Knowledge discovery in eco-evolutionary diversification-inspired federated networks (ROBHOOT)

1 Excellence

1.1 Radical vision of a science-enabled technology

Ecosystems' collapse around the globe is calling for technologies to discover novel ways of sustainable exploitation. Knowledge-based societies place great expectations on data-driven intelligent machines to face global sustainability challenges. In this regard, rapid and real-time discovery computation is currently a major issue revolving around data-driven intelligent machines and knowledge-based societies. Diversification of biological systems offers an unexplored avenue for inspiration of new computational approaches. However, diversifying ecosystems are not used for discovery computation yet, despite the rapid changes of traits and interactions observed in experimental and theoretical systems [1, 2]. Biological systems are characterized by feedbacks between the ecology and evolution of interacting traits, the eco-evolutionary feedbacks, to produce novel traits with new functionalities. This results in new computational properties, like new cooperation and competition strategies and information processing capabilities. Conventional Artificial Intelligence (AI) is rapidly moving towards explainable and discovery pattern inference [3] but often avoids evolutionary diversification for exploring new computing capabilities [4]. The same situation occurs for artificial neural networks that also make limited use of novel computing capabilities as a consequence of new interactions and traits [5]. **The goal of this project is to implement eco-evolutionary diversification-inspired solutions to perform discovery computation based on rapidly evolving traits and interactions**. The exploitation of evolving connections and traits will allow us to create novel discovery computation solutions for natural ecosystems facing sustainability challenges like overexploitation of the Oceans, where harvesting renewable resources are in the point of diminishing returns for many species [6].

Why should we go deeper into diversifying networks for discovery computation? With connections and traits represented in a spatially distributed network, as found in natural ecosystems, diversification is an avenue to harvest renewable resources. This allows considering not only evolutionary processes changing traits and agents but the formation of new entities to decipher new scenarios for sustainability. This also allows representing real-time solutions for ever-changing renewable resources, which is a key problem in many digital and natural ecosystems. To show the capabilities of the ROBHOOT approach, we will address full reproducibility, automation, visualization, and reporting (Figure 1). **The main impact of ROBHOOT is to provide a new technology to improve ecosystem sustainability relevant to community-rich digital and natural ecosystems.** To support this notion, we will perform eco-evolutionary diversification-inspired simulations along the whole life cycle of the project. The central goals of ROBHOOT are:

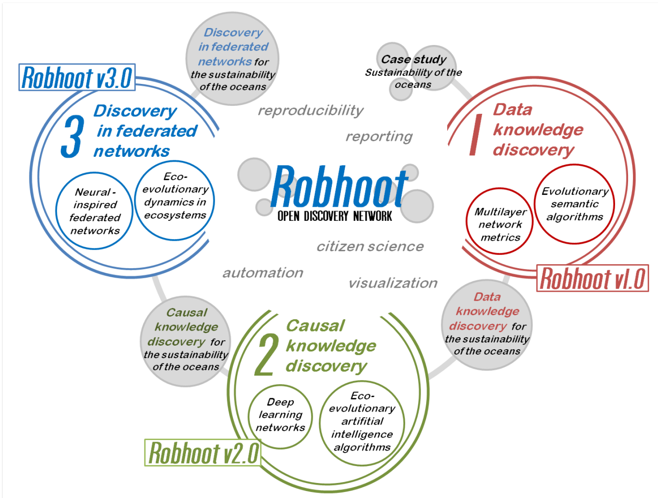

Figure 1: **Discovery in evolutionary diversification-inspired federated networks**. ROBHOOT target knowledge discovery when heterogeneous groups of species, humans and technologies share resources for a sustainable knowledge-based society: It introduces three science-enabled technologies: Evolutionary biology-inspired semantic algorithms for ROBHOOT v1.0 (ESA, data knowledge discovery, red), eco-evolutionary diversification-inspired AI models for ROBHOOT v2.0 (EEDA, causal knowledge discovery, green), and evolutionary neural diversification-inspired federated networks for ROBHOOT v3.0 (ENDI, discovery in federated networks, blue). ROBHOOT uses the sustainability of the Oceans case study in federated networks as an open-source technology with full reproducibility, automation, visualization and reporting for an open citizen science.

(G1) To extend existing theories of eco-evolutionary diversification and AI-inspired solutions to decipher the factors driving discovery computation in federated networks. This will allow us to identify novel solutions for ecosystem sustainability.

(G2) To investigate how spatiotemporal evolutionary diversification and AI-inspired networks mimic the empirical patterns of natural and socio-technological ecosystems when heterogeneous human groups and species coexist.

(G3) To develop fast, reproducible and automated eco-evolutionary biology-inspired discovery computation prototypes for real-time information processing tasks.

(G4) To obtain principles of discovery computing for prediction in federated networks when diversification in interactions and traits occurs in a large and heterogeneous set of species, technologies and human groups.

1.2 Science-to-technology breakthrough that addresses this vision

Data knowledge discovery (WP1)

Evolutionary biology-inspired semantic algorithms (ESA): Most studies of data discovery focus on advanced analytics functions to reveal insights, ignoring data source heterogeneity almost completely. Currently, only a few databases are semantically annotated from many data sources (e.g., gene ontology database, COVID-19). Ontology development is time-consuming and requires expert knowledge. It is also paired with data-driven research that checks the soundness of the ontology as it simultaneously seeks discovery. Thus, software tools for mapping and linking the terms between different ontologies accounting for many data sources are still not in place [8, 9].

Going beyond ROBHOOT will go beyond state-of-the-art to implement ESA. We will explore evolutionary-based functions to find datatype properties from ontologies and raw-data from non-semantic databases. ROBHOOT will also explore algorithms to gain an understanding of the replicability of data heterogeneity contrasting different evolutionary algorithms. ROBHOOT will explore the sustainability of the Oceans database started in 1965 and currently containing 9 million entries, 1612 species (i.e., 50 variables and traits per species), around 20 countries and 11 sampling methods (Figure 2).

