Ref. Ares(2024)7380590 - 17/10/2024

NEXOGENESIS
STREAMLINING WATER RELATED POLICIES

D4.2 Data Lake for data sharing

Lead: Lluís Echeverria (Eurecat)
Date : 30/04/2024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Project Deliverable

Project Number	Project Acronym	Project Title
101003881	NEXOGENESIS	Facilitating the next generation of effective and intelligent water-related policies, utilizing artificial intelligence and reinforcement learning to assess the water-energy-food-ecosystem (WEFE) nexus

Instrument:	Thematic Priority
H2020 RIA	LC-CLA-14-2020

Title
D4.2 Data Lake for data sharing

Contractual Delivery Date	Actual Delivery Date
M32: April 2024	M32

Start Date of the project	Duration
01 September 2021	48 months

Organisation name of lead contractor for this deliverable	Document version
EUT	1.0

Dissemination level	Deliverable Type
Public	**Demonstrator**

Authors (organisations)
Lluís Echeverria (EUT), Didac Colominas, (EUT) Nuria Nievas (EUT), and Chaymaa Dkouk (EUT)
Reviewers (organisations)
Lydia Vamvakeridou-Lyroudia (KWR)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Abstract

The WP4 deliverable D4.2 (**Data Lake for Data Sharing**) is classified as a "Demonstrator." This document is intended to accompany the deliverable, providing explanations about its design, functionalities, and technical aspects to complement the developed digital solutions. Ultimately, it can also be used as a guideline to develop a Data Lake or a Semantic Repository, since all the steps and stages are documented.

Data are critical to the Nexogenesis project, both for internal operations and for creating outputs for external publication and knowledge sharing. They support internal tasks and are essential for sharing results with external users and the general public, in line with the Nexogenesis commitment to Open Data and its Data Management Plan (DMP).

As a principal outcome of task T4.2, the **Nexogenesis Knowledge Repository** has been implemented and deployed. It is composed by two sub repositories, the **Nexogenesis Data Lake** (the Internal Data Repository), supporting intra-project data exchange, and the **Nexogenesis Semantic Repository**, for external data publication and nexus knowledge sharing. Together, both components are also known as the **Nexogenesis Data Sharing** tools pack, which establishes the basis for the Nexogenesis data pipeline.

Furthermore, the Nexogenesis Semantic Repository embraces both the **Nexogenesis Nexus Ontology** and **Semantic Repository services**. Additionally, the latter includes various components: the **Nexogenesis Contextualizer**, the **SPARQL server**, and the **Nexogenesis Data Explorer and Visualizer**.

The **Nexogenesis Nexus Ontology** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/ontology/.

The **Nexogenesis SPARQL server** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/semanticRepository/.

And the **Nexogenesis Data Explorer and Visualizer** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/visualizer/

Related Deliverables:
D4.1 Self-learning nexus engine specifications and technical design

Keywords
SLNAE; NEPAT; Data Lake; Semantic Repository; Nexus Ontology; Nexus Data and Knowledge Sharing, Exploration and Visualization

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Abbreviation/Acronyms

CS	Case Study
DMP	Data Management Plan
GUI/UI	Graphical User Interface/User Interface
HTML	Hyper Text Markup Language
HTTPS	Hypertext Transfer Protocol Secure
ICT	Information and Communication Technologies
IDR	Internal Data Repository
JSON	JavaScript Object Notation
NEPAT	Nexus Policy Assessment Tool
NXG	Nexogenesis project
NXGT	Nexogenesis tool
OWL	Web Ontology Language
SDM	System Dynamic Model
SLNAE	Self-Learning Nexus Assessment Engine
SH	Stakeholder
RDF	Resource Description Framework
RP	Reference Pathway
WEFE	Water-Energy-Food-Ecosystem
TTL	Terse RDF Triple Language
WP	Work Package

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

1. Introduction

Data play a critical role in the Nexogenesis project. They are essential for internal use to complete project tasks and outputs, and for final publication to external users and the general public. It all aligns with the Nexogenesis commitment to Open Data, the Open Data policies and the Nexogenesis DMP (Data Management Plan).

While the NXG DMP and the task T4.6 *Data Management Strategy* are in charge of the definition of the Nexogenesis Data Strategy, the task T4.2 *Data Sharing Tools* (to which this Deliverable is linked) is in charge of the development, deployment and maintenance of all the required digital ecosystem to support the internal and external data sharing activities.

Under this umbrella definition, within T4.2, the **Nexogenesis Knowledge Repository** has been implemented and deployed. It is composed by two sub repositories, the **Nexogenesis Data Lake** (the Internal Data Repository), supporting intra-project data exchange, and the **Nexogenesis Semantic Repository**, for external data publication and nexus knowledge sharing. Together, both components are also known as the **Nexogenesis Data Sharing** tools pack, as despised in Figure 1, which establishes the basis for the Nexogenesis data pipeline.

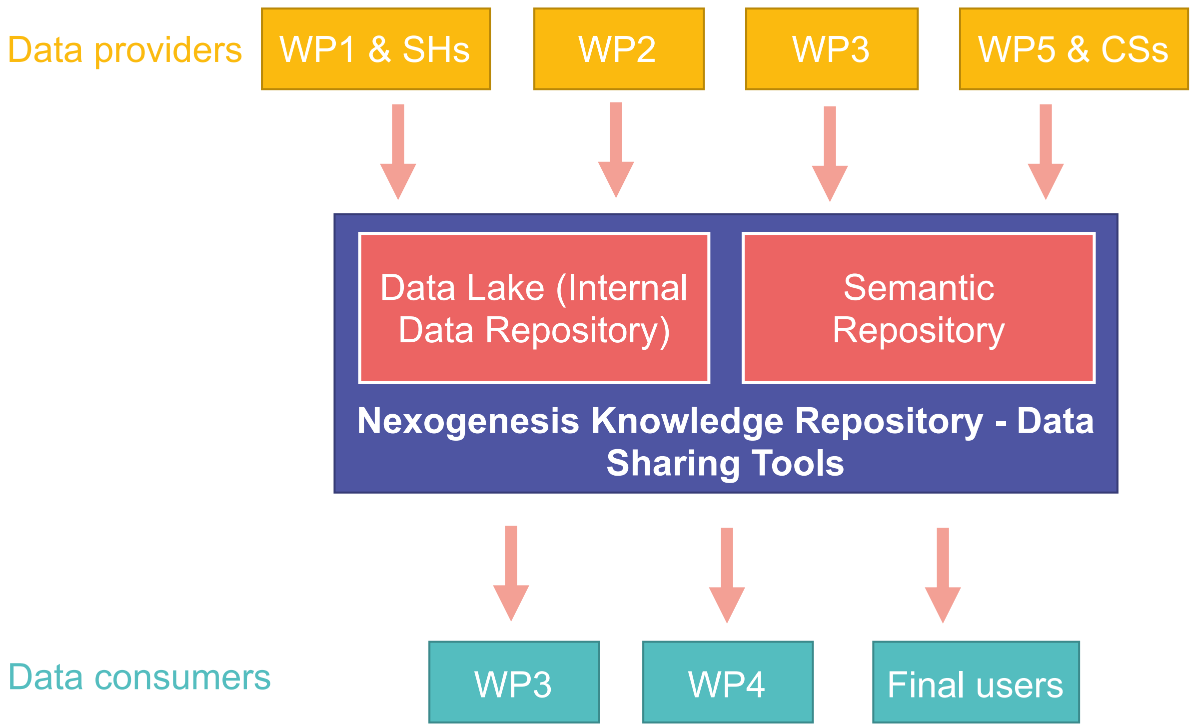


Figure 1. Nexogenesis data pipeline. Extended from D4.1.

The Nexogenesis Semantic Repository embraces both the **Nexogenesis Ontology** and **Semantic Repository services**. Additionally, the latter includes several components: the **Nexogenesis Contextualizer**, the **SPARQL server**, and the **Nexogenesis Data Explorer and Visualizer**.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

The **Nexogenesis Nexus Ontology** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/ontology/. The **Nexogenesis SPARQL server** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/semanticRepository/. And the **Nexogenesis Data Explorer and Visualizer** is publicly available at the following url: https://nepat-dev.nexogenesis.eu/visualizer/

All tools and services are operated by Eurecat (EUT). EUT holds the necessary security privileges and access credentials to log in to these systems and perform all management activities.

1.1. Disclaimer

The **Nexogenesis Knowledge Repository** has been successfully implemented and deployed, with several datasets already included. However, the repository is expected to be further supplemented with additional datasets that may be generated during the final phase of the project.

This process may continue until the project's completion and is in alignment with the Data Management Plan and Task T4.6.

1.2. Links to the ICT4WATER Cluster

WP4 is considered the ‘digital’ WP in the Nexogenesis project. Thus, it is the natural link between the project and the ICT4WATER Cluster[1].

The outcomes of the task T4.2 are specially linked to the ICT4WATER Digital Water Action Plan and its ‘Enabling Data Sharing’ Action Group[2], which is led by two of the NXG partners, UTH (WP3 leader) and EUT (WP4 leader, Data Protection Officer, and responsible of this deliverable).

Particularly, this deliverable and the digital services developed under T4.2 contribute to the ‘Data Sharing’ action (see https://ict4water.eu/wp-content/uploads/2023/06/Update-Digital-Water-Action-Plan-V7.pdf), specifically to its activities 1, 2, and 3, by implementing, deploying, offering and maintaining the Nexogenesis Semantic Repository service.

1.3. NEPAT: a new name for the SLNAE

[1] https://ict4water.eu/
[2] https://ict4water.eu/action-group-data-sharing/

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003881

During the initial period of the project, it was necessary to explain the definition and meaning of the SLNAE to non-technical audiences (e.g. stakeholders) on multiple occasions. The term "Self-Learning" is a technical concept related to AI and ML algorithms involved in its development, which can be difficult to understand and may generate confusion. Additionally, the acronym SLNAE and the full name are not easy to pronounce. Therefore, it was decided with the NXG consortium that a new name was needed.

WP4 led the initiative to define a new name that would be self-explanatory and avoid technical jargon, while incorporating Nexogenesis-related concepts such as "policy assessment" or "impact." Several options were proposed under the cocreation framework and put to a vote, and eventually, the name "Nexogenesis - Nexus Policy Assessment Tool (NEPAT)" was selected as the best option. The new name is simpler, easier to pronounce, and more reflective of the tool's and project's purpose. In order to be consistent with the GA and other official documentation, both names are valid to refer to the SLNAE.

Thus, the SLNAE tool is referred to as either SLNAE or NEPAT.

1.4. Document structure

The document is structured as follows: Section 2 describes the Nexogenesis Data Lake implementation and current usage. Next, Section 3 presents the Nexogenesis Semantic Repository, composed by the Nexogenesis Nexus Ontology and the Semantic Repository services. The latter comprises the Nexogenesis Contextualizer, the SPARQL server, and Nexogenesis Data Explorer and Visualizer. Finally, the section 6 ends with the conclusions and next steps.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

2. Data Lake

Following the decisions and designs proposed in Deliverable D4.1, the Nexogenesis Internal Data Repository, or Data Lake, has been organized and deployed using Microsoft OneDrive[3].

Although CKAN[4] (Comprehensive Knowledge Archive Network) was considered as an alternative solution, it was ultimately deemed unnecessary. CKAN is an open-source data management system used to create and maintain data portals and repositories. It is popular among governments, organizations, and research institutions for sharing and publishing datasets with the public.

However, CKAN was dismissed because Nexogenesis Data Lake users do not require its advanced functionalities. Principally, because the Data Lake purpose is for internal use. Furthermore, adopting CKAN would have introduced an unnecessary layer of complexity in terms of usage, management, and maintenance.

Other digital data repository solutions, such as PostgreSQL, MongoDB, and other relational and NoSQL databases, were also considered but eventually rejected for two main reasons:

- **Complexity**: Managing and accessing these types of software requires advanced computer science knowledge and resources. Most Nexogenesis partners, who are the primary users of the Data Lake, do not meet these requirements.
- **Advanced and Unnecessary Functionalities**: These repositories are designed to store and manage application data and offer specific features such as data manipulation, indexing, and complex querying. Since these capabilities are not required by the Nexogenesis Data Lake users, adopting these solutions would have added unnecessary overhead across all levels, from user interaction to system management.

Thus, Microsoft OneDrive was chosen as the optimal solution for the Nexogenesis Data Lake due to its simplicity and ease of use, aligning with the needs and capabilities of its users.

OneDrive is primarily a cloud storage service provided by Microsoft, designed for the storage, sharing, and synchronization of files and folders. It allows users to store files in the cloud, share files, and edit documents together. While not designed as a database, it does share some characteristics with NoSQL repositories, particularly in terms of data storage and retrieval:

1. **Schema-less Data Storage**: Similar to NoSQL databases, OneDrive does not require a fixed schema to store data. Files and folders can be of any type and structure, which is akin to the flexible schemas found in many NoSQL databases.
2. **Data Retrieval**: Files and data stored in OneDrive can be accessed via APIs, similar to how data is retrieved from NoSQL databases. This allows for the integration of

[3] https://www.microsoft.com/en/microsoft-365/onedrive/online-cloud-storage
[4] https://ckan.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

OneDrive storage with applications, albeit the operations and complexity are different from database queries.

3. **Scalability**: Like NoSQL databases, OneDrive is built on a scalable cloud infrastructure. It is designed to store and manage large volumes of data across a distributed network of servers.

