

## AnyForm OCR の稼働検証

今年の1月に株式会社ハンモック様にて OCR ハンズオン研修に参加いたしました。  
最近、巷で RPA とともに注目されている OCR ビジネスの製品の「AnyForm OCR」を ISC のデモ環境にて稼働検証を実施しました。

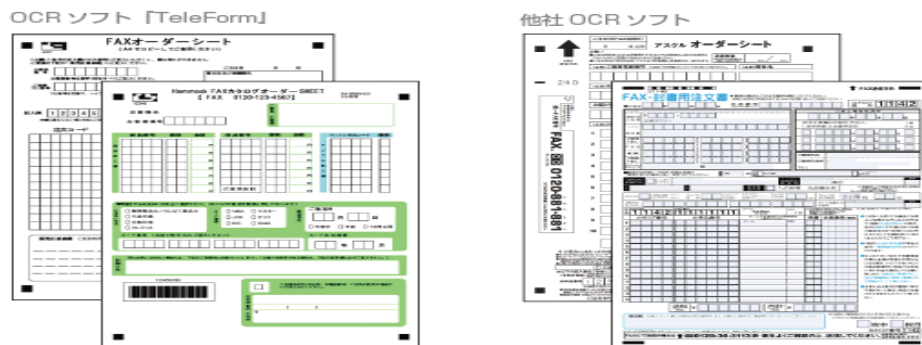
### 背景・課題

お客様がご利用中の FAX 受信や手書き伝票などの帳票をスキャナーなどを使用してスキャンし電子化（PDF、TIF、JPEG など）したもののや、すでに帳票から電子化されたものを Anyform OCR にて取込みデータ化して取込み精度の確認判断に伴う検証を弊社イグアスで実施しています。

#### ●一部の限られた帳票しか、OCR 処理しかなかった

従来 OCR の目的は、専用帳票で手書き文字を認識することでした。

自社で帳票を設計し、印刷したものを取引先に利用してもらう必要があるため、一部の限られた帳票だけしか利用できません。



しかし、取引先は独自の基幹システムを導入しており、そこから出力される帳票（既存帳票）のレイアウトは多種多様です。

OCR 活用の幅を広げるためにはこれらの帳票に対応する必要があります。

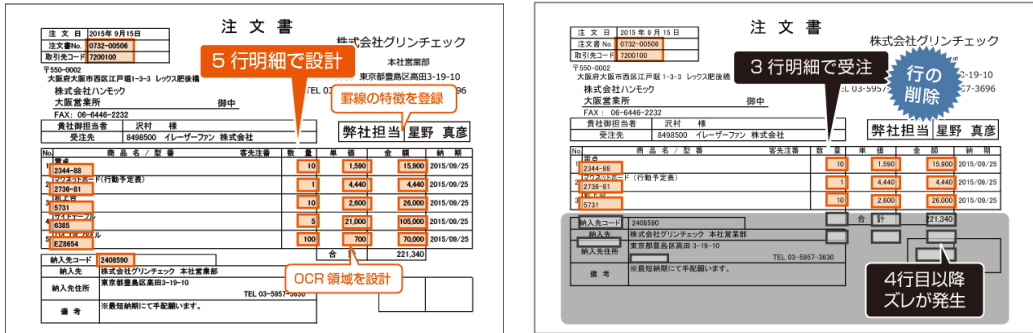
The image shows two examples of invoices. The left one is a '注文書' (Purchase Order) from '株式会社ハンモック' with a table of items. The right one is a '請求書' (Invoice) from '株式会社ハンモック' with a table of items and a total amount.

