## **Table of Contents**

What is ESGF Tools UI?	2
Getting started	2
Pre-requisites	2
Installation	2
Search of ESGF Climate Data	2
Selecting a index node	3
Selecting search parameters	3
Tree of parameters	4
Drop-down list of parameters	4
Configured parameters	5
Using free text search	5
Save a search	6
Select a saved search	6
Reset search panel	6
Metadata Harvesting of ESGF Datasets of a search	6
What is a ESGF Dataset?	7
Search Harvesting Panel	7
Exploring search harvesting	8
Harvesting view	9
Open metadata	10
Download Dataset - File selection window	10
Export to Metalink	11
Download a Search	11
Download of ESGF Datasets	13
Context menu of downloads	14
Services of context menu of downloads	14
ESGF Login : Obtaining credentials with ESGF Account	15
Join a group with the needed permissions for a unauthorized file	16

## What is ESGF Tools UI?

A tool that is integrated in <u>?ToolsUI NetCDF Java</u> for including a desktop client for <u>?ESGF services</u>, with features like search of data along multiple nodes, download manager and metadata aggregation for allowing a full exploration through dataset services.

The added functionality to ToolsUI NetCDF is located in the ESGF tab. This tab contains 3 sub-tabs, each one with an specific functionality.

If you find any issue, problem or want to make a comment, go to Issues

NetCDF (4.3) Tools								
<u>System M</u> odes <u>D</u> ebug <u>H</u> elp								
Viewer Writer NCDump Iosp CoordSys FeatureTypes THREDDS Fmrc GeoTiff Units NcML URLdump ESGF								
Login Remaining time of credentials: 70, 78 hours (Expire date: Sat May 31 10:19:37 CEST 2014)								
Saved searches & Harvesting Search Downloads								

## **Getting started**

#### **Pre-requisites**

JDK from Sun / Oracle (1.5 - newer) or OpenJDK 6

OpenJDK 7 is not supported

## Installation

- 1. Download the jar in <u>ESGFToolsUI-v0.6.1.jar</u> Other versions..
- 2. Go to download target directory:
  - In Windows:
    - Open ESGFToolsUI.jar
  - In Linux
  - In shell:

```
java -jar ESGFToolsUI.jar
```

## Search of ESGF Climate Data

A search in ESGF returns the records that match the search constraints after querying to an index node. The ESGF search service is always served by a index node and this node is able to do local and distributed searches. In this tool, the distributed capabilities of the ESGF are used so any selected index node will be used to query all nodes in the ESGF system.

The search panel of ESGF climate data is in the "Search" sub-tab (be patient, it takes a while for the panel to load). The search constraints are defined by parameter configuration in the panel or by introduction of a free-text search query. The result is the number of records finding in the federation that satisfy the constraints of search (in the bottom right area of the panel). A record is the physical replica of a climate dataset that is stored on a data node.

NetCDF (4.3) Tools			- 🗉 🗙
<u>S</u> ystem <u>M</u> odes <u>D</u> ebug <u>H</u> elp			
Viewer Writer NCDump losp CoordSys FeatureTypes THREDDS Fmrc GeoTiff Units	NcML	URLdump ESGF	
Login Remaining time of credentials: 69,74 hours (Expire date: Sat May 31 10:19:37 CEST 2014)			
Saved searches & Harvesting Search Downloads			
esgf-index1.ceda.ac.uk 💌 Free text:	Edit	IPSL-CM5A-LR(Zosga)	-
← Ensemble(6)       ← ↓         ← □ rlilp1(6)       ←         ← □ r2ilp1(6)       ←         ← □ r3ilp1(6)       ←         ← □ r4ilp1(6)       ←         ← □ r5ilp1(4)       ←         ← □ r6ilp1(4)       ←         ← ○ CMIP table(1)       ←         ← →       →         ▲ Add Parameters       ▼			•
Configured parameters: Model : [IPSL-CM5A-LR] Variable : [zosga] Experiment family : [Historica]] Experiment : [historica]] Project : [CMIP5]		Number of record	s: 32
Remove Remove all			

## Selecting a index node



different processing times for the response. For this reason, the user is allowed to select the index node where he/she prefers to send the search petition.

The index node can be configured in the top left drop-down list. After selecting a new index node, the search panel is updated.

## Selecting search parameters

The search parameters can be configured on the panel in two ways:

• By values selection in tree of parameters double click

By parameter selection in drop-down list

	_		
- Model(33)			-
➢ Project(1)	_	Experiment Experiment family	
<ul> <li>Time frequency(1)</li> </ul>		Lower/upper limit of the last timestamp	
► CMIP table(2)		id 3	=
<ul> <li>Product(2)</li> </ul>		index node Instance id	
<ul> <li>Institute(20)</li> </ul>		Institute	
- BCC(23)		Master id	-
- 🖬 BNU(4)			
- CCCMA(12)			
- CMCC(5)			
	-		
Add Parameters			

#### Tree of parameters

The tree of parameters shows the parameters with bounded values which are defined by the federation. In the first level is the name of the parameter and the number of its bounded values. In the second level are the parameter values, each one with the count of records that satisfy the previous configured constraints (restrictions that are listed in the "configured parameters" section) + the new constraint (parameter-value).

For each parameter can be selected more than one value doing double click in the checkbox. Finally, for adding the selected values on the tree in the search constraints you must click on "Add parameters" button. All values selected in a parameter are linked by logic OR and the selected parameters are linked by logic AND. Generally, the parameter configuration is as follows, where "P" is a parameter and "V" a value.

