

# 無害化したファイルの困ったを解決

無害化したPDFファイルをOffice、一太郎  
などの文書形式で編集できます

アンテナハウス株式会社

2017年2月 23日

# コンテンツ

- 1. アンテナハウス株式会社概要
- 2. PDF活用入門講座
- 3. PDF to Office変換関連製品紹介
  - 瞬簡PDF変換 9
  - PDF to Office変換ライブラリ
  - PDF Server V3.1

# アンテナハウス株式会社 会社概要



設立：1984年（昭和59年）8月

資本金：4,000万円

東京都中央区東日本橋2丁目1番6号

東日本橋藤和ビル5階

URL：<http://www.antenna.co.jp/>

問合せ：[sis@antenna.co.jp](mailto:sis@antenna.co.jp)

都営浅草線：東日本橋駅（徒歩約1分）

JR総武本線：馬喰町駅（徒歩約4分）

# アンテナハウス株式会社 事業内容

## ● 文書レンダリング技術

ワープロ,表計算,画像,Webファイル形式等を高精度で相互変換。

## ● PDF技術

PDFを紙の視認性と電子データの利便性を併せ持つ電子配信の基盤技術と考え、その生成、加工、編集、変換と高度な利用を推進。

## ● XML組版技術

W3CのXSL-FO仕様に準拠し、XMLを利用して多言語にも対応した高品位出版物を効率的に作成。

## ● 電子書籍サービス

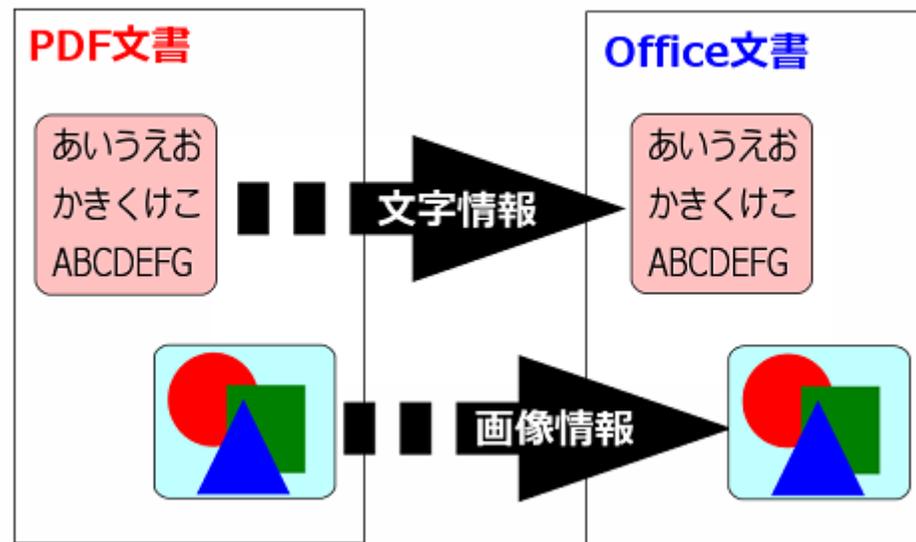
XML技術、PDF技術を利用し、電子書籍制作・変換Webサービス(CAS-UB)、書籍(PDF、EPUB)制作サービスを提供。

## 1.PDFからOfficeに変換する方法

PDFは、作成方法によって2種類

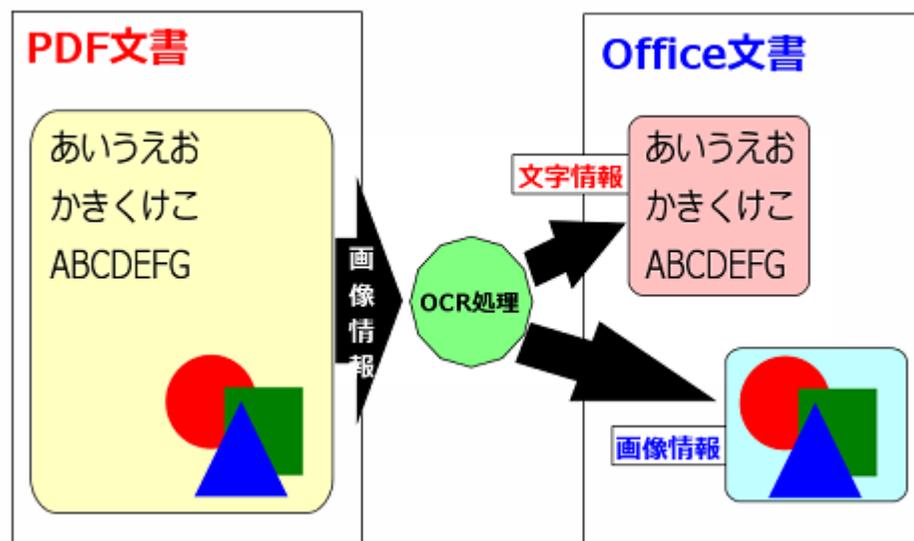
- 「電子文書」としてのPDF：テキストPDF

テキスト情報や、フォント設定など詳細なPDF内部のデータを解読して変換



- 「電子化文書」としてのPDF：画像PDF

スキャナーなどで取り込んだ画像をOCR機能で変換



## ●PDF「逆」変換ソフトの特長

### 内部解析型

- テキストを正確に変換できる...[○]
- 画像化されたテキストは変換できない...[×]

### OCR変換型

- 画像からテキストを変換できる...[○]
- テキストを誤認識する可能性がある...[×]

### 複合型

- テキストPDFからも画像PDFからもテキストを変換できる...[○]

# PDF活用入門講座

## 2.テキストPDFなのに文字化け!

The image shows two windows side-by-side. The left window is Adobe Acrobat Pro Extended displaying a PDF form titled '防災管理点検結果報告書' (Disaster Management Inspection Result Report). The form includes fields for '東京消防庁' (Tokyo Fire Department), '消防署長 殿' (Fire Station Chief), and '届出者' (Applicant) with fields for address, name, and phone number. It also has a table for '防災管理対象物' (Disaster Management Target Object) with columns for location, name, purpose, and construction type. The right window is Microsoft Word showing the same document but with garbled text, demonstrating the issue of text not being recognized as text in a PDF.

# PDF活用入門講座

理由はPDFの作成方法にあり！

PDFを作成する際に、文字の情報をPDFに格納する方法

- 1)文字コードをPDFに直接格納する。
- 2)文字の形状（グリフ）をPDFに埋め込む

文字コード対照表をPDFに用意するのは、PDFを作成する側のソフトウェアの責任。

**『瞬簡PDFシリーズ』をご利用下さい！**

OCR機能が用意された複合型変換ソフトがお役立ち。

## 3.配布用PDF作成条件・留意点

- 1)フォントの埋め込み
- 2)文字コードを正しく保存
- 3)長期保管：国会図書館などに保存（PDF/A）
- 4)アクセシビリティへの配慮  
「障害者差別解消法」への対応（PDF/UA）
  - タグ付きPDFへの変換
  - 文書を構造化する

## PDF to Office変換関連製品

### 1. デスクトップ製品

- ・瞬簡PDF変換9

### 2. システムコンポーネント製品

- ・ PDF to Office 変換ライブラリ

### 3. ソリューション製品

- ・PDF Server( Office変換オプション)

