Open call No EEA/MDI/14/005

ANNEX 8 to the tender specifications

Copernicus
Reference Data Access Node

CONCEPT PAPER

v.1.0

May 2014

Table of Contents

1. Introduction

1.1. Background

Copernicus is an EU-wide flagship programme for monitoring the Earth. Previously known as Global Monitoring for Environment and Security (GMES), it is coordinated and managed by the European Commission.

Copernicus integrates satellite and in-situ data with modelling to provide reliable and up-to-date information through a set of user-focused services. The services address six thematic areas: land, marine, atmosphere, climate change, emergency management and security. They support a wide range of applications, including environment protection, management of urban areas, regional and local planning, agriculture, forestry, fisheries, health, transport, climate change, sustainable development, civil protection and tourism.

Some services are already operational (land monitoring and emergency management) while others are still in a pre-operational mode (atmosphere monitoring and marine monitoring) or in a development phase (climate change monitoring and services for security applications).

While the technical implementation of operational Copernicus services is coordinated by EU bodies, service provision is contracted to the consortia of private companies, called "service providers".

Use of the Copernicus services is an integrated part of EEA's strategy to improve environmental information. The EEA plays a key role in the operation of the Copernicus services, in particular in the technical coordination and implementation of the Pan-European and Local component of the Land Monitoring Service and of the Copernicus Reference Data Access component. Under the Copernicus regulation[1] the Copernicus services cross-cutting in-situ coordination will be delegated to the EEA as part of the Copernicus programme in 2014-2020. In-situ data have been defined for the Copernicus programme as: "*all non-space-born data with a geographic dimension, including observation data from ground-, sea- or air-borne sensors as well as reference and ancillary data licensed or provided for use in Copernicus*". Geospatial reference data are a special category of in-situ data for Copernicus services providing a geographic framework to which other required in-situ data are referenced and maintained. Reference data are required by Copernicus services for creation, verification and validation of information products and services derived from satellite images.

Reference data in Europe are owned and hosted by national or regional authorities under differing specific license conditions and accessible through a variety of interfaces (e.g. national or regional INSPIRE spatial data infrastructures, institutional portals and other sources) in a number of data formats. In order to effectively incorporate these diverse data sources into the Copernicus platform, considerable resources and planning is required.

> COpernicus Reference Data Access (CORDA) is proposed as a single entry-point node hosted and maintained at EEA to facilitate quick and easy exploration and access to the national and/or regional geospatial reference data in EEA39 countries that have established free of charge web services (either fully free and open or free of charge but with restricted access).

[1] Regulation (EU) No 377/2014 of the European Parliament and of the Council of 3 April 2014 establishing the Copernicus Programme and repealing Regulation (EU) No 911/2010

It is expected that in the beginning this node will provide access to the pre-selected national and/or regional reference datasets linked to INSPIRE Annex I and II (addressing essential reference data needs of the Copernicus Land Monitoring and Emergency Management services), however the scope of the geospatial reference data accessible through the node operated by EEA will be extended over coming years as more data will become available free of charge for Copernicus services (as a result of both the progress in the implementation of INSPIRE, and EEA's in-situ coordination activities) and addressing as well the in-situ data needs of other Copernicus services (marine, atmosphere, climate change and security), hence linking to INSPIRE Annex III as well.

1.2. Added Value

CORDA should make Copernicus data exploration and access quick and easy for end users, and provide a number of additional benefits compared to data access methods in use today. Implementation of such a system must give direct benefits for solving current problems of data access, whilst maintaining the ability to evolve so that future data access challenges can also be met. As an example, many of the Copernicus datasets including Land Monitoring products such as Corine Land Cover have a temporal component, and as such are reproduced periodically to include the latest land cover data. Until now this reproduction has been a timely and costly process. CORDA aims to make recalculation of such products ever easier and more efficient as it will provide sustainable access to the reference datasets required for each production cycle. This extra dimension separates systematic data processing from one-time data production methodologies.

The system design needs to provide the following benefits:

Simplicity of use

Today end users require access to multiple national and/or regional reference datasets and services hosted by a number of data providers. Collecting these multiple datasets is no easy task, and often proves very time consuming (requiring licenses, usernames and passwords for a multitude of organisations). CORDA will provide a simple, single access point to collected national and regional reference datasets, giving each individual end user access to the latest available data through a single access node. This simple approach means end users will be able to direct their time and efforts so using the data, rather than struggling for hours to acquire it. Another advantage of this simplicity is that end users can then easily provide links to these datasets over the entire team to other end users.

High accessibility to CORDA

End users have well educated staff and develop technical skills based on the software they master. Therefore it is important to ensure that end users can harness these skills through their desired software for use in conjunction with CORDA, rather than having to learn new software. The same is true for data providers. Therefore CORDA should be accessible through any GIS software required. This is in line with other frameworks such as Inspire GeoPortal, ArcGis Online and the future ELF (European Location Framework) http://www.elfproject.eu/.

Round-the-clock availability

Amongst other Copernicus services the CORDA system will provide data for Copernicus Emergency Management service; therefore data should be available for access in the event of network problems at the data providers' side, or even in case of a major disaster, when master data sources may not be available. Therefore a data download should be available twenty-four hours a day, seven days a week, even in the event that data providers are temporarily offline. Having a backup capability is therefore very important to CORDA.

High reliability, efficiency and sustainability through centralisation

In the context of CORDA, centralisation means concentrating the system processes and control within the EEA. Centralisation will guarantee that end users are given access to verified, timely and authoritative geospatial reference data. Currently during every product update, there is a potential risk of losing knowledge of what reference data was used to improve or validate the product dataset generated in Copernicus. Centralization and monitoring will help to improve awareness of data gaps at data sources, such as broken links to data. Centralization will increase productivity by making data searching more efficient for end users, as everything will be discoverable in one place. This will reduce cost and time on product deliveries. Sustainability will be ensured through EEA management of reference data access between new contractors and update cycles.

1.3. Document Goals

The aim of this document is to describe our vision of CORDA, addressing various aspects of the design, architecture and functionality which we see as key to providing added user benefits compared to the existing methods for Copernicus data access. The document will include a series of user scenarios to help set out our vision.

2. CORDA System Description

2.1. System Objective

The above-mentioned benefits such as simplicity of use, high accessibility, round-the-clock availability, high reliability, efficiency and sustainability must be incorporated into CORDA for a successful system implementation. Translating these benefits in technical terms means that we need to build a single and secure access point to the required national, regional and European reference datasets for end users. The system should be developed on the server tier and be transparent for GIS software used by end users and data providers.

2.2. System Scope

Figure 1 provides an overview of the CORDA system (content of red boxes is considered within the system scope).

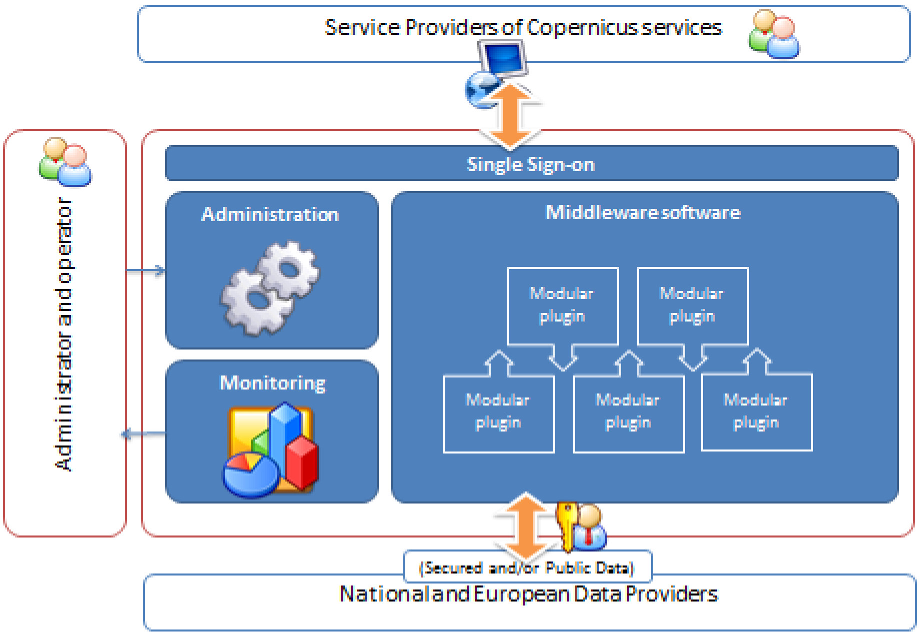


Figure 1 - CORDA System Overview

2.2.1. Administration of the System

> The operator of the CORDA system will play a key role in the maintenance of reference datasets. Round-the-clock system monitoring will ensure provision of a stable and reliable platform for end users