# $\begin{bmatrix} (P_1 - V_1) \lor (P_1 - V_2) \lor \dots \lor (P_1 - V_n) \end{bmatrix} \land \begin{bmatrix} (P_2 - V_2) \lor (P_2 - V_2) \lor \dots \lor (P_2 - V_n) \end{bmatrix} \land \begin{bmatrix} (P_n - V_1) \lor (P_n - V_2) \lor \dots \lor (P_n - V_n) \end{bmatrix}$

For example, if are selected:
 \* Institute
 \* BNU ?
 \* CCCMA ?
 \* Model
 \* BNU-ESM ?
Then the result will be the number of records that belong to BNU-ESM model and made by BNU or CCCMA institutes i.e.
 ((Institute,BNU)V(Institute,CCCMA))??((Model,BNU-ESM))

After adding new parameters the search panel is updated (?Add parameters? button). Now in the tree, all parameters that already have values configured only shows this values. This happens because the allowed values ??for a parameter are always those that do not result in an empty set if they are selected. To configure a parameter previously configurated you can uncheck the previously configured values and then click on "Add parameters" or you can remove them in "Configured parameters" section.

#### **Drop-down list of parameters**

The parameter selection in drop-down list (top right) allows values configuration of some parameters. In this list are contained the parameters without bounded values and that for this reason can't be configured in the tree of parameters. Also contains the parameters with bounded values. After selecting a list parameter, below of the list will be displayed a specific configuration panel for the selected parameter.

In the figure below you can see the panel that is displayed when ?temporal range of data coverage? is selected. Finally, to be added this configuration click in the bottom button "Add parameter" ("Add temporal range parameter" in this case).

ect temp	t temporal range of data coverage (Start-End)																
st	ta	rt –							en	d							
e	ne	его		-	*		19	20 ÷	en	его		-	*		20	00 ÷	
		lun	mar	mié	jue	vie	sáb	dom		lun	mar	mié	jue	vie	sáb	dom	
01	1				1	2	3	4	52						1	2	
02	2	5	6	7	8	9	10	11	01	3	4	5	6	7	8	9	
03	з [	12	13	14	15	16	17	18	02	10	11	12	13	14	15	16	
04	4	19	20	21	22	23	24	25	03	17	18	19	20	21	22	23	
05	5	26	27	28	29	30	31		04	24	25	26	27	28	29	30	
05 31																	
	_				A	dd te	empo	ral ra	ange	e pa	rame	ter					

Temporal range parameter selection has bugs in ESGF search service (still not fixed), and for that, maybe the files could not be properly selected. In this case, the files must be selected manually

#### **Configured parameters**

All configured parameters are displayed in a specific panel in bottom left. In this panel are listed the configured parameters with its values assigned.

In the figure below can see that the current search is configured for searching climate data that belong to CMIP5 project, that were made by BNU or CCCMA institute, that have a time frequency of six hours, belongs to "historical" experiment and that have data between 15-01-1920 and 20-01-2000, and finally contains at least one of this variables: "hus", "ps", "psl", "ta".

Configured parameters:	
Project : [CMIP5]	
end:2000-01-20T11:40:49Z	
start:1920-01-15T11:40:49Z	
Experiment : [historical]	
Variable : [hus, ps, psl, ta]	
Institute : [BNU, CCCMA]	
Time frequency : [6hr]	
	Remove all

This panel also allow to select a parameter and delete its configuration doing click on "Remove" and delete all parameters configuration doing click on "Remove all"

#### Using free text search

The free-text search o full text search allows to do a search rich in syntax, by arbitrary words, logic operations and wild-cards. In this case, the records searched will be records that contains metadata that are related with what is specified in the free text query.

In top center in the search panel can see the text box where can be introduced the query of free text search. The "Edit" button enables input by keyboard in the text box. The "Save" button adds the free text query parameter in the search. The new configured parameter will be displayed in the "configured parameters" section with the name "query".

ſ	Search	Search Harvesti	ng Downlo	ads Logir	1	
	esgf-index	1.ceda.ac.uk	▼ Free text:	R (id:*IPSL* AN	ID (model:IPSL-CM5A-LR OR model:IPSL-CM5B-LR))	Save

The free-text query may have:

- Logic operations (AND & OR) between words
- Wildcard \* for the words
- Parentheses ( ) as separators
- metadata\_name:word to specify the metadata name where must match the given word

```
For example, the query:
    institute:CCCMA OR (id:*IPSL* AND (model:IPSL-CM5A-LR OR model:IPSL-CM5B-LR))
Will search all records in ESGF whose institute is CCCMA OR that satisfy the following
points:
    l. Have an id that contains the word IPSL (id:*IPSL*)
    2. Belong to IPSL-CM5A-LR model OR IPSL-CM5B-LR model
```

#### Save a search

A configured search can be saved for later for to be able do a harvesting of metadata and services of datasets that satisfy the search constraints. The save section is in right bottom of search panel. You can overwrite the selected search with "Save search" button or save it with a new name clicking on "Save Search As..." button.

Save Search As	
Save Search	Enter a name for the search that will be saved.

The name of search must be unique. Duplicated names aren't allowed

#### Select a saved search

You can select a previously saved search from the right top drop-down list. After select a search, the search panel is updated.