| 品名コード | 商品名      | 数量  | 単価  | 金額     |
|-------|----------|-----|-----|--------|
| 0001  | りんご      | 100 | 100 | 10,000 |
| 0002  | バナナ      | 150 | 100 | 15,000 |
| 0003  | パイナップル   | 100 | 150 | 15,000 |
| 0004  | オレンジ     | 200 | 100 | 20,000 |
| 0005  | グレープフルーツ | 100 | 200 | 20,000 |
| 0006  | 梨        | 100 | 200 | 20,000 |
| 0007  | いちじく     | 100 | 200 | 20,000 |
| 0008  | みかん      | 100 | 200 | 20,000 |
| 0009  | りんごジュース  | 100 | 200 | 20,000 |
| 0010  | りんごジュース  | 100 | 200 | 20,000 |
| 0011  | りんごジュース  | 100 | 200 | 20,000 |
| 0012  | りんごジュース  | 100 | 200 | 20,000 |
| 0013  | りんごジュース  | 100 | 200 | 20,000 |
| 0014  | りんごジュース  | 100 | 200 | 20,000 |
| 0015  | りんごジュース  | 100 | 200 | 20,000 |
| 0016  | りんごジュース  | 100 | 200 | 20,000 |
| 0017  | りんごジュース  | 100 | 200 | 20,000 |
| 0018  | りんごジュース  | 100 | 200 | 20,000 |
| 0019  | りんごジュース  | 100 | 200 | 20,000 |
| 0020  | りんごジュース  | 100 | 200 | 20,000 |
| 0021  | りんごジュース  | 100 | 200 | 20,000 |
| 0022  | りんごジュース  | 100 | 200 | 20,000 |
| 0023  | りんごジュース  | 100 | 200 | 20,000 |
| 0024  | りんごジュース  | 100 | 200 | 20,000 |
| 0025  | りんごジュース  | 100 | 200 | 20,000 |
| 0026  | りんごジュース  | 100 | 200 | 20,000 |
| 0027  | りんごジュース  | 100 | 200 | 20,000 |
| 0028  | りんごジュース  | 100 | 200 | 20,000 |
| 0029  | りんごジュース  | 100 | 200 | 20,000 |
| 0030  | りんごジュース  | 100 | 200 | 20,000 |
| 0031  | りんごジュース  | 100 | 200 | 20,000 |
| 0032  | りんごジュース  | 100 | 200 | 20,000 |
| 0033  | りんごジュース  | 100 | 200 | 20,000 |
| 0034  | りんごジュース  | 100 | 200 | 20,000 |
| 0035  | りんごジュース  | 100 | 200 | 20,000 |
| 0036  | りんごジュース  | 100 | 200 | 20,000 |
| 0037  | りんごジュース  | 100 | 200 | 20,000 |
| 0038  | りんごジュース  | 100 | 200 | 20,000 |
| 0039  | りんごジュース  | 100 | 200 | 20,000 |
| 0040  | りんごジュース  | 100 | 200 | 20,000 |
| 0041  | りんごジュース  | 100 | 200 | 20,000 |
| 0042  | りんごジュース  | 100 | 200 | 20,000 |
| 0043  | りんごジュース  | 100 | 200 | 20,000 |
| 0044  | りんごジュース  | 100 | 200 | 20,000 |
| 0045  | りんごジュース  | 100 | 200 | 20,000 |
| 0046  | りんごジュース  | 100 | 200 | 20,000 |
| 0047  | りんごジュース  | 100 | 200 | 20,000 |
| 0048  | りんごジュース  | 100 | 200 | 20,000 |
| 0049  | りんごジュース  | 100 | 200 | 20,000 |