# 瞬時に簡単! 瞬間PDF変換9

## 新機能と使い方のご紹介

アンテナハウス株式会社

# 瞬簡PDF変換 - 製品の概要

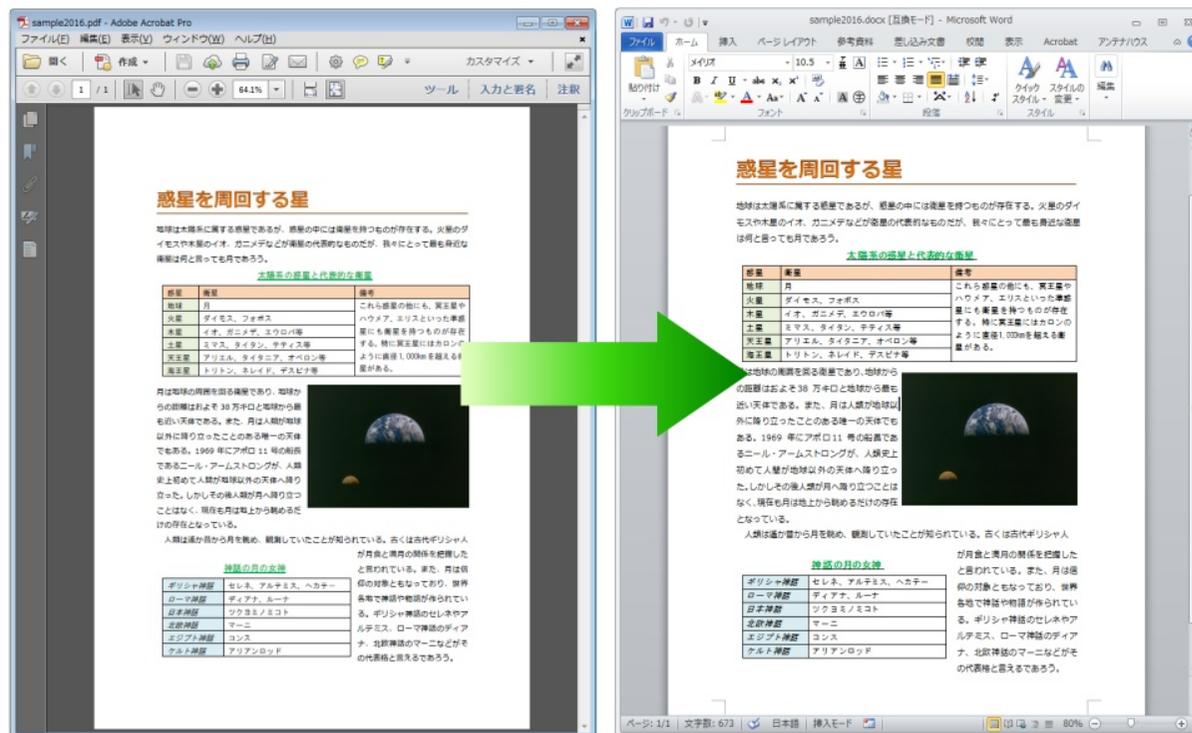
## PDFを編集可能なOffice文書に変換！

- PDFファイルからWordやExcel、PowerPoint、一太郎形式のファイルへ簡単な操作で変換（Office 2007以降のXML形式保存にも対応）
- スキャナーなどから読み込んだ画像PDFや画像ファイルも、OCR（文字認識）処理により変換



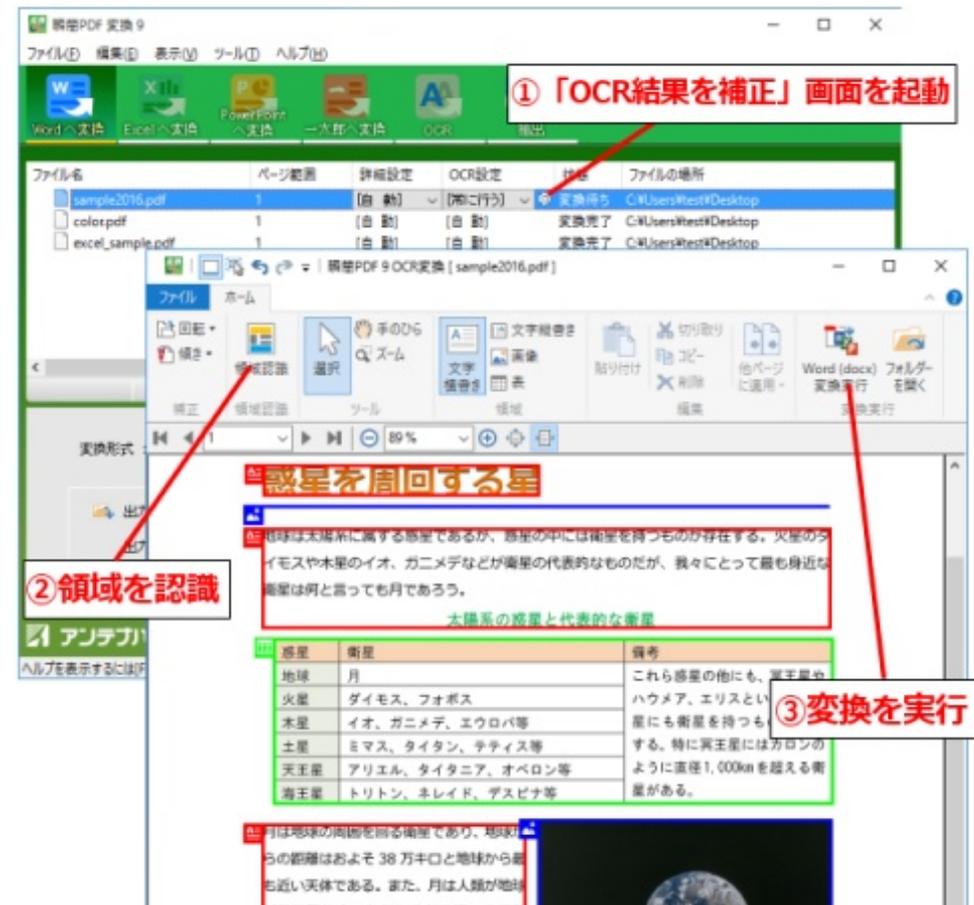
# 瞬簡PDF変換 - 新しくなったポイント

- Microsoft Office 2007以降のXML形式に対応
  - 従来の .doc/.xls/.ppt 形式への変換の他、ご要望の多かった Microsoft Office 2007以降のXML形式（拡張子.docx/.xlsx/.pptx）に対応！



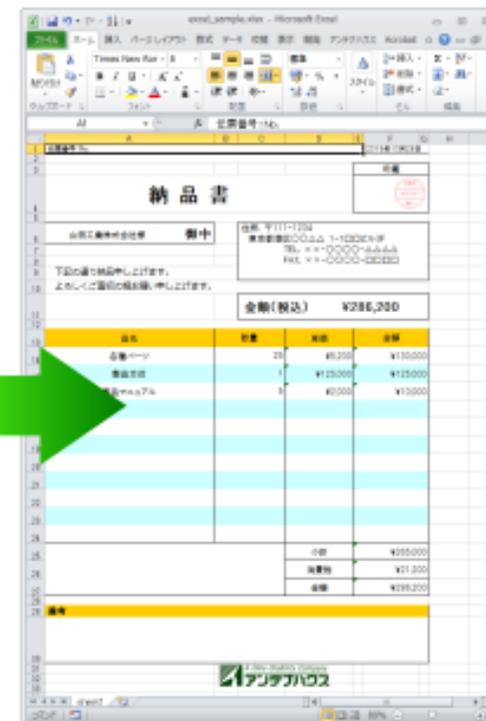
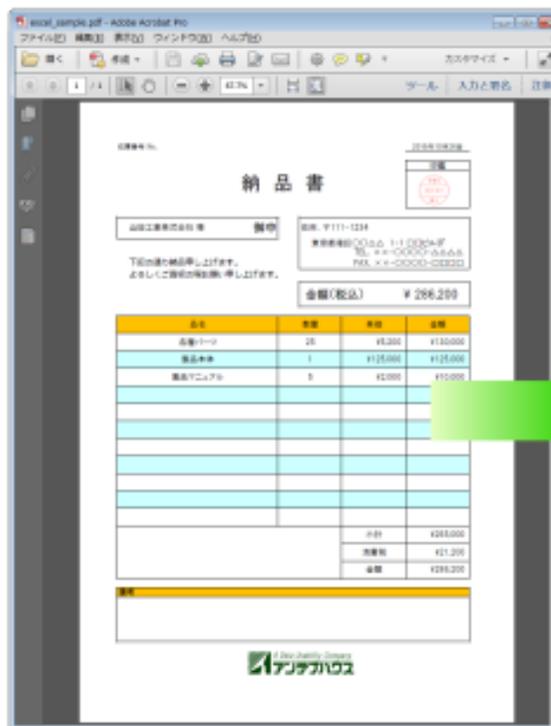
# 瞬簡PDF変換 - 新しくなったポイント2

- OCR変換をより確実にする「OCR補正機能」を搭載
  - 画像化されたPDFファイルの変換を行う際に、テキスト、表、画像にしたい箇所をそれぞれ変換前に設定することで、より正確な変換をサポート。



