



**InGeoCloudS**  
Inspired GEOdata CLOUD Services



## DELIVERABLE D4.2

**Grant Agreement number : CIP-297300**  
**Project acronym : InGeoCLOUDS**  
**Project title : INspired GEOdata CLOUD Services**

**Funding Scheme : Pilot B**

# Fully Operational InGeoCloudS Pilot

D4.2

Version 1.0

Reference D4.2-INGC

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

---

Contract Number : CIP-297300  
Document number : D4.2  
Document Title : Fully Operational InGeoCloudS Pilot  
Document version : 1.0  
Document status : Approved  
Date : 2013-10-30

WP contributing to the deliverable : WP4

Availability : Public

Authors : AKKA and InGeoCloudS partners

Approved by : InGeoCloudS Steering Committee

---

**Abstract**      *Starting October 24<sup>th</sup> 2013, the project advertised a Pilot2 of InGeoCloudS. This document accompanies the prototype (the actual deliverable) available on the Internet under the <http://portal.ingeoclouds.eu> domain name. It summarizes services that can be found in this prototype in the form of Release Notes with regards to Pilot1: new services and evolutions of already existing ones as well as some insights into changes concerning the technical exploitation of the pilot. Note that technical implementation details will be covered by D4.3 document.*

---

**Keyword List**      Prototype, data publication, maps, WFS/WMS, CSW, OGC, boreholes, groundwater, landslides, shake-maps, ground acceleration, Cloud,

**DOCUMENT CHANGE LOG**

Document Issue	Date	Reasons for change
Version 1-Draft 1	2013-09-09	Creation of the document with AKKA inputs.
Version1 – Draft2	2013-10-03	Integration of technical partners inputs and BRGM
Version 1 – FinalDraft1	2013-10-16	Integration of DataProviders inputs
Version 1 – FinalDraft2	2013-10-28	Re-ordering some sections about datasets and latest screenshots included
Version1-Approved	2013-10-30	Last proof-reading and edition for delivery.

**APPLICABLE AND REFERENCE DOCUMENTS (A/R)**

A/R and Reference	Title
[A1] ICT PSP Grant Agreement N° CIP 297300	InGeoCloudS Grant Agreement and its annex (including the description of work)
[A2] CA-INGC	InGeoClouds Consortium Agreement v1.0
[R1] D2.1-INGC	Use Cases for InGeoCloudS data and services
[R2] D2.2-INGC	Interface of Web services and models of data
[R3] D2.3-INGC	InGeoCloudS Web Services covering Use Cases
[R4] D3.2-INGC	Cloud Architecture, configuration and data access implementation
[R5] D3.3-INGC	InGeoCloudS maintenance plan and service profiling
[R6] D4.1-INGC	First Implementation of InGeoCLOUDS Pilot

**Table of contents**

<b>1. INTRODUCTION .....</b>	<b>6</b>
<b>2. INGEOCLOUDS DATA AND SERVICES .....</b>	<b>6</b>
2.1. InGeoCloudS Infrastructure Overview .....	6
2.2. InGeoCloudS Datasets .....	8
2.2.1. GEUS Data .....	8
2.2.2. GEO-ZS Data .....	8
2.2.3. BRGM Data .....	8
2.2.4. EPPO Data .....	8
2.2.5. EKBAAs Data .....	9
2.3. Users.....	9
2.3.1. Data Providers.....	9
2.3.2. Registered Users .....	9
2.3.3. Public.....	10
2.3.4. INGC Administration .....	10
2.4. Services Used by Data Providers.....	10
2.4.1. Datasets Pushed Into InGeoCloudS.....	10
2.4.2. Linked Open Data Support.....	11
2.4.3. DataProviders Toolkit.....	12
2.5. Services available for Public and Registered Users.....	13
2.5.1. Portal.....	13
2.5.2. Pesticides in GroundWater.....	13
2.5.1. Landslides Susceptibility Maps Application.....	14
2.5.2. GeoPublication and Web Mapping Application .....	14
2.5.3. ShakeMaps Application.....	14
2.5.4. GroundWater resources management in granular aquifers .....	15
2.5.5. Active Landslide Inventory Mapping and Susceptibility Zoning.....	16
2.5.6. SmartQueries Application.....	17
<b>3. OPERATING PILOT2 .....</b>	<b>19</b>
3.1. Administration Toolkit .....	19
3.2. User Management .....	19
3.2.1. Data providers .....	19
3.2.1. Registered Users .....	19
3.2.2. Public (anonymous users).....	19
3.2.3. INGC Administrators .....	20
3.3. Supervision of System Health .....	20
3.3.1. AWS Management Console .....	20
3.3.2. InGeoCloudS Specific Monitoring and Supervision .....	20
3.3.3. New Load Balancers .....	21
3.4. Support and Helpdesk .....	21



# Deliverable D4.2

## Fully Operational InGeoCloudS Pilot

Ref. : D4.2-INGC  
Version : 1.0  
Status : Approved  
Date : 2013-10-30  
Contract : CIP-297300

### List of Figures

Figure 1 : Logical View of InGeoCloudS Architecture as in Pilot2.....	7
Figure 2 : Management of Registered Users by Data Providers (view from the DataProvidersToolkit).....	10
Figure 3 : New login page on the Portal.....	13
Figure 4 : Sample screenshot of the “Shake-Maps” Application .....	15
Figure 5 : SmartQueries App: example of query illustrating UC6.....	16
Figure 6 : View of the SmartQueries App.....	17
Figure 7 : View of the SmartQueries App.....	19
Figure 8 : Example of dashboard on InGeoCloudS specific indicators.....	21
Figure 9 : Facility for leaving a message to InGeoCloudS helpdesk from the Portal .....	22

## 1. INTRODUCTION

This document describes Pilot2 updating and improving the document that accompanied Deliverable D4.1 “*First implementation of InGeoCloudS Pilot*” which was also of type “Prototype”.

***The document provides “Release Notes” highlighting changes of Pilot2 with regards to Pilot1.***

Some reminders about the InGeoCloudS infrastructure are first addressed: technical characteristics, cloud-based nature, the different components and their deployment.

The document then presents the different types of users and the possible interactions they can have with the system: main characteristics of the pilot are described in a summarized way according to various viewpoints:

- Users: different types of users supported; what they can do with the system
- Data and services: which datasets are available in this pilot and which services are available around them.
- Main Operations: how the Pilot2 is operated on a daily basis by consortium members.

Reference documentation on the wiki and more contextual online help as described in D5.1 “*InGeoCloudS Users documentation*” are both accessible directly from Pilot2.

## 2. INGEOCLOUDS DATA AND SERVICES

### 2.1. INGEOCLOUDS INFRASTRUCTURE OVERVIEW

In the figure below we show the component diagram of the InGeoCLOUDS platform. A more comprehensive description of the InGeoCLOUDS platform is given in D3.2 [R4].