| 0050  | りんごジュース  | 100 | 200 | 20,000 |
| 0051  | りんごジュース  | 100 | 200 | 20,000 |
| 0052  | りんごジュース  | 100 | 200 | 20,000 |
| 0053  | りんごジュース  | 100 | 200 | 20,000 |
| 0054  | りんごジュース  | 100 | 200 | 20,000 |
| 0055  | りんごジュース  | 100 | 200 | 20,000 |
| 0056  | りんごジュース  | 100 | 200 | 20,000 |
| 0057  | りんごジュース  | 100 | 200 | 20,000 |
| 0058  | りんごジュース  | 100 | 200 | 20,000 |
| 0059  | りんごジュース  | 100 | 200 | 20,000 |
| 0060  | りんごジュース  | 100 | 200 | 20,000 |
| 0061  | りんごジュース  | 100 | 200 | 20,000 |
| 0062  | りんごジュース  | 100 | 200 | 20,000 |
| 0063  | りんごジュース  | 100 | 200 | 20,000 |
| 0064  | りんごジュース  | 100 | 200 | 20,000 |
| 0065  | りんごジュース  | 100 | 200 | 20,000 |
| 0066  | りんごジュース  | 100 | 200 | 20,000 |
| 0067  | りんごジュース  | 100 | 200 | 20,000 |
| 0068  | りんごジュース  | 100 | 200 | 20,000 |
| 0069  | りんごジュース  | 100 | 200 | 20,000 |
| 0070  | りんごジュース  | 100 | 200 | 20,000 |
| 0071  | りんごジュース  | 100 | 200 | 20,000 |
| 0072  | りんごジュース  | 100 | 200 | 20,000 |
| 0073  | りんごジュース  | 100 | 200 | 20,000 |
| 0074  | りんごジュース  | 100 | 200 | 20,000 |
| 0075  | りんごジュース  | 100 | 200 | 20,000 |
| 0076  | りんごジュース  | 100 | 200 | 20,000 |
| 0077  | りんごジュース  | 100 | 200 | 20,000 |
| 0078  | りんごジュース  | 100 | 200 | 20,000 |
| 0079  | りんごジュース  | 100 | 200 | 20,000 |
| 0080  | りんごジュース  | 100 | 200 | 20,000 |
| 0081  | りんごジュース  | 100 | 200 | 20,000 |
| 0082  | りんごジュース  | 100 | 200 | 20,000 |
| 0083  | りんごジュース  | 100 | 200 | 20,000 |
| 0084  | りんごジュース  | 100 | 200 | 20,000 |
| 0085  | りんごジュース  | 100 | 200 | 20,000 |
| 0086  | りんごジュース  | 100 | 200 | 20,000 |
| 0087  | りんごジュース  | 100 | 200 | 20,000 |
| 0088  | りんごジュース  | 100 | 200 | 20,000 |
| 0089  | りんごジュース  | 100 | 200 | 20,000 |
| 0090  | りんごジュース  | 100 | 200 | 20,000 |
| 0091  | りんごジュース  | 100 | 200 | 20,000 |
| 0092  | りんごジュース  | 100 | 200 | 20,000 |
| 0093  | りんごジュース  | 100 | 200 | 20,000 |
| 0094  | りんごジュース  | 100 | 200 | 20,000 |
| 0095  | りんごジュース  | 100 | 200 | 20,000 |
| 0096  | りんごジュース  | 100 | 200 | 20,000 |
| 0097  | りんごジュース  | 100 | 200 | 20,000 |
| 0098  | りんごジュース  | 100 | 200 | 20,000 |
| 0099  | りんごジュース  | 100 | 200 | 20,000 |
| 0100  | りんごジュース  | 100 | 200 | 20,000 |