These three characteristics completely fulfill the Nexogenesis Data Lake user necessities.

2.1. Data Lake Objective

Effective and fluid communication between Work Packages (WPs) is essential for the success of the NXG project. Given this, several interdependent data pipelines that span multiple WPs were identified in task T4.1 (Figure 2). This situation underscores the need to not only establish a clear and specific plan to manage these interactions, such as the NXG co-creation framework (WP1) or the DMP (WP4), but also to implement digital services (the Data Lake) that can function as a bridge to facilitate communication among them.

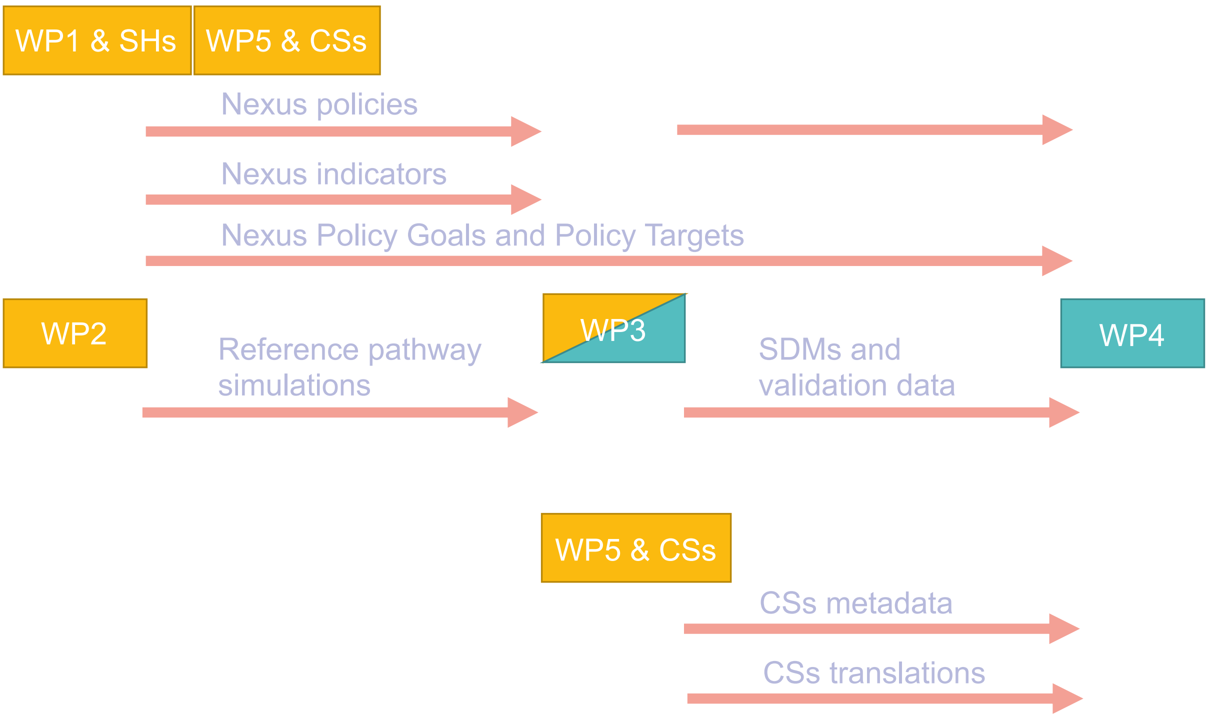


Figure 2. NXG cross-WP data pipelines in the Internal Data Repository. Source: D4.1.

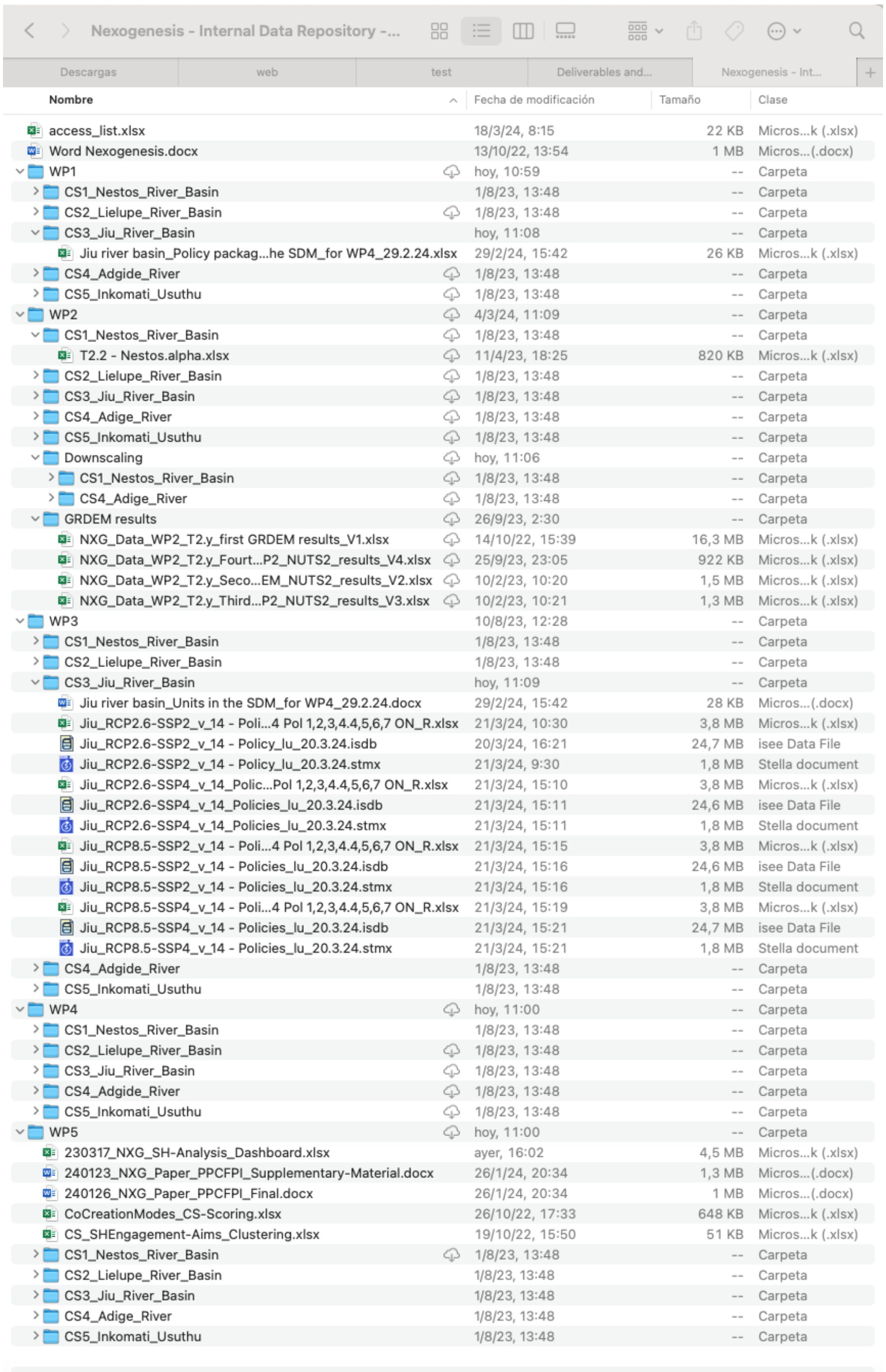
WPs that aim to produce data for other WPs within the project data pipeline were classified as data providers. Similarly, those WPs that need and use the results from other WPs to complete their tasks were identified as data consumers. In this framework, a WP can simultaneously serve as both a data consumer and a data provider, as is the case with WP3.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

2.2. Data Lake organization, protocol and datasets

The Data Lake is organized by folders (Figure 3), with the top-level structure based on WPs, and the second-level structure under each WP folder based on CSs. This setup allows for intuitive navigation throughout the Data Lake, reflecting the overall project organization.

A naming convention protocol has been established to set the rules for naming files. This protocol primarily applies to files that are subject to versioning and can be updated with new releases in the Data Lake.



Nexogenesis - Internal Data Repository -...

Descargas | web | test | Deliverables and... | Nexogenesis - Int...

Nombre	Fecha de modificación	Tamaño	Clase
access_list.xlsx	18/3/24, 8:15	22 KB	Micros...k (.xlsx)
Word Nexogenesis.docx	13/10/22, 13:54	1 MB	Micros...(.docx)
WP1	hoy, 10:59	--	Carpeta
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
CS2_Lielupe_River_Basin	1/8/23, 13:48	--	Carpeta
CS3_Jiu_River_Basin	hoy, 11:08	--	Carpeta
Jiu river basin_Policy packag...he SDM_for WP4_29.2.24.xlsx	29/2/24, 15:42	26 KB	Micros...k (.xlsx)
CS4_Adgide_River	1/8/23, 13:48	--	Carpeta
CS5_Inkomati_Usuthu	1/8/23, 13:48	--	Carpeta
WP2	4/3/24, 11:09	--	Carpeta
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
T2.2 - Nestos.alpha.xlsx	11/4/23, 18:25	820 KB	Micros...k (.xlsx)
CS2_Lielupe_River_Basin	1/8/23, 13:48	--	Carpeta
CS3_Jiu_River_Basin	1/8/23, 13:48	--	Carpeta
CS4_Adige_River	1/8/23, 13:48	--	Carpeta
CS5_Inkomati_Usuthu	1/8/23, 13:48	--	Carpeta
Downscaling	hoy, 11:06	--	Carpeta
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
CS4_Adige_River	1/8/23, 13:48	--	Carpeta
GRDEM results	26/9/23, 2:30	--	Carpeta
NXG_Data_WP2_T2.y_first GRDEM results_V1.xlsx	14/10/22, 15:39	16,3 MB	Micros...k (.xlsx)
NXG_Data_WP2_T2.y_Fourt...P2_NUTS2_results_V4.xlsx	25/9/23, 23:05	922 KB	Micros...k (.xlsx)
NXG_Data_WP2_T2.y_Seco...EM_NUTS2_results_V2.xlsx	10/2/23, 10:20	1,5 MB	Micros...k (.xlsx)
NXG_Data_WP2_T2.y_Third...P2_NUTS2_results_V3.xlsx	10/2/23, 10:21	1,3 MB	Micros...k (.xlsx)
WP3	10/8/23, 12:28	--	Carpeta
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
CS2_Lielupe_River_Basin	1/8/23, 13:48	--	Carpeta
CS3_Jiu_River_Basin	hoy, 11:09	--	Carpeta
Jiu river basin_Units in the SDM_for WP4_29.2.24.docx	29/2/24, 15:42	28 KB	Micros...(.docx)
Jiu_RCP2.6-SSP2_v_14 - Poli...4 Pol 1,2,3,4.4,5,6,7 ON_R.xlsx	21/3/24, 10:30	3,8 MB	Micros...k (.xlsx)
Jiu_RCP2.6-SSP2_v_14 - Policy_lu_20.3.24.isdb	20/3/24, 16:21	24,7 MB	isee Data File
Jiu_RCP2.6-SSP2_v_14 - Policy_lu_20.3.24.stmx	21/3/24, 9:30	1,8 MB	Stella document
Jiu_RCP2.6-SSP4_v_14_Polic...Pol 1,2,3,4.4,5,6,7 ON_R.xlsx	21/3/24, 15:10	3,8 MB	Micros...k (.xlsx)
Jiu_RCP2.6-SSP4_v_14_Policies_lu_20.3.24.isdb	21/3/24, 15:11	24,6 MB	isee Data File
Jiu_RCP2.6-SSP4_v_14_Policies_lu_20.3.24.stmx	21/3/24, 15:11	1,8 MB	Stella document
Jiu_RCP8.5-SSP2_v_14 - Poli...4 Pol 1,2,3,4.4,5,6,7 ON_R.xlsx	21/3/24, 15:15	3,8 MB	Micros...k (.xlsx)
Jiu_RCP8.5-SSP2_v_14 - Policies_lu_20.3.24.isdb	21/3/24, 15:16	24,6 MB	isee Data File
Jiu_RCP8.5-SSP2_v_14 - Policies_lu_20.3.24.stmx	21/3/24, 15:16	1,8 MB	Stella document
Jiu_RCP8.5-SSP4_v_14 - Poli...4 Pol 1,2,3,4.4,5,6,7 ON_R.xlsx	21/3/24, 15:19	3,8 MB	Micros...k (.xlsx)
Jiu_RCP8.5-SSP4_v_14 - Policies_lu_20.3.24.isdb	21/3/24, 15:21	24,7 MB	isee Data File
Jiu_RCP8.5-SSP4_v_14 - Policies_lu_20.3.24.stmx	21/3/24, 15:21	1,8 MB	Stella document
CS4_Adgide_River	1/8/23, 13:48	--	Carpeta
CS5_Inkomati_Usuthu	1/8/23, 13:48	--	Carpeta
WP4	hoy, 11:00	--	Carpeta
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
CS2_Lielupe_River_Basin	1/8/23, 13:48	--	Carpeta
CS3_Jiu_River_Basin	1/8/23, 13:48	--	Carpeta
CS4_Adgide_River	1/8/23, 13:48	--	Carpeta
CS5_Inkomati_Usuthu	1/8/23, 13:48	--	Carpeta
WP5	hoy, 11:00	--	Carpeta
230317_NXG_SH-Analysis_Dashboard.xlsx	ayer, 16:02	4,5 MB	Micros...k (.xlsx)
240123_NXG_Paper_PPCFPI_Supplementary-Material.docx	26/1/24, 20:34	1,3 MB	Micros...(.docx)
240126_NXG_Paper_PPCFPI_Final.docx	26/1/24, 20:34	1 MB	Micros...(.docx)
CoCreationModes_CS-Scoring.xlsx	26/10/22, 17:33	648 KB	Micros...k (.xlsx)
CS_SHEngagement-Aims_Clustering.xlsx	19/10/22, 15:50	51 KB	Micros...k (.xlsx)
CS1_Nestos_River_Basin	1/8/23, 13:48	--	Carpeta
CS2_Lielupe_River_Basin	1/8/23, 13:48	--	Carpeta
CS3_Jiu_River_Basin	1/8/23, 13:48	--	Carpeta
CS4_Adige_River	1/8/23, 13:48	--	Carpeta
CS5_Inkomati_Usuthu	1/8/23, 13:48	--	Carpeta

Figure 3. The Nexogenesis Data Lake. Screenshot of a local synchronization.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

To this day, around 80 datasets related to CSs' policies and goals (WP1 & WP5), Biophysical and socio-economic modelling (Reference Pathways) simulations (WP2), SDMs and SDMs outputs (WP3), Stakeholders (SHs) analysis (WP5) and other resources have been uploaded into the Data Lake and make available for other partners.