Regarding the deployment of Pilot 2, the Elastic DB, the Elastic File Server, the Elastic Web Server, the Elastic Map Server and the Data Integration & Linking components all exploit a dynamic pool of servers to sustain a possibly large volume of data requests (was limited to Elastic DB / File Server / Web Server components in Pilot1)

The OpenAM integration (IGC Management) now provides a complete solution to **manage users and security** in the InGeoCLOUDS platform. A Web console allows administrators managing users and security policies and monitoring sessions. The SPI (*Service Protocol Interface*) provides an authentication service that **allows implementing SSO** (Single Sign On) in the InGeoCLOUDS platform: the user connects once and accesses all applications and resources (depending on its rights) on a single virtual session. The SPI IdRepo (Id Repository, an OpenAM that manages all identities) allows applications to manage registered users. An LDAP directory contains all users. In Pilot2, a first implementation consists in the OpenAM-integrated LDAP server but migration to a more popular open source server (OpenLDAP) is planned by the Consortium.

The *geo-processing* API (Elastic Compute) provides a solution to manage on demand tasks into the InGeoCloudS platform. A specific set of APIs let a provider start an instance into the platform reacting to an event. The goal is to elaborate some data and then terminate the instance. For example, in the Shake-maps use case implemented by EPPO, strong motion recordings from the EPPO's Strong Motion Network

will be processed and a collection of ground motion and shaking intensity maps will be produced, in various formats, after significant earthquake events.

It is the responsibility of the provider to configure the instance for the on demand computation, to start it when it's required and stop it when the task ends.

The **Catalog** services are now fully integrated in the platform and can be used by Data Providers in conjunction with the setting of their OGC Web Services. It is based on the Geonetwork framework. SSO integration has been designed and integrated too for this component.

We deployed some **supervision tools** for a continuous monitoring of the platform. The proposed technical monitoring is defined in terms of indicators and alarms. The indicators and the collection of monitoring data was implemented within the Nagios framework. More details can be found in D3.3 document [R5]. A view on the user interface available to InGeoCloudS Administrators is presented in section 3.3.2 below.

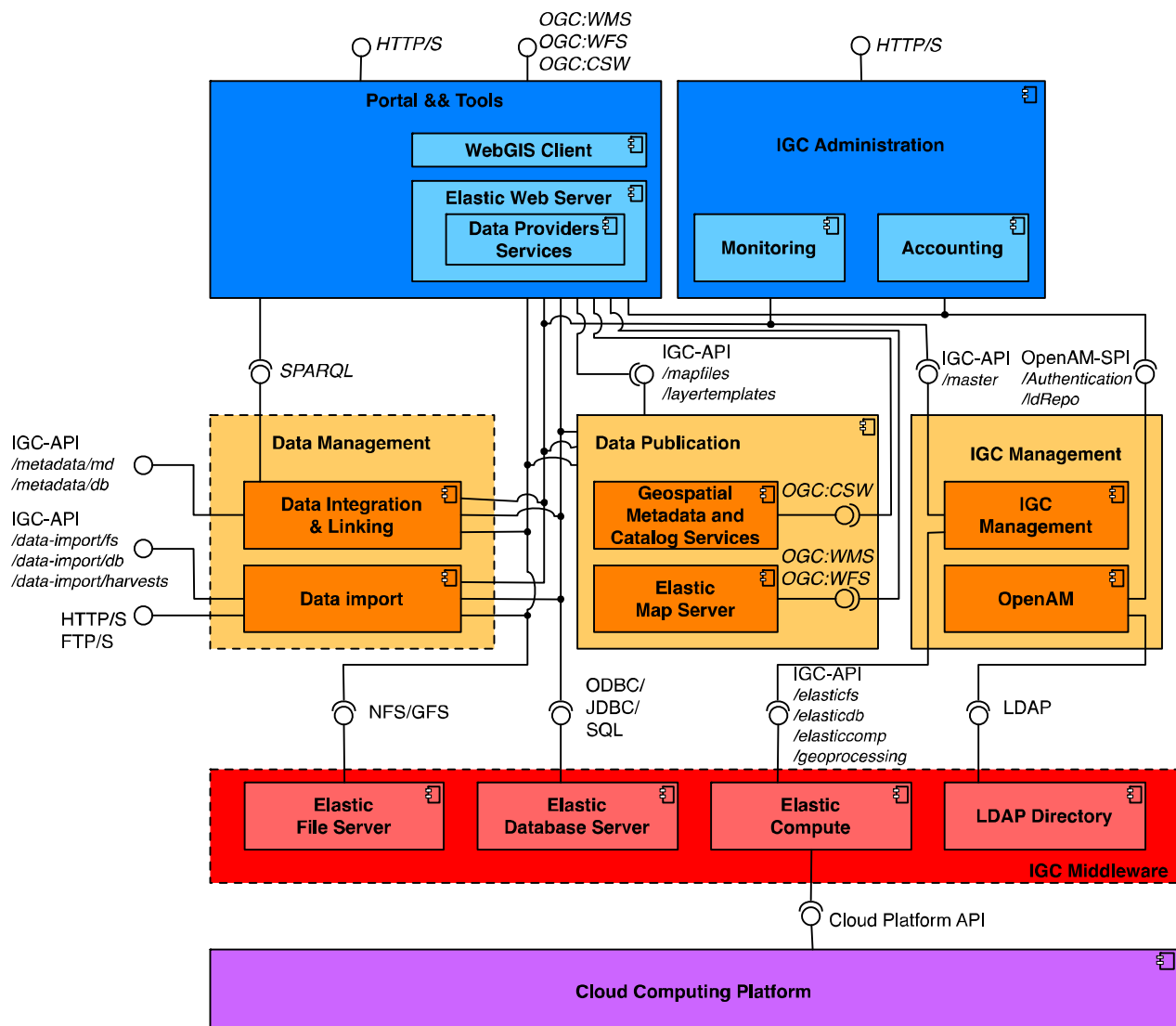


Figure 1 : Logical View of InGeoCloudS Architecture as in Pilot2

## **2.2. INGEOCLOUDS DATASETS**

### **2.2.1. GEUS DATA**

The second version of the InGeoClouds pilot does not include new datasets in the project.

The **GEUS groundwater dataset** consists of boreholes with a geographical location. For each borehole, there are classifications of the lithologic and lithostratigraphic layers the borehole penetrates. Within each borehole is one or more intakes from which groundwater can be extracted. All the intakes are hydrologically isolated from each other. Samples are periodically taken from each intake and analyzed for chemical and biological substances. In addition, the groundwater level for each intake is measured periodically.

### **2.2.2. GEO-ZS DATA**

The second version of the InGeoClouds pilot does not include new datasets in the project.

Aladin precipitation forecast are .asc files. All the other input data (precipitation values, landslide triggering threshold values) is raster format (GeoTiff or ArclInfo Binary Grid). The output format is also raster format (GeoTiff or ArclInfo Binary Grid). The final result is **WMS service with landslide prediction** created from previous input files.

### **2.2.3. BRGM DATA**

The second version of the InGeoClouds pilot does not include new datasets in the project.

Currently, the BRGM pushed 3 types of datasets in the project:

- the French groundwater database (10 Go) containing all groundwater quality measurements in France and groundwater level following by piezometric monitoring stations. The database contains more 20 millions of data and the content was indexed by FORTH in the context of the linked data process
- The French groundwater monitoring stations with their geospatial localization and in relation with the groundwater database. For some stations, the localization was blurred for security reasons.
- The geology dataset for France at the scale of 1 / 1 000 000 with information about lithology and age of rocks.

Moreover, the BRGM has developed and produced a WMS service as a “base” reference layer in Europe with the use of the OpenStreetMap data ([www.openstreetmap.org/](http://www.openstreetmap.org/)). The service could be used by all web GIS clients in the projet. The service has been pushed in the cloud infrastructure for the pilot 2.

