EU RIA H2020 Proposal Template

ROBHOOT

Abstract

In nature, biological interactions and traits diversify across multiple scales of organization, from neurons to populations and spatio-temporal scales maintaining a complex ecological balance. This never-ending eco-evolutionary diversification of traits and connectivity inspires a new computational approach for a global-sustainable knowledge-inspired society. We introduce evolutionary diversification-inspired and artificial intelligence solutions for sustainability in natural ecosystems. We validate our approach with a case study focusing on the sustainability of the Oceans. We implement federated networks, where many distinct groups of species, humans, and technologies coexist exploiting resources in complex ecosystems. Knowledge discovery running on a federated network encompasses a hybrid-technology to lay out the foundation of an open- and cooperative-science ecosystem for computation discovery in the face of global sustainability challenges. The project summarized here not only sets out to deliver knowledge discovery computation but also provides fully reproducible open-source software solutions of a science-enabled technology to connect knowledge-inspired societies to global sustainability challenges.

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

1 Excellence

1.1 Radical vision of a science-enabled technology

The ecosystems' collapse around the globe is calling for technologies to discover novel ways of sustainable exploitation of complex ecosystems. Knowledge inspired 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 inspired societies facing global sustainability challenges. 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 rapid changes in traits and interactions have been 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 interactions (i.e., cooperation, competition, antagonism, etc), and information processing and learning capabilities. Conventional Artificial Intelligence (AI) computation 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 the emergence of new interactions and traits [5]. **The goal of this project is to implement eco-evolutionary diversification-inspired solutions to make 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, communities, and ecosystems [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.

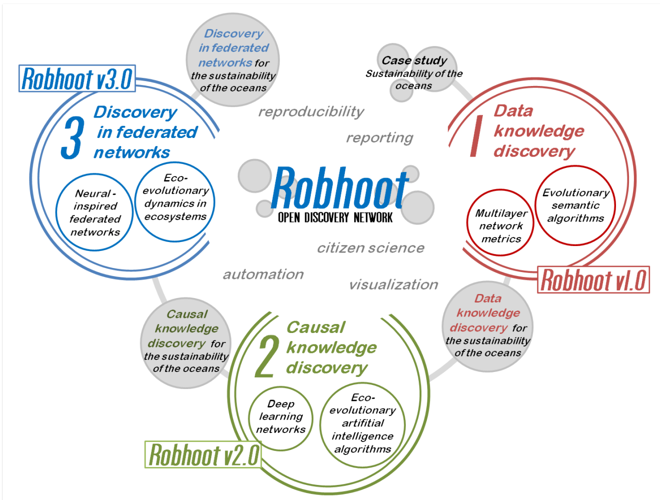

Figure 1: **Discovery in evolutionary diversification-inspired federated networks**. ROBHOOT target knowledge discovery when heterogeneous groups of species, humans and technologies share resources for a sustainable knowledge-inspired society: It introduces three science-enabled technologies: Evolutionary biology-inspired semantic algorithms for ROBHOOT v1.0 (data knowledge discovery, red), eco-evolutionary diversification-inspired AI models for ROBHOOT v2.0 (causal knowledge discovery, green), and evolutionary neural diversification-inspired federated networks for ROBHOOT v3.0 (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.

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 for community-rich digital and natural ecosystems.** To support this notion, we will perform eco-evolutionary diversification-inspired AI simulations along the whole life cycle of the project. The central goals of ROBHOOT are:

(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 paths of reliable 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 large and heterogeneous exploiting 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 arrive at powerful discovery computation principles for forecasting in federated networks when diversification in interactions and traits occur in a large and heterogeneous pool 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 [7, 8].

Going beyond ROBHOOT will go beyond state-of-the-art to implement ESA. We will explore insertions and deletions, different types of recombination, and other evolutionary-based functions to find datatype properties from ontologies and raw-data from non-semantic databases. ROBHOOT will also explore algorithms to gain understanding of the replicability of data heterogeneity contrasting different evolutionary algorithms.

Causal knowledge discovery (WP2)

Eco-evolutionary diversification-inspired AI algorithms (EEDA): Causal discovery from observable data has been extensively studied [9]. Many of these studies have used symbolic reconstruction of equations by symbolic regressions or evolutionary methods [10]. However, a common gap in much of the literature is that of reconstruction from partially known models, where the parameters represent eco-evolutionary and diversification processes, and thus, explainability can be reached more efficiently. The classical view on biology-inspired information processing technologies is to consider plasticity without structural changes, or without diversification among many interacting components [11]. Recent experimental evolution studies show that rapid trait changes with new information processing capabilities is far more complex because adaptations and speciation occurs forming new species and phenotypes [12]. 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 [13]. 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 [14].

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. This database started in 1965 and currently contains 9 million entries, 1612 species (i.e., 50 variables and traits per species), around 20 countries and 11 sampling methods (Figure 2). 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 [15, 16].

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 [18–21]. Yet, the implementation of ENDI type algorithms with learning 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, 15]. ENDI volutionary 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 [22], 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 [23, 24]. ROBHOOT v.3.0 connects knowledge discovery to diversification-inspired federated networks to study the properties of cooperative forecasting in the face of global sustainability challenges.

