CHARMe WP6

(Implementation in archives)

Image12.png

**CHARMe content providing**

written by: f.kratzenstein

DWD, Offenbach, 2015-03-06

Table of Contents

[Global Parameters / Glossary 2](#_Toc413654845)

[CM SAF - CHARMe content providing (Components) 4](#_Toc413654846)

[Details 5](#_Toc413654847)

[CM SAF - CHARMe content providing (Activity) 6](#_Toc413654848)

[Summary 6](#_Toc413654849)

[Details 7](#_Toc413654850)

[Endnote2CitoCsv.py 8](#_Toc413654851)

[Summary 8](#_Toc413654852)

[Details 9](#_Toc413654853)

[CMSAF\_Annotator\_LitDB.py 10](#_Toc413654854)

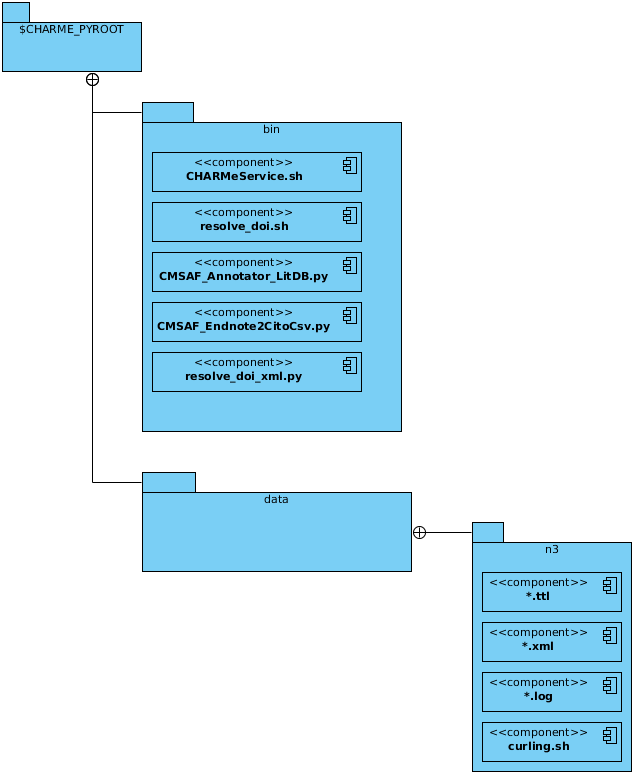
[Summary 10](#_Toc413654855)

[Details 11](#_Toc413654856)

# Global Parameters / Glossary

|  |  |  |
| --- | --- | --- |
| Name | Aliases | Documentation |
| CHARME\_NODE | CHARME\_SERVER, Triplestore |  |
| CHARME\_PYROOT |  | $CHARME\_ROOT/WP6/WP6x |
| CHARME\_ROOT |  | Quellverzeichnis des CHARMe Projektes: /media/nas/x18768/CHARMe | //ofnfa213/CHARMe |
| CHARME\_VOCAB\_SKOS | CHARME\_Vocab.xml | $CHARME\_PYROOT/skos/CHARMe\_vocab.xml |
| CHARMeAdmin |  |  |
| CITO\_CSV |  | a csv-file according to the CITO\_CSV\_FORMAT |
| CITO\_CSV\_FORMAT |  | CiTo csv format with at least the columns: CitingEntity, CitedEntity, CitationEvent and Motivation |
| CMSAF\_ANNOTATOR\_DOI | CMSAF\_Annotator\_DOI.py | The CMSAF\_Annotator\_DOI.py is a pyhton script which takes CMSAF\_CHARME\_DOI\_XML files as an input source.  From the CMSAF\_CHARME\_DOI\_XML file it extracts the CHARMe relevant information and transforms it into turtle files. The turtle files are posted to the CHARMe Node, where an insert operation is executed. |
| CMSAF\_CHARME\_DOI\_HTML |  | html-Snippet mit den CHARMe-Class Attributen zur Aktualisierung der CHARMe-enabled Webpage |
| CMSAF\_CHARME\_DOI\_XML |  | CM SAF WUI erweiterte DOI xml-Datei nochmals erweitert um die CHARMe Attribute |
| CMSAF\_CHARME\_DOI\_XSD | cmsaf\_charme\_doi.xsd | XML-Schema für die CHARMe Erweiterung der CM SAF WUI Doi xml-Dateien |
| CMSAF\_CHARME\_DOI\_XSLT | cmsaf\_doi2html.xsl | xslt-stylesheet for transforming the CMSAF\_CHARME\_DOI\_XML into CMSAF\_CHARME\_DOI\_HTML snippet |
| CMSAF\_DOI\_XML |  | DOI xml-Datei, wie sie im WUI benutzt wird |
| CMSAF\_DOIXML\_EDITOR | CMSAF\_DOIXML\_Editor.py | Python Skript um aus einem CMSAF\_DOI\_XML ein CMSAF\_CHARME\_DOI\_XML zu machen. Zusätzlich wird noch ein CMSAF\_CHARME\_DOI\_HTML erstellt. |
| CMSAF\_WUI\_DOI\_DIR |  | Quellverzeichnis für die CMSAF\_DOI\_XML Dateien: /cmsaf/cmsaf-ops2/CMSAF\_QM/png/doi |
| ENDNOTE\_LIB | Endnote Library | Endnote Library |
| ENDNOTE\_LIB\_XML | endnote\_lib.xml | xml-file exported from Endnote with the "Vancouver" Output-Style |
| WebEditor |  |  |

