



Project of Strategic Interest NEXTDATA

Final Requirements and System Architecture (ND-SoS-Ina)

Deliverable D2.7.3 General Requirements

Deliverable D2.7.4 General System and Interoperability Architecture

WP Coordinator: Stefano Nativi
CNR – IIA

Document identifier:	D2.7.3-D2.7.4
Date:	Apr 2015
Editors:	Stefano Nativi, Mattia Santoro, Enrico Boldrini, Fabrizio Papeschi
Activity:	SoSIna WP2 - Task 2.1 - 2.2 - 2.4
Lead Partner:	CNR
Document status:	final
Document link:	http://essi-lab.eu/nextdata/sosina/?page_id=68
Diffusion	Internal
Keywords	System Architecture, Requirements, Broker pattern, Brokering framework.
Abstract	This document presents the second and final version of the NextData SoSIna System requirements and Interoperability Architecture providing the necessary inputs for the Web portal design and development.

History

Document version	Date	Authors
1.0	Apr 2015	Stefano Nativi, Mattia Santoro, Enrico Boldrini, Fabrizio Papeschi (CNR-IIA)

Abstract

This document presents the second and final version of the **NextData SoSIna System requirements and Interoperability Architecture** providing the necessary inputs for the Web portal design and development.

Keywords

System Architecture, Requirements, Broker pattern, Brokering framework.

TABLE OF CONTENTS

1	<i>NextData Users' needs</i>	5
1.1	The Process	5
1.2	NextData e-infrastructure needs	5
2	<i>General System and Interoperability Architecture</i>	6
2.1	Leading Principles and system architecture	6
3	<i>NextData Multidisciplinary Requirements</i>	7
3.1	Functional Requirements	7
3.1.1	HCI functionalities	8
3.2	Non-functional Requirements	9
3.2.1	Data Sharing requirements	10
4	<i>Basic Use Case</i>	10
5	<i>Possible extensions</i>	11
6	<i>Conclusions</i>	12
7	<i>References</i>	12

1 *NextData Users' needs*

1.1 The Process

The introduced set of architectural requirements are stemming from the following actions:

- Interactions with NextData scientists and with many stakeholders working on Global Changes and Earth System Science.
- Considering the requirements coming from other relevant projects and initiatives (like GEOSS, INSPIRE, Copernicus, Belmont Forum, Earth Cube, and some FP7 and ESA funded programmes - including EuroGEOSS, GEOWOW, Med-SUV, MEDINA, Prod-Trees, GeoViQua, RECODE, EO-POWER, EU-BON, COOPEUS, ODIP, SeaDataNet II, etc.).
- Considering the assessment of the first NextData e-infrastructure prototype and its Web Portal functionalities.

1.2 NextData e-infrastructure needs

The following needs have been recognized and considered for the Architecture design and implementation:

- To achieve stepwise operating capabilities, including the development of the services that are necessary to make it possible to discover, view, access, and use the information resources made available by the Project partners in the different thematic areas.
- To undertake the research necessary to further develop thematic partner infrastructures to provide access not just to data but also to selected analytical models and workflows.
- To discover and access to resources made available through heterogeneous and distributed systems and DBs.
- In addition to the matching of discovery query clauses, to rank and classify discovered results coming from different distributed catalogs and systems. Resources could be ranked by considering: (i) the effective publication as open and free resources, (ii) the quality of the metadata (i.e. presence of title, description, URI, resource accessibility, availability of resource, etc.).
- To address the semantic heterogeneity among the terminology used by different scientific communities.
- To access data according to the needed spatial and temporal resolution, format encoding and Coordinate Reference System (from any of the available download services) using the service interface preferred by users.
- To access and harmonize data through controlled data schema mappings across different provider repositories.
- To implement a scalable and flexible infrastructure that seamlessly allows users to compose available processes to build complex business processes and execute them.
- To perform quality assessment of the resources discovered in the NextData infrastructure.
- To include provenance information and assure science reproducibility, where possible.
- To improve the Open Data availability, supporting easy registration, discovery and access of Open Data resources.

- To address access condition, through registration and authorization functionalities.
- To facilitate public participation and interaction with citizen-provided content.
- To create and implement an easy to use and flexible Web Portal.
- To address non-functional requirements such as scalability, performances and different security levels.

From these requirements, a set of functional and non-functional requirements has been recognized to design the NextData Web Portal and the NextData infrastructure architecture, more generally.

2 General System and Interoperability Architecture

It was introduced a list of high-level “system and software requirements”, building on the Users’ needs as well as on the interoperability requirements for interconnecting NextData infrastructure to significant international initiatives and programmes (i.e. GEOSS, INSPIRE, Copernicus, Belmont Forum, Earth Cube, EoE, Unep-live, WIS, etc.).

