

### スクリプト概要

本アプリケーションは複数のスクリプトを使用し、処理を行っております。

スクリプト名	種類	説明
ScannedPdf_To_Markdown_Claude_convert	メイン	LLM (Claude) を活用して、スキャンしたPDFファイルからテキストを抽出し、Markdown形式に保存
ScannedPdf_To_Markdown_Claude_validate_limits	サブ	スキャンしたPDFファイル変換の所要トークン数の制限値チェック
ScannedPdf_To_Markdown_Claude_get_max_page	サブ	スキャンしたPDFファイルのトータルページ数を取得
ScannedPdf_To_Markdown_Claude_convert_to_markdown	サブ	スキャンしたPDFファイルからテキストを抽出し、ファイルに保存

### LLM (Claude) 利用方法

本アプリケーションはAnthropic社のClaudeを用いて処理を行っております。  
ClaudeのAPI-Keyを利用するための準備作業をご説明いたします。

- Anthropic社のサイト > API から「Console log In」を選択してください。  
URL：https://www.anthropic.com/claude
- コンソールのログイン画面 > メールアドレスを入力し、画面の指示に従ってアカウント登録してください。
- 入力したメールアドレスにClaudeにアクセスするためのメールが届きます。クリックしてサインインしてください。
- サインインしたコンソール画面 > Get API Key をクリックしてください。
- API Keys の画面 > Create Key をクリックしてください。  
※クレジット購入前の場合、API-Key作成前にクレジット情報登録・クレジット購入が必要となります。
- Create API Key 画面でWorkspace、API-Keyの名前を入力してください。API-Keyが発行されます。

### コネクション設定

コネクタータイプ	REST接続
URL設定	https://api.anthropic.com

### 可変長書き込み設定

Claudeのレスポンスをファイルに書き出す際に、「可変長書き込みフォーマット定義」を利用します。  
後述の設定手順にて「可変長書き込みフォーマット定義」を作成のうえ、可変長書き込み処理コンポーネントのプロパティ設定ダイアログにて、フォーマットの設定を行ってください。

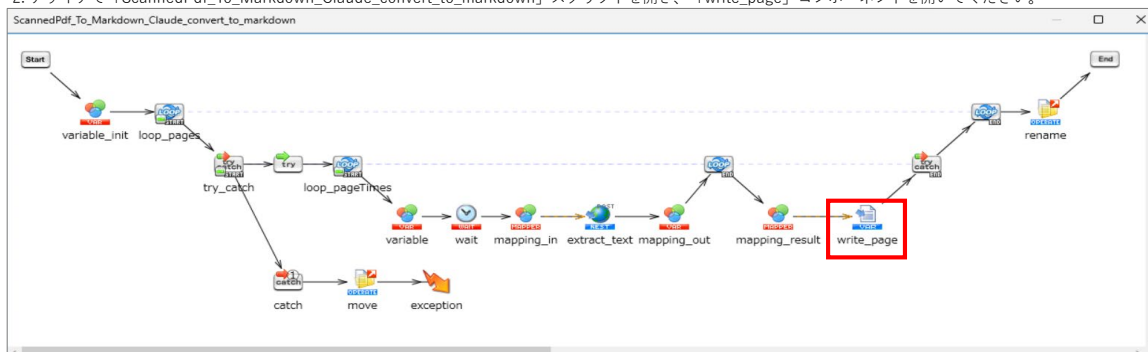
<可変書き込み 設定手順>

1. HULFT Squareサイト > データ形式 から「新規追加」を選択し、「可変書き込みフォーマット定義」を作成してください。

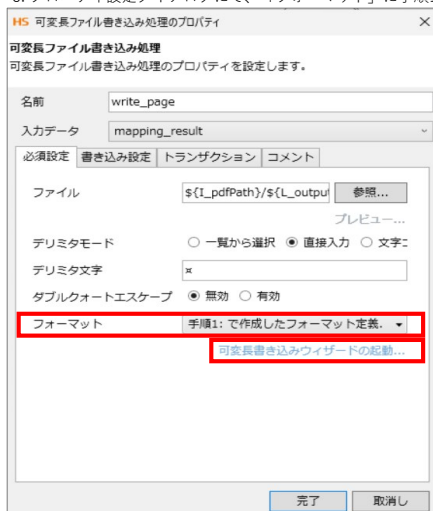
※ 「名前」「ワークスペース」には任意の値を設定してください



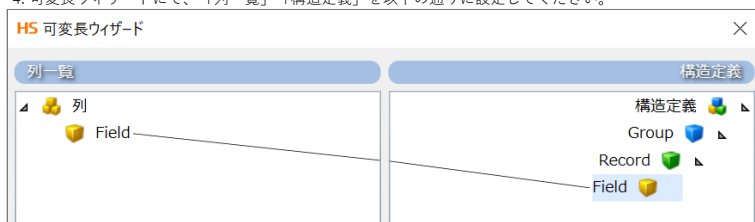
2. デザインで「ScannedPdf\_To\_Markdown\_Claude\_convert\_to\_markdown」スクリプトを開き、「write\_page」コンポーネントを開いてください。



