



Integrated Arctic Observation System

Research and Innovation Action under EC Horizon2020
Grant Agreement no. 727890

Project coordinator:
Nansen Environmental and Remote Sensing Center, Norway

Deliverable 5.2


IAOS Platform and Tools v1

Start date of project:	01 December 2016	Duration:	60 months
Due date of deliverable:	02 December 2018	Actual submission date:	28 November 2018
Lead beneficiary for preparing the deliverable:	TERRADUE		
Other contributing partners	AWI, ARMINES		

Lead author: Hervé Caumont

Reviewer: Pedro Goncalves

Version	DATE	CHANGE RECORDS	LEAD AUTHOR
0.1	16/03/2017	Template and ToC	Hervé Caumont
0.2	01/10/2018	First draft	Hervé Caumont
0.3	21/11/2018	Consolidated draft	Hervé Caumont
1.0	28/11/2018	Version 1.0	Hervé Caumont

Approval	Date:	Sign.
	29/11/2018	 Stein Sandven

DISSEMINATION LEVEL		
PU	Public, fully open	X
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

USED PERSON-MONTHS FOR WRITING THIS DELIVERABLE					
No	Beneficiary	PM	No	Beneficiary	PM
1	NERSC	-	24	TDUE	0,2
2	UiB		25	GINR	
3	IMR	-	26	UNEXE	
4	MISU		27	NIVA	
5	AWI	0,1	28	CNRS	
6	IOPAN		29	U Helsinki	
7	DTU	-	30	GFZ	
8	AU		31	ARMINES	0,1
9	GEUS	-	32	IGPAN	
10	FMI	-	33	U SLASKI	
11	UNIS		34	BSC	
12	NORDECO		35	DNV GL	
13	SMHI		36	RIHMI-WDC	
14	USFD		37	NIERSC	
15	NUIM	-	38	WHOI	
16	IFREMER		39	SIO	
17	MPG		40	UAF	
18	EUROGOOS		41	U Laval	
19	EUROCEAN		42	ONC	
20	UPM		43	NMEFC	
21	UB		44	RADI	
22	UHAM		45	KOPRI	
23	NORUT		46	NIPR	
			47	PRIC	

EXECUTIVE SUMMARY

The “Integrated Arctic Observation System” (INTAROS) is a 5-year project funded by Horizon 2020 under the Blue Growth Programme. The overall objective of INTAROS is to build an efficient integrated Arctic Observation System (iAOS) by extending, improving and unifying existing systems in the different regions of the Arctic. Within INTAROS, WP5 (Data integration and management) is in charge of designing and implementing a first version of the iAOS.

This document defines the **iAOS platform and tools** as part of WP5 activities. The iAOS platform has the objective to integrate multidisciplinary and distributed data repositories and provide a set of tools for data analysis, transformation and visualization. The iAOS will offer access to observations and derived parameters and seamless integration with geo-statistical methods for interpolation of spatiotemporal datasets, including the new observations from WP2-3-4, and store the generated datasets in iAOS enabled repositories.

The platform will facilitate access to multidisciplinary data, scalable allocation of data storage and computer power for integrative data processing and analysis. Selected applications from WP6 aim at demonstrating also the usefulness and functionality of the platform in service development. These applications address the significance of enhanced integration of data from Arctic observing systems, covering a range of remote sensing and in-situ platforms in geographically different locations covered by iAOS.

The iAOS platform and tools presented in this document cover the system and user essential functionality for data discovery view, download, transformation, and processing workflows, which efficiently address the challenges of Big Data (Volume, Velocity, Variety, Veracity, Value). With a capacity for Cloud deployment of new processing services for multi-source data, iAOS also benefits from a standard approach for defining processing chains (workflows) and assignment of DOIs for the newly generated datasets. The iAOS architecture is supported by a cloud infrastructure to facilitate the management of the applications’ lifecycle (integration, validation and deployment in production).

Table of Contents

Acronyms and abbreviations	6
1. Introduction	7
1.1. Using Terradue Cloud Platform	7
1.2. Accessing federated remote repositories	7
2. Processing Service Integration Guide	8
2.1. How to join the PaaS environment	9
2.1.1. My account (as a PaaS user)	12
2.1.2. Partner Support	14
2.1.3. User Forum	14
2.1.4. User Helpdesk	15
2.2. Using the Cloud Platform's core services	15
2.2.1. Data Agency	15
2.2.2. Cloud Dashboard	16
2.2.3. Virtual Private Network	16
2.2.4. Other Cloud Platform services	16
2.3. Using the Cloud Platform's Ellip solutions	16
2.3.1. Ellip Notebooks	17
2.3.2. Ellip Workflows	18
2.3.3. Ellip Launchpads	20
2.3.4. Ellip Infohubs	20
2.5. Cloud Platform SDK Handbook	24
3. Catalog client tools	24
3.1. Terradue Catalog client - OpenSearch for EO data	24
3.3. OpenDAP/THREDDS Catalog client	25
3.2. Pangaea Catalog client	26

Acronyms and abbreviations

API	Application Programming Interface
CPU	Central Processing Unit
DevOps	Development to Operations
DIAS	Data and Information Access Service
EGI	European Grid Infrastructure
EO	Earth Observation
FRAM	FRontiers in Arctic marine Monitoring project
HTTP	HyperText Transfer Protocol
IaaS	Infrastructure-as-a-Service
iAOS	integrated Arctic Observation System
ICT	Information and Communication Technologies
NERSC	Nansen Environmental and Remote Sensing Center
OGC	Open Geospatial Consortium
OpenDAP	Open Data Access Protocol
OSDD	OpenSearch Description Document
OWS	OGC Web Services
PaaS	Platform-as-a-Service
SDK	Software Development Kit
THREDDS	Thematic Real-time Environmental Distributed Data Services
VM	Virtual Machine
VPN	Virtual Private Network
WMS	Web Map Service
WPS	Web Processing Service

1. Introduction

This document provides the iAOS application integration and deployment guidelines, and describes how to use the platform and tools made available to the INTAROS partners to build the iAOS as a federation of distributed compute and data resources.

The guidelines include hands-on exercises that explain the Cloud Platform usage in iAOS, from the perspective of the providers of new processing services, to be integrated and deployed in order to support iAOS users with identified requirements for accessing and processing data.

1.1. Using Terradue Cloud Platform

With iAOS, a Hybrid Cloud approach is providing state of the art Platform services, accessible to the INTAROS partners for implementing applications.

The primary purpose of the Hybrid Cloud Platform is to facilitate the management of elastic compute resources, with low cost scale-out capabilities. It relies on the concept of an application integration environment (PaaS, or Platform-as-a-Service) and a production environment:

- **Integration environment:** Private Cloud infrastructure with specialized services for application integration, test, and packaging (PaaS from Terradue Cloud Platform)
- **Production environment:** Public Cloud infrastructures providing CPU and Storage as-a-Services (IaaS from EGI.eu federated Cloud, DIAS providers, etc...)

The transition from the integration environment to a selected production environment is technically supported by Terradue Cloud Platform (Cloud bursting APIs).

1.2. Accessing federated remote repositories

The iAOS approach to federate distributed repositories is based on interoperability conventions, implemented on both the service providers side and the service consumer side (data access protocols, catalogue services). The technical implementations rely on standard APIs. Examples of client software tools currently in scope to support iAOS partners are provided in this document, in the section 3. "Catalog Client tools".

The added value for iAOS to access federated remote repositories is that users can access data from multiple datastores from a single portal or client application, while the data remains within the control and ownership of the data provider.

2. Processing Service Integration Guide

The integration and deployment of new processing services for iAOS is demonstrated through a set of activities led by the INTAROS project partner NERSC. This capacity is delivered in INTAROS via Terradue Cloud Platform.

