

OCRシステム導入のご提案

バックオフィスの紙文章のデータ化 ～AI-OCRの能力と課題～

- バックオフィス業務の課題とOCR
- いいとこどりで活用
- いいとこどりのメリット
- コスト比較
- 総評

【モデルケース詳細説明】

～ DART 機能説明～

概要説明

概念図

モデルケース イメージ詳細

イメージカット機能

イメージ切出し ～DX Suite認識

Prexifort-OCR と DX Suiteの結果をマージ

補正画面(通常版)

補正画面(ボリューム処理 版)

置換テーブル

- AI-OCRの導入モデルのご紹介
e-recognitionシリーズ
- お問合せ先



- ・技術の進歩により、ペーパーレス化やデジタル化が促進。業務の効率化が進み。様々なツールが誕生！！
- バックオフィスでも、右のようなツールで業務の効率化が進む。
- ただ、事前の想定ほど改善が進まない状況・・・



なぜなら、バックオフィス業務では、紙書類は意外に多いから。

ツールを使用するために、紙文章を手入力でデータ化する必要がある。

電子化しづらい書類がある。
例) 押印が必要な契約書や、申請書など

・紙文章を電子化するツールとしての**OCR**が再注目される。

OCRとは：紙をデータ化する技術

固定帳票で高い読取率、そして、安価に利用可能。



さらに、AIの発展とOCRが融合し、AI-OCRの誕生。

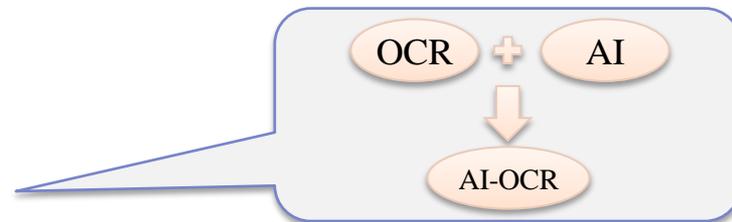


AI-OCRのディープラーニングに対して、個人情報保護のために、電子通信事業における個人情報保護ガイドラインが制定される。



AI-OCRでの特定個人情報を含む帳票の利用が制限される。

認識速度及び、運用コスト面でも課題もあり。



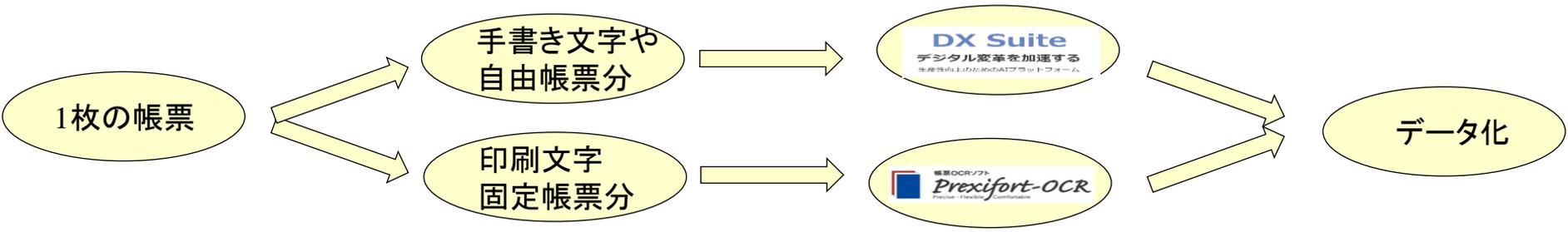
	従来型OCR	AI-OCR
認識率(印刷文字)	◎	◎
認識率(手書き)	△	◎
認識スピード	◎	△
コスト	◎	△

従来型OCR、AI-OCR、+αで得意分野を合わせて活用

従来型OCR	AI-OCR	モデルケース
<ul style="list-style-type: none"> 印刷文字や固定帳票に強い 非常に安価 代表的な商品↓ • Prexifort-OCR 	<ul style="list-style-type: none"> 手書き文字や自由帳票に強い 識字率が非常に高い 代表的な製品↓ • DX suite 	<ul style="list-style-type: none"> 帳票の切り分け作業 認識結果のマージ 置き換えなど • DART