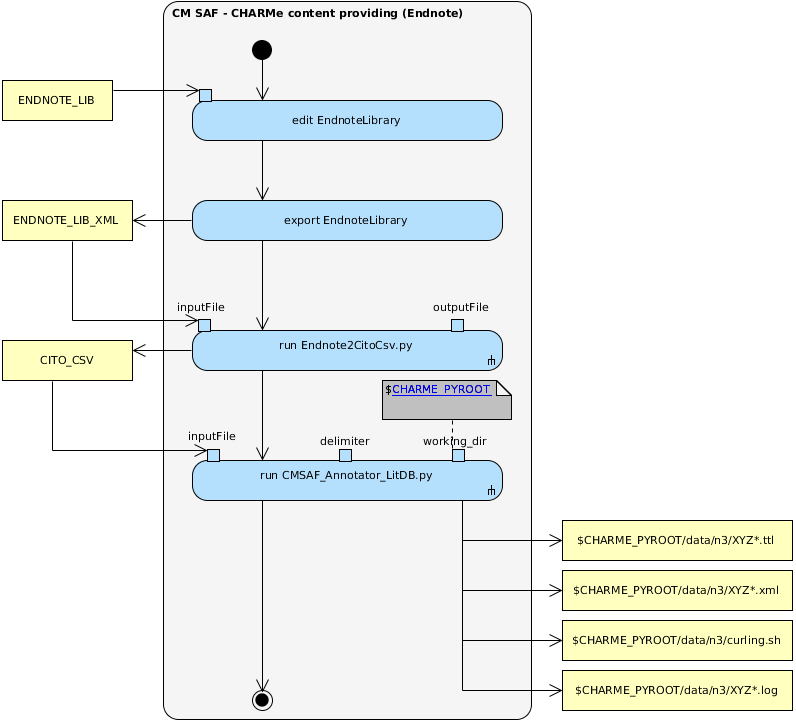
# CM SAF - CHARMe content providing (Components)



## Details

|  |  |
| --- | --- |
| Name | Documentation |
| Image1.png $CHARME\_PYROOT |  |
| Image1.png bin | Folder containing the executables |
| Image2.png CHARMeService.sh | script to call the CHARMe\_Node API |
| Image2.png resolve\_doi.sh | Script resolving the information of a doi |
| Image2.png CMSAF\_Annotator\_LitDB.py |  |
| Image2.png CMSAF\_Endnote2CitoCsv.py |  |
| Image2.png resolve\_doi\_xml.py | Python-wrapper of the resolve\_doi.sh |
| Image1.png data | Folder containing the data |
| Image1.png n3 | Folder containing the data, ready for tranfering to the CHARMe\_Node |
| Image2.png \*.ttl | Annotation as a turtle-file |
| Image2.png \*.xml | DOI XML-file |
| Image2.png \*.log | CHARMe\_Node operation log for an annotation turtle-file with the same name. |
| Image2.png curling.sh | Dynamically created script to transfer the \*.ttl-file to the CHARMe\_Node |