### **2.2.4. EPPO DATA**

The second version of the InGeoClouds pilot does not include new datasets in the project.

The EPPO dataset includes ground motion and shaking intensity maps (shake-maps) for specific significant earthquakes. Shake-maps are grids of data of various physical parameters (like peak-ground acceleration, peak-ground velocity, intensity and spectral response) for an area around the earthquake source. For the calculation of each map the data used includes Greek Accelerometers Network descriptive data (station cites, sensors, geology), earthquake source data and strong motion data. Triggered by earthquakes, the accelerometers network produces strong motion recordings, which are then processed and used for the calculation of the shake-maps. The maps are produced in a variety of formats like images, geo-referenced maps and data grids.

### 2.2.5. EKBAA DATA

#### **GroundWater resources management in granular aquifers**

These data are used in UC5 (see [R3]). The use case provides data from both field measurements (groundwater table fluctuation, pumping tests and pollution sources) and chemical analyses (major ions and trace elements).

#### **Active Landslide Inventory Mapping and Susceptibility Zoning**

These data are used in UC6 (see [R3]). The use case provides data from both field measurements (groundwater table fluctuation, pumping tests and pollution sources) and chemical analyses (major ions and trace elements). The datasets include an inventory map of the occurred landslide events as a point feature class containing a database of their characteristics – properties (e.g. location, classification, year of occurrence, implications etc), updated after every new recorded event. The database contains data concerning the landslides' characteristics (type of movement, season and year of occurrence etc), as well as any information available for the region of occurrence (geology, precipitation, altitude, slope etc).

## 2.3. USERS

### 2.3.1. DATA PROVIDERS

No change introduced in Pilot2 for the Data Providers management. Their prerogatives have been enriched with the management of registered users (see section 2.3.2 below) and access to GeoCatalogue services.

A *Data Provider* is a user (representing for example an institution) willing to contribute with his own data or with a novel service. The registration process for data providers is supervised by an InGeoCloudS administrator who must check and authorize the *Data Provider* account request (approved by Steering Committee). The platform creates a *Workspace* for each *Data Provider*. A *Workspace* simply provides the possibility to store (and access) his own data in a private database of the InGeoCLOUDS Elastic Database Server, or in a private folder of the InGeoCLOUDS Elastic File Server.

Pilot2 now also features a harvesting API allowing DataProviders to implement and invoke data update commands/scripts. The API can be used either programmatically or through a GUI part of the Data Provider Toolkit.

### 2.3.2. REGISTERED USERS

Registered Users are visitors of the platform that went through the registration process. These users can be recognized by InGeoCLOUDS, and they can thus potentially exploit some services that are not made available to the anonymous web user. Examples could be personalization of user experience and privileged access to certain services (e.g. data download).

In Pilot2, the registration is moderated by Data Providers: each Data Provider owns a group of users and is able to register and manage their user accounts through the InGeoCloudS Administration interface. A same user (same login) can pertain to several Data Providers groups.

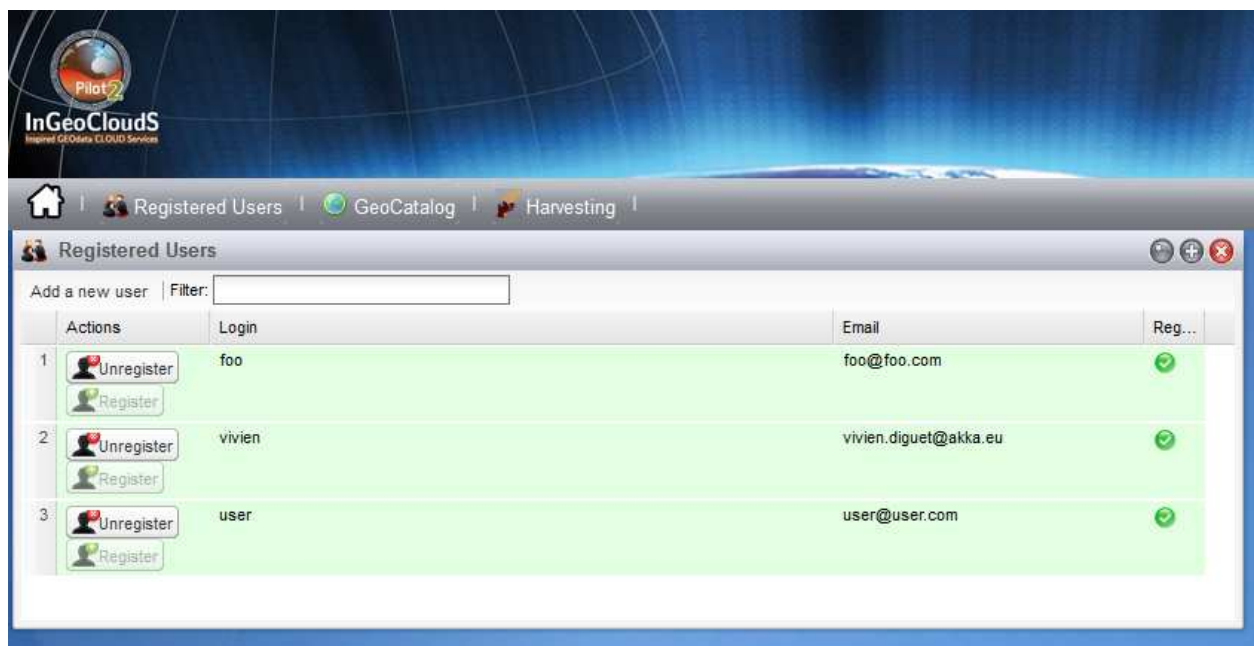


Figure 2 : Management of Registered Users by Data Providers (view from the DataProvidersToolkit)

### 2.3.3. PUBLIC

It designates all other types of users that browse the Portal without authentication/identification. In Pilot2, as it was the case in Pilot1, public users are allowed to access most of InGeoCloudS applications, maps, views and data without restrictions.

### 2.3.4. INGC ADMINISTRATION

INGC Administration groups a role of users that are able to both monitor the status of the infrastructure and also execute some management tasks on its different components (e.g. Projects integration in the Portal). The main facilities are described in chapter 3 below.

## 2.4. SERVICES USED BY DATA PROVIDERS

### 2.4.1. DATASETS PUSHED INTO INGEOCLOUDS

#### From BRGM:

The BRGM has created geospatial services with his datasets: groundwater monitoring station and geology layer at 1/ 1 000 000 scale.

#### From GEUS:

The GEUS groundwater database model has been extracted from GEUS internal database and implemented within the InGeoCloudS Elastic Database Service. A snapshot of the dataset has been extracted from GEUS and imported into the InGeoCloudS Elastic database service.

The data model for Pilot 1 has been extended with additional borehole information in Pilot 2. These additional data was necessary for the 'Borehole List' new functionality.

In Pilot2, the dataset is maintained by daily delta updates. The InGeoCloudS platform has scheduled this procedure and receives the delta updates by invocation of new harvesting services provided and hosted locally by GEUS. The new Update Log functionality shows the user about when the data was last successfully updated.

As for all other datasets, integration in GSOM and in InGeoCloudS triplestore is available (see section 2.4.2 below).

