

3 Beiträge / 1 neu

[Forenübersicht \(/de/forum\)](/de/forum/)

[Schnelle Antwort \(/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-form\)](/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-form)

[Erster ungelesener Beitrag \(/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-\)](/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-)

[Letzter Beitrag \(/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-1602\)](/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields#comment-1602)

[Beitrag abonnieren \(/de/notifications/subscribe/content_thread?0=118195&destination=node/118195×tamp=1540403167&skip=1&signature=5c860011b6cae84f50fbc6374fed6203\)](/de/notifications/subscribe/content_thread?0=118195&destination=node/118195×tamp=1540403167&skip=1&signature=5c860011b6cae84f50fbc6374fed6203)

01/09/2015

[#1 \(/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields\)](/de/forum/english-forums/docuware-questions-about-usage-and-configuration/using-xml-file-extract-index-fields)

[Mauricio Valencia](#) Using an XML file to extract index fields

Using an XML file to index scanned images

Punkte: 2

The new fiscal regulations in Mexico require to all private organizations the integration of electronic invoicing systems.

All approved systems generate an XML with all legal information and a graphic representation of

the invoice in PDF format.

The specific requirement of partners is to configure Docuware using DW Import to receive all vendor invoices (XML, PDF) and configure and autoindex environment to extract all index fields from the XML file to populate our database for future document retrievals and workflow approval processes

The XML files have the same legal structure regulated for the Fiscal authorities

Any similar experience? any recommendations?

REGARDS

Oben

01/09/2015

[#2 \(/de/comment/501#comment-501\)](#)

Mauricio Valencia Using an XML file to index scanned images

Punkte: 2

The new fiscal regulations in Mexico require to all private organizations the integration of electronic invoicing systems.

All approved systems generate an XML with all legal information and a graphic representation of the invoice in PDF format.

The specific requirement of partners is to configure Docuware using DW Import to receive all vendor invoices (XML, PDF) and configure and autoindex environment to extract all index fields from the XML file to populate our database for future document retrievals and workflow approval processes

The XML files have the same legal structure regulated for the Fiscal authorities

Any similar experience? any recommendations?

REGARDS

Oben

03/24/2016

[#3 \(/de/comment/1602#comment-1602\)](#)

Craig Williams

Using dwcontrol metadata feature

Punkte: 64

Importing Files using the control file -DWCONTROL-

posted Mar 24, 2016, 12:25 PM by Craig Williams

Docuware allows you to import records in ANY format if you use the XML file to point to the records. You must have Docuware Import to do this.

You will need to build a DWCONTROL file from the XML and images in that XML for example: I get a XML file and PDF from our client containing the Data and the Document (PDF) and sometimes an image (JPG) that all need clipped together and stored into Docuware with the correct indexes in place.

There is a PDF named: 635935515966691387_13cfdffe-98ef-40ad-acc2-c887be14714d_10061.pdf

An XML file named: 635935515966691387_13cfdffe-98ef-40ad-acc2-c887be14714d_10061.XML

The XML Looks like this:

```
"?xml version="1.0" encoding="UTF-8"?>
```

```
<ImportGroup xmlns="http://anothercompany/schemas/dataImport">
```

```
http://anothercompany/schemas/dataImport">)
```

```
<Document id="635935515966691387_13cfdffe-98ef-40ad-acc2-c887be14714d_10061"
```

```

name="635935515966691387_13cfdffe-98ef-40ad-acc2-c887be14714d_10061" domain="BPV"
index="Certification" typeName="Certificate" indexQueue="Data Import Queue"
convertFile="false" autoIndex="true" ocr="false">
  <Field name="LocationCity" value="SomeCity" />
  <Field name="LocationName" value="Some Company Name" />
  <Field name="LocationAddress" value="1122 Electric Ave" />
  <Field name="LocationState" value="IL" />
  <Field name="LocationZip" value="60201-4205" />
  <Field name="LicenseNumber" value="" />
  <Field name="JurisdictionNumber" value="B010XXXX" />
  <Field name="DocDate" value="03-14-2016" />
  <Field name="intRecordID" value="4564xxx" />
  <Field name="intTableID" value="12" />
  <File filePath="635935515966691387_13cfdffe-98ef-40ad-acc2-c887be14714d_10061.pdf"
  mimeType="application/pdf" />
</Document>
</ImportGroup>