Providing easy and coherent access to a large number of reference datasets from a collection of data providers is no small task. Therefore a permanent operator at the EEA will be required to manage, prepare and structure datasets for CORDA end users. A single pan-European reference dataset can consist of many individual components supplied by data providers, who will provide us with access points to the data (such as web services). Over time these access points might change or get restructured. An operator should be able to monitor these changes, and contact the data providers when access points fail. Other tasks for the operator will include searching, collecting and improving the existing reference datasets, to close data gaps and update CORDA when improved or updated information is made available. The work carried out by the operator of CORDA is therefore to be seen in the context of EEA's Copernicus in-situ coordination activities.

2.2.2. Monitoring

Constant system monitoring is required to ensure a stable and efficient platform. Good system monitoring techniques enable administrators to focus on improving areas of high demand or low performance, and to predict and solve problems before they occur. Therefore CORDA will have an in-built monitoring capacity in the form of a monitoring module, designed to constantly assess connectivity and performance at access points and data providers. The monitoring module will also gather user request statistics to help build up an understanding of who is using which datasets, at which times, and in which capacities. This will provide Copernicus with important information about the relative importance of every reference dataset hosted in the environment. The CORDA operator is the main user of these produced statistics.

2.2.3. Single Login

Each end user or data provider will register on CORDA, providing a login and password. The system should handle these credentials centrally and give the end user one single login to access all available data in CORDA. This single log in concept is designed to make life easier for users of the system, and to maximise accessibility and transparency, which are important added benefits of CORDA, compared to existing data access methods. Roles inside CORDA could be introduced to fine tune access to reference data for different users. This will be increasingly important for security products where some data may be restricted. The data provider's login and password for reference data would be stored securely inside CORDA and only accessible by the CORDA operator.

2.2.4. Middleware Software

Middleware is the heart of the CORDA and needs to perform a number of advanced processes for key functionalities, and to resolve some of the technical problems we will face (see Figure 2). The following example illustrates how middleware can be utilized pragmatically to solve a technical problem.

Each reference dataset should have one single access point, though technically a reference dataset may be compiled from multiple datasets combined together from several access points, to create a single mosaic dataset (See Figure 2). In this example, each dataset is retrieved directly from its respective data provider, to assure that the most up-to-date dataset is used for the reference mosaic. The more INSPIRE compliant the data providers are, the lower the technical barriers will be to creating this kind of on-the-fly operation. However we foresee that there may often be some challenges which cause integration issues such as the data being produced in non-compliant national or European projections, or compatibility issues between web access points (web protocols might be different in version or type). CORDA should allow end users to seamlessly integrate with the reference datasets provided through CORDA using their own GIS software of choice, and harmonisations of these technicalities is very important through middleware.

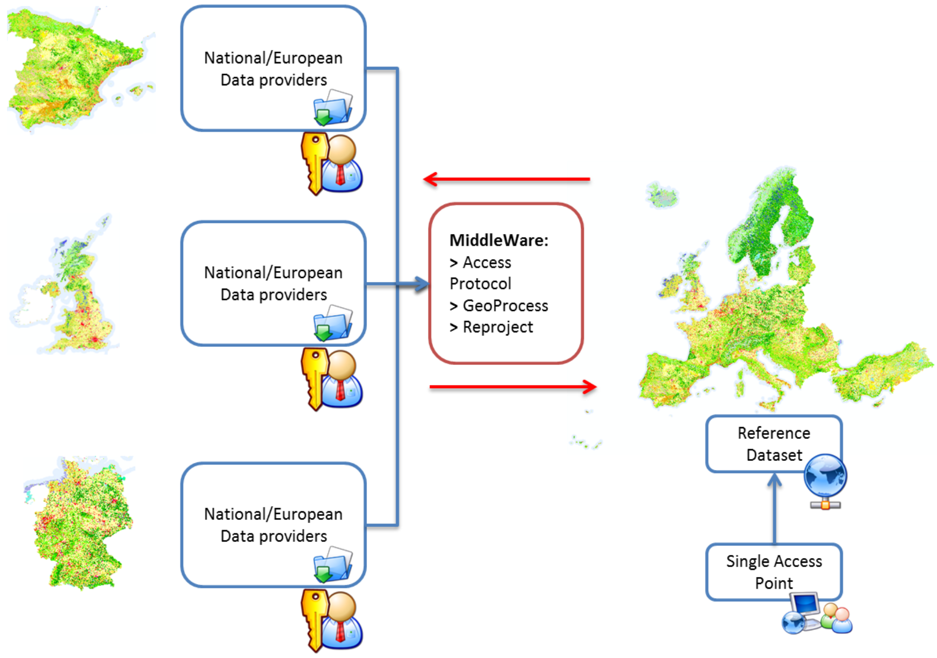


Figure 2 - CORDA Middleware used to create Reference Mosaic Dataset

It is important to note however that whilst CORDA will be very effective in creating seamless mosaics and mash-ups of reference datasets, it is **not** under the CORDA mandate to harmonise data, resulting in the creation of new source datasets. We see this as being one of the key functions of the upcoming ELF project, and we are confident that ELF will enhance many of the source datasets being explored and used through the CORDA platform.

It is anticipated that end users will use locally installed GIS software for data visualization, analysis and modification. Data will be fed to CORDA without any software implementation by national, regional or European dataset providers.

2.2.5. System Design

2.2.5.1. Modularity

Implementing a modular design, the complex CORDA system will be able to function as independent, modular parts to allow simple system implementation. Modularity will also make the system easier to maintain, and provide other benefits such as progressive implementation and bolt-on functionalities, as well as keeping the schema as simple as possible.

Implementations of some core system modules that compound the CORDA system are mandatory to obtain minimum functionality. For example, a security module for single login, and a data access module to query data repositories is necessary to achieve the minimum goals of the CORDA system – i.e. providing access for Copernicus end users to the required national, regional and European reference datasets.

The anticipated evolution of Copernicus will create new, sometimes unforeseen demands, which should be met through the CORDA system. Therefore it is vital that extra functionality can be ‘bolted on' by adding extra code modules to the CORDA core. With this in mind developers should adhere to good coding practices and transparent methods, and follow open standards where applicable.

Other modules could be implemented at later stages. For example, a geoprocessing module to transform data formats, or a language module to handle metadata language, might be considered as very important in the long term but not immediately needed to achieve the minimum goal of the system.

2.2.5.2. Flexibility

Flexibility will enable CORDA to evolve and adapt easily to future needs and requirements placed upon it by Copernicus. To achieve flexibility, we should aim to be as technologically transparent where possible using our development methods. Flexibility also means that CORDA should be accessible through a wide range of commonly used GIS software.

2.2.5.3. Scalability

Scalability is an important aspect of CORDA design. Good implementation of scalability will allow CORDA to effectively and efficiently deal with the expanding resources, users and demands that are foreseen for Copernicus. Although in the beginning CORDA will focus on meeting reference data needs for Copernicus Land Monitoring and Emergency Management services, other Copernicus services may become part of the CORDA system in the future. It is important that any required web services can be added to CORDA. If new functionality becomes necessary, adding new modules will not affect the existing system. It is our vision that CORDA should be entirely focused on the needs of Copernicus; it is not within the CORDA to extend to support other programs or projects.

