XML import manual for the European Pollutant Release and Transfer Register

Documentation

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **JOB NUMBER:** | | | **DOCUMENT REF:** | | | |
| **Version** | **Purpose of issue** | **Originated** | **Checked** | **Reviewed** | **Authorised** | **Date** |
| 1.0 | Initial Draft | Martin Holleufer-Sørensen |  |  |  | 25.05 2009 |
| 1.1 | Access import modified | Wolfgang Starzer |  |  |  | 03.05.2011 |
| 1.2 | Kml process implemented | Morten Hjelmsmark |  |  |  | 03.05.2011 |
| 1.3 | Detailed description of batch file and XML-upload | Wolfgang Starzer |  |  |  | 28.03.2012 |
| 1.4 | New XML import procedure | Helle Vollmer |  |  |  | 19.10.2010 |
| **ATKINS DANMARK** | | | | | | |

Table of Contents

[1 Summary 1](#_Toc266978299)

[2 Setup of the xml import environment 2](#_Toc266978300)

[3 Prepare the import of xml files 3](#_Toc266978301)

[4 Import xml file to eprtrxml database and validate data 5](#_Toc266978302)

[5 Copy data to eprtrmaster database 7](#_Toc266978303)

[6 Postprocessing imported data 8](#_Toc266978304)

# Summary

This document explains all steps needed to import xml files uploaded by the member states to the E-PRTR system. The xml files containing E-PRTR data are uploaded to the Central Data Register (CDR) at EEA by the member states. The data will first be imported and validated in a temporary database (EPRTRxml) and then the data will be moved to the E-PRTR master database making the data ready for review at the Review site. Atkins will handle the import and validation and will supply the member states with a feedback document in the member states envelope at CDR.

### Prerequisites:

* SQL Server 2008 R2 Native Client
* Microsoft .NET Framework 3.5 SP1
* XML\_Import program

### Parameters:

Global parameters:

|  |  |  |
| --- | --- | --- |
| **Name** | **Description** | **Value** |
| {SVN} | Location of import scripts | <https://svn.eionet.europa.eu/repositories/EPRTR/trunk/DataImport> |
| {EIONET} | Location of recently released XML files | <http://cdr.eionet.europa.eu/recent_eprtr> |

Currently active Environments:

|  |  |  |
| --- | --- | --- |
| **Name** | **Build (Atkins)** | **Test/Prod (EEA)** |
| {DB-SERVER} | Sdkcga6332 |  |
| {DB-SERVER-PHYSICAL} | Sdkcga6332 |  |
| {DB-USER} | gis |  |
|  |  |  |

Environment specific parameters for resubmission March 2012:

|  |  |  |
| --- | --- | --- |
| **Name** | **Sdkcga6332** | **EEA** |
| {Sqlscript\_dir} | D:\EPRTRimport\EIONET\SQLScripts |  |
| {Mapforce\_dir} | D:\EPRTRimport\XML-import\MAPFORCE |  |
| {Data\_dir} | D:\EPRTRimport\XML-import\XML\_DOWNLOAD\submission\_2012\_03\_01 |  |
| doValidate | true | False |

# Setup of the xml import environment

Download all files in directory MAPFORCE and SQLSCRIPT from SVN and install them on the import server at the location from where you want to use them during the import. During the import this location is found by specifying the respective parameters in the import batch file Validate\_and\_Import\_XML\_Files..

# Prepare the import of xml files

All member states are asked to upload their data to the CDR. At the CDR the member states create envelopes for each one of their obligations. For some obligations (e.g. the Pollutant Release and Transfer Report) it is possible to create several envelopes/upload data several times. In order to get hold of the right version of data Atkins need to log into CDR and manually seek out the approved envelopes with regards to the Pollutant Release and Transfer Report obligation. The authentication of data is done by EEA. This web site address is specified in parameter {EIONET}.

When the right envelopes have been identified please make sure that all the validations have been passed. The XML schema validation and the E-PRTR Compliance validation need to be passed without errors of any kind.