# 瞬簡PDF変換 - 新しくなったポイント3

- 変換エンジンの改良で、より高精度な変換を実現！
  - PDFからExcelへの変換例
  - PDFからPowerPointへの変換例



## 瞬簡PDF変換-その他主な機能

- 「瞬簡PDF 変換 9」の変換機能をMicrosoft OfficeのWord、Excel、PowerPointにアドイン。これにより、Office上で直接PDFを変換可能に。
- 画像PDFにOCR（文字認識）処理を行い、透明テキスト付きPDFを作成。これにより、文字情報を含まないPDFファイルでも本文の文字列検索が可能に。
- PDFからテキストや画像を取り出してそれぞれファイルに保存。これにより、PDF内のデータを別のソフトに取り込んで再利用することが可能に。

## 瞬簡PDF 変換 9 と他社製品で変換精度を比較

**瞬簡PDF 変換 9 は国産最高クラスの変換精度！**

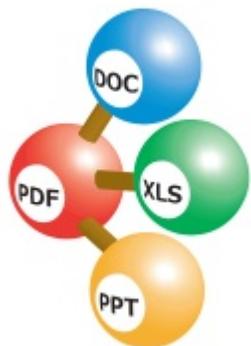
PDFからOffice文書（Word、Excel、PowerPoint）への変換機能について、変換処理の結果を比較しました。

- ① 複数のテキスト情報を含むPDFファイル
- ② 紙書類からスキャナー経由で作成した文字情報を持たないPDFはOCRを利用

（2016年11月 当社調べ）

変換処理の結果と、使用したファイルや採点方法、ファイル単位での比較結果は以下をご参照ください。

[http://www.antenna.co.jp/pdftooffice/conv\\_compare.html](http://www.antenna.co.jp/pdftooffice/conv_compare.html)



# Antenna House

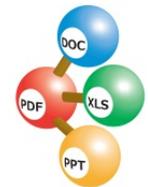
## PDF to Office変換ライブラリ

### のご紹介

アンテナハウス株式会社

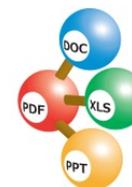
# PDF to Office-変換ライブラリの特長

- 本ライブラリは、PDFから以下のMicrosoft Office文書形式に変換します。
- ◆ Microsoft Word (2007-2016 /docx形式、97-2003/doc形式)
- ◆ Microsoft Excel (2007-2016 /xlsx形式、97-2003/xls形式)
- ◆ Microsoft PowerPoint (2007-2016 /pptx形式、97-2003/ppt形式)
- 国産最高クラスの強力な変換機能。
- 少ないステップでアプリケーションに組み込み可能。
- C/C++、.NET、Java、コマンドラインの各インタフェースを用意。
- 再頒布可能なライセンス形態 (OEM) でのご提供。



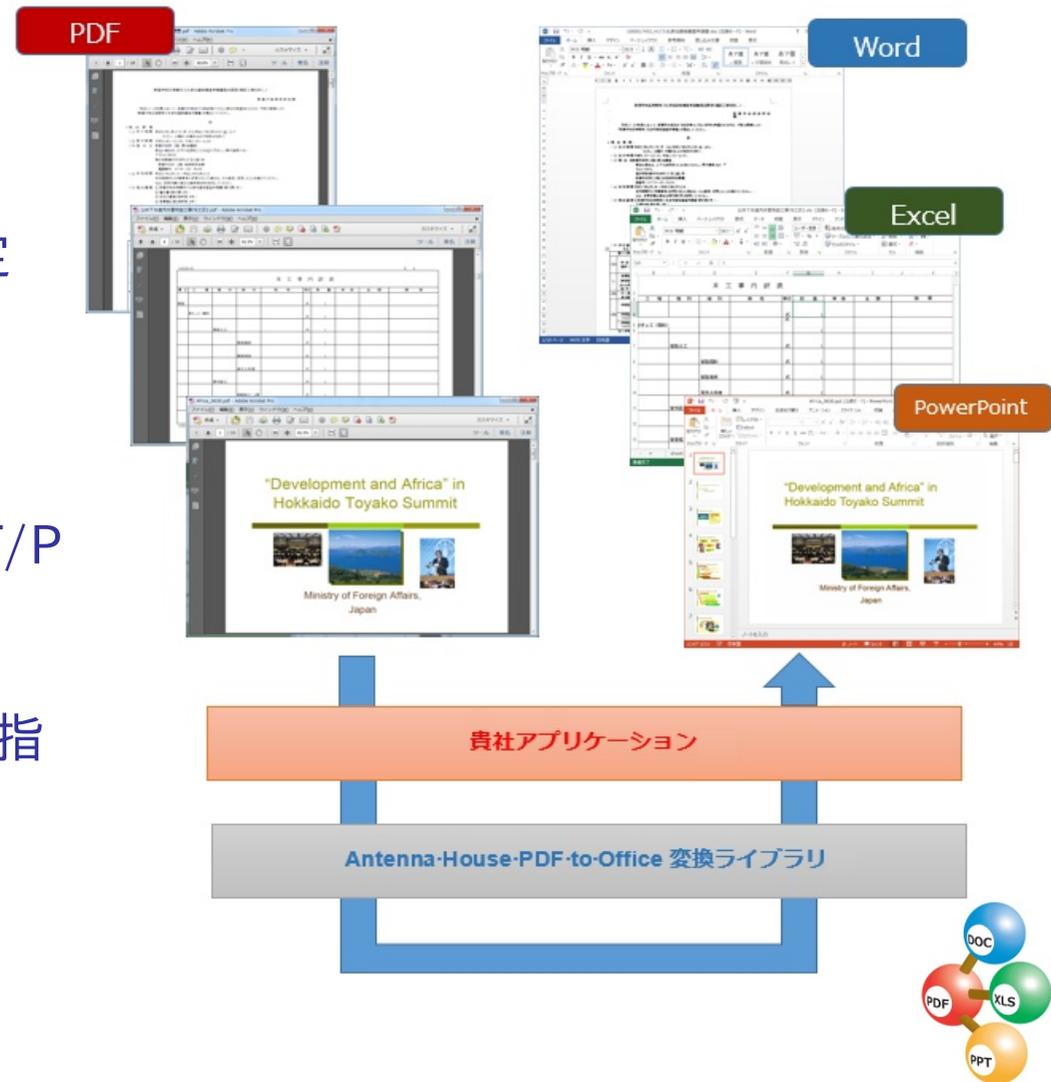
# PDF to Office-制限事項

- PDFのセキュリティ設定はパスワードのみ対応（解除にはパスワードが必要）
- OCR機能を追加（オプション）
- Open Office XML形式（docx、xlsx、pptx）に対応（実行環境にJavaランタイム（Java 8）のインストールが必要）
- 表現力の制限
  - システムに無いフォントや文字間調整の再現
  - 回転、カラー（パレット、カラースペースなど）
- 注釈やフォームには未対応



# PDF to Office-変換ライブラリの処理概要

- 1.変換元のPDFファイルを指定  
(必須)。
- 2.変換先のOfficeファイル名を指定  
(必須)。
- 3.変換先Officeファイルの種類  
(DOC/DOCX/XLS/XLSX/PPT/P  
PTX)を指定(必須)。
- 4.変換条件を設定したファイル名を指  
定(任意)。
- 5.変換を実行。