●なぜ、取引先システムから出力される帳票に OCR が活用できなかったのか？

① 可変明細に対応できない。

- ・罫線の位置や長さなどの特徴を元に、OCR 領域を設計します。  
しかし、設計した行数と異なる場合、帳票下部データでアジャストに失敗します。  
これにより OCR 項目の認識位置がズレてしまい、正確に文字を認識ができません。



② OCR 化に掛かるコスト（時間、費用）

- ・認識精度の高い帳票作成は難易度が高く、メーカーに依頼する必要があり、時間と費用がかかってしまいます。



③ 低い文字認識精度

- ・既存帳票 OCR は補正が弱いため、認識精度が低いという問題がありました。
- ・特に FAX 受信する帳票は、縮小や文字の擦れが発生してしまうため適しません。



---

## 製品概要

---

製品名 AnyForm OCR V5.1

『AnyForm OCR』は、取引先システムから出力される帳票を OCR 処理します。  
明細行がある帳票では、注文量に応じて行数が変更される場合があります（可変明細）。  
従来 OCR は、文字認識する場所が特定できないこのような帳票には利用することができませんでした。

しかし、『AnyForm OCR』は独自の画像認識処理により、従来 OCR の課題であった可変明細帳票への OCR 処理を可能とします。

高精度に文字を認識し、スピーディーなデータ化を実現します。



① 可変明細に対応！

独自の画像処理により、明細行の終端を認識することが可能。  
注文書や納品受領書などに対する OCR の活用範囲の拡大。



② 取引先帳票の OCR 化設計がかんたん

従来 OCR に比べ、非定型帳票を OCR 読み取りできるようにする設定が格段に簡略化。  
自社で OCR 帳票の設定ができるため、時間とコストを削減。



③ 高精度な文字認識率

画像補正や OCR 項目を正確に認識するための仕組みにより、  
FAX でも正確に文字認識することが可能。

## ● 4つのステップと機能

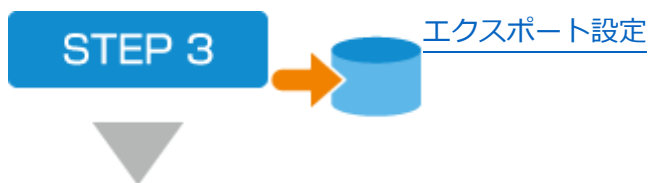
AnyForm OCR をご利用いただくための4つのステップと機能をご紹介します。



既存帳票に対して OCR 処理する項目を設定します。  
かんたん操作で短時間に設定できます。



帳票レイアウトに合わせて  
OCR 結果を確認/修正する画面を設計します。



OCR 結果をエクスポートする項目や書式を設定します。



FAX サーバーやスキャナーなど受付処理の設定や  
各種マスタの登録を行います。

## ●STEP1 – OCR 帳票化設計

AnyForm OCR は、OCR 帳票設計が不要な「キーワード指定方式」と、デザイナーツールでかんたんに OCR 帳票設計できる「座標指定方式」の2つの方法で OCR 項目の文字認識設定が行えます。