1つの帳票を、DARTにて、それぞれの種別ごとに分割し、自由帳票、手書き文字はAI-OCRで、印刷文字や、固定帳票は従来型OCRで処理。

さらに、特定個人情報を含む帳票は、特定個人情報だけを切り取って画像で認識させる。
→マスキング等と比較しても、画像自体がなくなる為、よりセキュリティの向上が図れる。

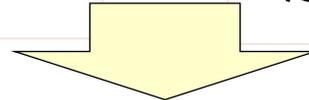


費用対効果の向上！

- AI-OCRと従来型OCRの組合せ
- 処理速度を下げることなく、認識精度が向上

誰でも処理が可能に！

- バックグラウンドで処理を行うため、オペレーションを簡略化。
- だれでも帳票のデータ化が可能に



90%以上の精読率

- つくば市、他4自治体とNTTデータの共同研究による。

データ入力業務の9割の代行を可能に

- データ化した文章をRPAやクラウド、マクロと組み合わせる

AI-OCRでも特定個人情報の認識が可能に

- 帳票データを切り出す為、セキュリティが高まる。

年間で**約50%のコスト削減！！**

(※年間約330万リクエストの場合)※処理状況により変動あり。



オペレーションも簡略化でき、人件費の削減も可能！
監視不要 ※バックオフィスで動作

認識できる帳票の種類が広がる！！
特定個人情報を含む帳票も利用可能に！



今まで以上に様々な文章の
データ化処理を効率化できる！

想定処理数

1日 300(枚) × 30(リクエスト) × 30(日間) × 12(ヶ月) = 324(万リクエスト) = 324(万円)

DX Suite

- 初年度利用料
- 初期費用20万円
- 追加利用料 324万円 **計344万円**
- 次年度以降 **324万円**

モデルケース

- 初年度利用料 326.5万円(リクエスト数が4分の1になったと想定。)
- 次年度以降 132.5万円+追加分21万リクエスト = **156.5万円**

次年度以降
約47.4%削減

課題

処理する枚数が多くなければ、費用対効果を向上させづらい。

	初期費用	年額利用料	追加リクエスト利用料	
DX suite Standard	¥200,000	¥1,200,000	¥0~1	5万リクエストまで無料 5万リクエスト以上は、0.5円~1円
DX suite Pro	¥200,000	¥2,400,000	¥0~1	20万リクエストまで無料 20万リクエスト以上は、0.1円~1円
Prexifort -OCR	¥500,000	¥125,000	0	
AI用Dart	¥1,000,000	0	0	

◆DX Suite: [(導入初期費用)【DX Suite Pro】]=20万円, [(月額費用)【DX Suite Pro】20万円×12か月]=240万円, [(追加リクエスト)324万リクエスト-(無償枠)]=84万円
[(初年度合計金額)]=344万円

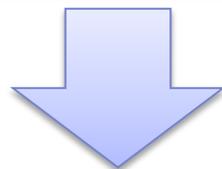
◆モデルケース: [(導入初期費用)【DX Suite Standard】]=20万円 [(AI用DART)]=100万円 [(Prexifort-OCR+保守料金)]=62.5万円 [(月額費用)【DX Suite Standard】]=10万円×12か月=120万円 [(追加リクエスト)81万リクエスト-(無償枠)]=24万円 (初年度合計金額)326.5万円

