EUROPEAN COMMISSION

Community research

Climate Information Portal

CLIPC DELIVERABLE (D -N°: 3.1)

Conceptual design of the CLIPC portal

File name: { ***D3 1_CLIPC_concept_design_portal*****.docx or .pdf}**

Dissemination level: PU (public)

Author(s): *Peter Thijsse e.a.*

Reviewer(s): *V. Bennett - STFC*

Release date for review: ***29/04/2015***

Final date of issue: ***18/05/2015***

Revision table			
Version	**Date**	**Name**	**Comments**
0.1	Feb 2015	P. Thijsse, W. Som de Cerff, E. de Vreede	concept
1.0	Apr 2015	W. Som de Cerff, M. Plieger, M. Juckes	Updates and corrections included
1.1	May 2015	P. Thijsse, V. Bennett	Review comments included

Abstract

This document describes conceptual design of the CLIPC portal. This includes the architecture on which the portal interface relies (chapter 2), followed by separate chapters for the specifications of the user interfaces required in the CLIPC portal, communication needed and first design sketches (chapter 3-8) This document also defines who creates which software components.

Table of contents

1. Introduction 6
2. Architectural summary underlying data service infrastructure 7
2.1 Administrative view 7
2.2 Services and user interfaces 8
3. Background principles 9
3.1 CLIPC portal design principles 9
3.2 Target user of CLIPC services 9
4. CLIPC portal website 11
4.1 Design portal homepage 12
4.2 Menu structure version 2 12
4.3 Content Management System 14
4.4 Portal releases 14
5. Climate data access 16
5.1 Climate data discovery 16
5.1.1 Concept 16
5.1.2 Implementation 17
5.1.3 Design examples 18
5.1.4 Partners involved and planning 20
5.2 CLIPC dataset catalogue 21
5.2.1 Concept 21
5.2.2 Implementation 22
5.2.3 Design examples 22
5.2.4 Partners involved and planning 23

6. Data viewing 24
6.1 Concept 24
6.2 Implementation 25
6.3 Design examples 25
6.4 Partners involved and planning 27
7. User authentication 29
7.1 Concept 29
7.2 Implementation 29
7.3 Design 30
7.4 Partners involved and planning 31
8. The binding factor: MyCLIPC data processing and visualisation 32
8.1 Concept 32
8.2 Implementation 32
8.2.1 Access to services and calculated data 33
8.2.2 Using PyWPS for processing 33
8.2.3 Summarising development approach 34
8.3 Design examples 34
8.4 Partners involved and planning 36
9. Guidance and education 37
9.1 Concept 37
9.2 Commentary information 37
9.2.1 Implementation commentary information 37
9.2.2 Design examples 38
9.3 Technical documentation 39
9.3.1 Implementation 39
9.3.2 Design examples 39
9.4 Glossary 40
9.4.1 Implementation 40
9.4.2 Design examples 41
9.5 Literature database 44
9.5.1 Implementation 44
9.6 Partners involved and planning 44
Annex 1: List of abbreviations and acronyms 46

Executive Summary

This document aims to describe the architecture on which the CLIPC portal interface relies (chapter 2), followed by separate chapters for the specifications of the user interfaces required in the CLIPC portal, communication needed and first design sketches (chapters 3-). It also defines who creates which software.

The document is written to guide the actual development process with respect to the integration and use of the various CLIPC services from the portal (services developed in the different work packages). In each chapter the functional specification for services have been drafted, plus design examples of the user interface. In year 2 of the project members from WP3, 4 and 5 will in close collaboration work on actual development of the services. The plans are not rigid because over time ideas might change and new techniques pop up. Therefore, during the development different solutions might be chosen. These changes will be documented in final deliverables.

Other work packages are urged to make use of the ideas of this document. Ideas and plans will be discussed with the relevant project members.

1. Introduction

The CLIPC portal is the linking pin from project results to the users. User interfaces will be provided to present underlying services (developed by, or incorporated under CLIPC) to the users.

This document describes the architecture on which the portal interface relies (chapter 2), followed by separate chapters (3-8) for the specifications of the user interfaces required in the CLIPC portal, communication needed and first design sketches. It also defines who creates which software.

The document is written to guide the actual development process with respect to the integration and use of the various CLIPC services from the portal (services developed in the different work packages). Other work packages are urged to make use of the ideas of this document. Ideas and plans will be discussed with the relevant project members.

2. Architectural summary underlying data service infrastructure

This section summarises the results of the Architecture Team document (no official deliverable but available under http://www.clipc.eu/the-project/public-deliverables-and-milestones). This document describes the concept of the components/services of the CLIPC data service infrastructure taking the initial ideas from the CLIPC description of work as basis.

2.1 Administrative view

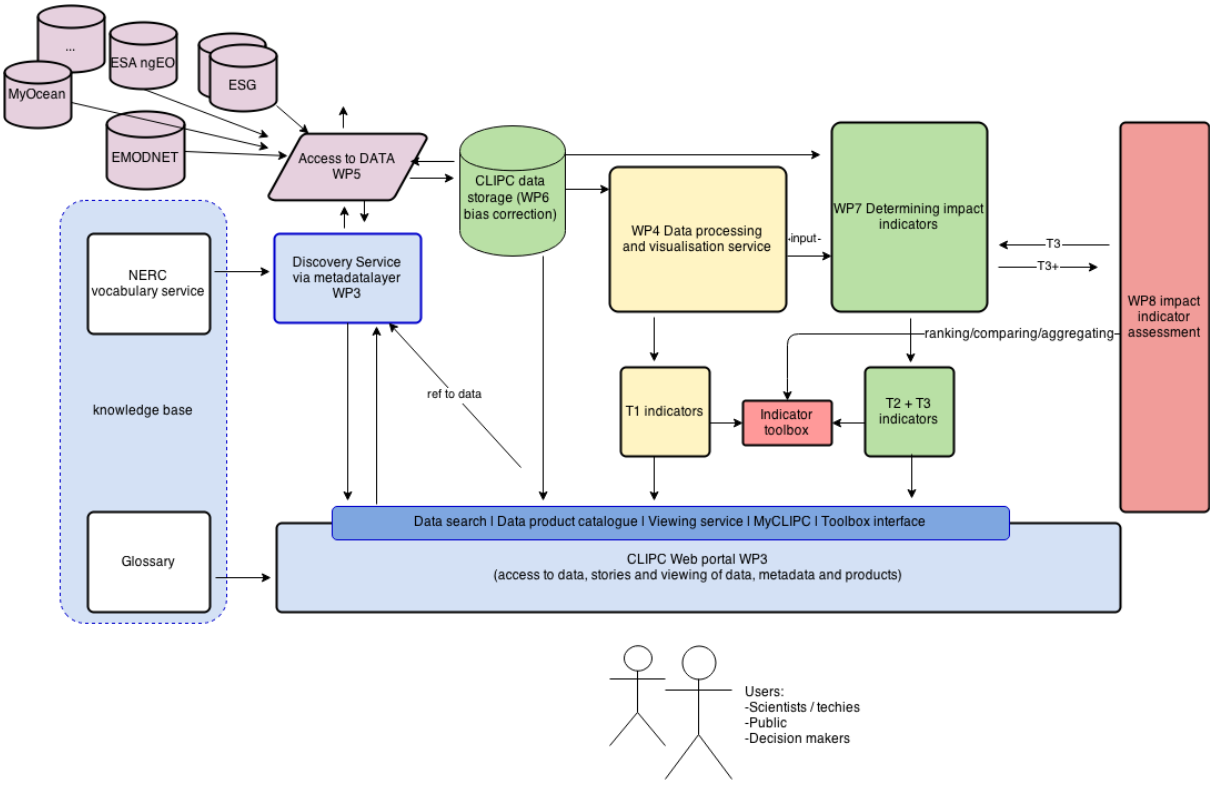

Figure 1: Administrative view of the CLIPC data flow

In CLIPC climate data flows from data source to user, underway transformed from raw data, validated and processed into climate impact indicators which can be presented to the user. Figure 1 presents a more or less administrative overview of CLIPC where the involvement and output of the various work packages is illustrated and combined with the software components required to support the flow of data.

2.2 Services and user interfaces

The CLIPC infrastructure will consist of the following services or group of services:

- Central web portal (Project background) with Content Management System
- Data search and access service
- MyCLIPC data processing service
- CLIPC Climate Impact Indicator toolbox
- Data viewing service
- Knowledge and information service (also known as knowledge base)
- User identification

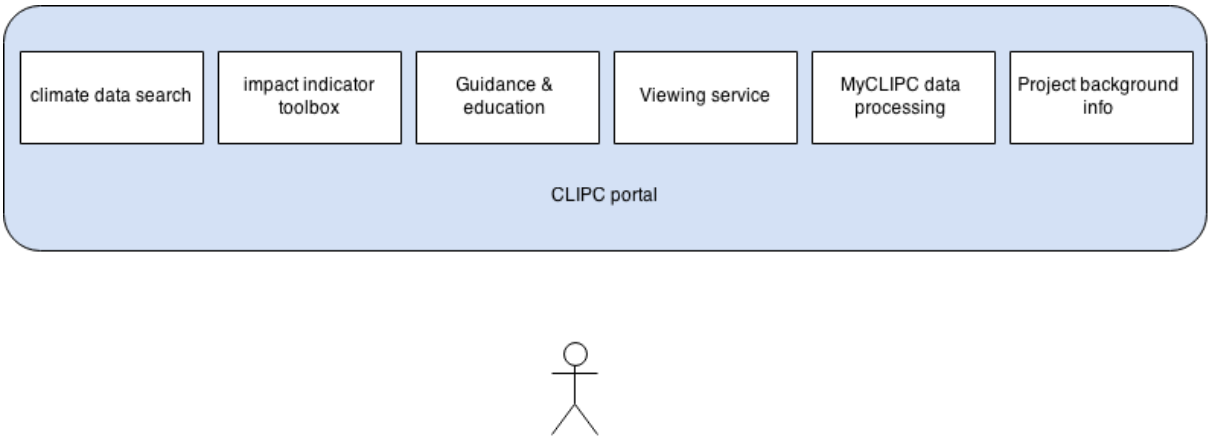

Figure 2: Portal component overview

The overarching webportal will provide a gateway onto these infrastructure components, as illustrated schematically in figure 2.

In this document the interface to each component is described in more detail: From user interface to underlying service, and independent of the work package. Where necessary the interrelation to other services is also mentioned.

The new CLIPC webportal will be developed in parallel to portal version 1 which is already deployed. After first release the « staging » version will remain and be the basis for development and testing. After successful tests new/upgraded components will be released to the operational portal.

3. Background principles

This chapter describes principles and target users of the CLIPC portal and underlying architecture based on the inventory of user needs.

3.1 CLIPC portal design principles

The CLIPC portal design aims to incorporate the following:

- Clear and as simple as possible
- Distinguishing interfaces served to different users without having users to choose which type of user they are in advance. The type of service offered will indicate them which they should need.
- The portal design needs to be “responsive” in order to be accessible from devices with varying screen resolutions.
- The HTML5 and CSS needs to be optimised for access by Google and other search engines.

3.2 Target user of CLIPC services

In CLIPC D2 .1 and D2.2 an extensive survey on the target users of CLIPC can be found. Quoting from D2.1 Chapter 4.2, priority user groups for CLIPC are:

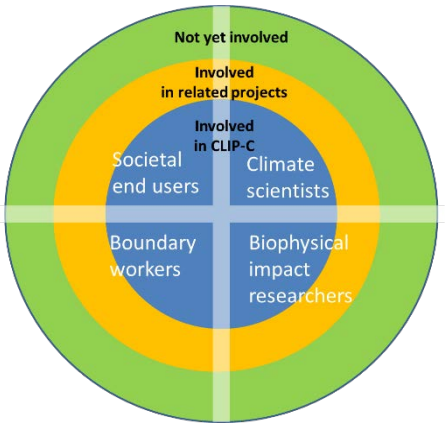

Figure 3 : Suggested classification of users groups in three circles dependent on connection to the project

First circle: Potential users already involved in projects of CLIPC partners;
Second circle: Users already involved in other similar European and national projects. The CLIPC partners participate in these related projects;
Third circle: Potential users of interest known by various partners but not necessary involved in any projects at the time.