NextData aims to reinforce the Italian contribution to the GEOSS tasks by interconnecting its e-infrastructure to the GEOSS Common Infrastructure (GCI) [R1]. Therefore, NextData was required to design, implement, and operate a multi-disciplinary e-infrastructure, applying the interoperability principles recognized by GEO and based on INSPIRE, Copernicus, and other international programs.

2.1 Leading Principles and system architecture

The NextData infrastructure is called upon to apply the following principles, characterizing the System of Systems (SoS) approach successfully introduced with the GCI [R2]:

1. Build on existing system/network infrastructures.
2. Supplement but not supplant system/network (i.e. Community) mandates and governance arrangements.
3. Address heterogeneity to lower entry barriers.
4. Be flexible and extensible to:
 - Interconnect new system/network infrastructures.
 - Sustain and advance the achieved interoperability.
 - Allow each system/network (i.e. Community) infrastructure to evolve.

Therefore, the set of high-level system and software requirements includes the need to build on the existing (or next coming) information systems published by the NextData “Resource Providers”.

To be interoperable, each NextData “Resource Provider” needs to publish one or more Internet based services/interfaces to make it possible to discover and access their shared resources (see D2.4 “Open Data Policy”). Virtually, there is no constraint on the types of published services/interfaces: semantically speaking, they can be either very simple (e.g. FTP, Web Accessible Folder, DB Web interface) or more complex (e.g. OGC/ISO Web services, Community SOAP-based/REST services, linked-data services).

On the top of these “Community” information systems, the NextData e-infrastructure is realized. The infrastructure is in charge of four overarching tasks:

- A. The interconnection of all the heterogeneous information systems provided by the NextData Partners, in a transparent way.
- B. The harmonization of the data and metadata provided by the NextData Partners.
- C. The implementation of a set of multidisciplinary discovery, evaluation, access, and simple transformation services.
- D. The development and publication of a general NextData Web Portal providing the Human-Interaction-Interface (HCI) functionalities to make use of the services in point C).

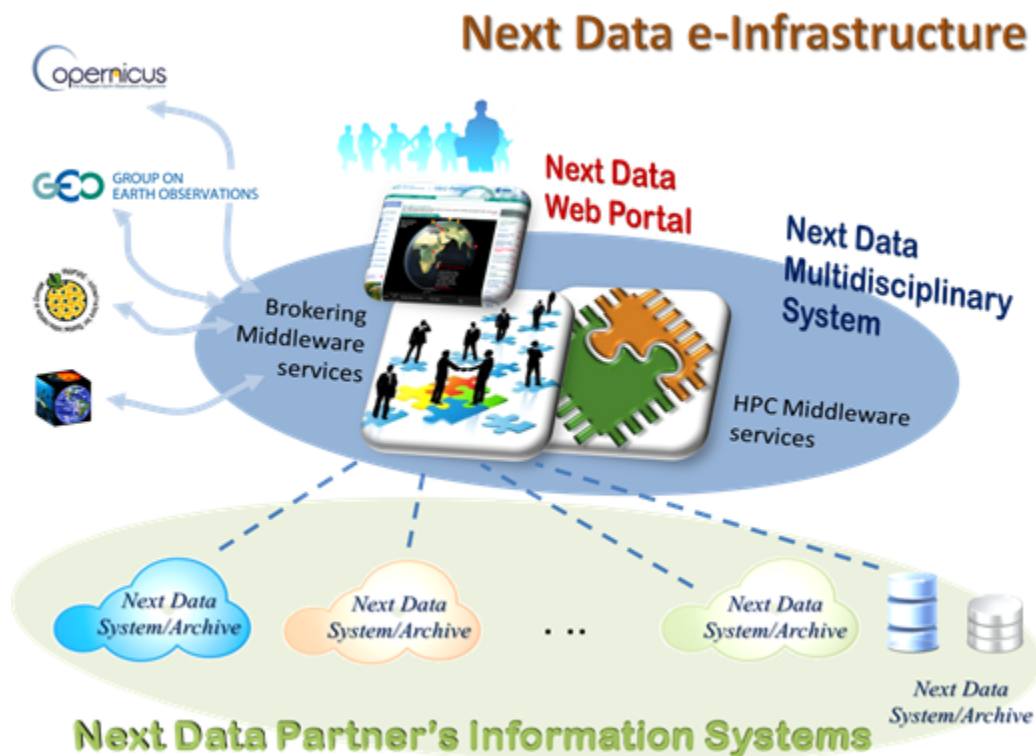


Fig. 1. NextData e-infrastructure architecture and the Web Portal.