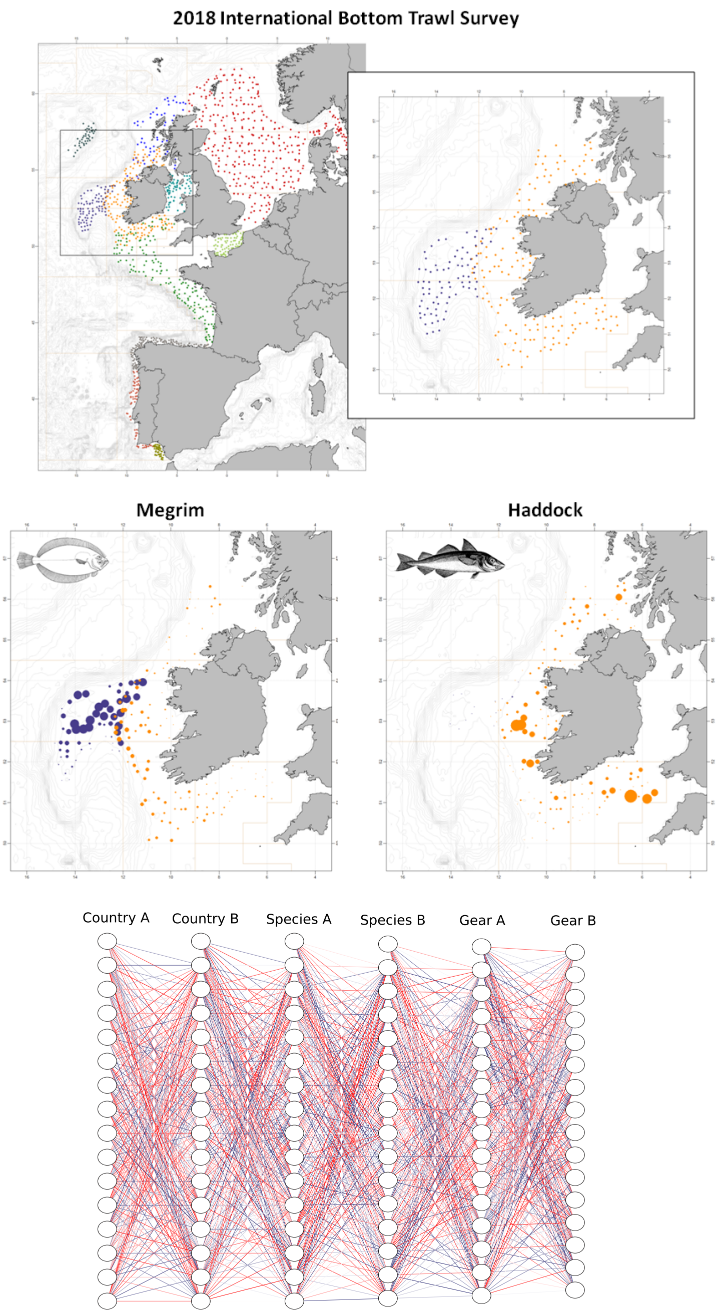

Figure 2: **Discovery for Sustainable Ecosystems**. ROBHOOT target discovery computation for new sustainability paths in complex ecosystems. The sustainability of the Oceans case study [17] will be enriched to validate the technology when many species, human groups and technologies exploit resources (**Top left**, data-color points represent a sampling from different countries. Zoomed in is the Irish ground Fish (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.

1.3 Interdisciplinarity and non-incrementality of the research proposed

The success of ROBHOOT relies on a team that includes different disciplines: 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 the 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, evolving and diversifying traits and interactions, cooperation, and dimensionality). They will explore novel paths to improve existing theories using 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 stages of 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-inspired 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**
Evolutionary semantic algorithms 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.
Eco-evolutionary dynamics of multiple traits 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.
Evolutionary neural diversification-inspired federated networks insufficiently developed: High	1-4	WP3	Alternative neural network models based on stochastic spiking neurons .
Cooperative forecasting mixing eco-evolutionary dynamics and neural nets in large scale federated networks insufficiently developed: High	1-4	WP3	Combine eco-evolutionary dynamics models with neural network models working 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 equations 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 discovery of novel evolutionary diversification-inspired algorithms to substantially improve solutions for sustainability in ecosystems. **Discovery of novel evolutionary-inspired algorithms in the context of diversifying traits, interactions, technologies and human groups for biodiversity maintenance have been hardly been investigated in this context so far**. Therefore, 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 novel technology to integrate heterogeneous sources of data and discover paths to make ecosystems sustainable. This will allow us to use the technology in public and private industry, for example, to generate rapid and 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 scenarios explaining the empirical patterns. 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. A reproducible, open-access, and automated tool will be developed for visualizing sustainability scenarios.

(IV) **Ecosystem health impact**: ROBHOOT focuses on discovery solutions for exploited ecosystems (e.g. fires, floods, droughts, overexploitation of natural resources). It uses a case study for the oceans where heterogeneous human groups share resources in a complex ecosystem 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 evolutionary biology-inspired 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. The PDE will also have a Dissemination and Exploitation Board (DEB).

(I) **Open Access**: Project reports and ISI journals 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.

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 will focus on project results and stakeholder engagement through the following activities.

(I) **Write position papers and relevant conferences**: High-impact scientific manuscripts are expected, together with the presentation of results in scientific conferences, as well as the organization of special sessions in international scientific and technological meetings.

(II) **Website**: A dedicated 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 and sharing updated versions with all target audiences.

(III) **Hackathons and robhacks**: we will organize joined activities with on-ongoing EU/International projects to attract multipliers and developers from the open-source community to the community who engage in data analytics and build hybrid evolutionary biology-inspired AI algorithms. At the end of the project we will organize a workshop specifically on *Next-generation evolutionary-biology AI-inspired solutions for global sustainability challenges*. This will help to disseminate 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.

(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.

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:(M=Milestone,D=Deliverable,R=Project Reporting,T=Task)

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	1	CSIC	48	1	18
2	Causal knowledge discovery	6	TARTU ULIKOOL	24	7	24
3	Discovery in federated networks	9	TU GRAZ	24	13	36
4	Dissemination	10	IEO	24	1	36
5	Management	6	EAWAG	31.2	1	36
			Total PMs	151.2		

Table 3.1b: **Work package description**

Work package number 1	**Lead beneficiary** EBD-CSIC				
Work package title	**Data knowledge discovery**				
Participant number	1	2	3	4	5
Short name of participant	EBD-CSIC	IFISC-CSIC	ICREA	SDSC	SCITE
Person month per participant	X	X	X	X	X
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: 1
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 [25–31].

Task T1.3: Based on the framework developed in T1.1 and T1.2, ICREA will derive automation procedures for data knowledge discovery (M15-M21) *Leader:* ***ICREA****. Contributors: 3*
Automation rules identify the ESA rules for data discovery [15]. ICREA 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:* ***SDSC****. Contributors: 4*
In this task the SDSC 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: 5*
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: 1,2,3,4,5*
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	6	7	
Short name of participant	EAWAG	TARTU	
Person month per participant	31.2	X	
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: 6*
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, 15, 32–35] will increase our ability to make stronger inferences about future sustainability challenges and solutions [36]. 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: 8*

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, ICREA will derive automation rules for causal discovery (M21-M27) *Leader:* ***ICREA****. Contributors: 3*

ICREA 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. ICREA 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:* ***SDSC****. Contributors: 4*

In this task the SDSC 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 ICREA 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: 5*

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: 6,7,8,3,4,5*

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	8	9	
Short name of participant	SRC	TU GRAZ	
Person month per participant	X	X	
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: 10*

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: 9*

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, ICREA will derive automation rules for knowledge discovery (M25-M36) *Leader: **ICREA**. Contributors: 3*

ICREA 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. ICREA 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: **SDSC**. Contributors: 4*

In this task the SDSC 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 ICREA 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: 5*

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.

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	10	11
Short name of participant	IEO	SCITE
Person month per participant	24	X
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: Elaboration, implementation and monitoring of *Dissemination and Exploitation Plan (DEP)* (M1-M36) *Leader: **IEO**. Contributors: 10, 11, 6*

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: 11, 10*

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: 11, 10*

A dedicated website and a public git ROBHOOOT repository, which is available (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 (e.g. Facebook, Twitter, Instagram, LinkedIn, Researchgate) and posts will be used to raise attention to project activities and achievements, adapted to the audience.