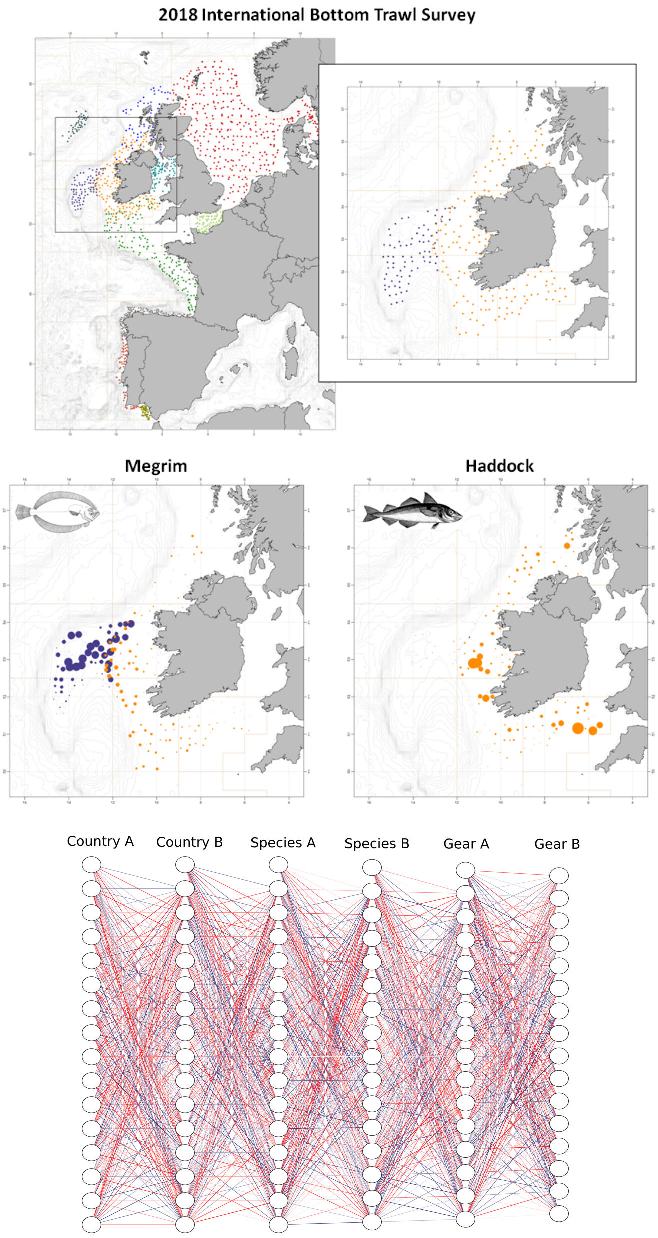

Figure 2: **Discovery for Sustainable Ecosystems**. ROBHOOT target discovery computation for new sustainability paths in complex ecosystems. The sustainability of the Oceans case study [7] will be enriched to validate the technology when many species, human groups and technologies exploit resources (**Top left**, data-color points represent a sampling from different countries. Zoomed in is the Irish Ground Fish Survey (IE-IGFS, Orange) and the Spanish Survey on the Porcupine Bank (SP-PORC, Blue). Countries produce strong bias in the distributions maps because they use different Gears according to their commercial interest (Megrim, *Lepidorhombus whiffiagonis*, consumed largely in Spain and France, **Center-left**) and Haddock, *Melanogrammus aeglefinus*, highly-priced in northern Europe, **Center-right**). This generates a strong bias for sustainability in natural resources. ROBHOOT integrates evolutionary biology-AI-inspired solutions represented as networks with many layers to discover sustainability paths with many coexisting species, human groups and technologies **Bottom**. Each country, species and gear is composed of many nodes: country contains fishery, environmental agency, stakeholders, etc. Species contains size-classes, habitat preference, species interactions, etc. Red and blue links mean competition and cooperation links connecting each pair of nodes.

Causal knowledge discovery (WP2)

Eco-evolutionary diversification-inspired AI algorithms (EEDA): Causal discovery from observable data has been extensively studied [10]. Many of these studies have used symbolic reconstruction of equations by symbolic regressions or evolutionary methods [11]. A common gap in much of the literature is one where parameters represent eco-evolutionary diversification processes, and thus, discovery can be explored broadly. The classical view on biology-inspired information processing technologies is to consider plasticity without structural changes, or without diversification among many interacting components [12]. Recent experimental evolution studies show that rapid trait changes with new information processing capabilities are far more complex because adaptation and speciation occur forming new species and phenotypes [13]. For example, eco-evolutionary dynamics strongly affect feedbacks between ecological and evolutionary processes, which in turn influences trait changes to open new properties with new information capabilities [14]. Furthermore, recent studies suggest that the interplay between trait dimensionality and adaptation is key to understand the emergence of new traits and information processing abilities to elaborate novel discovery computation strategies in ecosystems [15].

Going beyond ROBHOOT will, for the first time, employ EEDA to represent spatiotemporal causal inference in systems containing large heterogeneity and dimensionality (Figure 2). EEDA will be extended to deep process-based learning networks including traits and interactions driven by evolutionary changes to understand patterns in these systems. The search for causal knowledge discovery will be applied to the data knowledge discovery generated in WP1 for the sustainability of the Oceans, the largest ecosystem on Earth and key actor of climate change affecting biogeochemical and physical processes. Our approach will explore broad classes of evolving functions combining them to automated Bayesian machines ensuring the search, the evaluation of models, trading-off complexity, fitting to the data and quantify resource usage [16, 17].

Discovery in federated networks (WP3)

Evolutionary neural diversification-inspired federated networks (ENDI): Technologies in digital ecosystems around federated networks are rapidly increasing and mostly focus on decentralization, scalability and security fronts [18–20]. Yet, the implementation of ENDI type algorithms and their application to forecasting in global sustainability problems is still lacking. Recent studies have shown the importance of evolutionary search of mathematical and symbolic operations as building blocks to discover ML algorithms [4, 16]. ENDI will help to decipher how interactions among heterogeneous groups evolve and learn to solve complex sustainability problems. Evolutionary dynamics explore open-ended language of models with varying trait evolution functions to discover biologically inspired solutions in multidimensional systems [4]. ENDI accounts for heterogeneous agents to discover novel biology-inspired solutions for the sustainability of the Oceans federated network.

Going beyond: Our understanding of the outcomes from diversified information processing systems formed by highly heterogeneous groups, a kind of large-scale meta-learning in the federated setting [21], is currently quite limited. Therefore, new science-enabled approaches accounting for diversifying information processing in heterogeneous and highly dimensional systems are required. This allows the development of science-enabled technologies where heterogeneous agents with different interests find (non-optimal) solutions for the sustainable exploitation of ecosystems. Federated objects can be seen as "neural networks" containing many types of heterogeneous nodes with varying degrees of learning, connectivity and firing probabilities [22, 23]. ROBHOOT v.3.0 connects knowledge discovery to ENDI to study the properties of cooperative forecasting in the face of global sustainability challenges.

1.3 Interdisciplinarity and non-incrementality of the research proposed

The success of ROBHOOT relies on a multidisciplinary team: evolutionary biology, ecology, computational neuroscience, data science, complex systems and experts in communication and field studies in biodiversity. Data knowledge discovery will be developed by evolutionary biology, computer science and complex system members of the consortium (EBD-CSIC, IFISC-CSIC, and SDSC). Data discovery will be transferred to the causal domain by the other part of the consortium with expertise in ecology, evolutionary biology, data science and causal inference (EAWAG and TARTUR). The whole process will be enriched with full automation, reproducibility and visualization by ICREA, SDSC, and our company-partner SCITE, respectively. Conversely, scientists working on neurobiology and eco-evolutionary dynamics in ecosystems will feed information back on fundamental discovery computational challenges in federated networks (i.e., role of heterogeneity, expanding EEDA and ENDI and adding cooperation and dimensionality). They will explore novel paths to improve existing theories using EEDA and ENDI algorithms. This cross-fertilizing back-and-forth interaction will allow the project to keep high modularity within the WPs while keeping cross-interactions among the groups to run efficiently the project. Bringing together ESA, EEDA and ENDI algorithms require a long stride and this has not been attempted so far. This way, we expect to realize a truly novel, sustainability-driven knowledge-based society technology for which there are no predecessors. Thus, ROBHOOT will not be incremental, but a jump to a new direction for eco-evolutionary diversification-inspired discovery computation.