To get a more targeted picture of the user needs we subdivide each of these three groups into four categories, according to expected requirements and capabilities to handle climate:
A. Climate Scientists
B. Impact researchers
C. Boundary workers (or intermediary organizations) and socio-economic impact researchers
D. Societal end-users

Category A involves climate scientists but also includes the data providers themselves, who for instance need observational data to evaluate their model results or to make use of empirical ground truth in other processing in order to make their results more applicable (e.g. in mapping results).
The biophysical impact researchers group (category B) may overlap with the 'climate scientists' to some extent, but this category is not involved in developing and running climate models. This group may include the downscaling community who need data for empirical-statistical downscaling or bias correction, in addition to model validation. However, most of this group will be people from the hydrology, biology, agriculture, and engineering communities, with some experience dealing with data, and statistics.
'Boundary workers' in category C include consultants who work at the interface between the scientists and the societal end users. Boundary workers perform as intermediaries. Organizations such as the EEA but also consultant agencies and national portals can be considered boundary workers. Category C also involves socio-economic impact researchers who are assumed to apply climate data with less scientific literacy and numerical skill than the impact researchers.
The societal end users in category D represent policy makers (people involved in policy making within governmental institutions and business firms), decision-makers (people who are making the actual decisions like politicians) and practitioners (people involved in the implementation of adaptation such as NGOs, civil servants who often do not have a high climate or science literacy).
In CLIPC the focus will be on user groups A, B and C. The societal end-users will only be reached via the boundary worker, who first have to structure and represent the data products of CLIPC.

4. CLIPC portal website

As indicated in chapter 2 the CLIPC webportal will consist of several components that are developed and hosted in a distributed way. The central portal website is the binding factor that makes them all accessible via one interface, and from the central website the services of the components are triggered after a user action.
One of the components is the central portal website itself, which has an underlying database with content, is hosted on one of MARIS servers and serves the client interfaces to the user. The website is the connection from the user to the services.
See figure 4 for a simple schematic overview: The portal website at the top-right with CMS and database for specific content, and the visualization and processing services (KNMI servers) at the bottom. The user finds interfaces on the portal website to trigger (in this example) processing or visualization services, which are delivered by the portal website or directly to the client.

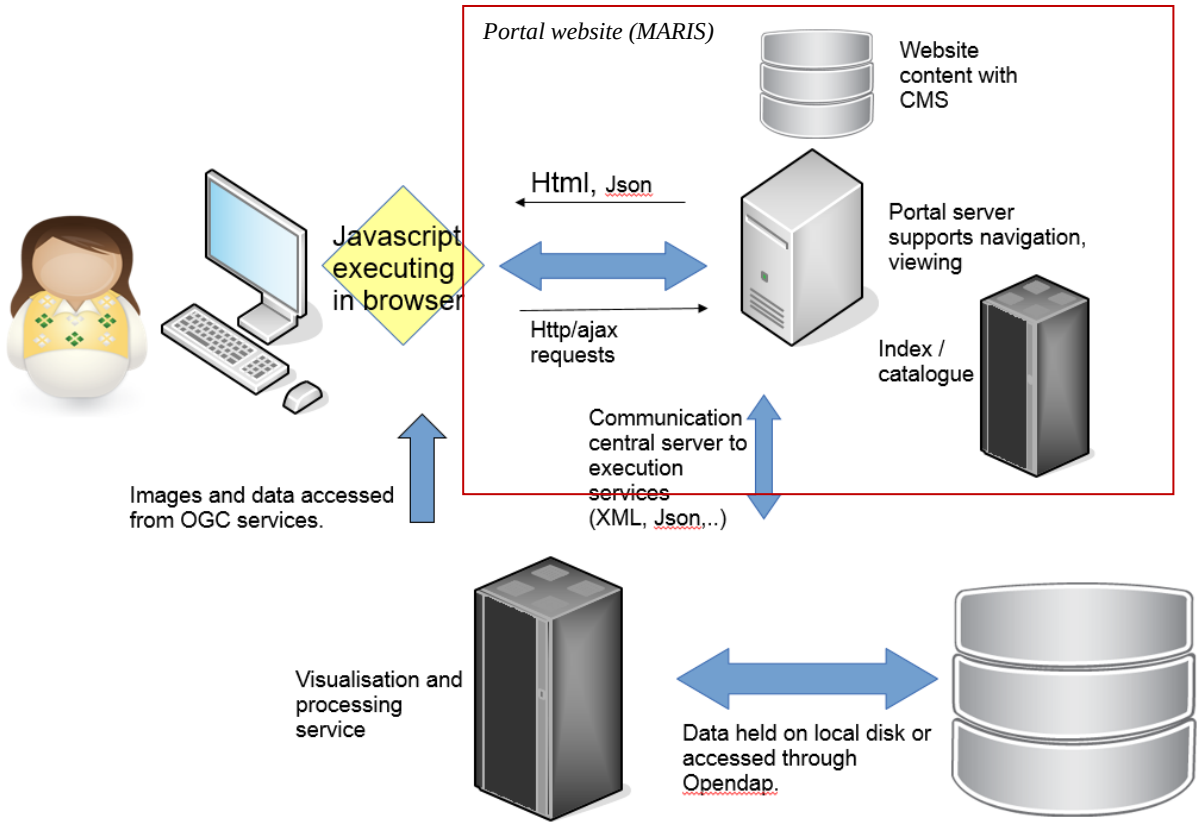

Figure 4: position of central portal website

4.1 Design portal homepage

As indicated in the portal design principles the portal homepage should immediately let the users focus on and select the CLIPC service relevant for them. Therefore the homepage needs to be clear and present the main services directly on the first screen (although scrolling should not be a problem). The services should be presented in a descriptive way, clear enough for the right user to want to click and try.

A first design (based on the existing CLIPC layout) is presented below in figure 5.Users are immediately directed to:

- "MyCLIPC" data processing services (for climate scientist, impact researcher)
- Climate data search (climate scientist)
- Climate impact indicator toolbox (boundary workers)
- Guidance (education, guidance and explanation to all, including general public).

Different users will understand which services have been developed for them, in that way the same portal can supply services for all users types, without questioning the users "which user type" they are (users won't know).

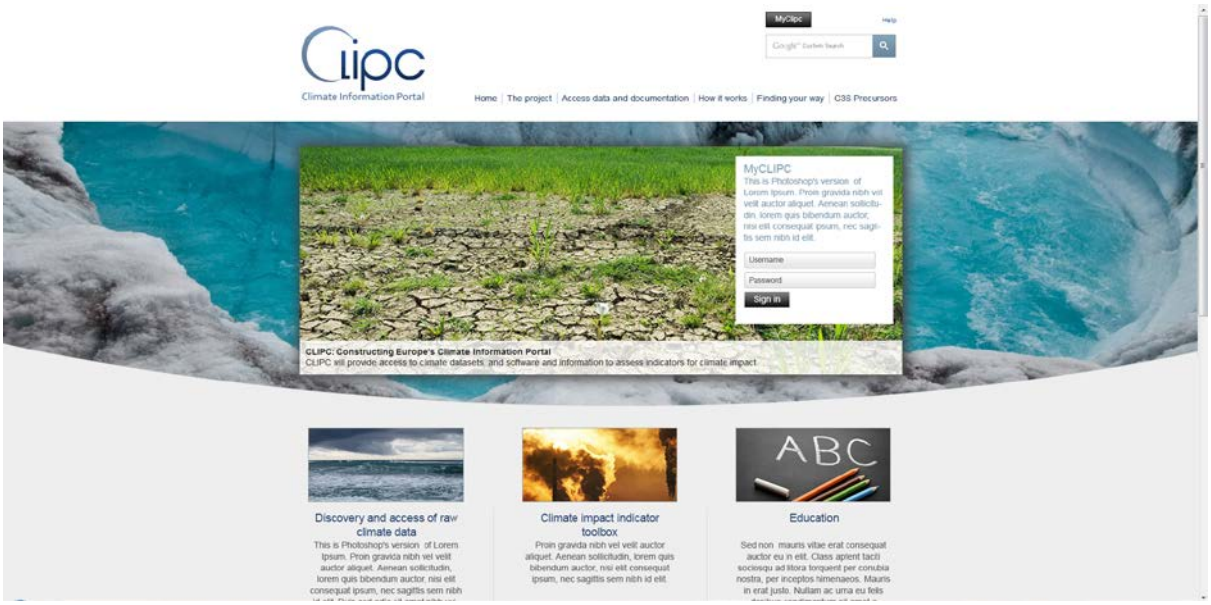

Figure 5: Indicative redesign of the CLIPC portal with main user interfaces centrally

4.2 Menu structure version 2

The first menu structure of the website (version 1) was aimed at supplying basic project information regarding workplan, project background, presenting the consortium, etc. However when CLIPC has developed services, these services should become central. Therefore not

only the homepage needs a new focus and design, also the menu options should fit this new goal.

Menu structure version 2

Climate data access

- Access to climate data
- CLIPC dataset catalogue

Impact indicator toolbox

- Use the toolbox
- How to use the toolbox

Guidance

- Climate impact use cases (Climate4Impact)
- Glossary
- Use of vocabularies
- Documentation of CLIPC models and processing services

Visualising data

- Viewing service (start plain)
- Use cases (described with link to a pre-defined set)

MyCLIPC data processing

- How to use

Project information

- Background
- Consortium
- Public deliverables
- Project resources
- News
- Project meetings

Footer menu (always visible):

- FAQ
- Help
- Disclaimer
- Partner extranet
- Contact

4.3 Content Management System

The main content of the central portal is maintained by a dedicated CLIPC website Content Management System developed by MARIS. Via this system all HTML content, the news items, images, menu, etc. can be managed.

Technically this system is developed in PHP and Javascript/AJAX in combination with a SQLServer database.

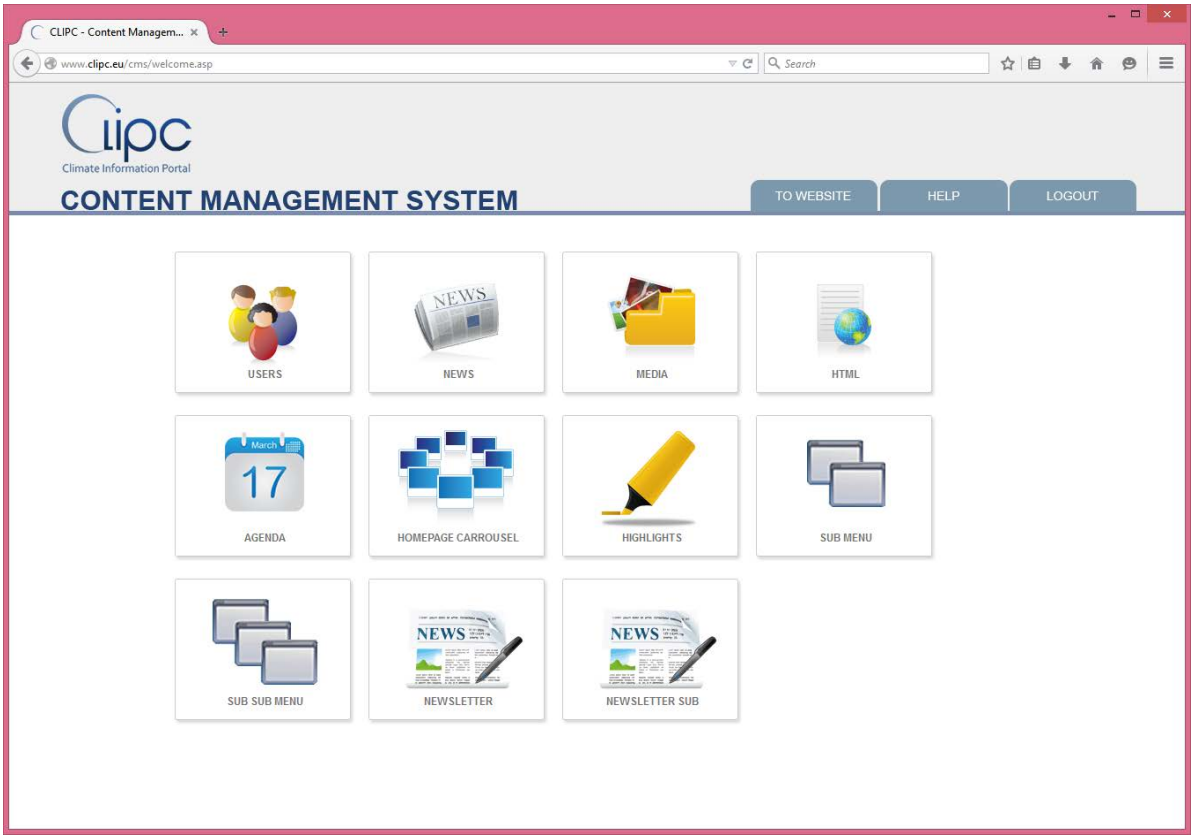

Figure 6: Illustration CLIPC CMS

4.4 Portal releases

The new design will be completed by MARIS after consultation with WP2 (user questionnaire results), and the consortium (or at least the work package leaders). After that it will be used as template for the new portal.