Task T4.4: Case study outreach: Sustainability of the Oceans (M11-M36) *Leader: **IEO**. Contributors: 10, 11*

Ocean research has been proven to capture the imagination of the public. Science communication specialists will work with the project researchers to develop and communicate a suite of outreach activities and products 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: 11, 10*

Three Hackathon events, coinciding with the three milestones (Robhoot v1.0, Robhoot v2.0 and Robhoot v3.0), are planned to attract multipliers and developers from the open-source community. At the end of the project, 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*

High-impact scientific publications are expected, together with the presentation of results in scientific conference, 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: 11, 10, 6*

This task will maximize market uptake of the results through the identification of all target groups and targeted communication actions to identified players of the value chain. An Exploitation Plan (D4.6) will be developed and maintained to ensure prompt utilization of project results by relevant stakeholders. This will include (i) SWOT (Strength, Weaknesses, Opportunities and Threats) analysis for the main project outcomes; (ii) definition of business models and plans (based on the lean business canvas); and (iii) efficient IP management to facilitate rapid commercial 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	2	6	
Short name of participant	EAWAG	IFISC-CSIC	
Person month per participant	X	X	
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: 2,6*

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: 2,6*

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: 2,6*

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: 2,6*

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: 2,6*

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: 2,6*

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

Continued on next page

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	ICREA	D	PU	21
D1.4	Reproducible demonstrator of evolutionary semantic rules for data discovery	WP1	SDSC	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	ICREA	D	PU	21
D2.4	Reproducible demonstrator of EEDA AI rules for causal discovery	WP2	SDSC	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	ICREA	D	PU	36
D3.4	Reproducible demonstrator of ENDI for discovery in federated networks	WP3	SDSC	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:
All partners of ROBHOOT are 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, France (expert in theoretical ecology and ecological networks), Mercedes Pascual, USA (expert in complex system modeling, to be confirmed), and Catherine Graham, Switzerland (expert in biogeography and ecological networks, to be confirmed)... have agreed to be members of the SAC.

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. Mention the IPR team... to set-up 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. **Keep elaborating about newcomers, gender balance, previous collaborations**

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:

(I) IFICS-EBD-CSIC: Data-driven modeling expertise of evolutionary processes including adaptation and coevolution and complex networks.

(II) EAWAG and TARTU: Theoretical and numerical expertise in eco-evolutionary dynamics and deep learning networks.

(III) SRC and TU GRAZ: Theoretical and numerical expertise in eco-evolutionary dynamics of communities and neural networks including synaptic plasticity, heterogeneity and diversification.

(IV) 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	Related work package(s)	Due data (months)	Verification
M1	Data knowledge discovery	WP1	28	OS-Software, Paper/Conf., Main website
M2	Causal knowledge discovery	WP2	30	OS-Software, Paper/Conf., Main website
M3	Discovery in federated networks	WP3	36	OS-Software, Paper/Conf., Main website

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

3.4 Resources to be committed

Total Budget: The ROBHOOT project is designed to run over 36 months. The total budget amounts to X €, which is the same as the requested EU contribution. Direct personnel costs are X €, other direct costs X €, and indirect costs X €. The total budget is well-balanced among partners according to their

Table 3.4a: Summary of staff effort	WP1	WP2	WP3	WP4	WP5	Total PM participant
1, EBD-CSIC	**24**	0	0	0	0	**24**
2, IFISC-CSIC	**24**	0	0	0	0	**24**
3, ICREA	6	6	**12**	0	0	**24**
4, SDSC	6	6	**12**	0	0	**24**
5, SCITE	6	6	11	**15**	3	**41**
6, EAWAG	0	**24**	0	0	7.2	**31.2**
7, TARTU	0	**24**	0	0	0	**24**
8, SRC	0	0	0	**24**	0	**24**
9, UNIGRAZ	0	0	**24**	0	0	**24**
10, IEO	0	0	0	**24**	0	**24**
Total PM (WP)	**66**	**66**	**59**	**63**	**10.2**	**264.2**
Total PM EU-funded (WP)	**66**	**66**	**59**	**63**	**10.2**	

roles in the project and provides sufficient resources to complete all tasks. Direct cost attributed to staff is X%. 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 X% of the total cost. Other major cost items of Other Cost cover travel and workshops (X% of the total cost, mostly for technical meetings and integration/evaluation stages: X €).

References

[1] G. F. Fussmann, M. Loreau, and P. A. Abrams, Functional Ecology **21**, 465 (2007).

[2] N. G. Hairston, S. P. Ellner, M. A. Geber, T. Yoshida, and J. A. Fox, Ecology Letters **8**, 1114 (2005).

[3] R. Iten, T. Metger, H. Wilming, L. Del Rio, and R. Renner, Physical Review Letters **124**, 1 (2020), ISSN 10797114, 1807.10300.

[4] E. Real, C. Liang, D. R. So, and Q. V. Le (2020), 2003.03384, URL http://arxiv.org/abs/2003.03384.

[5] J. Schmidhuber, Neural Networks **61**, 85 (2015), ISSN 08936080, URL https://linkinghub.elsevier.com/retrieve/pii/S0893608014002135.

[6] M. E. Mastrángelo, N. Pérez-Harguindeguy, L. Enrico, E. Bennett, S. Lavorel, G. S. Cumming, D. Abeygunawardane, L. D. Amarilla, B. Burkhard, B. N. Egoh, et al., Nature Sustainability **2**, 1115 (2019), ISSN 2398-9629, URL https://doi.org/10.1038/s41893-019-0412-1.

[7] *U.S. National Science Foundationś proposed Cyber-Infrastructure* (????), URL https://www.nsf.gov/div/index.jsp?div=OAC.

[8] *Knowledge Graph COVID-19* (????), URL http://www.odbms.org/2020/03/we-build-a-knowledge-graph-on-covid-19/.

[9] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, and A. Ramadhan, Proceedings of the National Academy of Sciences of the U.S.A. (2020).

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

[11] A. Darwish, Future Computing and Informatics Journal **3**, 231 (2018), ISSN 2314-7288.

[12] O. Seehausen, R. Butlin, and I. e. a. Keller, Nature Reviews Genetics **15**, 176 (2014).

[13] L. Govaert, E. A. Fronhofer, S. Lion, C. Eizaguirre, D. Bonte, M. Egas, A. P. Hendry, A. De Brito Martins, C. J. Melián, J. A. M. Raeymaekers, et al., Functional Ecology **33**, 13 (2019).

[14] M. I. O'Connor, M. W. Pennell, F. Altermatt, B. Matthews, C. J. Melián, and A. Gonzalez, Frontiers in Ecology and Evolution **7**, 219 (2019).