The Terradue Cloud Platform core services are hosted on Terradue's private cloud infrastructure, where the maintenance activities are directly performed by Terradue's DevOps team. This Private Cloud environment supports a Platform-as-a-Service (PaaS) solutions for application integration, and provides the capability to deploy in production (on a selected IaaS provider) the integrated/validated services.

Integration environment (PaaS from Terradue)

Terradue Cloud Platform builds on three major outcomes of the recent developments in Computer Science and Web technology - Cloud Computing, Open Data repositories, and Web Services interoperability. They constitute the foundation of the PaaS capabilities delivered to developers and service integrators. They provide the backbone and the core services on top of which Terradue Cloud Platform solutions are delivered to the PaaS users, and that we'll introduce in this guide:

- **Ellip Notebooks**, to create interactive laboratory notebooks
- **Ellip Workflows**, to design scalable processing chains
- **Ellip Launchpads**, to select a data processor and deploy it as a Web Processing Service, on a given Processing Center, hosted on a selected IaaS
- **Ellip Infohubs**, to connect a Web Processing Service to a custom Geobrowser client application, to be accessed as part of a user community.

The present Platform and tools Guide provides technical descriptions in support to the INTAROS Partners that are developing and delivering their Processing Service via the INTAROS Platform.

The Platform equips Application developers with a virtualized computing environment, to develop, test and validate their EO data processing applications. Developers benefit from a cost effective PaaS environment, to develop and simulate their Application behaviour on computing clusters. Once deployed, the resulting Application is made available as a highly scalable data processing service.

For each Processing Service, a final deployment in production is supported by the Platform, to run the Processing Service back-end onto a selected IaaS (providing distributed computing resources), and optionally to also run the other Processing Service components (e.g. Front-end software) on selected ad-hoc ICT resources hosted on a IaaS provider.

Production environment (IaaS from Cloud providers)

Deployments in production are managed from Terradue Cloud Platform, towards selected IaaS providers, in the perspective of the operational exploitation of these Applications at

scale (that will ingest large volumes of data to process), and for a competitive cost (typically on a commercial Cloud Computing cluster, or possibly on Cloud Computing resource available for research organizations).

EGI.eu is a INTAROS partner, providing IaaS resources for the INTAROS Applications when they are validated for production. Other Cloud providers are accessible from Terradue Cloud Platform and work is on-going to connect the DIAS providers, as additional IaaS providers capable to serve the INTAROS Partner needs for production environments on the Cloud (storage, data access, compute).

2.1. How to join the PaaS environment

Business Partners can join the PaaS environment as a service integrator, by creating a user account and getting an account upgrade.

For this, they can directly sign-up as a Terradue Cloud Platform user:

- <https://www.terradue.com/portal/signup> (cf. sign-up page registration form)

Once done and signed-in, users can access their Welcome page:

- <https://www.terradue.com/portal/welcome> (when signed-in)

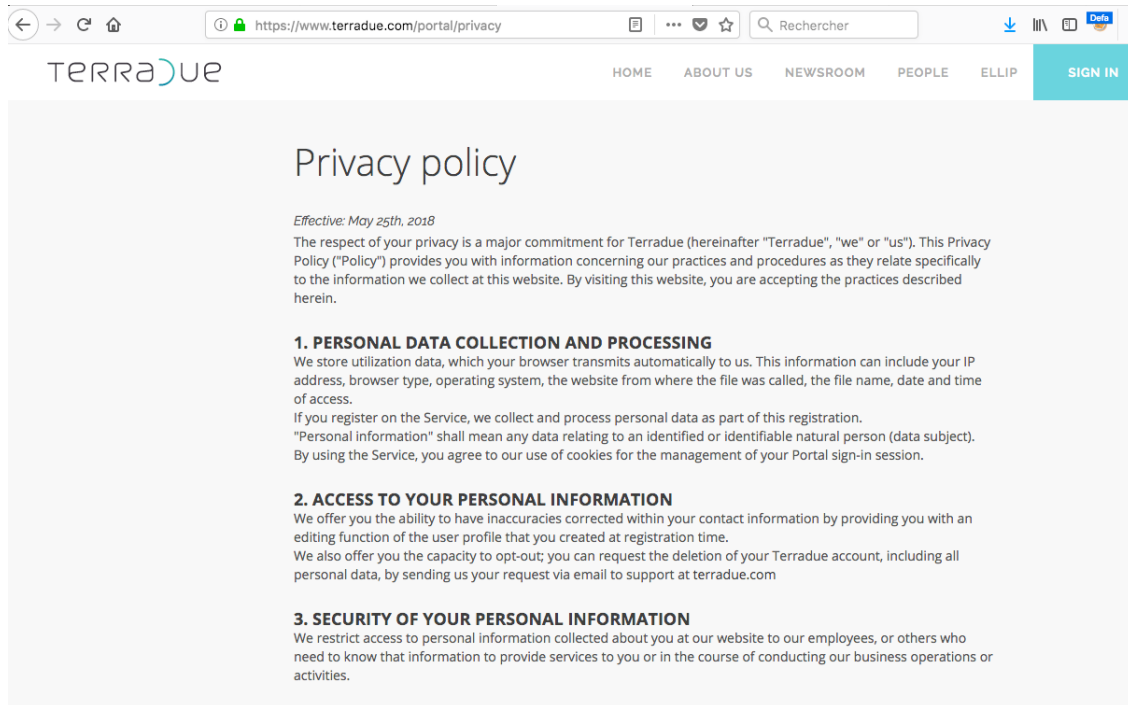
This pages provides the user with the links to the following services on the Terradue Cloud Platform:

- Access the Ellip Core Services and the Ellip Solutions that are part of your subscription level
- Request an account upgrade (see details further below)
- Access your user profile information (account settings):
 - **About you:** personal details (username, first name, last name, affiliation, country)
 - **Account:** user account details (username, email address, registered external authentication modes, password change)
 - **API Key:** used to exploit the API functions on Terradue Cloud Platform (e.g. use of the catalogue API to manage a personal/company catalogue index)
 - **Catalog:** creation and management of your personal catalogue index on Terradue Cloud Platform
 - **Storage:** creation and management of your personal data storage on Terradue Cloud Platform
 - **SSH keys:** used to access via SSH protocol your Virtual Machines created on Terradue Cloud Platform
 - **Github:** to link your github account and use it as version control tool for your developments.
 - **Subscription plans:** your subscription status to Terradue Cloud Platform services, and the "Contact us" function to request an account upgrade.

The screenshot displays the user profile settings interface on the Terradue Cloud Platform. At the top, the 'TERRADUE' logo is on the left, and navigation links for 'HOME', 'ABOUT US', 'NEWSROOM', 'PEOPLE', and 'ELLIP' are on the right. A teal button labeled 'Settings' is also visible. On the left side, a vertical menu under the heading 'Settings' includes 'About you' (highlighted in teal), 'Account', 'API key', 'Catalogue', 'Storage', 'SSH keys', 'Github', and 'Subscription plans'. The main content area is titled 'PUBLIC PROFILE' and contains a form for updating public information. The form includes the following fields: 'Cloud username' (with a blue cursor), 'First name' (filled with 'Herve'), 'Last name' (filled with 'Caumont'), 'Affiliation' (filled with 'Terradue'), and 'Country' (filled with 'France'). An orange 'Submit' button is located at the bottom of the form.