The following figure shows the portal release that are planned (the lower half of the figure indicates the timescales for interactions with users).

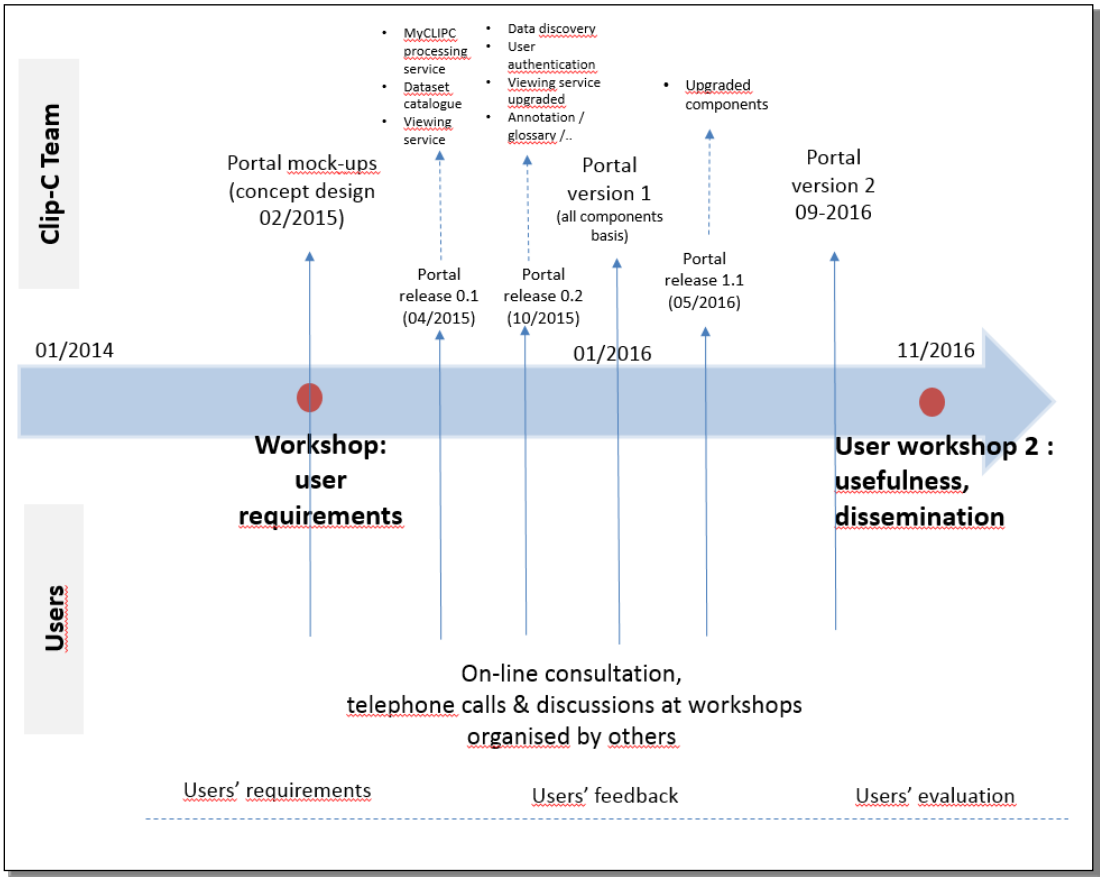

Figure 7: Planned portal releases and user consultation

5. Climate data access

CLIPC will facilitate the users in two ways to discover data and data products that are important for retrieval of climate data and information:

- Via a structured data search on selected infrastructures (chapter 5.1)
- Via the CLIPC data and data product catalogue (chapter 5.2)

The content of the two services is different and they are aimed at different users. The CLIPC data product catalogue is primarily aimed at providing validated datasets in an easy way to impact researchers and boundary workers, while the "raw" data search aims mainly at the climate scientist (although he/she will also be interested in the content of the catalogue). The CLIPC portal will guide the right users to the right service.

5.1 Climate data discovery

5.1.1 Concept

Climate data consists of satellite, in situ, re-analysis and climate model data. The CLIPC data discovery service aims to provide the users of 'raw' climate data (climate scientists mostly, but model data also for impact researchers) with an interface to search via metadata(!) for data from various connected data infrastructures.

Figure 8 shows a diagram of the discovery service to achieve harmonised search actions triggered by the user (bottom) to the diverse data resources (top). Discovery of data will be facilitated through a faceted search interface that triggers a search harmonisation layer that transforms the search actions to a request that fits the infrastructure (both technically as well as semantic).

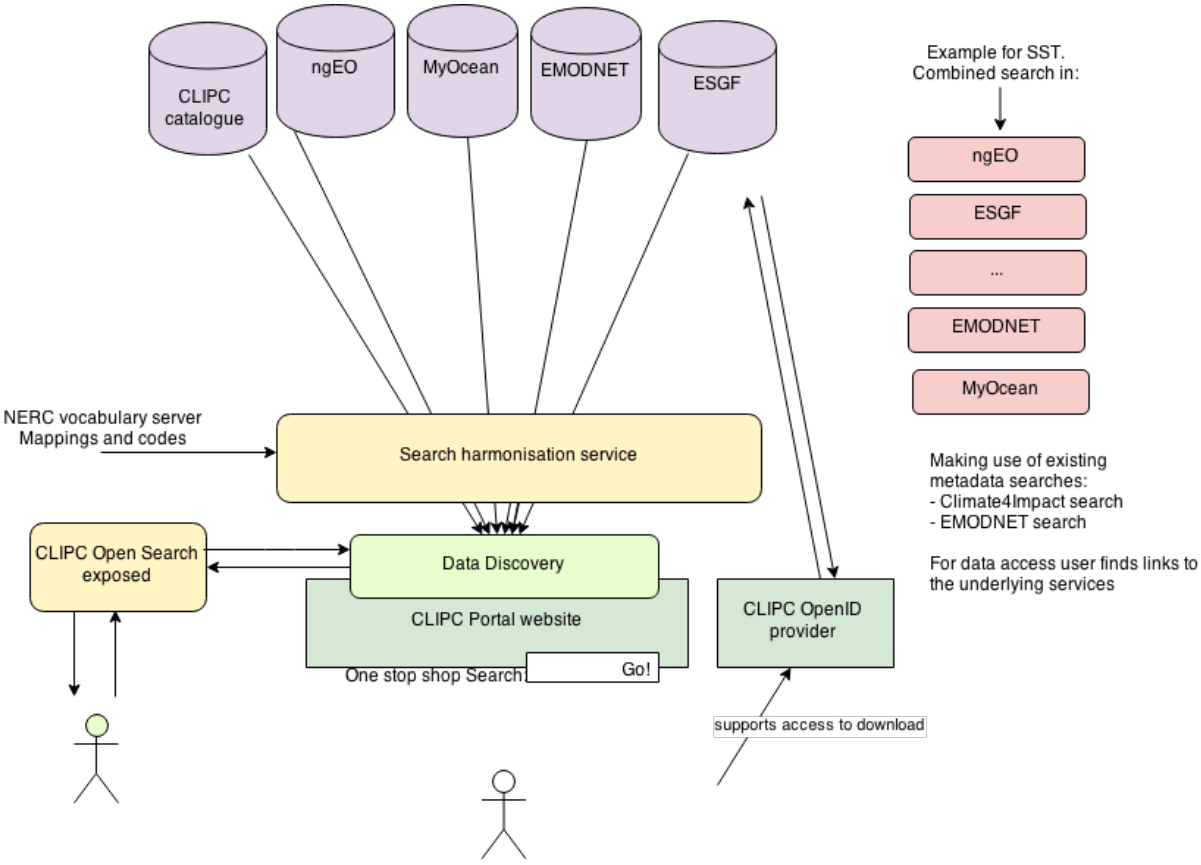

Figure 8: concept of climate data discovery service

5.1.2 Implementation

The implementation process involves the following steps:

- The search API's (or URL's of services) of all infrastructures will be registered in a CLIPC background catalogue.
- The metadata model of each infrastructure is analysed and common search fields are extracted. After first research the expected common fields are:
    - Geospatial (geographical box)
    - Parameter – via code/index/term list (Base is CF standard, or SeaDataNet standard)
    - Date from – to
    - Free search term (on all titles and codes – as far as infrastructure allows).

  The semantic model per infrastructure is different, especially for the parameters but the NERC Vocabulary Service (full specification in MS 10 document) will support in the required mappings.

- The search query on the portal is translated to the required API (or webservice) call per infrastructure. Where needed the NVS provides parameter mappings that are used on the fly.
- The NERC vocabulary service supports the "query harmonisation process" by supporting a mapping of the various controlled code lists into one (the one from NVS). Plus, the NVS provides for parameters a hierarchy of vocabularies to support discovery from discipline level to the detailed parameter term. The metadata layer contains minimum metadata per resource, plus links to the download of files or download system of the resource. More details can be found in specific MS10 document.
- The response per infrastructure is not combined in one list but presented in numbers per infrastructure. This prevents that a service with millions of datasets takes too much exposure in the results, and sets from other infrastructures cannot easily be found. When clicking the datasets from a certain infrastructure the user is directed to the specific portal where he will find the same results, with more details, and can order and download the data via that portal. (see chapter 5.1.3 for examples)
- Additional to the dedicated CLIPC search interface also a CLIPC OpenSearch interface will be developed. The CLIPC OpenSearch entry points will be accessible by overarching systems that use OpenSearch (www.opensearch.org) as standard.
- Security is an issue when moving to data download. CLIPC will reuse existing OpenID providers (Google, Linkedin, Twitter, EUDAT…) to allow user identification, and assist the CLIPC users on the CLIPC portal, and secure specific datasets. However it will not focus on a Single Sign On experience for the users. When the user finds datasets in the discovery services, he/she has to login according to the policy of that infrastructure.

5.1.3 Design examples

Via a search interface with options for a search by parameter, dates from-to and geographical area, users find a range of datasets that fulfils their requirements. The interface allows to "drill-down" step by step.

See design in figure 9 as an example of the portal user interface.

Figure 9: Data search interface (start)

The left column contains the query filters the user can select. The column on the right contains the infrastructures that are queried and the amount of datasets that match the query. As the user sets more filters the amount of records decreases (see figure 10).

Specific implementation details:

- Free text search possible on parameters, aim to use auto-complete by using the P02 (parameter category) or CF parameter list of terms used in the resources
- Unfold a tree according to the parameter hierarchy the NERC vocabulary server allows: P08 > P03 > P01/P02 (see http://seadatanet.maris2.nl/v_bodc_vocab_v2/vocab_relations.asp?lib=P08)
- Search field “Data type” will be investigated, which would indicate a search for: Model output, gridded / points timeseries observation, etc.
- More infrastructures could be added over time.
- Not necessary to log in for the Search interface.

Figure 10: Data search interface - query set

When the user selects “go to data” (icon on the right), the user is directed to the webpage of the infrastructure with the same pre-selection already applied.

