



Collaborative Climate Community Data and Processing Grid (C3Grid)

Creation of C3Grid ISO Metadata from CERA database

Work package:	WP3 – ISO Metadata
Autor(s):	Hans Ramthun
[Editor(s):]	H.Ramthun
Version:	1.0 (Feb.2008)
Publishing date:	February 2008
WP-Coordination:	
[Partners:]	
Contact:	H.Ramthun
Email:	hans.ramthun@zmaw.de

*project within the D-Grid community
promoted by BMBF*



C3 Grid Metadata

1. From CERA metadata to ISO metadata

This chapter describes how to create metadata XML (ISO standard: 19115 / 19139) from the CERA database.

To extract metadata from the CERA database a XSQL style sheet is used in conjunction with a XSL style sheet to retrieve data from the CERA metadata database and format them into a appropriate ISO 19115/19139 format. The database connection is made through a JDBC driver. The selection from the database is performed with SQL statements.

Actual all used XSQL and XSL templates are available on Anticyclone:
<http://anticyclone.dkrz.de:8080/oai/provider/xsql/>

These XSQL/XSL scripts are used inside a Python program for creating or updating ISO metadata.

This Python program (make_all_metadata.py, see appendix a) for actual listing) is controlled with an external property file (c3grid.metadata.xml.properties, see Appendix b) for actual listing). Appendix c) to e) show the listings of the processing scripts.

Help output (of 'make_all_metadata.py'):

Create ISO XML file, Version 1.3, 2007/2008

usage: [python] make_all_metadata.py -file 'properties file'



2. OAI Server with the jOAI software

Setup a OAI server means to download and install the dlese / jOAI software.
Download from here: http://www.dlese.org/dds/services/joai_software.jsp
You need Java 1.5 and a tomcat5 container.

Follow the install instruction:

- unzip to a directory of your choice
- copy the oai.war file to the webapps directory of the tomcat5 container
- restart tomcat5 container, that will deploy the OAI software
- installation must appear at: <http://localhost:8080/oai>
- then follow the instruction of the OAI documentation

2.1 Repository administration for the repository presenting data to the C3 Grid project.

Repository Information and Administration

Repository information

Update repository name, administrator's email, repository description or namespace identifier.

Repository name (required):	WDCC
Repository administrator's e-mail (required):	hans.ramthun@zmaw.de
Repository description (optional):	C3Grid OAI server at HH hosting ISO metadata for DKRZ/MPI-M/IFM-GEOMAR/M&D/GKSS
Repository namespace identifier (optional):	c3grid.dkrz



2.2 Metadata files presented by the OAI server

This is the actual list of metadata sets presented at the C3 Grid OAI server at the Hamburg cluster site.

Metadata Files Configuration

Metadata Directory	<u>Format</u>	<u>Num Files</u>	<u>Num Ready</u>	<u>Num Deleted</u>	<u>Indexing Errors</u>
ISO Meatadata FF DKRZ /opt/c3grid/oai/xml_iso_ff	iso	2	2		
ISO Metadata GKSS /opt/c3grid/oai/xml_iso_gkss	iso	11	11	9	
ISO Meatadat WDCC /opt/c3grid/oai/xml_iso_wdcc	iso	53	53	31	
	Totals:	66	66	40	



2.3 Metadata set configuration

To enable harvesting subsets of data sets are defined.

Sets Configuration

A set is a subgroup of the metadata files in a repository. A set is defined using metadata formats, directories, keywords and/or search queries. For more information about how sets are used in the OAI protocol.

Set name	SetSpec	Records Ready	Records Deleted
ISO FF	iso-ff	2	
ISO GKSS	iso-gkss	11	9
ISO WDCC	iso-wdcc	53	31



Appendix

a) Listing Python program 'make_all_metadata.py'

Listing of Python Programs to create ISO XML metadata from CERA data using the XSQL/XSL scripts on anticyclone:

http://anticyclone.dkrz.de:8080/xsql/cera_map_iso.xsql?id='CERA_ID'

b) Listing of Python program c3webservice.py

An additional Python service program: c3webservice.py from the base webservice tools for interface D

c) Example property file for process control

Example property file for the control of metadata access from the CERA metadata.

d) XSL(T) Stylesheet for output control

XSL Stylesheet for output control to meet the C3Grid ISO specifications.

e) XSQL Oracle database access script

XSQL script to access CERA database tables to retrieve metadata.