Each NextData Partners's Information Systems shall consider a set of data management principles (i.e. the GEO Data Management principles), as described in the NextData Open Data Policy.

3 NextData Multidisciplinary Requirements

3.1 Functional Requirements

The NextData e-infrastructure will implement a set of multidisciplinary functional requirements including:

(a) Architectural Interoperability:

- access to heterogeneous and distributed systems (supporting service-oriented, resource-oriented and event-based architectures).

- (b) Discoverability and Evaluation:
 - heterogeneous metadata access and harmonization to allow evaluation;
 - classification and ranking of discovered resources;
 - quality and provenance information support (including users' feedbacks).
- (c) Accessibility:
 - Heterogeneous data access and harmonization (where requested);
 - Transformation services to supplements data providers capabilities.
- (d) Semantics Interoperability:
 - Heterogeneous semantic assets access and use to support semantic discovery;
 - Semantic assets alignment (where requested).
- (e) Processing:
 - possibility to compose processes and scientific models;
 - community workspaces integration.
- (f) Organization Interoperability:
 - security and trustability services.

3.1.1 HCI functionalities

Figure 2 shows the main HCI functionalities.

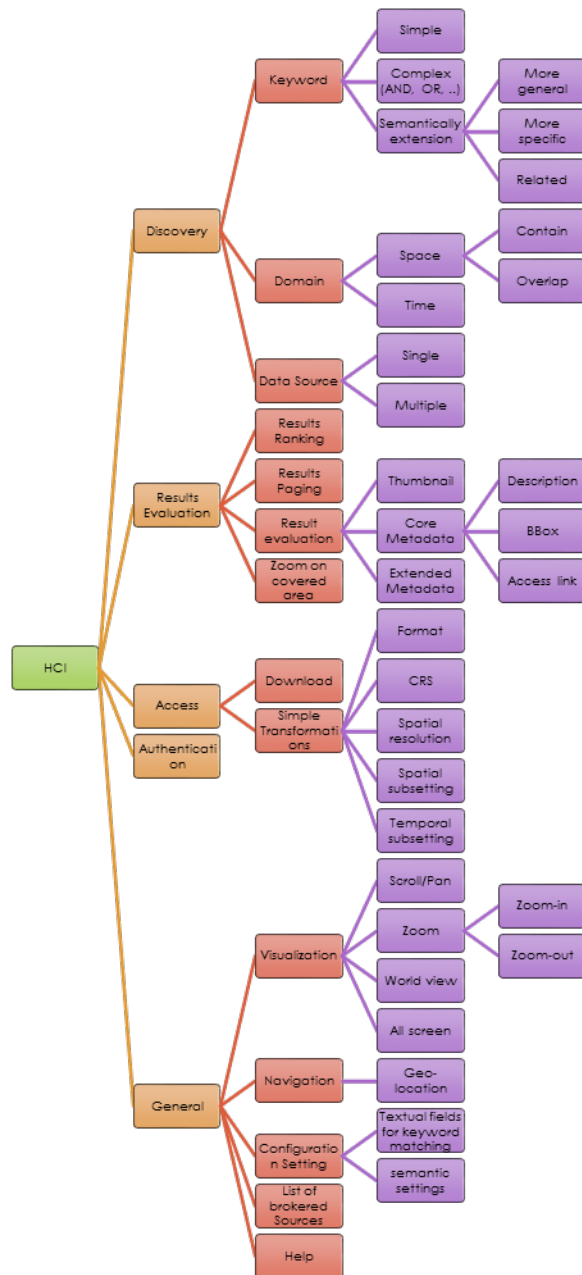


Fig. 2. Main HCI functionalities implemented by the NextData e-infrastructure.

3.2 Non-functional Requirements

The NextData e-infrastructure shall implement a set of non-functional requirements addressing following main challenges [R3]:

(a) User Interface and Human Factors

- Q: What type of user will be using the system? A: Scientists (e.g. Environmental and Data scientists).