NcML URLdump ESGF	
IPSL-Historical(CORDEX)	•
<< New search >> Historical- ACCESS1.0 BNU-Control(CMIP5) IPSL-Historical(CORDEX) Historical(Specific Humidity)CCMA/IPSL	

#### Reset search panel

To restart the search panel must be selected in the drop-down list the option: <<New search>>

## Metadata Harvesting of ESGF Datasets of a search

The metadata harvesting will allow, after completed, download the datasets from multiple data nodes, and will allow exploring climate datasets without having to download the dataset itself. So we can know its nature in detail before selecting them to download or we can know the services offered from ESGF to access and/or explore the dataset. Some of these services can be explored from ToolsUI NetCDF Java as will be described after.

The harvesting is doing at dataset level.

NetCD	DF (4.3) Tools	• ×						
<u>System M</u> odes <u>D</u> ebug <u>H</u> elp								
Viewer Writer NCDump Iosp CoordSys FeatureTypes THREDDS Fmrc GeoTiff Units NcML URLdump ESGF								
Login Remaining time of credentials: 69, 34 hours (Expire date: Sat May 31 10:19:37 CEST 2014)								
Saved searches & Harvesting Search Downloads								
New search								
historical-6hr(CMIP5)	100%							
experiment=[historical] institute=[NASA-GISS, BNU, BCC, CCCMA] project=[CMIP5] time_frequency=[6hr] variable=[psl, ta, hus]	remove reset harvested 29/29 Datasets (files:2333/4851 size:3,5 TB/7,1 TB)							
Explore search Edit Search	Completed Export to Metalink Download	=						
IPSL-CM5A-LR(1980-1990)-zosga	22%							
experiment=[decadal1980, decadal1985, decadal1990] model=[IPSL-CM5A-LR] variable=[zosga] Explore search	start remove reset harvested 4/18 Datasets (files: 4/172 size: 49, 4 KB/20, 2 GB so far) Edit Search							
		T I						
IDSI_Historical(CORDEX)	0%	•						

## What is a ESGF Dataset?

Dataset is a climate data in a specific version stored by ESGF. One dataset may have several records i.e. the records are the physical replica of a dataset in a data node.

Also, the datasets are formed of files, i.e. dataset are a virtual container of data, so that the information is contained in files and sometimes in aggregations. The versions of datasets are generated when errors are found in the datasets, giving rise to a new corrected version.

Datasets, files and aggregations have replicas in data nodes. And the ESGF services (THREDDS, LAS, HTTP, GridFTP and OPeNDAP) are always served at replicas level. That's the reason why the harvesting must be done before the download.

#### Search Harvesting Panel

The harvesting panel may be selected clicking on the sub-tab "Search Harvesting". In this panel are deployed a list of searches and their harvesting states. Also, provides several options for flow control, complete exploration of harvested datasets and the posibility to do a manual selection of files to put them in download queue.

Each harvesting for a search is an element in the list deployed in the harvesting panel. In the figure below can see one search harvesting. In the left are the search data (name of search and configured parameters). In the right are the state data of the harvesting and the flow options, also shows the number of files that are selected from the total number of files and their sizes in bytes.

One harvesting is always of a completed dataset (with all files). However, a search may include only some files of the total in a dataset (e.g a search with less variables that there is in the dataset or a search in a range of time). That is why, by default, the files selected to future downloads (in the application or by generation of metalink) are always the files that satisfy ALL search constraints.

Noteworthy that this application allows to manually select files to download in case that you want to download files that are not selected by default or deselected some of these.

historical-6hr(CMIP5)	24%
experiment=[historical] institute=[NASA-GISS, BNU, BCC, CCCMA] project=[CMIP5] time_frequency=[6hr] variable=[psl, ta, hus] Explore search	pause     remove     reset       harvested 7/29 Datasets (files: 71/128 size: 198,6 GB/412, 9 GB so far)       ETA:1 min, 14 sec       Edit Search

The flow control options are below progress bar:

- start start the harvesting
- pause pause the harvesting
- remove remove the search and its harvesting (but the harvesting remains in file system)
- reset remove all data of harvesting in file system and put the state of harvesting to zero)

In the center bottom there are the options to explore and download the harvested datasets:

Explore Search	is always visible and allows explore and put to download a individual dataset. Also, view a individual state of harvesting of a dataset
Edit Search	is always visible and allows edit a search in the Search Panel
Export to metalink	is visible when harvesting of search is completed. Allows to generate a metalink file with the files and its resources (replicas in data nodes)
Download	is visible when harvesting of search is completed. Allows to put to download all files that satisfy the constraints of the search or to put to download a set of files manually selected of the search

## **Exploring search harvesting**

To exploring the search harvesting of a search you must click on Explore Search option that is always visible in harvesting panel. And then a window pops up with the exploring options.

This window provides a paginated list of harvesting states of datasets that belong to the search. Each dataset are identify by its instance\_id. The last value comma separated is the version of dataset. The exploring options in this window are provided at dataset level. Each dataset have options, and these are only provided when the dataset harvesting is completed.

In ESGF the dataset are identify by id, instance\_id and master\_id. And in this application the "instance\_id" is the more important because identify all replicas across federation and is specific to each version.

For more information <a href="https://github.com/ESGF/esgf.github.io/wiki/ESGF\_Search\_REST\_API#identifiers">https://github.com/ESGF/esgf.github.io/wiki/ESGF\_Search\_REST\_API#identifiers</a>

In figure below can see the exploring window and its options.