In case of errors the data will not be imported and the member state need to resubmit a new version of the E-PRTR report. The Additional validation is just additional information to the member states about special data in their report that they need to be aware of. This additional information is not preventing the data from being imported. When the validations have been checked to see that they were passed, please follow the below steps to prepare the import:

1. Download the xml file containing the PRTR-report from the CDR system to { Data\_dir }. Prefix the name of the file with the member state initials and postfix it with the respective reporting year. E.g. the xml file from Austria for a resubmission for 2007 will be called "AT\_EPRTR-2007".
2. Identify the following information from the CDR system :
   * CdrURL for the envelope (e.g. http://cdrtest.eionet.europa.eu/az/eu/colsglhra/envshfmmq)
   * CdrUploaded (Date and time the file was uploaded e.g 2009-05-15 12:05:00)
   * CdrReleased (Date and time the envelope was released e.g 2009-05-18 13:56:11)
   * Resubmit reason (in some cases the member states attach documents (.doc/.txt) describing why the report was resubmitted).
3. Prepare a csv file with the above information for all datafiles to be imported.   
   The file must be formatted as defined by the template found in : {svn}\DataImport\ImportScripts\00\_config\_template.csv.  
   Important: The date must be given in the format shown in the template.  
   The csv file must be named 00\_config.csv and placed in { Data\_dir}.
4. Update paths and server names in Validate\_and\_Import\_XML\_Files. Also consider if validation is to be included (parameter "doValidate", see section 4). The script can run in singlemode by setting parameter "singleMode" = true. In singleMode the script will require user-interaction before continuing with the next file.

Running script “Validate\_and\_Import\_XML\_Files.bat” imports all XML files and generates a log file for each batch file respectively. Re-running the script will only import XML files where no log file already exists. Also a file can be excluded by prefixing the corresponding row in the csv file by #.

# Import xml file to eprtrxml database, validate data AND Copy data to eprtrmaster database

Besides the validation executed in CDR it is also important that the downloaded xml-files are validated manually with regards to which reporting year the xml file concerns and which xsd schema the member state have been using when creating the xml file.

The reporting year enclosed by tag <rsm:ReportingYear> in the PollutantReleaseAndTransferReport element of the xml file needs to correspond to the reporting year which is handled.

The xsd schema version is at the moment 2.0 so the header of the xml file needs to read:

<rsm:PollutantReleaseAndTransferReport xsi:schemaLocation="urn:eu:com:env:prtr:data:standard:2 http://www.eionet.europa.eu/schemas/eprtr/**PollutantReleaseAndTransferReport\_2p0.xsd**" xmlns:wgs84="urn:convertKKJ" xmlns:rsm="urn:eu:com:env:prtr:data:standard:2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

If this is not true then the report is not accepted and the member state will be notified about the error in the feedback report. After the above manual validation has been carried out successfully the next step is that a data validation of all reported data will be executed in the temporary EPRTRxml database.

So far the validation routine validates data for the following issues:

1. Validating previous nationalids and previous reporting year given in xml file
2. Validating if a facility in a previous report is referenced more than once in the new report
3. Validating if nationalids and previous nationalids are the same for new facilities
4. Validating if previous nationalid return multiple records
5. Validating if reported coordinates are within country polygons
6. Validating if reported coordinates are set to zero or not reported

Based on a successful run of the import script in the test environment and after examination of each validation log file respectively it needs to be decided whether or not a countries data is going to be accepted for the import into the EPRTRMASTER database in the production environment.

To run validation and import please do as described previously and place all country specific batch scripts together with script “Validate\_and\_Import\_XML\_Files.bat” (currently located in SVN at [https://svn.eionet.europa.eu/repositories/EPRries TR/trunk/DataImport](https://svn.eionet.europa.eu/repositories/EPRries%20TR/trunk/DataImport)) at the very location on your server from where you want to run the import. Before running the script please have a look at the parameters defined within the script and adjust them. This parameters are:

|  |  |
| --- | --- |
| sqlscript\_dir: | Defines the location of the SQL scripts that are used by the import. |
| mapforce\_dir: | Defines the location of the SQL MAPFORCE routines. |
| SET data\_dir: | Defines the location of the XML files. |
| SET doValidate: | Allows to jump over the validation step. Possible values: true/false . Having validated the data in the test environment it is typically not necessary to run the validation also in the production environment. Thus setting this parameter “false” saves time |
| SET singleMode | If true, the user must interact with the script to continue to the import of the next file. Normally this is false |

.