#### From GeoZS:

All the data is pushed to the cloud using new harvesting mechanisms available in Pilot2 (DataImport module):

- Precipitation forecast data (daily upload),
- Landslide susceptibility model (static: uploaded only once),
- Landslide triggering threshold values (static: uploaded only once).

After all data for predicting landslides is uploaded the data for creating WMS service needs to be uploaded (all the files that mapserver needs to create WMS service).

As for all other datasets, integration in the geo-scientific observation model (GSOM) presented in [R2] and the InGeoCloudS triplestore is available (see section 2.4.2 below).

#### From EPPO:

Data used for the calculation of shake-maps can be considered as “static” and “dynamically produced”. The first category includes Greek Accelerometers Network descriptive data (station cites, sensors, geology) which is uploaded in InGeoCloudS platform once and used for service setup. The second category includes earthquake source data and strong motion data. As soon as this data is prepared, after earthquake events, it is automatically uploaded to the InGeoCloudS platform through the *Data Import Service*. At the same time, data is pushed in the InGeoCloudS triplestore implemented by the *Data Integration&Linking* component. The upload event triggers the calculation of shake-map data which is also stored in the platform.

#### From EKBA:

##### Groundwater resources management in granular aquifers

Data about groundwater resources management in granular aquifers has been pushed in InGeoCloudS triplestore only, implemented by the *Data Integration&Linking* component. The data and its representation according to the geo-scientific observation model (GSOM) presented in [R2] are available for queries through the SmartQueries application (new in Pilot2) and also programmatically through the Linked Data Management API (see next section).

Active Landslide Inventory Mapping and Susceptibility Zoning. Data about Active Landslide Inventory Mapping has been pushed in InGeoCloudS triplestore only, implemented by the *Data Integration&Linking* component. The data and its representation according to the geo-scientific observation model (GSOM) presented in [R2] are available for queries through the SmartQueries application (new in Pilot2) and also programmatically through the Linked Data Management API (see next section). Data used are:

- Landslides Inventory map (active upload: after the occurrence of a new event),
- Landslide density map (active upload: recalculated after the occurrence of a new event)
- Landslide susceptibility model (static: uploaded only once),

#### 2.4.2. LINKED OPEN DATA SUPPORT

The GSOM meta-model developed during Pilot 1 has been extended to cover all possible themes corresponding to the data available from the data providers involved in the project; thus now it includes support for boreholes, groundwater data, landslides data (model has been extended) data description. Finally earthquake data descriptions were inserted and the model has been validated. The model extension followed the guideline of being as much as possible compliant to the corresponding INSPIRE model(s).

Concerning the mapping from the providers' datasets to the linked data, apart from the existing relational-to-linked data mapping mechanism, an additional method has been implemented in the Linked

Data Management API which allows mapping XML-based data sets to GSOM-compliant linked data through exploiting a XSLT mapping specification provided as input. In this way, there are now three methods through which linked data can be created and updated out of the original data of the data providers:

(a) the direct importing method where the data provider is responsible for creating the linked data from his/her original data, importing them in the underlying Triple Store as well as updating them through calling the respective methods of the Linked Data Management API,

(b) the indirect mapping method which exploits a R2RML mapping specification and the underlying mechanisms which guarantee that the relational data of the providers do not only need to the generation of the respective linked data but also any updates on the relational are reflected on the linked data, and

(c) the new indirect mapping method where a XSLT specification is exploited and run over a XML-based database in order to create the respective linked data of a data provider.

Please consider that in the latter method, specific synchronization mechanisms between the XML and linked data can be easily realized by just calling the appropriate Linked Data Management API method with the new XML data sets as input. Such mechanisms have been found adequate for the realization of the specific use cases which have selected to exploit the Linked Data Management API and concern the usage of XML-based data sets.

Apart from the ability to pose provider-specific, cross-provider as well as cross-theme queries, the Linked Data Management API was extended with new methods which:

(a) enable production of feature collection specifications in various formats (i.e., GeoJSON, SHAPE, GML, and KML) out of SPARQL query results,

(b) to pose geo-spatial queries in GeoSPARQL,

(c) to export theme-specific data in XML-based format following the INSPIRE model(s) by also enabling the capability to provide filtering constraints so as to reduce the amount of information that is returned by e.g. restraining the location of the features in a particular theme as well as constraining some characteristics regarding particular measurements for a specific theme, and

(d) to transform linked data representations from one format to the other. Through the offering of the above methods, not only the management of linked data has reached a good and sufficient level but also suitable and added-value functionality has been incorporated which can be immediately exploited by the respective use cases in the project.

#### 2.4.3. DATAPROVIDERS TOOLKIT



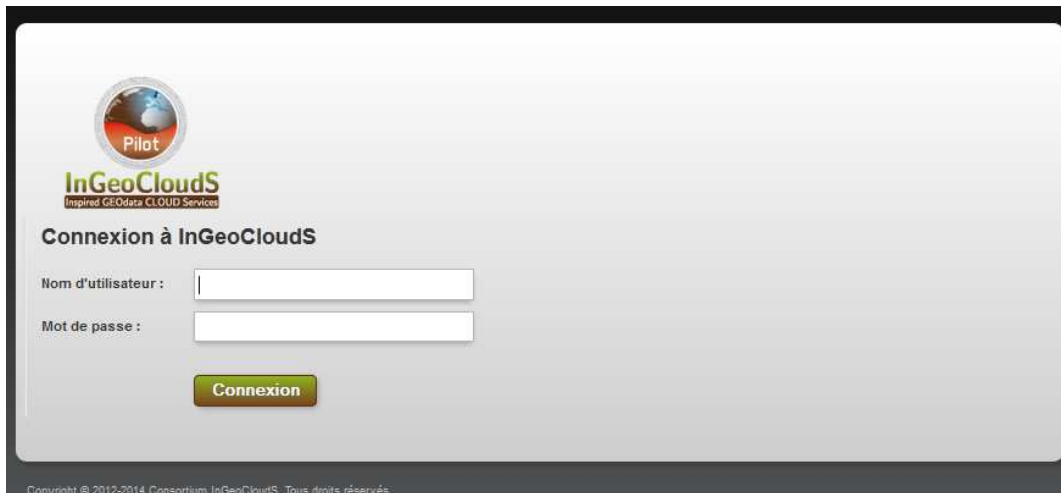
As a mean for grouping access to reserved applications/tools for data providers a *Toolkit* project has been integrated in the portal for Pilot2, giving access to:

- Management of Registered Users (see 2.3.2 above).
- GeoCatalog (Geonetwork) with catalogue edition roles.
- A GUI for CRUD operations on data harvesting tasks as specified by the *DataImport* module, thus facilitating overview and management.

## 2.5. SERVICES AVAILABLE FOR PUBLIC AND REGISTERED USERS

All users, Public or Registered, have access to the InGeoCloudS portal. From there, they are re-directed to the application of their choice.

### 2.5.1. PORTAL



*Figure 3 : New login page on the Portal*

The portal provides ordered access to the different domains of services. As explained in [R4] the portal is based on *Sitools2 framework*. The framework offers advanced back-office facilities for configuring and customizing the different “*Projects*” in the portal, help pages, simple roles management.

In Pilots, three (3) main scenarios are taken into consideration for integrating an existing service GUI into the Sitools2 portal framework:

- Integration of a “view” on an external application as an *iFrame*.
- Displaying background map and WMS-type layers on top
- More complete integration where application can be customized using panel templates developed and integrated in Sitools2 for integrating specific functionalities.