Table 1.4a: Critical risks for the research approach

Description of risk	**Objective**	**WP**	**Proposed risk- mitigation measures**
ESA insufficiently developed: Medium	2	WP1	Use traditional non-semantic genetic algorithms to infer data connections.
Low number of training data available: Medium	2,3	WP2	Alternative methods focusing on matrix decomposition.
Automated evolutionary-inspired expressions for causal knowledge discovery insufficiently developed: Medium	2,3	WP2	Symbolic regression methods for causal discovery accounting for evolutionary rules.
Extended EEDA in species-rich ecosystems insufficiently developed: Medium	1-4	WP3	Mean-field approximations using classical ODE systems and novel universal differential equations for scientific machine learning.
ENDI in federated networks insufficiently developed: High	1-4	WP3	Alternative neural network models based on deterministic spiking neurons.
Cooperative forecasting mixing EEDA and ENDI in federated networks insufficiently developed: High	1-4	WP3	Combine EEDA and ENDI on a smaller spatiotemporal scale.

1.4 High risk, plausibility and flexibility of the research approach

ROBHOOT represents a novel approach for discovery computation in ecosystems. The transfer of eco-evolutionary diversification-inspired principles onto fully reproducible and automated software, progressing towards a process-based discovery technology, will be a major qualitative step, defining ROBHOOT as a high-risk project, fitting into FET-Open. To achieve the ambitious goals, we will combine expertise from all the involved areas, gradually mitigating risk, following a strict line and increasing step-by-step in the complexity of the problems addressed. We will start with evolutionary biology-inspired semantic algorithms for data discovery applying them to the sustainability of the Oceans case study. This is followed by more complex eco-evolutionary diversification-inspired AI modeling to infer causality in our case study. Then, we will advance to more complex situations, where evolutionary neural diversification-inspired modeling will expand the search along many distinct forecasting schemes to discover sustainability paths. To keep the project technically feasible, and to be able to connect properties from data and causal discovery computation to discovery in federated networks, we will limit methods to three main approaches. All of the above will be done by combining data-driven simulations, theoretical work and numerical simulations with our sustainability of the Oceans case study crossing them all. The knowledge gained will allow us to present ROBHOOT as a compact science-enabled open-source technology. We will use fast computing languages to implement Agent-Based Models (ABM) along with all the theoretical development of the proposal. We will contrast the ABM with differential/difference stochastic equation methods when a large number of agents, traits, and interactions diversify in time and space. This feature represents a very desirable fallback in case of speed and convergence problems for multidimensional and nonlinear systems (Table 1.4a, Critical risks for the research approach). Our implementation activities are all complemented by numerical investigations contrasted for speed and robustness with the sustainability of the Oceans case study (Figure 2). The success of ROBHOOT would represent a breakthrough in the current discovery computation with direct application to the sustainability of ecosystems and beyond. The combination of rapid, heterogeneous database and cooperation for discovery computation based on open-source algorithms will lead to fast implementations of the demonstrators with high flexibility that will permit a rapid transit to the public.

2 Impact

2.1 Expected impact

(I) **Contribution to the foundation of a new future technology**: ROBHOOT uses the discovery of novel evolutionary diversification-inspired algorithms (EEDA and ENDI) to substantially improve solutions for sustainability in ecosystems. **Discovery of novel EEDA and ENDI in the context of diversifying traits, interactions, technologies and human groups for biodiversity maintenance have been hardly been investigated in this context so far**. Predictors related to biodiversity, technological and social analysis will be tested and further developed to enable robust predictions. Altogether, this project will lay the foundation for future sustainability studies.

(II) **Future social/economic impact or market creation**: Our approach uses a novel technology to integrate many data types and discovery paths to make ecosystems sustainable. This will allow us to use the technology in public and private industry to generate robust scenarios when facing complex problems including global sustainability challenges (e.g., global health, food and feed production, ecosystems degradation).

(III) **Impact on transparency**: Decision making and governance at local, regional and global scales require access to reproducible information containing viable sustainability scenarios. ROBHOOT consortium brings together different partners in the fields of computer science, neurobiology, complex system, biology, social sciences, evolutionary ecology and one SME, all focusing on reproducibility, automation, visualization and reporting scientific data to different audiences.

(IV) **Ecosystem health impact**: ROBHOOT focuses on discovery solutions for exploited ecosystems. It uses a case study for the Oceans and provides solutions for ecosystem sustainability, thereby connecting ecosystem sustainability and ecosystem health. This feature aligns with the EU Reflection paper towards a Sustainable Europe by 2030 and the UN's Sustainable Development Goals. ROBHOOT can be seen as a horizontal enabler for a scientific-based transition to sustainability-based technology on large amounts of heterogeneous data, artificial intelligence and EEDA solutions.

(V) **Building leading research and innovation capacity across Europe**: This consortium brings together excellent partners from the fields of computer science, machine learning, deep learning networks, neurobiology, complex systems, experimental biology, biology and evolutionary ecology, physics, theory and applications of complex systems in social networks, delivering a highly innovative science-enabled technology focusing on sustainability solutions. All consortium partners exhibit a long-standing experience in interdisciplinary research across the boundaries of the individual disciplines. A web-based sustainability discovery portal will be produced (WP4), which will allow researchers, NGOs, managers and the public to train students in the discovery process to manage overexploited ecosystems. This will also allow us to scale up the number of people participating in the sustainability process thus mobilizing forward-thinking researchers and excellent young researchers to work together and explore what may become a novel paradigm in sustainability research.

2.2 Measures to maximize impact

Dissemination and exploitation

A plan for dissemination and exploitation (PDE) will be developed and managed under WP4. It will address the project strategy and concrete actions related to: i) Dissemination: Open Access format; ii) Data Management: how data will be handled; iii) Protection: IPR strategy; iv) Exploitation, namely "business models", and v) Communication, particularly the different actions to communicate the project's results and demonstrators to key groups of end-users.

(I) **Open Access**: Project reports and ISI journal publications will be under the Open Access format. Following the Open Science principles, software and scientific publications will be deposited in the online institutional repositories and on the EC Participant Portal. ZENODO (http://zenodo.org), recommended by the European Research Council and the EC, and supported by EUs OpenAire platform (https://www.openaire.eu/) will be also used for dissemination and communication purposes (publications, presentations, datasets, images, videos/audio and interactive materials such as lectures).

(II) **Open access to research data**: recommended data repositories (e.g. PANGAEA, NASA Goddard Earth Sciences Data and Information Services Center) will be used to share the generated data and software. Open-source codes and analysis of standardized inputs/outputs and software will be made public through an online platform with the aim of converting it in *the Reference Point* for any future research in knowledge discovery.

(III) **Data management**: Good research data management practice will ensure all the data is registered, stored, made accessible for use, managed over time and/or disposed of, according to legal, ethical, funder requirements and good practice. This management will provide benefits such as reducing the risk of data loss, improving data workflows and data availability and discovery, visibility of research outputs, attracting new collaborators and research partners, strengthening the research environment and infrastructure. A data management plan (DMP) will be created by the Project Coordinator in close cooperation with the partners and approved by the Steering Board at the start of the Project. The DMP will follow the FAIR principles. The document will describe how to collect, organize, manage, store, secure, back-up, preserve, and where applicable, share data.