				_ 0 X			
Dataset: cordex.	output.EUR-11.IPSL-INERIS.IPSL-IPSL-	CM5A-MR.historical.rlilpl.WRF	331F.v1.mon.rlus.v20140301				
Completed	size: 188,5 MB	Harvesting View	Open metadata	Download dataset			
Dataset: cordex.	output.EUR-44.IPSL-INERIS.IPSL-IPSL-	CM5A-MR.historical.rlilpl.WRF	331F.vl.mon.pr.v20140301				
Completed	size: 15,2 MB	Harvesting View	Open metadata	Download dataset			
Dataset: cordex.	Dataset: cordex.output.EUR-11.IPSL-INERIS.IPSL-IPSL-CM5A-MR.historical.r1i1p1.WRF331F.v1.mon.rsdt.v20140301						
Harvesting							
Dataset: cordex.	Dataset: cordex.output.EUR-11.IPSL-INERIS.IPSL-IPSL-CM5A-MR.historical.rlilp1.WRF331F.vl.day.sfcWindmax.v20140301						
Harvesting							
Dataset: cordex.	output.EUR-11.IPSL-INERIS.IPSL-IPSL-	CM5A-MR.historical.rlilpl.WRF	331F.vl.mon.tasmin.v201403	301			
Harvesting							
				05/19			

## Harvesting view

Harvesting View option allows to explore a dataset in a tree view. Here you can explore the metadata for datasets that include files, replicas of datasets, replicas of files and finally, ESGF services offered for explore them (by URL endpoint).

- LAS and THREDDS are services of datasets offered at replica level.
- HTTP, OPeNDAP and GridFTP are services of files offered at replica level.

	3 )	ž.
ip5.output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.0mon.r5i1p1.v20111119 {}		•
<pre>replicas: [4]     [0] http://dias-esg-rp.tkl.iis.u-tokyo.ac.jp {}     [1] http://abedo2.dkrz.de {}     [1] http://wesg.ipsl.fr{}     [2] http://wesg.ipsl.fr{}     [2] http://wesg.ipsl.fr{}     [2] http://wesg.ipsl.fr/las/getUI.do?catid=2775C793BD2CBFA8D451089FC11EC4C6_ns_cmip5.output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.     CATALOG: http://wesg.ipsl.fr/thredds/esgcet/4/cmip5.output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.     or [3] http://esgf-data1.ceda.ac.uk {}     datetime_start: 1850-01-16T12:00:00Z     model: [IPSL-CM5A-LR]</pre>	Omo utl.	
type: Dataset	-	_
datetime stop: 2005-12-16T12:00:00Z		
drs id: [cmip5_output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.0mon.r5i1p1]		
version: 20111119		
timestamp: 2014-03-28T23:47:55Z		
title: project=CMP5, model=PSI-CM5A4R, Institut Pierre-Simon Laplace, experiment=historical, time frequency=mon, modeling realm=ocean, ensemble	e=r	
realm: [ccean]		
number of aggregations: 86		
Hamber_o_genergenerations of the second s		
description (c. 00.001)		
experiment (instantia) variable long name: [/]		
variable_initiation ( + 0) dataset id template + (rmin5%(nodult)%%(institute)%%(model)%%(experiment)%%(time frequency)%%(realm)%%(cmor table)%%(ensemble)%		
actact is control to the second state of the s		
files: [64]		
<ul> <li>[0] cmip5.output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.Omon.r5ilp1.v2011119.hfevapds_Omon_IPSL-CM5A-LR_historical_r5ilp1_185001-200512.</li> <li>[1] cmip5.output1.IPSL.IPSL-CM5A-LR.historical.mon.ocean.Omon.r5ilp1.v2011119.rsntds_Omon_IPSL-CM5A-LR_historical_r5ilp1_185001-200512.nc</li> <li>[4] [0] http://dias-esg.rp.tkl.iis.u-tokyo.ac.jp{}</li> <li>[1] http://dias-esg.rp.tkl.iis.u-tokyo.ac.jp{}</li> </ul>	nc { {}	
T I The phylometeo 2 on 2 de T     Herein Construction	005	
Services {}	rsnt mo /rsn	
		-

#### Open metadata

Open metadata option allows to explore an abstract of metadata of a dataset in a table view.

Metadata	Values				
CF_STANDARD_NAME	[surface_air_pressure, specific_humidity, air_temperature,				
	eastward_wind, northward_windj				
CMOR_TABLE	[6hrLev]				
DATASET_ID_TEMPLATE_	[cmip5.%(product)s.%(institute)s.%(model)s.%(experiment)s.% (time_frequency)s.%(realm)s.%(cmor_table)s.%(ensemble)s]				
DATETIME_START	1850-01-01T03:00:00Z				
DATETIME_STOP	2005-12-31T21:00:00Z				
DESCRIPTION	[IPSL-CM5A-LR model output prepared for CMIP5 historical]				
DRS_ID	[cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r 5i1p1]				
EAST DEGREES	356.25				
ENSEMBLE	[r5ilp1]				
EXPERIMENT	[historical]				
EXPERIMENT_FAMILY	[All, Historical]				
FORCING	[Nat,Ant,GHG,SA,Oz,LU,SS,Ds,BC,MD,OC,AA]				
FORMAT	[netCDF, CF-1.4]				
INSTANCE_ID	cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r 5ilp1.v20111119				
INSTITUTE	[IPSL]				
LATEST	true				
MASTER_ID	cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r 5ilp1				
METADATA FORMAT	THREDDS				
MODEL	[IPSL-CM5A-LR]				
NORTH_DEGREES	90.0				
NUMBER_OF_AGGREGATIONS	8				
NUMBER_OF_FILES	68				
PRODUCT	[output1]				
PROJECT	[CMIP5]				
REALM	[atmos]				
Close					

#### **Download Dataset - File selection window**

Download Dataset... allows to manually select the files that will be put to download. In the top of window can see the id of dataset, the number of files selected, the total number of files of the dataset, the size in bytes of the sum of selected files and the total size of the whole dataset.