To make sure that the result of the import at Atkins test environment and EEA’s production environment will be identical, it is always a good idea to start the validation and import procedure by restoring the latest backup of the EPRTRmaster from EEA at Atkins.

Running script “Validate\_and\_Import\_XML\_Files.bat” executes all batch files and generates a log file for each batch file respectively. The logfile (e.g. slovenia\_2007\_imp.bat.log) contains information about the import to the temporary xml database, the validation of data (if parameter “doValidate” is set to true) and the result of inserting all the new data into Database EPRTRmaster. This information from the import and validation log needs to be given as feedback to the member state in the CDR system.

Coping data from database EPRTRxml to EPRTRmaster the following attributes are set automatically during the import-procedure based on the information entered in the batch files:

* POLLUTANTRELEASEANDTRANSFERREPORT.CDRURL 🡪 URL of envelope in CDR
* POLLUTANTRELEASEANDTRANSFERREPORT.CDRUPLOADED 🡪 Date and time when the xml file was uploaded
* POLLUTANTRELEASEANDTRANSFERREPORT.CDRRELEASED 🡪 Date and time when the envelope in CDR was released
* POLLUTANTRELEASEANDTRANSFERREPORT.RESUBMITREASON 🡪 Voluntary description from the envelope in CDR
* FACILITYREPORT.LOV\_RIVERBASINDISTRICTID 🡪 RiverBasinDistrictID of RiverBasinDistrict based on spatial analysis
* FACILITYREPORT.LOV\_NUTSRegionID 🡪 NUTSRegionID of NUTSRegion based on spatial analysis
* FACILITYREPORT.LOV\_STATUS 🡪 Valid, Outside or Missing with regards to the coordinates of the facility

Besides the above attributes a couple of default values that are not required in the xml file are set during xml import:

* POLLUTANTRELEASE.LOV\_TOTALQUANTITYUNITID is set to 2 which refers to KGM
* POLLUTANTRELEASE.LOV\_ACCIDENTALQUANTITYUNITID is set to 2 which refers to KGM POLLUTANTTRANSFER.LOV\_QUANTITYUNITID is set to 2 which refers to KGM
* WASTETRANSFER.LOV\_QUANTITYUNITID is set to 3 which refers to TNE

# Postprocessing imported data

## Copy data to Review Site

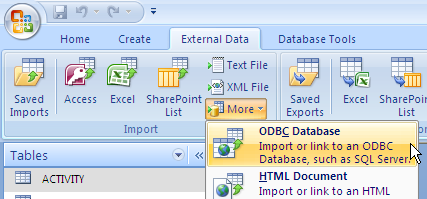
In order to copy the imported data to the Review Site download the file copyEPRTRmasterViews2EPRTRwebTables.sql from {SVN}/SQLScripts. Please replace every instance of the word EPRTRweb with EPRTRreview in the script. Also make sure that the "set Published date" part of the script is commented out and the "set ForReview date" part of the script will be executed. When everything is ready execute the script in a sql-editor.

## Copy data to Public Site

In order to copy the imported data to the Public Site download the file copyEPRTRmasterViews2EPRTRwebTables.sql from {SVN}/SQLScripts. Please replace every instance of the word EPRTRweb with EPRTRpublic in the script. Also make sure that the "set ForReview date" part of the script is commented out and the "set Published date" part of the script will be executed. When everything is ready execute the script in a sql-editor.