2.3. Security, Authentication, Access and Management

The security and authentication layer for the Data Lake relies on services provided by Microsoft. To gain access to the Data Lake, partners must contact the EUT team. Currently, access has been granted to over 40 people.

Finally, EUT offers the Microsoft OneDrive license as part of its technology stack.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

3. Semantic Repository

The Nexogenesis Semantic Repository consists of the Nexogenesis Ontology and Semantic Repository services. Furthermore, the latter comprises the Nexogenesis Contextualizer, the Jena Fuseki SPARQL server, and Nexogenesis Data Explorer and Visualizer.

3.1. Nexogenesis Ontology Service

The Nexogenesis Nexus Ontology service comprises an online tool that enables users to explore the Nexus ontology implemented in Task 4.2, representing information pertinent to the project's scope and its five case studies.

The implemented ontology serves as a foundational framework within the Nexogenesis Knowledge Repository, integrating established ontologies to organize and harmonize data effectively. Leveraging key ontologies such as SAREF, SAREF4WATR, W3C Time, W3C QUDT, SIM4NEXUS, and others, it provides standardized representations for measures, properties, temporal information, geospatial data, units of measure, and nexus components and variables.

It is motivated by the need to integrate information from different data sources and types referring the Nexus and their economics of different agencies and uses-cases around EU. The proposed ontology should be able to harmonize the high variability of terms referring to the same topic or domain (in ontology language called 'Thing'). The ontology will provide a common data understanding of this information based on the context (their use). Moreover, semantic enriched data representation should be aligned with current standards from the integrated domains.

Considering these aspects, the Nexogenesis Nexus Ontology main purpose is to provide a complete ontology (or semantic model) about all knowledge and decision-making processes around water, energy, land, climate and food production including economics and ecosystems.

The scope of the Nexogenesis Nexus Ontology is to represent the information from multiple systems that are observing an heterogeneous number of water, land, energy and climate change variables. These variables will conform some policy evaluation scenarios. The type and the number of these variables will vary from one scenario to others. So, one of the main aspects of the ontology is to make understandable the information collected from the scenarios and to harmonize it under a common model. For that, the ontology needs to follow standard definitions and terms adoptions from representative organizations (WMO, OGC, INSPIRE, W3C, etc).

Moreover, the ontology service needs to support the scenario definition and the human-interaction with the Nexogenesis SLNAE https://nepat-dev.nexogenesis.eu/ (NEPAT) in order

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

to facilitate the evaluation of different policies and the subsequent policy recommendation as an output of the scenario simulation.

The service is publicly available at the following urls: https://nepat-dev.nexogenesis.eu/ontology/ and https://slnae-dev.nexogenesis.eu/ontology/

3.1.1. Ontology

An ontology in the context of knowledge representation is a way to describe and categorize the structure of a domain – that is, the types of entities within that domain and the ways they relate to one another. Ontologies are used to facilitate information sharing and integration across different systems and in areas like artificial intelligence, semantic web, software engineering, and information architecture.

Technically speaking, ontologies are defined based on the following components:

1. **Class**: A class (or a concept) represents a group or a category of things within the domain. For example, in an ontology about the nexus, there might be classes like Nexus Sector, Policy, Economic Property or Biophysical Property. Each class encapsulates entities that share certain characteristics. In object-oriented terms, a class is akin to a blueprint from which instances (individual members of the class) can be created.
2. **Object Property**: Object properties are relationships between two classes or instances of classes. They describe how instances of one class relate to instances of another class. For example, an object property could be "hasProperty", linking the class Nexus Sector to the class Economic Property. This tells you that an instance of Sector may has a property instance of Economic Property.
3. **Data Property**: Data properties (sometimes called datatype properties) link instances of classes to data values. These properties are used to add attributes that describe the class's instances using literal values like strings or numbers. For example, a data property could be "hasValue", which would relate an instance of a Economic Property to a value like 10.
4. **Named Individual**: A named individual is a specific instance of a class. It represents an individual entity with a unique identity within the domain. For example, if the class is Nexus Sector, a named individual could be "Water", a particular sector with characteristics that distinguish it from other sectors.

These elements help define and structure the knowledge within a domain, making it possible to perform automated reasoning, manage information, and integrate data across different systems effectively. Ontologies are essential for semantic technologies because they provide a structured framework that machines can interpret and on which they can act, facilitating tasks such as data integration, natural language processing, and complex decision-making.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

The Nexogenesis Nexus Ontology aims to serve as a foundational framework for organizing and harmonizing data within the Nexogenesis Knowledge Repository. Here are the main features of the Nexogenesis Nexus Ontology:

1. **Integration of Existing Ontologies**: The Nexogenesis Nexus Ontology leverages terms and concepts from several well-established ontologies, including SAREF[5], SAREF4WATR[6], W3C ontology[7] and SIM4NEXUS ontology[8]. This integration ensures compatibility and interoperability with existing semantic resources, facilitating seamless data exchange and integration across different domains.
2. **Representation of Measures and Properties**: Building upon the SAREF ontology and its extension SAREF4WATR, the Nexogenesis Nexus Ontology provides a standardized representation for measures and properties, encompassing variables relevant to the project's scope and its 5 Case Studies. This model enables precise characterization and categorization of data related to various domains, ensuring consistency and clarity in data interpretation and analysis.
3. **Temporal Information Management**: Utilizing the W3C Time ontology[9], the Nexogenesis Nexus Ontology incorporates robust mechanisms for representing timestamps and temporal information. This capability enables the effective handling of time-related data, supporting temporal reasoning and analysis within the knowledge repository.
4. **Units of Measure Standardization**: Drawing from the W3C QUDT ontology[10], the Nexogenesis Nexus Ontology provides a standardized representation of units of measure associated with corresponding variables. This ensures uniformity and consistency in the expression of quantitative data, facilitating accurate measurement and comparison across different datasets and domains.
5. **Nexus Components and Variables Representation**: Inspired by the SIM4NEXUS ontology, the Nexogenesis Nexus Ontology includes models to represent nexus components and variables. This aspect enables the comprehensive characterization of interdependencies and interactions within complex socio-ecological systems, supporting holistic analysis and decision-making processes.

In the following sections, the main design decisions for the Nexus knowledge representation are explained.

3.1.1.1. Property class

One of the key terms that has been modelled in the ontology is the Property, which represents identifiable qualities that can be observed. Based on the saref:Property class, from the saref

[5] https://saref.etsi.org/core/v3.2.1/
[6] https://saref.etsi.org/saref4watr/v1.1.1/
[7] https://www.w3.org/TR/vocab-org/
[8] https://seriousgame.sim4nexus.eu/ontology
[9] https://www.w3.org/TR/owl-time/
[10] https://qudt.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

ontology standard, specific properties (Figure 4) have been included. Several properties available in other standard or domain-specific ontologies, such as s4watr:WaterProperty or s4n:LandUseProperty, have been considered, and additional properties have been created when required (nxg:EcosystemsProperty).

Figure 4. Property types

Furthermore, additional sub properties have been created that allow for a low-level domain specification. For example, the saref4watr:WaterProperty (Figure 5) has been disaggregated into the nxg:ActualWaterUseProperty, nxg:EvapotranspirationProperty, nxg: PotentialWaterUseProperty, nxg:WaterCycleComponentsProperty and nxg:WaterStorageAndMoisureProperty sub properties.

Figure 5. nxg:WaterProperty subtypes

A property is linked to a nxg:Sector through the nxg:hasProperty object property.

3.1.1.2. Measurement class

A saref:Property relates to a saref:Measurement through the saref:relatesToMeasurement object property. Similarly, a saref:Measurement is related to a saref:Property through the saref:relatesToProperty object property.

A saref:Measurement represents the measured value obtained over a property. It is also associated with a unit of measure (saref:UnitOfMeasure), through the saref:isMeasuredIn object property, which specifies the unit in which the value is expressed, along with the timestamp of the measurement.

Furthermore, the saref:Measurement type has been extended to encompass additional information corresponding to different future scenario projections. Consecuently, the saref:Measurement object has been complemented with additional data properties (e.g. nxg:fromClimateDriver, nxg:fromModel, nxg:fromRCPSSP or nxg:fromRun) to determine this information (Figure 6).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

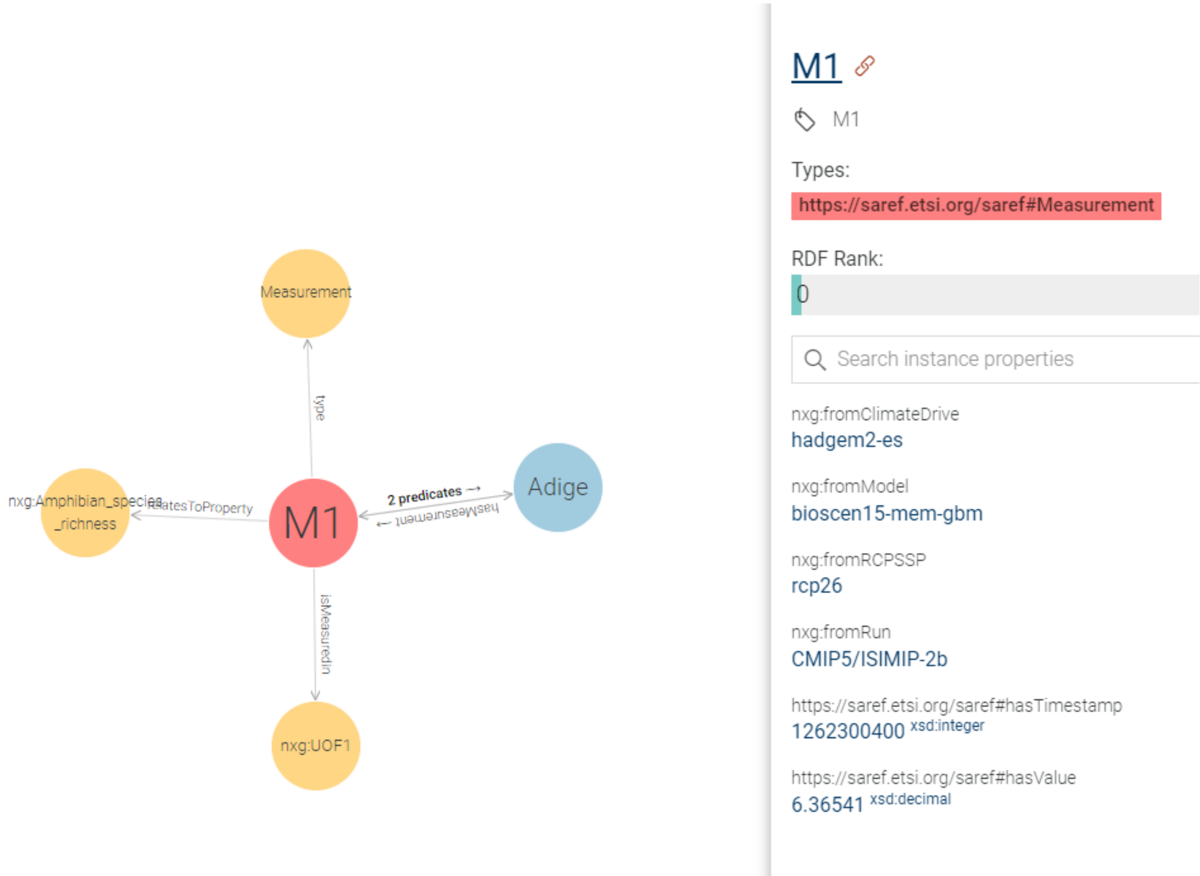


Figure 6. saref:Measurement Named Individual example

Finally, each saref:Measurement is also linked to a Nexogenesis CS or region to georeference its location and source. In these cases, the nxg:measuredInCS and nxg:measuredInRegion object properties have been defined. Similarly, in order to gather all the measurements corresponding to a CS or a region, the saref:hasMeasurement object property has been included as well.

3.1.1.3. FeatureOfInterest class

A feature of interest represents any real world entity from which a property or a state may be acted upon, such as observed and controlled. An instance of saref:FeatureOfInterest represents one specific real world entity.