※キーワード指定方式で設定し、文字認識精度をさらに上げたい場合には座標指定方式で設計を行う。

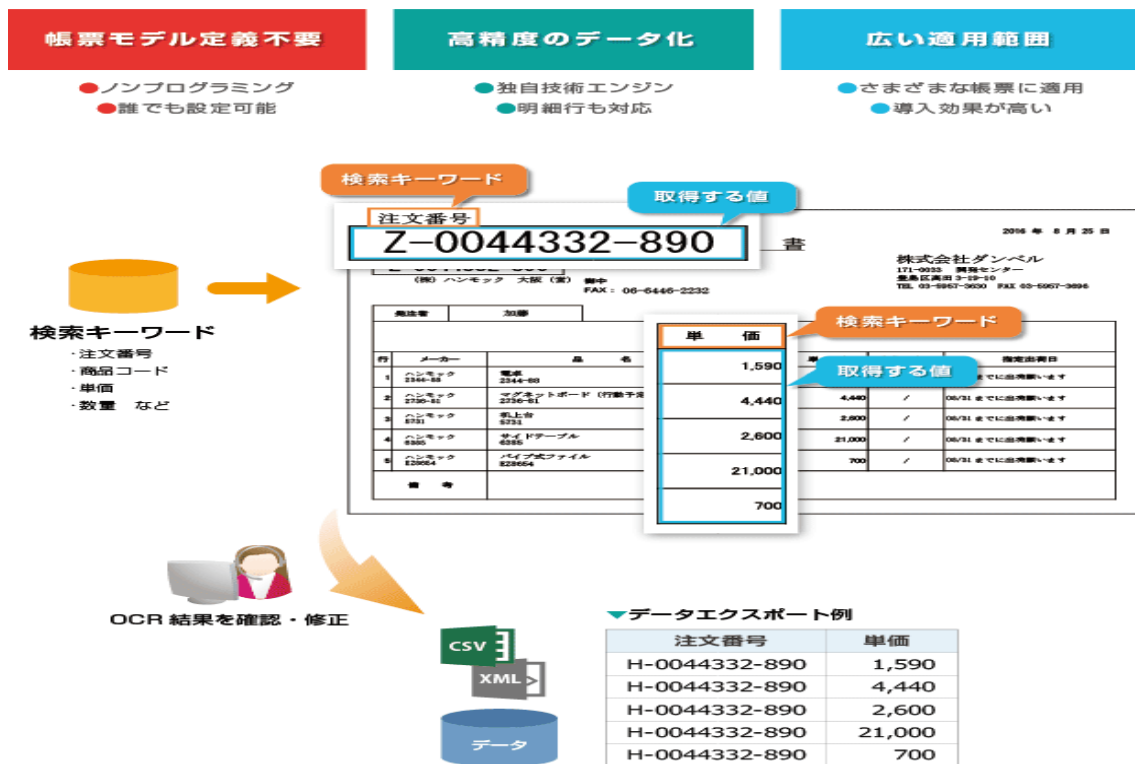
### キーワード OCR 設定 (キーワード指定方式) New

帳票設計不要！OCR 処理したい項目名称をキーワードとして登録する方法。

さまざまな帳票に活用でき、大量な帳票をすばやくデータ化します。

帳票画像ファイルの中から、事前登録した検索キーワードを認識し必要な値を取得します。

従来 OCR の仕組みとは全く異なり、レイアウトに依存しないため、さまざまな帳票にご活用いただけます。



## OCR 帳票設計 (座標指定方式)

OCR 処理したい項目のエリア (座標) を指定して、OCR 帳票を設計する方法。

非定型帳票のイメージファイルに対して OCR 項目を設定できます。

OCR 項目をマウスでドラックし、必要項目を選択するだけでかんたんに作成できます。

## 帳票レイアウト自動認識

取引先帳票 (既存帳票) をスキャナーで読み取りイメージファイルにします。

イメージファイルを AnyForm OCR デザイナーで取り込むことで**帳票レイアウトを自動認識**します。

## OCR 項目設定

OCR 項目を選択し、読み取るエリアをマウスで指定します。

ノンプログラミングで設計できるため、誰でも操作できます。

②読み取るエリアを指定

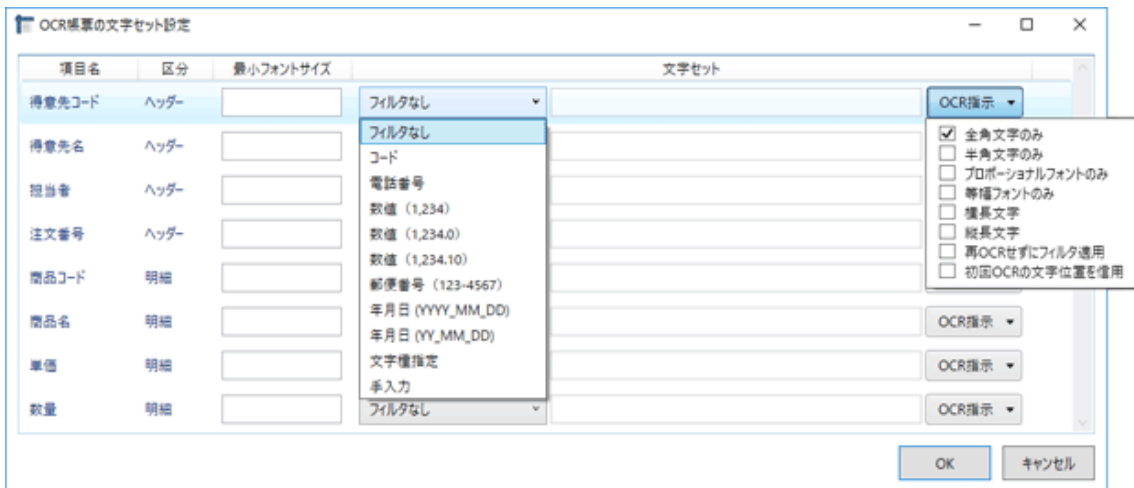
①OCR項目を選択

| 行 | メーカー   | 品名      | 数量 | 単価    | 計 |
|---|--------|---------|----|-------|---|
| 1 | パナソニック | 2394-88 | 10 | 1,598 |   |
| 2 | パナソニック | 2736-81 | 1  | 4,548 |   |
| 3 | パナソニック |         | 10 | 2,600 |   |

OCR処理するエリアを  
マウスでドラッグするだけ!