2.2.6. Architecture

The CORDA system relies heavily on web services, so it will be designed as a modular **service oriented** and **server side system** compromising of:

- a distributed server infrastructure to host all software modules
- a Single Sign-On system
- backup infrastructure to host replicated data for Emergency Management service

Service Oriented Architecture (SOA) is a well-known IT architecture and aligns well with the modular, reusable and scalable system approach described above. SOA allows creation of highly scalable information systems and provides a well-defined form of exposure and calling of services (commonly web services), which will facilitate interaction with other systems (for example, national and/or regional SDI OGC services) or clients (platform independent, for example ArcMap or QGIS).

Web services are extremely versatile in a constantly changing technological world. They simplify complete separation between presentation of information, and actual data layers, which facilitates integration with any client. Web services also benefit end users because they have a fast response to required and requested changes.

Figure 3 illustrates the architecture of the CORDA system, which can be presented in the usual three layer architecture:

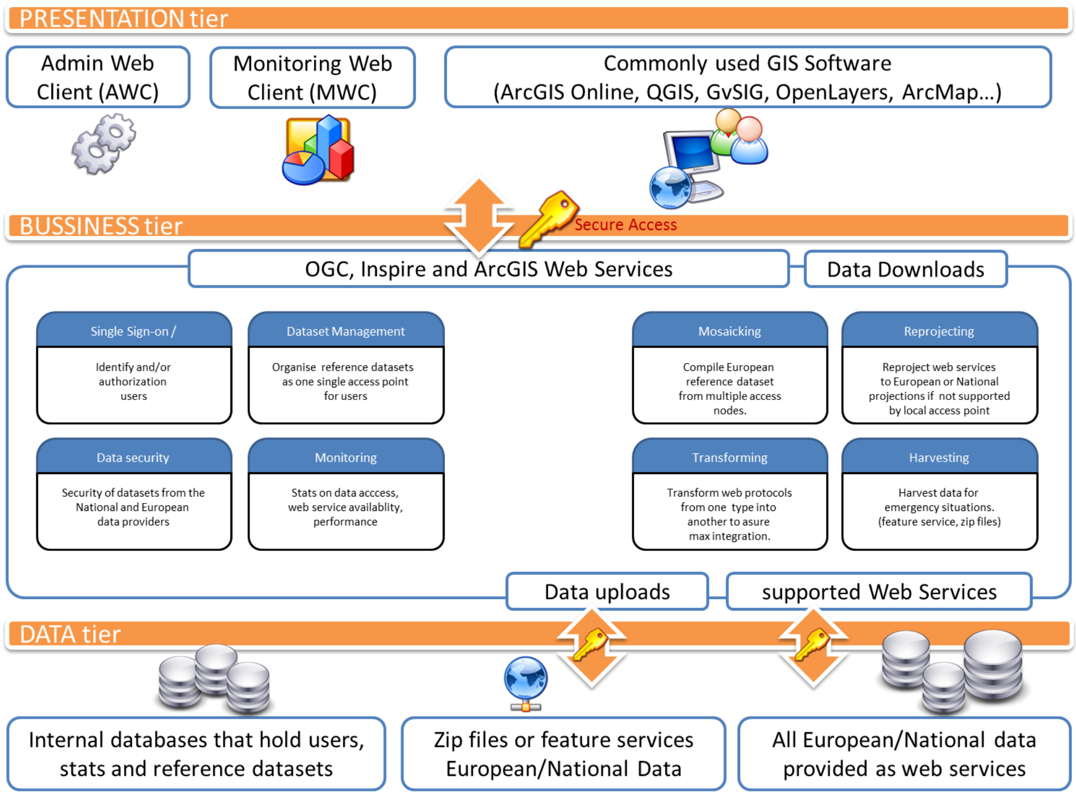


Figure 3 - CORDA System Architecture

2.2.6.1. Presentation Tier

End users will simply require an internet connection to access the CORDA platform from the GIS software of their choice, which is in-line with the INSPIRE directive. End users should be able to interact with web map services to visualize or even download geospatial reference data to desktop in order to work with it locally should this be necessary. This will provide each end user with the maximum flexibility and capabilities for their specific needs. During the CORDA system implementation, access from various desktop and/or web GIS clients will be demonstrated.

The CORDA Administration client is a web based application (please refer to the Administration module for more information) to be used by the CORDA Operator.

The CORDA Monitor tool is a web mapping application (please refer to the Monitoring module for more information) to visualize system, user and data statistics to allow monitoring and improvement of CORDA workflows.

> Presentation of spatial data layers of the CORDA system will be integrated inside commonly used desktop GIS software clients as used by end users including ESRI, MapInfo, QGIS, gvSIG, and commonly used web GIS clients such as ArcGIS Online or Open Layers.

2.2.6.2. Business Tier

CORDA will be hosted and maintained inside the EEA SDI (Spatial Data Infrastructure) as a fully integrated production environment. For this reason technology choices on one hand should follow the technological developments and on the other hand be compatible with the existing infrastructure at the EEA. This assures high level knowledge of the given platforms, and should allow for easy integration due creating low inertia to adoption. Based on the key functionality required the business logic rules will be best implemented on the server side.

In utilising SOA architecture the aim is to ease development costs by enabling platform independent software modules to communicate where possible using web services. Where non-existent server side modules and business logic need to be developed, standard frameworks will be used, for example .Net or Java.

REST (REpresentational State Transfer) protocol will be used to communicate between modules and expose and call service functionalities wherever possible. This approach will facilitate the addition of new capabilities, and integration of other Copernicus services, with little or no cost to the existing CORDA system.

Functionality in this business tier will include (but not be limited to) retrieving the appropriate data for a precise product and region, preparing and elaborating the information in the most suitable format, automation of data updates, and enabling end users to identify and report missing data.

The INSPIRE directive features heavily in CORDA design and implementation, which will focus on enrichment and utilization of well-known implemented INSPIRE policies, format rules and protocols. OGC standard geographic services and integration with ARCGIS REST API will also be considered in CORDA design to ensure maximum compliance with data and technological standards.

2.2.6.3. Data Tier

The aim of the CORDA is to provide a well-defined operational system that can be used by end users, as opposed to being a generic "all-purpose" system. As shown in Figure 3, geospatial reference master data will be provided by national and/or regional authorities, when possible from national and/or regional SDI web services. When these services are not available, CORDA could also communicate with other formats and sources previously defined and approved by the CORDA operator (For example, files manually managed by the operator).

Datasets and metadata hosted by European and national and/or regional authorities will be pre-registered and approved by the CORDA Operator. Metadata information such as covered geographic area, STRS, format, security access data will also be detailed in the system and stored in the Admin Database.

Simplicity of use is one of the objectives of this system and as such even if data is secured at data provider level, the system should make it transparent to the end user operator. To safeguard security of restricted data end users will need to complete a registration procedure to be granted permissions to specific data.

To ensure timely Copernicus Emergency Management response, datasets regardless of source or format will be duplicated and securely warehoused by the CORDA system, upon joint agreement with the data provider and CORDA operator. Backed up data can either be made available as a zip file or as a Web Feature Service (WFS), both located at the data provider. This backed up data, when possible, will be automatically updated as explained in section "2.6.2.2. *Off-line Scenario*".

The CORDA system will be primary fed by national, regional and European data providers. Backed up data can be used when data is not available because of major emergency disasters at data providers.

Administrative and monitoring data will be stored in the CORDA system databases.

2.3. System Modules

In each of the tiers described in Figure 3 we have identified a number of modules that we need to develop in order to achieve our goal. This chapter will go into detail on what each module should do and how they relate to the overall system.

2.3.1. Secure Access Modules

Secure access needs to take place in two areas. Firstly, access towards national and/or regional and European data providers and secondly, access for end users.