The following integration modes are used:

- iFrame integration: Shake-Maps and Web Mapping applications
- Background maps and WMS-type layers: Landslides susceptibility Maps
- Use of panels templates (more “complete” integration): Pesticides in GroundWater services

The Ground Water Resources Management in Granular Aquifers (UC5) and Landslides inventory mapping (UC6) are mainly served by the new *SmartQueries* application based on the Linked Open Data API (see sections 2.5.4, 2.5.5 and 2.5.6 below).

### 2.5.2. PESTICIDES IN GROUNDWATER

The GEUS application shows the locations of the boreholes with additional borehole information available as click info. One of the additional borehole information available is a graphical illustration of the lithologic layers through which the borehole passes. When the user selects compounds and/or rock types, boreholes where these are not found are removed from the graphical illustration.

The user is able to filter the shown boreholes even further by providing a time period and/or depth interval of interest. If one or more compounds are selected, the additional borehole information (click info) will also

contain a time series graph of the found amounts for each of the selected compounds. If the user ticks the water level measurement checkbox, only boreholes where groundwater level measurements are present will be shown. If water levels measurements are enabled, the click info will also contain a time series graph of the water level for each intake.

As new facilities in Pilot2, in addition to the graphical view, there is also a list view containing the exact same filters and displaying the boreholes in list form. The user can at any time jump from the graphical view to the list view maintaining the current filter settings – but not the other way. By going to list view, you remove the geometric constraints of the view port.

By selecting one or more boreholes – in graphical or list view – the user is able to download all the borehole data in csv format.

### 2.5.1. LANDSLIDES SUSCEPTIBILITY MAPS APPLICATION

GeoZS is building an early warning system that is based on: expert analysis of existing landslide events, geology (landslide triggering threshold values) and optimization of model with forecast calculations for known events in the past.

The system will predict (in a best possible way) the areas where the probability of triggering of landslides will be increased due to higher precipitation levels.

Our main goal is to make fully automated system (automatically pushing data in to the cloud, GIS modeling for predicting landslide events, creating WMS map services and displaying WMS services in

### 2.5.2. GEOPUBLICATION AND WEB MAPPING APPLICATION

The web mapping application developed by BRGM is integrated in the general portal as a link to the application. In the first version, the system allowed a register user (data-provider or others) pushing his dataset (using InGeoCloudS API) and to create a map, defines the interface of the map (button, icon...) and publish services with the web-mapping editor. The map could be shared with a simple link (URL).

The second part of the application allows for all users (public access) to display the interactive map with a “classical” web-mapping interface.

In the second version integrated in Pilot2, the Web Mapping application includes new functionalities, in particular:

- a download service as WFS service
- a download service as ATOM files as required by Inspire guideline for the download services.
- a link between metadata (see below) and the webmapping application with the capacity to tag each datasets and maps.
- New functions for the creator of the map and web services and improvements and bug corrections for the GeoPublication platform.
- The integration of the new version of the InGeoClouds API and the inGeoClouds authentication.

The second improvement of the web mapping application is the GeoCatalog. The GeoCatalog is a catalog defined to store and share metadata about each geospatial datasets and services published in the InGeoClouds project. The catalog is provided by the GeoNetwork opensource solution with an installation and a deployment in the cloud infrastructure and the InGeoClouds architecture. The GeoCatalog is available from the portal platform. Roles of InGeoCloudS users (Anonymous/Public, Registered users, Data Providers and Administrators) are mapped to those of geonetwork for giving access to their specific user interfaces and functions.

### 2.5.3. SHAKEMAPS APPLICATION

The Shake-Maps application is integrated in the Portal as an *iFrame*. It displays a chronologically ordered archive of shake-maps for the Greek Region. The user can locate the desired shake-map by browsing through the archive. For each shake-map the visitor can select to display maps of different physical

parameters (like peak-ground acceleration, peak-ground velocity, and intensity). Each map displays the distribution of the values of a specific parameter in an area around the earthquake source using iso curves. Moreover, the positions of the recording stations are drawn over the map and the user can select a station to see the specific recorded values.

The user can download shake-map data in various formats which include images, raw grid data in xml and ascii and GIS files (shape files, KML and ESRI raster files). Additionally, for each shake-map the user can download data about the recording network in ascii and xml format and metadata about the processing / production process.

The shake-maps application calculates and publishes a new shake-map each time data from a new significant earthquake becomes available. The new data is imported through the *Data Import* module and stored in the *Elastic File System* in EPPO's workspace. The calculation is performed by an instance managed via the InGeoCloudS geo-processing facility (see section 2.1) and the results are stored again in the *Elastic File System*. Finally, the new shake-map data is published through the InGeoCloudS *Web Server*.

