Objectives

- Develop a evolutionary biology-inspired semantic framework for data discovery
- Derive semantic functionality rules required for data computation discovery
- Adaptive learning rules and data discovery properties for the sustainability of the Seas case study

Description of work

Task T1.1: Semantic evolving algorithms (M1-M18) *Leader:* ***EBD-CSIC****. Contributors: 1*
Algorithms with insertions and deletions (i.e, homonymous molecular biology techniques) and other methods used to explore gene and phenotype functions will find classes, object properties and datatype properties from ontologies, and raw data from non-semantic databases. They will infer semantics on the raw data to link them to the ontological terms. We will translate the semantically-annotated databases to a *Neo4j* graph database by mapping classes to nodes, object properties to links between nodes, and datatype properties to nodes' attributes, despite current software tools to infer semantics have not been fully developed nsf. In addition, the graph database has an architecture flexible enough to get high scalability [4] to accommodate big data and to infer properties using multilayer metrics (T1.2). T1.1 provides semantic evolutionary algorithms that will allow WP2 and WP3 to implement semantics in causal knowledge discovery and discovery in federated networks

Task T1.2: This task extends T1.1 into multilayer network metrics for general principles of data discovery (M1-M18) *Leader:* ***IFISC-CSIC****. Contributors: 2*
Multilayer network metrics [12, 10] for evolutionary semantic algorithms will focus on large pools of data heterogeneity to explore how data configurations, privacy requirements, formats, dimensions, biases and spatiotemporal resolution affect data discovery [5, 1, 3, 2]. Victor, Emilio: Keep elaborating

Task T1.3: Based on the framework developed in T1.1 and T1.2, ICREA will derive automation rules for data discovery (M15-M21) *Leader:* ***ICREA****. Contributors: 3*
Automation rules [20] for evolutionary semantic algorithms and multilayer network metrics search and rules transformation for data discovery. Roger: Keep elaborating

Task T1.4: Reproduce (M15-M21) *Leader:* ***SDSC****. Contributors:*
In this task the SDSC will merge the work done in T1.1 and T1.2 into reproducible and replicable data knowledge graphs

Task T1.5: Visualize (M15-M21) *Leader:* ***SME****. Contributors:*
In this task the partner SME will apply visualization algorithms to the work done in T1.1 and T1.2 Charles and Miguel: Keep elaborating

Task T1.6: All participants apply results from evolutionary semantic algorithms and multilayer network metrics into a fully automated, reproducible and animated sustainability exploitation of the Seas case study (M15-M24) *Leader:* ***EBD-CSIC****. Contributors: 1,2,3,4,5*
Evolutionary semantic algorithms and multilayer network metrics will search and transform many source-data (i.e., Fishery data using the (global fishing watch), species interactions, environemental data and social and stakeholders groups with different interests within each of the countries, etc, together with the sustainability of the Seas database started in 1965 contains around 9 million entries, 1612 species, 20 countries and 11 sampling methods (Figure 2).

Deliverables

D1.1 Semantic evolving software for data discovery **(M18)**
D1.2 Report on definition of multilayer network metrics applied to data discovery **(M18)**
D1.3 Automated demonstrator of evolutionary semantic rules for data discovery **(M21)**
D1.4 Reproducible demonstrator of evolutionary semantic rules for data discovery **(M21)**
D1.5 Visualization demonstrator of evolutionary semantic rules for data discovery **(M21)**
D1.6 Demonstrator all parts for the sustainability exploitation of the Seas case study **(M24)**