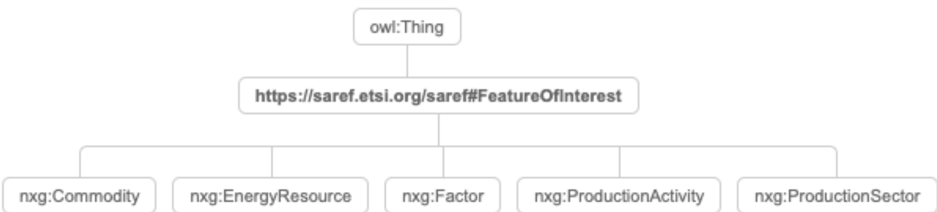


Figure 7. saref:FeatureOfInterest class and its subclasses

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

On this basis, additional sub classes (Figure 7) have been created to represent different types of entities related to the Nexus and required by the project. These classes are nxg:Commodity, nxg:EnergyResource, nxg: Factor, nxg:ProductionActivity and nxg:ProductionSector. Next, the required named individuals of these types have been created (Figure 8).

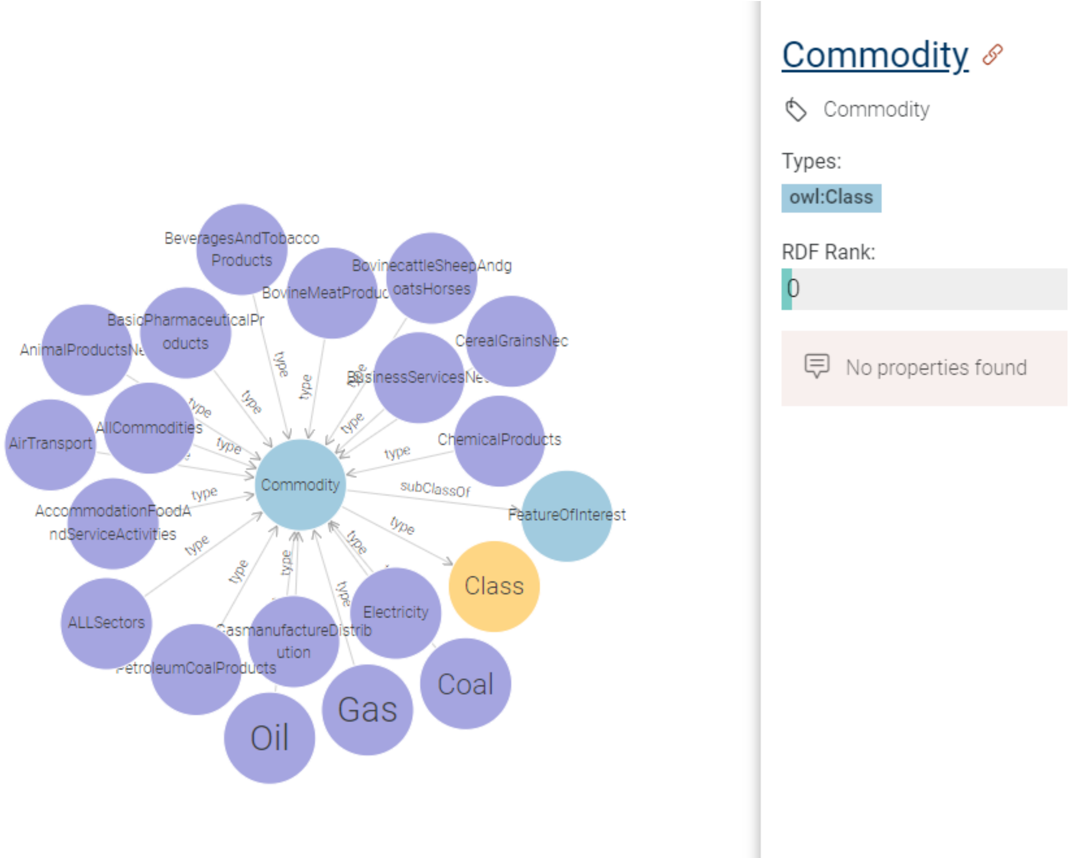


Figure 8. nxg:Commodity named individuals

A property can apply to different features of interest. This link is implemented through the saref:isPropertyOf object property, and its inverse saref:hasProperty.

Furthermore, a feature of interest is also linked to zero or several measurements. This link is implemented through the saref:hasMeasurement object property, and its inverse saref:isMeasurementOf.

3.1.1.4. Policy and PolicyGoal classes

The s4n:Policy and s4n:PolicyGoal objects have been updated (additional data properties have been included and others are not used) in order to fit the Nexogenesis necessities for knowledge representation. A s4n:Policy and a s4n:PolicyGoal are linked to a nxg:Sector through the nxg:hasSector object property and the inverse link is represented by nxg:hasPolicy and nxg:hasPolicyGoal respectively. A s4n:Policy and a s4n:PolicyGoal are linked to a nxg:CaseStudy through the nxg:hasCaseStudy object property and the inverse link is

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

represented by nxg:hasPolicy and nxg:hasPolicyGoal as well. Finally, a s4n:Policy is linked to a s4n:PolicyGoal through the saref4city:isKPIOf object property and the inverse link is represented saref4city:hasKPI.

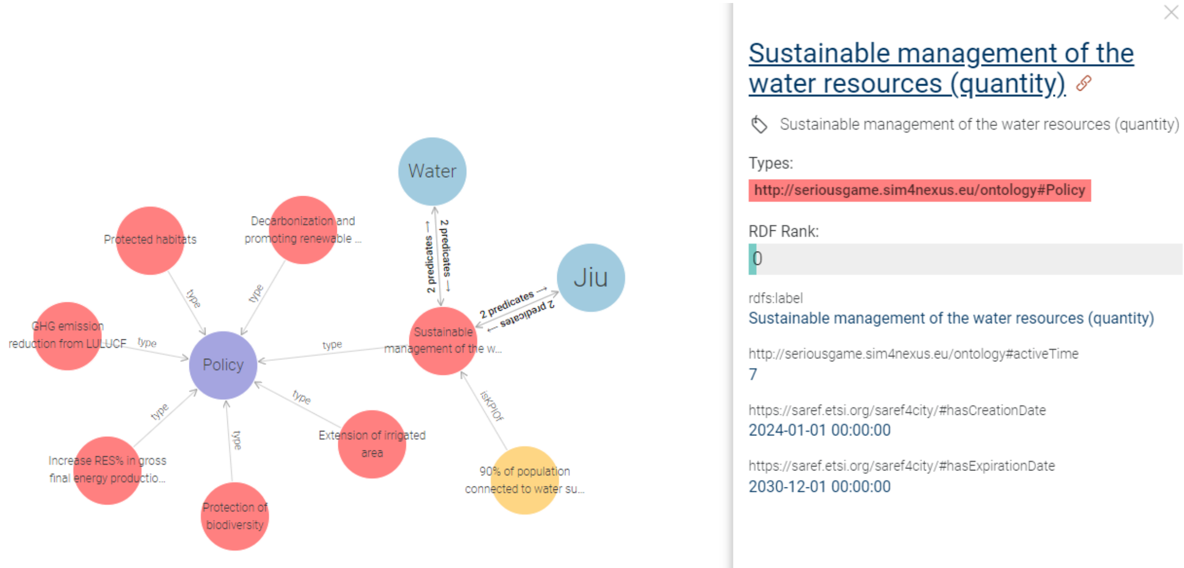


Figure 9. s4n:Policy class and named individuals for Jiu CS

3.1.1.5. Case Study and Region classes

The nxg:CaseStudy and nxg:Region classes have been introduced to georeference the information and classify it, aligning with the 5 Nexogenesis CSs. Several object properties link them with other classes as it has been introduced previously. Additionally, a case study has regions. This link is implemented through the nxg:hasRegion object property. Also, a region may have sub-regions (represented as nxg:SubRegion). This link is implemented by the nxg:hasSubRegion object property.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

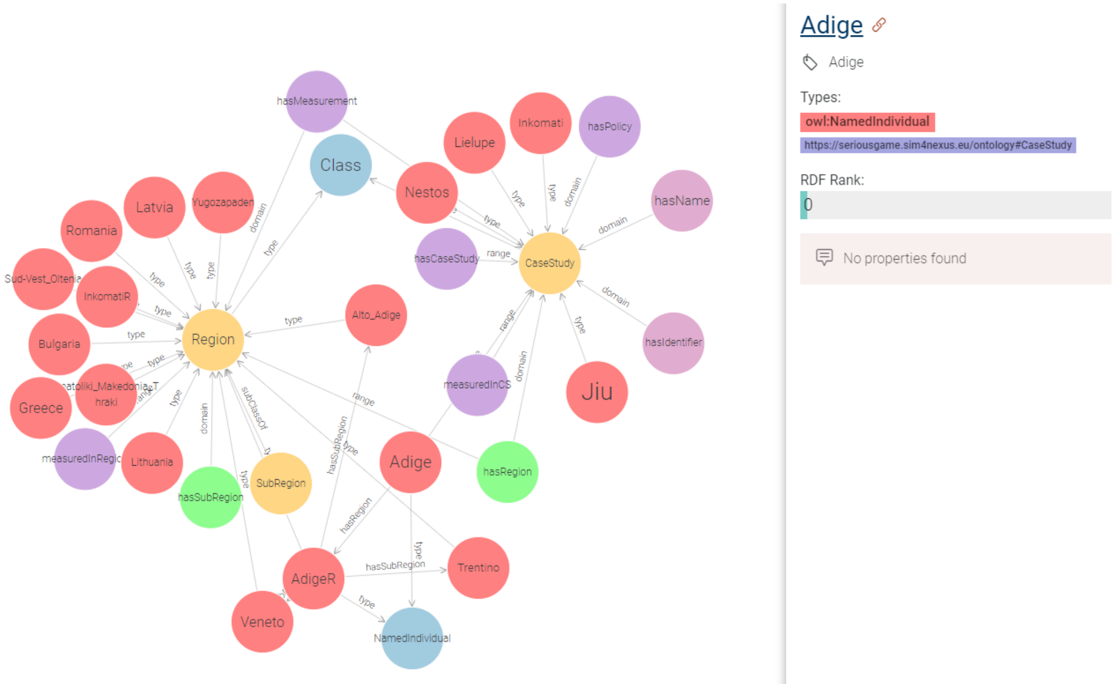


Figure 10. nxg:CaseStudy and nxg:Region classes

3.1.1.6. Sector class

Finally, the nxg:Sector type represents a Nexus Sector. This type acts as a knowledge core entity from which other classes make domain sense. It is linked to almost all the other classes as it has been described in each of the previous sections. The nxg:Biodiversity, nxg:Biome, nxg:Economy and nxg:Ecosystems named individuals been created to complement those available in s4n ontology, s4n:Agriculture, s4n:Energy, s4n:Climate, s4n:Food, s4n:Water and s4n:Land. Finally, a sector may be a sub-sector of another sector; for instance, biome is a subsector of ecosystems. This relation is represented through the nxg:hasSubSector object property.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

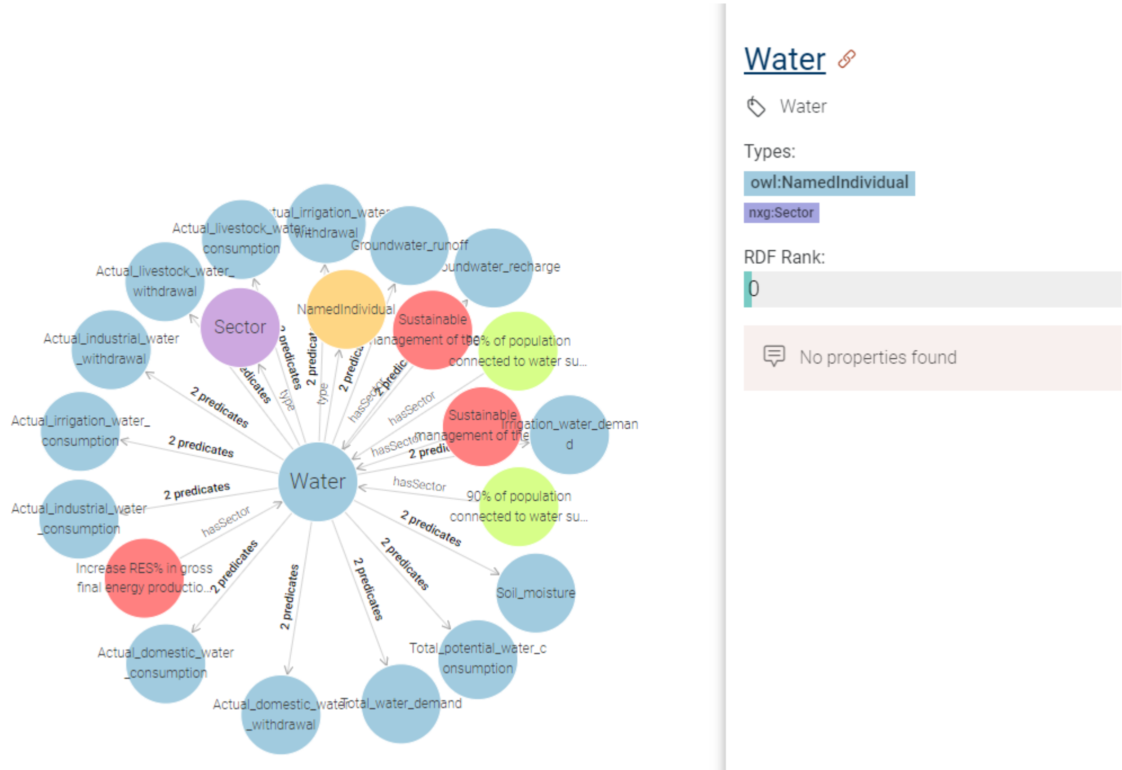


Figure 11. nxg:Water named individual of type nxg:Sector

3.1.2. Service design, architecture and technology behind

The ontology has been designed in compliance with industry standards, ensuring semantic interoperability and reusability across different domains. To achieve this, well-established frameworks like the Web Ontology Language (OWL) [1], and the Resource Description Framework (RDF) [2] have been used, which support the formal representation of data structures and relationships. In addition, key concepts were defined using a robust taxonomy, providing a structured hierarchy that facilitates logical inference and consistent data interpretation. This development process involved thorough schema validation and extensive use of standard vocabularies, enhancing the ontology's capacity to integrate seamlessly with various systems and applications.