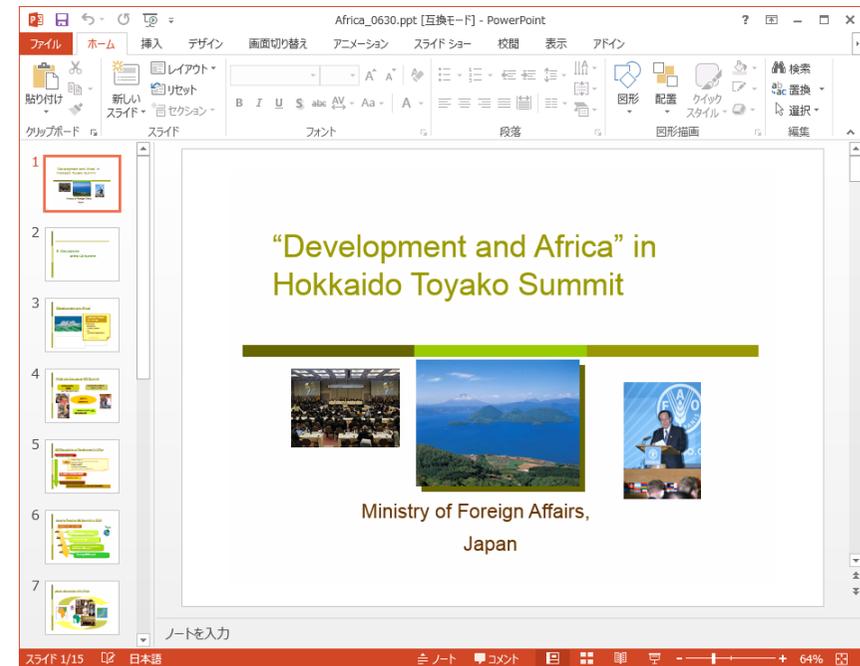






# PDF to Office-変換ライブラリの変換例(3)

- PDFからPowerPointへ



# PDF to Office-変換オプション (1/3)

- 変換オプション設定ファイル (XML)
  - オプション設定ファイルは以下のような書式で変換条件を記述。

```
<?xml version="1.0"?>
```

```
<pto-interface>
```

```
<doc-option ... ></doc-option>
```

Word への変換条件

```
<xls-option ... ></xls-option>
```

Excel への変換条件

```
<ppt-option ... ></ppt-option>
```

PowerPoint への変換条件

```
</pto-interface>
```



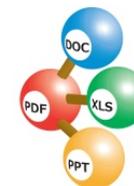
# PDF to Office-変換のオプション (2/3)

## Word変換のオプション

- from
  - 開始ページを指定する
- to
  - 終了ページを指定する
- layout
  - レイアウト優先で変換する
- listnum
  - 段落番号を変換する
- underline
  - 下線、取り消し線を変換する
- vertoff
  - 縦書き書式を変換しない

## EXCEL変換のオプション

- from, to
  - 開始ページと終了ページの指定
- shmerge
  - 複数シートを1シートにまとめる
- noimg
  - 線画・画像を出力しない
- vcmerge
  - セルの縦結合
- tables <xls-option>の子要素
  - 表変換範囲指定



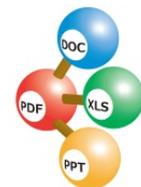
# PDF to Office-変換オプション (3/3)

## PowerPoint変換オプション

- from
  - 開始ページを指定する
- to
  - 終了ページを指定する
- listnum
  - 段落番号を変換する
- underline
  - 下線を変換する

## ◆ 「PDF to Office」 Webサイト

- <http://www.antenna.co.jp/pdftooffice-lib/>
- 評価版あり（機能制限なし、メールにて個別対応）



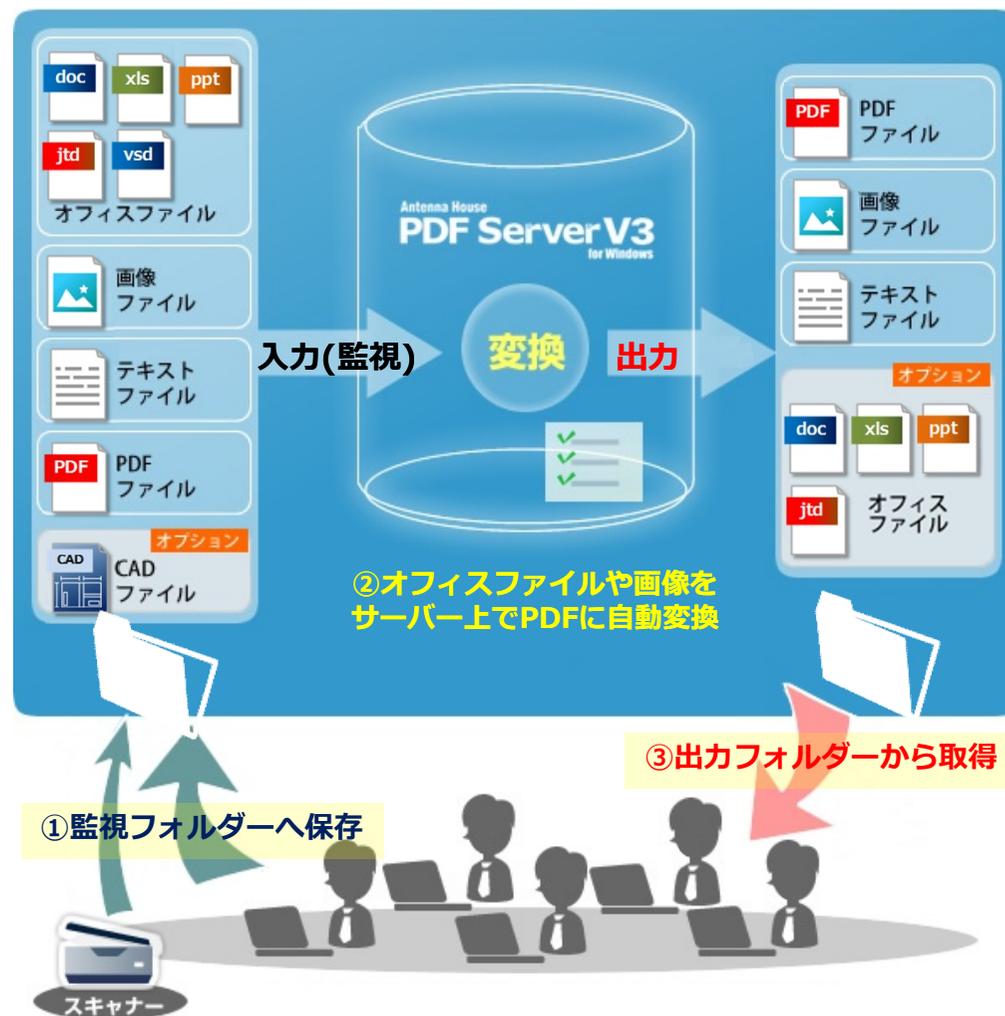
PDF自動変換・編集  
サーバーソリューションソフトウェア  
**PDF Server V3.1 のご紹介**

アンテナハウス株式会社

# AH PDF Serverについて(1/2)

- 『AH PDF Server』とは

- サーバーOS上で動作する、ファイル変換サーバーソリューションを提供するソフトウェアです
- MS-Office文書や画像ファイルなどを予め設定した変換設定に従って、自動でPDF等に変換・出力します
- フォルダ監視による一括変換から、コマンドライン、Webブラウザを使ったファイル変換など、多様な運用が展開可能です



