

***X-Center
Solution
Portfolio***

Improved Invoice Capture and Data extraction for SAP VIM



Our solution aims to automate in a fully reliable way the customer order capture and extraction processes.

As a result, the Sales Administration clerks can rely on data accuracy and directly focus on the Sales Order creation control in SAP

1. The Pain Points

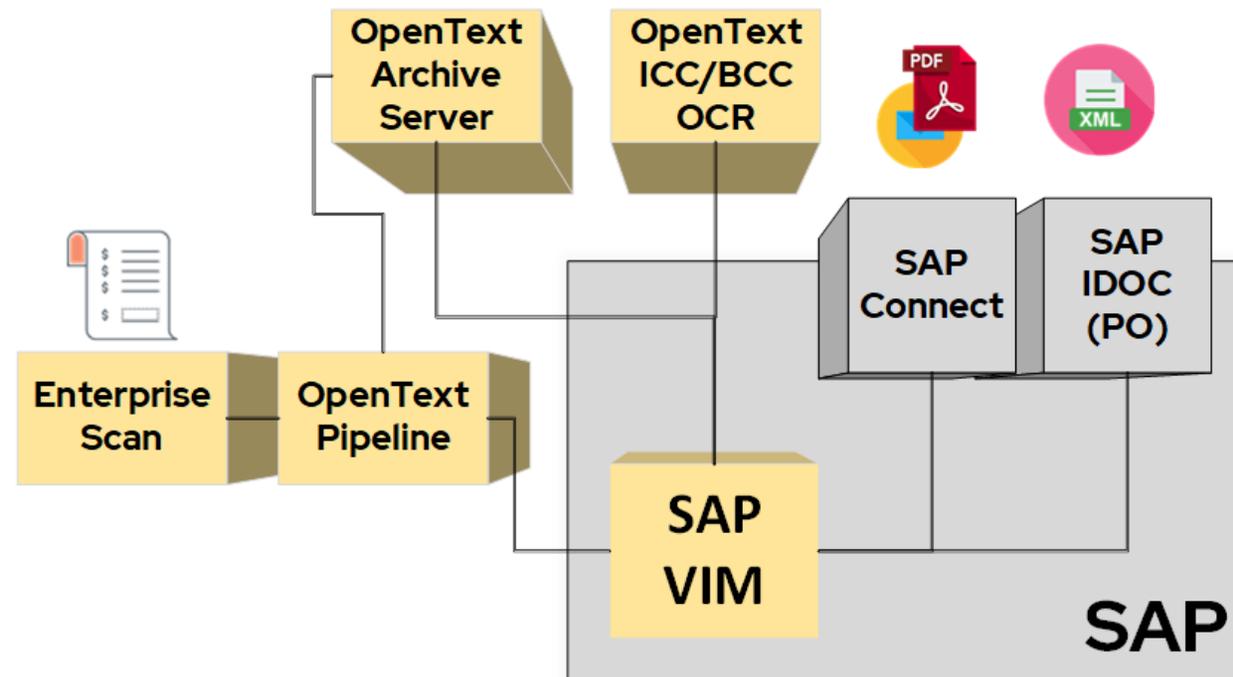
In the B2B arena more than 70% of customers are not received from EDI or electronic transmission, and this percentage is not poised to decrease quickly. Customer preference goes for the PDF files attached to eMails, as it combines all the advantages for themselves : flexible + fast + legal cheap + straight forward.

But for the vendors, who are expected to please their customers by all means, it means:

- Intensive workload: process incoming eMails, detach and classify PDF orders, manual customer input in his ERP system
- Costly errors in the order creation process (customer selection article numbers, quantities ...)
- Disgruntled customers, when the delivery does not match the order due to errors
- Backlog of orders awaiting process, with delayed turnover as a result



2. VIM solution – Capture and extraction components



INVOICE CAPTURE



Document Capture

VIM recommends using **EnterpriseScan** for the **Invoice Paper Capture**. This is an excellent Scan application, very well design for vendor invoice capture. But rather than triggering the capture from the SAP transaction OAWD, we propose a “pipeline” solution which “pushes” the scanned images to SAP VIM, and avoid the following disadvantages:

- Need of SAPgui for the Scan operator
- Documents transferred “on the fly” to the Archive Server
- No possibility to define some values during the scanning process

- No possibility to feed a local Cache Server

VIM recommends using **SAP-Connect** for the **Vendor eMail Capture**. This is a very basic solution which properly detaches the attached PDF documents. But this does not address the different challenges due to particular AP clerks requirements and the diversity of the eMail contents.

EDI capture and integration via **IDOC**'s is the perfect solution, once all the significant agreements and efforts have been realized. But it does not address two basic requirements : generate a readable image out of the EDI file, adapt EDI file content when it does not fully comply with accounting requirements.

