

TECNOLOGÍA

D2dSBI SMALL BUSINESS INVOICES

EXTRACCIÓN AUTOMÁTICA DE DATOS Y VERIFICACIÓN DE FACTURAS DE PROVEÏDORES

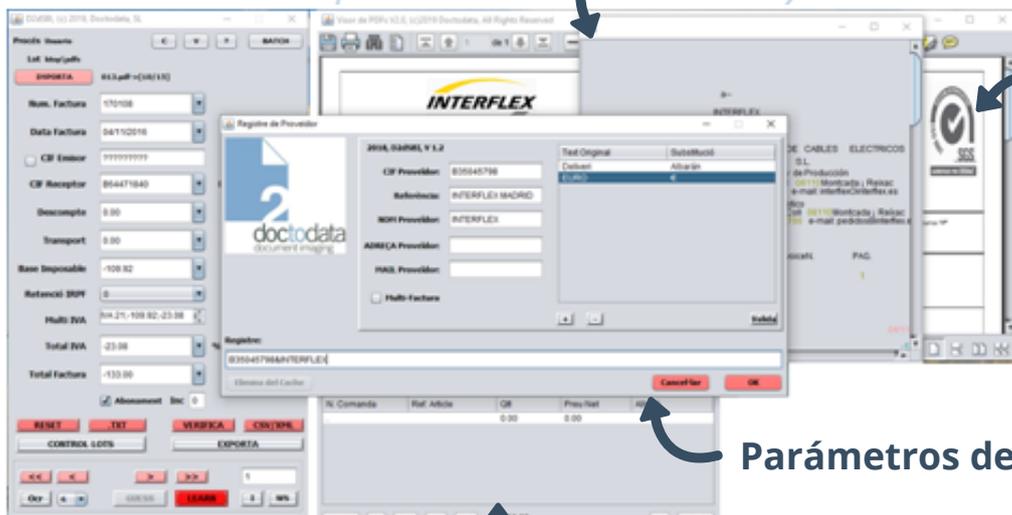
Datos de cabecera de la factura

Modelización y segmentación del documento por tipos de datos

Visor de la factura amb possibilitat de navegació i de fer anotacions

Parámetros de aprendizaje

Líneas de detalle



FUNCIONALIDAD BÁSICA

Importación de documentos

- En cualquier formato (Pdf, Tif, Jpg, Png)
- Origen de documentos diverso (File System, Unidad remota, Samba, FTP/sFTP, Mail, SQL, WS-RESTFul/SOAP, Scan-To-Folder, Hot Folder)
- Interactiva y desatendida ('Cron')
- Separación de facturas por código de barras, logo, Registro de texto, Registro o imagen-huella
- Crea 4 representaciones complementarias del documento

Pre-procesado de las imágenes

- Filtro de mediana para uniformar el fondo de página
- Perfilado y mejora del contraste del texto según histograma
- Escalado
- Diversos algoritmos de interpolación
- Limpieza de ruidos (diversas estrategias)
- 'Deskew' y recorte automático de la imagen
- Diversos algoritmos de Binarización (Otsu)

OCR

- Multi-idioma
- Entrenable y configurable
- Posibilidad de adaptar otros motores OCR comerciales (Nuance, ABBYY, ...) para funcionamiento por votación
- Aplicación de correctores contextuales locales o vía WS
- Funcionamiento a través de la distancia de Levenshtein
- Ocr 'On-The-Fly'



Extracción de Datos

- Automática con aprendizaje
- Por exploración de la imagen y aplicación de patrones: 'Regex'
- Flexible, permite establece posiciones absolutas i relativas, expresiones regulares, delimitadores de campos, 'Lehv. Distance'
- Registro de documentos: Por texto, Logo, Correlación de imagen
- Adaptado al client: maestro de proveedores/ WSs

Proceso / Verificación

- El proceso de extracción puede ser 100% desatendido
- WS de acceso a Albaranes, Pedidos y Maestro de proveedores
- Campos de captura con 'suggestors' a partir de datos obtenidos
- Ocr Zonal
- Ocr 'On-The-Fly'

Exportación (Opciones posibles, a definir y a valorar en cada caso)

- Formatos y compresiones de salida
- Nombres de los ficheros
- Destino: 'File System', SQL, WSNotificaciones vía MailPdf/a2-1b
- Firma electrónica con certificado
- Inserción de datos de la factura en el Pdf de salida (XMP)
- Inserción de códigos QR con los datos extraídos de la factura
- Marcas de agua
- Lotificación, logs y traza de actividad