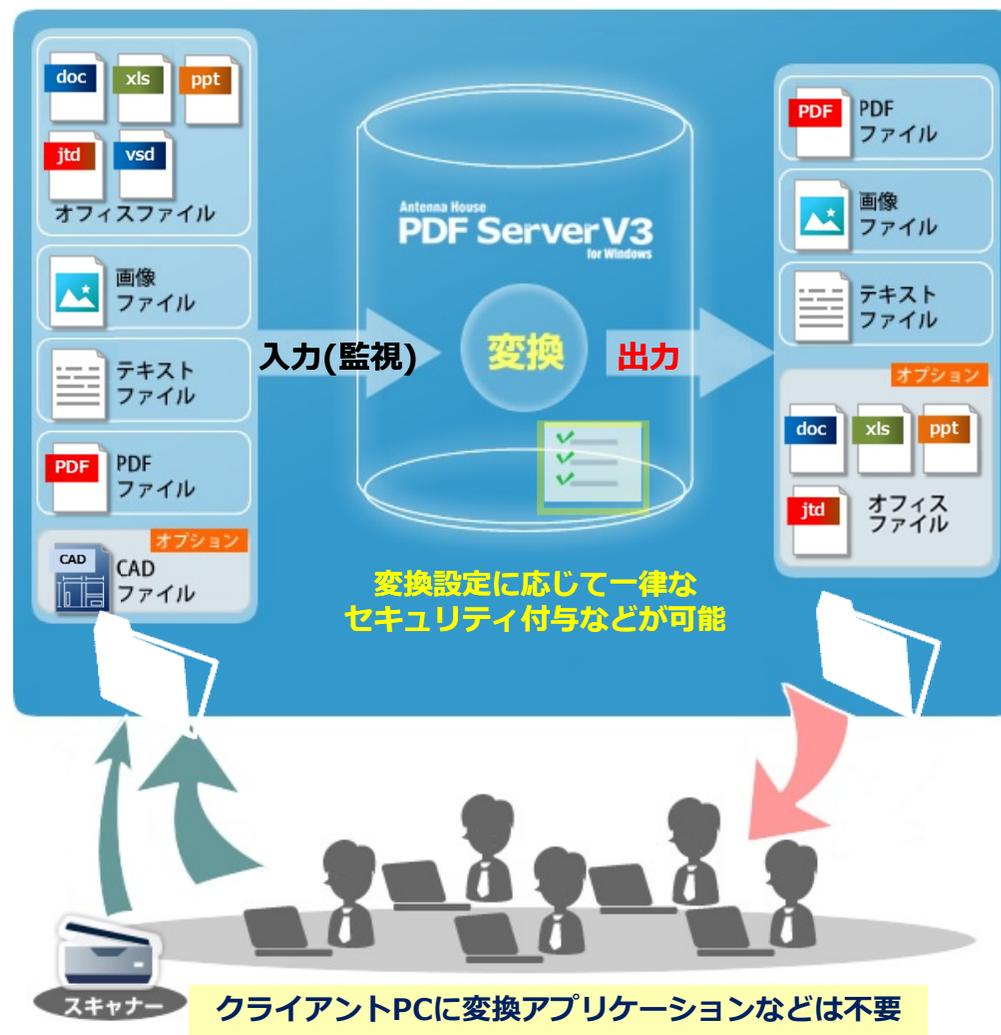
# AH PDF Serverについて (2/2)

## 『AH PDF Server』導入のメリット

- サーバー側で変換するため、クライアントPCにアプリケーションのインストールや管理が不要です
- ライセンス料金は接続PCの台数に影響を受けないため低コストで導入可能です。

※MS-Officeのライセンスは必要。

- サーバーで変換を一元管理  
PDF変換時に設定に応じたヘッダ・フッタ/ウォーターマーク/セキュリティの付与などが可能。社内ルールに則った一律な変換も行えます



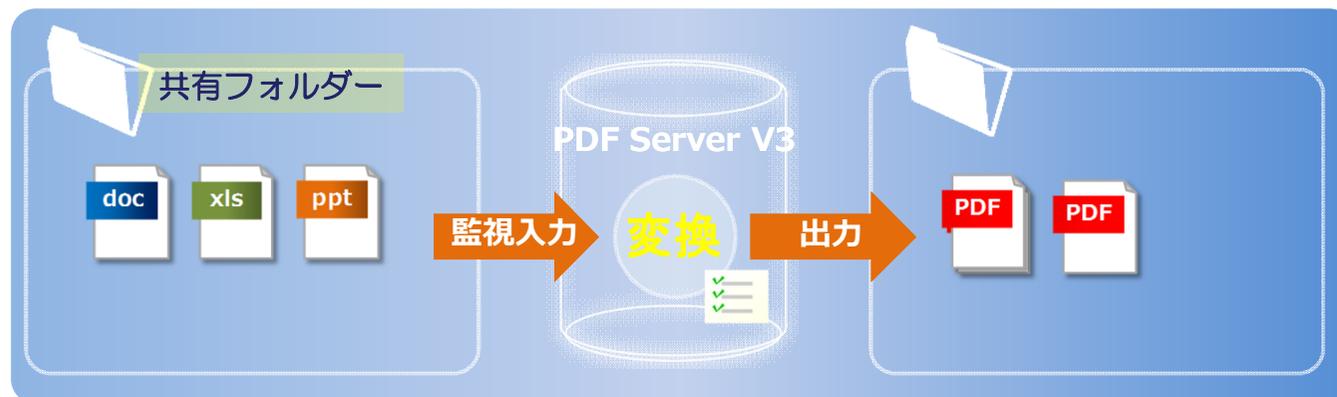
# AH PDF Server - 主要な機能と利用イメージ

- フォルダー監視によるPDF変換
  - MS-Office文書の変換
  - スキャナー連携
- コマンドラインを利用した変換
  - コマンドライン概要
  - Webアプリケーションからの利用
- CAD版とCADオプション
  - ダイレクト変換
  - 図面管理システム用にTIFFを出力
  - CAD変換の補足
- Office変換オプション
  - 既存PDFをMS-Office文書に逆変換

# AH PDF Server-フォルダー監視によるPDF変換 (1/2)

- MS-Office文書の変換

- 共有フォルダーをPDF Serverの監視フォルダーに指定する標準的な運用例です。



- タスクと変換設定を複数用意することで、投入するフォルダーによって変換設定を使い分ける運用が可能です。

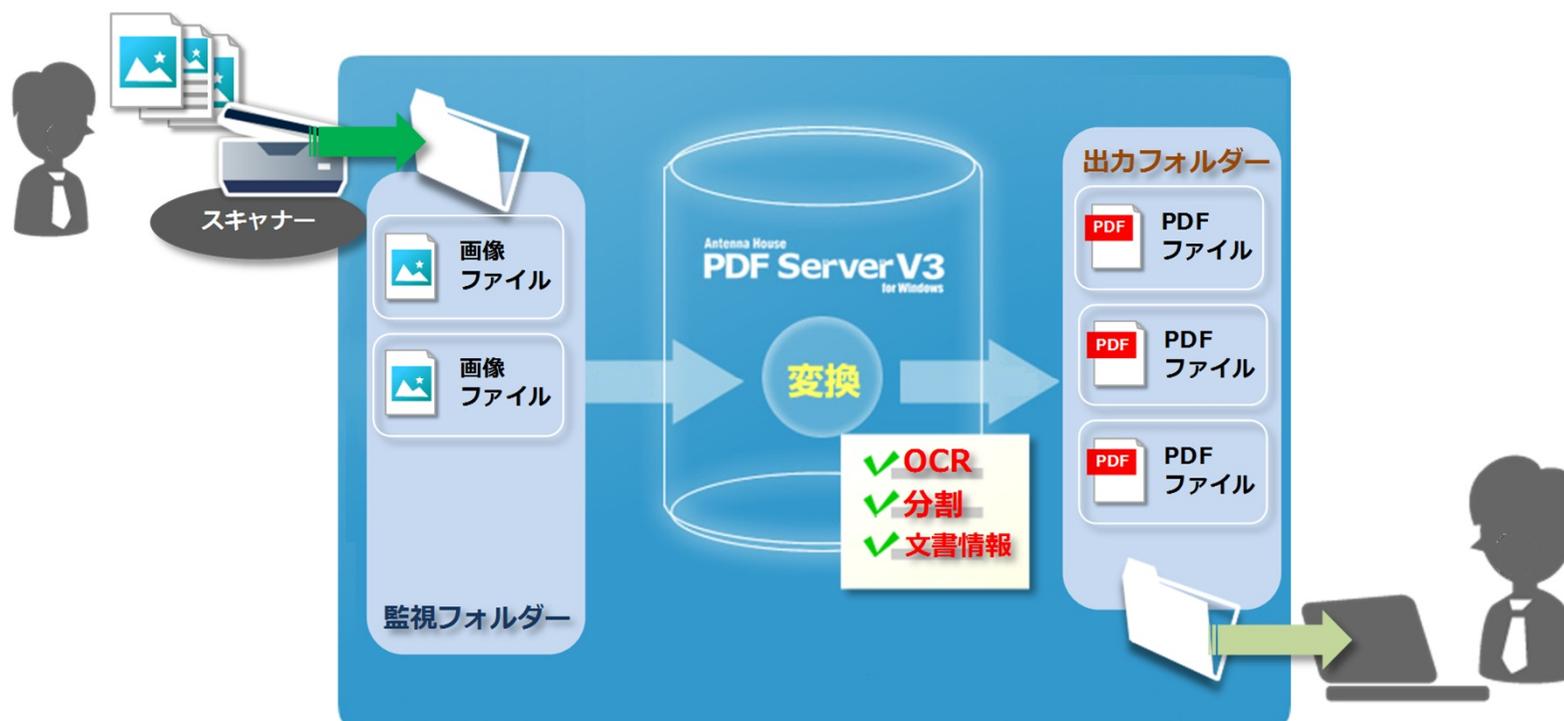
設定例：

- 社内文書フォルダー：
  - 共通文書情報の付与など
- 社外提出資料フォルダー：
  - 変換ファイルの結合
  - 専用の文書情報の付与
  - ウォーターマーク（すかし）の付与
  - セキュリティ付与