5.1.4 Partners involved and planning

The development of this interface involves (at least) the following partners:

Partner	Actions	Timing
MARIS	Analysis datamodels infrastructures (BODC /KNMI/STFC support)	March – June 2015
MARIS	Development of prototype search interface	June – October 2015
BODC	Implement extended mapping of parameters (STFC support)	June 2015 – May 2016
MARIS	Development of final search interface	October – May 2016

5.2 CLIPC dataset catalogue

5.2.1 Concept

The CLIPC dataset and data product catalogue is the key tool to provide information about the CLIPC datasets. It will provide the users metadata of the datasets, and via the metadata access to the datasets that are validated and processed in the project in order to create the climate impact indicators. The indicators themselves will be described as well, plus reference to download. Summarising the catalogue will contain metadata of:

- The bias corrected datasets from within CLIPC (WP6)
- Processed /calculated datasets
- Tier 1/2/3 climate impact indicators as created, used and collected (lots of datasets from outside) in the project.
- Other climate datasets made available through CLIPC services

The catalogue has three main functions:

- Users can search the catalogue, view metadata details and get links to downloads and visualize the datasets.
- Datasets found in the catalogue can also be selected to be used in the MyClipc processing toolkit (see chapter 8) where the user can start processing datasets.
- Make the metadata accessible by web crawlers so datasets can also be found using standard search engines (like Google).

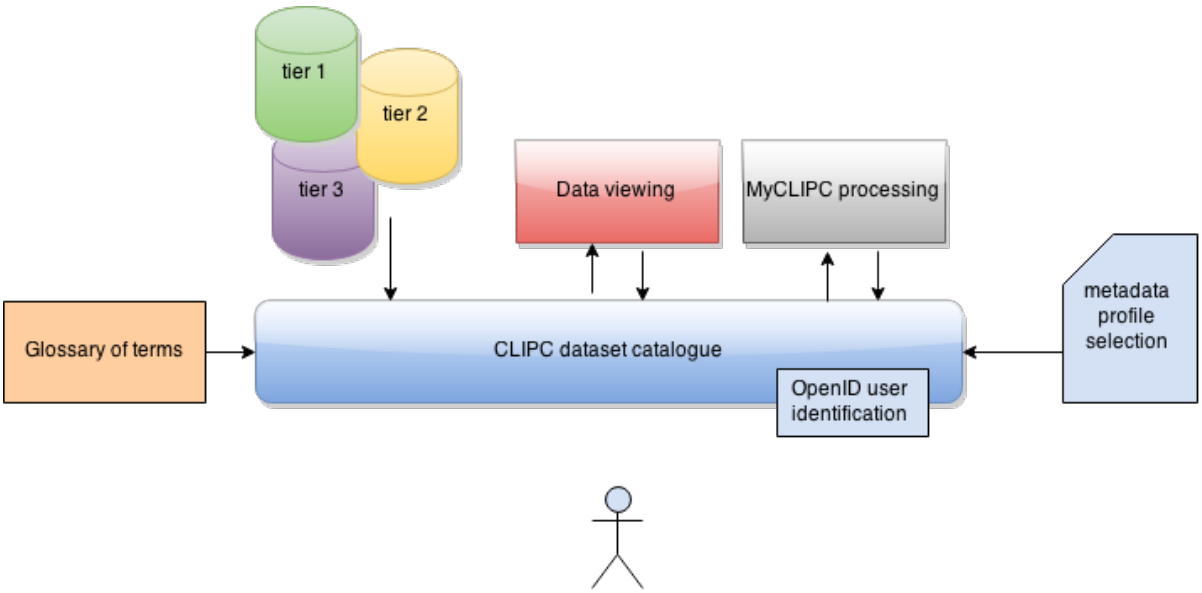

Figure 11: Data product catalogue overview

5.2.2 Implementation

The catalogue will be built up with ISO19115 / ISO19139 metadata, the correct profile has to be selected based on current experiences. It will be built up with a back-end (management tool/editor) and a front-end (user interface). For the editor a selection will be made out of existing solutions. GeoNetwork software might be an option with an advantage of having also front end software and being able to harvest and be harvested in one package. But other options are also possible, e.g. using a specific editor fitting the profile selected which will have XML records as output that are harvested by the CLIPC GeoNetwork application.

The user interface will allow the user to search, browse and view metadata. The catalogue content will initially be small and mainly consist of the datasets related to the first storyline, but will be expanded later on when more climate impact indicators are created. All datasets will be available for download in the catalogue and when possible can be visualised by WMS/WCS services via the CLIPC viewer (see chapter 6).

Regarding the second function of the catalogue: The datasets can be put in a basket and be processed in the toolkit (based on Climate4Impact developments). The results can be used again by the user in the viewing services that are available in the portal.

Browsing the catalogue is open for every user, but when the user want to select a dataset for processing and add it to his ‘basket’ the user has to login first. (More on user identification see chapter 7)

5.2.3 Design examples

The design of the catalogue user interface allows the users to filter in the left column (drill-down method again), and browses the results in the main part of the screen. A small map shows the geographic coverage of the datasets in the selection.
The datasets in the result list can be downloaded, viewed in the CLIPC viewer (chapter 6) and selected to be processed in the MyCLIPC processing service (chapter 8).

Figure 12 shows the design of the search and result list screen of the catalogue

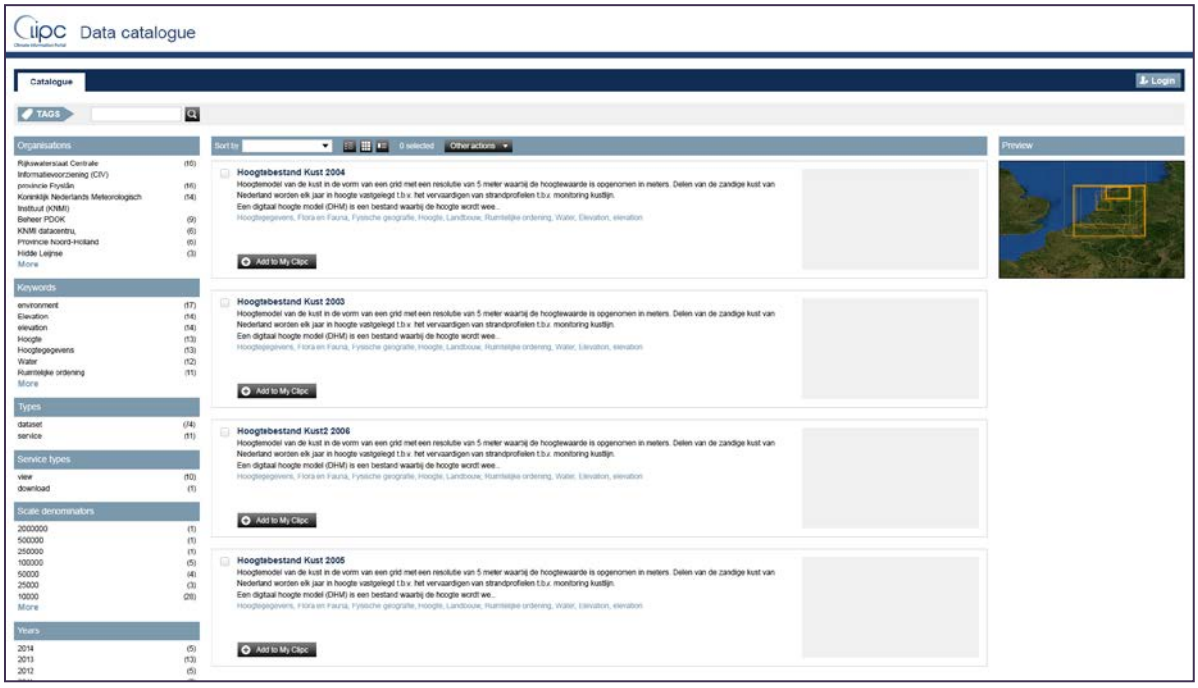

Figure 12: Design of dataset catalogue search and result screen

5.2.4 Partners involved and planning

Partner	Actions	Timing
KNMI + STFC + MARIS	Analyse existing data models and profiles for datasets. Select proper one for CLIPC	March/April – 2015
MARIS	Development version 1 of central CLIPC catalogue and user interface	June – October 2015
MARIS	If necessary develop catalogue editor (preferably online), otherwise use existing editor	April – October 2015
MARIS	Development of version 2 of catalogue interface	October – May 2016

6. Data viewing

In CLIPC datasets will be collected, processed and modelled, after which the user would like to visualize the datasets for a first quick inspection.

6.1 Concept

In order to view datasets the system needs a server side and a client side.

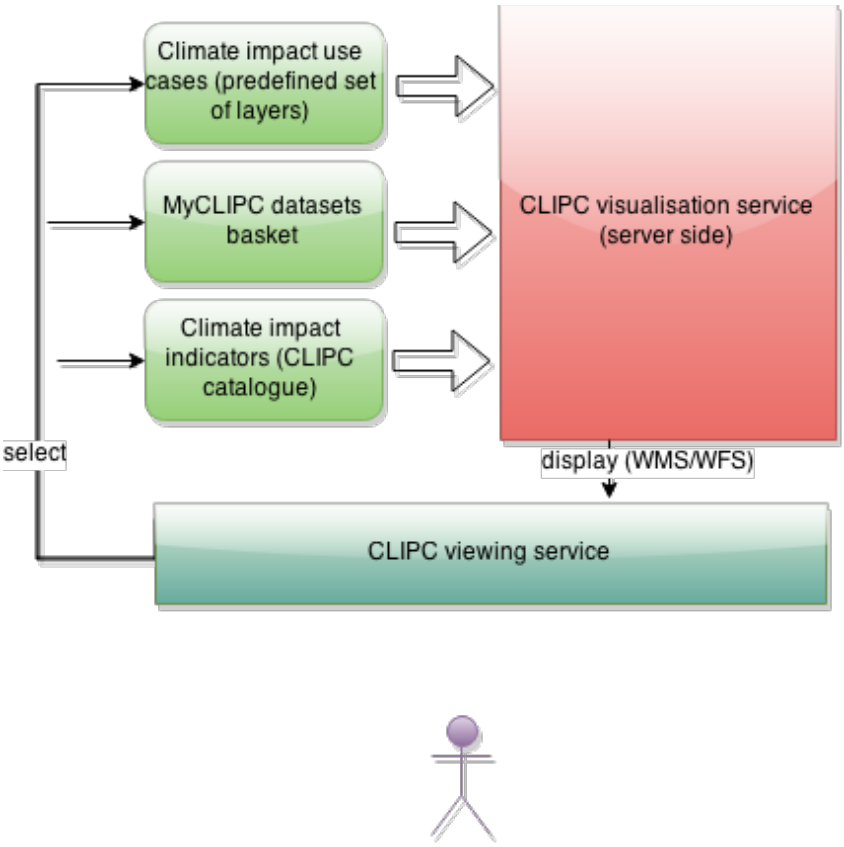

Figure 13: Visualisation (server side) and viewing service (client side)

Server side:
Datasets are stored in the CLIPC infrastructure (KNMI server) to allow visualisation. KNMI will implement the server side component of KNMI's ADAGUC software for this purpose. This software is able to visualise NetCDF files from OpenDAP servers and serves out OGC compliant dataformat (WMS, WCS).
The client triggers the visualisation of selected datasets from various use interfaces: From the datasets catalogue, from the MyCLIPC userbasket, or climate impact use cases. (see later chapters)

Client side:

The CLIPC portal will have its own central viewing service (MARIS development) to allow users to go from browsing the catalogue or basket, to viewing the datasets via available mapping services.
The viewing service will trigger or (receive from) the CLIPC server side visualisation software (ADAGUC, see above) and also allows to combine the datasets, climate impact indicators and (processed) model output, with the maps from external mapservers. Client side developments will make use of specific software that use OGC standards for interoperability.

CLIPC server side visualisation will be described in detail by WP4 documentation. The client side user interfaces will be described in more detail in this chapter.

6.2 Implementation

Users on the CLIPC portal will have access to a viewing service based on OpenLayers 3 (www.openlayers.org) in combination with a JavaScript interface. This way of applying OpenLayers is flexible and can combine maps that the user generates in his/her “user playground” (calculated datasets, impact indicators tier 1 – 3), plus other maps from selected datasources. The datasets in the CLIPC catalogue can also be visualised via the CLIPC viewing service, and combined in the same user session.