The GDPR notice related to user control over his/her personal information managed on the Portal is accessible from the Terradue website footer:

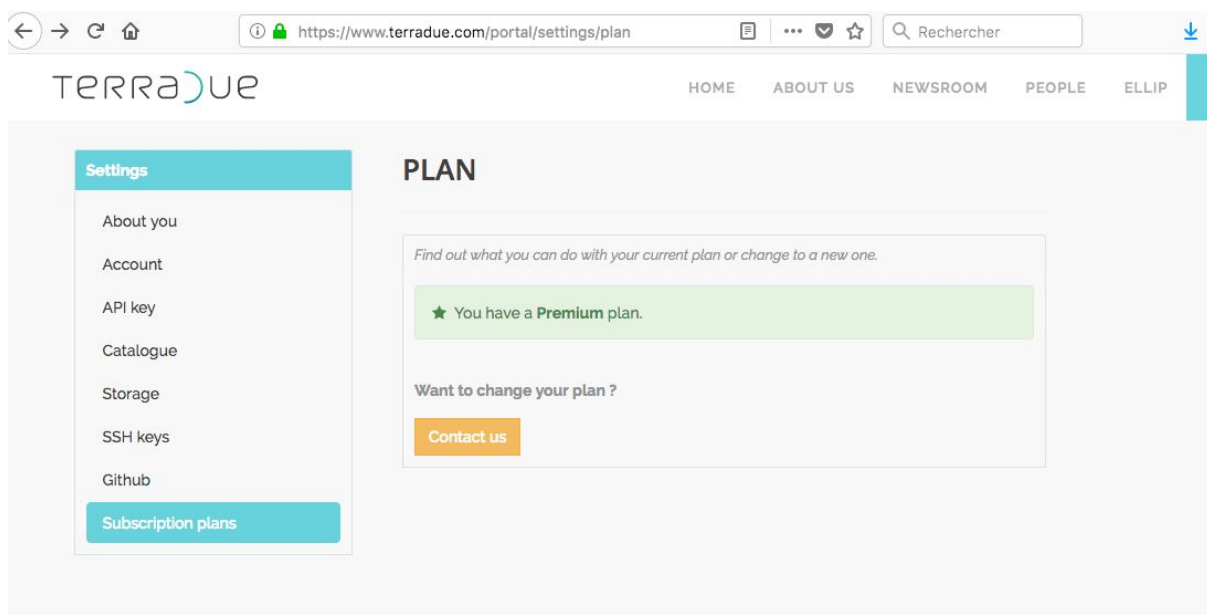
- <https://www.terradue.com/portal/privacy>



How to request a Terradue account upgrade ?

After the account creation, one needs to get an account upgrade, with the proper access rights granted. This can be requested via the Terradue Portal account settings page:

- <https://www.terradue.com/portal/settings/plan> (when signed-in, "Subscription plans" menu entry)



2.1.1. My account (as a PaaS user)

Once signed-in (see sections above), users of the Integration environment (PaaS from Terradue) can access their user account from here:

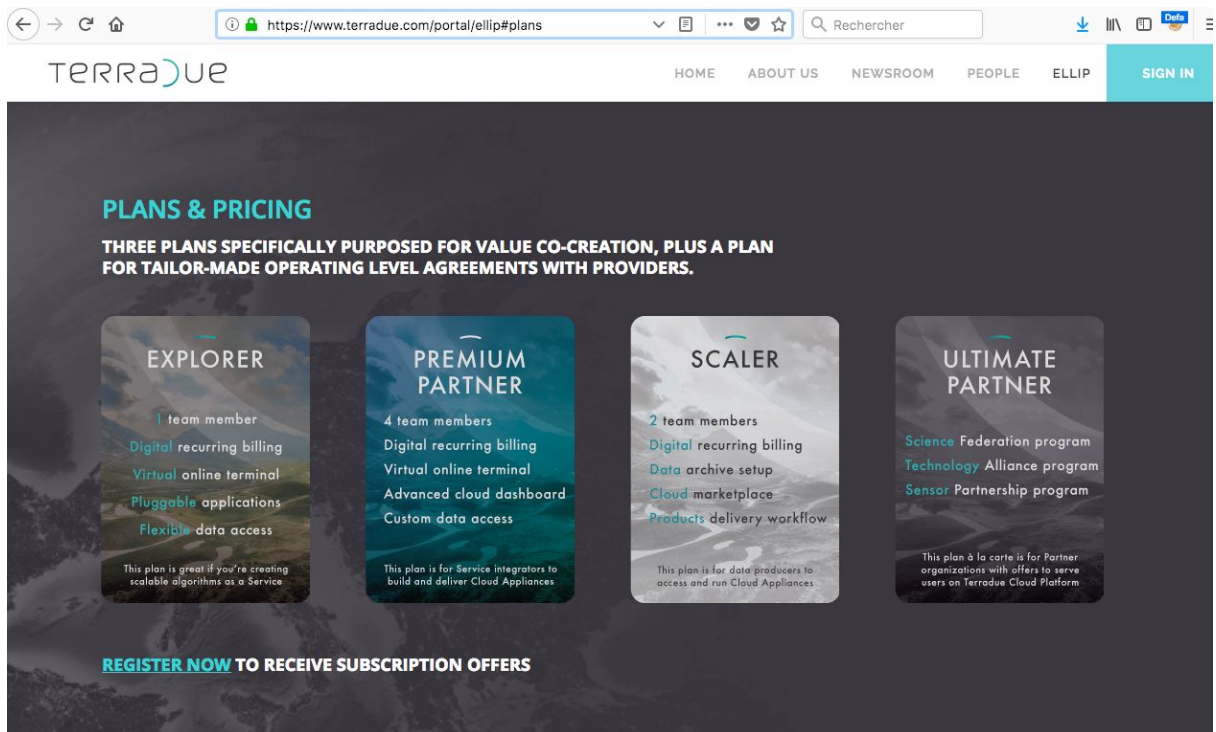
- <https://www.terradue.com/portal/settings/profile> (when signed-in)

Depending on their account upgrade status, user accounts are assigned a set of credentials which are defined as follows:

- **No plan:** after initial registration, accounts do not have a subscription plan associated, they imply by default 'Visitor' accounts
- **Membership:** this is a user account upgrade meant for the users of external, Ellip-powered systems. A membership account means the possibility to create on Terradue Cloud Platform a personal Cloud storage space, a personal catalog index, as well as an API key, that are linked from and used by the Portal functions of a given external, Ellip-powered system.
- **Explorer:** this subscription plan is meant for creating processing services as Cloud Appliances, either based on Jupyter Notebook executables or on scalable algorithms packaged for a data processing framework (e.g. Hadoop)
- **Scaler:** this subscription plan is meant for data producers to access and run Cloud Appliances previously prepared on Terradue Cloud Platform
- **Premium Partner:** this subscription plan is meant for Service integrators to build and deliver client side user services (e.g. Web Portal) and server side Cloud Appliances on Terradue Cloud Platform, as well as deploy and operate these on third party IaaS providers (production environments).
- **Ultimate Partner:** this subscription plan 'à la carte' is meant for organizations with offers to serve users on Terradue Cloud Platform (data providers, technology providers, data processing module providers).