The ontology has been developed (Figure 12) with Protégé[11], a versatile, open-source software platform designed for creating, editing, and managing ontologies and knowledge bases. Developed at Stanford University, Protégé offers a user-friendly interface and supports various ontology languages, such as OWL and RDF. It facilitates the visualization of complex data structures and relationships, allowing users to define classes, properties, and instances within a logical framework. With its plug-in architecture, Protégé is extensible, enabling developers to

[11] https://protege.stanford.edu/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

add custom functionality to meet specific needs. It's a widely used tool in the Semantic Web community for ontology development, knowledge engineering, and semantic data integration.

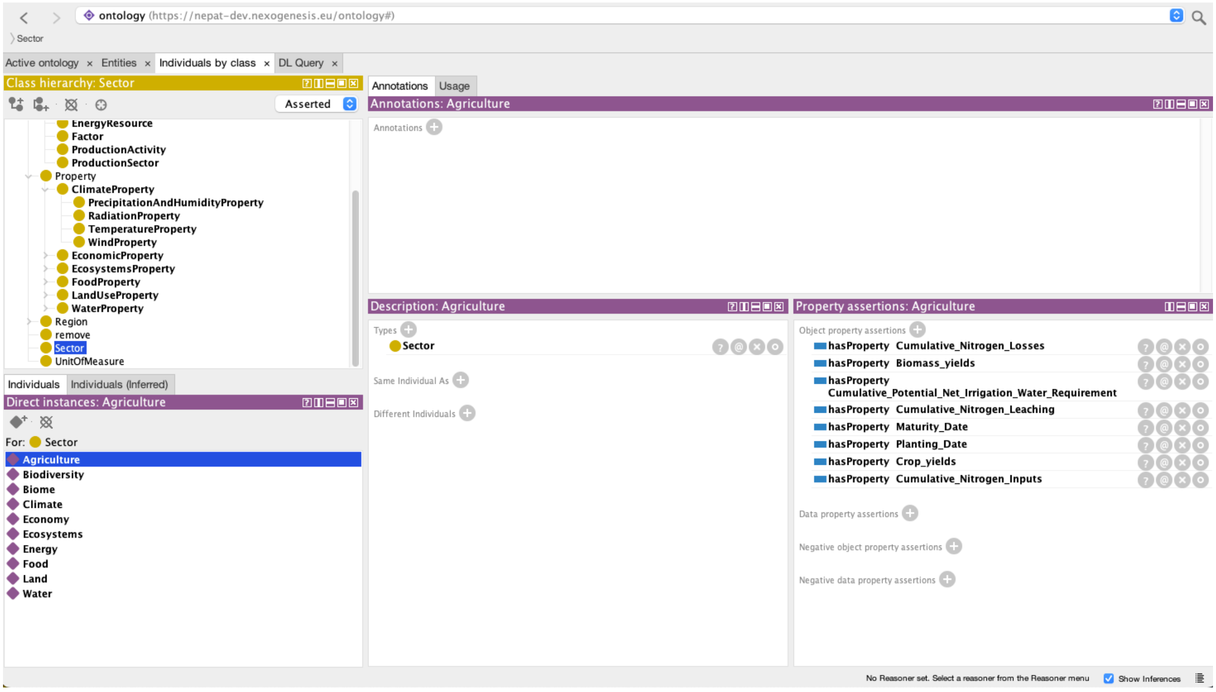


Figure 12. Nexogenesis Nexus ontology development with Protégé software

The ontology is persisted in .ttl format (Terse RDF Triple Language), a concise text-based syntax used for representing RDF data. It simplifies the notation for RDF triples, where each triple consists of a subject, predicate, and object, forming the fundamental structure of RDF data. The TTL format is designed to be more human-readable than other RDF formats like RDF/XML, making it easier to write and debug RDF statements. It supports compact representation through abbreviations, namespaces, and the use of common prefixes, streamlining the process of defining complex data relationships in semantic web applications and ontologies.

Next, the ontology documentation is generated as an HTML document through the Widoco software[12]. HTML (HyperText Markup Language) is the standard markup language used to create and structure content on the web, allowing for the inclusion of text, images, links, and multimedia elements to build interactive and visually appealing web pages. Widoco is a comprehensive documentation tool designed for generating user-friendly documentation for ontologies. It simplifies the process of creating detailed and structured descriptions of ontologies, focusing on clarity and accessibility. Widoco can automatically generate human-readable documentation that includes ontology metadata, descriptions of classes, properties, and individuals, as well as diagrams illustrating their relationships. This tool supports various export formats and is particularly useful for maintaining and sharing ontologies with a wider audience.

[12] https://github.com/dgarijo/Widoco

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

Finally, the generated HTML document is accessible behind the NEPAT Nginx inverse proxy, which manages the incoming requests to the https://slnae-dev.nexogenesis.eu or https://nepat-dev.nexogenesis.eu subdomains, both directing to the NEPAT (SLNAE) tool. The ontology documentation is accessible via the /ontology resource, thus it can be accessed at the URLs https://slnae-dev.nexogenesis.eu/ontology and https://nepat-dev.nexogenesis.eu/ontology.

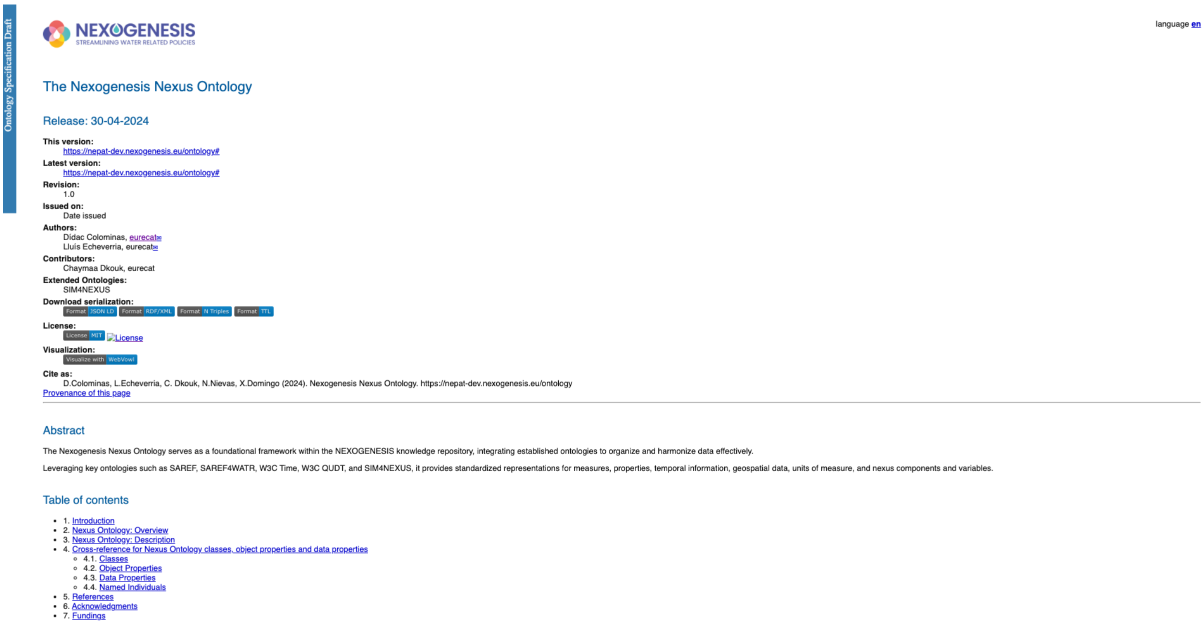


Figure 13. Nexogenesis Nexus ontology online documentation

Additionally, alongside the documentation, there are links to download the ontology in JSON-LD, RDF/XML, NTriples, or TTL formats. The ontology is under the MIT licence and can be cited using the following reference: D.Colominas, L.Echeverria, C. Dkouk, N.Nievas, X.Domingo (2024). Nexogenesis Nexus Ontology. https://nepat-dev.nexogenesis.eu/ontology

3.2. Semantic Repository Service

The Nexogenesis Semantic Repository service aims to provide a common platform for stakeholders to interact with Nexus data, fostering improved integration, contextualized data analysis, and better decision-making.

It enables a specialized database system and mechanisms designed to manage and retrieve data based on semantic relationships defined by the Nexogenesis Nexus Ontology. Its primary objective is to provide a structured representation of data, allowing for enhanced data integration and interoperability. By using semantic metadata, the repository can define the relationships between different pieces of information, facilitating seamless integration from multiple sources. This structure also enables complex query capabilities, allowing users to retrieve information based on semantic contexts, which would be challenging to achieve in traditional databases.

In addition to structured data representation, the Semantic Repository service supports advanced knowledge management, promoting knowledge exploration, sharing and collaboration. It helps capture and store semantic relationships, making it easier to understand, search, and retrieve relevant information.

The Nexogenesis Semantic Repository is composed by The Nexogenesis Contextualizer, the SPARQL server, and the Nexogenesis Data Explorer and Visualizer.

So far, the available data in the Semantic Repository corresponds to Policies, Goals and biophysical and socio-economic modelling simulations. Additional data will be uploaded as they become available.

3.2.1. Semantic Repository architecture

The architecture of the Semantic Repository Service comprises three key components: the Contextualizer, a SPARQL server, and the data Explorer and Visualizer. Each component plays a distinct yet integral role in facilitating the storage, management, and visualization of semantic data within the repository.

- **Contextualizer:** At the starting point of the architecture is the Contextualizer, tasked with the responsibility of transforming raw data from diverse project sources (e.g. biophysical and socio economic data from WP2) into a semantically enriched format. Through diverse data processing mechanisms, the Contextualizer identifies relevant entities, attributes, and relationships within the datasets and maps them to corresponding concepts and properties within the domain ontology. This process establishes the necessary contextual framework that underpins meaningful data interpretation and analysis within the repository.
- **SPARQL Server**: Serving as the backbone of the Semantic Repository Service is the SPARQL (SPARQL Protocol and RDF Query Language) server renowned for its robust support of Semantic Web technologies. SPARQL is a query language and protocol used to retrieve and manipulate data stored in Resource Description Framework (RDF) format, commonly used in semantic web technologies and linked data applications. The Nexogenesis SPARQL service is based on the Jena Fuseki [13]software, and provides the infrastructure for storing and querying semantic data using the SPARQL query language. Leveraging its capabilities, users can execute complex queries against the repository to retrieve specific information, explore relationships between entities, and extract valuable insights from the underlying dataset. The Fuseki server ensures efficient data storage and retrieval, enabling seamless integration with other components of the architecture.
- **Data Explorer and Visualizer**: Completing the architecture is the data Explorer and Viualizer, an intuitive interface designed to facilitate user interaction and exploration of

[13] https://jena.apache.org/documentation/fuseki2/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

the semantic data stored within the repository. It is implemented by GraphDB[14] software. Through visually engaging displays and interactive features, the Visualizer empowers users to navigate the repository, visualize complex data relationships, and gain deeper insights into the underlying dataset. Whether through graphical representations, charts, or customizable dashboards, the Visualizer provides users with the tools and capabilities to interact with semantic data effectively, facilitating informed decision-making and knowledge discovery.

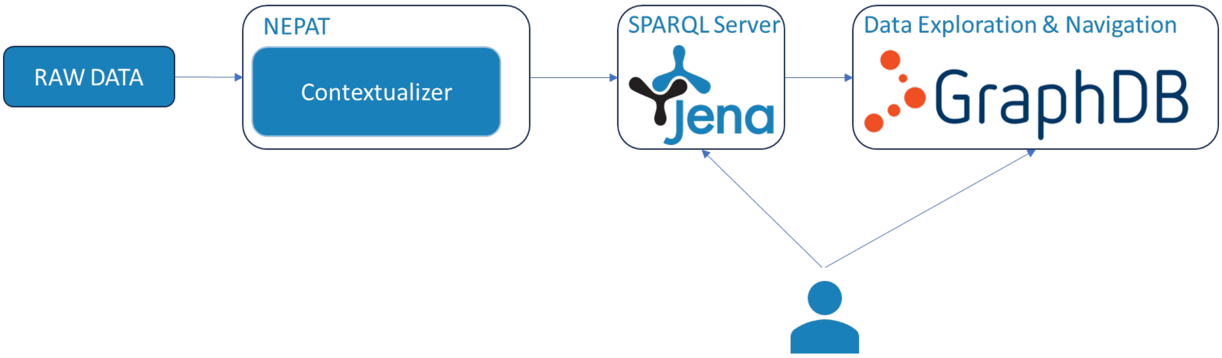


Figure 14. Semantic Repository architecture

3.2.1.1. Nexogenesis Contextualizer

The Contextualizer, as part of the NEPAT project, serves as a distinct module within our system, designed specifically to extract data from various project sources. Its primary function is to generate the contextual framework necessary for efficient data management and knowledge disovery.