OpenLayers 3 (http://openlayers.org/) is fully OGC compliant and can therefore visualise the maps generated from the ADAGUC server side software via WMS or WCS services. MARIS will make use of the latest version of this software as basis of the viewing service, and will implement it via JavaScript and PHP code into a user friendly interface.

6.3 Design examples

Example designs of the CLIPC viewing service interface are given below.

The interface has a fixed toolset on the top left. This provides the basic map functions:

- Zoom in/out
- Pan
- Identify
- Make distance measurement
- Export / print

On the right the user finds the layer panel. A set of preselected layers are presented as basis: Background map (satellite or street map), and other maps that CLIPC wants to show as a minimum. Layer transparency can be set by the slider function. Layers can be moved up/down to influence the way they are plotted on the map.

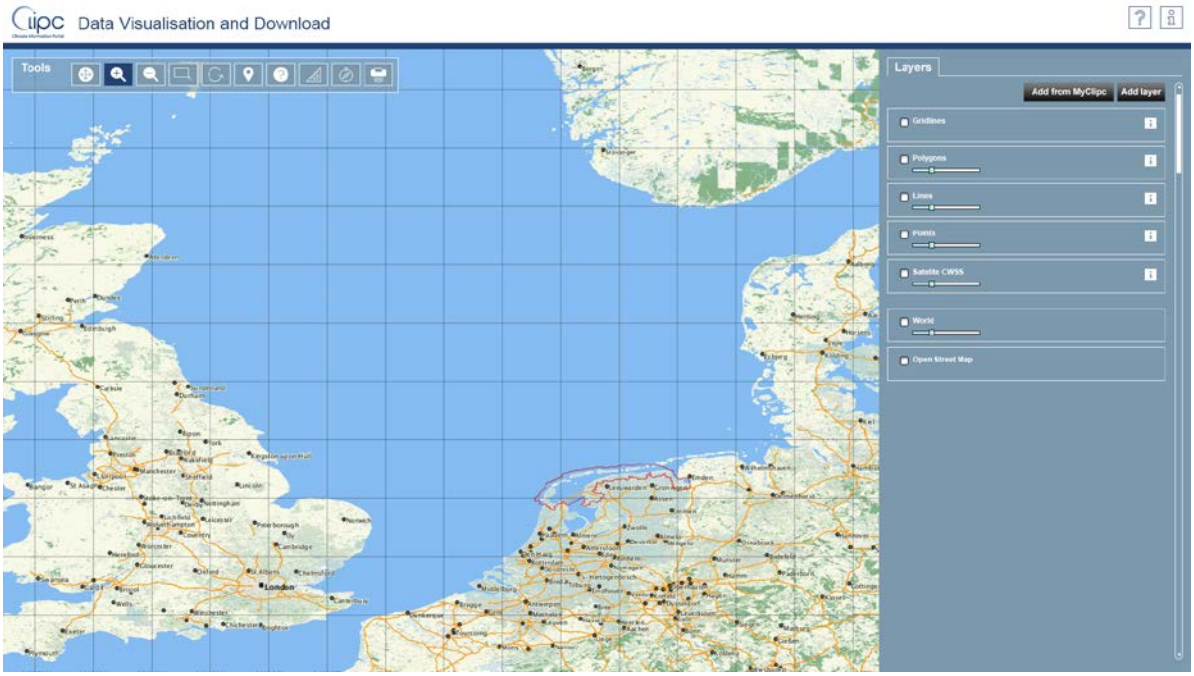

Figure 14: Design CLIPC data viewer

Apart from the fixed layers the user can add layers in two ways to build up his/her own map:

- Add external layer: User can select maps from external WMS and WFS services to include in his map (session). They will be added to the layer menu, just for this user (see figure 15).
- Add layer from CLIPC catalogue: In the same way that the user can add external datasets, he/she can also select datasets from the CLIPC dataset catalogue. If they allow map view they can be added to the users map view.

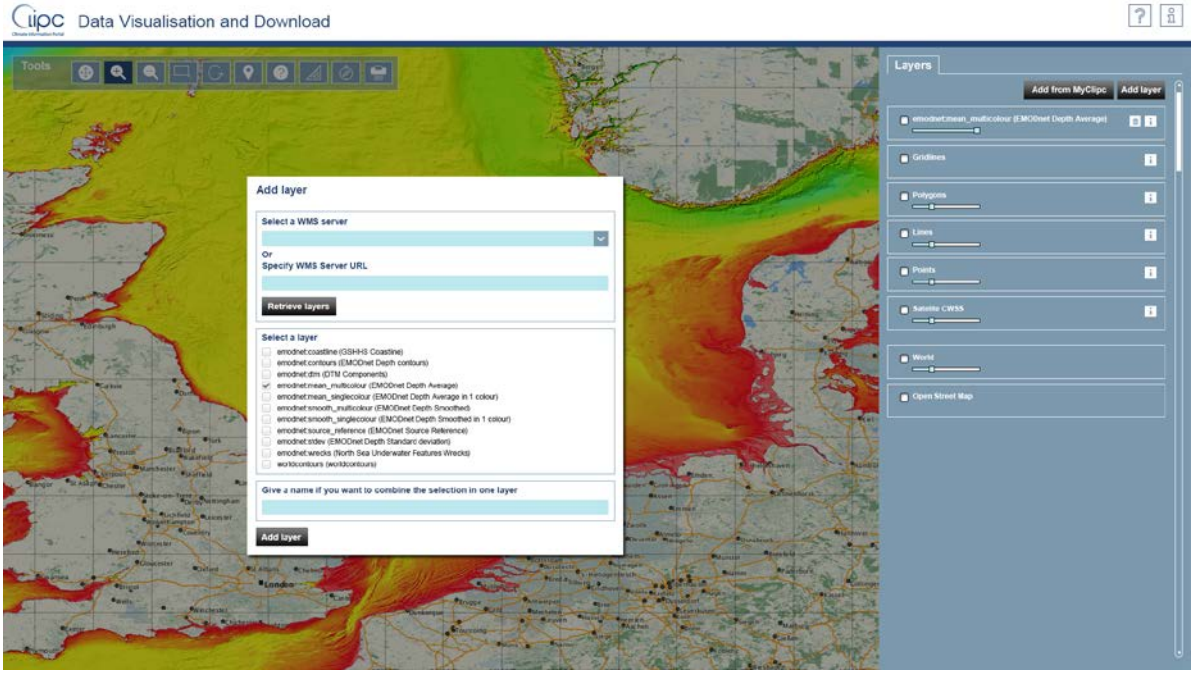

Figure 15: Design data viewer - add layer

The viewing service will also be used in the project to build up a thematic map with preselected datasets on a specific topic. E.g. there will be maps for “flooding risk” in certain area to illustrate which datasets are needed to make a judgement of this, same for other types of climate impact. This action is directed by WP8 who supply the datafiles (impact indicators) by Alterra, datasets selected and provided by PIK, processed and stored by KNMI, and in the end map view implemented by MARIS on the portal.

6.4 Partners involved and planning

Partner	Actions	Timing
KNMI	Set up CLIPC ADAGUC map server plus first datasets for test.	March - April 2015
MARIS	Develop version 1 of CLIPC maps viewer on top of CLIPC map server.	March – April 2015
MARIS	Development of version 2 of map viewer with connection to MyCLIPC and thematic map view as defined by Alterra (WP8)	May 2015 – May 2016

7. User authentication

7.1 Concept

Discovery of data via the CLIPC data discovery service and well as browsing the catalogue (plus part of viewing / visualisation) is open and no user registration or login will be needed. Only when the user is accessing the MyCLIPC processing services, « the user playground » the user needs to login to the CLIPC services. This user identification is required and functional, because the system will have to know who the user is to provide him/her the correct content in the « basket » of datasets, also when returning to the system later.

7.2 Implementation

The CLIPC portal will make use of OpenID-Connect standard for user authentication. OpenID-Connect is the logical follow up on OpenID. OpenID is deprecated and is not supported by major identity providers like Google. OpenID-Connect is an extension on the Oauth2.0 standard and is more widely supported than OpenID.

The first version of the portal will make use of Google as OpenID-Connect provider. Later connection with EUDAT, CHARME, LinkedIn, Twitter, Facebook will be made. Google will be the starting point for the first version of the portal because Google supports OpenID-Connect and many people have a Google account. On the CLIPC portal a user will be able to login using Google. The first time the user has to approve that user's identity (email and unique identifier) is being used by the CLIPC portal. The second time this option is remembered by Google and will not be shown.

Communication between the Climate4Impact portal and the CLIPC portal requires user information to be passed, e.g. identity is required for accessing climate4impact services. The climate4impact portal needs to know which user is doing what operation. Passing identity information, or delegation of credentials, is achieved by passing a signed JWT Bearer token to climate4impact. The JWT Bearer token is part of the OpenID-Connect standard. The JWT token is issued by Google during login at the CLIPC portal and can be passed around the application. The JWT Bearer token consists of a header, a payload and a signed key. The payload contains the user identifier. The last part, the signed key, needs to be used to verify the header and payload of the JWT Bearer token. It is critical that this token is verified at the Climate4Impact portal; user identity information can only be extracted from the token when the token is valid,.

The Climate4Impact portal will be able to verify JWT Bearer tokens, so safe delegation of credentials is ensured. Access to climate4impact services is granted when a valid JWT Bearer token is posted over HTTPS to one of the services. These services can be WMS, WCS, WPS

and basket access. Besides authentication using OpenID Connect, x509 certificates (used by ESGF) can also be used to pass credentials.

Climate4impact is connected to the ESGF network. In order to access ESGF data, a valid ESGF OpenID account is needed. At climate4impact the possibility will be offered to link an existing ESGF account with an OpenID-Connect account. This allows to access ESGF services via Climate4impact to the CLIPC portal, this enables for example processing and visualization on CMIP5 and CORDEX datasets in the CLIPC portal.

CLIPC will make use of services developed by Climate4Impact that already can be accessed by other systems or portals.

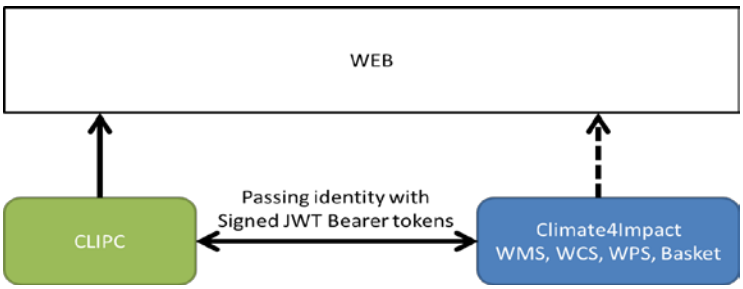

Figure 16: Possible collaboration of CLIPC using climate4impact processing capacity

By using OpenID-Connect, the CLIPC system will have a light user management system, providing as few barriers to the user as possible (only functional). The authorisation will be implemented on WMS, WCS, WPS and file storage requests. MARIS and KNMI will focus on this as early in the development process as possible. Detailed information about OpenID-Connect can be found at https://developers.google.com/identity/protocols/OpenIDConnect

7.3 Design

The design of the login screen is fairly straightforward. The user selects his/her OpenID provider, allows CLIPC to use at least the email address and after that the session is started. The user gets confirmation of login success and finds his/her name at the top right: Logged in <name>.