National, regional and European Data Providers (Data Security Module)

Data providers can provide secure services for use by CORDA users. These services are secured behind username and password based web-authentication, or a token based authentication environment. Through agreement between CORDA and the data provider, data will be registered by the CORDA operator and stored securely inside the CORDA system.

End Users (Single Sign-on Module)

End Users will be able to access CORDA by one single Sign-on. When users access secured services they will do so transparently by a proxy mechanism inside CORDA controlling who will perform the secure authentication and retrieve the requested content. Through keeping central control of access credentials for end users, CORDA will be able to remove old users when they are no longer involved in projects. Access towards datasets inside CORDA is managed by roles. Registered users will be granted access permissions assigned to these roles, ranging from limited access to mid and high level access. This functionality shall be implemented anticipating that some national and/or regional data providers may limit access to their data only to allow use by specific (but not all) Copernicus services.

2.3.2. Dataset Management Module

The dataset management module organizes Copernicus reference datasets from a European perspective. Datasets and metadata hosted by and national, regional and European data providers are registered as reference datasets, by the CORDA operator. Every dataset integrated exists in at least one map service. The operator will decide on which layers from a map service need to be kept for each specific reference dataset, and then prepare the map service for that specific content.

The preferred map service protocol is WMS Inspire, but where this protocol is not provided other protocols should be accepted as well. Every dataset will have authentication parameters and scale based visualization.

One European reference dataset can exist out of multiple (national or regional) map services. A collection of map services can represent a full coverage (performed by the mosaicking module).

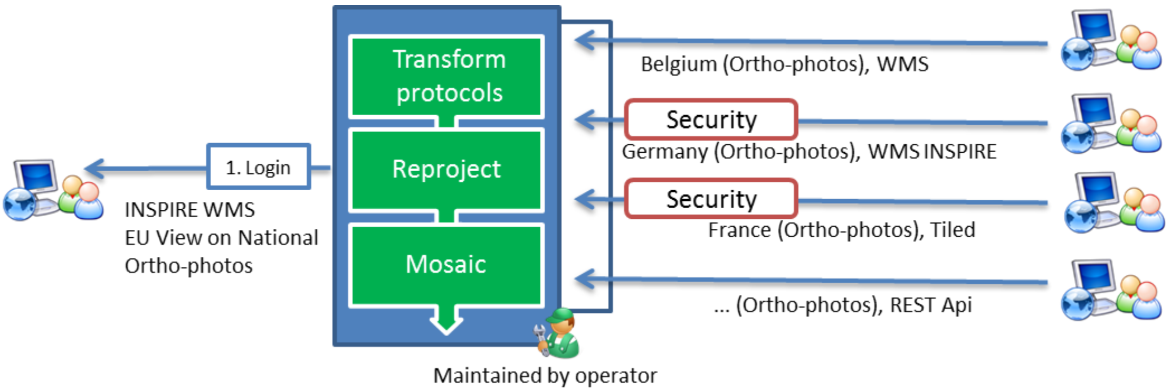


Figure 4. MiddleWare use in dataset mashup from multiple sources

For emergency purposes it is crucial that CORDA has the possibility to hold a copy of all available datasets. Every dataset has the possibility to be linked towards a feature service or web-url that points to an archived file located at the data provider that can be harvested on a regular basis (using the harvesting module).

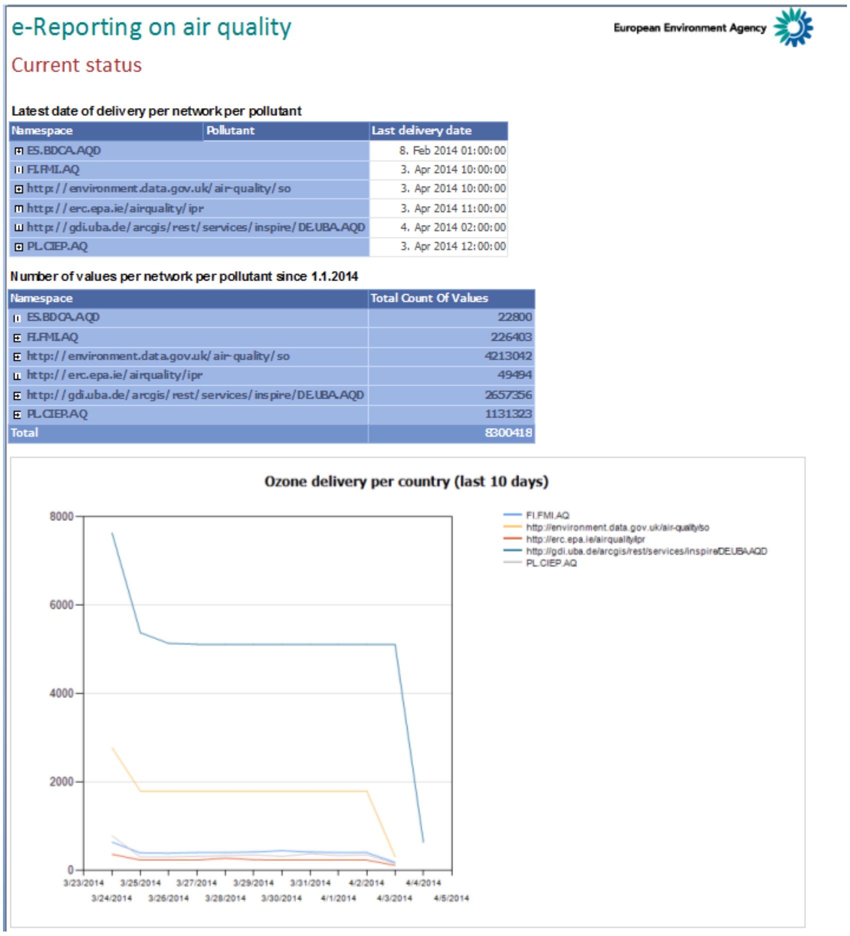
> The dataset management module will provide an accurate, accountable and up-to-date inventory of European datasets available for end users.

The dataset management module should be stored inside the internal database of the CORDA system, and will automatically harvest and store relevant information. This should include (but not be limited to) metadata of map services, and feature services and tables used inside the monitoring module, transformation and re-projection module.

2.3.3. Monitoring Module

The monitoring module is used to provide auditing on user and data access as well as system performance (See Figure 5 for an example of how this could be achieved). The aim of the module is to improve the system performance and also help refining Copernicus product quality through evaluation of requests and data fitness for use. The monitoring module will contribute user and data statistics, date and time requests, operations, services and URLs as well as identifying the most heavily used or requested products.

> Monitored information should also include a geographic component in order to visualize the spatial component of information access requests.



e-Reporting on air quality

European Environment Agency

Current status

Latest date of delivery per network per pollutant

Namespace	Pollutant	Last delivery date
ES.BDCA.AQD		8. Feb 2014 01:00:00
FI.FMI.AQ		3. Apr 2014 10:00:00
http://environment.data.gov.uk/air-quality/so		3. Apr 2014 10:00:00
http://erc.epa.ie/airquality/ipr		3. Apr 2014 11:00:00
http://gdi.uba.de/arcgis/rest/services/inspire/DE.UBA.AQD		4. Apr 2014 02:00:00
PL.CIEP.AQ		3. Apr 2014 12:00:00

Number of values per network per pollutant since 1.1.2014

Namespace	Total Count Of Values
ES.BDCA.AQD	22800
FI.FMI.AQ	226403
http://environment.data.gov.uk/air-quality/so	4213042
http://erc.epa.ie/airquality/ipr	49494
http://gdi.uba.de/arcgis/rest/services/inspire/DE.UBA.AQD	2657356
PL.CIEP.AQ	1131323
Total	8300418

Figure 5. An example of an EEA monitoring tool based on Microsoft SQL reporting.

The Monitoring Web Client (MWC) will consume web services provided by the Monitoring module and act as the interface to spatially visualize the results of monitoring.