OCR 処理する項目に対して文字種の設定を行います。

文字種を指定することで文字認識精度が向上します。

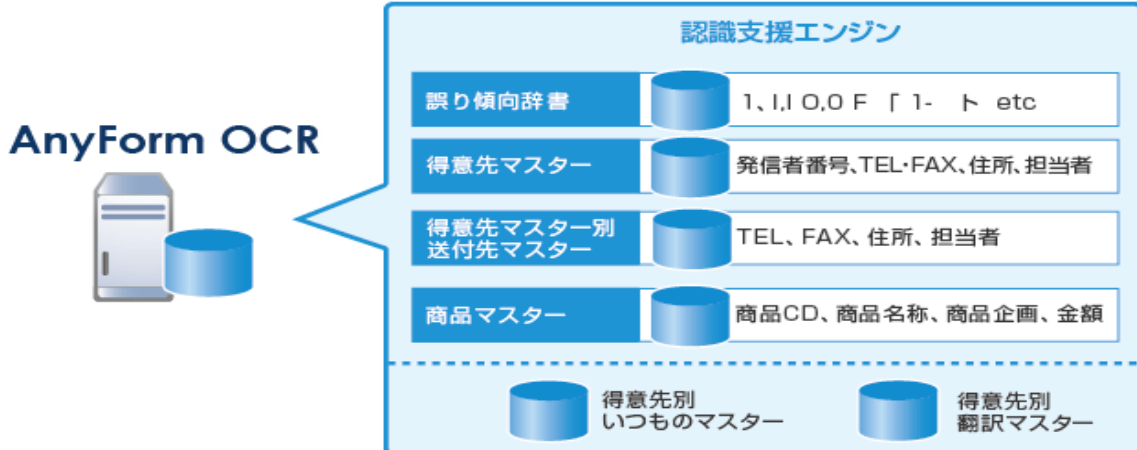


【フィルタ】 数字のカンマ区切りや日付の桁数などを指定

【OCR 指示】 全角/半角やフォントの飾り文字などを指定

## 文字認識支援

非定型帳票から読み取った値を正確に文字変換するための認識支援エンジンを搭載。



---

## OCR 検証したパターン

---

- ① パターン 1 (ハンズオンサンプル帳票)
  - ・ヘッダー項目、明細項目で固定枠有り
  - ⇒キーワード OCR または、座標 OCR にて対応
- ② パターン 2
  - ・ヘッダー項目、明細項目、固定枠無し
  - ⇒座標 OCR にて対応
- ③ パターン 3 (手書き帳票)
  - ・ヘッダー項目、明細項目
  - ⇒座標 OCR で手書き対応
- ④ パターン 4
  - ・ヘッダー項目が 1 つで、明細項目が左右 2 列の場合
  - ⇒座標 OCR 対応で明細の読み込み順序を考慮して対応
  - ※通常は上から左右の順番だが、左側の列処理後、右側の列の順番に設定可能。
- ⑤ パターン 5
  - ・ヘッダー項目と明細項目が 1 ページに 2 つずつ存在
  - ⇒2 つのヘッダー要素と、明細を 1 つ要素として 1 レコード出力するイメージ。
  - ※CSV 出力後、分割する作業を RPA などでも考慮が必要なので注意願います。
- ⑥ パターン 6
  - ・ヘッダー項目、明細項目が可変明細の場合
  - ⇒座標 OCR で終端文字列識別にて対応
  - ※終端文字列識別にて可変明細も対応可能です。
- ⑦ パターン 7
  - ・ヘッダー項目、明細項目が非定型明細の場合
  - ⇒座標 OCR で「いつものマスター」機能にて対応
  - ※「いつものマスター」とは、明細項目を固定登録表示し、数量のみを入力する形態。

※共通の認識向上処理としては、「マスター一覧」や「認識支援辞書」使用して対応。



## 所感

ハンモック様のハンズオン研修を約 3 時間程度受講することで、基本的な設計することが出来るようになります。またサポート体制もあり、質問には的確に回答いただけます。

イメージ画像の傾き、文字の大きさや濃さにも柔軟に対応されているが補正機能で問題はありません。

認識支援辞書やマスター一覧機能により、さらに認識率を解消できること。

ベリファイ画面では、誤認識のコードの場合、修正入力しマスター一覧を参照して関連情報を入力することなく展開されます。

働き方改革の観点から、1 から入力した作業に比べれば作業効率化が図れると考えます。

また、OCR 機能で出力されたデータを RPA と連携して、基幹システムへのデータエントリ処理の連携自動化することでさらに効率化が見込まれると考えます。

### ※AnyForm OCR + AutoMate 連携イメージ