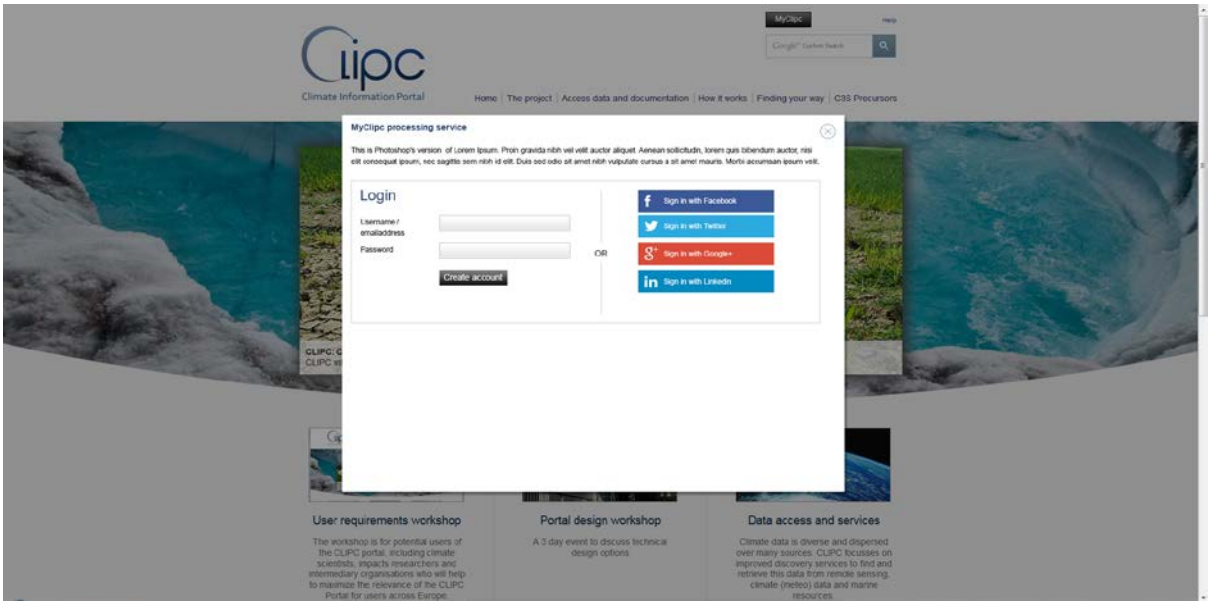

Figure 17: Design CLIPC login screen

7.4 Partners involved and planning

Partner	Actions	Timing
KNMI	Implement OpenID-Connect secured WMS, WFS and WPS services. Develop in close communication with MARIS.	March - December 2015
MARIS	Implement OpenID-Connect authentication on CLIPC portal and send secured requests to KNMI server.	March – December 2015

8. The binding factor: MyCLIPC data processing and visualisation

8.1 Concept

In CLIPC there will be both pre-computed products (indicators) and a processing service. The processing tools will serve both to support science staff to maintaining pre-computed products as well as to provide users with the capability to explore.

The exploration will be in a MyCLIPC "playground" area of the portal and is not to be confused with the pre-computed products that will have more scientific authority.

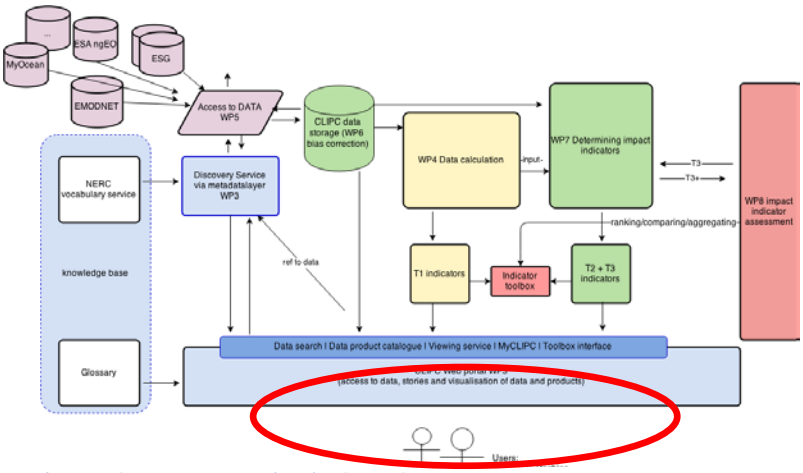

Figure 18: Data processing in CLIPC

Allowing the CLIPC users to explore and process datasets andcalculate climate impact indicators (tier 1, 2 and 3) is an important CLIPC development. CLIPC will provide an environment in which the user can select validated datasets, launch calculations and processing to the datasets, and afterwards visualise the results. CLIPC will not start from scratch but make use of methods and techniques that have investigated and tested under Climate4Impact.

8.2 Implementation

CLIPC will provide (following KNMI developments in Climate4impact) a graphical user interface for processing. See figure on the right.

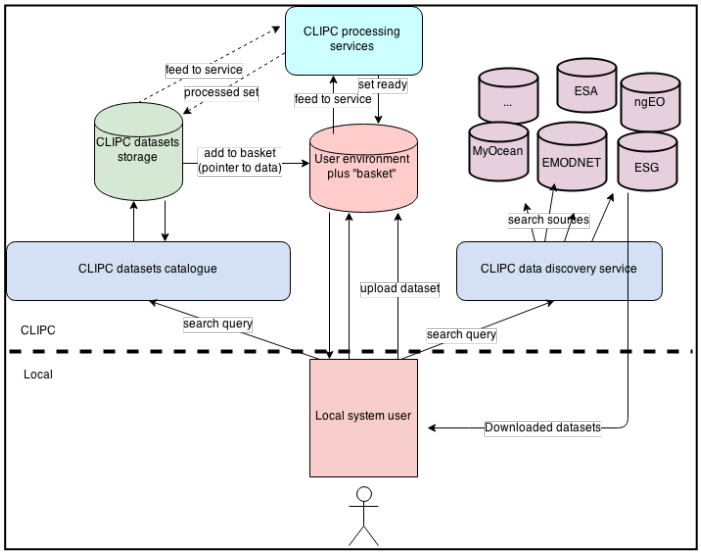

Figure 19: Use of user environment / basket

The user has access to a « basket of data » and can select which files to use. The user can select files from the CLIPC catalogue to his basket, plus upload files downloaded from other infrastructures (e.g. via the CLIPC discovery services) to his own "playground" basket. This is all on the MARIS servers, but data is stored on KNMI servers, close to the

processing services which also run on the KNMI servers.

8.2.1 Access to services and calculated data

The MyCLIPC processing services can be used to do initial generic computations, like Tier 1 indices. The user selects in the portal environment (after login) which datasets from the catalogue he/she wants to use for the calculation from a « basket of data ». Calculated datasets for a specific users will be published/saved in the catalogue of the user, but not made public (only if the user allows this). From his private catalogue the user can visualise his datasets using the CLIPC portal viewing service, as well as create new calculations, visually compare etc.
On the other hand official CLIPC datasets, calculated datasets by CLIPC researchers to create climate impact indicators (e.g. the datasets related to the storyline) will be publicly available in the CLIPC catalogue to every user.

More specific application example: Tier1 and tier 2 calculation of climate indices can be achieved using the processing functionalities of the CLIPC portal. Climate4impact has developed some basic calculation mechanisms which will be used as basis. It is for example possible to calculate the number of Tropical days on any CMIP5 and CORDEX dataset containing tasmax. The processing is based on the OGC Web Processing Service (WPS) standard and uses the PyWPS framework. The service is accessible as a webservice for other portals. It is secured using the same security mechanism as used for ESGF (OpenID or x509). The basis for the calculation of valid climate impact indicators has to be validated and bias corrected datasets. Therefore only the bias corrected (and other selected) datasets as created/selected under WP6 will be included in a catalogue of source data.

8.2.2 Using PyWPS for processing

PyWPS is an open source implementation of the OGC Web Processing framework. Processes are modular within PyWPS, they are simple individual scripts. Any process or algorithm that can be described with a set of input and output parameters can be used. PyWPS does not care about what the process is doing, as long as the process periodically gives status information. This can be used to display progress information to the user. The actual process remains a black box for PyWPS. Any tool, like NCL, CDO, R or an executable can be used. In Climate4Impact this mechanism is used to provide indices calculation based on ICCLIM on datasets accessible over OpenDAP, e.g. datasets in ESGF.
PyWPS can also be used for calculation framework for the Tier 2 and Tier 3 indices. It is to be discussed if Tier 3 calculations are to be made interactive or adjustable for the CLIPC researchers. This will determine if PyWPS implementation work for T3 is justified. Current viewpoint is that only Tier 1 and Tier 2 indices will be processed via the PyWPS framework

and Tier 3 indices will be not be interactively calculated but « Precooked » by CLIPC. Interactivity could be provided in the final visualisation, created by CLIPC.

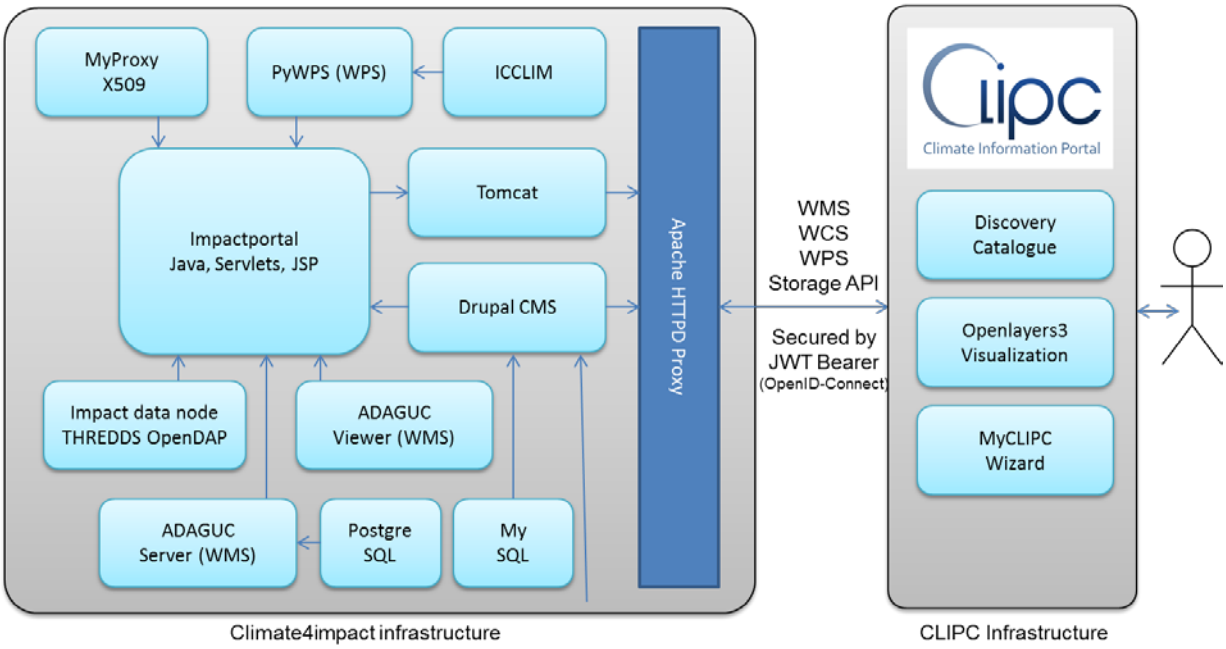

Figure 20: CLIPC using http services from Climate4Impact infrastructure via proxy

8.2.3 Summarising development approach

The development approach is as follows:

- Processing services as well as data storage will be done on the Climate4Impact servers (running at KNMI). The ADAGUC web map service will generate visualizations of the processed datasets.
- The portal catalogue, user basket and visualisation service will run on MARIS servers as part of the CLIPC portal.
- The development will start small (AGILE approach), only focussing on the selected datasets and processed datasets (for tier 1, 2 and 3), plus the processing services related to the storyline. All services will be developed and tested for this test case, and afterwards extended for the other storylines and with the ability for users to upload their own datasets.

8.3 Design examples

Some designs have been made to illustrate and support the MyCLIPC developments.

First the user has to login (see previous section), after that the user has access to his/her MyCLIPC user basket. Figure 20 shows the content of the basket. The dataset can be selected

from e.g. the CLIPC catalogue or downloaded first via the data search services and after that uploaded directly during the users browsing sessions.