# AH PDF Server –フォルダー監視によるPDF変換 (2/2)

- スキャナー連携

- スキャナーや複合機からの出力フォルダーを、PDF Serverの監視フォルダーに設定することで、PDF Serverがタスクや変換設定に基づいてOCR処理やページ分割などを自動で実行。紙の文書のPDF化が簡単に行えます。



# AH PDF Server - コマンドラインを利用した変換 (1/3)

## • コマンドライン概要

- 『PDF Server』の変換と結合/分割をコマンドライン上で実行出来るコンソールアプリケーションで、プロフェッショナル版/コマンドライン版で利用可能です。
- 任意のタイミングでPDF変換を実行でき、「バッチ処理による一括変換」や「WebシステムへPDF変換の組み込み」などで利用が可能です。

### 主な機能の一覧：

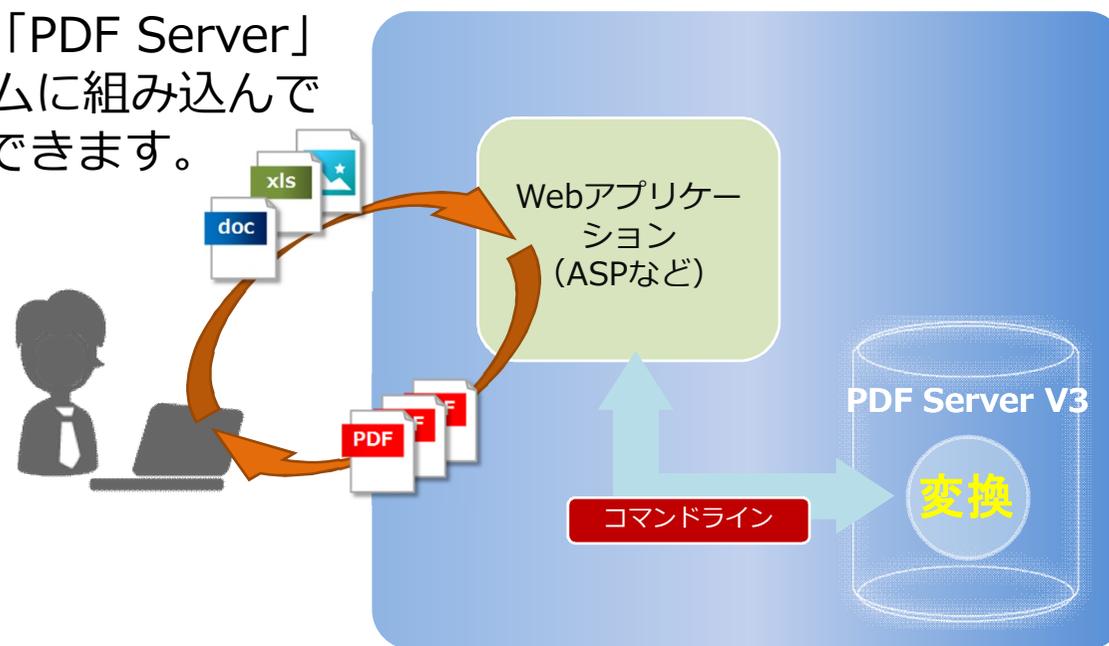
| パラメーター設定・目的           | 解説  |
|-----------------------|---|
| 入力ファイル名を指定する標準的なPDF変換 | PDF Serverのデフォルト設定(単純なPDF変換)で、入力ファイルをPDF変換します。変換可能ファイルは、「PDF Server」のフォルダ監視(Windowsサービス)と同じです。        |
| OCRを実行し透明テキスト付PDFを作成  | 引数(入力ファイル名)の後にOCRの実行オプションを付加すると、PDF変換時にOCRが実行され透明テキスト付PDFを作成できます。<br>※この機能は、入力ファイルがイメージ(画像)のときに利用できます |
| PDFドライバの設定を指定したPDF変換  | 引数(入力ファイル名)の後に、PDFドライバオプション、ドライバの設定名を入力すると、そのドライバの設定内容に従ったPDFが作成されます。<br>※この機能は、Office文書のPDF変換で利用できます |
| 既存の変換設定を利用したファイル変換    | 引数(入力ファイル名)の後に、変換設定のオプションと変換設定名を指定することで、ファイル変換時、指定された変換設定の内容に従った変換処理を行うことができます。                       |
| 一般的なアプリケーション文書のPDF変換  | プロフェッショナル機能の一つに、Office製品以外のアプリケーションを利用したファイル変換設定がありますが、コマンドラインからも利用可能です。<br>【例】一太郎のPDF変換              |
| 出力ファイル種別と出力先フォルダーの指定  | 引数(入力ファイル名)の後に、出力ファイルを指定するオプションを付加し、その後ろに目的のフォルダパスを指定すると、指定されたフォルダーに、変換された出力ファイルが格納されます。              |
| 画像ファイル出力時の詳細の指定       | オプションとしてTIFF/JPEG出力時に解像度や画質などを指定できます。<br>これらはマルチプロセス対応の為、パラメーターとして指定できるようV3.1で追加されました。                |

# AH PDF Server - コマンドラインを利用した変換 (2/3)

- Webアプリケーションからの利用

「PDF Server」のオプションにラインナップされている『Web変換オプション』は非常にシンプルなWeb変換を実現していますが、この『Web変換オプション』もPDF Serverの「コマンドライン」を利用して作成されています。

新たにPDF変換を行うWebシステムを構築する際、あるいは既存のWebシステムのカスタマイズとして「コマンドライン」を利用することで、「PDF Server」の変換処理をWebシステムに組み込んで利用していただくことができます。

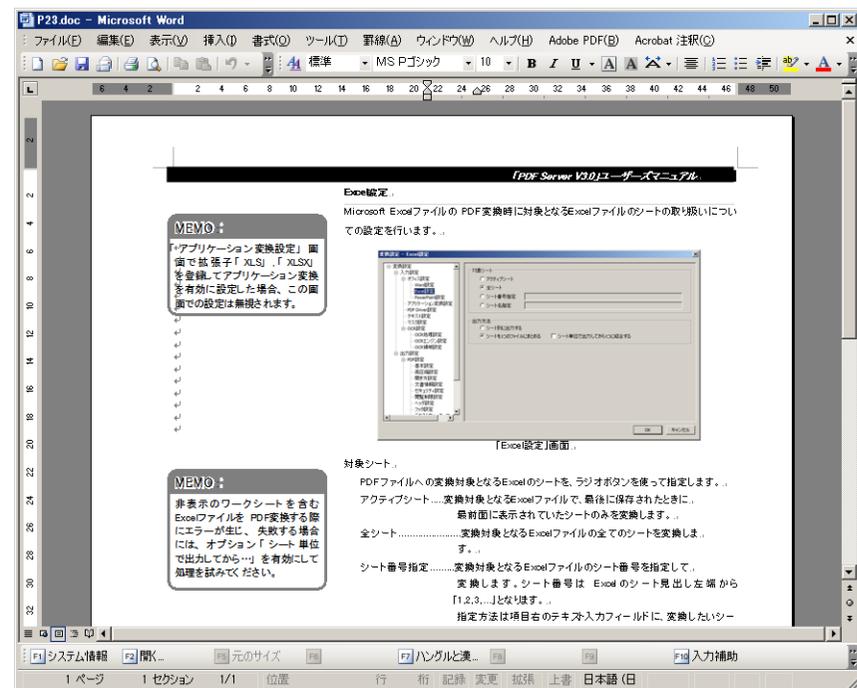
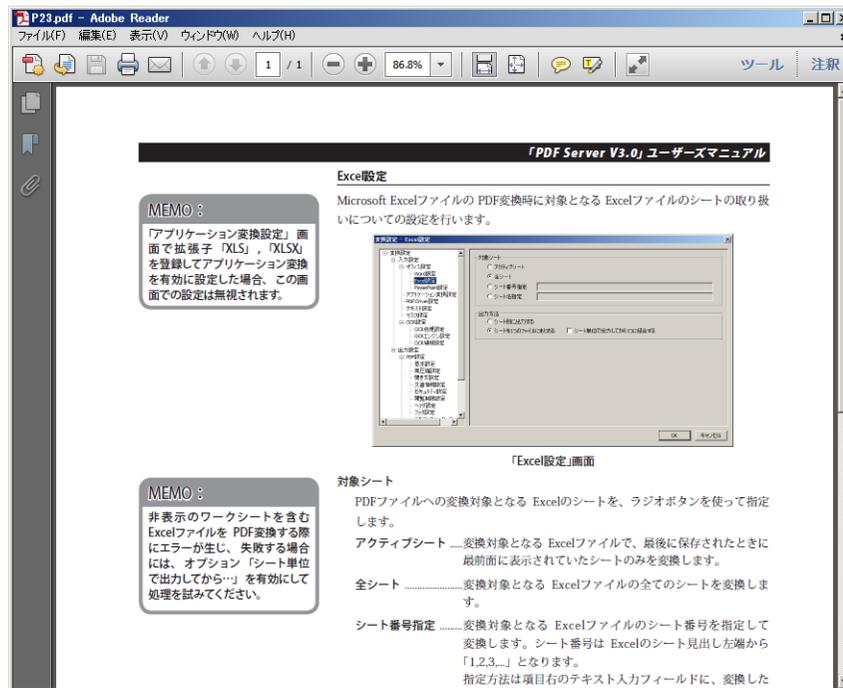