The subscription plans to access the tools and services of the Integration environment (PaaS) are presented on the Terradue website:

- <https://www.terradue.com/portal/ellip#plans>

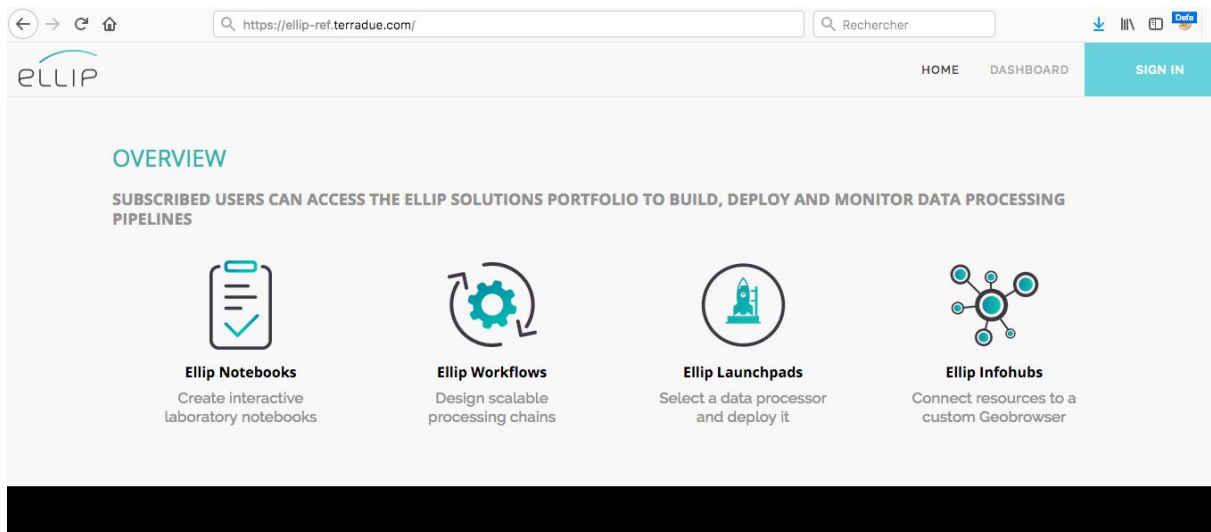


Once subscribed to a plan, you can access the Ellip Dashboard from here:

- <https://ellip.terradue.com/> (Ellip Dashboard interface to access the subscribed Ellip Solutions)

As of today (this is subject to change in the future), access control is provided as follows:

- Solutions subscribed to by **Explorer** users:
 - Ellip Notebooks
 - Ellip Workflows
- Solutions subscribed to by **Scaler** users:
 - Ellip Launchpads
 - Ellip Infohubs
- Solutions subscribed to by **Premium Partners**:
 - Ellip Notebooks
 - Ellip Workflows
 - Ellip Launchpads
 - Ellip Infohubs



The online documentation for the Ellip solutions will be made available from here:

- <http://docs.terradue.com> (entry point for the documentation about Terradue Cloud Platform and the Ellip solutions).

2.1.2. Partner Support

Partner users with “Premium Partner” or “Ultimate Partner” subscription plans can access a dedicated support space in order to interact with Terradue’s DevOps team:

- <http://support.terradue.com> (entry point for the service integrators, data providers, technology providers, data processing module providers that are integrating value-added services on Terradue Cloud Platform via the provided APIs).

2.1.3. User Forum

All Terradue users can access the public user Forum, and also start contributing to it (after login). Several sections on the User Forum are already contributed by user communities. All the sections are under the editorial control of Terradue (as a moderator), and the Forum is a source of shared knowledge about using the Platform tools and services, either as part of the application integration activities, or as part of the exploitation of Ellip-powered services.

- <http://discuss.terradue.com> (entry point for users having a Terradue account, to sign-in and comment or ask questions on existing posts, or provide their feedback, hints & tips and advices to others about using the Platform tools and services).

2.1.4. User Helpdesk

Terradue users having a Membership can access the User Helpdesk service as a 'customer' requesting support, for the projects they have been granted. For these projects, they can receive professional support services in relation with an Ellip-powered system they are using.

Terradue users having a Subscription plan can access the User Helpdesk service as an 'agent' delivering support, for the projects they are managing. For these projects, they can deliver professional support services in relation with an Ellip-powered system they are operating.

- <http://helpdesk.terradue.com> (entry point for users having a Membership or a Subscription plan to benefit from a professional service desk environment supporting their exploitation goals on Ellip-powered systems).

2.2. Using the Cloud Platform's core services

We present hereafter the list of the Ellip core services, with their access points, and their access conditions for Terradue Cloud Platform users in order to build, test and deploy data processing pipelines.

2.2.1. Data Agency

The Data Agency is made of two core services for the management of federated data repositories and scalable data staging operations:

- <https://store.terradue.com>
Access to your personal storage on the store service follows this template:
 - <https://store.terradue.com/<username>>
 - <username> is provided via your Terradue user accountThe protocol to access the store functions is based on HTTP:
 - <http://docs.terradue.com/t2-api/data/storage.html>
- <https://catalog.terradue.com>
Access to your personal index on the catalog service follows this template:
 - <https://catalog.terradue.com/<username>/search?apikey=<YOUR-API-KEY>>
 - <username> is provided via your Terradue user account
 - <YOUR-API-KEY> is provided via your Terradue user accountThe protocol to access the catalog functions is based on OpenSearch:
 - <http://docs.terradue.com/t2-api/data/publication.html>

2.2.2. Cloud Dashboard

The Cloud Dashboard service is used for the creation and management of Containers and Virtual Machines supporting the application integration work done within the PaaS environment, as well as for the deployment of Cloud Appliances in production (cloud bursting):

- <https://cloud.terradue.com> (for developer users to create and manage their VMs)
- <https://sandbox-ip/dashboard> (for developer users to access their VM dashboard and work on it)

2.2.3. Virtual Private Network

The Virtual Private Network service is used for establishing a secured communication from a client (the user desktop/laptop machine) to the Virtual Machines that the user has created on Terradue Cloud Platform:

- <https://access.terradue.com> (for developer users to activate their VPN connexion, for their access to VMs managed on Terradue's private Cloud)

2.2.4. Other Cloud Platform services

Additional services are being configured on Terradue Cloud Platform in order to serve the PaaS users needs. They will be documented in a later stage of the Terradue Cloud Platform developments.

- geo.terradue.com (Geoserver instance)
- recast.terradue.com (Data Casting service)
- repository.terradue.com (Software repository)
- anaconda.org/Terradue (Software repository)

Upcoming: as part of the iAOS user needs, a dedicated MATLAB Compiler service is foreseen as a Platform service made available to subscribers.

2.3. Using the Cloud Platform's Ellip solutions

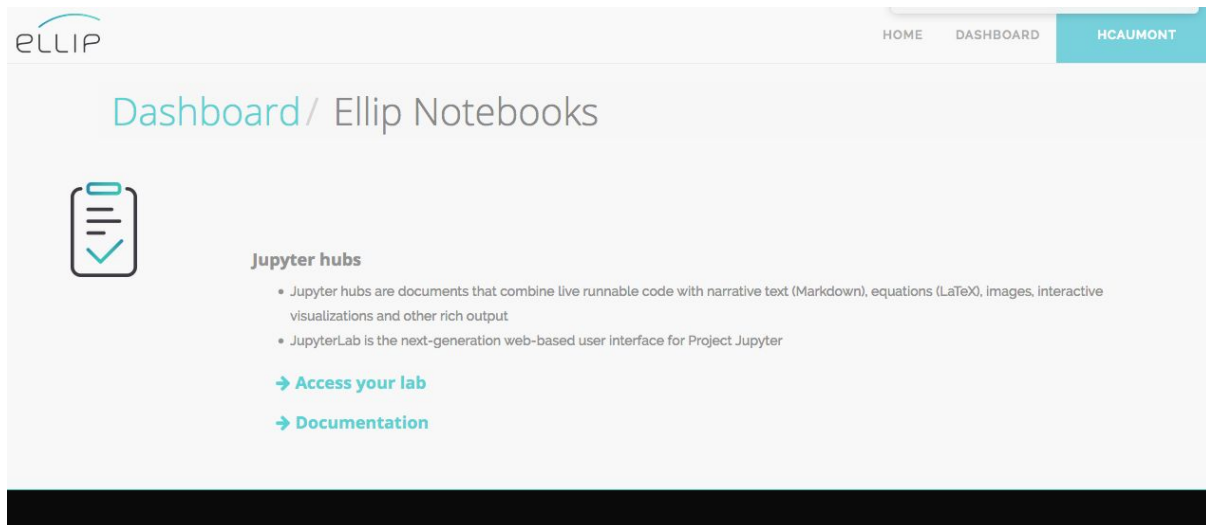
The **Ellip Solutions** are empowering the PaaS users to create value as part of the build, test and deploy phases of a data processing pipeline lifecycle. They can select the way (open / paid) to share their assets on Terradue Cloud Platform to other users, in a self-service mode.