(IV) **Innovation and IPR**: The Consortium will benefit from the innovation and technology transfer environment in place at EAWAG and CSIC, which will assist in the patent application process. Support is also available to assist the realization of innovative ideas into efficient business concepts. The necessary precautions will also be taken to protect the IPR of individual institutions. A Consortium Agreement will be signed before the beginning of the project to take into account the different interests of the partners, in particular how to treat pre-existing know-how, the ownership of the results and the intellectual property rights to prevent conflicts during the project. The Steering Board will ensure that all innovations and generated data are exploited to the benefit of the involved partners.

(V) **Exploitation, including business models**: The project's results will be showcased in trade shows (e.g. WebSummit), by communicating through specialized trade press media, and also to a targeted audience (policy makers, funding agencies, industry and SMEs). A detailed business plan will be prepared during the project work in collaboration with the SME and academic partners involved, with the ultimate goal of creating a Start-Up at the end of ROBHOOT. The value proposition of the project is to develop computation discovery solutions for rapidly diversifying traits and complex interactions that improve the sustainability of exploited natural ecosystems.

YEAR					2021				2022				2023			
MONTH					Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
MILESTONE												MS1		MS2		MS3
WP	WP Name	PROGRESS	START	END												
WP1	**Data knowledge discovery (DKD)**															
T1.1	Evolutionary semantic algorithms	17%	1/1/21	31/6/22						D1.1						
T1.2	Multilayer network metrics	34%	1/1/21	31/6/22						D1.2						
T1.3	Automation DKD	51%	1/4/22	1/8/22							D1.3					
T1.4	Reproducibility DKD	68%	1/4/22	30/9/22							D1.4					
T1.5	Visualization DKD	85%	1/4/22	30/9/22							D1.5					
T1.6	Data knowledge discovery SUSO	100%	1/4/22	31/12/22								D1.6				
WP2	**Causal knowledge discovery (CKD)**															
T2.1	Eco-Evolutionary diversification AI Algorithms	17%	1/5/22	31/12/22								D2.1				
T2.2	Eco-Evolutionary Deep Learning	34%	1/5/22	31/12/22								D2.2				
T2.3	Automation CKD	51%	1/10/22	31/3/23									D2.3			
T2.4	Reproducibility CKD	68%	1/10/22	31/3/23									D2.4			
T2.5	Visualization CKD	85%	1/10/22	31/3/23									D2.5			
T2.6	Causal knowledge discovery SUSO	100%	1/10/22	30/6/23										D2.6		
WP3	**Discovery in federated networks (DFN)**															
T3.1	Eco-evolutionary diversification-inspired	17%	1/1/22	31/12/23										D3.1		
T3.2	Evolutionary neural diversification-inspired	34%	1/1/22	31/12/23												D3.2
T3.3	Automation DFN	51%	1/1/23	31/12/23												D3.3
T3.4	Reproducibility DFN	68%	1/1/23	31/12/23												D3.4
T3.5	Visualization DFN	85%	1/1/23	31/12/23												D3.5
T3.6	Discovery in federated networks SUSO	100%	1/1/23	31/12/23												D3.6
WP4	**Dissemination, Knowledge Transfer and Outreach**															
T4.1	Dissemination and exploitation Plan	10%	1/1/21	31/12/23	D4.1							D4.4				D4.5
T4.2	Branding and communication guidelines	15%	1/1/21	31/3/21	D4.2											
T4.3	Website and social media	20%	1/1/21	31/12/23		D4.3										
T4.4	Case study outreach	45%	1/10/21	31/12/23								D4.4				D4.5
T4.5	Knowledge Transfer	70%	1/7/22	31/12/23								D4.4				D4.5
T4.6	Publications and Conferences	90%	1/7/22	31/12/23								D4.4				D4.5
T4.7	Exploitation	100%	1/4/23	31/12/23												D4.6
WP5	**Management**															
T5.1	Project initiation	50%	1/1/21	30/6/21												
T5.2	Other management task (R = Reporting)	100%	1/1/21	31/12/23				R				R				R

Gantt chart:(MS=Milestone,D=Deliverable,R=Project Reporting,T=Task)

Communication activities

ROBHOOT has very general communication targets, from general public to scientists, decision-makers and to the business community. ROBHOOT's general dissemination measures focus on project results and stakeholder engagement through the following activities.

(I) **Write position papers and relevant conferences**: White papers will be produced showcasing the results receive feedback from stakeholders about implementations. Results will also be shown in relevant conferences where scientific, industry, NG and NGO entities meet.
(II) **Website**: A website to reach the general public through social media (Instagram, Twitter, Facebook, LinkedIn, Video-channels), press releases/TV/radio and a public git ROBHOOOT repository (ROBHOOT git repository), will be used for communicating results to all target audiences.
(III) **Hackathons**: we will organize joined activities with on-ongoing EU/International projects to attract multipliers and developers from the community who engage in data analytics. At the end of the project we will organize a workshop specifically on *Next-generation evolutionary-biology AI-inspired solutions for global sustainability challenges.*
(IV) **Testnet**: ROBHOOT will launch a testnet to disseminate the results of discovery in federated networks. The launch will have invited NGOs and GO across disciplines and social, economical and technological sectors. The ROBHOOT Open Discovery Network will be launched as a Biodiversity and sustainability network. It will offer solutions for the sustainability of the Oceans and to integrate additional public databases and data collections into the open discovery network to facilitate NGOs, GOs and other organizations transparency and governance in ecosystem management.

3 Implementation

3.1 Research methodology and work plan, work packages and deliverables

The project consists of five work-packages (WP1-WP3: R&D, WP4: Dissemination and WP5: Management). WP1 deals with evolutionary semantic algorithms for data knowledge discovery, WP2 addresses eco-evolutionary diversification-inspired AI models to infer causal knowledge discovery with an implementation for the sustainability of the Oceans' sustainability case study, WP3 addresses evolutionary neural biology-inspired for knowledge discovery to provide cooperative forecasting in federated networks. WP3 also provides an empirical case implementation of cooperative forecasting for the sustainability of the Oceans. **Demonstrators**: The project will create three demonstrators of increasing complexity all containing full reproducibility, automation and visualization capabilities:
1. ROBHOOT v1.0 Software demonstrator with evolutionary semantic algorithms to decipher ontologies for the sustainability of the Oceans data knowledge discovery case study (MS1);
2. ROBHOOT v2.0 Software demonstrator with evolutionary diversification-inspired AI modeling for spatiotemporal causal pattern knowledge discovery (MS2);
3. ROBHOOT v3.0 Software demonstrator with evolutionary neural diversification-inspired modeling for discovery in federated networks (MS3).

The inference of causal mechanisms and the discovery of spatiotemporal patterns in federated networks is a generic problem found in e.g. many agents sharing resources, sustainability, eco-evolutionary networks, biodiversity maintenance, or social networks. Thus, the discovery computation of spatiotemporal patterns represents a ubiquitous computational problem in digital and natural ecosystems, where many evolving and heterogeneous agents and interactions share information to reach sustainability goals. In the demonstrators of ROBHOOT, we will consider different scenarios for each of the software implementations such that agents contain many evolving traits and interactions (MS1, MS2 and MS3). This allows, for example, finding patterns of trait and interaction changes to improve sustainability as a function of the observed empirical patterns in our Oceans' sustainability case study. In the course of the project, more complex context-dependent trait changes of agents and interactions together with different learning functions will be considered to explore how they affect sustainability properties in federated networks.