```

We grab this data using a VB.NET program we wrote to create matching XML and files and we change the name of the files to reflect the BATCHID (time stamp:YYMMDDHHMMSS), A DOC_ID which is 4 digit Alpha starting with AAAA and going to ZZZZ (492,804 unique Documents Converted from 18,954-511,758) and a FILE_ID which is a 3 digit number incrementing with every attached file.(000-999 inclusive) This makes the files related to each other by Batch ID, in an order by DOC_ID, and a FILE_ID) they look like this:

An PDF file Named:160323090155AIJO001.PDF

An XML File Named:160323090155AIJO000.dwcontrol

It looks like this:

```

<ControlStatements xmlns="http://dev.docuware.com/Jobs/Control"
(http://dev.docuware.com/Jobs/Control) xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"> (http://www.w3.org/2001/XMLSchema-instance>)
<Document>
<InsertFile path='160323090155AIJO001.pdf'/>
</Document>
<Page>
<Field dbName="DOMAIN" type="Text" value="BPV"/>

```

```

<Field dbName="INDEX" type="Text" value="CERTIFICATION"/>
<Field dbName="TYPENAME" type="Text" value="CERTIFICATE"/>
<Field dbName="PRAESES_ID" type="Text" value="635935515966691387_13CFDFFE-
98EF-40AD-ACC2-C887BE14714D_10061"/>
<Field dbName="CMCITRAX" type="Text" value="B:160323090155 D:160323090155AIJO
F:PRAESES_FTP"/>
<Field dbName="LOCATIONCITY" type="Text" value="SomeCity"/>
<Field dbName="LOCATIONNAME" type="Text" value="Some Company Name"/>
<Field dbName="LOCATIONADDRESS" type="Text" value="1122 Electric Ave"/>
<Field dbName="LOCATIONSTATE" type="Text" value="IL"/>
<Field dbName="LOCATIONZIP" type="Text" value="60201-4205"/>
<Field dbName="LICENSENUMBER" type="Text" value=""/>
<Field dbName="JURISDICTIONNUMBER" type="Text" value="B010XXXX"/>
<Field dbName="DOCDATE" type="Date" value="03-14-2016" culture="en-US" format="MM-dd-
yyyy"/>
<Field dbName="INTRECORDID" type="Text" value="4564xxx"/>
<Field dbName="INTTABLEID" type="Text" value="12"/>
</Page>
</ControlStatements>

```

If there were more than 1 file it would have looked this this instead:

```

<ControlStatements xmlns="http://dev.docuware.com/Jobs/Control"
(http://dev.docuware.com/Jobs/Control) xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"> (http://www.w3.org/2001/XMLSchema-instance>)
<Document>
<InsertFile path='160323090155AIJO001.pdf'/>
<InsertFile path='160323090155AIJO002.xml'/>
<InsertFile path='160323090155AIJO003.jpg'/>
<InsertFile path='160323090155AIJO004.tif'/>
<InsertFile path='160323090155AIJO005.pdf'/>
<InsertFile path='160323090155AIJO006.pdf'/>
</Document>
<Page>
<Field dbName="DOMAIN" type="Text" value="BPV"/>
<Field dbName="INDEX" type="Text" value="CERTIFICATION"/>

```

```

<Field dbName="TYPENAME" type="Text" value="CERTIFICATE"/>
<Field dbName="PRAESES_ID" type="Text" value="635935515966691387_13CFDFFE-
98EF-40AD-ACC2-C887BE14714D_10061"/>
<Field dbName="CMCITRAX" type="Text" value="B:160323090155 D:160323090155AIJO
F:PRAESES_FTP"/>
<Field dbName="LOCATIONCITY" type="Text" value="SomeCity"/>
<Field dbName="LOCATIONNAME" type="Text" value="Some Company Name"/>
<Field dbName="LOCATIONADDRESS" type="Text" value="1122 Electric Ave"/>
<Field dbName="LOCATIONSTATE" type="Text" value="IL"/>
<Field dbName="LOCATIONZIP" type="Text" value="60201-4205"/>
<Field dbName="LICENSENUMBER" type="Text" value=""/>
<Field dbName="JURISDICTIONNUMBER" type="Text" value="B010XXXX"/>
<Field dbName="DOCDATE" type="Date" value="03-14-2016" culture="en-US" format="MM-dd-
yyyy"/>
<Field dbName="INTRECORDID" type="Text" value="4564xxx"/>
<Field dbName="INTTABLEID" type="Text" value="12"/>
</Page>
</ControlStatements>

```

NOTE the InsertFile path is SINGLE QUOTES
 DOCUWARE HELP has them listed as DOUBLE but I have never had any luck with that.....
 I changed them to single quotes and it works....
 I KNOW THIS WORKS
 DO WHAT WORKS!