All the value creation (assets) done by using an “Ellip Solution” will become referenced on the **Ellip Exchange**, a new cloud platform core service, that will be developed during the coming year, and will enforce Ellip as “a collaborative workplace” where partners can co-design their Applications.

2.3.1. Ellip Notebooks

The Ellip Notebooks solution is based on Jupyter Hub. It provides individual users with a Jupyter Notebook server.

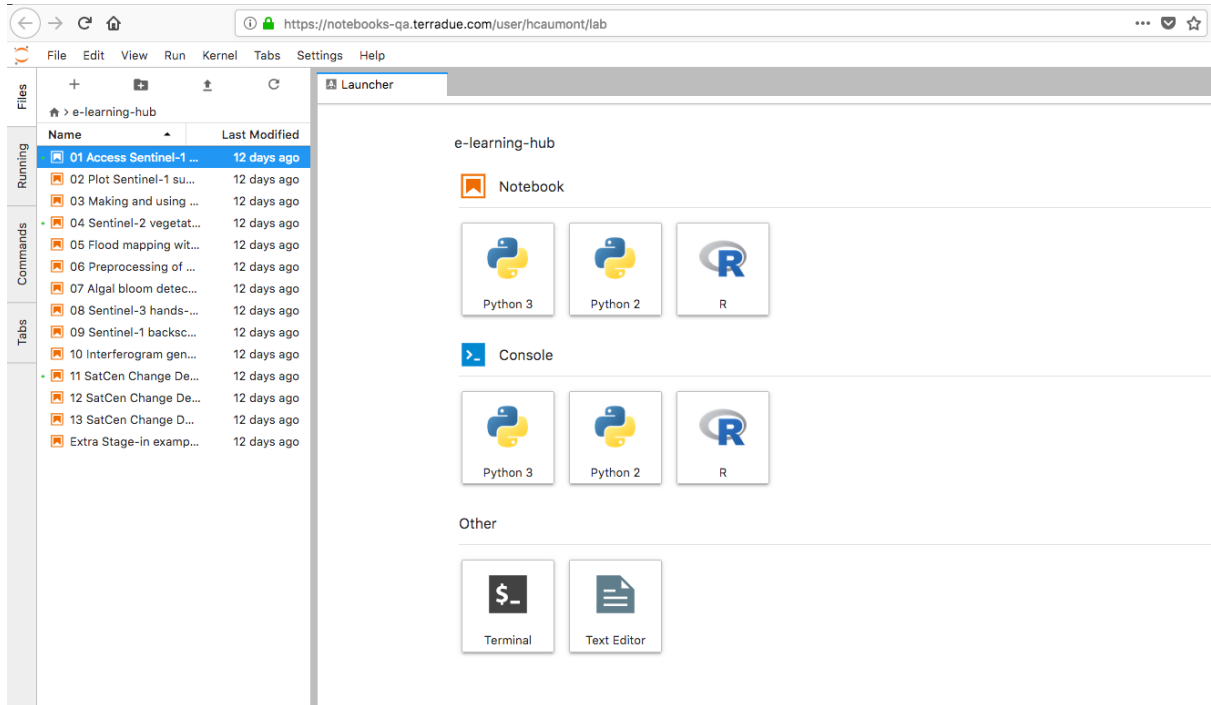
It is made available to Explorer plan subscribers & Premium Partners.



The Ellip Notebooks solution allows users to integrate, test and share simple data processing functions in a very interactive way. These functions are implemented a Notebooks, which can further on become unitary nodes of a data processing pipeline.

It is also suitable for users to easily integrate client software contacting data access or data processing Web Services.

The Ellip Notebooks Solution provides a user with a dedicated access to a Jupyter Notebook server instance, managed on the Platform’s JupyterHub cluster.



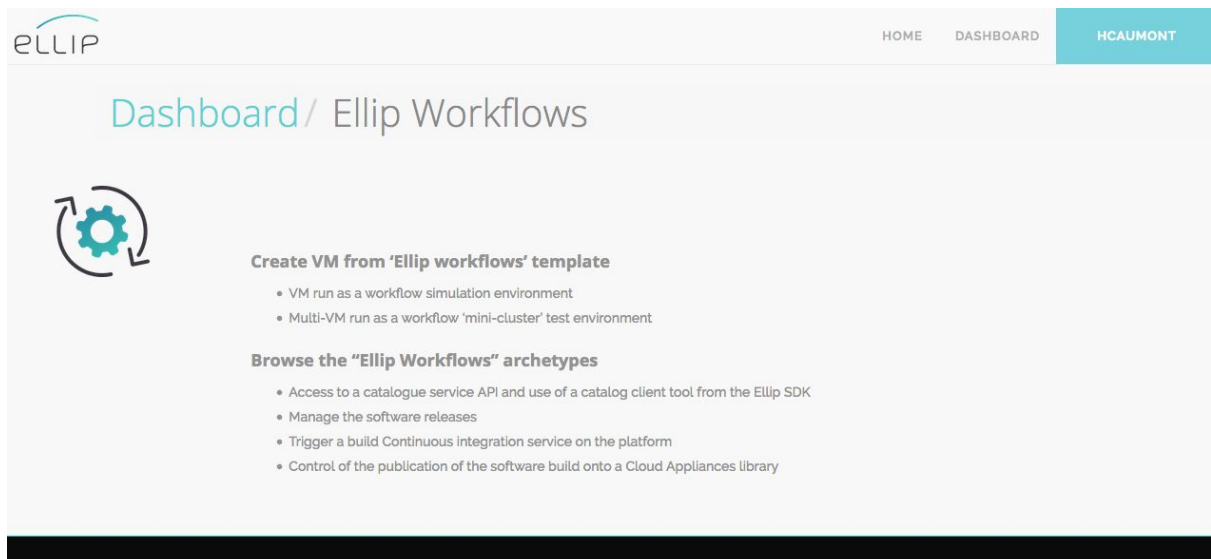
The user documentation for the Ellip Workflows solution is available online:

- <https://docs.terradue.com/ellip/solutions/notebooks>

2.3.2. Ellip Workflows

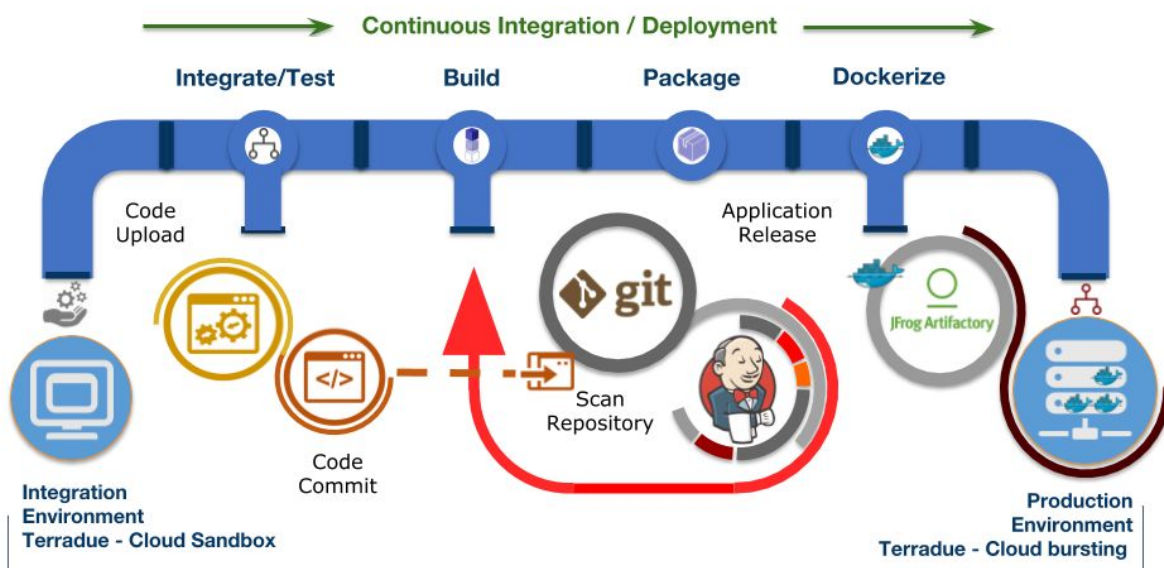
The Ellip Workflows solution is based on a framework including the Hadoop MapReduce streaming API as well as ad-hoc data staging tools and services ('ciop' tools) and Web Processing Service interface (WPS).

It is made available to Explorer plan subscribers & Premium Partners.



The Ellip Workflows solution allows users to integrate, test and validate/package scalable data processing workflows, that can be deployed as highly interoperable data processing services. The application Integration Environment (Cloud Sandbox) is accessed by developer users within a Virtual Machine (VM). It consists in a software framework and tools for processor integration and testing, that allows to run initially the applications in a simulation environment (sandbox mode) that can readily scale when deployed in production (cluster mode) without any further modification. The auto-scalability feature enabled by the provided framework and tools is a major asset of the Terradue PaaS.

A new interface is furthermore made available within the Ellip Workflows solution, also based on the Jupyter Notebook technology, which allows to integrate data processing services within the provided framework in a very interactive way.



The Cloud Sandbox environment allows to plug scientific applications written in a variety of languages (e.g. Java, C++, IDL, Python, R), then test, validate and package these applications. The Application Packaging output is a file in RPM format that, once validated, can be easily installed on a production host or cluster, including via the management of Docker images. The package provides both the Application and the software dependencies required for executing the Application itself on the Production environment.

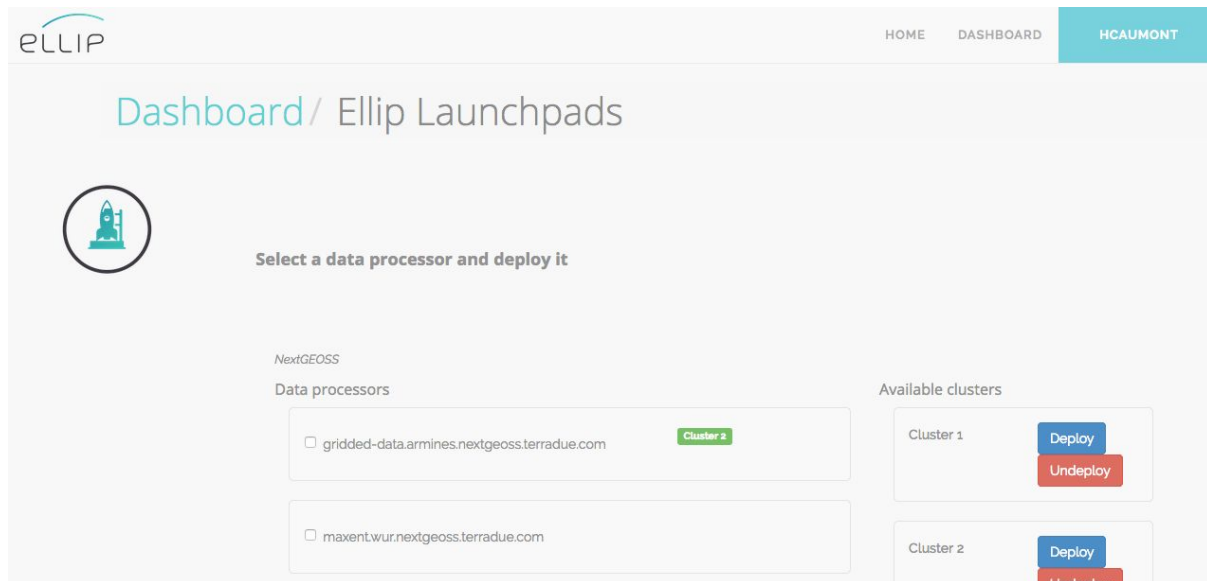
Overall, the Continuous Integration capacity on Terradue Cloud Platform provides the automation layers to manage versioned application packages, in order to deploy (cloud bursting) these from Terradue Cloud Platform, towards a selected Production environment, where they can be accessed as scalable data processing services, in a Software-as-a-Service mode.

The user documentation for the Ellip Workflows solution is available online:

- <https://docs.terradue.com/ellip/solutions/workflows>

2.3.3. Ellip Launchpads

It is made available to Scaler plan subscribers & Premium Partners.



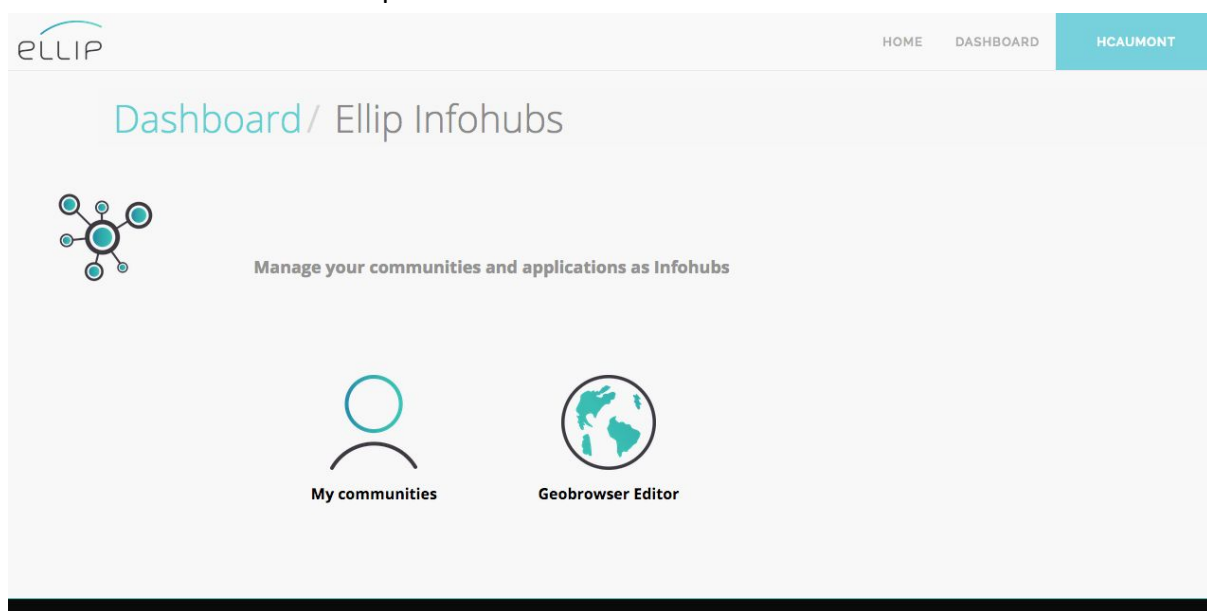
The Ellip Launchpads solution will be made available to users in a later stage of the Terradue Cloud Platform developments.