Table 3.1a: **List of work packages**

WP	**Work package title**	**Lead No.**	**Lead Name**	**PMs**	**Start Month**	**End Month**
1	Data knowledge discovery	2	CSIC	48	1	18
2	Causal knowledge discovery	4	TARTU	24	7	24
3	Discovery in federated networks	6	TU GRAZ	24	13	36
4	Dissemination	7	IEO	24	1	36
5	Management	1	EAWAG	31.2	1	36
			Total PMs	151.2		

Table 3.1b: **Work package description**

Work package number	1	Lead beneficiary	CSIC		
Work package title	**Data knowledge discovery**				
Participant number	2	2	8	5	3
Short name of participant	EBD-CSIC	IFISC-CSIC	URV	EPFL	SCITE
Person month per participant	24	24	6	6	6
Start month	1				
End month	24				

Objectives

- To develop an evolutionary biology-inspired semantic framework for data discovery
- To derive semantic functionality rules required for data computation discovery
- To apply data discovery properties for the Oceans' sustainability case study

Description of work

Task T1.1: Evolutionary semantic algorithms (ESA) (M1-M18) *Leader:* ***EBD-CSIC****. Contributors: 2*

ESA will find classes and datatype properties from ontologies, and raw data from non-semantic databases. ESA 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. The graph database has an architecture flexible enough to get high scalability to accommodate many source data and to infer its properties using multilayer metrics (T1.2). T1.1 provides ESA to allow WP2 and WP3 to implement the models for 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 for ESA will focus on data heterogeneity to explore how data configurations, privacy requirements, formats, dimensions, biases and spatiotemporal resolution affect data discovery properties [24–26].

Task T1.3: Based on the framework developed in T1.1 and T1.2, URV will derive automation procedures for data knowledge discovery (M15-M21) *Leader:* ***URV****. Contributors: 8*

Automation rules identify the ESA rules for data discovery [16]. URV will complement T1.1 and T1.2 to obtain posterior probabilities of evolutionary expressions that represent the empirical patterns of the data knowledge graph generated in T1.1 and T1.2.

Task T1.4: Reproducible data knowledge graphs (M15-M21) *Leader:* ***EPFL****. Contributors: 5*

In this task the EPFL will integrate the work done in T1.1 and T1.2 into reproducible and replicable data knowledge graphs. T1.4 samples the data sources to obtain the robustness of data heterogeneity. Robustness will be analyzed working closely to the IFISC-CSIC partner in T1.2.

Task T1.5: Visualize (M15-M21) *Leader:* ***SCITE****. Contributors: 3*

In this task the partner SCITE will apply visualization algorithms to the patterns obtained in T1.1 and T1.2. Data knowledge graphs will be represented in static (figures) and dynamic (animations) visualizations using cutting-edge graphic libraries like D3.js, LightGraphs.jl. All animations will be used by SCITE to strengthen the dissemination, communication and exploitation activities.

Task T1.6: All participants apply results from ESA and multilayer network metrics into a fully automated, reproducible and animated Oceans' sustainability case study (M15-M24) *Leader:* ***EBD-CSIC****. Contributors: 2,3,5,7,8*

ESA and multilayer network metrics will generate the sustainability of the Oceans data knowledge graph integrating many data sources. Fishery data (i.e., global fishing watch), species interactions data, environmental data and social and stakeholders groups data with different interests within each country, etc, will be merged into the sustainability of the Oceans database started in 1965 containing 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 Oceans' sustainability case study **(M24)**

Work package number	2	Lead beneficiary	TARTU		
Work package title	**Causal knowledge discovery**				
Participant number	1	4	5	8	3
Short name of participant	EAWAG	TARTU	EPFL	URV	SCITE
Person month per participant	24	24	6	6	6
Start month	7				
End month	30				

Objectives

- To develop an evolutionary-diversification AI-inspired framework for causal discovery
- To derive functionality rules required for causality-based computation discovery
- To apply diversification rules to mimic the empirical patterns for the Oceans' sustainability case study

Description of work

Task T2.1: Develop EEDA algorithms (M7-M24) *Leader:* ***EAWAG****. Contributors: 1*

T2.1 provides process-based algorithms with diversifying traits and interactions for species, human groups and technologies to allow WP2 to implement this feature in causal knowledge discovery. Causal modeling is particularly relevant in Earth, Ecosystem and Sustainability science where rapid progress of AI in explainable technology [4, 16, 27, 28] will increase our ability to make stronger inferences about future sustainability challenges and solutions [29]. EEDA solutions will be required to explore a broad range of sustainability scenarios, particularly relevant to find diversification rates in species, technologies and human strategies that best represent the empirical observations for the sustainability of the Oceans data knowledge discovery generated in WP1.

Task T2.2: This task extends T2.1 into EEDA deep learning networks metrics for general principles of causal discovery (M7-M24) *Leader:* ***TARTU****. Contributors: 4*

Using as input the knowledge graphs extracted from WP1, the goal of this task is to use deep learning technology to infer sustainability paths. In particular, this task will train graph neural networks (specifically designed to handle network data) to predict the effects of the complex interactions between species, human groups and exploitation technology. Simulations and the Ocean case will be first studied and calibrated. Once the deep learning model is trained, causal perturbations on the inputs and biases towards sparse models will be implemented to give an explainable account for the key causal interactions. Finally, the model will be optimised with respect to human interventions to aim the ecological system under study towards plausible sustainability paths.

Task T2.3: Based on the framework developed in T2.1 and T2.2, URV will derive automation rules for causal discovery (M21-M27) *Leader:* ***URV****. Contributors: 8*

URV will complement T2.1 and T2.2 to obtain scenarios of EEDA that represent the causal knowledge discovery graphs that best represent the empirical patterns. URV will work together with T2.1 and T2.2 to address the fit-complexity trade-off and to obtain the posterior probabilities for the rules and expressions generated with the EEDA.

Task T2.4: Causal reproducible knowledge graphs (M15-M21) *Leader:* ***EPFL****. Contributors: 5*

In this task the EPFL will integrate the work done in T2.1 and T2.2 into reproducible and replicable causal knowledge graphs. T2.4 samples the causal graphs to obtain the robustness of the inference. Robustness will be analyzed working closely to the URV partner in T2.3.

Task T2.5: In this task SCITE will apply visualization algorithms to T2.1 and T2.2 (M21-M27) *Leader:* ***SCITE****. Contributors: 3*

Spatial and networks patterns will be represented in static (figures) and dynamic (animations) visualizations using cutting-edge graphic libraries like $D3.js$, $Vega.jl$, $NetworkD3.js$, Leaflets, and $ggplot2$. Animations will represent the EEDA and deep learning networks patterns. Storytelling techniques will be applied in order to effectively communicate those findings.

Task T2.6: All participants apply results from EEDA and deep learning networks into a fully automated, reproducible and animated Oceans' sustainability case study (M21-M30) *Leader:* ***EAWAG****. Contributors: 1,3,4,5,7,8*

EEDA and deep learning networks will generate the sustainability of the Oceans causal knowledge graph (Figure 2).