- Q: What sort of training will be required for each type of user? A: No specific training; however, a guideline document and a mailing list will be provided to help system Users.
- (b) Documentation
- Q: What kind of documentation is required? A: Guidelines for Users.
 - Q: What audience is to be addressed by each document? A: The target is the Scientific Community.
- (c) Performance Characteristics
- Q: Are there any speed, throughput, or response time constraints on the system? A: NextData will adopt the same quantitative characteristics recommended by GEOSS.
 - Q: Are there size or capacity constraints on the data to be processed by the system? A: NextData will adopt the same quantitative characteristics recommended by GEOSS.
- (d) Quality Issues
- Q: What are the requirements for reliability? A: They are covered by the data management principles specified by the NextData “Open Data Policy”.
 - Q: What is the acceptable system downtime per 24-hour period? A: 0%
- (e) System Modifications
- Q: What parts of the system are likely candidates for later modification? A: the mediation and brokering functionalities to interconnect new (and heterogeneous) data systems.
 - What sorts of modifications are expected?
- (f) Resources and Management Issues
- Q: How often will the system be backed up? A: every week.
 - Q: Who will be responsible for system maintenance? A: CNR and CINECA.

3.2.1 Data Sharing requirements

In keeping with the Data policy of NextData [R4], the multi-disciplinary e-infrastructure will apply the following sharing principles:

- There will be full and open exchange of data, metadata and products shared within NextData, recognizing relevant international instruments and national policies and legislation.
- All shared data, metadata and products will be made available with minimum time delay and at minimum cost.
- All shared data, metadata and products being free of charge or no more than cost of reproduction will be encouraged for research and education.

4 Basic Use Case

Several use cases are possible with the NextData e-infrastructure; however, the basic one to be supported is the following (see Figure 3.):

1. NextData User generates and issues a discovery request (i.e. query).
 - a. The NextData e-infrastructures answers with a set of matching results coming from all the brokered resources.
2. NextData User browses and evaluates the result set, asking the system to visualize dataset information (i.e. metadata) and a portrayal.
3. NextData User requires to access one or more evaluated datasets, specifying a common spatial and temporal domain and data format (i.e. common grid environment) to improve their usability.
 - a. The NextData e-infrastructures finalizes all the transformations required to return the selected datasets according to the specified common grid environment.

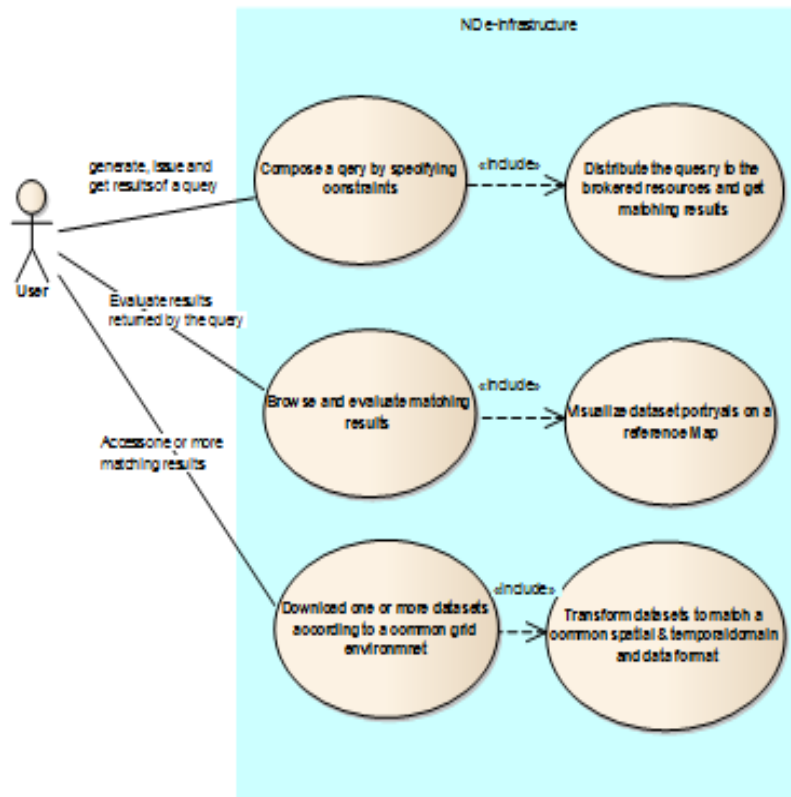


Fig. 3. The NextData “discovery & access” basic Use Case

5 Possible extensions

In an AGILE software development approach, requirements are often refined and new ones may be introduced during all the Project life-time. Therefore, even if this is the final version of the architectural requirements, it is still possible that this document will be updated in the case a major development cycle would be necessary.

6 Conclusions

In this document, we have presented the high-level system and software requirements for the architecture of the NextData e-infrastructure.

Specific functional and non-functional needs are listed and briefly discussed.

7 References

R1	GEOSS Web Site: http://www.earthobservations.org/geoss.shtml
R2	Brokering Approach: NATIVI S. et AL., (2013): IEEE JSTARS manuscript.
R3	Non-functional requirements: CSEE, University of Maryland
R4	Policy Data of NextData