# AH PDF Server – Office変換オプション

- 既存PDFをOffice文書へ高精度に逆変換
  - PDFファイルのレイアウトを保ったまま、Office文書（Word／Excel／PowerPoint）のOffice文書や一太郎のファイルに変換。  
※変換できるのは doc／xls／pptとjtd 形式。  
docx／xlsx／pptx形式での出力には非対応。
  - PDFファイルだけでなく、スキャナーなどから読み込んだ画像PDFや画像ファイルも、OCR（文字認識）処理により、Office文書や一太郎のファイルへ変換が可能。
  - これらPDFファイルからの逆変換には、Word／Excel／PowerPoint／一太郎の各ソフトウェアは不要。  
※オフィス・一太郎ファイルのPDF変換では各ソフトウェアは必要。
  - 当社の実績のある変換エンジンを採用。  
国産ソフト最高の変換精度（2012年4月当社調べ）

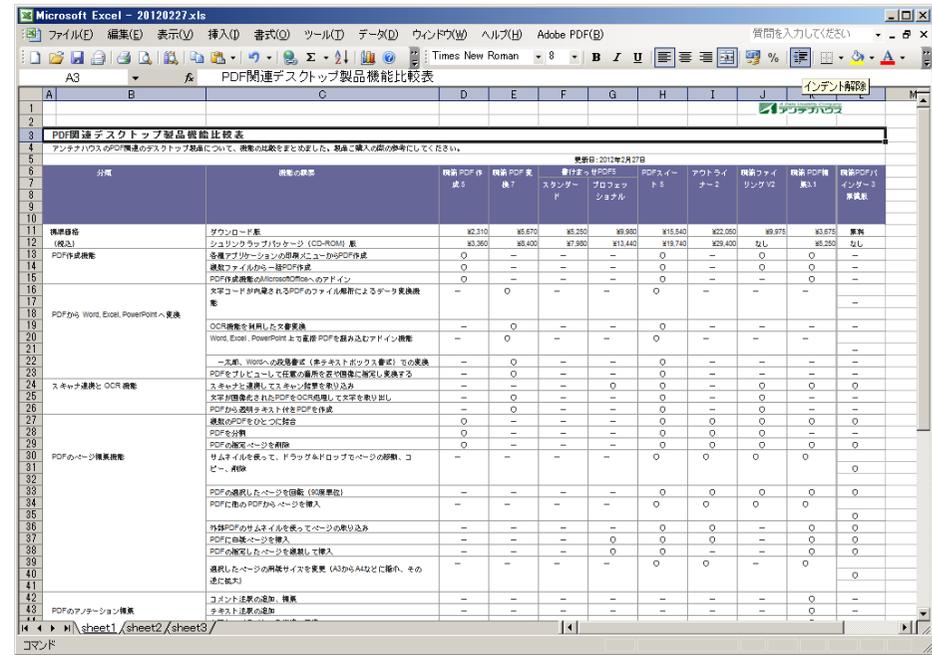
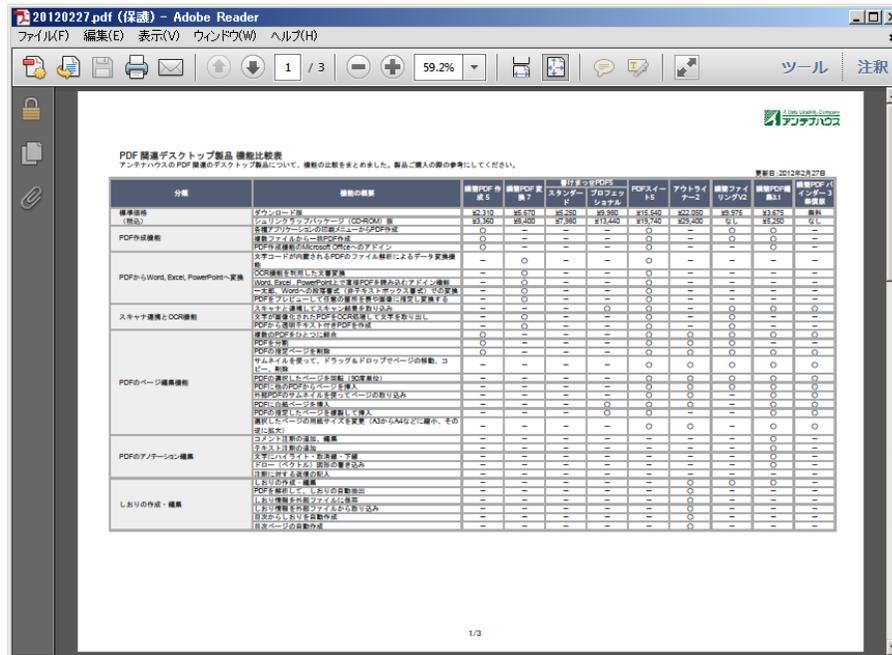
# AH PDF Server – Office変換オプション ～Word例

