●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ご提供を行います。

(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ご提供を行います。

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。

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。
休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 / RTF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 〒604-8155 京都市中京区錦小路通室町東入占出山町308 ヤマチュウビル1F
TEL 075-254-8780 FAX 075-254-8790

URL : http://www.dehenken.co.jp/ E-Mail : info_ml@dehenken.co.jp

高精度・超高速テキスト抽出ソフトウェア

DocCat

Dehenken

●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。
●半角かな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使用されますが、DocCat は自動判定 (一部優先判定) しますので、文字コードの統一に役立てることができます。

●日本語文字コードの扱いに精通

DocCatの内部処理の文字コードは UTF-16 を使用しています。出力時には、さまざまな日本で使用されている文字コードにすることができます。出力時の文字コードは JIS X 0213:2004 (デフォルト)を指定により Windows-31J(CP932) に切り替えることができます。

文書ファイルからの高精度・超高速テキスト抽出コマンド

DocCat

DocCat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ソフトウェアです。UNIX のcat コマンドのような使い勝手のイメージで、テキスト化を行うことができます。本ソフトウェアは、文書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。全文検索や添付ファイルのテキスト化など、多くのシーンでご使用いただいているテキストフィルタのコマンドです。DocCatは、テキスト抽出コマンドとして2000年にリリースされて以来、SIベンダ様を通じて多数のお客様のシステムにご導入いただいております。

●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、文書ファイルの最大値の制限 (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
	＊MS-PowerPoint for Mac は、読み取りパスワードによる保護の機能がありません。	
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)(*2)
	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 DocCat 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。	
	*2 Acrobat X で作成されたものは未対応です。	
	詳しくはお問い合わせください。	

■対応 OS

Red Hat Linux
AS3 / ES3 / WS3 /AS4 / ES4 / WS4 / EL5 / EL6
Solaris 9 / 10
※他のOSに関しましては、別途お問い合わせください。

■構成

メモリ 1GB以上
(推奨2GB以上。メモリ量が多い方が大きな文書ファイルに対応できます)
HDD利用量 10MB以上

■対応文書

Microsoft Word
95 / 97 / 98 / 2000 / 2002 (XP) / 2003 / 2007 / 2010
Microsoft Excel
95 / 97 / 2000 / 2002 (XP) / 2003 / 2007 / 2010
Microsoft PowerPoint
95 / 97 / 2000 / 2002 (XP) / 2003 / 2007 / 2010
Microsoft Visio
2002(XP) / 2003 / 2007 / 2010
Microsoft Word for Mac
98 / 2001 / 2004 / 2008 / 2011 for Mac
Microsoft Excel for Mac
98 / 2001 / 2004 / 2008 / 2011 for Mac
Microsoft PowerPoint for Mac
98 / 2001 / 2004 / 2008 / 2011 for Mac
Microsoft XPS 1.0
JustSystems 一太郎
Ver.5 - Ver.13 / 2006 - 2012
Adobe Systems Acrobat
4.0 / 5.0 / 6.0 / 7.0 / 8.0 / 9.0 / X
(一部未対応の形式があります)
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 1.1 / 1.2 (Writer / Calc / Impress)
OpenOffice 3.0 / 3.1 / 3.2 / 3.3
LibreOffice 3.4

DocCat