Basic monitoring Information including daily user logs, national end point access and the total number of access requests could also be visualized in the MWC. Other information relevant to the system operator such as performance of data services, and technical problems with failing services should also be exposed.

2.3.4. Harvesting Module

The harvesting module will enable secure storage of data registered within the CORDA (from national, regional and European data providers) in the cloud, to be accessible in cases of emergency management.

> Cloud hosting is a safe, reliable and suitable method for providing data retrieval when primary data is not available from data providers due to network problems or disasters. Cloud based solution shall be physically on EU-28 territory and managed under the combined legislation of the hosting member state and the European Union.

Automated data backup should be periodically triggered by the system. Configuration of the backup module ought to be made available to the system operator so that an up to date backup is always available.

As data backup could potentially be very demanding of system resources, performance and efficiency enhancing strategies should be developed to harmonise the process. Ultimately all registered and suitable data in the CORDA system could be stored in a single basic format.
The packaging module will act to wrap all data received for a requested bundle. This will be client specific based on the types of service and data formats. One example would be an mxd with all data layers retrieved; another might be a web map to visualize web services on ArcGIS Online, or an Inspire compliant download. Again, some research should be conducted into packaging techniques, and where available existing packaging methods (e.g. Esri map pack) should be employed.

2.3.5. Transforming Module

This module assists transposing data between protocols like WFS to WMS or WMS Inspire Compliant.

Transposition services would be very useful in case specific clients do not support one or more of the services or formats required for retrieval. For example the transpose module (with a proxy) could enable an ArcGIS Online viewer to display WFS services.

The transpose module would not involve transposing all protocols in a generic way, and will more likely be task specific, keeping in mind the fundamental goal of the CORDA system. The capabilities and limitations for transposition should be further studied and specified, e.g. querying WFS services through the ArcGIS REST API, or transposing WMS to WMTS with MapProxy.

2.3.6. Reprojection Module

The reprojection module will facilitate data reprojection, format conversion, or combination so that services from different sources can be realised as layers in the end service. A number of geoprocessing tasks should be made available, wherever possible using existing API's and methods. One example of a geoprocessing operation could be to harmonise projections between differently projected datasets. Another valuable geoprocessing task could be to blend layers from different WMS services into a single WMS service (using WMS cascading). Also, if data is retrieved from WFS services it could be converted to shape files, file geodatabases etc. Existing tools like Esri Geoprocessing Services, as well as FME or GDAL libraries are suitable for this purpose.

2.4. User Scenarios

CORDA will be a server-side based focal point for searching geospatial reference data for a specific region and a specific product. As such it is worth repeating that is not in the scope of the CORDA system to interact directly with the end users, but to provide services that commonly used GIS software can access (i.e. ArcMap, QGIS, MapInfo, ArcGIS Online, Open Layers Web Viewer).

The data returned from CORDA should be in commonly used formats and protocols. Ideally these would include (but not be limited to) WMS, WFS, WMS Inspire compliant services and formats, WMTS, ARCGIS Rest services, zip files containing shapefiles. This approach aims at simplifying the system development and easing user interaction. The different scenarios provided below demonstrate how an end user would interact with CORDA and how this will be related to the above described modules.

2.4.1. Land Monitoring

2.4.1.1. Basic Scenario

This example explains a scenario where basic modules have been deployed (See Figure 6).

An end user, using ArcMap connects to the CORDA system using ArcGIS for Desktop plugin. He/she logins into the system and then requests available geospatial reference data for the Abruzzo region in Italy, for the High Resolution Layer (HRL) Forest production. The operator has registered all data available for that region so the CORDA system retrieves a list of the data to the ArcMap plugin. The end user then selects the needed datasets. This is the required geospatial reference data for HRL Forest production:

Administrative units - scale 1:25 000
DEM high resolution - scale 1:25 000
Forest inventory data (spatial) - scale 1:5 000-1:25 000
Hydrographical information - water bodies - scale 1:25 000
Orthophotos - accuracy 0.5 m

After composing the request, CORDA returns the five raw datasets as provided by the Italian authorities: administrative units, the forest inventory and the hydrographical information of water bodies as shapefiles to download, the orthophotos as a WMS service and the DEM high resolution layer as a WCS service.

The following diagram illustrates the above described information:

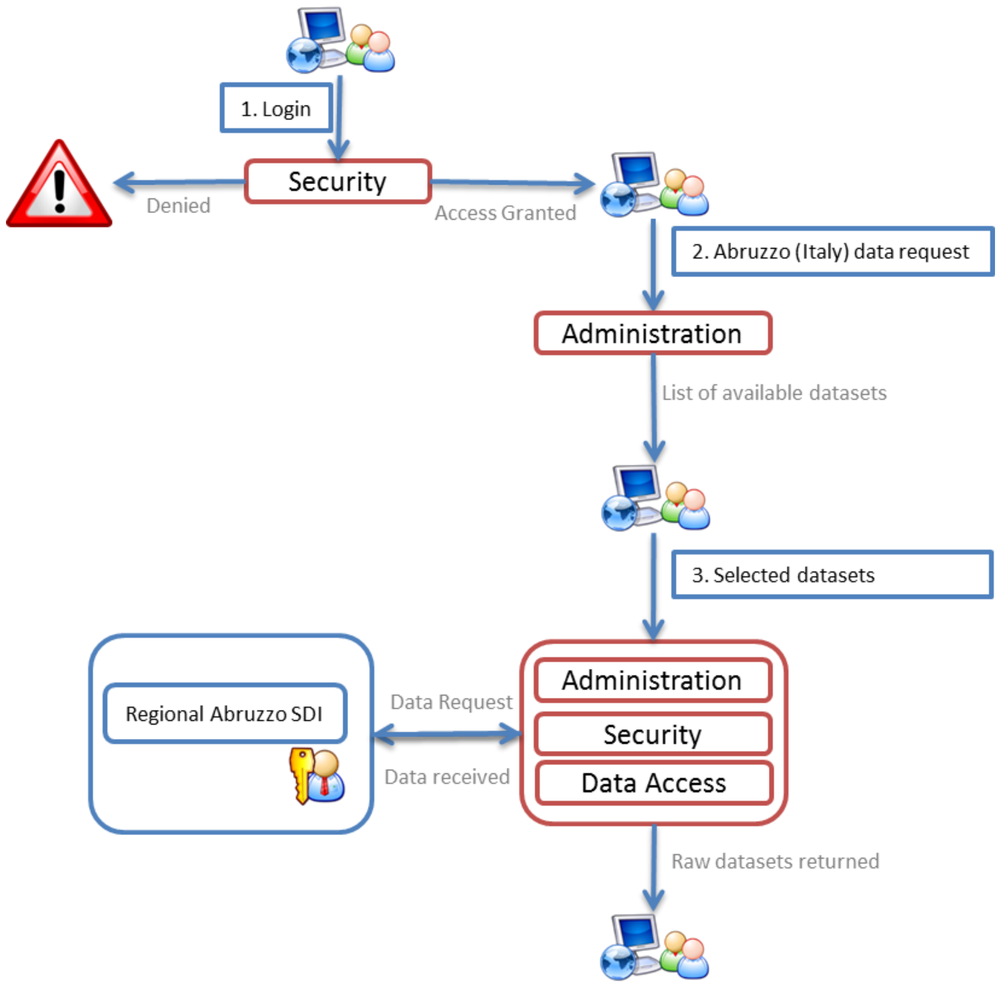


Figure 6 – Basic Example Diagram

Red boxes include CORDA modules inside the scope of the system. Text inside blue squares represents system interaction with the application/user.

When data is returned, the operator will have to handle the data in ArcMap in order to visualize it.

2.4.1.2. Complete Scenario

In this example an end user connects to the CORDA system using ArcGIS for Desktop plugin (See Figure 7). He/she logs in to the system and then requests data for Abruzzo region for the HRL Forest production.