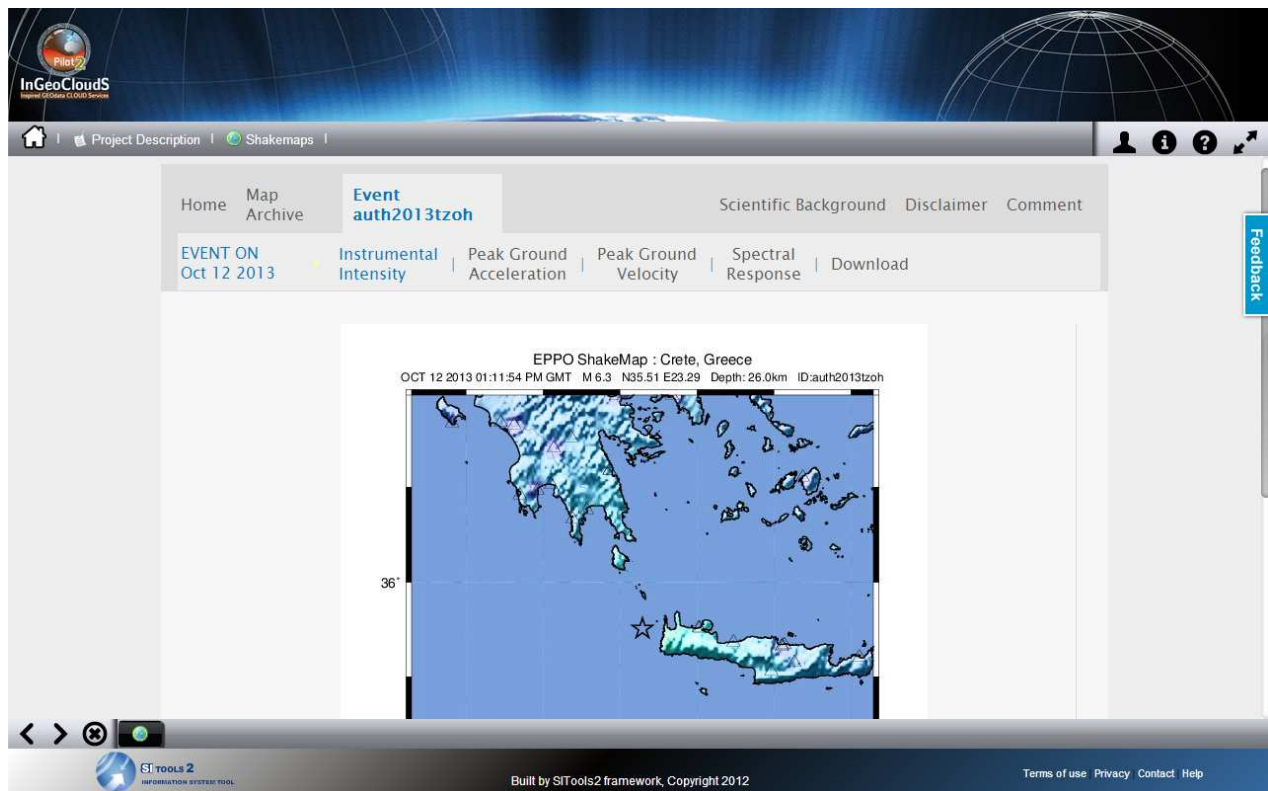


Figure 4 : Sample screenshot of the “Shake-Maps” Application

#### 2.5.4. GROUNDWATER RESOURCES MANAGEMENT IN GRANULAR AQUIFERS

The EKBA application shows the locations of the boreholes with additional borehole information available as click info. One of the additional borehole information available is the geological formation, the X-Y coordinates in WGS84 (geographical info), the drilling depth, the aquifer and the water district. Also, other useful information is stored such as the pumping type and the topology (e.g. valley, hill). Moreover, chemical analyses are linked to boreholes showing the concentration of several chemical compounds. When the user selects compounds, boreholes where these are found are highlighted. End user is able to make different queries such as the chemical compound's concentration within different time periods (range, up to, below or above a threshold value etc) the number of boreholes within a rectangle, circle or polygon, the drilling geological formation for every available borehole, the groundwater table (range, up to, below or above a threshold value etc). Finally, if the end user enables the groundwater table layer, the

boreholes where groundwater level measurements exist will be highlighted. By selecting one or more boreholes, the user is able to download all the borehole data in csv format.

### 2.5.5. ACTIVE LANDSLIDE INVENTORY MAPPING AND SUSCEPTIBILITY ZONING

Taking into account the data in the InGeoCloudS triplestore, a division of the land into homogeneous areas is produced according to the landslides' occurrence percentage (Landslides' density map). The calculation of the spatial probability created a susceptibility zoning map. The map was based on the fact that landslides in the future will occur under the same circumstances that they occurred in the past. Maps used for the above analysis are also available (e.g. digital terrain model, land use map etc). All input data except inventory map are in raster format. The output format is vector (point feature class for the landslides) or raster format (for density, susceptibility and background maps). The final result is WFS service for the inventory map and WMS service for landslides' density and susceptibility maps.

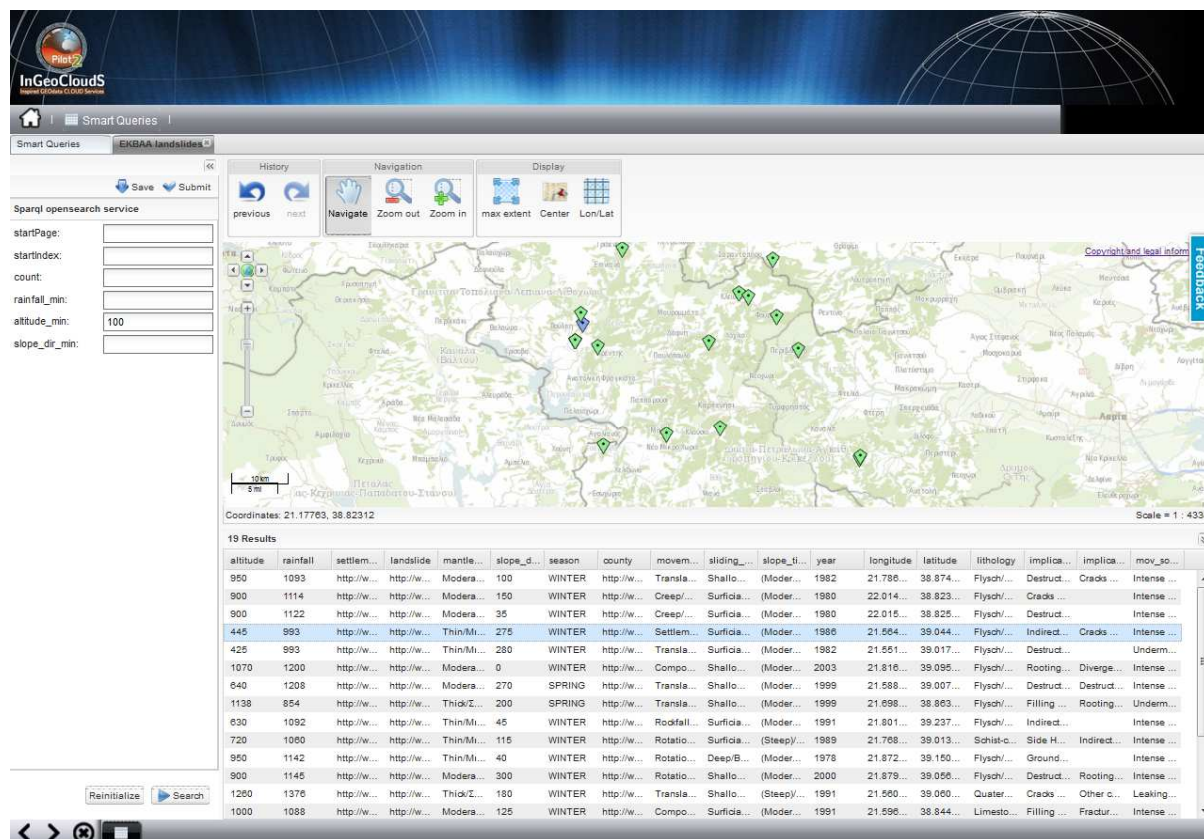
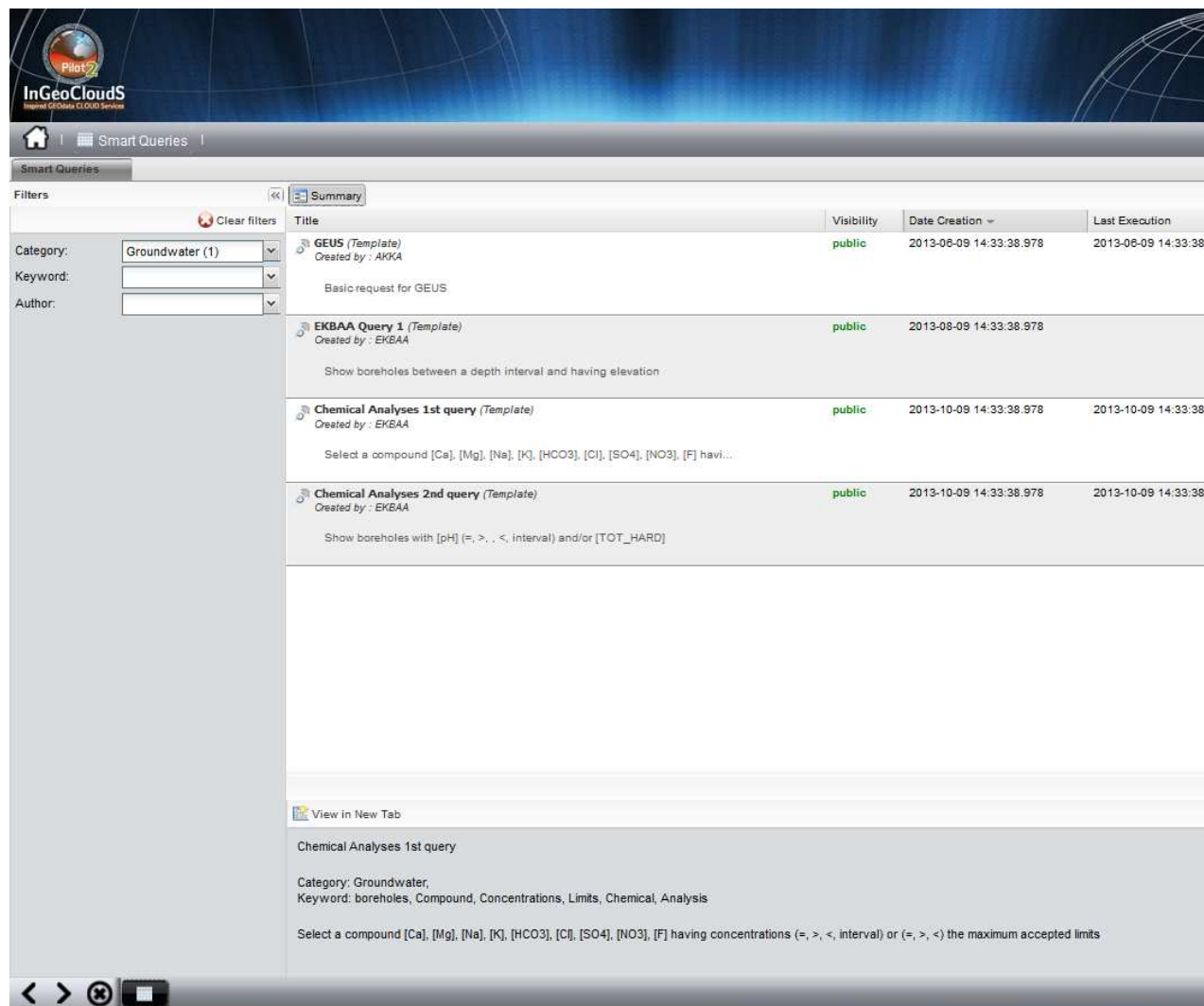


