 TF ライブラリ for Java の場合】**
**Java JDK** 1.6 以降 (Windows / Linux にて動作確認済)
( Oracle Java / Zulu Java / Open JDK に関しては、お客様要望により確認します )

## ■構成

**メモリ** 1GB 以上
**HDD 利用量** 500MB 以上

※Windowsは、x86 または x64 を対応に含めます。
※Linuxは、x86,x64 または Arm64 を対応に含めます。
※他のOS・コンパイラ・開発環境下でご不明な点は、お問い合わせください。
※ハードウェアの搭載メモリは推奨2GB以上で、メモリ量が多い方が大きな文書に対応できます。

## ■対応文書

**Microsoft Word**
97 / 98 / 2000 - 2016 / 2019 / 2021 / 2024
**Microsoft Excel**
97 / 2000 - 2016 / 2019 / 2021 / 2024
**Microsoft PowerPoint**
97 / 2000 - 2016 / 2019 / 2021 / 2024
**Microsoft Visio**
2002(XP) / 2003 - 2016 / 2019 / 2021 / 2024
**Microsoft Word for Mac**
98 / 2001 - 2016 / 2019 / 2021 / 2024
**Microsoft Excel for Mac**
98 / 2001 - 2016 / 2019 / 2021 / 2024
**Microsoft PowerPoint for Mac**
98 / 2001 - 2016 / 2019 / 2021 / 2024
**Microsoft 365**
年 2 回動作確認
**Microsoft XPS** 1.0
**JustSystems 一太郎**
Ver.5-Ver13 / 2004 - 2025
**Adobe Systems Acrobat**
4.0 / 5.0 / 6.0 / 7.0 / 8.0 / 9.0 / X / XI / DC / 2017 / 2020
**PDF** 1.2 / 1.3 / 1.4 / 1.5 / 1.6 / 1.7
**RTF** 1.0 -1.9
**テキスト文書**
JIS (ISO-2022-JP) / EUC-JP / Shift_JIS / UTF-8 / UTF-16
**マークアップ言語**
HTML / XML / SGML
**ODF(Writer / Calc / Impress)** 1.1 / 1.2

※詳細は、マニュアルをご参照ください。