Deliverables

D2.1 Report on definition of EEDA rules for causal discovery **(M18)**
D2.2 Report on definition of EEDA deep learning networks applied to causal computation discovery **(M18)**

D2.3	Automated demonstrator of EEDA rules for causal discovery **(M21)**
D2.4	Reproducible demonstrator of EEDA AI rules for causal discovery **(M21)**
D2.5	Visualization demonstrator of EEDA for causal discovery **(M21)**
D2.6	Demonstrator all parts for the sustainability of the Oceans' **(M24)**

Work package number	**3**	**Lead beneficiary**		**TU GRAZ**	
Work package title	**Discovery in federated networks**				
Participant number	9	6	5	8	3
Short name of participant	SRC	TU GRAZ	EPFL	URV	SCITE
Person month per participant	24	24	12	12	11
Start month	13				
End month	36				

Objectives

- To develop an evolutionary-diversification inspired framework for discovery in federated networks.
- To derive diversification rules required for computation discovery in federated networks.
- To apply rules to discover novel paths for Oceans' sustainability.

Description of work

Task T3.1: Extend EEDA for discovery in federated networks (M13-M36) *Leader: **SRC**. Contributors: 9*

This task extends EEDA in T2.1 for general principles of discovery in federated networks. T3.1 provides process-based models with diversifying traits and interactions along node heterogeneity gradients. Federated networks are represented as interacting species, human groups and technologies containing heterogeneity gradients and multidimensional properties. Mean field deterministic models will be developed and contrasted to the stochastic counterparts (T3.2). These features will allow WP3 to implement diversification rules when heterogeneous groups interact and share resources in ecosystems. Extensions of EEDA solutions are required to discover novel diversification rates in species, technologies and human strategies that improve sustainability paths in comparison to the observed empirical patterns in our case study.

Task T3.2: Develop ENDI in federated networks (M13-M36) *Leader: **TU GRAZ**. Contributors: 6*

T3.2 provides computation algorithms for neural diversification-inspired processes to allow WP3 to implement this feature for discovery in federated networks. Neurons will be represented as algorithms along heterogeneity and/or complexity gradients. Links represent cooperation, learning or competition. A federated neural network is composed by types of neurons: Species, human groups and technology all containing heterogeneity along many dimensional traits. The goal is to discover new rules representing high sustainability values defined broadly as a high degree of coexistence among many species, diversified technologies and human groups. The dynamics of interacting neurons will be essentially stochastic. The following are the rules governing the dynamics: we start with a population of algorithms extended in a landscape and fitness functions determine birth and deaths of algorithms. We will consider stochastic spiking neurons within nodes to compute how information processing evolves in the network. We will consider a variety of scenarios from strong selection to neutral federated networks to explore sustainability paths. Which scenario provides higher sustainability-efficiency information processing in federated networks? T3.2 and T3.1 will interact to strengthen deterministic and stochastic solutions for EEDA and ENDI implementations in federated networks.

Task T3.3: Based on the framework developed in T3.1 and T3.2, URV will derive automation rules for knowledge discovery (M25-M36) *Leader: **URV**. Contributors: 8*

URV will complement T3.1 and T3.2 to obtain the scenarios of ENDI that represent the knowledge discovery paths for sustainability trajectories not observed in the empirical patterns of the sustainability of the Oceans case study. URV will work together with T3.1 and T3.2 to automate and discover expressions and rules generated by ENDI.

Task T3.4: Reproducible discovery knowledge graphs (M25-M36) *Leader: **EPFL**. Contributors: 5*

In this task the EPFL will integrate the work done in T3.1 and T3.2 into reproducible and replicable discovery knowledge graphs. T3.4 samples the discovery paths to obtain the robustness of these roads to sustainability. Robustness will be analyzed working closely to the URV partner in T3.3.

Task T3.5: In this task SCITE will apply visualization algorithms to T3.1 and T3.2 (M25-M36) *Leader:* ***SCITE****. Contributors: 3*
Spatial and networks patterns will be represented in static (figures) and dynamic (animations) visualizations using cutting-edge graphic libraries like *D3.js*, *Vega.jl*, *NetworkD3.js*, Leaflets, and *ggplot2*. Animations will represent the ENDI and the extension of EEDA to federated networks. Storytelling techniques will be applied in order to effectively communicate those findings.

Task T3.6: All participants apply results from extended EEDA and ENDI into a fully automated, reproducible and animated Oceans' sustainability case study (M21-M36)
Leader: ***TU GRAZ****. Contributors: 3,5,6,7,8,9*
Extended EEDA and ENDI will generate the sustainability of the Oceans discovery knowledge.

Deliverables

D3.1 Demonstrator on EEDA for discovery in federated networks **(M36)**
D3.2 Demonstrator on ENDI for discovery in federated networks **(M36)**
D3.3 Automated demonstrator of ENDI for discovery in federated networks **(M36)**
D3.4 Reproducible demonstrator of ENDI for discovery in federated networks **(M36)**
D3.5 Visualization demonstrator of ENDI for discovery in federated networks **(M36)**
D3.6 Demonstrator ENDI from all parts for the Oceans' sustainability federated network case study **(M36)**

Work package number	**4**		**Lead beneficiary**	**IEO**
Work package title	**Dissemination, Knowledge Transfer and Outreach**			
Participant number	7	3		
Short name of participant	IEO	SCITE		
Person month per participant	24	15		
Start month	1			
End month	36			

Objectives

- Ensure effective external communication, dissemination and optimal knowledge transfer of ROBHOOT results

Description of work

Task T4.1: ***Dissemination and Exploitation Plan (DEP)*** **(M1-M36)** *Leader:* ***IEO****. Contributors: 3,7*
A DEP will be put in place immediately upon project commencement. Yearly, the plan will be evaluated for effectiveness and adjusted if needed.

Task T4.2: Branding and communication guidelines (M1-M3) *Leader:* ***SCITE****. Contributors: 3,7*
Create a visual project identity with a logo and templates for presentations, posters, and deliverable documents. Brochures introducing the project, aims and expected results will also be produced.

Task T4.3: Website and social media (M1-M36) *Leader:* ***SCITE****. Contributors: 3, 7*
A dedicated website and a public git ROBHOOOT repository, (https://github.com/RobhooX/Robhoot), will be used for communicating results and sharing updated versions with all target audiences. Social media accounts will be created and posts will be used to raise attention to project activities and achievements, adapted to the audience.

Task T4.4: Sustainability of the Oceans (M11-M36) *Leader:* ***IEO****. Contributors: 3, 7*
Ocean research has been proven to capture the imagination of the public. We will work with researchers to communicate outreach activities to ensure engagement with European Citizens. Effective dissemination to the general public will also be achieved through press releases announcing project start and key milestones to provide a public media dimension.

Task T4.5: Knowledge Transfer (M18-M36) *Leader:* ***SCITE****. Contributors: 3, 7*
Hackathon events, coinciding with the three milestones (Robhoot v1.0, Robhoot v2.0 and Robhoot v3.0), are planned to attract multipliers and developers. We will organize a workshop specifically on "Next-generation evolutionary-biology AI-inspired solutions for global sustainability challenges" for disseminating our results to a broad set of groups and experts in fields related to global sustainability for assessing future exploitation potential, inviting partners from academia as well as industry. Trade media articles targeting companies and end-users will also be published in artificial intelligence magazines, as well as in magazines focused on ocean sciences and sustainability. Once the final prototype is developed, it will be presented in trade shows, such as Web Summit, World Ocean Summit, among others.