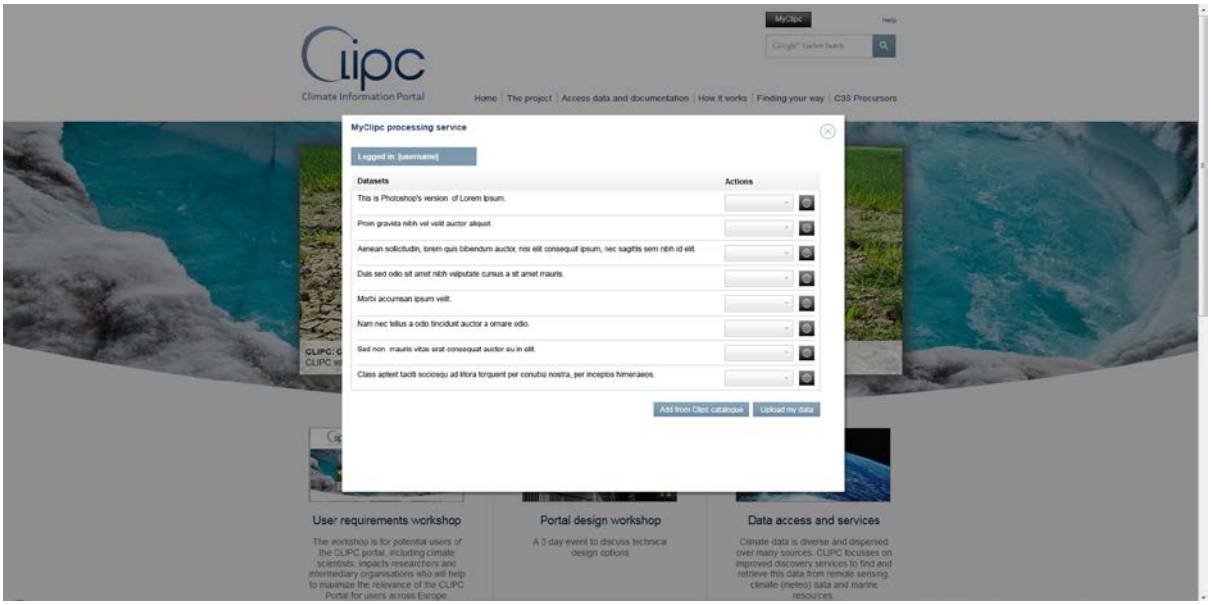

Figure 21: Design MyCLIPC user dataset basket

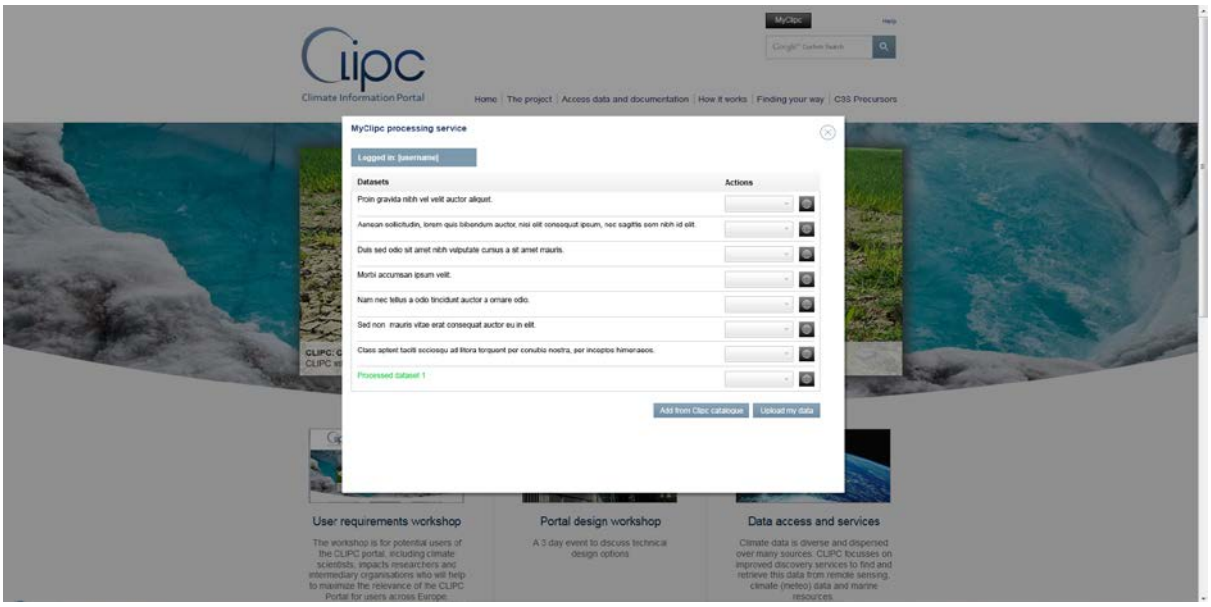

Figure 22: Design MyCLIPC user dataset basket - processed dataset added

Figure 21 illustrates that a dataset is added to the basket after being processed. Any dataset in the basket can be visualised in the CLIPC viewer (all in the same session of the user). Pressing the “visualise” button, automatically adds the dataset as a layer in the viewer and opens the viewer. See figure 22.

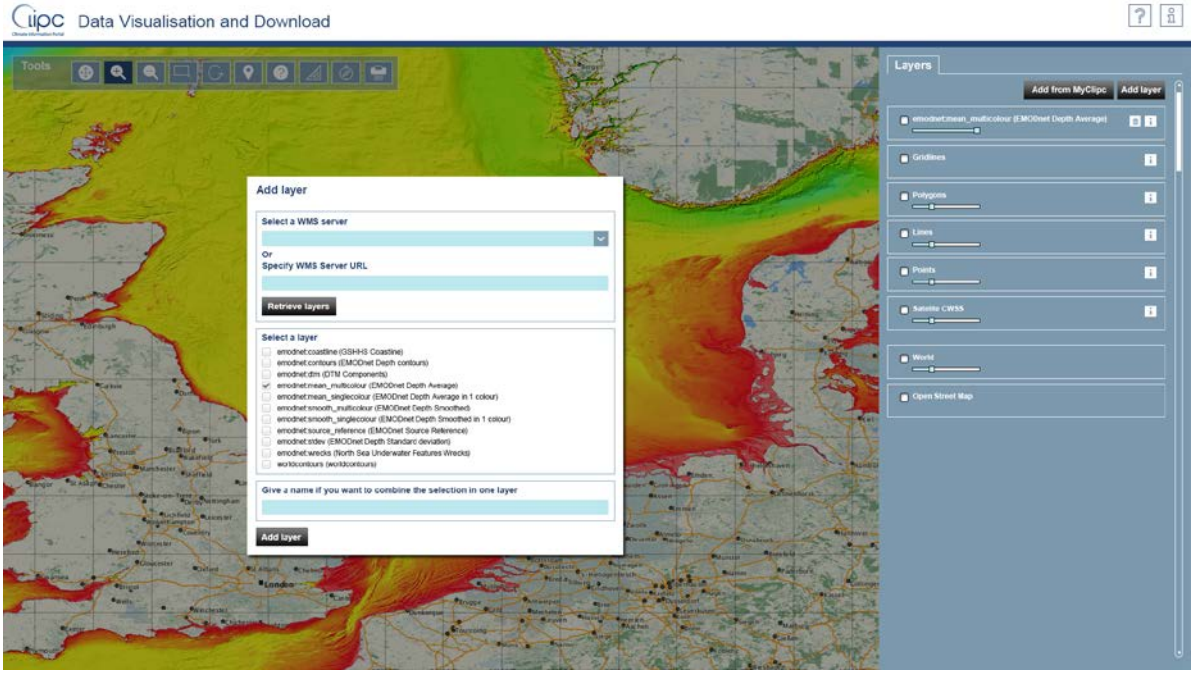

Figure 23: Design MyCLIPC user dataset - Dataset added to viewer - user session

8.4 Partners involved and planning

Partner	Actions	Timing
KNMI	Server side development of processing services, to be triggered by CLIPC client.	Version 1: March – Sept 2015 Version 2: Oct 2015 – Jan 2016
MARIS	Client side development MyCLIPC user basket.	Version 1: March – Sept 2015 Version 2: Oct 2015 – Jan 2016

9. Guidance and education

9.1 Concept

The services that together form the “knowledge base” are an important core component of the CLIPC services. They provide the user guidance while working with the services and education/explanation about the terminology in use.

To structure the content the following categories of knowledge base service were identified during user meetings and workshops:

1. CLIPC metadata catalogue (of scientific information of datasets)
2. Commentary information
3. Technical documentation / guidance
4. Glossary of terminology
5. Literature

The CLIPC metadata catalogue has already been explained. The others will be detailed in the following chapters.

9.2 Commentary information

9.2.1 Implementation commentary information

The commentary information is supplied by two components :

- Frequently Asked Questions (FAQ) section: This will be implemented as a set of webpages managed and updated via the CLIPC CMS. In this section issues will be explained which users often run into when using data or applications on the CLIPC website. The FAQ section will also provide general help on using the CLIPC portal and tools. It will contain a web form to contact user support for reporting questions, errors, remarks, etc.
- Annotation to URL’s to data (or better via “Handle” = group of URL’s) using CHARMe methodology. CHARMe allows users to view or create annotations that describe how climate data has been used and what has been learned. This information includes citations, results of assessments, provenance, external events and supplementary dataset quality information. The CHARMe system collects and share annotations associated with climate datasets. This information is called *commentary metadata*. Users are able to add or view commentary metadata, complementing existing information from the data providers.
    - Implementation on the portal will follow the rules of http://charme.org.uk/

- During retrieval of a dataset the CLIPC server will query the CHARMe system if annotation is available.

9.2.2 Design examples

FAQ section will simply be webpages managed from the CMS.

The CHARME commentary metadata will be implemented using specific icons linked to pop-ups showing the URLs linked to the comments. As an example, commentary metadata as implemented on ECA&D[1]is shown in Figure 23

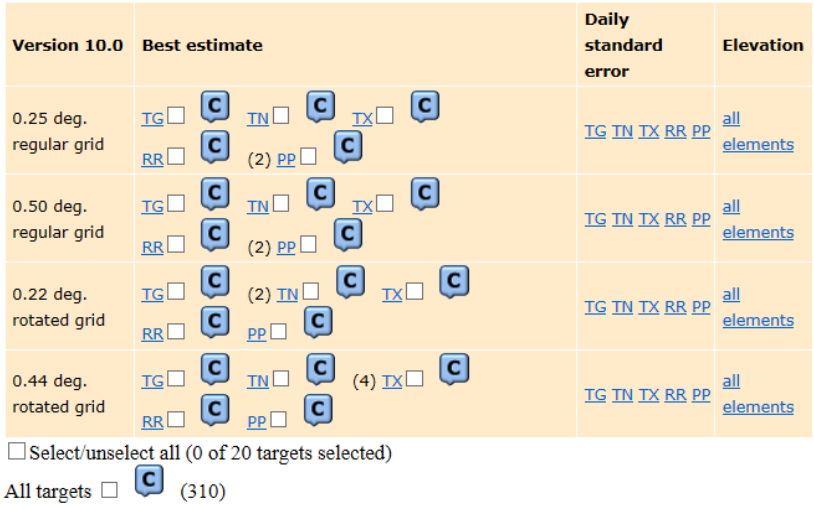

Figure 24 "Example of commentary metadata on EOBS gridded dataset in ECA&D"

The "C" image indicates the availability of commentary metadata on specific parameters in the EOBS dataset. Clicking on the 'C' behind TG will show a pop-up containing a list of URLS to documentation, as shown in Figure 24.

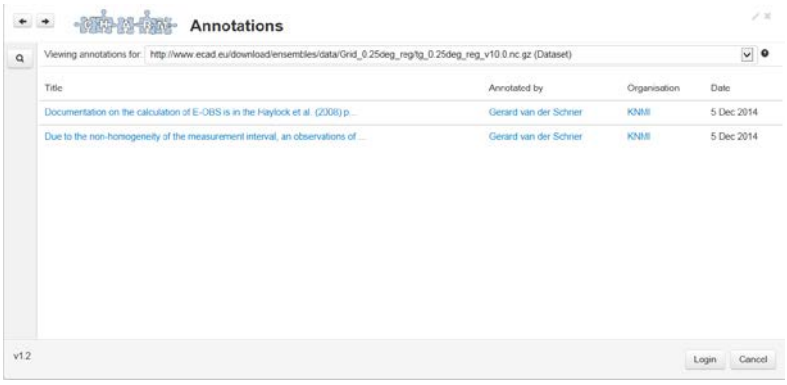

Figure 25 'Pop up containing comment URLs'

[1] http://www.ecad.eu/download/ensembles/download-charme.php

9.3 Technical documentation

9.3.1 Implementation

CLIPC will provide as much technical documentation as possible, explanation of terminology used, and links to existing technical documentation. This part is covered by several services:

- Providing definitions and documentation of the calculation and processing services implemented in the portal to generate the Tier 1, Tier 2 and Tier3 data products.
- Providing definitions of the search terms in the data discovery service
- Use references to CCI documentation, which is loosely structured. Planned to import definitions from BODC vocabularies via SKOS.
- ESDOC and CMOR climate model documentation

Links to definitions and documentation about the models and calculations in the processing services will be done via a specific CLIPC CMS module in which webpages will be created with documentation about algorithms and documentation of services. These pages will be linked to from the interface.

