

D2d

PdfDataExtractor

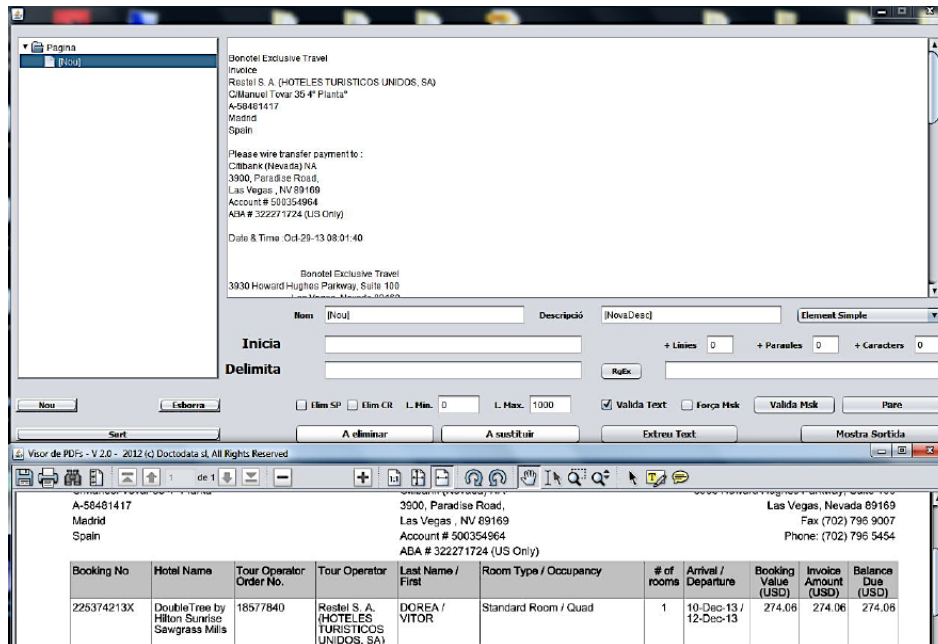
CLASSIFICACIÓ I EXTRACCIÓ AUTOMÀTICA DE
DADES CONTINGUDES EN EXCELS I PDF'S A
PARTIR DE PLANTILLES

The image shows the PdfDataExtractor (c)2016 Doctodata Ver 3.12 interface on the left and a Microsoft Excel spreadsheet on the right. The software interface includes fields for 'pdf/excel' (AD567640.59 22062018 T.xlsx), 'plantilla' (novaPlantilla.xml), and buttons for 'EDITA PLANTILLA', 'SALVA PLANTILLA', 'LOGIN', 'Importa', 'Valida', 'Visor ICE', and 'ISO-8859-1'. It also has a 'Batch' section with 'Carpeta Pdfs ó Excel', 'Plantilla', 'Carpeta Pntils', 'Carpeta Errors', 'Carpeta Sortida', 'Esborra OUT i ERRORS', and 'Go Batch'. The Excel spreadsheet shows a table with columns A through S, containing data such as 'PUJOLASOS CASAD', 'ABADIAS MORALES', 'DE JESUS MOLINA', 'HEBA DEL VALLE', 'MOLINA COLOMER', 'GUTIERREZ PAULS', 'JULIA BALLES IR', 'COIET BOTYA ALI', 'FORA BORRELL JO', 'ALVAREZ AUGUST', 'NAVARRO CASAS O', 'PEREZ ARAGON MA', 'HARTINEZ PUIG S', 'QUINONES VILCA', 'IGUAL MOSCARDO', 'LLASAT SOLER JO', 'HIDROZ TRIADO M', 'GARRIDO SERRANO', 'SALVADOR DIOS L', 'SALVADOR DIOS L', 'GARCIA ORTIZ TR', 'GARCIA ORTIZ AL', 'SEQUERA SAMPEDE', 'SEQUERA SAMPEDE', 'ORELLANA DIEZ J', 'MAS GRAN MIRIA', 'TURRILLO SOCARR', 'BOCA GOMEZ REI', 'SANCHEZ BLAZQUE', 'HONCUNILL SAEZ', 'ALARCÓN FERRAS', 'HASSO ZAPATER A', 'PEÑA GOMEZ ROGE', 'MILLAN CUADRA L', 'MUR TORRENTO DA', 'NEILON ZARZA ANU', 'MORENO ARSENOL', 'ESQUITUS CORTES', 'AMICO PEREZ EST', 'AZHAY BOROWSKA', 'AMILL ALBA', 'FIGUERAS GARCIA', 'FIGUERAS MINN J', 'BALLBE BLASI JA', 'JANARDIS IZUA', 'CANO SAGARRA MA', and 'ARIS AMARO M PA'.

Doctodata SL.
C/ Fígols, 19-21
08028 Barcelona
Tel. 934.211.824
www.doctodata.com

CARACTERÍSTIQUES

- ✓ Multi-idioma
- ✓ Basat en Java, compatible amb qualsevol plataforma
- ✓ No té límit de número de documents
- ✓ Utilitza components OpenSource amb llicència de lliure distribució (BSD, Apache 2.0)
- ✓ Mode de funcionament Interactiu i també ('Batch')
- ✓ Incorpora les llibreries de referència 'Open Source'



FUNCIONALITAT BÀSICA

IMPORTACIÓ DE DOCUMENTS

- En format (PDF, Excel)
- Suporta diversos visors i codificacions de caràcters
- Origen de documents divers
- Interactiva i desatesa ('Cron')
- Separació de pàgines opcional
- Generació de xmls amb dades de cada document
- Inventari CSV amb dades globals del procés
- Reanomenat dels Pdfs de Sortida Opcional



EXTRACCIÓ DE DADES

- Automàtica a partir de plantilla
- Per exploració del document i l'aplicació de regles patrons: 'Regex'
- Flexible. Permet establir: Posicions absolutes i Relatives, Expressions regulars, Delimitadors de camps, 'Lehv. Distance', etc
- Registre de documents: Per text
- Adaptat al client: FTP / WSs

EXPORTACIÓ

- Formats i compressions de sortida
- Nombres dels fitxers
- Notificacions via Mail
- Signatura electrònica amb certificat
- Inserció de dades extretes en el mateix Pdf de sortida (XMP)
- Inserció de codis QR amb les dades extretes
- Marques d'aigua
- Lotificació, logs i traça d'activitat

APLICACIONS

- Extracció de dades de:
 - Factures i albarans
 - Contractes
 - Llistat de facturació de mútues mèdiques
 - Formularis