[15] R. Guimerà, I. Reichardt, A. Aguilar-Mogas, F. A. Massucci, M. Miranda, J. Pallarès, and M. Sales-Pardo, Science Advances **6**, eaav6971 (2020), ISSN 23752548.

[16] C. Steinruecken, E. Smith, D. Janz, J. Lloyd, and Z. Ghahramani (2019), pp. 161–173, ISBN 9783030053185.

[17] ICES Scientific Reports (2019).

[18] N. Durov, pp. 1–132 (2017).

[19] E. Androulaki, A. Barger, V. Bortnikov, S. Muralidharan, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Murthy, C. Ferris, et al., Proceedings of the 13th EuroSys Conference, EuroSys 2018 **2018-Janua** (2018), arXiv:1801.10228v2.

[20] Ocean Protocol Foundation, BigchainDB GmbH, and DEX Pte. Ltd, pp. 1–51 (2018), URL https://oceanprotocol.com/.

[21] BigchainDB GmbH, BigchainDB. The blockchain database. pp. 1–14 (2018), URL https://www.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf.

[22] J. Dilley, A. Poelstra, J. Wilkins, M. Piekarska, B. Gorlick, and M. Friedenbach (2016), 1612.05491, URL http://arxiv.org/abs/1612.05491.

[23] W. Maass, Proceedings of the IEEE **102**, 860 (2014), ISSN 00189219.

[24] W. Maass, Proceedings of the IEEE **103**, 2219 (2015), ISSN 00189219.

[25] *Openstreet* (????), URL https://www.openstreetmap.org/{#}map=5/51.509/-0.110.
[26] *Blue-cloud* (????), URL https://www.blue-cloud.org/.
[27] *HOT* (????), URL https://www.hotosm.org/.
[28] *Elixir open data* (????), URL https://elixir-europe.org/.
[29] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas, Nature Communications **6**, 1 (2015), ISSN 20411723, arXiv:1011.1669v3.
[30] G. Bianconi, *Multilayer Networks* (Oxford: Oxford University Press. 402 p., 2018).
[31] E. Cozzo and Y. Moreno (2018), arXiv:1311.1759v4.
[32] A. O'Hare, Journal of Computational Biology **22**, 997 (2015), ISSN 1066-5277, URL http://online.liebertpub.com/doi/10.1089/cmb.2015.0086.
[33] Y. Gil and B. Selman (2019), 1908.02624, URL https://bit.ly/2ZNVBVbhttp://arxiv.org/abs/1908.02624.
[34] K. Cranmer, J. Brehmer, and G. Louppe, pp. 1–10 (2019), 1911.01429, URL http://arxiv.org/abs/1911.01429.
[35] G. Futia and A. Vetrò, Information (Switzerland) **11** (2020), ISSN 20782489.
[36] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and &. Prabhat, Nature (????), ISSN 0028-0836, URL www.nature.com/nature.

4 Members of the consortium

4.1 Participants (applicants)

EAWAG

Dr. Carlos J. Melian will work for EAWAG in ROBHOOT.
Carlos J Melian takes the official coordinator lead of the ROBHOOT project and takes care of the management.

The Swiss Federal Institute for Aquatic Science and Technology (EAWAG) is an independent research institute within the Swiss Federal Institute of Technology (ETH) domain. As such it is an independent partner in a network of exceptionally strong research and education institutions (2 federal universities and 4 federal research institutes). EAWAG is a world-leading water research institute. EAWAG hosts over 300 research fellows, postdocs and PhD students, who are supported by technical and support staff.

Contributions to ROBHOOT ...

Dr. Carlos Melián is a tenured researcher in Theoretical Evolutionary Ecology at EAWAG and associate professor at the University of Bern. Dr. Melián is widely recognized as an expert in Eco-evolutionary networks where he has contributed with novel approaches combining stochastic modeling and empirical patterns to study the interaction between ecological and evolutionary dynamics in multispecies assemblages. Dr. Melian has made important contributions to the fields of Ecological Networks (e.g. De Laender and Melián, 2014, Ecol Lett; Melián and Křivan, 2015, AmNat), Eco-evolutionary networks (e.g. Melián et al., 2011, Adv Ecol Res; Andreazzi and Melián, 2018, PRSB), and Diversification on eco-evolutionary networks (e.g. Melián et al., 2012, PLoS Comput Biol; Leprieur, Melián, Pellissier, 2016, Nat Commun). Most of his contributions combine stochastic modeling, large empirical datasets, and Bayesian approximations, to quantify the impact of intra- and inter-specific trait variation on species interactions, divergence and the macroscopic properties of ecological networks. He has been Principal Investigator in 15 projects obtained in 5 different countries (Spain, USA, UK, Germany and Switzerland) with a total of approx. 1 Million Euro. He has successfully co-supervised 5 PhD students and supervised 7 postdocs. The feasibility of this proposal is firmly established by his track record further reinforced by his solid and active international network of collaborators. Among others, he works with Prof. S. Allesina (U Chicago, USA), Dr. A. Eklöf (Linköping U, Sweden), Prof. P. Guimares (U Sao Paulo, Brazil), Prof. M. O'Connor (U Vancouver, Canada), and Dr. F. De Laender (U Namur, Belgium). Dr. Melián has expertise combining skills in networks and experience in modeling complex multi-scale eco-evolutionary networks. He also has combined basic and applied-oriented research; 2) Integrating a range of methodologies: he is experienced with statistical and mathematical modeling, and has analytical and advanced programming skills; 3) Extensive experience collaborating with theorists and empiricists: he has collaborated with researchers of diverse fields i.e. mathematics, ecology, evolutionary biology, conservation science.

List of publications

• Melián C, et al. 2018. Deciphering the interdependence between ecological and evolutionary networks. Trends in ecology evolution 33,7: 504-512.
• Andreazzi C, Guimaraes P, Melián C. 2018. Eco-evolutionary feedbacks promote fluctuating selection and long-term stability of antagonistic networks. Proc. R. Soc. B 285: 20172596.
• Melián C, Seehausen O, Eguiluz V, Fortuna M, Deiner K. 2015. Diversification and Biodiversity Dynamics of Hot and Cold Spots. Ecography 38, 393-401.
• Melián C, et al. 2015. Dispersal dynamics in food webs. American Naturalist 185, 2: 157-168.

• Melián C., et al. 2014. Individual trait variation and diversity in food webs. Advances in Ecological Research. Vol. 50. Academic Press, 207-241.