Definitions of search terms (parameter disciplines, category and group) will be provided from the NERC vocabulary server. This will be applied in the same way as the Glossary (next chapter 9.4) showing a link to the term that shows the definition on “mouse-over”.

For the data discovery service as well as technical documentation the integration and extension of the NERC vocabulary services is a key development within the CLIPC project. More information on the background and use is provided in the MS10 document.

9.3.2 Design examples

Via a script in the HTML pages upon display the content will be screened for terms that have a definition in Glossary or Vocabulary Services. When a definition is available the user will find the term underlined, and a mouse-over will show the definition. See figure below.

Figure 26: Design mouse-over for definition of term (in data search service in this case)

9.4 Glossary

9.4.1 Implementation

Next to the technical documentation also “softer” documentation about terminology will be available via several glossaries.

- **Glossary of terminology**
  The glossary created by EUPORIAS is validated and used in the IS-ENES website. It will be integrated in the CLIPC web portal and extended. EUPORIAS makes use of a Google doc as source for the Glossary and can be shared to other websites (Within the EUPORIAS project a Drupal module was developed for this). Terms in the Glossary will get a “href-link” in the HTML webpage.
  CLIPC will load the EUPORIAS Glossary in a database table, extend the glossary and develop an extra glossary for the Climate Impact indicator terminology reusing the IPCC defined glossaries
  https://www.ipcc.ch/publications_and_data/publications_and_data_glossary.shtml.
  The different glossaries can be used at the same time.

Integration of the Glossary in the CLIPC portal will be done via the RDF-a technique . When printing the requested page on the client, terms in the website will be underlined. The terms in the HTML code will be marked in a specific way which is very beneficial for Google ranking. Example page how to apply RDF-a in HTML for links to vocabs:
http://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41479c42-4dfb-4da9-be97-4c532ce13922/.
There are plenty of plugins (e.g. RDF Detective) which can extract the RDF from the page.

- **Use case glossary**
  The Climate4Impact website contains a lot of useful documentation pages regarding use cases for climate model data:
  http://climate4impact.eu/impactportal/documentation/guidanceandusecases.jsp. This information will be reused as-is in the CLIPC portal, by using a scraping technology (dynamic) having only the Climate4Impact website as main source.

9.4.2 Design examples

The following designs support the implementation process.

Glossary of terminology

The glossary will be built up in database tables with fields “term” and “definition”. The content not originating from EUPORIAS will be managed via the CLIPC Content Management System.
The glossary itself will get its own webpage under the website to view it (see figure 26).

Figure 27: Glossary listing page - alphabetically

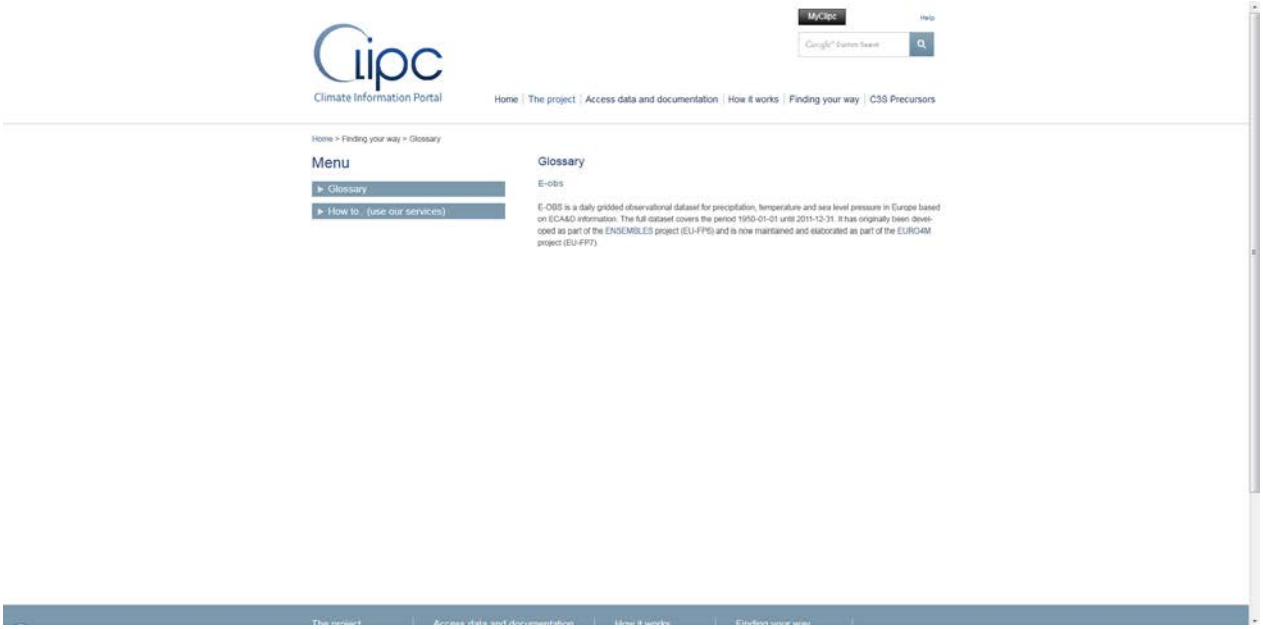

Figure 28: Glossary - definition of term

When the user clicks the term in the list, the full definition is displayed, see figure 27.

However, the most important use of the glossary is in all webpages of the portal. This will work in an identical way to the use of the NERC vocabulary terms: Via a script in the HTML pages upon display the content will be screened for terms that have a definition in Glossary or Vocabulary Services. When a definition is available the user finds an underlining for the term and mouse-over shows the definition. See figure below.

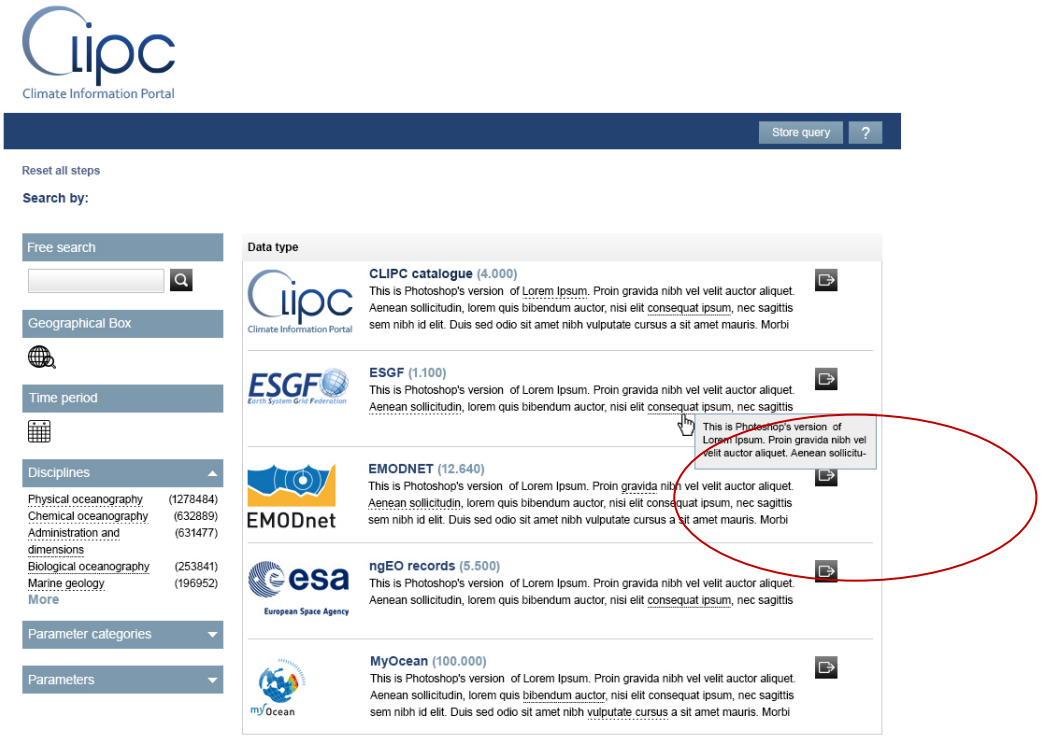

Figure 29: Design mouse-over for definition of term (in data search service in this case)

Use case glossary

CLIPC will re-use the documentation of the thematic climate impact use cases as developed on the Climate4Impact website. The content will be scraped and published under the CLIPC website – section "Education".

Figure 30: Overview of climate impact thematic use cases

Figure 31: Details of thematic climate impact use case

9.5 Literature database

9.5.1 Implementation

For literature references a database will be set up, managed via the CLIPC CMS. As a first step the CLIPC deliverables and milestone documents will be entered into a skeleton system. For each document we would need to enter, e.g., document type, title, abstract, authors, date, URL, keywords, possibly following the Dublin Core metadata format.
It will also be possible to add items by providing a cross-ref DOI and keywords. Via the DOI a script will harvest title etc, and abstract would then be optional. The added value would be linking keywords to documents so that they become part of the « linked data » infrastructure.

9.6 Partners involved and planning

Partner	Actions	Timing
KNMI	Commentary information. KNMI test out CHARMe annotation developments. Document for usage in CLIPC. KNMI set up hosting server	March – Oct 2015
MARIS	Apply annotations in portal via script	May – December 2015
MARIS	Technical documentation implementation – ref to definitions NVS in data search service	May – December 2015

MARIS	Glossary of terminology – Harvest EUPORIAS glossary into CLIPC CMS	April – June 2015
SMHI/PIK/KNMI	Expand existing Glossary of terminology	June 2015 – June 2016
MARIS	Re-use existing use case Glossary in Climate4impact in CLIPC portal v2	May – July 2015
STFC/MARIS	Develop literature database in CLIPC v2 CMS, plus develop user interface	Oct 2015 – April 2016
SMHI/STFC /KNMI	Specify literature database format and content demands	March – Dec 2015

Annex 1: List of abbreviations and acronyms

ADAGUC	- Atmospheric Data Access for the Geospatial User Community (OpenSource OGC WMS/WCS implementation) – http://adaguc.knmi.nl/
CMIP5	- Coupled Model Intercomparison Project - Phase 5
CORDEX	- Coordinated Regional Downscaling ExperimentCCII
CSW	- Catalogue Service for the Web (ISO standard for access to catalogue data)
EMODNET	- European Marine Observation Data NETwork (can be seen as data access infrastructure, mostly built upon SeaDataNet standards)
ESGF	- Earth System Grid Federation - http://esgf.org/
EUDAT	- Research Data Services, Expertise & Technology Solutions - http://www.eudat.eu/
GeoServer	- Open Source map server software
MyO	– MyOcean, FP7 project for marine/ocean modelling.
NcWMS	- Software able to create WMS (graphical map) from NetCDF files
ngEO	– Next Generation Earth Observation ESA
OGC	- Open Geospatial Consortium
PyWPS	- Open source python implementation of the OGC Web Processing Service standard.
OAI-PMH	- Catalogue access standard, forthcoming from library domain.
OpenID	- Open source and widely used user management service
OpenLayers	- Open source viewing service code set, see Openlayers.org
SeaDataNet	- Pan European marine data infrastructure focussing on metadata and data standards for marine data.
UMN Mapserver	– Map server software developed by Minnesota University
WCS	- OGC Web Coverage Service
WMS	- OGC Web Mapping Service
X509	- Public key infrastructure (PKI), security mechanism used in ESGF