The operator has registered all data available for that region and also identified compulsory datasets needed to create the product.

The Product module manages the information for HRL Forest production:

Required geospatial reference dataset	Criticality of the data
Administrative units - scale 1:25 000, EEA39 coverage	Essential to the product
DEM high resolution - scale 1:25 000, EEA39 coverage	Essential to the product
Forest inventory data (spatial) - scale 1:5 000-1:25 000, EEA39 coverage	Essential to the product
Hydrographical information (water bodies) - scale 1:25 000, EEA39 coverage	Essential to the product
Orthophotos - accuracy 0.5 m, EEA39 coverage	Essential to the product
DEM low to med resolution - scale 1:100 000, EEA39 coverage	Desirable to the product
Land use - scale 1:25 000, EEA39	Desirable to the product
National land cover inventories - scale 1:25 000, EEA39 coverage	Desirable to the product
Road network - scale 1:25 000, EEA39 coverage	Desirable to the product
Cadastral parcels -scale 1: 1 000, EU27 coverage	Useful to the product
City maps - scale 1:5 000, EU27 coverage	Useful to the product
Land Parcel Identification System (LPIS) - scale 1: 10 000, EEA39 coverage	Useful to the product
Protected/designated areas - accuracy 1:25 000, EEA39 coverage	Useful to the product

All of the essential datasets are available for the Abruzzo region, so the Data Access module retrieves them.

If this hadn't been the case, the system would have needed input on what available datasets to retrieve and also would have monitored this case, to identify required improvements in data availability for the geographical area. Further focus will be on the essential datasets returned from national data providers. The administrative units, the forest inventory and the hydrographical information on water bodies are returned as shapefiles. The orthophotos - as a WMS service and the DEM high resolution - as a WCS service. For each dataset, the GeoProcess module reprojects the data from the national data projection to WGS84. Then in parallel, metadata language is retrieved in English by the Language module.

All the datasets are packed appropriately so that a webmap is created (internally shapefiles are served and appropriate transposition is done for orthophotos and DEM layers services) so that all information is visualized on a template.

The following diagram illustrates the above described information:

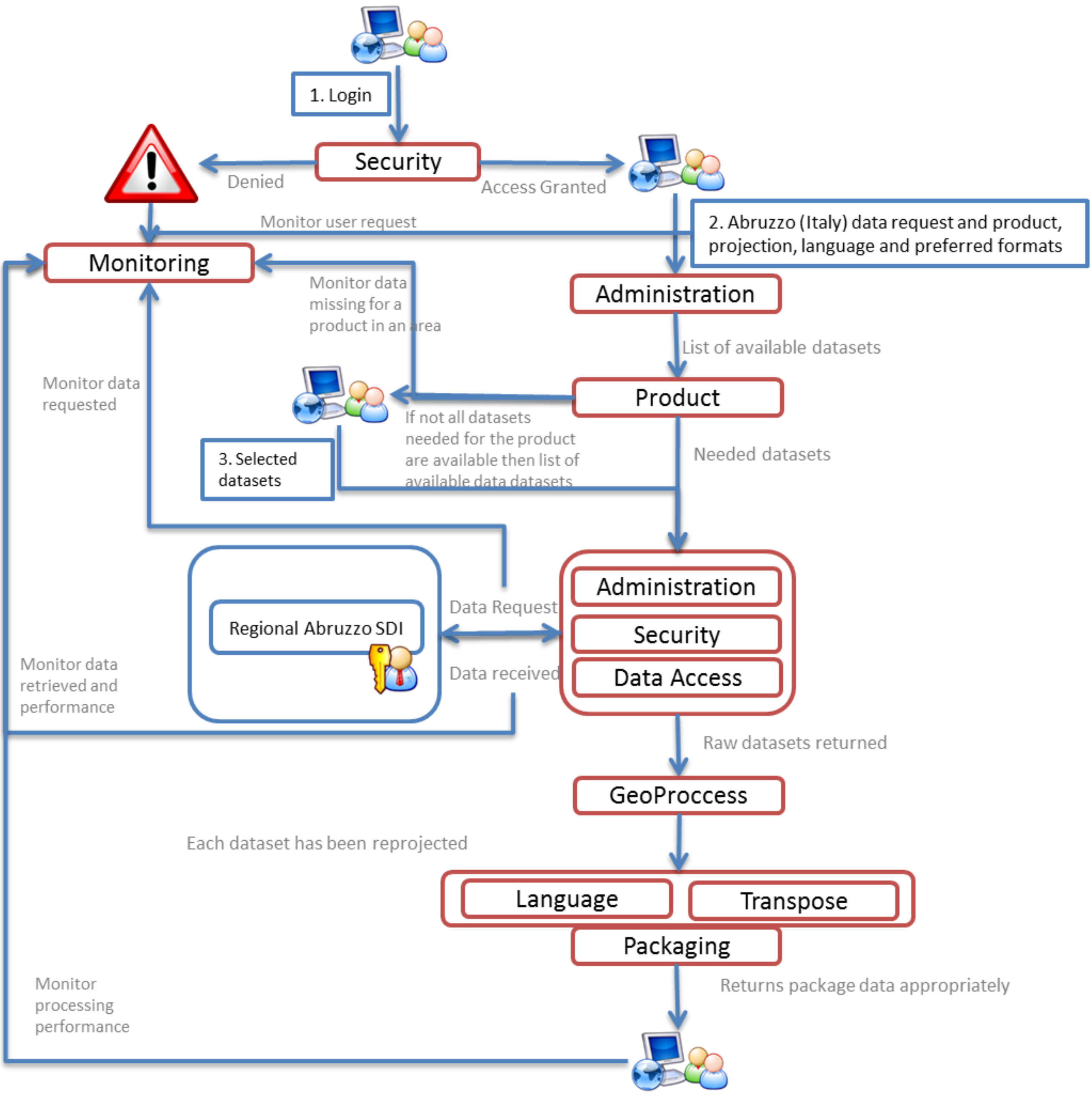


Figure 7 – Complete Scenario Diagram

2.4.2. Emergency Management

The main difference in CORDA procedures between Copernicus Land Monitoring and Emergency Management is that, for the former, national web services might not be available or accessible during emergency events, therefore the geospatial reference data registered in the system must be also stored in another secure location (See Figure 8).

In this case an essential module to ensure the system can respond to the end users' needs is the cloud storing module called Backup.

2.4.2.1. Online Scenario

The following scenario is called “online” because is activated manually.

Data is requested for a rush-mode emergency activation. The product to be delivered is a delineation map. Reference datasets required for the delineation maps are not available at national level (perhaps due to network problems), so data stored in the cloud backup is retrieved instead.

The end user logins into the system and then requests data for Abruzzo region.

The operator has registered all data available for that region and also identified compulsory (essential) datasets needed for the product.

The Product module manages the information for the product:

Required geospatial reference dataset	Criticality of the data
Administrative boundaries - scale better then 1: 50 000, global coverage	Essential to the product
Built-up areas/settlements - scale better than 1: 10 000, global coverage	Essential to the product
Critical infrastructures - public services - scale better than 1: 10 000, global coverage	Essential to the product
Critical infrastructures - utilities - scale 1: 10 000, global coverage	Essential to the product
Hydrographical information (water bodies) - scale better than 1:50 000, global coverage	Essential to the product
Key infrastructure - airports - scale better than 1: 10 000, global coverage	Essential to the product
Key infrastructure - ports - scale better than 1:10 000, global coverage	Essential to the product
Land cover information - scale better than 1: 50 000, global coverage	Essential to the product
Settlement locations and toponyms - scale 1: 50 000, global coverage	Essential to the product
Topographic maps - scale 1:25 000, global, EEA39 coverage	Essential to the product
Transport networks - railways scale better than 1:50 000, global coverage	Essential to the product
Transport networks - roads - scale better than 1: 50 000, global coverage	Essential to the product
Aerial photography - accuracy 0.2 m, for disaster affected areas	Desirable to the product
DEM high resolution - accuracy 1-10 m, EEA39, global coverage	Desirable to the product
Forestry maps - scale 1:50 000, European coverage	Desirable to the product
Landuse information - scale 1: 50 000, global coverage	Desirable to the product
Protected areas - accuracy 1:50 000, global coverage	Useful to the product
Soil information/maps - scale 1:100 000, global coverage	Useful to the product

All the essential data are available for the Abruzzo region, so the Data Access module retrieves them.