The user documentation for the Ellip Launchpads solution is not yet available online:

- User documentation: upcoming

2.3.4. Ellip Infohubs

It is made available to Scaler plan subscribers & Premium Partners.



The Ellip Infohubs Solution provides to its users the ability to create ready-to-use geobrowser apps, and save their definition as standardised OWS Context files in the Platform’s catalog, allowing collaborative work to take place by customising existing templates, and sharing the resulting Apps with selected group of users.

Step 1: define properties

Properties | Base Maps ▲ | Map features | Data collections | Data features | WPS Services | Save ▲

Basic properties of your app

Title
My demo app

Add new field ▾

Summary
This is my demo app created during the NextGEOSS second training session.

Icon Url
https://pbs.twimg.com/profile_images/99491530064159

Authors

Name	Email	Uri
Enguerran Boissier	enguerran.boissier@terradue.com	https://www.terradue.com

➕ Add a new author

Step 2 & 3: select a base map and map features (like existing WMS layers)

Properties | Base Maps | Map features | Data collections | Data features | WPS Services | Save ▲

Background maps of your app

Selected your Base Maps from available base maps box (The first is the default one)

default
Dark map | Light map

Available maps

Default | Natural Earth | Google Road map

Step 4: configure data collections

The screenshot shows the 'Data collections' configuration page. At the top, there is a navigation bar with tabs for 'Properties', 'Base Maps', 'Map features', 'Data collections' (selected), 'Data features', 'WPS Services', and 'Save'. Below the navigation bar, a light blue box contains the following text: 'Data collections available in your app. Group collections using a first level context. Each context with several collections will appear as a dropdown menu button on the geobrowser'. Below this, there is a table with two columns: 'Menu Path' and 'Description url'. The table contains three rows of data. To the right of the table is a 'Tree Menu Preview' showing a hierarchical structure of the menu items. At the bottom left of the configuration area, there is a button labeled 'Add new data context'.

Menu Path	Description url
EO Data/Sentinel-1	https://catalog.terradue.com/sentinel1/description
EO Data/Sentinel-2	https://catalog.terradue.com/sentinel2/description
EO-based products/Sentinel-1 Medium-Resolution InSAR Browse	https://catalog.terradue.com/dlr-insarql/search?pt=DPHASE

Tree Menu Preview

- EO Data
 - Sentinel-1
 - Sentinel-2
- EO-based products
 - Sentinel-1 Medium-Resolution InSAR Browse

Step 5: configure the data management features

The screenshot shows the 'Data features' configuration page. At the top, there is a navigation bar with tabs for 'Properties', 'Base Maps', 'Map features', 'Data collections', 'Data features' (selected), 'WPS Services', and 'Save'. Below the navigation bar, a light blue box contains the text: 'Additional data features available in your app'. Below this, there are two main sections: 'Data search' and 'Feature Basket'. The 'Data search' section has a sub-header 'Data search' and a description: 'Allow to search over defined data collections and list search results within a Search Result tab'. There is a checkbox labeled 'Enable Search Result Tab' which is checked. The 'Feature Basket' section has a sub-header 'Feature Basket' and a description: 'Allow user to temporarily save datasets returned by search results into a dedicated Feature basket'. There is a checkbox labeled 'Enable Feature Basket Tab' which is checked. To the right of the 'Feature Basket' section, there is a preview of the 'Features Basket' interface. The preview shows a header 'Features Basket' with 'Total results 4' and buttons for 'see all', 'inv.sel.', 'Remove all', and 'Save'. Below the header, there are four items listed in a list view, each with a search icon and a title: '(Aci-576/Cali-657) Fuego Volcano in Guatemala', '(Aci-575/Cali-656) Fire in China', '(Aci-574/Cali-655) Cyclone Mekunu in Yemen', and '(Aci-573/Cali-654) Flood in Sri Lanka'.

Data search

Allow to search over defined data collections and list search results within a Search Result tab

Enable Search Result Tab

Feature Basket

Allow user to temporarily save datasets returned by search results into a dedicated Feature basket

Enable Feature Basket Tab

Features Basket

Total results 4 | [see all](#) | [inv.sel.](#) | [Remove all](#) | [Save](#)

- (Aci-576/Cali-657) Fuego Volcano in Guatemala
- (Aci-575/Cali-656) Fire in China
- (Aci-574/Cali-655) Cyclone Mekunu in Yemen
- (Aci-573/Cali-654) Flood in Sri Lanka

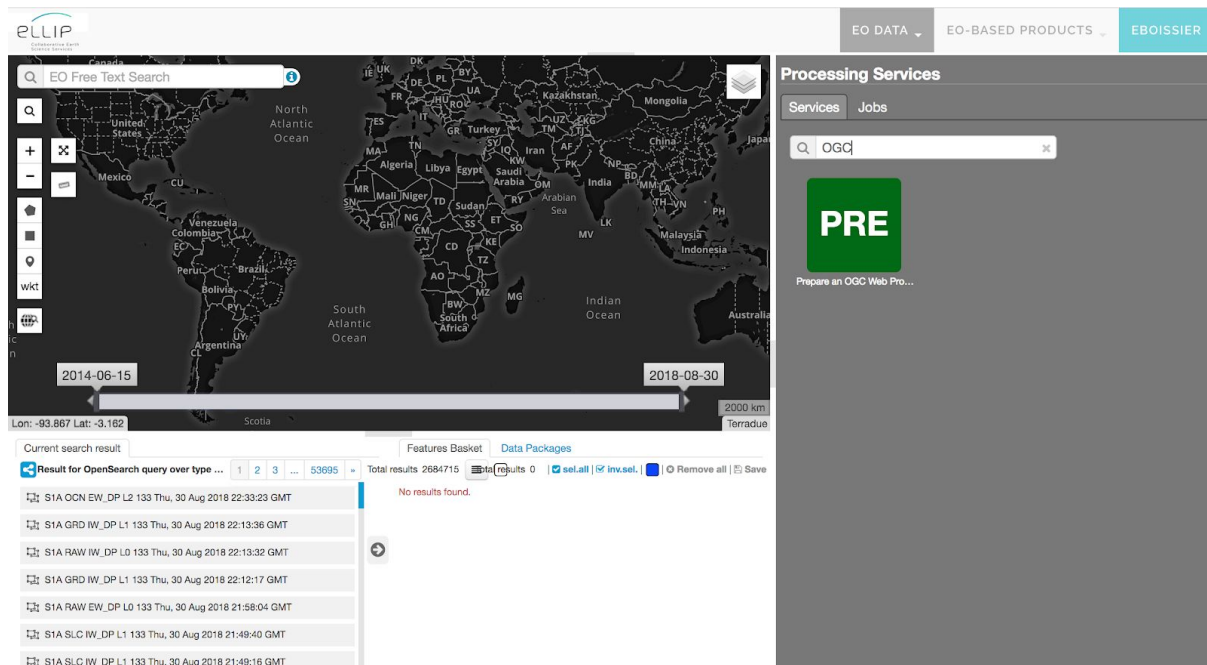
Step 6: configure the WPS services to be accessed

The screenshot shows the 'WPS Services' configuration tab in a software interface. At the top, there is a navigation bar with tabs for 'Properties', 'Base Maps', 'Map features', 'Data collections', 'Data features', 'WPS Services', and 'Save'. Below the navigation bar, a light blue banner contains the text: 'WPS processing functionalities available in your app'. The main section is titled 'WPS Processing' and includes a descriptive paragraph: 'Allow user to process new jobs from a list of available processing services, visualize the results on the map and share it with other users, as well as access shared jobs from other users'. A checkbox labeled 'Enable WPS Services Tab' is checked. Below this, a yellow banner states: 'You can use filters to select only specific WPS services. All filters are optional.' The 'Filter with domain' section has a sub-header and a description: 'Returns only WPS belonging to the selected domain'. It features two input fields, both containing the text 'eboissier'.