DATA EXTRACTION

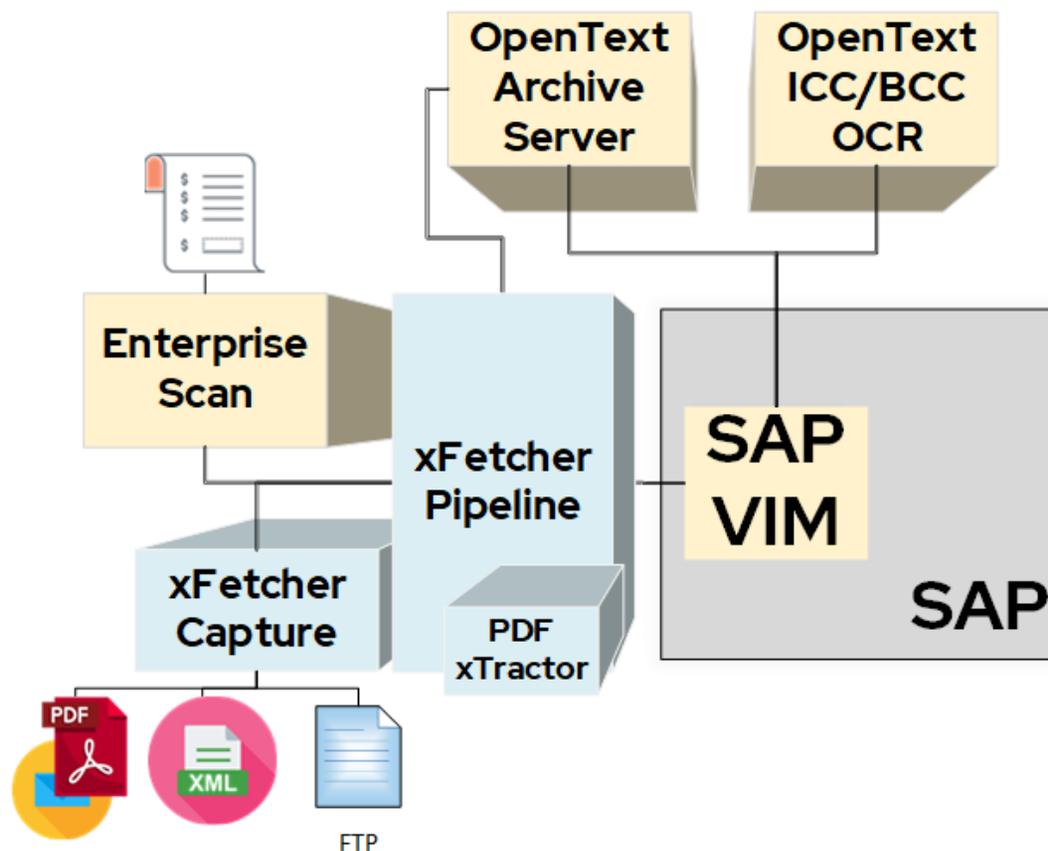
The quality of data extraction based on OCR, depends on several factors, which clearly define the limits of this technology:

- OCR recognize characters based on pixels, with a lot of risks of wrong interpretation
- Some fields are identified with the help of keywords, which may be relevant for most vendors but not for all
- OCR training is time consuming, not transparent and subject to a lot of limitations



**Data
extraction**

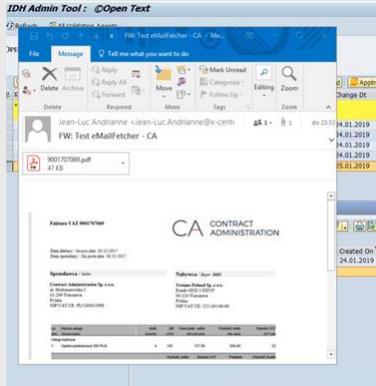
3. Our proposition



INVOICE CAPTURE

Paper invoices: we propose to add an indexing screen with some scripts inside the EnterpriseScan, to send out the scanned images to the Archive Server. This way the Scan operator can select all necessary fields during the scanning process and does not need access to SAP.

eMail invoices: we propose a Windows solution **xFetcher Capture**, which can be connected to multiple mailboxes where the vendors eMail their invoices.



```

<?xml version="1.0" encoding="utf-8"?>
<Invoice xmlns:udt="urn:un:unescefact:data:draft:UnqualifiedDataTypeschemaM-
<cbc:UBLVersionID>2.1</cbc:UBLVersionID>
<cbc:ContractReferenceID>urn:www.cenbii.eu:transaction:bi1tsma010:ver2.0:extended:u
<cbc:ProfileID>urn:www.cenbii.eu:profile:bi1l05:ver2.0</cbc:ProfileID>
<cbc:ID>2190638</cbc:ID>
<cbc:IssueDate>2019-05-31</cbc:IssueDate>
<cbc:InvoiceTypeCode>listID="UNCL1001">380</cbc:InvoiceTypeCode>
<cbc:Note />
<cbc:TaxPointDate>2019-05-31</cbc:TaxPointDate>
<cbc:DocumentCurrencyCode>listID="ISO4217">EUR</cbc:DocumentCurrencyCode>
<cbc:OrderReference>
<cbc:ID>80450003342</cbc:ID>
</cbc:OrderReference>
<cbc:AdditionalDocumentReference>
<cbc:ID>TAXI_JENNY_BVBA_VER_2190638(7).pdf</cbc:ID>
<cbc:DocumentType>CommercialInvoice</cbc:DocumentType>
<cbc:Attachment>
<cbc:EmbeddedDocumentBinaryObject mimeCode="application/pdf" fileName="TAXI
</cbc:Attachment>
<cbc:AdditionalDocumentReference>
<cbc:ID>EFPF</cbc:ID>
<cbc:DocumentType>Maibooks on Mab 2.4.0</cbc:DocumentType>
</cbc:AdditionalDocumentReference>
<cbc:AccountingSupplierParty>
<cbc:Party>
<cbc:EndpointID schemeID="URI:CBCE">BE0445791412</cbc:EndpointID>
<cbc:PartyName>
<cbc:PartyName>
</cbc:PartyName>
</cbc:PartyName>
<cbc:StreetName>Diestsesteenweg 458</cbc:StreetName>
<cbc:AdditionalStreetName>/</cbc:AdditionalStreetName>
<cbc:CityName>REPERE-LO (LEWOPB)</cbc:CityName>
<cbc:PostalZone>3010 </cbc:PostalZone>

```

It includes a very extended functionality:

- eMail also archived and attached to the SAP DP process
- unique sequential number can be added on each invoice image
- URL invoices can also be downloaded automatically
- eMails without PDF files are forwarded to AP clerks
- PDF files which are not invoices, are not sent to the OCR engine
- number of pages to be sent to the OCR engine, can be limited by sender

The Windows application consists of a Web-based application, to perform the customizing , monitoring and reporting.

EDI invoices: due to the new EDI UBL2 format commonly adapted in all European countries and worldwide, a growing number of vendors do not see the advantage to send their invoices to the Service Providers and prefer to send their EDI + PDF invoices via eMails to their customers, or make them accessible via sFTP server or Web service.

Our **xFetcher Pipeline** solution can also process eMails which include EDI + PDF files, or EDI files only, with some additional functionality

- generate a readable PDF file if necessary
- process UBL2 and Fattura format
- Add unique sequential number of PDF file first page
- Fine-tine extraction (e.g. : capture the purchase order numbers even if not present where they should be present)

much on the company. It can be defined by number ranges (e.g. « 45 » + 8 digits) or preferably it should be matched with the list of active purchase order numbers present in SAP. This field is very useful because:

- it directly influences the financial posting line items
- it can be used to « double-check » other data (vendor, company code, currency ...)

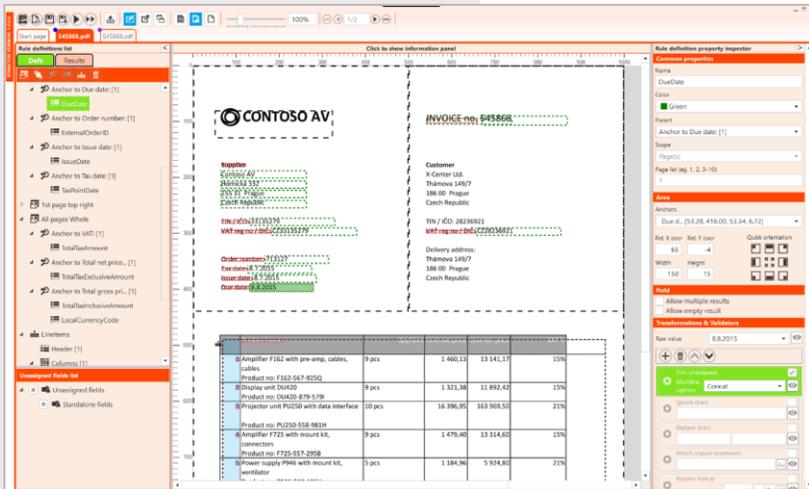
DATA EXTRACTION

Why do we propose to combine another data extraction method to the existing

ones, when no EDI agreement was made yet ?

- The vendors with large invoice volume most of time send their invoice as PDF file attached to eMail, and not paper invoices
- These PDF files are most of time "Readable PDF". It does not make sense to translate these Readable PDF into graph files, in order to process them with OCR ... with the risk not to recognize characters or digits properly
- EDI is technically the ideal format, but the numerous barriers will prevent from becoming the most widespread format soon

Our solution **xTractor** is aimed to process the vendors with large volumes, who send their PDF invoices via eMail. **xTractor** helps create extraction template by vendor, with the help of a graphical interface. We can guarantee 100% accuracy, even on line item extraction.



Our solution **xFetcher** can combine the vendor invoices to be processed by **xTractor** (e.g. the 20% largest vendors who usually generate about 80% of the invoices), and OCR for the vendors with low invoice volume (the long tail).

INTEGRATION WITH SAP-VIM

The main objective of the document capture and data extraction automation, is to help prepare the financial document preparation accurately and smoothly.

We propose to help SAP financial document preparation as follows:

- also archive the eMail and make it accessible from SAP DP VIM
- invoice header data : also transfer additional information like payment reference and vendor IBAN numbers (in order to check vendor master data)
- invoice line item: transfer line item fields which can be used to generate financial document line items