Task T4.6: Publications and Conferences (M18-M36) *Leader: **IEO**. Contributors: All*
Around 6 to 8 high-impact scientific publications are expected, as well as organization of special sessions in international scientific and technological meetings. To raise awareness about this topic for young generations, an article will also be published in ”Frontiers for Young Minds”.

Task T4.7: Exploitation (M28-M36) *Leader: **SCITE**. Contributors: 3, 7*
We will identify all target groups and key players of the value chain. An Exploitation Plan will be developed and maintained to ensure prompt utilization of project results by relevant stakeholders. This will include (i) SWOT analysis; (ii) definition of business models and plans (based on the lean business canvas); and (iii) efficient IP management to facilitate rapid exploitation. Key stakeholders from all levels of society will be identified as potential beneficiaries of the project outcomes and engaged at an early project stage.

Deliverables

D4.1 Dissemination and Exploitation Plan **(M3)**
D4.2 Brand book, templates and brochures **(M3)**
D4.3 Launch website **(M5)**
D4.4 Progress Report on Dissemination, Knowledge Transfer and Outreach **(M24)**
D4.5 Final Report on Dissemination, Knowledge Transfer and Outreach **(M36)**
D4.6 Business plan **(M36)**

Work package number 5	**Lead beneficiary** EAWAG	
Work package title	**Management**	
Participant number	1	3
Short name of participant	EAWAG	SCITE
Person month per participant	7.2	3
Start month	1	
End month	36	

Objectives

- Management deliverables of the project during the contractual period.
- Administrative and financial management of the project.
- Ensure the delivery of the project on time and on budget.
- Coordinate the technological and scientific orientation of the project with the scientific and communication advisory teams.
- Secure the quality of the work and of the delivered documents and software

Description of work

Task T5.1: Project initiation (M1-M6) *Leader: **EAWAG**. Contributors: 1,3*
Initiation of the project involving a kick-off meeting, definition of the details of the early work (details that go beyond the aspects written here). Planning the early deliverables.

Task T5.2: Coordination and organization of deliverables and communication (M1-M36) *Leader: **EAWAG**. Contributors: 1,3*
EAWAG is responsible for the overall coordination, administration and organization of the management structure and will deploy the organizational structure and procedures to ensure the smooth and efficient operation of the project from both the strategic and tactical perspectives.

Task T5.3: Project management (M1-M36) *Leader: **EAWAG**. Contributors: 1,3*
To ensure working efficiency, there will be day-to-day operational project management of the project's work overseen by the Project Manager. The Project Manager is overall responsible for the success and smooth running of the project.

Task T5.4: Work package management (M1-M36) *Leader: **EAWAG**. Contributors: 1,3*
Coordination and organization of the work package management. The work itself is part of the individual work packages. The main players in this are: WP Leaders and WP teams.

Task T5.5: Risk management (M1-M36) *Leader: **EAWAG**. Contributors: 1,3*
Management of all risks and issues which are identified in the project. All project participants and external advisors will be responsible for raising any risk or issue they perceive. As the consortium is small, risks can and will be immediately addressed via discussions of the partners.

Task T5.6: Project reporting (M1-M36) *Leader: **EAWAG**. Contributors: 1,3*
EAWAG AND IFISC-CSIC will be responsible for producing the formal project reporting documents including the yearly progress reports and the final project report.

Table 3.1c: Deliverable list

Table 3.1b: Deliverable list

Deliverable number	Deliverable name	WP no.	Lead participant name	Nature	Dissemination Level	Delivery date (proj. month)
D4.1	Dissemination and Exploitation Plan	WP4	IEO	R	PU	3
D4.2	Brand book, templates and brochures	WP4	SCITE	O	PU	3
D4.3	Launch website	WP4	SCITE	O	PU	5
D1.1	Semantic evolving software for data discovery	WP1	EBD-CSIC	R	PU	18
D1.2	Report on definition of multilayer network metrics applied to data discovery	WP1	IFICS-CSIC	R	PU	18
D2.1	Report on definition of EEDA rules for causal discovery	WP2	EAWAG	R	PU	18
D2.2	Report on definition of EEDA deep learning networks applied to causal computation discovery	WP2	TARTU	R	PU	18
D1.3	Automated demonstrator of evolutionary semantic rules for data discovery	WP1	URV	D	PU	21
D1.4	Reproducible demonstrator of evolutionary semantic rules for data discovery	WP1	EPFL	R	PU	21
D1.5	Visualization demonstrator of evolutionary semantic rules for data discovery	WP1	SME	R	PU	21
D2.3	Automated demonstrator of EEDA rules for causal discovery	WP2	URV	D	PU	21
D2.4	Reproducible demonstrator of EEDA AI rules for causal discovery	WP2	EPFL	R	PU	21
D2.5	Visualization demonstrator of EEDA for causal discovery	WP2	SCITE	R	PU	21
D1.6	Demonstrator all parts for the Oceans' sustainability case study	WP1	EBD-CSIC	R	PU	24
D2.6	Demonstrator all parts for the sustainability of the Oceans'	WP2	EAWAG	R	PU	24
D4.4	Progress Report on Dissemination, Knowledge Transfer and Outreach	WP4	IEO	R	PU	24
D3.1	Demonstrator on EEDA for discovery in federated networks	WP3	SRC	D	PU	36
D3.2	Demonstrator on ENDI for discovery in federated networks	WP3	TU GRAZ	D	PU	36
D3.3	Automated demonstrator of ENDI for discovery in federated networks	WP3	URV	D	PU	36
D3.4	Reproducible demonstrator of ENDI for discovery in federated networks	WP3	EPFL	D	PU	36
D3.5	Visualization demonstrator of ENDI for discovery in federated networks	WP3	SCITE	D	PU	36
D3.6	Demonstrator ENDI from all parts for the Oceans' sustainability federated network case study	WP3	TU GRAZ	D	PU	36
D4.5	Final Report on Dissemination, Knowledge Transfer and Outreach	WP4	IEO	R	PU	36
D4.6	Business plan	WP4	SCITE	R	PU	36

3.2 Management structure, milestones and procedures

Management procedures and structure:

ROBHOOT is organized by the Project Manager, with a Steering Board (SB), an external Scientific Advisory Committee (SAC), and a Dissemination and Exploitation Board (DEB). The SB, which will consist of one representative from each partner and the Project Manager, will meet at least once a year. The SB will have the overall responsibility for the technical, financial, administrative, legal, and risk analysis. The SB will also have responsibility together with the DEB for all the dissemination and outreach of the project. The SAC, headed by the coordinator, will consist of senior experts in the respective fields. Prof. Elisa Thebault (female), France (expert in theoretical ecology and ecological networks) and Prof. Catherine Graham (female), Switzerland (expert in biogeography and ecological networks) have agreed to be members of the SAC. There are 11 people named for this project. There are 7 men and 4 women. With 36% of the fraction females is above the range being usual in the technical fields. The goal of ROBHOOT is to help younger non-PI and PI females to develop a research career in academia or in the industry (see section four for Gender balance statement for each institution involved). ROBHOOT includes four out of nine first time participants to FET under Horizon 2020, two young researchers and one SME: SCITE is the SME with all the team within the young researchers age, and the other three are IEO, EAWAG, and SRC, with a substantial contribution to the development of ROBHOOT.