List of relevant projects

2020 Melián, C. J. and Ferrão Filho, Aloysio S. Granted: Brazilian-Swiss Joint Research Programme SNSF, Title: Feedbacks between coevolving predator-prey interactions and the functioning of aquatic ecosystems. Period: 24 Months, SFr 228k
2018 Melián, C. J., Andreazzi, C., and Astegiano, J. SNSF, Scientific exchange program, Title: Biodiversity Dynamics in Coevolutionary Metaecosystems. Period: 3 Months, SFr 20k
2016 Melián, C. J., Matthews, B., Seehausen, O., and Harmon, L. J. Granted: Swiss National Science Foundation, International exploratory workshops. Title: Interactions on Trees. Period: 1 Week, SFr 21k.
2015 Kalinkat, G., and Melián, C. J. Granted: German Academic Exchange Service (DAAD). Germany. Title: Analysing the interplay between allometric constraints and intraspecific trait variation to predict food web dynamics. Period: 6 Months, SFr 19k.
2015 Melián, C. J. Granted: Swiss National Science Foundation, Division III. Switzerland. Title: A theory for next- generation food web data. Period: 2 years (Postdoc), SFr 161k.

Infrastructure relevant to the proposed work

EAWAG in Kastanienbaum Lucerne offers excellent office, meeting rooms, laboratory and testing facilities in modern, state-of-the-art buildings. EAWAG provides access to first-class research facilities that regularly offer training for the use of equipment, tools and software. Of particular relevance for this research project is the access to two computing clusters “Leonhard” and “Euler” with more than 50 000 processor cores available for scientific computations, and training for their use offered by ETH Zürich.

CSIC

The Spanish National Research Council (CSIC) is Spain’s largest public research institution and ranks third among Europe’s largest research organizations. Attached to the Spanish Ministry of Science and Innovation, the CSIC plays a key role in scientific and technological policy in Spain and worldwide. According to its Statute (Article 4), it has 4 main missions:

- to foster multidisciplinary scientific and technological research,
- knowledge transfer to industry and society,
- education and training of scientific and technical staff,
- creation of Technology-Based Companies (spin-offs).

The CSIC has more than 10 000 employees, including nearly 4 000 staff researchers. Currently, it has 120 institutes spread across the country, of which 67 of them are fully-owned institutes and 53 are Joint Research Units in partnership with other Spanish universities or research institutions.

The CSIC supports research and training across a wide range of knowledge, from the most basic or fundamental aspects of science to the most complex technological developments; from human and social sciences to food science and technology, including biology, biomedicine, physics, chemistry and materials, natural resources and agricultural sciences. It carries out research in all fields of knowledge, distributed in three global areas: Life, Society and Materia.

The CSIC has a Vice-presidency of International Affairs dedicated to the planning and promotion of international relations, including the participation in international and European programmes, especially the Framework Programme of Research and Innovation of the EU. It also has a delegation in Brussels that holds the institutional representation of the CSIC before the institutions of the EU and other relevant organizations and forums, and fosters and potentiates relationships with the representatives of foreign research organizations similar to the CSIC.

The CSIC produces 20% of the national scientific output (over 10 000 publications in high impact international journals in 2017) and remains the first institution in Spain in the generation of patents, with around 200 patent applications in 2017. The Vice-presidency for Technology Transfer assists the CSIC's researchers with patent evaluation and application processes, commercialization of the CSIC's technology offer and with the creation of start-ups.

The CSIC has also a broad experience managing large and singular infrastructures. For instance, it provides services to the entire scientific community through the management of several Singular Scientific and Technological Infrastructures (ICTS) such as the "Calar Alto" Astronomical Observatory, the "Doñana" Biological Station, the European Synchrotron Radiation Facility, the "Hesperides" Ocean Research Vessel, the Integrated Micro and Nanoelectronics Clean Room, the "Juan Carlos I" Antarctic Base, the "Max Von Laue-Paul Langevin" Institute and the "Sarmiento de Gamboa" Ocean Research Vessel. In addition, the CSIC has a broad experience in conducting RI projects funded by national, European and international public and private entities.

The CSIC is a major player in the development of the European Research Area (ERA) and therefore a significant contributor to the European integration process. Within the 7th Framework Programme the CSIC is listed as the 1st organization in Spain and the 4th in Europe within the research organizations, with a total of 726 signed actions and a contribution of over 264 million euros (E-CORDA).

As of December 2019, the CSIC has obtained 643 projects in H2020, with a total EU financial contribution of 270 million euros and is listed as the 1st organization in Spain and the 4th participant by number of projects (E-CORDA).

The CSIC is a main actor in the ERC programme, with a total of 112 projects signed as Host Institution in all areas of knowledge and is also an active member in the European Institute for Innovation and Technology's (EIT) Knowledge and Innovation Communities (KICs), currently participating in the EIT Raw Materials and EIT Food as core partner. As for European and International Programmes other than H2020, the CSIC has achieved 27 LIFE (2014 – 2020) grants, 42 INTERREG V projects, mainly SUDOE, POCTEP, POCTEFA, MED, ATLANTIC or MAC. Additionally, the CSIC has obtained 30 contracts resulting from call for tenders published by EU institutions, agencies and other bodies (e.g. ESA, EASME, JRC etc.) with a total EU financial contribution of 1,5 million euros. Regarding International projects i.e. those projects which are not financed with funds from the European Commission, in recent years the CSIC has received funds mainly from the following funding entities:

- USA funding entities: The Center for Produce Safety (CPS), the National Institute of Health (NIH), the
  - Michael J. Fox Foundation, the U.S. Department of Energy Office of Science, the Gordon and Betty Moore
  - Foundation and the Muscular Dystrophy Association.
- European funding entities: The University of Southampton, the Research Council of Norway, the Swiss
- National Science Foundation, the Rothschild Foundation and the Novo Nordisk Foundation.
- Iberoamerican funding entities: The Administrative Department of Science, Technology and Innovation of Colombia (Colciencias) and the Ibero-American Programme on Science and Technology for Development (CYTED).

Biological Donana Station (EBD)

Brief description of EBD ...

Contributions to ROBHOOT ...

Dr. Miguel A. Fortuna is an ecologist and evolutionary biologist turned network scientist who thinks differently about problem-solving. He conducts interdisciplinary research by combining mathematical models, computer simulations, and database analysis, to answer questions that go beyond the traditional boundaries among disciplines, merging ecology with evolution, sociology, genetics, software design, and artificial life.

His current research line builds on his previous research and is among the few trying to understand how evolution in complex networks of interactions can help us control human diseases. This research line combines, with a solid methodology, community ecology and evolutionary biology in a new fresh way. It has implications in at least three burgeoning fields of biotechnological and biomedical research: 1) cancer research (i.e., recent advances have shown that tumours—like species striving for survival—harbor intricate population dynamics, which suggests the possibility to exploit the ecology of tumours for treatment), 2) phage therapy (i.e., recent findings are showing the success of using phage cocktails to fight antibiotic resistance), and 3) human microbiome (i.e., the manipulation of evolving interactions among bacteria to restore unbalanced human microbial ecosystems).