The Contextualizer undertakes the task of transforming raw data, often sourced from different formats, schemas and origins, into a structured format conducive to semantic analysis. This transformation involves translating data from formats such as Excel spreadsheets or CSV (Coma Separated Value) into triples, a fundamental data structure in semantic web technology, facilitating seamless integration and interoperability.

Moreover, the Contextualizer goes beyond mere data translation by establishing meaningful connections between the extracted information and the overarching domain Nexogenesis Nexus Ontology, detailed in previous sections. By linking data entities to the ontology, the Contextualizer enriches the semantic understanding of the dataset, establishing relationships and contextual associations that enhance the overall coherence and depth of analysis.

[14] https://graphdb.ontotext.com/

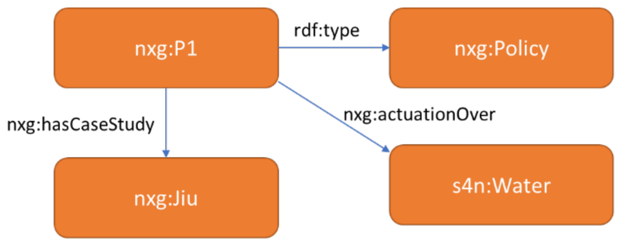
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

```
Starting contextualizing POLICIES
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
                            prefix owl: <http://www.w3.org/2002/07/owl#>
                            prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
                            PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
                            PREFIX s4n: <http://seriousgame.sim4nexus.eu/ontology#>
                            PREFIX saref: <https://saref.etsi.org/saref#>
                            PREFIX saref4city: <https://saref.etsi.org/saref4city/#>
                            PREFIX nxg: <https://nepat-dev.nexogenesis.eu/ontology#>

INSERT DATA {
  nxg:P1 rdf:type s4n:Policy ;
  nxg:hasCaseStudy nxg:Jiu;
  nxg:hasSector s4n:Water ;
rdfs:label "Sustainable management of the water resources (quantity)";
s4n:activeTime "7";
saref4city:hasCreationDate "2024-01-01 00:00:00";
saref4city:hasExpirationDate "2030-12-01 00:00:00";
};
```

Figure 15. Example of RDF code generated by the Contextualizer to contextualize a Policy entity

In the process of contextualizing data (Figure 15 and Figure 16), one of the principal considerations is identifying and establishing the appropriate relationships and entities that effectively link to the ontology. This task is crucial for ensuring the integrity and coherence of the semantic model, as well as enabling meaningful data analysis and interpretation.



As an example

	property	value
1	rdf:type	s4n:Policy
2	nxg:hasCaseStudy	nxg:Jiu
3	nxg:hasSector	s4n:Water
4	rdfs:label	"Sustainable management of the water resources (quantity)"
5	s4n:activeTime	"7"
6	saref4city:hasCreationDate	"2024-01-01 00:00:00"
7	saref4city:hasExpirationDate	"2030-12-01 00:00:00"

Figure 16. Policy entity contextualization

As part of the NEPAT project, the Nexogenesis Contextualizer is deployed alongside the NEPAT tool. When a new data resource is detected, either corresponding to a new creation or a modification of a previous resource, EUT triggers a new contextualizer process to upload it into the Semantic Repository. If necessary, additional data processing steps are developed and integrated into the Contextualizer.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

3.2.1.2. Nexogenesis Semantic Repository Core

For the deployment of the Nexogenesis Semantic Repository Core, a robust solution centred around an Apache Jena Fuseki server has been deployed and configured. This choice was made for its comprehensive support of SPARQL 1.1 protocols[15], facilitating both query and update functionalities, alongside the SPARQL Graph Store protocol[16], crucial for managing graph data effectively.

The deployment configuration integrates Fuseki seamlessly into the Semantic Repository broader architecture, encompassing the Contextualizer, the Semantic Repository core, and the data Explorer and Visualizer components.

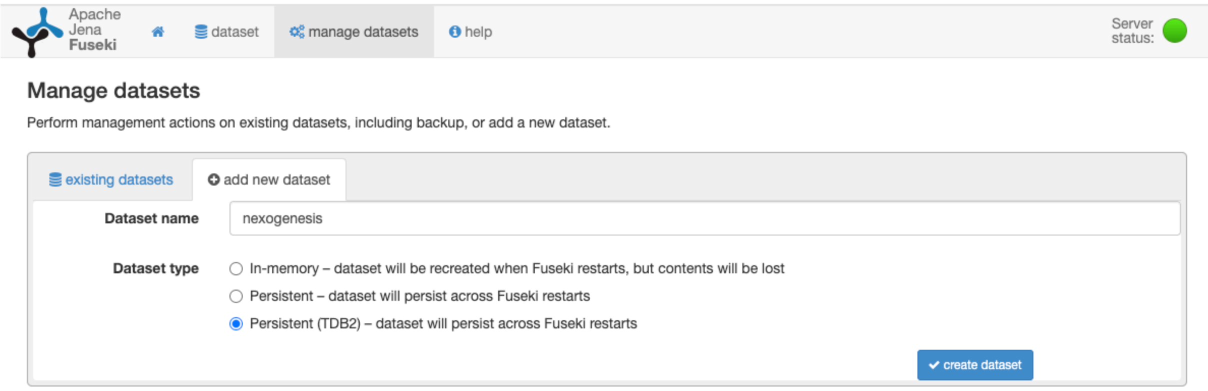


Figure 17. The nexogenesis TDB2 dataset initialization

The ‘nexogenesis’ dataset has been created to store the semantic information (Figure 17). It is a persistent dataset of TDB2 type[17]. TDB2 can be used as a high-performance RDF store on a single machine.

[15] https://www.w3.org/TR/sparql11-protocol/
[16] https://www.w3.org/TR/sparql11-http-rdf-update/
[17] https://jena.apache.org/documentation/tdb2/

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003881

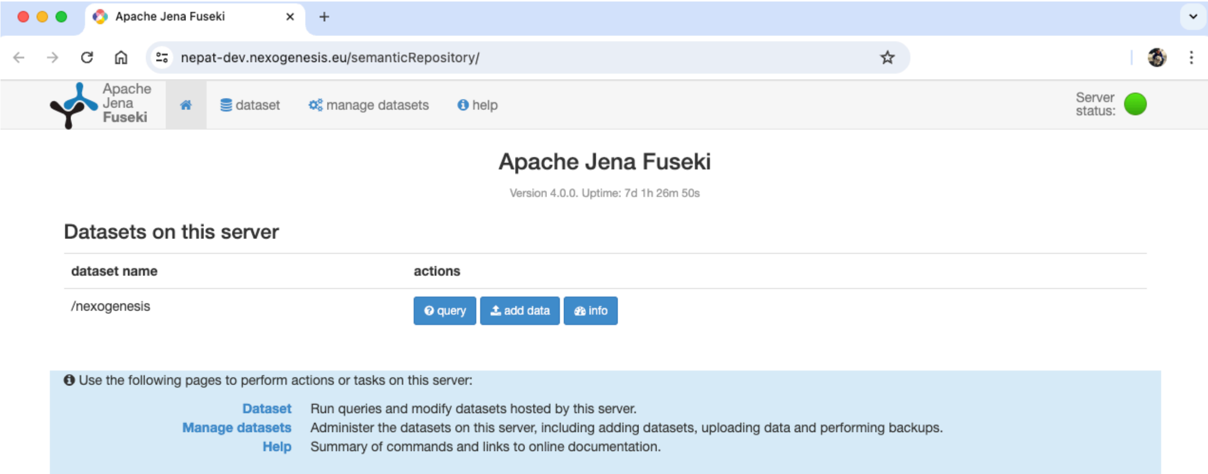


Figure 18. Nexogenesis SPARQL server

The SPARQL server access is managed by the NEPAT Nginx inverse proxy as well. The server is accessible via the /semanticRepository resource, thus it can be accessed at the URLs https://nepat-dev.nexogenesis.eu/semanticRepository/ and https://slnae-dev.nexogenesis.eu/semanticRepository/ .

3.2.1.3. Nexogenesis Data Explorer and Visualizer

A data exploration and navigation tool has been deployed, based on a GraphDB instance, which connects directly to the SPARQL server to access and retrieve the semantic data.

The tool facilitates the exploration and navigation of data by interfacing seamlessly with the SPARQL server, which acts as the primary repository for the semantic data. By connecting to the SPARQL server, the GraphDB instance can fetch data dynamically, allowing users to perform in-depth data analysis and visualization through advanced querying capabilities.

The data exploration tool offers several mechanisms that facilitate comprehensive analysis:

- **Graph Overview**: This section provides users with a high-level view of the default graph and enables targeted searches for specific graphs within the dataset. It serves as an entry point for exploring the dataset's structure and contents.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

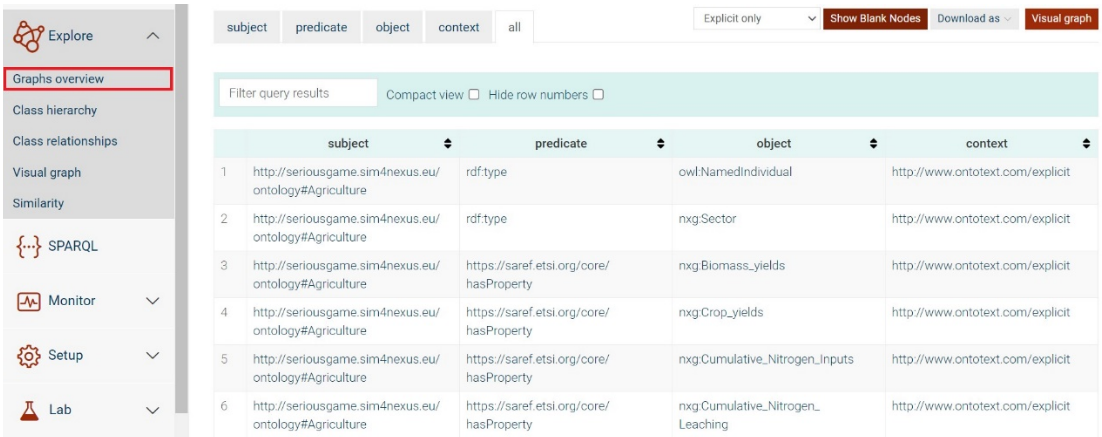


Figure 19. Graph Overview

- **Class Hierarchy**: Users can delve into the hierarchical structure of classes based on the number of instances associated with each class. Parent classes are represented by larger circles, while nested subclasses are depicted as smaller circles. This visualization aids in understanding the organization and distribution of data across different class hierarchies.

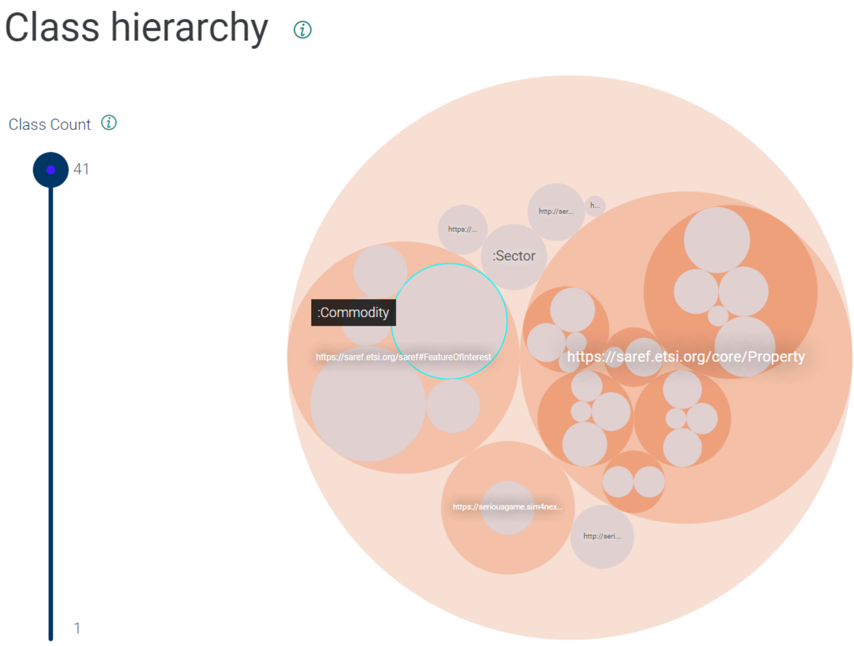


Figure 20. Class hierarchy. Named individuals for saref:Mesurement are not shown

- **Class Relationships**: This section illustrates the relationships between various classes within the dataset. Each relationship is depicted as a link between individual instances of two classes, forming RDF statements. The subject represents an instance of one class, the object denotes an instance of another class, and the link serves as the predicate. This visualization elucidates the interconnectedness of different data entities within the dataset.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