This windows displays a list of files and its size in bytes. You can select manually the files that you want download doing click in the check-box associated in each file in the list.

cmip5.output1.CCCma.CanESM2.historical.6hr.atmos.6hrPlev.r4i1p1.v20120618 files: (14/28) size: (11,6 GB/29,0 GB)	
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_194101010000-195012311800.nc (456, 6 MB)	-
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_195101010000-196012311800.nc (456, 6 MB)	=
☑ psl_6hrPlev_CanESM2_historical_r4i1p1_196101010000-197012311800.nc (456,6 MB)	
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_197101010000-198012311800.nc (456, 6 MB)	
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_198101010000-199012311800.nc (456, 6 MB)	
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_199101010000-200012311800.nc (456, 6 MB)	
✓ psl_6hrPlev_CanESM2_historical_r4i1p1_200101010000-200512311800.nc (228, 3 MB)	
✓ ta_6hrPlev_CanESM2_historical_r4i1p1_194101010000-195012311800.nc (1, 3 GB)	-
Deselect all         Select all         Filter by constraints of search         Download selected         Classical	ose

This window also provides the next options:

- Deselect all
- Select all
- Filter by constraints of search select only files that satisfy the constraints of search
- Download selected put to download the files selected in this window

When you click on Download selected a window asks for destination path.

	Do you want to change path of downloads?					
The files will be downloaded at default path: /home/ **** /ESGFDATA Do you want to change the path of the downloads?						
	Yes <u>N</u> o					

#### **Export to Metalink**

To export the search to Metalink you must click on "Export to Metalink". This option is is visible when harvesting of search is completed. And then a window pops up with the save options.

The files that will be include in the Metalink file are only the files that satisfy the constraints of search.

	Guardar		
Guardar en: 📑 me	talink	-	a 🗇 🗖 🔡 🗄
/	оте		
	***** ] metalink		
<u>N</u> ombre de archivo:	BCC-CSM1-1_psl_ta_hus.me	talink	
Archivos de <u>t</u> ipo:	Todos los Archivos		-
		Guar	dar Cancelar

Metalink is an extensible metadata file format that describes one or more computer files available for download. It specifies files appropriate for the user's language and operating system; facilitates file verification and recovery from data corruption; and lists alternate download sources (mirror URIs). For more info: <u>?Metalink, ?Metalink-Wikipedia</u>

Exist many clients that supports Metalink: <u>?Aria2</u>, <u>?GetRight</u>, <u>?DownloadThemAll</u>, <u>?Orbit Downloader</u>, etc.

Note that if you want use a external client to download ESGF files you must configure the client for using the ESGF certificates.

#### **Download a Search**

To put the files that belongs to a search in downloads queue you must click on Download .... This option is visible when harvesting of search is completed. And then a window pops up with the download options.

In this window (that you can see below) are displayed the list of datasets that satisfy the constraints of search. Each dataset have: id, description, the number of files that satisfy the search constraints with respect the total number of files and the size in disk of the selected files with respect the total size in disk of dataset.

Id: cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r3i1p1.v20110406 Description: [IPSL-CM5A-LR model output prepared for CMIP5 historical] Number of files selected: 16/68 (312,8 GB/1,2 TB)	
Select files to download	
Id: cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r6i1p1.v20120526 Description: [IPSL-CM5A-LR model output prepared for CMIP5 historical] Number of files selected: 16/68 (312,8 GB/1,2 TB)	E
Select files to download	
Id: cmip5.output1.IPSL.IPSL-CM5A-LR.historical.6hr.atmos.6hrLev.r5i1p1.v20111119 Description: [IPSL-CM5A-LR model output prepared for CMIP5 historical] Number of files selected: 16/68 (312,8 GB/1,2 TB)	=
Select files to download	
Id: cmip5.output1.CCCma.CanESM2.historical.6hr.atmos.6hrLev.r1i1p1.v20120410 Description: [CanESM2 model output prepared for CMIP5 historical] Number of files selected: 56/280 (89,8 GB/361,8 GB)	
Select files to download	,
	[
Download Close	

For download all files that satisfy the constraints (without manually selection) you must click on "Download" button that is in the bottom of window. Then a dialog box will be showed to confirm the download and select the destination path.

	Select an option	
?	Are you sure that you want to download all files that satisfy the const	raints?
	Number of files to download: 184 of 824 files	
	Size of download: 2,5 TB	
	Ye <u>s</u> <u>N</u> o Cancel	
	Do you want to change path of downloads?	
?	The files will be downloaded at default path: /home/ **** /ESGFDATA Do you want to change the path of the downloads?	
	Yes <u>N</u> o	

For manually selection you may click on "Select file to Download", that displays a file selection window as explained in File selection window

## Download of ESGF Datasets

To download ESGF Datasets you must done previously a harvesting of a search. The downloads panel is in "Download" tab.

You can add to the download queue:

- A complete search all files of datasets that satisfy the constraints of the search (view Download a Search) •
- A personalized file selection of the set of datasets that satisfy the constraints of search (view File selection window)

The most of data in ESGF require credentials to access them. To access the user must have:

1. A ESGF Account (view ESGF Login)

2. This account have to be authorized to access the desired data. This authorization is done by control groups (each account can belong a many groups). Each group is authorized for download a set of data.