List of publications

1. Fortuna et al. (2019). Coevolutionary dynamics shape the structure of bacteria-phage infection networks. Evolution, 73:1001-1011.

This is the first time that the influence of coevolutionary dynamics (i.e., arms race dynamics versus fluctuating selection) on the architecture of an ecological network is quantified in an experimental host-parasite system. This study is the starting point towards integrating coevolution into ecological network approaches.

2. Fortuna et al.(2017). Non-adaptive origins of evolutionary innovations increase network complexity in interacting digital organisms. Philosophical Transactions of the Royal Society B., 372:20160431.

This was the first study on webs of self-replicating and coevolving computer programs (i.e., digital organisms) aimed at disentangling the ecological and evolutionary mechanisms shaping species interaction networks. I found that host resistance traits arising as exaptations increase the complexity of host-parasite networks.

3. Fortuna et al. (2010). Nestedness versus modularity in ecological networks: two sides of the same coin? Journal of Animal Ecology, 79:811-817.

This paper was published at a time when researchers thought that the two most pervasive structural properties of species interaction networks, nestedness and modularity, were opposed to each other. I showed that only when the complexity of the network is large enough, network structure can be either modular or nested.

4. Fortuna et al. (2009). Networks of spatial genetic variation across species. Proceedings of the National Academy of Sciences, USA, 106:19044-19049.

In this paper I explored the consequences of habitat fragmentation for the maintenance of genetic variation. It was the first time that the structure of genetic variation across different species inhabiting the same landscape was compared, opening new research paths in landscape genetics.

5. Fortuna Bascompte. (2006). Habitat loss and the structure of plant-animal mutualistic networks. Ecology Letters, 9:281-286.

I developed here the first spatially-implicit model to describe the dynamics of mutualistic metacommunities interacting in realistic ways. This study paved the way toward studying the biogeography of species interactions.

List of relevant projects

1. Name of the project: An eco-evolutionary network of biotic interactions. Entity where project took place: Scientific Research Network (WOG) City of entity: Brussels, Belgium Name principal investigator (PI, Co-PI....): Dries Bonte; Luc Brendonck; Erik Matthysen; Hans Jacquemyn; Filip Volckaert; Lander Baeten; An Martel; Frederik Hendrickx; Ellen Decaestecker; Frederik De Laender; Nicolas Schtickzelle; David G. Angeler; Florian Altermatt; Rampal S. Etienne; Rosemary Gillespie; Mark Urban; Erik Svensson; Mathew A. Leibold; Joel White; Alison Duncan; Miguel A. Fortuna; Kerstin Johannesson; Steven Declerck; Michael Begon; Justin Travis. N° of researchers: 25 Funding entity or bodies: Research Foundation Flanders (FWO) Start-End date: 01/01/2016 - 31/12/2020 Total amount: 62.500

2. Name of the project: Dinámica espacio-temporal de redes de flujo génico: unidades de conservación y propagación de enfermedades de anfibios (RNM-8147) Entity where project took place: Estación Biológica de Doñana (EBD-CSIC) City of entity: Seville, Spain Name principal investigator (PI, Co-PI....): Jordi Bascompte; Andrew P. Dobson; Miguel A. Fortuna; Jaime Bosch Nº of researchers:4 Funding entity or bodies: Proyecto de Investigación de Excelencia, Junta de Andalucia Start-End date: 01/02/2013 - 31/01/2016 Total amount: 134.242

3. Name of the project: Síntesis Ecológica — Postdoctoral Fellowship JAE-Doc (JAEDOC025) Entity where project took place: Estación Biológica de Doñana (EBD-CSIC) City of entity: Seville, Spain Name principal investigator (PI, Co-PI....): Jordi Bascompte; Miguel A. Fortuna Nº of researchers: 2 Funding entity or bodies: Ministry of Economy and Competitiveness European Social Fund Start-End date: 01/09/2012 - 31/08/2015 Total amount: 82.620,72

4. Name of the project: Unifying ecological and evolutionary networks — Marie Curie International Outgoing Fellowship (IOF) Entity where project took place: Princeton University (USA) City of entity: Princeton, New Jersey, United States of America Name principal investigator (PI, Co-PI....): Miguel A. Fortuna; Simon Levin; Jordi Bascompte Nº of researchers: 3 Funding entity or bodies: European Community (International Outgoing Fellowship (IOF) Type of entity: 7th European Community Framework Programme Start-End date: 01/05/2009 - 30/04/2012 Total amount: 225.036,19

5. Name of the project: Integrando redes espaciales y genética de poblaciones: conservación de dos especies de anfibios autóctonas de Andalucia (RNM-02928) Entity where project took place: Estación Biológica de Doñana (EBD-CSIC) City of entity: Seville, Spain Name principal investigator (PI, Co-PI....): Jordi Bascompte; José A. Godoy; Peter Buston; Miguel A. Fortuna Nº of researchers: 4 Funding entity or bodies: Proyecto de Investigación de Excelencia, Junta de Andalucia Start-End date: 01/08/2008 - 01/07/2011 Total amount: 124.330,12

Instituto de Fisica Interdisciplinar y Sistemas Complejos (IFISC-CSIC)

IFISC (Institute for Cross-Disciplinary Physics and Complex Systems) is a joint research Institute of the University of the Balearic Islands (UIB) and the Spanish National Research Council (CSIC) created in 2007. IFISC has been awarded in 2018 the "Unit of Excellence María de Maeztu" distinction, entering the selective SOMMa Alliance and thus consolidating IFISC as a reference institute in the research field of complex systems. The award has been granted by the Spanish National Agency (AEI), Ministry of Science, Innovation and Universities. Emerging from a backbone transversal research line of exploratory nature on Complex Systems, Statistical and Nonlinear Physics, IFISC has 5 research lines of transfer of knowledge in the interface with other disciplines (Quantum Technologies, Information and Communication Technologies, Earth Sciences, Life Sciences and Social Sciences). These are: i) Biocomplexity, ii) Dynamics and collective phenomena of social systems, iii) Transport and Information in Quantum Systems, iv) Nonlinear Photonics, v) Nonlinear dynamics in fluids.

Dr. V.M. Eguíluz is a complex systems' scientist with an interest in interdisciplinary applications at the interface between Physics, Biology and Social Sciences. Our early studies on co-evolution networks showed the relevance of network plasticity on the emergence of cooperation, and as a generic mechanism leading to fragmentation transitions. The extensive study of the voter model on complex networks is an example of the micro-macro connection in social collective phenomena: how to link microscopic rules to macroscopic emergent phenomena. Recently we combined census data and election results to present the first model based on microscopic rules compatible with the patterns of voting. In Biology, we introduced the first large scale functional network of the brain. Our expertise on complex networks, on the one hand, and the more recent research activity on the connection between ecological-human activity-and environmental factors In connection to the current project supports his contribution to the current proposal. Current research includes the characterization and modeling of the structure-function relationship of real systems. He is Associate Editor of frontiers in Physics (since 2016) and Advances in Complex Systems (since 2007), Young Researcher Award from the Spanish Royal Physical Society (2003).