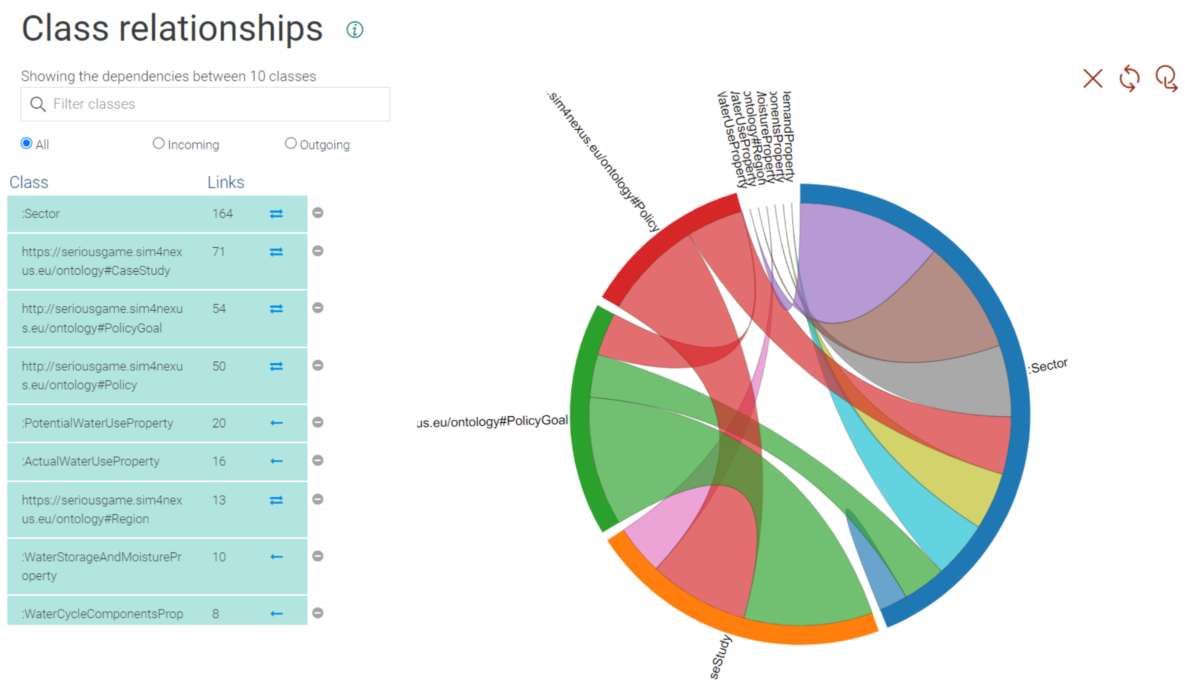


Figure 21. Class relationships

- **Visual Graph**: Users can explore a visual representation of specific parts of the graph. The visualization begins from a single resource or from the result of a graph query. By double-clicking on resources, users can expand the graph to reveal their connections, facilitating seamless navigation through the dataset. This interactive feature empowers users to explore the dataset's structure and relationships intuitively.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

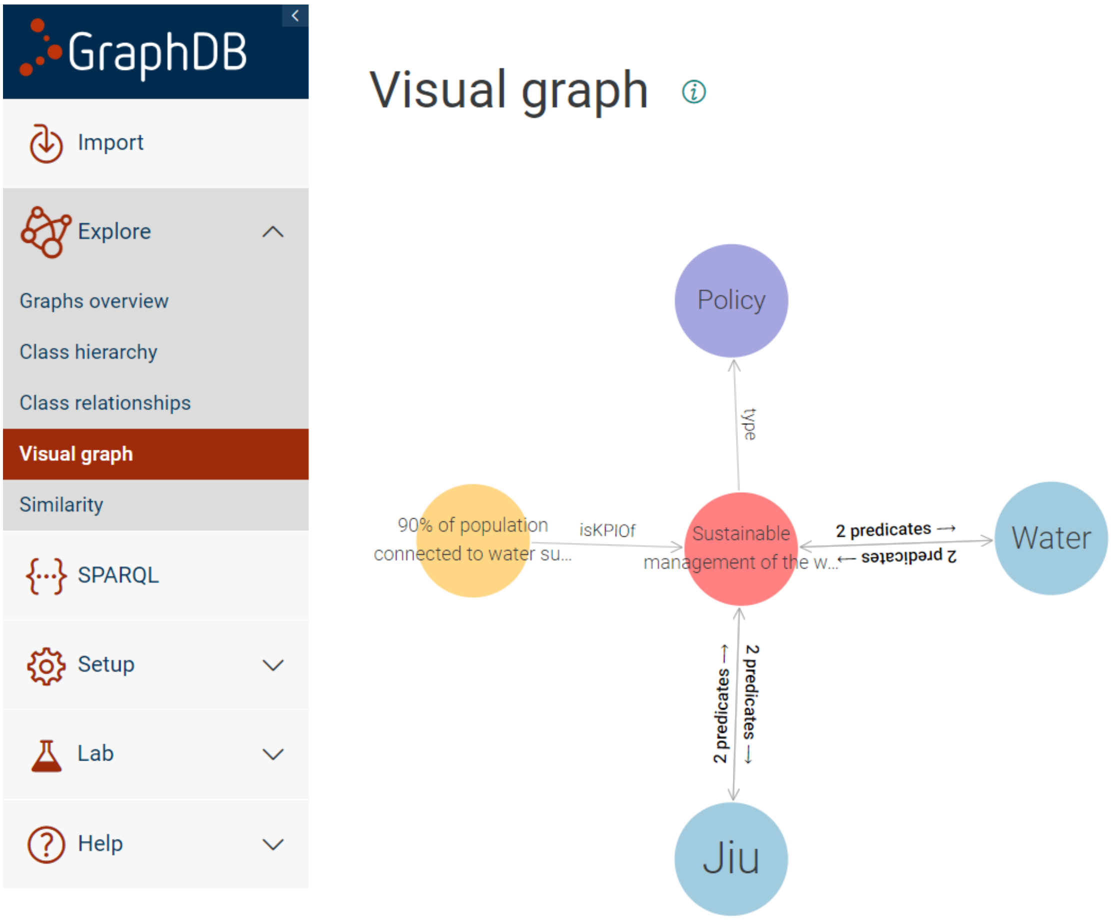


Figure 22. Visual Graph for entity nxg:P1

The Data Explorer and Visualizer access is managed by the NEPAT Nginx inverse proxy as well. The server is accessible via the /visualizer resource, thus it can be accessed at the URLs https://nepat-dev.nexogenesis.eu/visualizer/ and https://slnae-dev.nexogenesis.eu/visualizer/ .

3.2.1.4. Architecture and systems security

Stringent security measures have been configured to safeguard the various aspects of data edition within the system and architecture, including data insertion, management, and removal. These critical functionalities are fortified by dual layers of protection: access to the system User Interfaces (UIs) over HTTPS, and via secured SPARQL endpoints. HTTPS (HyperText Transfer Protocol Secure) is a secure version of HTTP that encrypts data transmitted between a user's browser and a website, providing confidentiality and integrity through SSL/TLS protocols, two cryptographic protocols that enable secure network communication.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

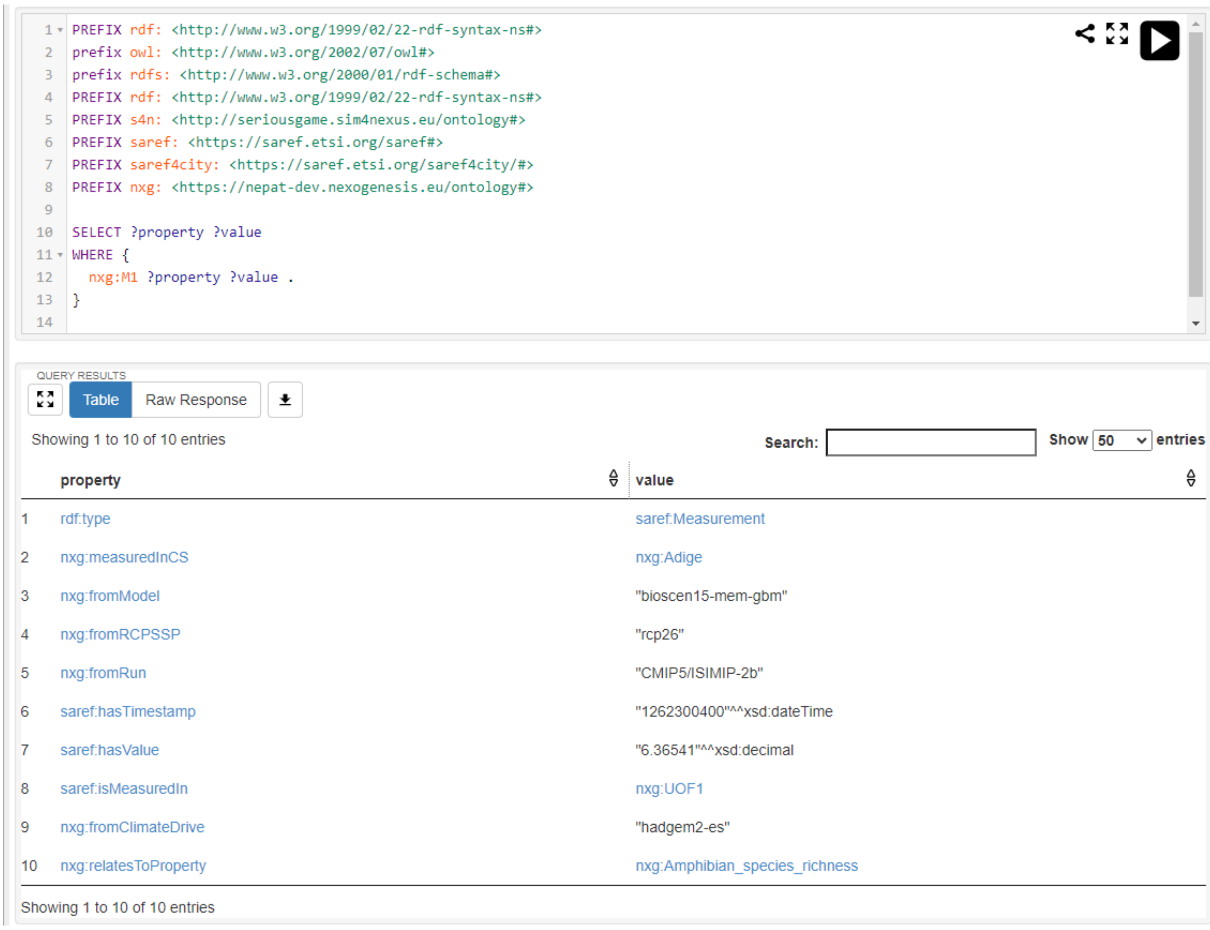


Figure 23. SPARQL query example through an open Jena Fuseki UI section

To facilitate seamless access to data querying functionalities, an open approach (Figure 23) has been enabled for accessing the endpoint dedicated to data consultation. This endpoint is freely accessible to anonymous users, enabling the general public to effortlessly retrieve information through either the data Explorer and Visualizer UI or by directly executing SPARQL queries.

EUT possesses the admin user credentials for system management (Figure 24 and Figure 25).

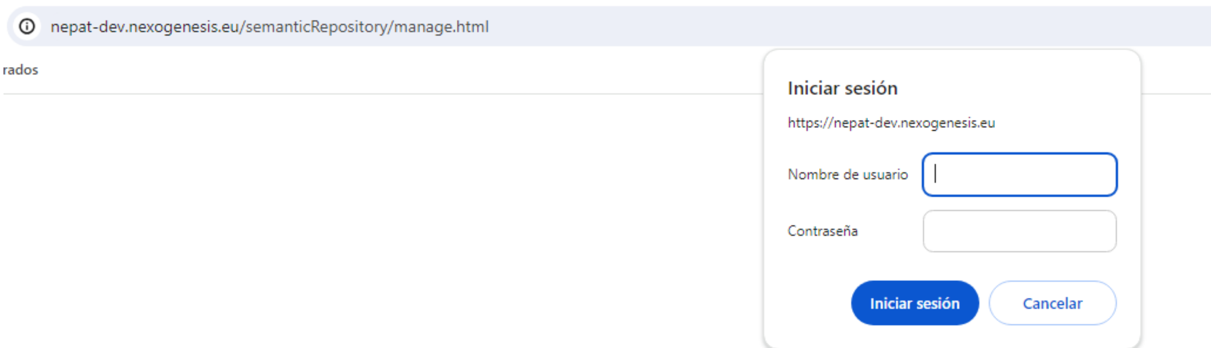


Figure 24. Secure SPARQL server section

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

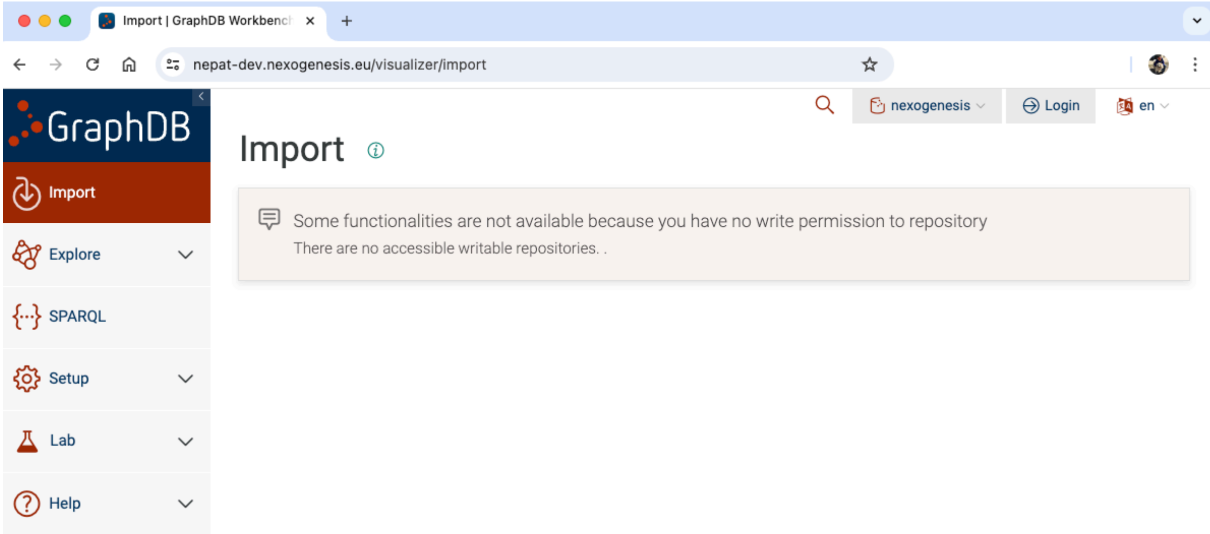


Figure 25. Secure Data Explorer and Visualizer section

3.2.2. The Nexogenesis Semantic Repository for Nexus exploration

For data exploration, users will primarily interact with the system through two distinct GUI interfaces, although they also have the option to perform SPARQL requests directly to the endpoint.