#### To read more about register in ESGF go to (?How to register and download data from esgf)

<u>System M</u>odes <u>D</u>ebug <u>H</u>elp

NetCDF (4.3) Tools

TUREDDE Contiff Unite NeMI (URIdu

DOWNLOADING 1,28 GB

45.67 MB

1.23 GB

0 B

0 B

FINISHED

WAITING

DOWNLOADING

1,<mark>60</mark> GB

45.67 MB

1.60 GB

1,<mark>60</mark> GB

dapp2p.cccma.ec.gc...

dapp2p.cccma.ec.gc.

dapp2p.cccma.ec.gc..

cmip3.dkrz.de

viewer writer NCD	Jump Tosp Coordsy	s reaturerypes in	EDDS FIIIC Geofin	UTILS NUML	OKLUUIIIP	ESGF	
Login 🚦 Remaining	time of credentials: 70,	54 hours (Expire date: F	ri Jun 13 10:13:48 CEST 201	4)			
Saved searches & Han	vesting Search C	ownloads					
Start all downloads	Pause all downloads	Remove all download	s				
	Name		Progress	Status	Current Size	TotalSize	Data node
cmip5.output1.CCCm	a.CanESM2.historical.day.a	atmos.day.rli1p1.v2012	100%	FINISHED	104,31 GB	104,31 GB	
💡 🛅 cmip5.output1.CCCm	a.CanESM2.historical.6hr.a	atmos.6hrLev.r1i1p1.v20	8%	DOWNLOADING	29,63 GB	361,81 GB	
🗋 ua_6hrLev_CanESi	M2_historical_r1i1p1_1961	01010000-1961123118	100%	FINISHED	1,60 GB	1,60 GB	dapp2p.cccma.ec.gc
🗋 va_6hrLev_CanES!	M2_historical_r1i1p1_1998	01010000-19981231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	cmip3.dkrz.de
🗋 ua_6hrLev_CanESi	M2_historical_r1i1p1_1976	01010000-1976123118	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	cmip3.dkrz.de
🗋 ta_6hrLev_CanESN	M2_historical_r1i1p1_1972(	01010000-19721231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ps_6hrLev_CanESI	M2_historical_r1i1p1_1968	01010000-1968123118	100%	FINISHED	45,67 MB	45,67 MB	cmip3.dkrz.de
🗋 va_6hrLev_CanES!	M2_historical_r1i1p1_1956	01010000-19561231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ta_6hrLev_CanESN	M2_historical_r1i1p1_2002(	01010000-20021231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ta_6hrLev_CanESN	M2_historical_r1i1p1_1953(	01010000-19531231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	cmip3.dkrz.de
🗋 va_6hrLev_CanES!	M2_historical_r1i1p1_1977	01010000-19771231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ps_6hrLev_CanESI	M2_historical_r1i1p1_1971	01010000-1971123118	100%	FINISHED	45,67 MB	45,67 MB	cmip3.dkrz.de
🗋 ua_6hrLev_CanESi	M2_historical_r1i1p1_1999	01010000-1999123118	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
hus_6hrLev_CanE	SM2_historical_r1i1p1_197	101010000-197112311	100%	FINISHED	1,60 GB	1,60 GB	cmip3.dkrz.de
🗋 ta_6hrLev_CanESM	M2_historical_r1i1p1_19520	01010000-19521231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ua_6hrLev_CanESi	M2_historical_r1i1p1_1974	01010000-1974123118	100%	FINISHED	1, <mark>60 G</mark> B	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 va_6hrLev_CanESI	M2_historical_r1i1p1_2005	01010000-20051231180	100%	FINISHED	1,60 GB	1,60 GB	dapp2p.cccma.ec.gc
🗋 va_6hrLev_CanESI	M2_historical_r1i1p1_1996	01010000-19961231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
ta_6hrLev_CanESN	M2_historical_r1i1p1_1979(	01010000-19791231180	100%	FINISHED	1,60 GB	1, <mark>60 G</mark> B	cmip3.dkrz.de
🗋 ta_6hrLev_CanESM	M2_historical_r1i1p1_1994(	01010000-19941231180	Checking checksum	DOWNLOADING	1,60 GB	1,60 GB	cmip3.dkrz.de
🗋 ua_6hrLev_CanES	M2_historical_r1i1p1_1977	01010000-1977123118	82%	DOWNLOADING	1,32 GB	1,60 GB	dapp2p.cccma.ec.gc

hus\_6hrLev\_CanESM2\_historical\_r1i1p1\_1977010100000-197712311.. WAITING 1,60 GB cmip3.dkrz.de ta\_6hrLev\_CanESM2\_historical\_r1i1p1\_195001010000-19501231180. WAITING 1,60 GB dapp2p.cccma.ec.gc. 0 B The downloads panel displays a list of downloads. Each element of the list is the grouping of files by dataset. The files of each dataset can be showed or

hided. Each element have a progress bar and file elements displays the data node from where the download is acceded.

The files can be in five states:

- No completed: When the download isn't complete. It is displayed with a blue progress bar and a percentage ٠
  - CREATED: File has just been added
  - PAUSED: File has paused
  - WAITING/READY: File is in waiting for be downloaded

ua\_6hrLev\_CanESM2\_historical\_r1i1p1\_199501010000-1995123118...

 ps\_6hrLev\_CanESM2\_historical\_rli1p1\_199101010000-1991123118...

 hus\_6hrLev\_CanESM2\_historical\_rli1p1\_195101010000-195112311...

 ta\_6hrLev\_CanESM2\_historical\_rli1p1\_195001010000-19511231180...