# CM SAF - CHARMe content providing (Activity)



## Summary

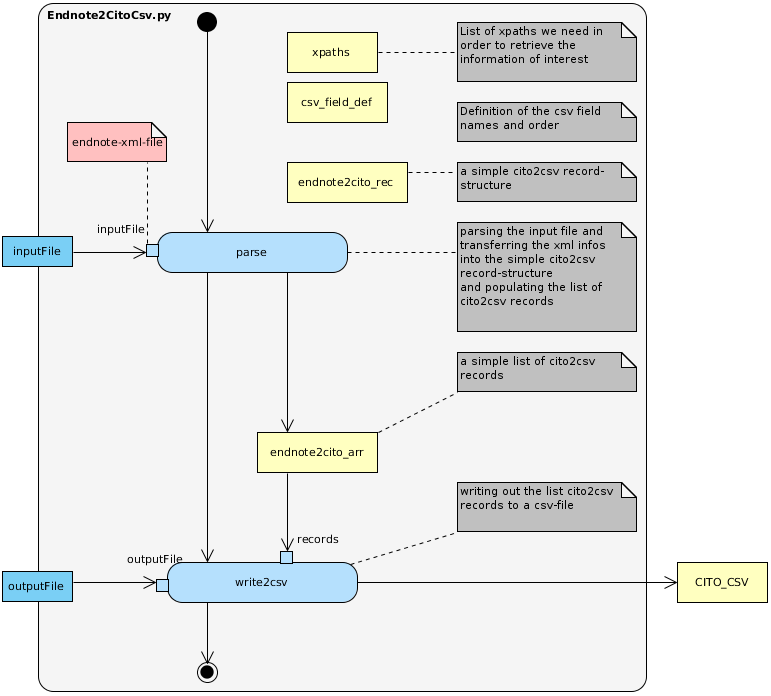
Die Aktivität "CM SAF - CHARMe content providing" beschreibt den Prozess, wie über eine CITO\_CSV Datei mehrere Annotations auf den CHARME\_NODE übertragen werden können.

Die Erstellung einer CITO\_CSV Datei erfolgt in der Regel manuell, für das Microsoft Programm Endnote kann ein Konverter, der Endnote2CitoCsv.py, genutzt werden.

## Details

|  |  |
| --- | --- |
| Name | Documentation |
| **Image5.png edit EndnoteLibrary** | add the dataset DOI as Reviewed Item to the References of the Endnote Library item of interest. |
| **Image5.png export EndnoteLibrary** | export the records of interest from the EndnoteLibrary as xml with output style “Vancouver”. |
| **run**  **Endnote2CitoCsv.py** | converts the ENDNOTE\_LIB\_XML file into a CITO\_CSV file. |
| **Image5.png run CMSAF\_Annotator\_LitDB.py** | transfers the records from a CITO\_CSV to the CHARME\_NODE and performs the DML operation. |

# Endnote2CitoCsv.py



## Summary

Das pyhton skript Endnote2CitoCsv.py übernimmt die Aufgabe eine aus Endnote exportierte XML-Datei (Vancouver Output-Style) in eine Datei mit dem CITO\_CSV\_FORMAT zu konvertieren.

Pre-Conditions:

Die xml-Datei wurde aus Endnote mit der Option für den Output-Style = "Vancouver" exportiert.