- **Graphical User Interface**: The system provides a graphical interface that interfaces with the GraphDB instance. This GUI offers intuitive tools and visualizations to facilitate exploration and analysis. Users can navigate through different sections such as the graph overview, class hierarchy, and visual graph representation to gain insights into the dataset's structure and relationships.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

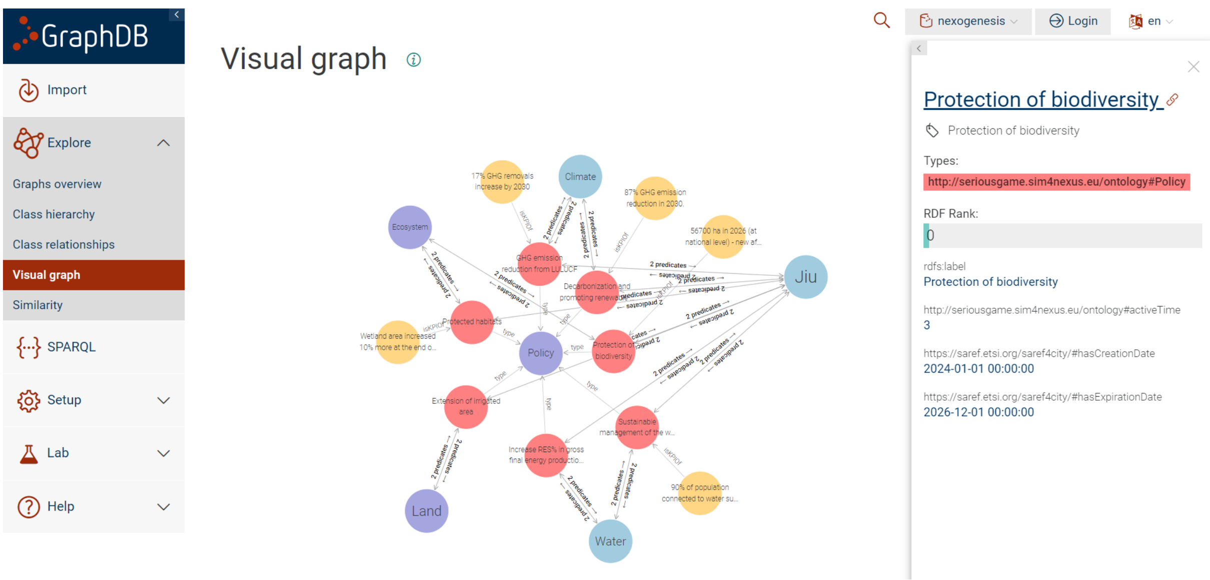
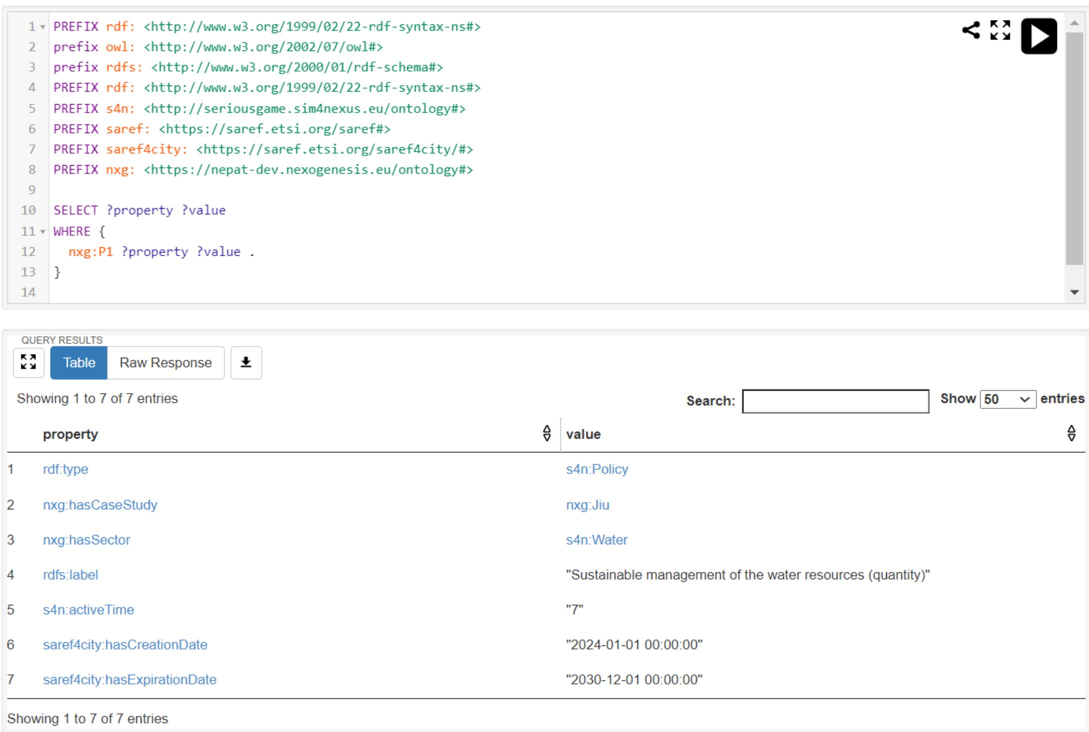


Figure 26. Visualizing Policies (Extending from the class)

- **SPARQL Requests via GUI**: Additionally, users can leverage the GUI interface of Jena Fuseki to execute SPARQL queries directly against the dataset. This functionality enables users to formulate custom queries to retrieve specific information from the dataset. Common queries include obtaining a list of different classes present in the dataset or retrieving data related to a single entity of interest.



```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix owl: <http://www.w3.org/2002/07/owl#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX s4n: <http://seriousgame.sim4nexus.eu/ontology#>
PREFIX saref: <https://saref.etsi.org/saref#>
PREFIX saref4city: <https://saref.etsi.org/saref4city/#>
PREFIX nxg: <https://nepat-dev.nexogenesis.eu/ontology#>

SELECT ?property ?value
WHERE {
  nxg:P1 ?property ?value .
}

```

QUERY RESULTS

Table | Raw Response

Showing 1 to 7 of 7 entries

Search: Show 50 entries

	property	value
1	rdf:type	s4n:Policy
2	nxg:hasCaseStudy	nxg:Jiu
3	nxg:hasSector	s4n:Water
4	rdfs:label	"Sustainable management of the water resources (quantity)"
5	s4n:activeTime	"7"
6	saref4city:hasCreationDate	"2024-01-01 00:00:00"
7	saref4city:hasExpirationDate	"2030-12-01 00:00:00"

Showing 1 to 7 of 7 entries

Figure 27. Get data from single entity via SPARQL Request

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

A complex query is presented in in Figure 28, where all the available saref:Measurement corresponding to the nxg:Amphibian_species_richness property and the nxg:Adige CS are gathered.

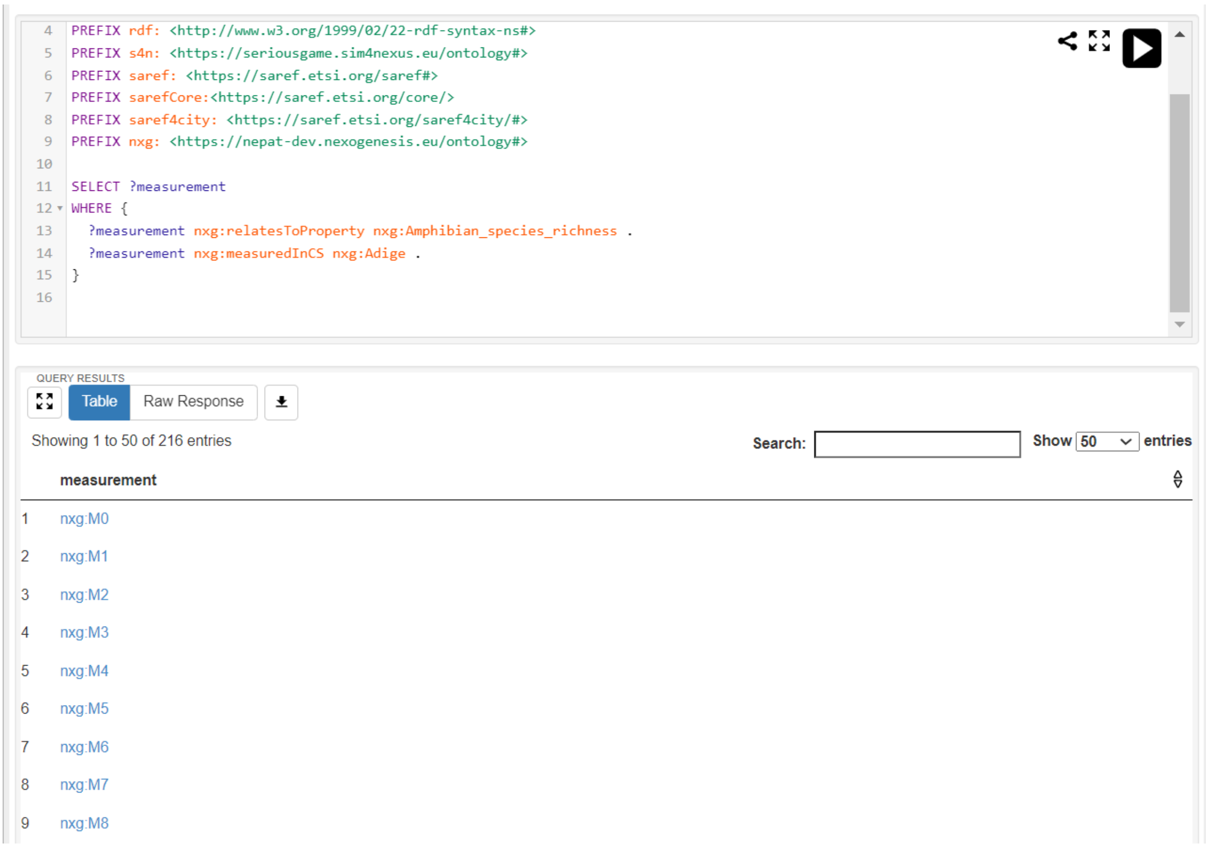


Figure 28. Get all Measurements from the property Amphibian Species Richness from the Case Study Adige

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

4. Conclusions

The **Nexogenesis Knowledge Repository** has been successfully implemented and deployed, reaching one of the main objectives of the task T4.2 and WP4. It is composed by two sub repositories, the **Nexogenesis Data Lake** (the Internal Data Repository), supporting intra-project data exchange, and the **Nexogenesis Semantic Repository**, for external data publication and nexus knowledge sharing. Together, both components are also known as the **Nexogenesis Data Sharing** tools. The Nexogenesis Semantic Repository embraces both the **Nexogenesis Ontology** and **Semantic Repository services**. Additionally, the latter includes several components: the **Nexogenesis Contextualizer**, the **SPARQL server**, and the **Nexogenesis Data Explorer and Visualizer**.

The second and final objective of T4.2 is to publish Nexogenesis and Nexus-related data in the Semantic Repository. Currently, the data available include information on Policies, Goals, and biophysical and socio-economic modeling simulations. Additional data will be uploaded as it becomes available during the remaining project period.

The **Nexogenesis Knowledge Repository** public tools are available at the following urls:

- **Nexogenesis Nexus Ontology**: https://nepat-dev.nexogenesis.eu/ontology/.
- **Nexogenesis SPARQL server**: https://nepat-dev.nexogenesis.eu/semanticRepository/.
- **Nexogenesis Data Explorer and Visualizer**: https://nepat-dev.nexogenesis.eu/visualizer/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881

5. References

[1] Bechhofer, S. (2009). OWL: Web Ontology Language. In: LIU, L., ÖZSU, M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_1073

[2] Beckett, Dave, and Brian McBride. "RDF/XML syntax specification (revised)." *W3C recommendation* 10.2.3 (2004).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003881