If this hadn’t been the case, the system would have needed input on what available datasets to retrieve and also would have monitored this case, to identify required improvements in data availability for the geographical area. For each dataset, the GeoProcess module reprojects the data from the national data projection to WGS84.

Then in parallel, metadata language is retrieved in English by the Language module.

All is packed appropriately, for example in an MXD file. Because all the information has been stored in the cloud all data files are downloaded locally and the MXD points to them.

The following diagram illustrates the above described information:

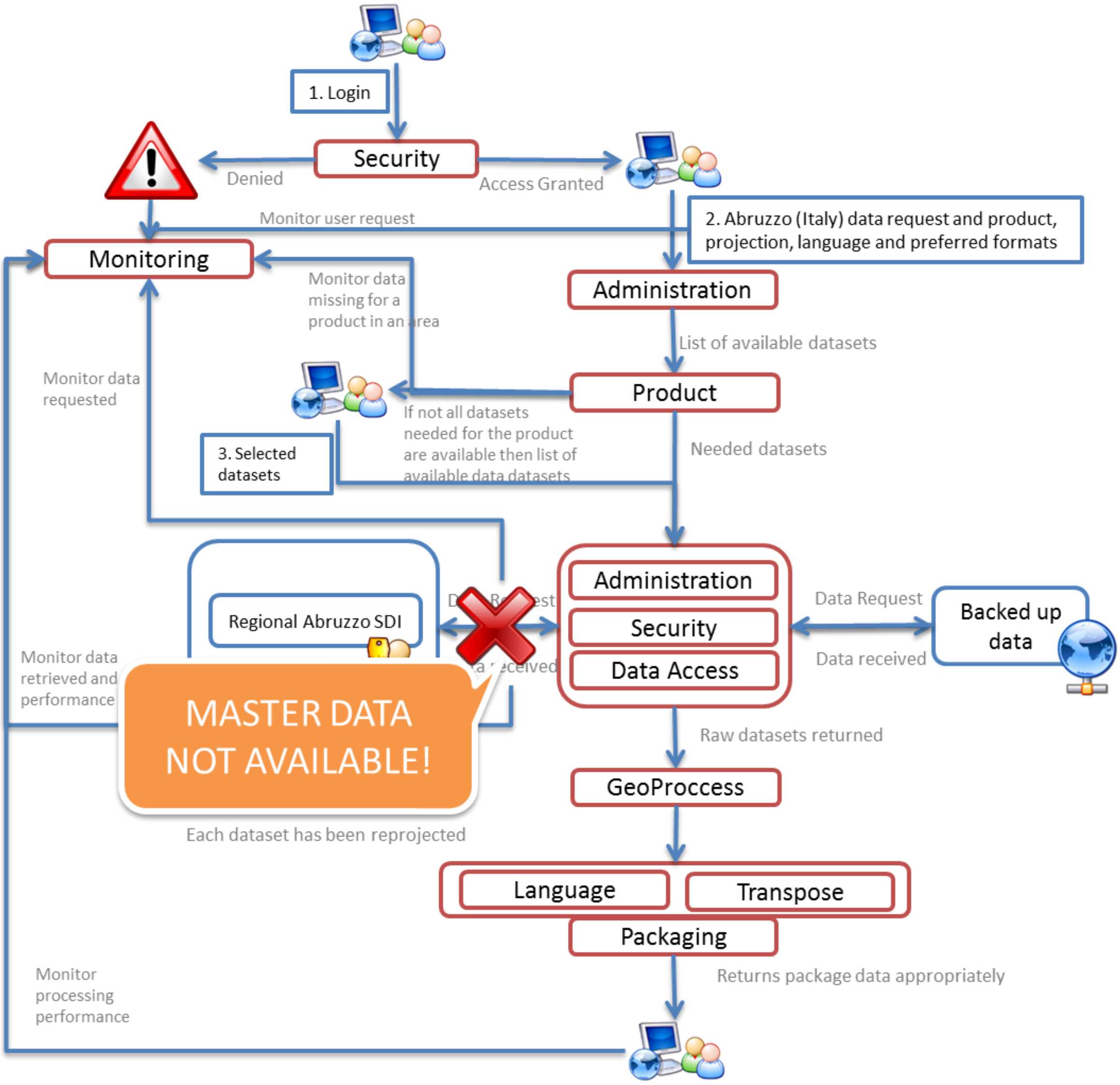


Figure 8 – Online Scenario Diagram

2.4.2.2. Offline Scenario

The following scenario is called offline as it is triggered regularly by the CORDA system in order to update the cloud backup data from the national and European data providers (See Figure 9). For each dataset registered in the system, a request is composed by the Data Access module and data is returned and stored in the cloud.

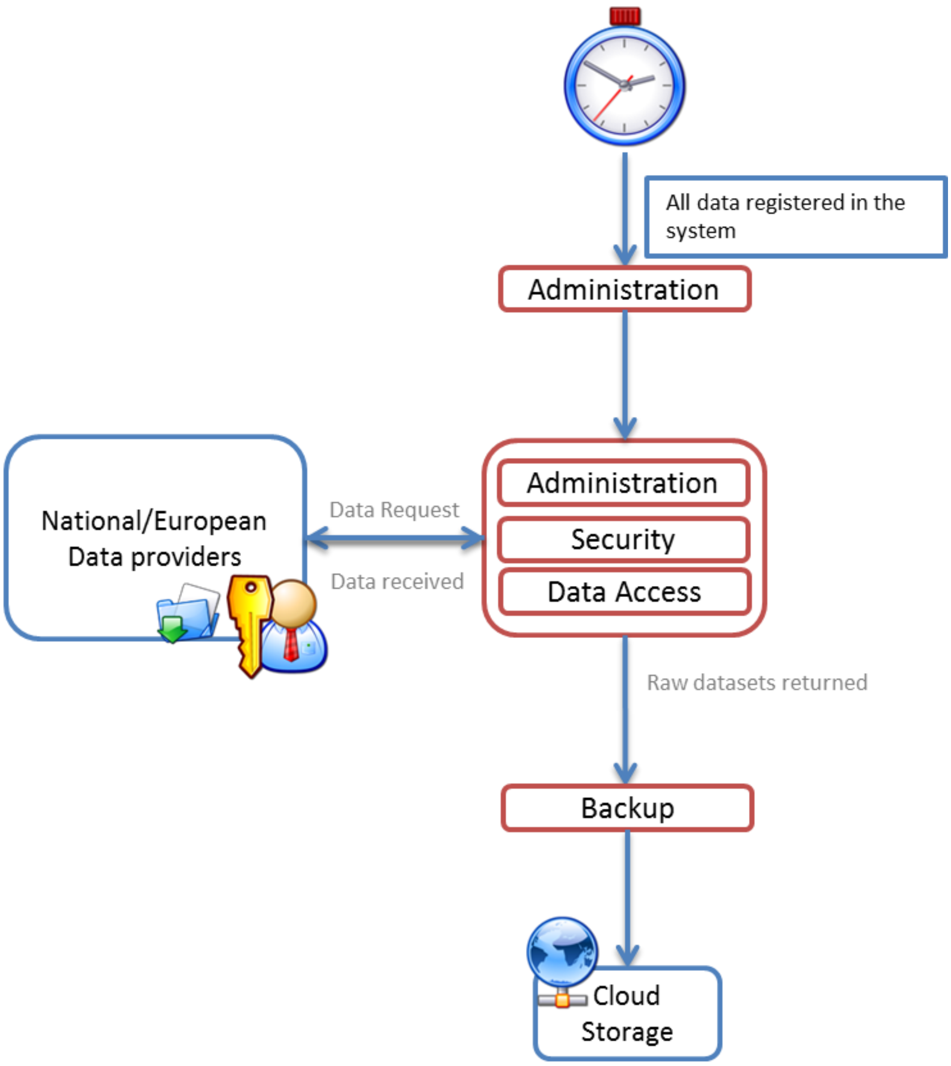


Figure 9 – Offline Scenario Diagram

2.5. Operator

2.5.1. Tasks of the Operator

The CORDA operator introduced in 2.2.1 will be responsible for:

- Informing data providers (national and European authorities) about technical details of the CORDA system
- Engaging data providers and registering CORDA system datasets available from the providers
- Maintaining the system, which includes:
  - Adding/Modifying/Deleting Regions
  - Adding/Modifying/Deleting Data layers
  - Adding/Modifying/Deleting User (End User) access
  - Adding/Modifying/Deleting Products (relation between Copernicus products and datasets needed)

2 Managing Cloud Backup

The operator of CORDA is also a data expert responsible for registering datasets in the CORDA system, data management, and fixing problems with the data, such as broken links in services.

2.5.2. Operator Scenario

The operator should build up an encyclopaedic knowledge of the Copernicus products and services, and source datasets used for them. In the following example we will explore the role of the operator in creating a hypothetical Copernicus product, called ‘Forest Fire Risk Assessment’.

The operator looks at the product description, and sees that the Forest Fire Risk Assessment is a Pan-European multi layered raster output consisting of three layers; a forest raster, a road network and a river network.

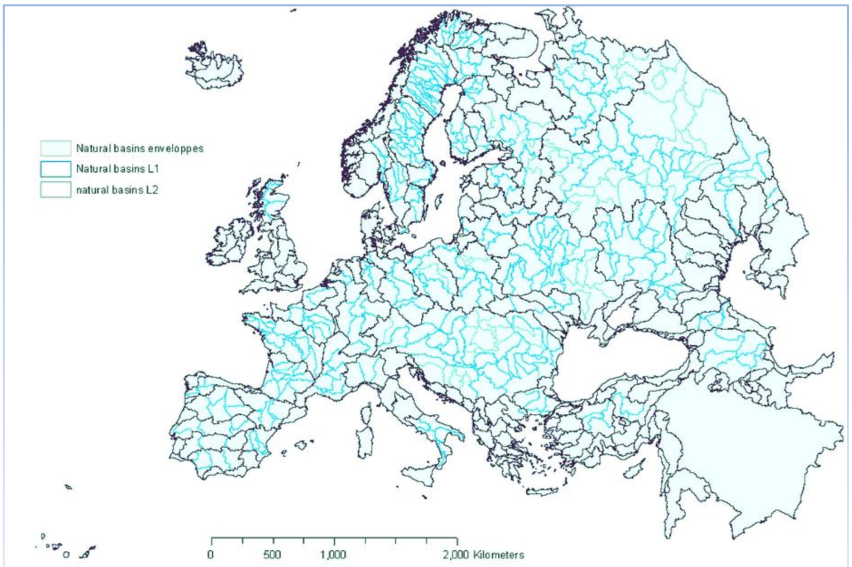
The operator looks for the three source datasets he will need in the Copernicus inventory, and finds three different data products which contain suitable datasets:

Corine Land Cover

The Corine Land Cover inventory was initiated in 1985 (reference year 1990). Updates have been produced in 2000 and 2006, and the latest 2012 update is under production. It consists of an inventory of land cover in 44 classes. CLC uses a Minimum Mapping Unit (MMU) of 25 ha for areal phenomena and a minimum width of 100 m for linear phenomena. The time series are complemented by change layers, which highlight changes in land cover with an MMU of 5 ha.

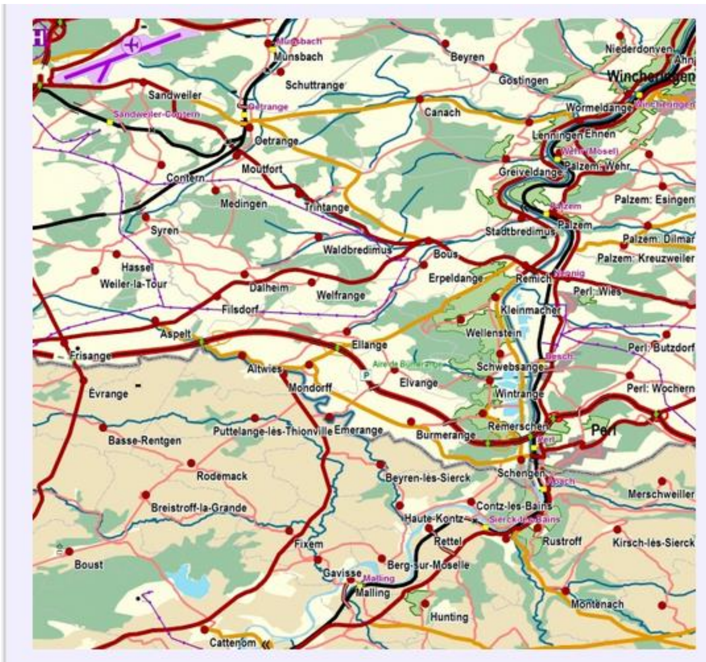
CLC is produced by visual interpretation of high resolution satellite imagery. The Eionet networks National Reference Centres Land Cover (NRC/LC) are producing the national CLC databases, which are coordinated and integrated by EEA. The 2012 version of CLC is the first one embedding the CLC time series in a structural context of the Copernicus programme.

ECRINS hydrography dataset



ECRINS is acronym for European catchments and Rivers network system. It is a geographical information system of the European hydrographical systems with full topological information. ECRINS is a composite system made from the CCM developed by the JRC, Corine land Cover, WFD reporting elements, etc. It is organised from a layer of 181,071 “functional elementary catchments (FECs)” which average size is ~62 km2, fully connected with explicit identifier (ID) relationships and upstream area. Catchments are grouped as sub-basins, river basin districts (actual and functional to meet hydrographical continuity). The catchments are as well organised according to their sea shore of emptying to meet Marine Strategy delineations. Catchments are drained by 1,348,163 river segments, sorted as “main drains” (connecting together the FECs) and secondary drains (internal to a FEC). river segments mimic the natural drainage, however fulfilling the topological constraint of “0,1 or 2 upstreams, single or 0 downstream”. Each segment is populated with distance to the sea, to ease further processing.

EuroRegionalMap



EuroRegionalMap v6.0 is a pan-European dataset containing topo-geographic information at the scale 1:250 000 covering 33 European states: 26 EU member states (Bulgaria not included), 4 EFTA states (Liechtenstein, Norway, Iceland and Switzerland), Moldova, Serbia, and Georgia. It is a seamless and harmonised data and is produced in cooperation by the National Mapping and Cadastral Agencies of Europe, using official national databases.

EuroRegionalMap is ideal for a wide range of uses, including spatial analysis, cartographic publishing and backdrop visualisation, or in combination with other datasets for marketing planning and socio-economic analysis, environmental analysis, and transport management.

The operator sees the three datasets contain many layers, not just those required for the Forest Fragmentation Indicator. For example, Corine Land Cover contains 44 land cover classes grouped into three hierarchical levels, in vector and raster format. The operator then selects the Forest CLC The operator chooses CLC 3.1 – Forests raster, from each **National Reference Centre Land Cover (NRC/LC)**. The operator makes similar selections from the remaining datasets - Ecrins River Segments, and EuroRegional Road and Rail Network.

The operator now creates the Forest Fragmentation Indicator mashup dataset, registering the output node as a web service to make the indicator available, and hooking up the three layers from each of the national reference datasets. Here, the latest data available is used in the indicator, and made available to the Copernicus end user.