- DOWNLOADING: File is downloading
- UNAUTHORIZED: When the user hasn't been authorized to download a file. It is displayed with a yellow progress bar with the message "UNAUTHORIZED".
- Completed : When the download is completed.
  - FINISHED: Download is successful. It is displayed with a green progress bar.

- FAILED : When happens some error in the current data node. It is displayed with a red progress bar.
- CHECKSUM\_FAILED: When the validity algorithm of file failed. It is displayed with a gray progress bar with the message "CHECKSUM\_FAILED".

psl_6hrPlev_ACCESS1-0_historical_r1i1p1_1970010106-1980010100.nc	24%	376,6 MB 1,5 GB	pcmdi9.llnl.gov
psl_6hrPlev_ACCESS1-0_historical_r1i1p1_1980010106-1990010100.nc	FAILED	380,2 MB 1,5 GB	pcmdi9.llnl.gov
ta_6hrPlev_ACCESS1-0_historical_r1i1p1_1960010106-1961010100.nc	CHECKSUM_FAILED	463,2 MB 463,2 MB	albedo2.dkrz.de
ta_6hrPlev_ACCESS1-0_historical_r1i1p1_1950010106-1951010100.nc	100%	462,0 MB 462,0 MB	albedo2.dkrz.de
psl_6hrPlev_ACCESS1-0_historical_r1i1p1_1990010106-2000010100.nc	UNAUTHORIZED	387,1 MB 1,5 GB	pcmdi9.llnl.gov

If you is logged in ESGF and there are files with "Unauthorized state" is because the groups they are associated with the user account in the federation do not have permissions for that data.

To join a group with the needed permissions for a unauthorized file, view Join a group

## Context menu of downloads

The datasets and the files have a context menu. This menu are displayed after right-clicking in a dataset or file element. Depending on the state of download are enabled different options:

Option	Description
Start download/Start all file downloads	Start the download
Pause download/Pause all file downloads	Pause the download
Reset	To reset all download. Remove file/files of file system (disk) and reset its/their status
Retry/Retry download in failed files	To retry download. In a file element, you must select a data node and then the download is reinitialized. In a dataset element, the download is retry in all files with failed state
Remove	To remove a file or a dataset and their files of the downloads queue. This option doesn't delete the files from disk
Open download URL in browser (only file)	Open the url of the HTTP service from which the file is being downloaded.
File info/Dataset info	To display info of file/dataset
Services	To access services of files/datasets (the services will be described in the next table)

🗋 ta_6hrLev_CanESM2_historical_r1i1p1_198201010000-198	21231180	85%	DOWNLOADING	1,38 GB	1, <mark>60</mark> GB	dapp2p.cccma.ec.gc
ps_6hrLev_CanESM2_Start download	7123118	100%	FINISHED	45,67 MB	45,67 MB	cmip3.dkrz.de
ps_6hrLev_CanESM2	123118	100%	FINISHED	45,67 MB	45,67 MB	cmip3.dkrz.de
ta_6hrLev_CanESM2	1231180	78%	DOWNLOADING	1,26 GB	1,60 GB	cmip3.dkrz.de
ua_6hrLev_CanESM2 Reset	.123118	45%	DOWNLOADING	744,31 MB	1,60 GB	dapp2p.cccma.ec.gc
ta 6hrLev CanESM2 Retry	1231180	42%	DOWNLOADING	690,66 MB	1,60 GB	cmip3.dkrz.de
ta_6hrLev_CanESM2 Remove of downloads queue	1231180	9%	DOWNLOADING	152,29 MB	1,60 GB	dapp2p.cccma.ec.gc
va_6hrLev_CanESM2 Open download URL in browser	1231180	0%	WAITING	0 B	1,60 GB	cmip3.dkrz.de
ps_6hrLev_CanESM2 File info	123118	0%	WAITING	0 B	45,67 MB	dapp2p.cccma.ec.gc
va_6hrLev_CanESM2 Services	THREDDS	0%	WAITING	0 B	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 ta_6hrLev_CanESM2_historical_r1i1p1_199801010000-199	LAS	0%	WAITING	0 B	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
🗋 va_6hrLev_CanESM2_historical_r1i1p1_196601010000-196	Local	0%	WAITING	0 B	1, <mark>60 G</mark> B	cmip3.dkrz.de
🗋 va_6hrLev_CanESM2_historical_r1i1p1_200201010000-200	OPeNDAP	Open in Viewer Panel	▶ G	0 B	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
ps_6hrLev_CanESM2_historical_r1i1p1_195001010000-195	HTTP N	Open in Features Types P	anel b G	0 B	45,67 MB	cmip3.dkrz.de
ps_6hrLev_CanESM2_historical_r1i1p1_199001010000-199	GridETP	Open LIBL in browser	G	0 B	45,67 MB	dapp2p.cccma.ec.gc
🗋 va_6hrLev_CanESM2_historical_r1i1p1_198001010000-198	01251100	Copy UPL to cliphoard	G	0 B	1, <mark>60 G</mark> B	dapp2p.cccma.ec.gc
nta 6hrLev CanESM2 historical rlilpl 196001010000-196	01231180[	copy one to cipboard	G	0 B	1.60 GB	dapp2p.cccma.ec.gc