- Word／一太郎：文字・文字飾り・段落書式・表書式・イメージ・線画を変換できます



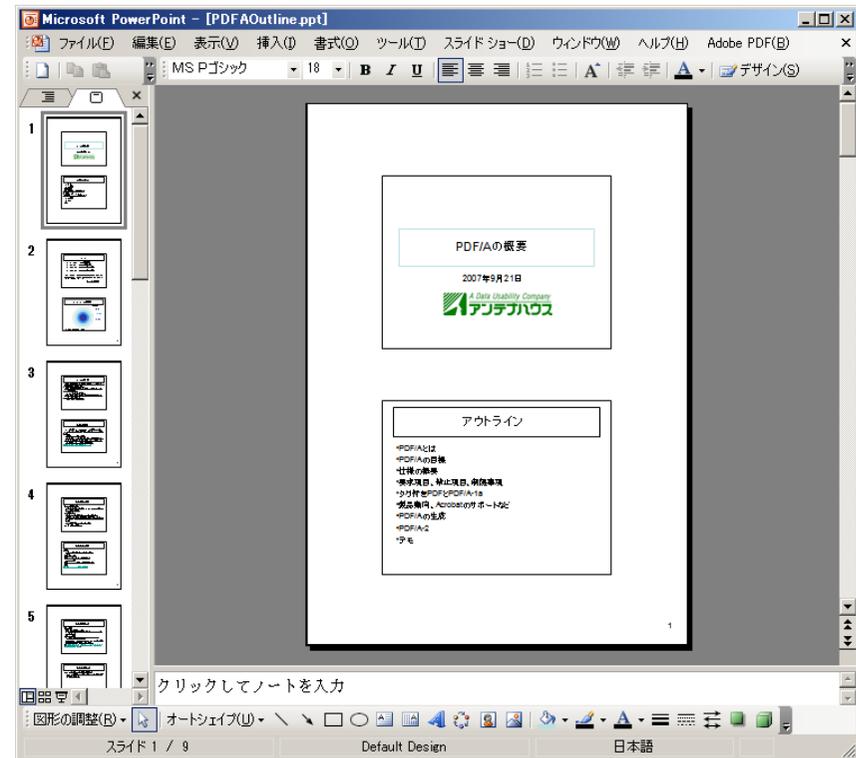
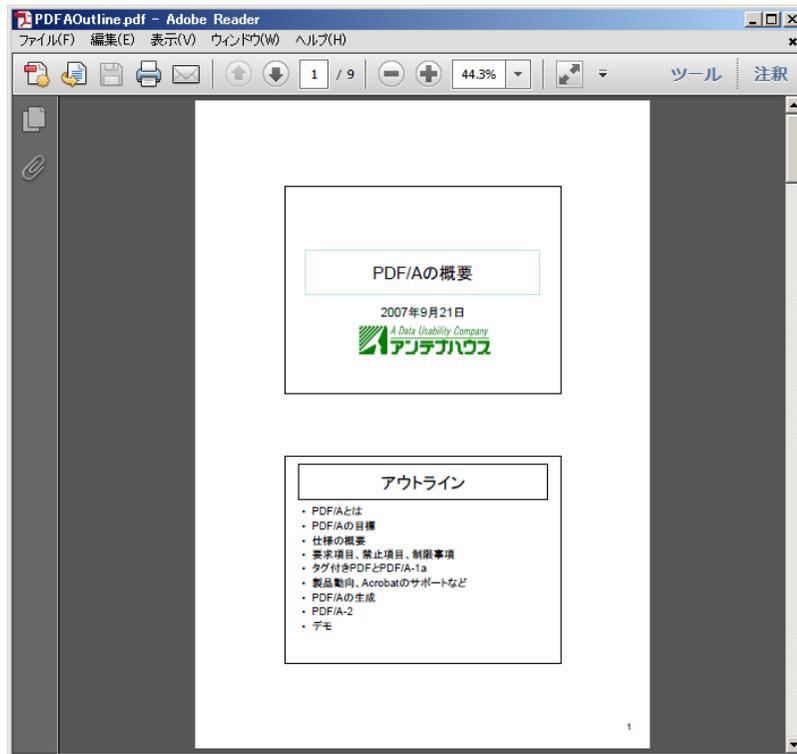
# AH PDF Server – Office変換オプション ～Excel例

- Excel : 文字・表書式・イメージ・線画を変換できます。PDFの1ページをExcelの1シートに変換します



# AH PDF Server – Office変換オプション ～PowerPoint例

- PowerPoint：文字・表書式・イメージ・線画を変換できます。  
PDFの1ページをPowerPointの1スライドに変換します



# AH PDF Server –ライセンスと価格

| 製品名         | 説明  | 価格（税別）   | オプション対応 |   |   |
|-------------|---|----------|---------|---|---|
| プロフェッショナル版  | 基本機能（スタンダード版）に加え、IN/OUTモード・コマンドライン機能が利用可能。<br>同時に起動できるタスク数は「無制限」          | 470,000円 | ○       | ○ | ○ |
| スタンダード版     | 通常のフォルダ監視が可能なエディションで、PDF変換機能はプロフェッショナル版と同等。<br>同時に起動できるタスクの上限は「5」         | 250,000円 | ○       | ○ | × |
| コマンドライン版    | プロフェッショナル版からコマンドライン機能と変換設定アプリケーションを抜粋したエディション<br>PDF変換機能の他システムへの組み込み向け。   | 360,000円 | ○       | ○ | ○ |
| CAD版        | 基本機能に「CAD変換オプション」を標準搭載し、フォルダ監視でのCADファイルのダイレクト変換が可能。<br>同時に起動できるタスクの上限は「1」 | 629,000円 | —       | ○ | × |
| CAD変換オプション  | CADファイルをダイレクトにPDF変換する機能を追加するオプション。プロフェッショナル版、スタンダード版、コマンドライン版で利用可能。       | 540,000円 |         |   |   |
| オフィス変換オプション | PDFや画像ファイルからオフィスファイルへ逆変換する機能を追加するオプション。全てのエディションへ追加可能。                    | 50,000円  |         |   |   |
| Web変換オプション  | Webブラウザからファイルをアップロードして変換する機能を追加するオプション。プロフェッショナル版、コマンドライン版へ追加可能。          | 98,000円  |         |   |   |

※初年度の年間保守費用は価格に含まれております

# AH PDF Server – 動作環境など (1/2)

- 動作環境

『PDF Server』をインストールしてご利用いただくには以下の動作環境が必要です。

- ハードウェア

|         |   |
|---------|---|
| CPU     | Pentium4 1.2GHz以上（もしくは100%互換を持つプロセッサ）<br><b>【推奨】マルチコアCPU 2.0GHz以上</b> |
| メモリ     | 512MB以上<br><b>【推奨】2GB以上</b>   |
| ハードディスク | システムドライブに500MB以上の空き領域<br><b>【推奨】2GB以上の空き領域</b>                        |

※上記ディスク容量には、PDF Server が作成するPDFファイルなどの容量は含みません。

# AH PDF Server – 動作環境など (2/2)

## – ソフトウェア

|   |   |
|---|---|
| <p>オペレーティングシステム<br/>(OS)<br/>(すべて日本語)</p> | <p>Microsoft Windows Server 2012 R2 (64bit)<br/>Microsoft Windows Server 2012 (64bit)<br/>Microsoft Windows Server 2008 R2 (64bit)<br/>Microsoft Windows Server 2008 (32bit/64bit)<br/>Microsoft Windows Server 2003 R2 SP2以降 (32bit/64bit)<br/>Microsoft Windows Server 2003 SP2以降 (32bit/64bit)</p>   |
| <p>その他注意事項</p>                            | <ol style="list-style-type: none"> <li>64bit OSではPDF Server 本体は32bit アプリケーションとして動作します。</li> <li>Windows Server 2003/2003 R2 の場合、AH PDF Server インストール前に「Windows イメージングコンポーネント」のインストールが必要です。(32bit/64bitのいずれか)</li> <li>インストール時に下記のソフトウェアがインストールされます。 <ul style="list-style-type: none"> <li>.Net Framework 4 Client Profile</li> <li>VC++ 2010 再頒布パッケージ</li> </ul> </li> </ol>   |
| <p>Office文書のPDF変換に必要なOffice製品</p>         | <p>Microsoft Office2003/2007/2010/2013 (Excel/Word/PowerPoint/Visio) のいずれか1つ。<br/>ただし、Microsoft Office2010/2013 は32bit 版のみ対応。<br/>※該当するOffice文書変換時に必要。正規版以外は変換保証出来ません。</p>  |
| <p>Web変換オプション利用時に必要な環境</p>                | <p>「.NET Framework4」 (インストール時に無ければインストールを行います)</p> <p>「IIS (Internet Information Services) 6.0/7.0/7.5/8.0」のいずれかのバージョン ※1</p> <p>「Microsoft SQL Server 2008 (R2) /2008R2 (SP2) /2012 (Expressを含めたすべてのエディション) ※2</p> <p>その他</p> <ul style="list-style-type: none"> <li>「オフィス変換オプション」と「CAD変換オプション」は利用できる状態になっている場合は自動で適用します。</li> <li>クライアント側の対応ブラウザ ※3 <ul style="list-style-type: none"> <li>Internet Explorer 6.0 以降</li> <li>Fire Fox 2.0 以降</li> </ul> </li> </ul> |

- ※1 IISはOS毎にバージョンが決まっています。
- ※2 SQL Serverは運用規模でどのエディションを使うか決めて下さい。
- ※3 バージョンによって若干デザイン (見え方) が異なるケースもあります。

# 瞬簡PDFシリーズのご紹介

## デスクトップ製品の冠ブランド「瞬簡PDF」

2010年から弊社デスクトップ製品に採用

- ブランドの強化 ⇒ 「PDFならアンテナハウス」を目指す
- 旧クセロ社のPDF事業、ブランドを継承  
⇒ 現在はすべてアンテナハウス製

**5つのソフトが入って大変お得！**