Management activities:
The project coordinator (CJ Melian, EAWAG) will coordinate the work and its scientific input, communicate with EC, and organize the project reviews with the EC. The Project Manager (To be named) will work on administrative, financial and dissemination activities, and risk management. IPR set-up will be regulated by a Consortium Agreement. WP leaders will be responsible for WP planning, scientific and WP activities. WP groups will meet for the specific needs of each WP.

Methods for monitoring and reporting progress:
Meeting and reporting schedule is planned as follows: Every 3 months (oral and video-conferences) WP leaders report to the coordinator. Every 6 months the coordinator summarizes overall status to the SB. Every 6 to 12 months the coordinator setups SB meetings to review the progress of the project and to critically review the outlook for effective communication and deliverables. At months 12, 24 and 36 the SB prepares consolidated management and annual activity reports and also the coordinator and the Project Manager setup SAC meetings to obtain advice and feedback.

3.3 Consortium as a whole

Core Expertise: The ROBHOOT consortium has been designed to represent the four central project requirements and is, thus, composed of groups with long-standing track records in: (1) IFICS-EBD-CSIC: Data-driven modeling expertise of evolutionary processes including adaptation and coevolution and complex networks; (2) EAWAG and TARTU: Theoretical and numerical expertise in eco-evolutionary dynamics and deep learning networks; (3) SRC and TU GRAZ: Theoretical and numerical expertise in eco-evolutionary dynamics of communities and neural networks including synaptic plasticity, heterogeneity and diversification, and (4) SCITE and IEO: Expertise in data collection for the Oceans' sustainability case study and communication strategy for large and complex projects.

Cross-Expertise: Partners IFISC-EBD/CSIC have worked extensively in the last years on big data and complex spatiotemporal metrics, as well as in co-evolutionary processes shaping resource-consumer interaction networks allowing linking WP1 with WP2. Partners EAWAG and TARTU complement each other in network approaches. They will build eco-evolutionary process-based deep learning networks for causal knowledge discovery allowing linking WP2 with WP3. Partners EBD-CSIC, IFISC-CSIC, EAWAG, TARTU, SRC, and TU GRAZ are all familiar with abstract models allowing linking WP1 with WP2 and WP3.

Table 3.2a: List of milestones

Milestone number	Milestone name	WP	Due data (months)	Verification
MS1	Data knowledge discovery	WP1	28	OS-Soft, Paper/Conf., website
MS2	Causal knowledge discovery	WP2	30	OS-Soft, Paper/Conf., website
MS3	Discovery in federated networks	WP3	36	OS-Sof, Paper/Conf., website

3.4 Resources to be committed

Total Budget: The ROBHOOT project run over 36 months. The total budget amounts to 2380960€, which is the same as the requested EU contribution. Direct personnel costs are 1721059€ and indirect costs 476192€. The budget is well-balanced among partners according to their roles in the project and provides sufficient resources to complete all tasks. Direct cost attributed to staff is 73%. This project is open-source software-heavy, as three full open-source software will be built, which is well connected to the dissemination part from our communication partner SCITE with about 10% of the total cost. Other major cost items of Other Cost cover travel and workshops (7% of the total cost, mostly for technical meetings and integration/evaluation stages).

Table 3.2b: Critical risks for implementation		
Description of risk	**WP**	**Proposed risk-mitigation measures**
Unforeseen changes in cross-expertise synchronization	WP5	Requirements to coordinate in time end-to-start tasks across WPs in an efficient manner
Dissemination message is not under- standable by the targeted audience	WP4	The consortium as a whole will agree on the main message to transmit to the targeted audience and, for this, elaborate the appropriated material
Unforeseen changes in the WPs	WP5	Requirements to provide short monthly reports to the coordinator to allow spotting looming changes soon

Table 3.4a: Summary of staff effort	WP1	WP2	WP3	WP4	WP5	Total PM
1, EAWAG	0	**24**	0	0	7.2	**31.2**
2, CSIC	**48**	0	0	0	0	**48**
3, SCITE	6	6	11	**15**	3	**41**
4, TARTU	0	**24**	0	0	0	**24**
5, EPFL	6	6	**12**	0	0	**24**
6, TU GRAZ	0	0	**24**	0	0	**24**
7, IEO	0	0	0	**24**	0	**24**
8, URV	6	6	**12**	0	0	**24**
9, SRC	0	0	**24**	0	0	**24**
Total PM (WP)	**66**	**66**	**83**	**39**	**10.2**	**264.2**

References

[1] G. F. Fussmann et al., Functional Ecology **21**, 465 (2007).

[2] N. G. Hairston et al., Ecol. Lett. **8**, 1114 (2005).

[3] R. Iten et al., Phys. Rev. Lett. **124**, 1 (2020).

[4] E. Real et al. (2020), arXiv:2003.03384.

[5] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[6] M. E. Mastrángelo et al., Nat. Sust. **2**, 1115 (2019).

[7] ICES, ICES Scientific Reports (2019).

[8] *U.S. National Science Foundationś proposed cyberinfrastructure*, URL https://www.nsf.gov/div/index.jsp?div=OAC.

[9] *Knowledge graph covid-19*, URL http://www.odbms.org/2020/03/we-build-a-knowledge-graph-on-covid-19/.

[10] C. Rackauckas et al., Proc. Natl. Acad. Sci. U.S.A. (2020), arXiv:2001.04385.

[11] J. R. Koza, *Genetic Programming: on the programming of computers by means of natural selection* (MIT Press, Cambridge, Mass., 1992).

[12] A. Darwish, Future Computing and Informatics Journal **3**, 231 (2018).

[13] O. Seehausen et al., Nat. Rev. Gen. **15**, 176 (2014).

[14] L. Govaert et al., Functional Ecology **33**, 13 (2019).

[15] M. I. O'Connor et al., Frontiers in Ecology and Evolution **7**, 219 (2019).

[16] R. Guimerà et al., Science Advances **6**, 6971 (2020).

[17] C. Steinruecken et al., *The Automatic Statistician* (2019).

[18] E. Androulaki et al., Proceedings of the 13th EuroSys Conference, EuroSys 2018 (2018).

[19] O. P. Foundation (2018), URL https://oceanprotocol.com/.

[20] BigchainDB (2018), URL https://www.bigchaindb.com/.

[21] J. Dilley et al. (2016), arXiv:1612.05491.

[22] W. Maass, Proceedings of the IEEE **102**, 860 (2014).

[23] W. Maass, Proceedings of the IEEE **103**, 2219 (2015).

[24] M. De Domenico et al., Nature Communications **6**, 1 (2015).

[25] G. Bianconi, *Multilayer Networks* (Oxford: Oxford University Press. 402 p., 2018).

[26] E. Cozzo and Y. Moreno (2018), arXiv:1311.1759v4.

[27] A. O'Hare, Journal of Computational Biology **22**, 997 (2015).

[28] K. Cranmer et al., pp. 1–10 (2019), arXiv:1911.01429.

[29] M. Reichstein et al., Nature **566**, 195 (2019).