## Services of context menu of downloads

Service	Description
TOTAL (Day states)	Open in 1982200 Jonal Alexa select a TARECCO Antonio in a dis- nale la las di inte tati Catalog Oncoré al Tanàci RecCP. Ana Open III. in la mavan Quen TARECCO ESCP service URL in Interna- Copy III. in a rispinsari Capy TARECCO ESCP service URL in Spinsari
List (Drly dataset)	Open URL in Internet OpenLAXESDP service URL Information Open URL to a Fliphmand Copy LAXESDP service URL in alphaned
Local	Open in: Viscour Exact (styl) line # PEDERC study To last listed line in the information of Tashida NetCDP Jacob Open in: Functionary Open Jacob (styl) line in PEDERC study To had listed line in the test for Annuary Open of Tashida NetCDP Jacob Open Al sectory Open Restlement (styl) Res(Corp) Res (in Spatia) Open Al sectory Open Restlement (styl) Res(Corp) Res (in Spatia)
coverso (Soly Re)	(see in S. Venere 2 and 7) has fits the EDP to OVERF in the MFT income of Media/NetCore Laws. In the Internet Property of the Media Sector Laws. Of PhetoPhetoPhetoPhetoPhetoPhetoPhetoPheto
8772 (Dag Be)	Open in Viewer Janei Tahadila han EEDF by HTTP is beind. Viewer 4 Indial MeCCV Jane Open in Fauran Open Janei Tahadila han EEDF by HTTP is Beind Fauran Open Janei Tahadila CCF Jane Jahas Jaho HTTP is Beind Fauran Open Hall in Internet Open HTTP EEDF service UK is between Open UK. In Internet Open HTTP EEDF service UK. In Internet
Generation (Delynter)	Open III. in Innuer OpenDidPTP EDDP service URL in Innuer Oppy III. in a cliphoard Opp DidPTP EDDP service URL in Innuer

NetCDF (4.3) Tools				
<u>System Modes Debug H</u> elp				
Viewer Writer NCDump losp CoordSys FeatureTypes THREDDS Fmrc GeoTiff Units NcML URLdump ESGF				
Catalog Chooser Catalog Crawler Catalog Enhancer Catalog Copier TDS Configure				
Catalog URL http://esg.bnu.edu.cn/thredds/esgcet/1/cmip5.output1.BNU.BNU-ESM.esmControl.mon.land.Lmon.rl 🗸 Connect 🔄 🙀				
project=CMIP5, model=BNU-ESM, College of Global Chang	i Sy 🕶			
<ul> <li>baresoilFrac_Lmon_BNU-ESM_esmControl_rlilp1_1450</li> <li>cmip5.output1.BNU.BNU-ESM.esmControl_rnon.land.Lmp</li> <li>clitter_Lmon_BNU-ESM_esmControl_rlilp1_145001-170</li> <li>cmip5.output1.BNU.BNU-ESM.esmControl_rlilp1_145001-170</li> <li>clitterAbove_Lmon_BNU-ESM_esmControl_rlilp1_145001-170</li> <li>clitterBelow_Lmon_BNU-ESM_esmControl_rlilp1_14500</li> <li>clitterBelow_Lmon_BNU-ESM_esmControl_rlilp1_145001-170</li> <li>comip5.output1.BNU.BNU-ESM.esmControl_rlilp1_145001-1707</li> <li>cmip5.output1.BNU.BNU-ESM_esmControl_rlilp1_145001-1707</li> </ul>	E =			
Cariables:				
Consider Enterprise Structure of the principal structure of the princip				
• cmip5.output1.BNU.BNU.ESM.esmControl.mon.land.Lmo	-			
Status text				
Open File Open CoordSys Open dataset Source				
fidnmwtt BdLL-ESM.esmControl.mon.land.Lmon.r1i1p1.v20120503.xml#cmip5.output1.BNU.BNU-ESM.esmControl.mon.land.Lmon.r1i1p1.v201	20503			

# ESGF Login : Obtaining credentials with ESGF Account

In the top of ESGF panel there is a login toolbar. This toolbar shows a red icon if you aren't logged or a green icon if you are logged. Login allows retrieve the credentials needed to download the most of data in the federation. To login you must select the "Login" button.

Login 📕 Ren	naining time of cre	dentials: 70,78 hours (Expire date: Sat May 31 10:19:37 CEST 2014)	
Saved searche	s & Harvesting	Search Downloads	

After clicking in "Login", a window is displayed, in this window if you are logged in the bottom of panel is displayed You are logged along with information validity time remaining for the credentials that are being used by the application. Else if you still aren't logged or you want to switch accounts, you must select the identity provider node from the dropdown list and after type in the box the account user name.

Id Provider: pcmdi9.llnl.gov 👻
https://pcmdi9.llnl.gov/esgf-idp/openid/userName
password: ••••••
Login
Not logged.
Ok

You can also paste the OpenID URL directly if selected in the drop down list << Custom OpenID URL >>.

Id Provider: << Custom OpenID URL >> 🔻
https://[IdPNodeName]/esgf-idp/openid/[userName]
password: ······
Login
Not logged.
Ok

The session starts after pressing the "Login" button. If the login failed then It will be notified by a message on the bottom of the window.

## Join a group with the needed permissions for a unauthorized file

Check that you are logged in. If you are logged and the Unauthorized state remains then you must do the follow steps

- 1. Deploy the context menu of the file in state "Unauthorized?
- 2. Select "Open download URL in browser" option
- 3. A ESGF site will be loaded in the browser. Select join to the group.
- 4. Go to "Login" Panel in Tools UI
- 5. Login with you account again (to retrieve new credentials)

If you are logged and you have the specific group permissions. Then, probably is an error in data node.

Reset the download and try in another data node