List of publications

GC Hays et al, Key questions in marine megafauna movement ecology, Trends in Ecology & Evolution 31 (6), 463-475 (2006).

F Vazquez, VM Eguíluz, M San Miguel, Generic absorbing transition in coevolution dynamics, Physical Review Letters 100, 108702 (2006).

A Cózar et al, The Arctic Ocean as a dead end for floating plastics in the North Atlantic branch of the Thermohaline Circulation, Science Advances 3 (4), e1600582 (2017).

AF Rozenfeld et al, Network analysis identifies weak and strong links in a metapopulation system, Proceedings of the National Academy of Sciences 105, 18824-18829 (2008).

N Queiroz et al. Global spatial risk assessment of sharks under the footprint of fisheries Nature 572, 461–466 (2019).

List of relevant projects

Coupled Animal and Artificial Sensing for Sustainable Ecosystems (CAASE). The Red Sea as a CAASE study Project OSR-KAUST. From 2016 to 2020 Coordinator: C. Duarte (KAUST); PI IFISC: V.M. Eguíluz.

SPASIMM Spatiotemporality in sociobiological interactions, models and methods. Project FIS2016-80067-P of the MINECO (Spain). From 2017 to 2020. PI: V.M. Eguìluz, K. Klemm.

LASAGNE: multi-LAyer SpAtiotemporal Generalized Networks. FP7-ICT-2011-8 Collaborative Project, Grant Agreement 318132. From 2012 to 2015 Coordinator: S. Thurner (Vienna University).

MODASS: Modeling and analysis of social systems: structural evolution, temporal correlations and opinion propagation) Project FIS2011-24785 of the MICINN (Spain). From 2012 to 2015. PI: V.M. Eguíluz.

IBESINC: Network on Dynamics and synchronization in networks. Complementary action FIS2010-09832-E (subprogram FIS) of MCINN (Spain). 2011. PI: J. M. Buldú (U. Rey Juan Carlos); Coordinator IFISC: V.M. Eguíluz.

SPANISH INSTITUTE OF OCEANOGRAPHY (IEO)

The Spanish Institute of Oceanography (IEO), founded in 1914, is a public research organization dedicated to research in marine science, especially in relation to scientific knowledge of the oceans, sustainability of fishing resources and the marine environment. Attached to the Spanish Ministry of Science and Innovation, the IEO is an autonomous body, with its own personality and legal assets, with a workforce of approximately 700 people, of which 80% are research and research support personnel. The IEO budget exceeds 65 million Euros and has a wide geographical coverage and important facilities. It has a central office in Madrid, nine coastal oceanographic centers and five marine culture experimentation plants. Its oceanographic fleet, with more than twenty vessels, has five important oceanographic vessels. The IEO not only conducts basic and applied research, but also provides scientific and technological advice to administrations in matters related to oceanography and marine sciences, being the research and advisory body for the fisheries sector policy of the Spanish Government. Furthermore, it is the scientific and technological representative of Spain in most of the forums and international organizations related to the sea and its resources.

Dr. Francisco Baldó works for the Spanish Institute of Oceanography in Cadiz as a Senior Researcher. His main scientific motivation is to understand the biodiversity of marine ecosystems and contribute to the sustainable management of fisheries resources. During many years, he has been involved in long-term ecological research in the Guadalquivir estuary, which has set up a time series of monthly sampling of the aquatic communities of the estuary. The basis for his PhD was theoretical individual-based models based on a high resolution individual-based food web data-set of the Guadalquivir estuary. Since 2008 he is the cruise leader of the Porcupine bottom trawl survey, which takes place every September at the Porcupine bank (West of Ireland) and a member of the ICES International Bottom Trawl Surveys Working Group (IBTSWG), which coordinates the European bottom trawling surveys in the North Sea and the North East Atlantic. Overall, he has extensive experience in field work including over 50 oceanographic cruises. The combination of the empirical and theoretical facets is, perhaps, the most outstanding characteristic of his background.

List of publications

• Barros-García D, et al. 2020. Phylogeography highlights two different Atlantic/Mediterranean lineages and a phenotypic latitudinal gradient for the Deep-Sea Morid Codling Lepidion lepidion (Gadiformes: Moridae). Deep-Sea Research Part I: Oceanographic Research Papers 157: 103212.
• Bañón R, De Carlos A, Alonso-Fernández A, Ramos F, Baldó F. 2020. Apparently contradictory routes in the expansion of two fish species in the Eastern Atlantic. Journal of Fish Biology 96, 1051–1054.
• Carvalho-Souza GF, et al. 2019. Natural and anthropogenic effects on the early life stages of European anchovy in one of its essential fish habitats, the Guadalquivir estuary. Marine Ecology-Progress Series 617-618: 67-79.
• Melián CJ, Baldó F, et al. 2014. Individual trait variation and diversity in food webs. Advances in Ecological Research. Vol. 50. Academic Press, 207-241.
• Melián CJ, et al. 2011. Eco-evolutionary dynamics in individual-based food webs. Advances in Ecological Research. Vol. 45. Academic Press, 225-268.

List of relevant projects

• PIMETAN: The role of penguins in the biogeochemical cycles of trace metals in the Southern Ocean. Project RTI2018-098048-B-I00 of the Spanish Program for Research, Development and Innovation. From 2019 to 2021. 153,670 Euros.
• ERDEM: Evaluation of demersal resources by direct methods in the ICES area. European Maritime and Fisheries Fund (EMFF). From 2018 to 2020. 70,420 Euros.
• DILEMA: Oceanografic dynamics in the Gulf of Cadiz and its influence on the planktonic ecosystem. Project CTM2014-59244-C3-2-R of the Spanish Program for Research, Development and Innovation. From 2015 to 2018. 121,000 Euros.
• ECOBOGUE: Ecology of early stages of anchovy Engraulis encrasicolus: The role of the Guadalquivir Estuary - Gulf of Cadiz coupled system in the recruitment of the species. Project RNM-7467 of the Andalusian Program for Research. From 2013 to 2015. 114,885 Euros.
• TPEA: Transboundary Planning in the European Atlantic. DG MARE (European Commission). From 2012 to 2014. 1,250,000 Euros.