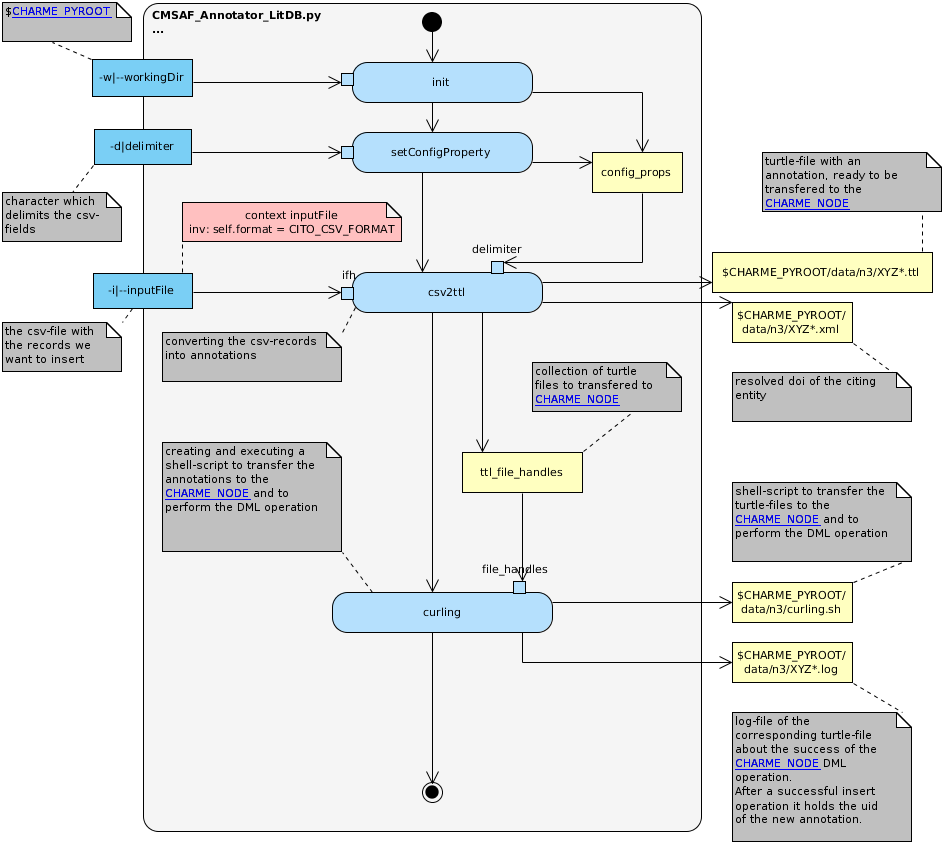
Post-Conditions:

In dem Verzeichnis $CHARME\_ROOT/data ist eine Datei mit dem Namen wie in outFile spezifiziert vorhanden. Diese Datei ist aufgebaut, wie es das CITO\_CSV\_FORMAT vorgibt.

## Details

|  |  |
| --- | --- |
| Name | Documentation |
| **Image5.png parse** |  |
| **Image5.png write2csv** |  |

# CMSAF\_Annotator\_LitDB.py



## Summary

The CMSAF\_Annotator\_LitDB.py is a pyhton script which takes a csv-file as an input source.

The csv-file must be formatted according to the rules defined at CITO\_CSV\_FORMAT.

Each record of the csv-file is transformed into an annotation and is stored as an individual turtle file.

During the transformation the DOI of the CitingEntity is resolved and its title is used as content of the annotation body.

The turtle files containing the annotations are posted to the CHARME\_NODE, where an insert operation is executed.The response

from the CHARME\_NODE is stored in a log-file. Each turtle file should have its corresponding log-file. If the insert operation

has been successful the log-file contains the assigned uid of the annotation otherwise it contains the error message of the CHARME\_NODE.

## Details

|  |  |
| --- | --- |
| Name | Documentation |
| **Image5.png init** | initializing the base structure and the config properties |
| **Image5.png setConfigProperty** | setting a config property; using the user defined character as the delimiter on the inputFile |
| **Image5.png csv2ttl** | converting the records of the csv-file into a turtle-file containing the annotation |
| **Image5.png curling** | creating and executing a shell-script to transfer the annotations to the CHARME\_NODE |

Step by step instructions:

* Change to /media/x18768/CHARMe/WP6-cmsaf/cmsaf-annotator/bin
* Load python: module load python
* Export python path:
  + module use "/cmsaf/nfshome/routcm/Modules"
  + module load python/2.7.9

currently:

* + export PATH=/cmsaf/cmsaf-ops3/sfinkens/usr/bin:$PATH
  + export LD\_LIBRARY\_PATH=$LD\_LIBRARY\_PATH:/cmsaf/cmsaf-ops3/sfinkens/usr/lib:/cmsaf/cmsaf-ops3/sfinkens/usr/lib/python2.7/site-packages
  + export PYTHONPATH=/cmsaf/cmsaf-ops3/sfinkens/usr/lib/python2.7/site-packages
* using endnote: create csv-file from endnote xml file
  + python CMSAF\_Endnote2CitoCsv.py -i ../templates/testEndNote.xml -o ../data/test..csv
* export to CHARME DB:
  + python CMSAF\_Annotator\_LitDB.py -w ..--i test.csv -d ;