Figure 5 : SmartQueries App: example of query illustrating UC6

### 2.5.6. SMARTQUERIES APPLICATION

The “Smart Queries” is a new application for public and registered users in Pilot2.



*Figure 6 : View of the SmartQueries App*

It provides access to a database of pre-defined queries that can be made on the whole set of data currently handled by InGeoCloudS in its triplestore following GSOM model. In particular, it provides implementation of EKBAQ’s use cases UC5 and UC6 as described in D2.3 [R3].

Behind the scene, queries are saved using standard Opensearch<sup>1</sup> formats and are executed on the InGeoCloudS triplestore using LinkedData management API (see section 2.4.2). It thus hides SPARQL syntax complexity to end-users while providing optimization and more efficient execution.

The Opensearch/SPARQL support that currently exists in the InGeoCloudS portal has the following functionalities available to all users:

<sup>1</sup>

<http://www.opensearch.org/Home>

- Query using a simple form
- Display results on a map and a in grid
- For Registered Users and data providers only: Save query parameters on the form of a new, private "Smart Query"
- For Data Providers only: possibility to make private Smart Queries public.

Opensearch is a collection of simple formats for sharing search results. The opensearch description document is a simple XML file which describes the way to make queries including query parameters, type of document retrieved and other metadata about the search engine.

The smart queries module asks for the opensearch.xml file. That file contains the URL template of the query with the parameters.

The module creates a simple form from those parameters and the user can enter unformatted values. The SITools2 server transforms that request into calls to the LinkedData management API and returns the result as GeoJSON which is displayed on a map.

### **3. OPERATING PILOT2**

New Utilities Projects in the Portal dedicated to Administrators and Data Providers

#### **3.1. ADMINISTRATION TOOLKIT**



*Figure 7 : View of the SmartQueries App*

#### **3.2. USER MANAGEMENT**

Users are POSIX accounts managed in a LDAP directory.

User management in Pilot2 relies on OpenAM. This open source software provides a Web console to manage users and groups. Moreover, exposed RESTful services allow implementation of Single Sign On (SSO) in the platform. A connected user is identified by a token. All applications use the SSO token to ensure the user is connected and to retrieve user information. LDAP directory is not directly exposed anymore.

Full documentation of the API is available at <http://ingeoclouds-api.isti.cnr.it/>

Not done in Pilot2 at time of writing: User related services will also incorporate subscriptions and notifications mechanisms as specified in [R2].

##### **3.2.1. DATA PROVIDERS**

No major changes in Pilot2:

The Pilot allows managing data providers using RESTful web services of the master API additionally to user management console available in the *Administration toolkit*. The API ensures user's workspace and user's database are created or deleted together with the data provider account. Creation, Deletion and Getting information methods are available.

##### **3.2.1. REGISTERED USERS**

InGeoCloudS support for data providers' applications the management of registered users. Pilot 2 allows managing registered users using the master API on one hand and a tool made available to data providers in the *Data Providers Toolkit*. The tool relies on the API. The API ensures for example that a user is not created multiple times by different data providers.

##### **3.2.2. PUBLIC (ANONYMOUS USERS)**

In Pilot2 anonymous users can access the InGeoCloudS portal without restrictions.

The Data Publication application, Data Providers toolkit (Management of registered users, Geo Catalogue, Monitoring) and Administration tools require authentication.

### 3.2.3. INGC ADMINISTRATORS

Pilot 2 does not provide any service or API to manage INGC Administrators programmatically: their management is available through usage of specific (and hidden) services of the User Management module of INGC Administration. As any kind of user, INGC administrators can be added, modified and deleted using the OpenAM Web console.

For the moment, a single INGC administrator exists in the platform to allow accessing administration services.

### 3.3. SUPERVISION OF SYSTEM HEALTH

Amazon provides useful services to supervise system health: monitoring EC2 instances with *CloudWatch* and ensuring reliability and stability of the system with *Elastic Load Balancers*.

Additionally, InGeoCloudS provides monitoring services to focus on specificities of the platform, like fine-grained monitoring of **Elastic Web Server** or **Elastic DBMS**. See section **Erreur ! Source du renvoi introuvable.** for an overview.

#### 3.3.1. AWS MANAGEMENT CONSOLE

An *AWS Identity & Access Management (IAM)* is required to access the *AWS Management Console*. You must sign in with user name and password assigned by the system administrator of the InGeoCloudS *AWS account*. Note that this user is not managed by InGeoCloudS, but by Amazon. The *EC2 Dashboard* summarizes the EC2 resources used by the InGeoCloudS *AWS account*. From there the supervisor can select 'Instances' that allows listing all EC2 instances. Filter with '*Pilot2*' keyword to see only the instances running the Pilot 2 platform. All instances have a distinguishing name.

CloudWatch is a monitoring solution for the amazon cloud resources. It gives tools to developers and system administrators to collect and track metrics so that they can instantly react to unexpected problems or load peaks to keep their applications running smoothly. The tools gives you access to the service via API, command line tools or AWS management console. In the latter case, for any EC2 instance, a 'Monitoring' tab displays alarms and a graphical representation of the metrics defined for the EC2 instance.