Your IMPORT Configuration (this is the web configuration not local or administrator) must be set
 to IMPORT PDF with METADATA files
 Docuware (.DWCONTROL)
 Metadata specifies the document

Do not forget to set the permissions correctly for the users or they will not see them at
 all. Docuware will CLIP these all together.....
 Docuware does have some help on this subject but there are pieces of it spread around in several
 different places on their help.
 And it is at time incorrect probably a format change by an editor The IMPORTANT things to note

are: Follow all XML RULES for illegal characters and filter them (items like ' and & etc [there are over 200 bad chars](https://en.wikipedia.org/wiki/List_of_XML_and_HTML_character_entity_references#Predefined_entities_in_XML)) (https://en.wikipedia.org/wiki/List_of_XML_and_HTML_character_entity_references#Predefined_entities_in_XML) of them must be handled they will not get picked-up or imported
Make CERTAIN you follow the rules for DATES and AMOUNTS
Make CERTAIN the file name is EXACTLY the same if not it will not work! Example:
Your dwcontrol shows <InsertFile path='1234ABC.PDF'/>
The ACTUAL file name is 1234ABC.pdf (PDF is lower case)
THIS WILL NOT IMPORT BE SPECIFIC Docuware will not tell you that it is not picking up the records or WHY it looks like it is LOCKED UP but most of the time it is BAD XML

The Index field data dbname is the SQL name and not the Display name look that up in Administrator Field Details. Do not get confused with the LABEL
I have a previous blog entry that can help [DWCONTROL FIELDS](http://www.commicrofilm.com/home/about/help/docuwareandthecontrolfile-dwcontrol-) (<http://www.commicrofilm.com/home/about/help/docuwareandthecontrolfile-dwcontrol->) outlines some of these things. Once you work out the bugs you can have the importer read the data into Docuware and speed may be an issue, so do not be impatient. It will take a moment to start importing and it may not be as fast as you think it should be but ignore the feedback and instead watch the file location as the files are removed and stored. Lastly the JOBS folder is where these are kept if they fail.

JOBS FOLDER IS located at : C:\Users\All Users\DocuWare\Jobs

Make a note of this as I have never found documentation on this but it is IMPORTANT.

In every version before 6.9 the log says pass or fail and nothing else. The XML in the Job location is of no value so the file names become very important to you. Here are LINKS to other help on the subject.....[DocuWare Control commands with DWControl Font](https://www.docuware.com/support_faq/index.php?action=artikel&cat=183&id=1431&artlang=en) (https://www.docuware.com/support_faq/index.php?action=artikel&cat=183&id=1431&artlang=en)
[Import with XML Metadata using .DWCONTROL XML Tags to Clip Documents](https://www.docuware.com/support_faq/index.php?action=artikel&cat=187&id=2500&artlang=en&highlight=dwcontrol) (https://www.docuware.com/support_faq/index.php?action=artikel&cat=187&id=2500&artlang=en&highlight=dwcontrol)
[Adding a Metadata File with Indexing](http://help.docuware.com/en/#b57696t60689n63710) (<http://help.docuware.com/en/#b57696t60689n63710>)
[XML and HTML character entity references \(Wikipedia\)](https://en.wikipedia.org) (<https://en.wikipedia.org>)

[/wiki/List of XML and HTML character entity references#Predefined entities in XML\)](#)
[DocuWare Control \(http://help.docuware.com/en/#b57696t59700n63693\)](http://help.docuware.com/en/#b57696t59700n63693)
[Defining Language and Region \(http://help.docuware.com/en/#b57696t59751n63695\)](http://help.docuware.com/en/#b57696t59751n63695)
[Define date format \(http://help.docuware.com/en/#b57696t59821n63696\)](http://help.docuware.com/en/#b57696t59821n63696)
[Sample file from Docuware \(https://www.docuware.com/support_faq/dwdl/dwdl.php?file=DWControl_DWImport_Sample.zip\)](https://www.docuware.com/support_faq/dwdl/dwdl.php?file=DWControl_DWImport_Sample.zip)

REMEMBER! USE SINGLE QUOTES when importing files.

Oben

Neuen Kommentar schreiben

Ihr Name

César Javier Fois Acosta

Betreff

Kommentar *

[Rich-Text deaktivieren](#)

Bilder

Neue Datei hinzufügen

Hochladen



No file selected.

Die Dateien müssen kleiner als 100 MB sein.

Zulässige Dateierweiterungen: png gif jpg jpeg.

Bilder müssen zwischen 220x220 und 2000x2000 Pixel haben.

Speichern >

Vorschau >