## Export data to Full and Public Access databases

In order to export E-PRTR data to Access databases used by ETC and EEA dataservice download the two Access databases found in {SVN}/Export2Access and follow the procedure given below:

1. Open up the Access database
2. Open up a ODBC connection to the EPRTRmaster database
3. Select “Import the source data into a new table.....” and click OK
4. Click New (Data Source), choose SQL Server Native Client 10.0 or SQL Server and click Next. The following steps 5-8 only has to be done once. Next time just select the Data Source already set up.
5. Enter a name for the new Data Source (e.g. tetrasql) and click Next and click Finish
6. Enter the appropriate Server (e.g. TETRASQL) and click Next
7. Choose SQL Server authentication and type in login ID (e.g. gis) and password (e.g XXXXXX) and click Next
8. Change the default database to EPRTRmaster and click Next, Finish, Test Data Source, and OK
9. Now select the Data Source that was just created and click OK. Type in password and click OK
10. If the Access database opened is the Full\_Database.mdb select all views prefixed with FULL (e.g. dbo.FULL\_ACTIVITY) (12 views in all) and click OK. If the Access database opened is the Public\_Database.mdb select all views prefixed with PUBLISH (e.g. dbo.PUBLISH\_ACTIVITY).
11. Run Macro FINISH\_DATABASE (calls macro RENAME\_TABLES, Add\_Primary\_Keys and Add\_Foreign\_Keys)
12. Change manually name of table dbo\_FULL\_UPLOADEDREPORTS to UPLOADEDREPORTS.
13. Save the database and exit Access
14. Change the name of the database to Full\_Database\_YYY-MM-DD, respective Public\_Database\_YYYY-MM-DD.
15. Zip the databases and put them on EEA ftp server (<ftp://ftp.eea.europa.eu/Ose3%20GIS/EPRTR/>) username: eeaftp

An outdated version of the documentation of the two databases can be found in {SVN}/documents/TechnicalSpecification/, a newer version is published by EEA on: <http://www.eea.europa.eu/data-and-maps/data/member-states-reporting-art-7-under-the-european-pollutant-release-and-transfer-register-e-prtr-regulation-3>

# Creating KMZ file

This is the procedure to follow when a kml file has to be recreated.

The scripts etc. can be found on *svn.eionet.europa.eu/repositories/EPRTR/trunk/Eprtr\_kml\_solution*, or in the *C:\GISLibrary\Python\EPRTR\_KML* folder on *EEAPC8662.*

If new data is issued for the same year as the previous run, it should be enough just to run the script again. The script can be run by double clicking the **Create\_EPRTR\_KML.bat** file located in the Script folder (on EEAPC8662 it located in C:\GISLibrary\Python\EPRTR\_KML\Scripts).

If the new data doesn’t show in GoogleEart or if data is covering a new year, you need to update the base layer file. Locate the Layer file EPRTR\_style\_basic\_reportid.lyr (on EEAPC8662 it located in C:\GISLibrary \Python\EPRTR\_KML\Data) and open it in ArcMap.

* Update the data source
* Make sure the joins still works
* In the definition query tab, update query to the correct reportingyear
* In the fields tab reset the list of implemeted fields – remember to **Apply**!

|  |  |
| --- | --- |
| **Layer name** | **Alias** |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.FacilityReportID | FacilityReportID |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.FacilityName | FacilityName |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.ReportingYear | ReportingYear |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.Address | Address |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.City | City |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.PostalCode | PostalCode |
| EPRTRPUBLIC.SDE.FACILITYSEARCH\_MAINACTIVITY\_GEOGRAPHICALCOORDINATE.IASectorCode | IASectorCode |
| EPRTRPublic.dbo.LOV\_COUNTRY.Name | Country |
| EPRTRPublic.dbo.LOV\_ANNEXIACTIVITY.Code | Sector code |
| EPRTRPublic.dbo.LOV\_ANNEXIACTIVITY.Name | Sector |

* When done, save the layer by using ’Save As Layer file..’ and save the lyr file in the data folder

After the layer file is updated, make sure the ..\scripts\ EPER\_fill\_mxd.py points at it.

#Path to the base lyr file

inpPth = solPth + "\Data\EPRTR\_style\_basic\_reportid.lyr"

Then run the **Create\_EPRTR\_KML.bat** file located in the Script folder (on EEAPC8662 it located in C:\GISLibrary\Python\EPRTR\_KML\Scripts).

For more information read the document **EPRTR kml solution documentation.docx**, which can be found in the svn repository *svn.eionet.europa.eu/repositories/EPRTR/trunk/Eprtr\_kml\_solution*