#### 3.3.2. INGEOCLOUDS SPECIFIC MONITORING AND SUPERVISION

The InGeoCloudS API includes specific methods for querying the current status of every service running on the platform. The RESTful web services developed for the API (<http://ingeoclouds-api.isti.cnr.it/>) include specific methods we developed for precise and customized monitoring of the technical components.

As examples:

- **/geoprocessing/status** allows getting information on a specific geo-processing instance running in the cloud.
- **/elasticsearch/status** method returns a description of the current status of the elastic file system, including storage capacity and storage used, number of servers and some other information on a provider basis.
- **/elasticsearch/status** returns various information about the database status including Number of DB exploited servers, current size of all databases (sum of all tables and all indexes), number of users created in Postgres Server
- **/elasticsearch/status** returns the number of servers used by the service (healthy and unhealthy instances) as well as the hostname of the service (so that a client can query it).
- Etc.

We extended such functionalities and integrated them in a comprehensive monitoring framework based on Nagios. D3.3 (ref: [R5]) describes monitoring and supervision put in place in Pilot 2. The proposed technical monitoring is defined in terms of indicators and alarms. Each indicator measures the status of the platform according to a specific point of view, while an alarm signifies an event that requires some

intervention or further investigation (e.g., excessively large response time). For each alarm, we discuss the possible correcting actions to be undertaken in order to adjust the behaviour of the platform. The set of indicators allow a continuous monitoring of the InGeoCloudS platform and of its software components. The set of alarms defines the proposed maintenance plan, i.e., the correcting actions to be undertaken over time. See D3.3 [R5] for more details.

The Administration Toolkit integrated provides access to the supervision dashboard (nagios WebClient) with comprehensive views about the infrastructure health.

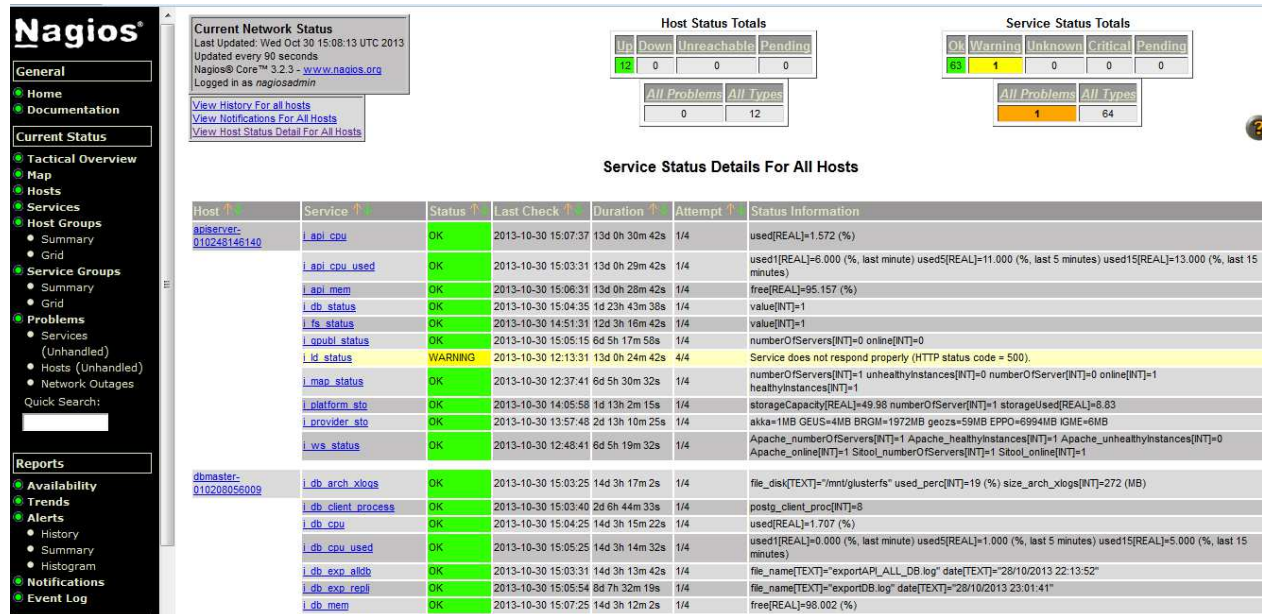


Figure 8 : Example of dashboard on InGeoCloudS specific indicators.

### 3.3.3. NEW LOAD BALANCERS

Load balancer is a core element in a cloud infrastructure. It provides a solution to distribute the incoming traffic of a service across a pool of EC2 instances hosting the same application, as well as to detect unhealthy instances inside the pool rerouting consequently the request to healthy instances only. By doing that, a load balancer prevents any server from becoming a single point of failure, increasing the application availability and responsiveness. Load balancer also improves the resource utilization.

Pilot2 now integrates load balancing on several of the main services:

- Elastic WebServer
- ElasticDB
- ElasticFS
- Elastic Map Server
- Data Integration and Linking

### 3.4. SUPPORT AND HELPDESK

No major change on that level: users and partners are provided with several tools for reporting issues or for getting information about usage of Pilot2:

- An email address for helpdesk: [ingeoclouds-support@ingeoclouds.eu](mailto:ingeoclouds-support@ingeoclouds.eu)
- A feedback form (*credits to feedebuz.de for their free widget*) always accessible in the portal that allows to easily acquiring help or sending a request to InGeoCloudS team.

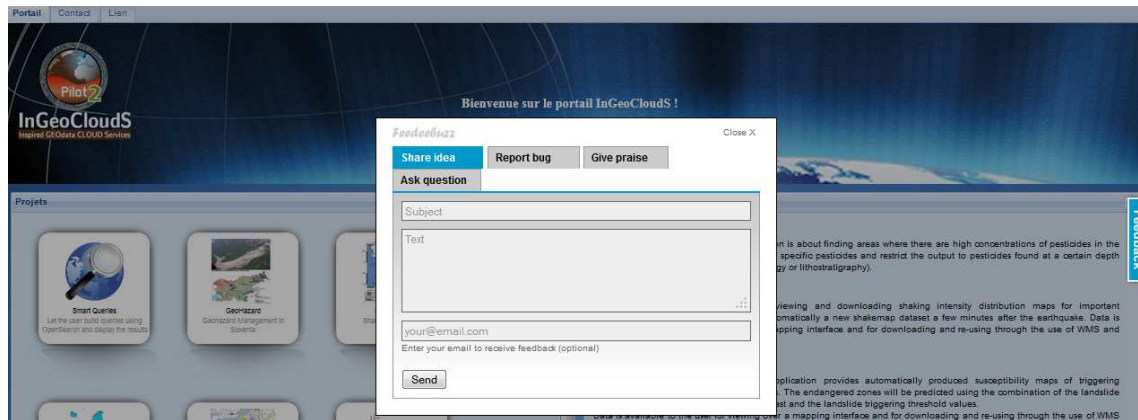


Figure 9 : Facility for leaving a message to InGeoCloudS helpdesk from the Portal

- Help pages directly accessible from the Pilot's portal pages <http://portal.ingeoclouds.eu>. The help material is incrementally populated with reference pages, FAQs, documents to be downloaded etc.
- A bug and issues management tool available at: <http://atlas.akka.eu>. This tool is open to project's partners only: it is integrated in support/helpdesk processes as explained in D3.3 chapter 4.

\*\*\* End of the document \*\*\*