・バックオフィスの紙文章がすべてデータに移行することはありません。

むしろ、取扱い注意が必要な書類が今後増えることが予想されます。

OCR処理は、精度とコスパの両立が必須です。

従来型OCR、AI-OCR ともに人の目のチェックは必須です。



モデルケースの導入で、精度とコスパの両立を！

※ 当社オリジナルアプリ「DART」を活用したモデルケースをご紹介します

【モデルケース詳細説明】

KSIオリジナルアプリ

「DART」機能紹介



～OCR・AI-OCRの精度向上と最適化を図り、入力処理業務の効率化を実現する～

KSI 株式会社KSインターナショナル

DART機能説明

- **イメージカット機能**
- **マージ機能**
- **認識結果訂正機能**
- **結果置換機能**
- **置換データの登録/修正機能**

Prefixort-OCR で認識できない部分だけを切り出し、DX Suite にアップロードする。切り出すことで、認識速度も改善される。

認識結果の訂正データを蓄積することにより、置換機能の精度をより向上させていく。

「DART」とは、Prefixort-OCR と DX Suite の得意な部分を組み合わせ、認識精度、コストパフォーマンス両面の向上をはかる機能の集合体。※各機能の組み合わせにより、様々な仕様に対応が可能です。

①認識率の向上に

Prefixort-OCRで認識できない部分を DX Suite にアップロードし結果をマージさせることが出来るため認識率が向上します。

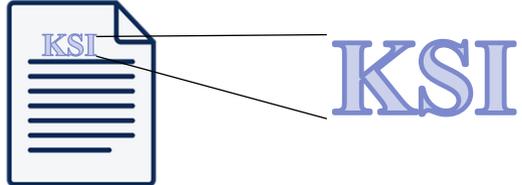
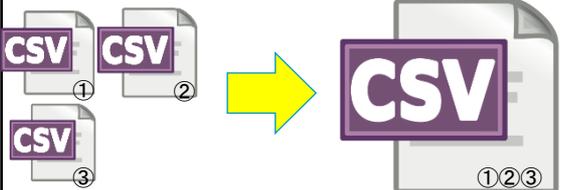
②操作性のしやすさ

DARTでは、イメージ切り出しから認識結果のマージまでバックグラウンドで行うため知識がない方でも簡単に操作が可能です。

③コストの削減

DX Suiteは従量課金制の為、イメージカットで部分的にDX Suiteを使用することでコストを抑えることが出来ます。

DART・・・株式会社 KSインターナショナル製の Prexifort-OCRのオプションツールです。
DX Suite と連携させることにより**低価格で認識率の向上**を望むことが可能です。

機能	機能概要	イメージ
イメージカット	<ul style="list-style-type: none"> 手書き文字など Prexifort-OCR で認識率が望めない部分をイメージカットツールで切り出します。 	
CSV一覧化	<ul style="list-style-type: none"> DART で修正、補正することにより CSVファイルを1つにまとめることが可能になります。 	
データマージ	<ul style="list-style-type: none"> DX Suite で認識したデータを Prexifort-OCR のデータと融合・統一させます。 	

DART 特性

- ※DART 独自の補正画面を備え、CSVファイル出力前にイメージ画面を確認しながらのデータ修正が可能。
- ※DART 独自の学習機能を持つ置換テーブルを備え、過去の補正データを蓄積し、置換修正が可能。

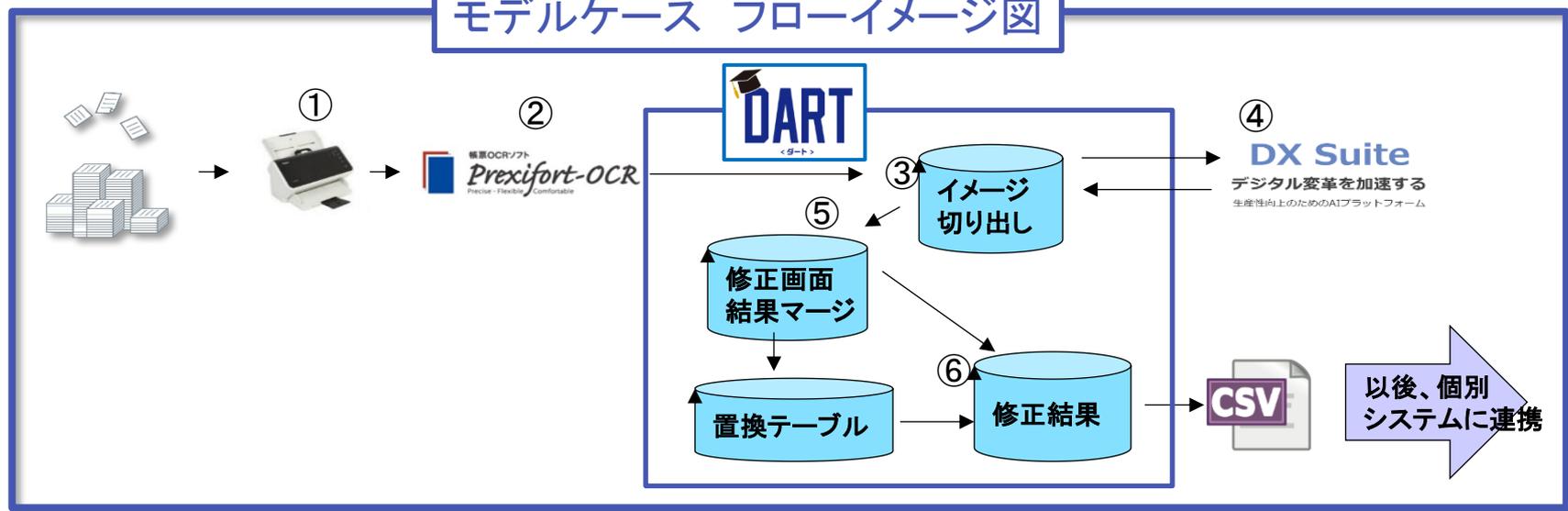
【処理フロー】

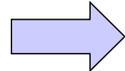
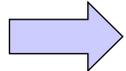
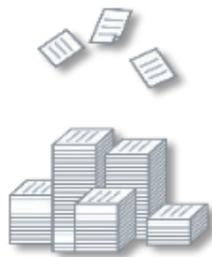
- ①紙資料をスキャナーで読み取る。
 - ②Prexifort-OCRで認識を行う。
 - ③DARTにてDX Suiteに読み込みイメージをカット
 - ④DX Suiteにアップロードを行って認識する。
 - ⑤DARTでPrexifort-OCRとDX Suite の結果をマージする。
 - ⑥マージした結果をCSV形式で出力する。
- ※誤読が多い箇所は、蓄積データから置換することが可能

使用するツール

- Prexifort-OCR
※国内最大シェアを誇るOCR
- AI-OCR
→ DX Suite シェア6割以上のAI-OCR
- DART
→弊社開発ソフト

モデルケース フローイメージ図





Tiff形式 300dpi
白黒でスキャン

Prexifort-OCRで帳票を判別し
読み取り箇所を認識する。

位置情報を使って、
DARTがイメージ切り出し



氏名

所得者の扶養控除等（異動）申告書	
(フリガナ) 田中 一郎	あなた
あなたの氏名 田中 一郎	印
あなたの個人番号 44996437198972	あなた
あなたの住所 (郵便番号)	
又は 居所	

田中 一郎

田中, 郎

候補文字: ↓

個人番号

4 9 9 6 4 3 7 1 9 8 9 7 2

4 ? ? 6 4 3 7 2 8 9 7 2

所得者の	月 1 日	証するに いての 等申告書 提出して は、○ でください
(フリガナ)		
あなたの氏名		
あなたの個人番号 44996437198972	あなたとの職所	
あなたの住所 (郵便番号)	配偶者の有無	
又は 居所		

かつ、あなた自身が障害者、寡婦、寡夫又は勤労学生のいずれにも該当しない場合には、以下の各欄に記入する必要があります

(フリガナ)	タナカ イチロウ	あ
あなたの氏名	田中 一郎 印	い
あなたの個人番号	419191614379189712	あ
あなたの住所 又は居所 (郵便番号)	-	

2. あなた自身が障害者、盲聾、盲又又は働かざる者の養育者である

DARTが
イメージ切り出し



田中 一郎 印

田中 一郎

田中一郎

DARTが
それぞれを別画像で
DX Suite へ アップロード

DX Suite

デジタル変革を加速する
生産性向上のためのAIプラットフォーム

419191614379189712

419191614379189712

499643798972

令和2年分 給与所得者の扶養控除等(異動)申告書

氏名	田中 一郎
個人番号	419191614379189712
住所	東京都江東区333

0/4項目完了 (1/1枚目)

ご請求書

下記の通りにご請求申し上げます。

ご請求金額 **506,000 円**

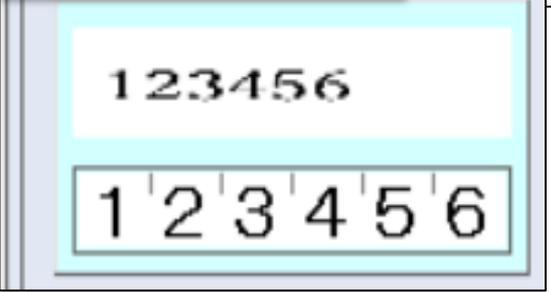
No	商品番号	品名	数量	単価	合計
1	123456	スキャナー	1	300,000	300,000
2	22	パソコン	1	100,000	100,000
3	33	ソフト	1	60,000	60,000

DX Suite
デジタル変革を加速する
生産性向上のためのAIプラットフォーム



手書き文字や、複数候補がある場合は
DX Suite で認識！

帳票OCRソフト
Prexifort-OCR
Precise · Flexible · Comfortable

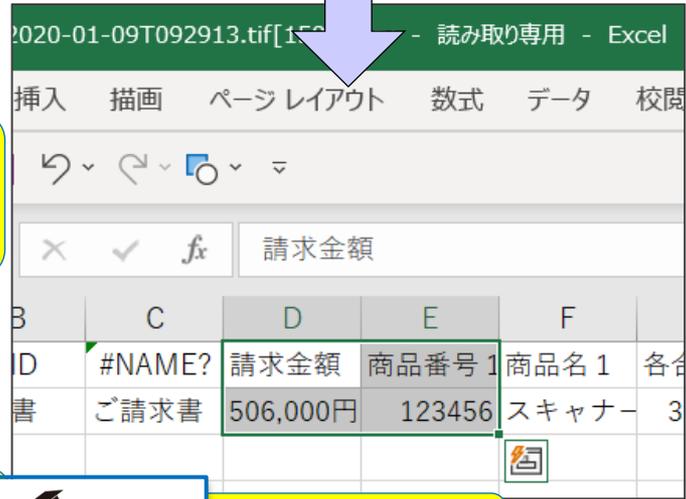


印刷文字や、固定文字の場合は、
Prexifort-OCRで認識！

切り出し画像のアップロード
認識データのダウンロード
データのマージ



置換テーブルの作成
データ補正画面作成



CSVで出力

【DART(通常版)の補正画面】

DARTなら データ補正用の
修正画面を作成できる



OCR結果訂正画面

506,000 円

品名	数量	単価	合計
スキャナー	1	300,000	300,000
パソコン	1	100,000	100,000
ソフト	1	60,000	60,000
		小計	460,000
		消費税	46,000
		合計	506,000

フォルダ: デモ20220915133229 ファイル: K0001.tif.xml

請求書

項目名	OCR結果	訂正内容
請求金額	506,000円	506,000円
品番番号1	123456	
品番番号2	234567	
品番番号3	345678	
品名1	スキャナー	
品名2	パソコン	
品名3	ソフト	
数量1	1	
数量2	1	
数量3	1	
単価1	300,000	
単価2	100,000	
単価3	60,000	
合計1	300,000	

DARTでは、修正画面を作成できるため、CSVファイル出力前にイメージ画像を見ながらの修正が可能。

※通常、Prexifort-OCRのバッチ版を使用する場合は、修正画面が無く、データ補正はCSV出力後になる

【DART(ボリューム処理版)の補正画面】

ボリューム処理版なら、
更に補正しやすい修正画面を作成できる



OCR結果訂正画面

フォルダ: KSI20200901

ファイル: 00001.tif.xml, 00002.tif.xml, 00003.tif.xml, 00005.tif.xml, 00006.tif.xml, 00008.tif.xml, 00010.tif.xml

選択中のフォルダ: KSI20200901

選択中のファイル: 00008.tif.xml

フォルダ内エラー: 49 ファイル内エラー: 6 エラーサーチ

請求書 確定時に次エラー表示

項目名	OCR結果	訂正内容
ご請求金額	506,000	
商品番号1	123456	
商品番号2	224466	
商品番号3	333333	
品名1	スキャナー	
品名2	パソコン	
品名3	ソフト	
数量1	1	
数量2	1	
数量3	1	
単価1	300,000	
単価2	100,000	
単価3	60,000	
合計1	300,000	
合計2	100,000	
合計3	60,000	
小計	460,000	
消費税	46,000	
総合計	506,000	

ご請求金額: 506,000

No	商品番号	品名	数量	単価	合計
1	123456	スキャナー	1	300,000	300,000
2	224466	パソコン	1	100,000	100,000
3	333333	ソフト	1	60,000	60,000
小計					460,000
消費税					46,000
合計					506,000

銀行名: 千葉銀行 神崎支店
口座番号: 普通 123456
口座名義: KS インターナショナル

以上、よろしくお願いたします。

縮小 25% 拡大

終了 CSV一括出力 訂正確定

大量の帳票データにも対応するDARTボリューム処理版の補正画面。
DART通常版の補正画面と比較すると、エラーサーチ機能、
ファイル内エラー数の表示があるため、更に補正を行いやすい。
また、DX Suiteの認識フラグを色分けで表示するのでエラーの発見が容易になります。

【DARTの置換テーブル】

DARTなら 補正データを蓄積し
置換テーブルを作成できる



OCR処理後の修正結果に対して
ログを集計。

置換テーブルで置換を行う
判定項目を登録する。

登録後は、判定項目に対して
置換処理が行われる。

置換テーブル作成指示

置換ログ登録 ログ参照/テーブル登録 ログメンテナンス

作成状況

- フォルダ数 3 件
- ファイル数 8 件
- 項目数 16 件

置換テーブル作成指示

置換ログの登録が完了しました

OK

置換テーブル作成指示

置換ログ登録 置換テーブル登録 ログメンテナンス

帳票名 請求書

ファイル名 デモ20200206111127/2020-01-09T092913.tif.xml

項目名	変更前/判定条件	変更後/置換内容	判定項目	置換項目
請求金額	? 06,000円	506000円	<input type="checkbox"/>	<input checked="" type="checkbox"/>
商品番号1	123456		<input type="checkbox"/>	<input type="checkbox"/>
商品名1	abodefg		<input type="checkbox"/>	<input type="checkbox"/>
各合計1	300,000		<input type="checkbox"/>	<input type="checkbox"/>
商品番号2	224466		<input type="checkbox"/>	<input type="checkbox"/>
合計	506,000		<input checked="" type="checkbox"/>	<input type="checkbox"/>
各合計2	100,000		<input type="checkbox"/>	<input type="checkbox"/>
商品番号3	333333		<input type="checkbox"/>	<input type="checkbox"/>
商品名3	ソフト		<input type="checkbox"/>	<input type="checkbox"/>
各合計3	60,000		<input type="checkbox"/>	<input type="checkbox"/>
小計	460,000		<input type="checkbox"/>	<input type="checkbox"/>
商品名2	パソコン		<input type="checkbox"/>	<input type="checkbox"/>

帳票名を判定項目として登録する

登録時に同じ判定項目のログを削除する (XMLファイル単位)。

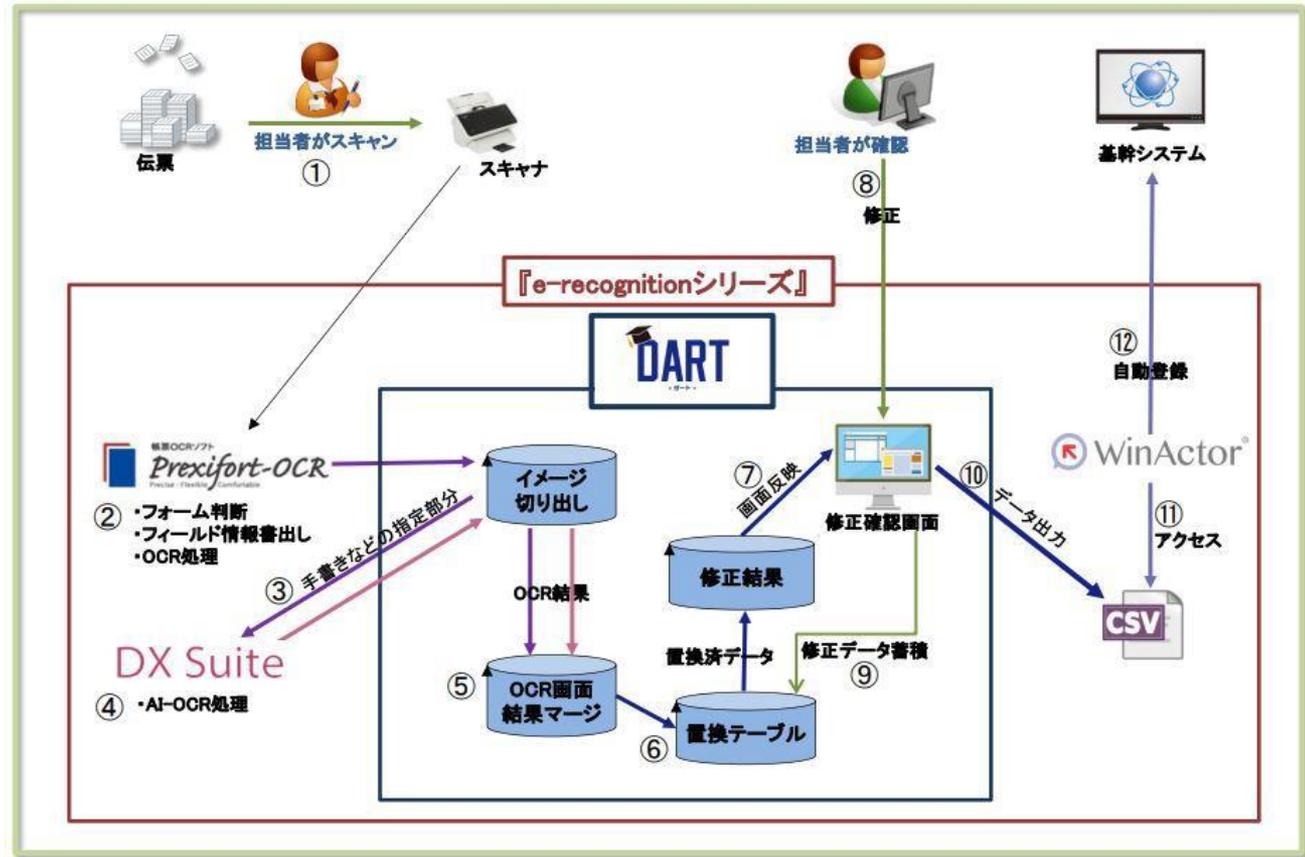
登録 戻る

終了

【導入モデル例】 e-recognitionシリーズ

※従来型OCR (Prexifort-OCR) と AI-OCR (DX Suite) のメリットを、
 当社オリジナルアプリ「DART」が組み合わせ、
 RPA (WinActor) で基幹システムへの登録も自動化するパッケージプラン

【モデルケース フローイメージ図】



※紹介動画を配信中
 イーレコシリーズ
 イーレコBPO で検索！

※料金等、詳細は
 「NTTデータ
 DXマーケットプレイス」
 (<https://nttdata-mp.com/>)
 でご確認ください！

※WinActorは、NTTアドバンステクノロジー株式会社の登録商標です。
 ※DX Suiteは、日本国内におけるAI inside株式会社の登録商標です。

見積りや、効果検証などは下記へお問い合わせください

- 問合せ先

TEL 03-6666-4106

✉ Info@ksin.jp

- 株式会社 KSインターナショナル

- <https://ksin.jp/>

東京都江東区扇橋2丁目1-7 クレープメトルハウス 3F

【営業担当者】

矢崎 裕大

✉ Yazaki.yuta@ksin.jp