3. プロパティ設定ダイアログにて、「フォーマット」に手順1.で作成した可変書き込みフォーマット定義を選択し、「可変書き込みウィザードの起動...」を押してください。



4. 可変長ウィザードにて、「列一覧」「構造定義」を以下の通りに設定してください。

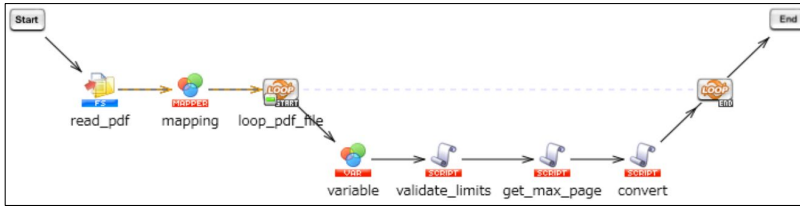


可変書き込み設定は以上です。

**スクリプト名**

ScannedPdf\_To\_Markdown\_Claude\_convert

LLM (Claude) を用いて、スキャンしたPDFファイルからテキストを抽出し、Markdown形式に変換します。



**初期設定項目(スクリプト変数)**

項目名	必須/省略可	説明
l_anthropicAPIKey	必須	事前に取得したAPIキー
l_ITPM	必須	インプットトークン数/秒 (Anthropicのコンソールから取得するITPM)
l_modelName	必須	利用するAIモデル名
l_requestInterval	省略可	APIを実行する時間間隔 (初期値: 120秒)
l_pdfPath	必須	PDFファイルを格納するフォルダのフルパス
l_pageConvertTimes	省略可	出力精度向上のため処理を繰り返す回数(初期値: 3)
l_onlyFilled	省略可	記入済み項目のみを抽出するスイッチ(0: オフ 1: オン 初期値: 1)

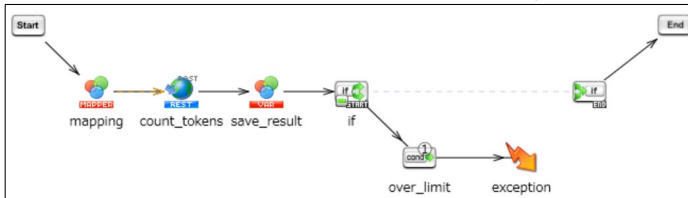
**備考**

スクリプトコピー後、実行前にスクリプト呼び出しコンポーネントを開き[完了]ボタンを押下

**スクリプト名**

ScannedPdf\_To\_Markdown\_Claude\_validate\_limits

スキャンしたPDFファイル変換の所要トークン数の制限値をチェックします。



**初期設定項目(スクリプト変数)**

項目名	必須/省略可	説明
なし	-	-

**スクリプト名**

ScannedPdf\_To\_Markdown\_Claude\_get\_max\_page

スキャンしたPDFファイルのトータルページ数を取得します。



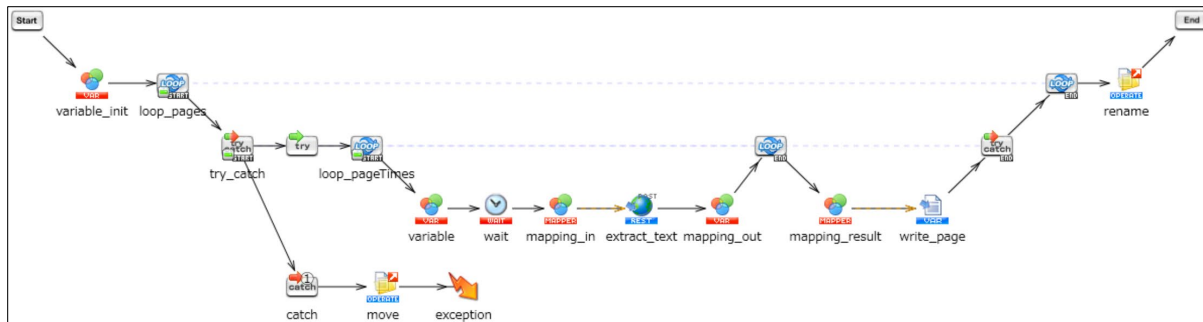
**初期設定項目(スクリプト変数)**

項目名	必須/省略可	説明
なし	-	-

スクリプト名

ScannedPdf\_To\_Markdown\_Claude\_convert\_to\_markdown

PDFのページごとに、Claudeにリクエストしたテキスト抽出結果を可変長ファイルとして出力し、Markdownファイルに変換します。



初期設定項目(スクリプト変数)

項目名	必須/省略可	説明
なし	-	-