Swiss Data Science Center (SDSC)

Dr. Christine Choirat is the Chief Innovation Officer of the Swiss Data Science Center (SDSC, https://datascience.ch/) where she provides leadership over the lifecycle of sponsored projects in the domains of environmental science, health science and technology, personalized medicine, and open science. She also fosters engagement with partners to facilitate the adoption of FAIR data and workflow sharing platforms nationally and internationally. At SDSC, she leads the strategic development and outreach efforts of the Renku platform (https://renkulab.io/) for reproducible, collaborative and open data-driven science. Dr. Choirat has over 15 years of experience in data science, computational statistics and in industry-standard software development. She is passionate about education in data science and reproducible research. She created a module for the HarvardX MOOC “Principles, Statistical and Computational Tools for Reproducible Science” and is also the instructor of “Computing for Big Data” at Harvard University. Dr. Choirat will advise on best practice in data science to organize FAIR data management, and on creating and distributing high-quality software tools.

The mission of the Swiss Data Science Center (SDSC) is to accelerate the adoption of data science and machine learning techniques within academic disciplines between the Swiss academic community at large, and the industrial sector. In particular, it addresses the gap between those who create data, those who develop data analytics and systems, and those who can extract value from it. The Center is composed of a large multi-disciplinary team of data scientists and computer scientists, and experts in select domains, with offices in Lausanne and Zurich (www.datascience.ch).

List of Publications

Wasfy, J. H., Zigler, C. M., Choirat, C., Wang, Y., Dominici, F., Yeh, R. W. (2017). Readmission rates after passage of the hospital readmissions reduction program: A pre-post analysis. Annals of Internal Medicine, 166(5), 324–331. https://doi.org/10.7326/M16-0185 Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., . . . Schwartz, J. D. (2017). Air Pollution and Mortality in the Medicare Population. New England Journal of Medicine, 376(26), 2513–2522. https://doi.org/10.1056/NEJMoa1702747 Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J. D., Dominici, F. (2017). Association of short-term exposure to air pollution with mortality in older adults. JAMA - Journal of the American Medical Association, 318(24), 2446–2456. https://doi.org/10.1001/jama.2017.17923 Zigler CM, Kim C, Choirat C, Hansen JB, Wang Y, Hund L, Samet JM, King G, and Dominici F (2016) Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations. The Health Effects Institute, Cambridge, MA. http://pubs.healtheffects.org/view.php?id=453 Henneman, L. R. F., Choirat, C., Zigler, C. M. (Forthcoming). Accountability assessment of health improvements in the United States associated with reduced coal emissions between 2005 and 2012. Epidemiology.

List of relevant projects

1. R01 ES026217 (PI: Zigler) - NIH
   Causal Inference with Interference for Evaluating Air Quality Policies
   02/01/2016 - 01/31/2021

   Public health interventions routinely target upstream determinants of health to advance the health of pop- ulations, but methods for causal inference to evaluate their effectiveness are limited by a current focus on clinical investigations of individual-level therapies. This work develops methods for bipartite causal inference with interference for the evaluation of complex public health interventions. We deploy the newly-developed methodology to compare the effectiveness of regulatory policies designed to reduce health burden associated with pollution emissions from power plants across the US.
   Role: Co-Investigator

2. R01 ES028033 (PI: Laden) - NIH
   Relationship Between Multiple Environmental Exposures and CVD Incidence and Survival: Vulnerability and Susceptibility
   12/15/2017 - 11/30/2022

   The major goals of the proposed project are to study associations of lmultiple environmental exposures on cardiovascular disease (CVD), mortality and survival after a non-fatal CVD event in the context of multiple confounders and effect modifications. We will be developing new statistical methods, assessing air pollution (particulate matter, nitrogen dioxide, and ozone) and weather (e.g. temperature variability) as main effects, and evaluating effect modification by contextual, lifestyle and genetic factors.
   Role: Co-Investigator

3. 83587201-0 (PI: Koutrakis) - EPA grant (ACE Center)
   12/01/2015 - 11/30/2020
   Regional Air Pollution Mixtures: The Past and Future Impacts of Emission Controls and Climate Change on Air Quality and Health

The overarching goal of this Center is to generate new scientific knowledge on past and future US in air quality and the associated health impacts. Specifically, we will investigate the sources, composition, trends and effects of regional air pollutant mixtures across the US over a relatively long chronological period span- ning past and future years (2000-2040), and will examine the influence of climate, socioeconomic factors, policy decisions, and control strategies on air pollution, human health and economic outcomes.
Role: Co-Investigator

4. swissuniversities (PI: Choirat)
   01/01/2020 - 21/31/2020
   Easy FAIR: Supporting the adoption of FAIR and reproducible digital scholarship with Renku

   The World Wide Web and digital technologies are fundamentally changing how scientific knowledge is produced, disseminated and preserved. This transformation represents an opportunity to make the scientific en- deavor more transparent, inclusive, collaborative, reproducible and impactful. However, research in the digital age requires new standards, tools and infrastructures, as well as a new set of research skills. The change is reflected by new requirements from funders, journals, from the research community in general, as well as society at large. Supporting researchers active in Swiss institutions in their adoption of digital best practice is necessary to guarantee they produce research results of the highest quality and impact.
   Role: Principal Investigator

5. EPFL Open Science Fund (PI: Unser)
   01/08/2020 - 02/28/2021
   Interdisciplinary Collaborations in Imaging at EPFL: A Pilot Project with RENKU

   The present project is a collaboration between Imaging@EPFL and the Swiss Data Science Center (SDSC). It aims at the evaluation of how the open-source platform RENKU can facilitate the sharing of data, meta- data, and code within the imaging community at EPFL. Concretely, we shall extensively test and improve RENKU in a series of collaborations that will involve two (or more) imaging laboratories with complementary skills. In parallel, we shall build a repertoire of reproducible, reusable, and well-documented image-processing workflows, and make them accessible to the whole imaging community. A range of new features in RENKU will be developed throughout the project to support these objectives. Of central importance is the definition of a common language (ontology) to describe the vast and heterogeneous world of imaging at EPFL. The widespread adoption of a transparent environment for interactive research can have a huge impact for EPFL, as the imaging community represents about a quarter of its laboratories (80+ groups). Role: Co-Principal Investigator

4.2 Third parties involved in the project (third party resources)

5 Ethics and Security

5.1 Ethics

5.2 Security

[1]

[1] Article 37.1 of the Model Grant Agreement: Before disclosing results of activities raising security issues to a third party (including affiliated entities), a beneficiary must inform the coordinator – which must request written approval from the Commission/Agency. Article 37.2: Activities related to "classified deliverables" must comply with the "security requirements" until they are declassified. Action tasks related to classified deliverables may not be subcontracted without prior explicit written approval from the Commission/Agency. The beneficiaries must inform the coordinator – which must immediately inform the Commission/Agency – of any changes in the security context and –if necessary – request for Annex 1 to be amended (see Article 55).