Step 7: save the result as an OWS Context file

The screenshot shows the 'Save Application' configuration tab in the same software interface. The navigation bar at the top is identical to the previous screenshot. A light blue banner at the top of the main area says 'Save your app'. Below it is a 'Summary of your app' section with a '(show)' link. The 'Save Application' section contains two input fields: 'Identifier' with the value 'mydemoapp-nextgeoss' and a checkmark, and 'Index' with the value 'eboissier'. A red 'Save' button is located at the bottom of this section.

Finally, users can access the instantiated Infohub via a group management feature:



The user documentation for the Ellip Infohubs solution is not yet available online:

- User documentation: upcoming

2.5. Cloud Platform SDK Handbook

The Ellip Software Development Kit (SDK) is providing functions for using the Terradue Cloud Platform APIs, performing Data management, Catalogue access, Workflow simulation and invocation as well as parameters retrieval and error logs management.

The user documentation for the Cloud Platform SDK is available online:

- <https://docs.terradue.com/ellip/developer/sdk>

3. Catalog client tools

INTAROS partners have currently the possibility to use three standard ways to retrieve datasets from distributed repositories, that are connected to the iAOS via an approved protocol: the OGC OpenSearch interface, the PANGAEA data download service interface and the OpenDAP THREDDS interface.

3.1. Terradue Catalog client - OpenSearch for EO data

The OpenSearch specification defines an OpenSearch Description Document (OSDD) format in order to expose the capabilities of a search engine, that can be used by search client applications.

The OpenSearch client provided on Terradue Cloud Platform as part of the Ellip SDK uses the OSDSD to learn about the public interface of the server. The OSDSD contains parameterized URL templates, that indicate how the search client should place search requests.

The usage of the OpenSearch client tool is documented online here:

https://docs.terradue.com/ellip/developer/sdk/bash_commands_functions/catalogue/opensearch-client.html

3.3. OpenDAP/THREDDS Catalog client

The PyDap client software is made available on Terradue Cloud Platform as part of the Conda repository (Python software libraries).

Hereafter, an example to query a THREDDS catalog (example of the OpenDAP server instance provided by IMR in the context of iAOS):

The following code demonstrates use of THREDDS catalog, to find list of datasets available for processing

```
def
load_all_data(geo_bounds,min_date,max_date,min_depth,max_depth,server_type="thredds",verbose=0,show
_time=0):
    #Download sub-catalog
    if server_type is "thredds":
        catalog_url = "http://opendap1-test.nodc.no/thredds/catalogs/physics/physics_point_yearly.xml"
    else:
        catalog_url = "http://opendap1.nodc.no/opendap/physics/point/yearly/catalog.xml"
    contents = urllib.request.urlopen(catalog_url).read()

    #Parse xml
    e = xml.etree.ElementTree.fromstring(contents)

    #Find dataset list
    root_dataset = e.find('{http://www.unidata.ucar.edu/namespace/thredds/InvCatalog/v1.0}dataset')
    for child in root_dataset:
        url = "http://opendap1.nodc.no"+child.get('ID')
        if server_type is "thredds":
            url = "http://opendap1-test.nodc.no/thredds/dodsC/"+child.get('urlPath')
        else:
            url = "http://opendap1.nodc.no"+child.get('ID')

        min_time = date2time(min_date)
        max_time = date2time(max_date)
        time_depth_bounds = geometry.box(min_time,min_depth,max_time,max_depth)
        load_data(url,geo_bounds,time_depth_bounds,verbose,show_time)
```

The complete guidance in using this client library in the context of iAOS will be made available to users in a later stage of the project developments.

The user documentation for this software client is under works.

- User documentation: upcoming

3.2. Pangaea Catalog client

The PANGAEA data download service allows a client application to query the FRAM data warehouse and filter by geocodes and parameters. The service returns values in tab delimited text files.

The usage of the PANGAEA data download service interface is documented online here: <http://ws.pangaea.de/dds-fgp/>

In summary:

Mandatory search parameters	Optional search parameters
<p>@westBoundLon description: west bound longitude example: westBoundLon=-96.5667</p> <p>@southBoundLat description: south bound latitude example: southBoundLat=-13.4251</p> <p>@eastBoundLon description: east bound longitude example: eastBoundLon=-96.9987</p> <p>@northBoundLat description: north bound latitude example: northBoundLat=-12.9052</p> <p>@pangaeaParameterIDs description: the parameters to retrieve (internal PANGAEA parameter ID's need to be provided) example: pangaeaParameterIDs=717,716,80103,756,128305</p>	<p>@dateTimeMin description: datetime minimum (ISO 8601) example: dateTimeMin=2011-04-14T00:00:00</p> <p>@dateTimeMax description: datetime maximum (ISO 8601) example: dateTimeMax=2011-04-16T00:00:00</p> <p>@depthMin description: minimal water depth (metres) example: depthMin=45</p> <p>@depthMax description: maximal water depth (metres) example: depthMax=55</p> <p>@depthMinSed description: minimal sediment/rock depth (metres) example: depthMinSed=0.01</p> <p>@depthMaxSed description: maximal sediment/rock depth (metres) example: depthMaxSed=0.15</p> <p>@projectID description: internal PANGAEA project ID example: projectID=4055</p>

Service call examples:

<http://ws.pangaea.de/dds-fgp/rest/dwhquery?westBoundLon=-96.9984&southBoundLat=-13.41512&eastBoundLon=-96.9987&northBoundLat=-13.41124&pangaeaParameterIDs=717>

Description: service call using only the mandatory search parameters, which returns all temperature data archived at PANGAEA within the specified bounding box.

<http://ws.pangaea.de/dds-fgp/rest/dwhquery?westBoundLon=-96.5667&southBoundLat=-13.42512&eastBoundLon=-96.9987&northBoundLat=-12.9052&pangaeaParameterIDs=717,716,80103,756,128305&depthMin=45&depthMax=55&dateTimeMin=2011-04-14T00:00:00&dateTimeMax=2011-04-16T00:00:00>

Description: service call using mandatory and optional search parameters, which returns all values archived at

PANGAEA for the specified PANGAEA parameter ID's within a bounding box defined in horizontal space (latitude, longitude), vertical space (water depth) and time.

<http://ws.pangaea.de/dds-fgp/rest/dwhquery?pangaeaParameterIDs=24645&westBoundLon=-15&southBoundLat=75&eastBoundLon=15&northBoundLat=82&dateTimeMin=2000-01-01T00:00:00&dateTimeMax=2001-01-01T00:00:00&projectID=4055>

Description: service call using mandatory and optional search parameters, which returns all values archived at PANGAEA for the specified PANGAEA parameter ID's within a bounding box defined in horizontal space (latitude, longitude) and time. The internal PANGAEA project ID is used as an additional filter.

<http://ws.pangaea.de/dds-fgp/rest/dwhquery?pangaeaParameterIDs=157811,4870&westBoundLon=-4&southBoundLat=79&eastBoundLon=5&northBoundLat=80&projectID=4055&depthMinSed=0.001&depthMaxSed=0.1&dateTimeMin=2009-07-15T06:45:00&dateTimeMax=2009-07-15T06:47:00>

Description: service call using mandatory and optional search parameters including sediment depth, which returns all values archived at PANGAEA for the specified PANGAEA parameter ID's within a bounding box defined in horizontal space (latitude, longitude) and time. The internal PANGAEA project ID is used as an additional filter.

----- END of DOCUMENT-----



INTAROS

This report is made under the project
Integrated Arctic Observation System (INTAROS)
funded by the European Commission Horizon 2020 program
Grant Agreement no. 727890.



Project partners:

