EU RIA H2020 Proposal Template

ROBHOOT

Abstract

Eco-evolutionary biology teaches us how interactions and traits evolve and diversify across levels of biological organization, from neurons to populations. Evolving networks in nature with ever changing traits and connectivity patterns can inspire a new computing discovery for a global-sustainable knowledge-inspired society. Many studies have shown global sustainability could be achieved by strengthening transparency, communication, and rapid access to discovery technologies. Sustainability goals, however, strongly depend on global access to discovery-based knowledge. Yet, science-enabled technologies targeting knowledge discovery to reach sustainability goals are not in place. We propose an eco-evolutionary biology-inspired computing discovery technology for a knowledge-inspired society. We introduce evolutionary biology-inspired and artificial intelligence solutions to explore sustainability of the Seas in federated networks, networks composed by many distinct groups of individuals within species, humans and technologies exploiting resources in complex ecosystems. Knowledge discovery running on a federated network encompass a hybrid-technology to lay out the foundation of an open- and cooperative-science ecosystem for computing discovery in the face of global sustainability challenges. The project summarized here is not only set out to deliver knowledge discovery computation in federated networks, but also to provide the architecture of a science-enabled technology, as a proof-of-principle, to connect knowledge-inspired societies to global sustainability challenges.

Knowledge discovery in eco-evolutionary biology-inspired federated networks ROBHOOT

1 Excellence

1.1 Radical vision of a science-enabled technology

Rapid, real time, data heterogeneity- and cooperation-based, discovery computation is currently a major issue revolving around data-driven intelligent machines and knowledge inspired societies. Several of these properties are found in evolving networks being these changes occurring in dynamic connectivity patterns and/or traits in neurons, and populations in natural ecosystems. However, evolving networks are not used for discovery computing yet, despite rapid trait evolution has been observed in experimental and theoretical systems [15, 17]. For example, evolving networks are characterized by feedbacks between the ecology and evolution of interacting traits, the eco-evolutionary feedbacks, to produce novel trait changes with new functional properties in ecosystems. This results in new computational properties, like novel interaction types (i.e., cooperation, competition, antagonism, etc), morphologies and/or evolving learning capabilities among agents to add discovery computing properties to the network. Conventional Artificial Intelligence (AI) computation is rapidly evolving towards explainable and discovery pattern inference [22] but often avoids evolutionary changes for exploring new computing capabilities [27]. The same situation occurs for artificial neural networks that also make only limited used of novel computing capabilities as a consequence of evolutionary changes in interactions and traits [29]. However, the rapid novel properties of evolving connections and traits, the diversifying power of biological systems, that make evolutionary-biology inspired networks highly plastic and resilient, have not yet been exploited in discovery computation. **The goal of this project is to implement eco-evolutionary-biology inspired solutions to make discovery computation a cooperative game of rapidly evolving traits and interactions**. The exploitation of evolving connections and traits will allow us to create novel types of discovery computation solutions for natural ecosystems facing sustainability challenges like overexploitation of the Seas, where harvesting renewable resources are in the point of diminishing returns for many species, communities and ecosystems (refs +++).

Why should we go deeper into evolving and diversifying information processing systems for discovery computation? With connections and traits (i.e., nodes and links in networks) represented in a spatially distributed network, as found in natural ecosystems, it is possible to untangle mapping of many spatiotemporal inputs onto many output functions considering learning among the interacting and heterogeneous traits and agents to decipher new solutions for harvesting renewable resources. This allows representing real-time solutions for spatiotemporal ecosystems with renewable resources, which is a key problem in many digital and natural ecosystems.

To show the capabilities of the ROBHOOT approach, we will complement the novel implementations of evolutionary biology-AI discovery computation with full cycle reproducibility, automation and visualization to trigger its properties at large-scale (Figure 1). **The main impact of ROBHOOT is that we provide novel discovery computation solutions to substantially improve ecosystem sustainability especially relevant for community-rich digital and natural ecosystems**. To support this notion, we will perform eco-evolutionary biology-AI network inspired simulations of multiple data-heterogeneity based networks. The central goals of $\mathcal{ROBHOOT}$ are:

1. To extend existing theories of eco-evolutionary biology-AI inspired in networks to obtain understanding of the factors and their interactions underlying discovery computation in cooperative federated networks. This will allow us to identify novel paths of reliable solutions for ecosystem sustainability.
2. To investigate how spatiotemporal evolutionary biology-AI-inspired networks can mimic the empirical patterns of natural and socio-technological ecosystems when large and heterogeneous exploiting groups and species coexist.
3. To develop fast, reproducible and automated discovery eco-evolutionary biology-inspired computation prototypes for real-time information processing tasks.

4. To arrive at powerful discovery computing principles for cooperative forecasting in federated networks, models of evolutionary neural biology-inspired to investigate cooperative forecasting for ecosystem sustainability when changes in learning, interactions and traits occur in a large and diverse pool of species, technologies and human groups.

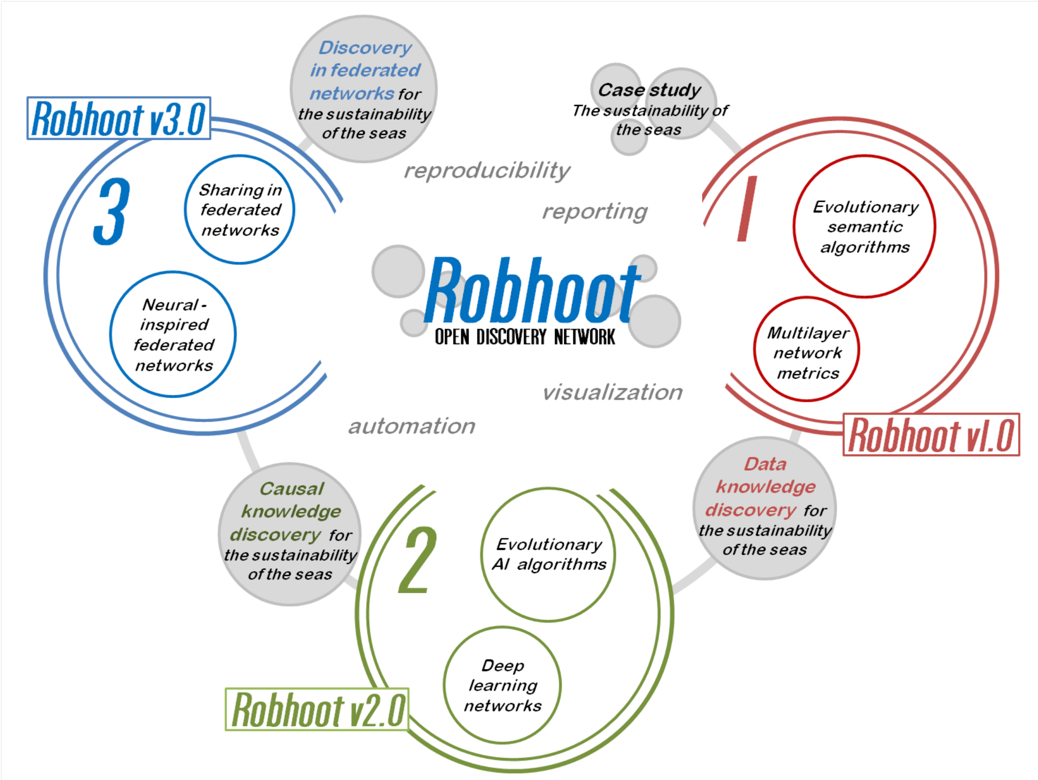

Figure 1: **Discovery in Evolutionary Biology-Inspired Federated Networks**. ROBHOOT targets knowledge discovery in federated networks. Federations are composed by highly heterogeneous groups sharing ecosystem resources for a sustainable knowledge-inspired society: ROBHOOT introduces three science-enabled technologies: Evolutionary biology-inspired semantic algorithms for ROBHOOT v1.0: data knowledge discovery (red), evolutionary-biology inspired AI-deep neural networks for ROBHOOT v2.0: causal knowledge discovery (green), and evolutionary neural biology-inspired ROBHOOT v3.0: discovery in federated networks (blue).

1.2 Science-to-technology breakthrough that addresses this vision

Data knowledge discovery (WP1)

Evolutionary biology-inspired semantic algorithms (WP1): The majority of studies of data discovery focus on add-hoc algorithms, ignoring ecological and/or evolutionary biology-inspired solutions. Currently, only a few databases are semantically annotated (e.g. , gene ontology database, COVID-19). This is because ontology development is time-consuming, requires expert knowledge and community commitment, and is ideally paired with data-driven research that iteratively checks the soundness of the ontology as it simultaneously seeks discovery. Thus, software tools for mapping and linking the terms between different ontologies are still to be developed, although Semantic Web technologies are included in programs such as the U.S. National Science Foundation's proposed CyberInfrastructure (refs +++). **Going beyond** ROBHOOT will go beyond state-of-the-art to implement evolutionary computation concepts (i.e., genetic algorithms with different rules and selection modes) to investigate new data properties along many data-sources. ROBHOOT will provide a detailed understanding of the replicability of accounting for many data-sources on the global data architecture map contrasting different evolutionary algorithms.

Causal knowledge dicovery (WP2)

Eco-evolutionary biology-AI-inspired discovery computation algorithms (WP2): Many studies of causal discovery (i.e., explainable or interpretable discovery), focus on genetic programming (refs

+++) and symbolic regression (refs +++), ignoring eco-evolutionary biology-inspired computation from where causal inference can be obtained. On the other side, experimental evolution data shows the importance of rapid structural trait changes beyond plasticity for new functional information processing capabilities of the interactions and traits (refs +++). The classical view on biology-inspired information processing is to consider plasticity without structural changes, or without co-evolution among many interacting components (refs +++). Recent studies indicate that rapid trait changes and information processing as a consequence of these changes is far more complex (refs +++). For example, eco-evolutionary dynamics strongly affect feedbacks between ecological and evolutionary processes, which in turn influences trait changes to open new functional properties of populations with new information capabilities (i.e., new adaptations to new habitat or niche biotic or abiotic conditions, refs +++). Furthermore, recent studies suggest that the interplay between trait dimensionality (biotic, abiotic, migration traits, etc), and adaptation is key to understand the emergence of new traits and information processing abilities to elaborate new computation strategies in ecosystems (Box 1, and refs, +++).
Going beyond ROBHOOT will, for the first time, employ eco-evolutionary biology-inspired solutions to implement AI process-based methods to create spatiotemporal causal inference in systems containing large heterogeneity and dimensionality (Figure 2). Using the above models, this will be extended to deep process-based learning networks including trait and interactions as evolutionary changes and coevolution to implement diversification patterns in these systems. The search for causal knowledge discovery will be applied to the sustainability of the Seas case study containing 9 million entries, 1612 species (around 50 variables and traits per species), around 20 countries and 11 sampling methods (Figure 2). Our approach will explore broad classes of evolving functions from evolutionary biology-AI-inspired algorithms combining them to automated Bayesian machines ensuring the search, the evaluation of models, trading-off complexity, fitting to the data and quantify resource usage (Deliverable D2.3, [20, 30]).

Discovery in federated networks (WP3) Integrating data and causal knowledge graphs provide a mechanistic understanding of how the balance of cooperation vs. competition might alter sustainability in our exploration of the Seas case study. However, causal knowledge graphs are not enough if they only represent isolated contributions and can not “learn to learn” to find novel, emergent solutions in neural biology-inspired networks composed by highly heterogeneous groups. In this regard, 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]. Technologies in digital ecosystems around federated networks are scarce and mostly focus on decentralization, scalability and security fronts [8, 9, 13, 14, 19, 25]. In the science ecosystem, only a few applications of open decentralized technologies exist [21]. Yet, the discovery of novel algorithms in biology-inspired federated networks for cooperative forecasting of global sustainability problems when heterogeneous groups learn and share from each other is currently not in place. Recent studies have shown the importance of evolutionary search of mathematical and symbolic operations as building blocks to discover ML algorithms ([20, 27]). Evolutionary biology-inspired search for algorithmic discovery can help to decipher how interactions among heterogeneous groups evolve and learn to solve complex sustainability problems. For example, evolutionary dynamics can explore open-ended language of models with varying trait evolution functions to discover biologically inspired solutions in multidimensional systems ([27],+++). ROBHOOT v.3.0 deploys biology-inspired federated networks accounting for heterogeneous agents to discover novel biology-inspired solutions for the exploration of the Seas federated network (Deliverables D3.1 and D3.2, Tables 3.1a-c)
Going beyond: Our understanding of the outcomes from evolved information processing systems formed by highly heterogeneous groups, a kind of large-scale meta-learning in the federated setting [13], is currently quite limited. Therefore, new science-enabled approaches accounting for information processing with diversification of heterogeneous and highly dimensional systems in federated networks are required to develop science-enabled technologies where heterogeneous agents with different interests find (non optimal) solutions that allow sustainable explotation of ecosystms. ROBHOOT v.3.0 con-

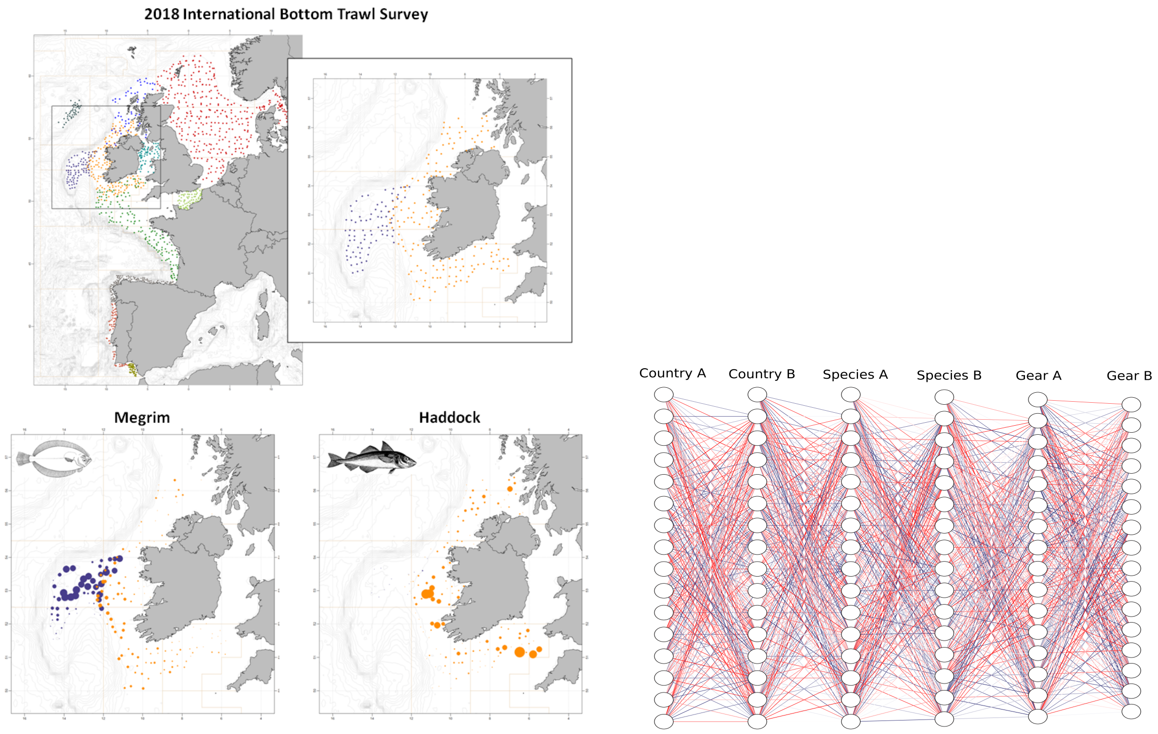

Figure 2: **Figure 2: Causal Knowledge Discovery for Sustainable Ecosystems**. **Top left**) The Irish Ground Fish Survey (IE-IGFS, Orange) and the Spanish Survey on the Porcupine Bank (SP-PORC, Blue) were part of the 2018 International Bottom Trawl Survey, coordinated by the International Council for the Exploration of the Sea [7]. Ireland and Spain use different Gears: The GOV gear has a larger vertical opening (Ireland, 3-4 m) respect to the Baka used on the Porcupine Bank (Spain, 2-3 m). This makes catchability different for fish species, such as Megrim (*Lepidorhombus whiffiagonis*, **Center left**) and Haddock (*Melanogrammus aeglefinus*, **Center right**), in which both countries have very different commercial interests. Haddock is a species of the cod family, highly prized in northern Europe, while Megrim is a species of flatfish, consumed largely in Spain and France. Spain catches Megrim better than Haddock and viceversa for Ireland. This generates a strong bias in the distribution maps (compare Megrim vs. Haddock map, Center) with potential implications for biodiversity management and sustainability in natural ecosystems. **Right** Causal knowledge discovery graph representing the 2-countries, 2-species and 2 gears for the example above. The whole data set for 2018 contains 11 countries, 461 fish species (approx. 200k individuals sampled), and 5 gears. Each country, species and gear is composed by many nodes: For example 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.

nects knowledge discovery to biology-inspired federated netwoks to study the properties of cooperative forecasting and strong inference in the face of global sustainability and biodiversity challenges (Figure 2 and Tables 3.1.a-c).

1.3 Interdisciplinarity and non-incrementality of the research proposed

To succeed with ROBHOOT, it is essential to build an interdisciplinary team that includes scientists from different disciplines, including evolutionary biology, ecology, computational neuroscience, computer science, data science, complex systems and experts in biodiversity sampling methods and the infrastructure related to international protocols for sampling the Seas. Data knowledge discovery gained by analysis and modelling of the computation discovery capabilities of evolutionary-inspired semantic algorithms by the evolutionary biology, computer science and complex system members of this consortium (EBD-CSIC, IFISC-CSIC, SDSC) can be transferred to the causal domain addressed by the other part of the consortium with expertise in evolutionary biology, data science and causal inference

(EAWAG and TARTUR). This will be enriched with full automation, reproducibility and visualization supported by ICREA, SDSC, and our company-partner (SME), respectively. Conversely, those 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 traits and interactions, cooperation, learning functions, and dimensionality) encountered in their implementations to explore to what degree this is reflected also in eco-evolutionary biology-inspired and neurobiology inspired discovery computation models to augmented their models. This cross-fertilizing back-and-forth interaction will allow the project to keep high modularity within the work packages while keeping functional interactions among the groups to run efficiently the different stages of the project. To bring together adaptive biology-inspired semantic algorithms for data discovery and evolutionary-neurobiology-inspired discovery in federated networks requires 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 leap opening a new direction for eco-evolutionary biology-inspired discovery computation.

1.4 High risk, plausibility and flexibility of the research approach

ROBHOOT represents a novel approach for complex, adaptive and multidimensional discovery computation. The transfer of eco-evolutionary biology-inspired principles onto fully reproducible and automated software, progressing from fragmented- and pattern-based to integrated- and 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 involved areas, mitigating risk in a gradual way, following a strict line and gradually increasing in complexity of the problems addressed. Figure 3 shows that we will start with evolutionary biology-inspired semantic algorithms for data discovery in the context of the Sustainability of the Seas case study. This is followed by investigation and implementation of more complex eco-evolutionary biology-inspired AI and deep learning network modeling to infer causality in the sustainability of the Seas case study (i.e., dimensionality, nonlinearities). Then we will advance to more complex situations, where the evolutionary neurobiology-inspired modeling will expand the search along many learning and cooperative forecasting schemes to find scenarios for the sustainability of the Seas case study. To keep the project technically feasible, and to be able to identify the mechanisms and their 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 theoretical work and numerical simulations with a real empirical case for the Sustainability of the seas. The knowledge gained along these three lines will allow us to compactly represent all the steps into a unified science-enabled technology. This leaves open the option to work with fast computing languages to develop low-level Agent Based Models along all the theoretical development of the proposal (i.e., Julia, C++), instead of differential/difference equations methods when a large number of agents and interactions change in time and space. This feature represents a very desirable fallback in case of speed and convergence problems for multidimensional and nonlinear systems (Table 1.41 Critical risks for implementation). Our implementation activities are all complemented by numerical investigations contrasted for speed and robustness with the sustainability of the Seas case study started in 1965 and containing around 9 million entries, 1612 species, 20 countries and 11 sampling methods (Figure 2). The success of ROBHOOT would represent a breakthrough in the current discovery computation with direct application to sustainability of ecosystems. It exploits eco-evolutionary biology-inspired computational capabilities of evolving traits and interactions to discovery and transfers their properties to natural ecosystems. The combination of rapid, data heterogeneity and cooperation for discovery computation based on, mostly open-source languages, will lead to fast implementations of the demonstrators with high flexibility that will permit a rapid transit to the public.

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	Consider more developed genetic programming methods to infer data interactions.
Multilayer metrics accounting for spatiotemporal patterns along many datasets insufficiently developed: Low	2	WP1	Implementation of more standard complex networks metrics to characterize data knowledge discovery.
Low number of training data available: Medium	2,3	P2	Alternative methods focusing on matrix decomposition methods.
Automated evolutionary-inspired expressions for causal knowledge discovery insufficiently developed: Medium	2,3	WP2	Symbolic regression methods to full automation 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 neurobiology-inspired federated networks insufficiently developed: Medium	1-4	WP3	Spiking neural network models as alternatives to evolutionary neural biology-inspired algorithms in federated networks.
Cooperative forecasting mixing eco-evolutionary dynamics and neural nets in large scale federated networks insufficiently developed: Medium	1-4	WP3	Mix eco-evolutionary dynamics models with less alternative neural nets models working a smaller spatiotemporal scales.

2 Impact

2.1 Expected impact

- **Scientific and technological contribution to the foundation of a new future technology**: ROBHOOT targets novel approaches towards sustainable ecosystems. One of the tasks in WP3 focus on the discovery of novel evolutionary neurobiology-inspired algorithms to provide results for sustainability fisheries. Solutions around WP3 ultimately depend on merging WP3 with the rest of WP's in the proposal. For example, it is known that sustainable ecosystems strongly depend on many data sources collected by different groups using different technologies (refs +++). ROBHOOT discover data interactions combining fisheries, stakeholders, and technology data, the data knowledge discovery graph, as a first step towards the discovery process. ROBHOOT also infer the technological and environmental changes and the processes underlying the empirical patterns, the causal knowledge discovery, to provide the existing sustainability status in a human-disturbed ecosystem. Altogether, this project will lay the foundation for future sustainability studies. Discovery of novel evolutionary-inspired algorithms for biodiversity maintenance have been hardly been investigated in this context so far. Therefore, several predictors related to biodiversity, technological and social times series analysis will be tested and further developed to enable robust prediction of sustainability. The discovery of new solutions not observed in the empirical data, but containing plausible scenarios for maintaining species-rich and sustainable ecosystems, will be the basis for estimation of the severity of overfishing and sampling bias when many groups enter in commercial conflict of interest... Such a targeted sustainability proxies would be of great interest not only for the biodiversity maintenance but also from an economic and social point of view, as it would save costs for future generations. Sustainability challenges are related to the development of future sustainable societies, which according to (Organization...) **Keep elaborating**

- **Potential for future social or economic impact or market creation**: Collapse of ecosystems can lead to serious long term economic and ecological disfunctionalities (refs +++). However, there are not well established metrics for the characterization of sustainability in complex ecosystems. Our approach accounts for heterogeneous sources of data, the (evolving) mechanisms underlying technological, environmental and social changes required to make ecosystems sustainable and novel rules that could impact positively the maintenance of biodiversity by developing cooperative forecasting strategies among the many (international) groups involved. Such a risk assessment would not only be of great interest to the groups exploiting the resources, but also from an economic and ecological point of view, as having less bias in the field data provides more accurate measures from the observed time series for planning fish stocks for a large number of species. Finally, ROBHOOT contributes towards knowledge-inspired societies in need of radically tackling new societal and global environmental challenges: it provides reproducible and transparent methods for making sustainability goals achievable and reproducible across many sectors and economies.

  In the medium-term this technology may also have interesting applications in public and private industry. For example, access to discovery with cooperative forecasting might suggest new paths and solutions that are key to generate rapid and robust scenarios when facing complex problems including global sustainability challenges (i.e., global health, ecosystems degradation, biodiversity loss, etc). First, evolutionary biology-inspired AI algorithms deciphering open-ended search of interpretable mechanisms underlying the targeted complex systems for private and public industry facing highly heterogeneous data sources. Second, cooperative forecasting challenges existing fragmented responses to emergent global sustainability problems by compactly offering reproducible forecasting emerging from many-to-many human and machine cooperative discovery, and third, open-access explainable and automated information generation account for global data-arquitecture allowing individuals and companies to address scenarios of future strategies in highly fluctuating local and global market conditions.

- **Impact on transparency and reproducibility**: Decision making and governance at local, regional and global scales require access to transparent and reproducible information containing the interpretable factors and their plausibility to explain the empirical patterns. In this regard, the ROBHOOT consortium brings together excellent partners from the fields of computer science, neurobiology, complex system, biology, social sciences, evolutionary ecology and including one SME focusing on reproducibility, automation, visualization and reporting along its whole developmental life cycle (Dissemination plan below and Figure 3). At the same time, all groups composing the consortium exhibit a long-standing experience interdisciplinary research across the boundaries of the individual disciplines (Figure 3). The subsection on related projects shows that this is a novel constellation in Europe and possibly worldwide (section 4). This consortium is also at the leading edge of developing novel evolutionary biology-AI inspired solutions to automation and reproducibility in complex systems facing sustainability challenges.

- **Ecosystem health impact**: Ecosystem sustainability and ecosystem health are usually used as metaphors to describe the mechanisms that maintain functional and diverse systems and the condition of an ecosystem, respectively. Ecosystem sustainability and condition can vary as a result of many disturbances like fire, flooding, drought, extinctions, invasive species, climate change, mining, overexploitation in fishing, farming or logging, chemical spills, and a host of other reasons. ROBHOOT focus on novel discovery solutions for ecosystems under a varying degree of disturbances. ROBHOOT introduces a case study for overexploited ocean ecosystems when highly heterogeneous social groups with different interests exploit limited and shared resources. Thus, ROBHOOT is a technology designed to provide novel discovery solutions paths for ecosystem sustainability, improving the underlying discovery paths that allow draw novel connections between ecosystem sustainability and ecosystem health. This feature aligns to the EU Reflection paper towards a Sustainable Europe by 2030 focusing on the need of investing in sustainable growth and spur action by goverments, institutions and citizens, leading the way for the rest of the world using the UN's Sustainable Development Goals (SDGs). Specifically, ROBHOOT can be seen as an horizontal enabler for the sustainability transition to make Europe sustainable by 2030. It introduces evolutionary biology- and artificial intelligence-inspired solutions to benefit ecosystem health and people's lives and work. By being able to process large amounts of heterogeneous data instantaneously, artificial intelligence and evolutionary-biology inspired solutions have the potential to significantly increase productivity in environmental sustainability and ultimately make informed decisions to enhance food security [1].
- **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 and in particular evolutionary biology-inspired federated networks both from a theoretical and an experimental point of view, Physics, theory and applications of complex systems in social networks and one highly innovative science-based reproducibility, automation, reporting and communication focusing on sustainability solutions. Many of the components of the consortium are first-time participants to FET under Horizon 2020 (Section 4). The use of advanced evolutionary biology-inspired and complex networks-based analyses to characterize and predict novel discovery in systems formed by heterogeneous and evolving groups and interactions combined with the implementation of intelligent learning discovery in federated networks and the development of a reproducible and automated protocol user friendly interface go much beyond the current state-of-the-art in science-based discovery technologies. All consortium partners exhibit a long-standing experience in interdisciplinary research across the boundaries of the individual disciplines (Figure 3). The subsection on related projects shows that this consortium is at the leading edge of innovation and interdisciplinarity (Tables 3.1a-c). A significant value proposition of the project is to increase the research on large-scale sustainable federated networks where many heterogeneous agents share resources embedded in complex ecosystems. This will produce valuable information and data about how federated networks work under broad set of socio-ecological scenarios, similar to natural ecosystems consoritiums where many paths produce coexistence of heterogneous populations and high biodiversity (refs ++). It is important to consider that all ecosystems facing many human pressures are all across the world and discovery technologies facilitating the solutions in large-scale federated networks could inspire new developments improving our understanding of sustainability at global scale. For in-home, we also expect an explosion of discovery knowledge approaches and future publications, which will place Europe at the top of sustainability in federated networks.

  Moreover, in WP3, we propose the generation of a web-based sustainability discovery portal that will allow researchers, NGO, managers and the public to train students in the discovery process to manage over-exploited ecosystems, allowing to scale up the number of people participating in the sustainability process by an order of magnitude thus mobilising forward thinking researchers and excellent young researchers to work together and explore what may become a new technology paradigm in sustainability research. Members of the consortium already have experience in generating such types of training tools that are currently available online (check github repository RobhooX). This approach would provide an unprecedented capability for the access to a multitude of people interested in sustainability discovery tools that will result in facilitating consensus and a valuable source of information for science-enabled technologies in ecosystem sustainability and management.

2.2 Measures to maximize impact

Dissemination

- **The Plan for diseminating and exploiting the project results**: ROBHOOT allocates three research groups along its whole developmental life cycle to guarantee dissemination, transparency and easy exploitation of the technology **(when)**. **(what)** The three milestones of the project, data knowledge discovery, causal knowledge discovery and discovery in federated networks (Table 3.2a) will be fully automated and reproducible to facilitate visualization, reporting and full scalability. **(who)** Automated discovery will be implemented along Bayesian machine scientist to facilitate open-ended search during the development of the three milestones (Tables 3.1a-b). Reproducible knowledge discovery graphs will be developed in the Renku open-source software (Swiss Data Science Center, SDSC). Visualization and reporting will be fully implemented in the Julia computing language for its speed and unique features (SME, Codes will be available in the public git Robhoot repository. Having the whole developmental life cycle as reproducible and automated knowledge discovery graphs facilitates the reuse and the dissemination of the technology as a whole in any platform and OS. Full reproducibility, automation, visualization and reporting provide to ROBHOOT legal and financial transparency and reproducibility in social governance a feature for easy replication of the discovery process by third parties, a property that can be used to facilitate reporting for governance public policy, NGO, society and thinktank in the face of local and global sustainability challenges. **why, how and which journals, conferences and with which preliminary results**.

- All the data, codes and outputs generated during ROBHOOT development will be open access stored in public git repositories. The project will collect data from many sources (i.e., fisheries, environmental and social data, technology data). generate data knowledge discovery graphs, causal knowledge graphs and the data and algorithms generated from the discovery in federated networks for the sustainability of the Seas case study **Keep elaborating**

Communication activities

- The full open-source developmental life cycle strategy of reproducibility, automation, and reporting generation of ROBHOOT targets the search of societal relevance and long-term economic impact of open and transparent science. Underlying to this strategy is to build support for future research and innovation funding, by ensuring uptake of results within the scientific community, and opening up potential business opportunities for novel products or services, and potentially contributing to better decision-making processes and valuable input for public policies formulation. ROBHOOT has very general dissemination targets, from scientists and decision-makers, to the business community and the public. ROBHOOT's general dissemination measures will focus on project results and stakeholder engagement (stakeholder consultation processes; workshops to raise awareness, etc.) through:
  . The project website is to be set up within the first three months of the project. There is already a public git Robhoot repository.
  . Up to date information material, e.g. brochures, presentationslides, will be distributed at events

to increase awareness about the ROBHOOT project.
. General other publication means will be used such as newspapers, YouTube, TV and radio, social networks as well as targeted mailing lists (e.g., evodir, AI-worldwide).
Scientific publications for the scientific community. We will target high-level journals with open access (i.e., Science, Nature Communication, etc.)
. The consortium will visit conferences in the related scientific fields and interdisciplinary conferences in order to interactively present and discuss our results with others researchers, groups and institutions. Among other activities, the consortium will organize special sessions at several conferences in different countries. Additionally, some targeted, specific dissemination actions will be considered: We will organize hackatons and robhacks activities to attract multipliers and developers from the open-source community to the community who engage in data processing and build hybrid evolutionary biology-inspired and AI algorithms. This will be achieved by a "traveling salesman" approach using personal visits and invitations to demonstrate how ROBHOOT works. At the end of the project we will organize a workshop specifically on "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.
. ROBHOOT will launch a testnet to help disseminate the main results of discovery in federated networks (Section 3.1.3). The launch will have invited NGO's and GO across disciplines and social, economical and technological sectors. The ROBHOOT Open Discovery Network will be launched as a Biodiversity and sustainability open discovery network to offer the solutions for the exploration of the Seas case study and to integrate additional public databases and data collections into the open discovery network to facilitate NGOs, GOs and other organizations transparency, reproducibility, and governance in Biodiversity management.

- ROBHOOT strictly adheres to the Open Access Policy of the Commission and all publishable (non-protected) results will follow the green or gold OA policy. Software as well as hardware protocols will be made openly available through standard computer science repositories. The ROBHOOT public git repository is already active Robhoot. Data (measured data), as such, will not be acquired by ROBHOOT. 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. Open access to publications will be granted under the terms and conditions laid down in the Grant Agreement, in accordance with the Rules for participation and dissemination in Horizon 2020. The beneficiaries will deposit an electronic copy of the published version or the final manuscript accepted for publication of a scientific publication relating to foreground in an institutional or subject-based repository at the moment of publication, e.g., via the OpenAIRE portal (www.OpenAIRE.eu). In addition, beneficiaries will make their best efforts to ensure that this electronic copy becomes freely and electronically available to anyone through this repository (i.e., that it becomes "open access"): immediately, if the scientific publication is published "open access", i.e., if an electronic version is also available free of charge via the publisher, or within 6 months of publication.

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 evolutionary biology-AI-inspired models to infer causal knowledge discovery with an implementation for the exploration of the Seas case study, WP3 addresses evolutionary neural biology-inspired for knowledge discovery to provide cooperative forecasting in federated networks. WP3 also provides a empirical case implementation of cooperative forecasting for the exploration of the Seas.

Demonstrators: The project will create three demonstrators of increasing complexity all containing full reproducibility and automation capabilities:

- $\mathcal{RH}$v1.0 Software demonstrator with evolutionary semantic algorithms to decipher ontologies along many data-sources for the exploration of the Seas data knowledge discovery case study (MS1);

- $\mathcal{RH}$v2.0 Software demonstrator with evolutionary biology-AI-inspired modeling for spatiotemporal causal pattern knowledge discovery (MS2);

- $\mathcal{RH}$v3.0 Software demonstrator using evolutionary neural biology-inspired networks for spatiotemporal discovery in federate networks (MS3).

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

Work package No.	**Work package title**	**Lead No.**	**Lead Short Name**	**PMs**	**Start Month**	**End Month**
1	Data knowledge discovery	1	CSIC	XX	1	18
2	Causal knowledge discovery	6	TARTU ULIKOOL	XX	7	24
3	Discovery in federated networks	9	UNIGRAZ	XX	13	36
4	Dissemination	10	IEO	XX	1	36
5	Management	6	EAWAG	XX	1	36
			Total PMs	XXX		

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 an 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 $\mathcal{ROBHOOT}$, we will consider at least different scenarios for each of the software implementations such that agents contain many evolving traits and interactions can also evolve along different signs and effects (M1, M2 and M3). This allows, for example, finding trait and interaction changes patterns that improve sustainability scenarios with respect to the observed empirical patterns in the exploration of the Seas 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.

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

YEAR					2021				2022				2023			
MONTH					Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
MILESTONE												M1		M2		M3
WP	WP Name	PROGRESS	START	END												
WP1	Data knowledge discovery (DKD)															
T1.1	Evolutionary semantic algorithms	17%	1/1/21	31/06/22						D1.1						
T1.2	Multilayer network metrics	34%	1/1/21	31/06/22						D1.2						
T1.3	Automation DKD	51%	1/4/22	1/8/22							D1.3					
T1.4	Reproducibility DKD	68%	1/8/22	1/13/22							D1.4					
T1.5	Visualization DKD	85%	1/2/21	1/4/21							D1.5					
T1.6	Data knowledge discovery EX	100%										D1.6				
WP2	Causal knowledge discovery (CKD)															
T2.1	Eco-Evolutionary AI Algorithms	17%	1/3/21	1/7/21								D2.1				
T2.2	Eco-Evolutionary Deep Learning	34%	1/5/21	1/10/21								D2.2				
T2.3	Automation CKD	51%	1/10/21	1/13/21									D2.3			
T2.4	Reproducibility CKD	68%											D2.4			
T2.5	Visualization CKD	85%	1/10/21	1/12/21									D2.5			
T2.6	Causal knowledge discovery EX	100%	1/10/21	1/13/21										D2.6		
WP3	Discovery in federated networks (DFN)															
T3.1	Sharing federated networks	17%	1/16/21	1/21/21										D3.1		
T3.2	Evolutionary neural networks	34%	1/22/21	1/26/21												D3.2
T3.3	Automation DFN	51%	1/27/21	2/1/21												D3.3
T3.4	Reproducibility DFN	68%	2/2/21	2/6/21												D3.4
T3.5	Visualization DFN	85%	1/27/21	1/31/21												D3.5
T3.6	Discovery in federated networks EX	100%														D3.6
WP4	Dissemination															
T4.1	Dissemination activities	17%	date	date				D4.1				D4.2				D4.5
T4.2	Analysis Exploitables	34%	date	date												
T4.3	Business plan	51%	date	date												D4.6
T4.4	Hackaton Robhoot 1.0	68%	date	date								D4.3				
T4.5	Data management	85%	date	date								D4.4				
T4.6	Exploration of the Seas outreach	100%														D4.7
WP5	Management															
T5.1	Project initiation	50%	date	date												
T5.2	Other management task (R = Reporting)	100%	date	date				R				R				R

Table 3.1b: **Work package description** Please note that the basic technical considerations for the setups are described in side boxes...

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	SME
Person month per participant	9	6	3	3	3
Start month	1				
End month	24				

Objectives

- Develop a evolutionary semantic framework for data discovery
- Derive semantic functionality rules required for data computation discovery
- Adaptive learning rules and metrics for the sustainability of the Seas data discovery

Description of work

Task T1.1: Develop evolutionary semantic algorithms to investigate data extraction and transformation in ontologies (M1-M18) *Leader: **EBD-CSIC**. Contributors: 1*

Evolutionary driven ontologies will be mapped to a graph-based data architecture flexible enough to get high scalability (i.e. Neo4j) and to infer multilayer metrics (T1.2). Currently software tools for mapping and linking the terms between different ontologies are still to be developed [5] and/or avoid adaptive algorithms for doing data extraction and transformation (refs ++, add as ref Semantic Web technologies are included in programs such as the U.S. National Science Foundation's proposed CyberInfrastructure). T1.1 provides adaptive algorithms to allow WP2 and WP3 to implement this feature in causal knowledge discovery and discovery in federated networks. Miguel: Keep elaborating

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 [10, 12] for evolutionary semantic algorithms will focus on large pools of data heterogeneity to explore how data configurations, privacy requirements, formats, dimensions, biases and spatiotemporal resolution affect data discovery [2–4, 6]. Victor, Emilio: Keep elaborating

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

Automation rules [20] for evolutionary semantic algorithms and multilayer network metrics search and rules transformation for data discovery. Roger: Keep elaborating

Task T1.4: Reproduce (M15-M21) *Leader: **SDSC**. Contributors:*

In this task the SDSC will merge the work done in T1.1 and T1.2 into reproducible and replicable data knowledge graphs

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

In this task the partner SME will apply visualization algorithms to the work done in T1.1 and T1.2 Charles and Miguel: Keep elaborating

Task T1.6: All participants apply results from evolutionary semantic algorithms and multilayer network metrics into a fully automated, reproducible and animated sustainability exploitation of the Seas case study (M15-M24) *Leader: **EBD-CSIC**. Contributors: 1,2,3,4,5*

Evolutionary semantic algorithms and multilayer network metrics will search and transform many source-data (i.e., Fishery data using the (global fishing watch), species interactions, environemental data and social and stakeholders groups with different interests within each of the countries, etc, together with the sustainability of the Seas database started in 1965 contains around 9 million entries, 1612 species, 20 countries and 11 sampling methods (Figure 2).

Deliverables

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

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	9 (Provisional)	6 (Provisional)	
Start month	7		
End month	30		

Objectives

- Develop a evolutionary-biology-AI inspired framework for causality discovery
- Derive functionality rules required for causality-based computation discovery
- Adaptive learning rules to mimic the empirical patterns for sustainability of the Seas

Description of work

Task T2.1: Develop eco-evolutionary dynamics modeling ... (M7-M24) *Leader: **EAWAG**. Contributors: 6*

... T2.1 provides computation algorithms with evolving traits and interactions to allow WP2 to implement this feature in causal knowledge discovery. This is particularly relevant in Earth, Ecosystem and Sustainability science. The rapid progress of AI as an automated and explainable technology ([11, 16, 18, 20, 26, 27],+++) will increase our ability to make stronger inferences about future sustainability challenges and solutions [28]. Yet, eco-evolutionary biology-AI-inspired computation discovery solutions will be required to explore a broader range of scenarios with changing functions and .... Carlos:Keep elaborating

Task T2.2: This task extends T2.1 into evolutionary biology-inspired deep learning networks metrics for general principles of causal discovery (M7-M24) *Leader: **TARTU**. Contributors: 8*

Raul:Keep elaborating

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

Automation rules [20] for evolutionary semantic algorithms and multilayer network metrics search and rules transformation for data discovery.Roger:Keep elaborating

Task T2.4: Reproduce (M21-M27) *Leader: **SDSC**. Contributors: 4*

In this task the SDSC will merge the work done in T2.1 and T2.2 into reproducible data knowledge graphs Christine:Keep elaborating

Task T2.5: Visualize (M21-M27) *Leader: **SME**. Contributors: 5*

In this task the partner SME will apply visualization algorithms to the work done in T2.1 and T2.2 Charles and Miguel:Keep elaborating

Task T2.6: All participants apply results from eco-evolutionary AI algorithms and deep learning networks into a fully automated, reproducible and animated sustainability of the Seas case study (M21-M30) *Leader: **EAWAG**. Contributors: 6,7,8,3,4,5*

0.05 in

Deliverables

D2.1 Report on definition of eco-evolutionary biology-AI-inspired rules for causal discovery **(M18)**
D2.2 Report on definition of eco-evolutionary process-based deep learning networks applied to causal computation discovery **(M18)**
D2.3 Automated demonstrator of eco-evolutionary biology-AI-inspired rules for causal discovery **(M21)**
D2.4 Reproducible demonstrator of eco-evolutionary biology-AI-inspired rules for causal discovery **(M21)**
D2.5 Visualization demonstrator of evolutionary semantic rules for data discovery **(M21)**
D2.6 Demonstrator all parts for the sustainability exploitation of the Seas case study **(M24)**

Work package number	**3**	**Lead beneficiary**	**UNIGRAZ**
Work package title	**Discovery in federated networks**		
Participant number	8	9	
Short name of participant	SRC	UNIGRAZ	
Person month per participant	X	X	
Start month	13		
End month	36		

Objectives

- Develop a evolutionary-biology inspired framework for discovery in federated networks
- Derive functionality rules required for computation discovery in federated networks
- Adaptive learning rules to discover novel paths for sustainability of the Seas

Description of work

Task T3.1: Develop eco-evolutionary biology-inspired modeling for discovery in federated networks (M13-M36) *Leader:* ***SRC****. Contributors: 10*

This task extends eco-evolutionary biology-inspired modeling for general principles of discovery in federated networks... Jon:Keep elaborating

Task T3.2: Develop evolutionary neurobiology-inspired algorithms... (M13-M36) *Leader:* ***UNIGRAZ****. Contributors: 9*

... T3.2 provides computation algorithms with evolving neurons with (many) traits and interactions to allow WP3 to implement this feature in discovery in federated networks.... Wolfgang:Keep elaborating

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

Automation rules for eco-evolutionary and neurobiology-inspired modeling for discovery in federated networks Roger:Keep elaborating

Task T3.4: Reproduce (M21-M27) *Leader:* ***SDSC****. Contributors: 4*

In this task the SDSC will merge the work done in T3.1 and T3.2 into reproducible and replicable discovery in federated networksChristine:Keep elaborating

Task T3.5: Visualize (M21-M27) *Leader:* ***SME****. Contributors: 5*

In this task the partner SME will apply visualization algorithms to the work done in T3.1 and T3.2Charles:Keep elaborating

Task T3.6: Sustainability of the Seas federated network (M21-M30) *Leader:* ***UNIGRAZ****. Contributors: 6,7,8,9,10*

All participants apply results from eco-evolutionary and neurobiology-inspired algorithms into a fully automated, reproducible and animated sustainability of the Seas federated network case study

Deliverables

D3.1 Demonstrator on eco-evolutionary biology-inspired rules for discovery in federated networks **(M30)**
D3.2 Demonstrator on evolutionary neurobiology-inspired rules for discovery in federated networks **(M36)**
D3.3 Automated demonstrator of for evolutionary biology-inspired rules in federated networks **(M36)**
D3.4 Reproducible demonstrator of evolutionary rules in federated networks **(M36)**
D3.5 Visualization demonstrator of evolutionary rules for discovery in federated networks **(M36)**
D3.6 Demonstrator all parts for the sustainability of the Seas federated network case study **(M36)**

Work package number	4	**Lead beneficiary**	IEO
Work package title	**Dissemination**		
Participant number	10	11	
Short name of participant	IEO	SEM	
Person month per participant	X	X	
Start month	1		
End month	36		

Objectives

- This WP deals with the system entire scope of dissemination of results in the research community and for the general public. Connection to SME for visualization

Description of work

Task T4.1: Paco:Keep elaborating... (M7-M24) *Leader:* ***IEO****. Contributors: 10*

Task T4.2: Paco:Keep elaborating (M7-M24) *Leader:* ***IEO****. Contributors: 10*

Task T4.3: Paco:Keep elaborating (M21-M27) *Leader: **IEO**. Contributors: 10*

Task T4.4: Miguel:Keep elaborating (M21-M27) *Leader: **SME**. Contributors: 11*

Task T4.5: Miguel:Keep elaborating (M21-M27) *Leader: **SME**. Contributors: 11*

Task T4.6: Miguel:Keep elaborating (M21-M30) *Leader: **SME**. Contributors: 11*

Deliverables
D4.1 (M18)
D4.2 (M18)
D4.3 (M21)
D4.4 (M21)
D4.5 (M21)
D4.6 (M24)

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	6	6	
Start month	7		
End month	30		

Objectives
- Management and work process 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.
- Co-ordinate the technological and scientific orientation of the project.
- Secure the quality of the work and of the delivered documents and softwar

Description of work

Task T5.1: Carlos:Keep elaborating (M1-M36) *Leader: **EAWAG**. Contributors: 8*

Task T5.2: Victor:Keep elaborating (M36-M27) *Leader: **IFISC-CSIC**. Contributors: 1*

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)**
D1.1	Software evolutionary semantic rules 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 eco-evolutionary biology-AI-inspired rules for causal discovery	WP2	EAWAG	R	PU	18

Continued on next page

D2.2	Report on definition of eco-evolutionary process-based deep learning networks applied to causal computation discovery	WP2	TARTU	R	PU	18
D4.1		WP4	IEO	R	PU	18
D4.2		WP4	IEO	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 eco-evolutionary biology-AI-inspired rules for causal discovery	WP2	ICREA	D	PU	21
D2.4	Reproducible demonstrator of eco-evolutionary biology-AI-inspired rules for causal discovery	WP2	SDSC	R	PU	21
D2.5	Visualization demonstrator of evolutionary semantic rules for data discovery	WP2	SME	R	PU	21
D4.3		WP4	SME	D	PU	21
D4.4		WP4	SME	R	PU	21
D4.5		WP4	SME	R	PU	21
D1.6	Demonstrator all parts for the sustainability exploitation of the Seas case study	WP1	EBD-CSIC	R	PU	24
D2.6	Demonstrator all parts for the sustainability exploitation of the Seas case study	WP2	EAWAG	R	PU	24
D4.6		WP4	SME	R	PU	24
D3.1	Demonstrator on eco-evolutionary biology-inspired rules for discovery in federated networks	WP3	SRC	D	PU	30
D3.2	Demonstrator on evolutionary neurobiology-inspired rules for discovery in federated networks	WP3	UNIGRAZ	D	PU	36
D3.3	Automated demonstrator of for evolutionary biology-inspired rules in federated networks	WP3	ICREA	D	PU	36
D3.4	Reproducible demonstrator of evolutionary rules in federated networks	WP3	SDSC	D	PU	36
D3.5	Visualization demonstrator of evolutionary rules for discovery in federated networks	WP3	SME	D	PU	36
D3.6	Demonstrator all parts for the sustainability of the Seas federated network case study	WP3	UNIGRAZ	D	PU	36

Table 3.2a: List of milestones

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) and an external Scientific Advisory Committee (SAC). The SB, which will consist of one representative from

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

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, dissemination aspects of the project, and risk analysis. 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 netowrks, 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: 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 meeting 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**

Table 3.2b: Critical risks for implementation

(TO BE DONE)

3.3 Consortium as a whole

ROBHOOT is a science-enabled multi-feature technology designed with a highly modular structure. Modularity allows to gain module functionality while maintaining cross-functional features among the different parts to produce a science-enabled interdisciplinary technology (Figure 1, WP one to three and milestones one to three, red, green and blue, respectively): Data knowledge discovery's team requires skills in evolutionary biology, evolutionary computation, computer science and the physics of complex systems (Section 3.1.1, Table 3.2a). ROBHOOT v.1.0 work mixes expertise in semantic algorithms, evolutionary computation algorithms and multilayer network metrics to create novel evolutionary-biology inspired ontology annotations along heterogeneous data-sources into one data knowledge discovery. EBD-CSIC team takes care of data knowledge graphs introducing novel evolutionary semantic algorithms to decipher ontologies and interactions among many data-sources (D1.1,

Description of risk	WP	Proposed risk- mitigation measures
Evolutionary semantic algorithms insufficiently developed: Medium	WP1	Consider more developed genetic programming methods to infer data interactions.
Multilayer metrics accounting for spatiotemporal patterns along many datasets insufficiently developed: Low	WP1	Implementation of more standard complex networks metrics to characterize data knowledge discovery.
Low number of training data available: Medium	WP2	Alternative methods focusing on matrix decomposition methods.
Automated evolutionary-inspired expressions for causal knowledge discovery insufficiently developed: Medium	WP2	Symbolic regression methods to full automation for causal discovery accounting for evolutionary rules.
Eco-evolutionary dynamics of multiple traits in species-rich ecosystems insufficiently developed: Medium	WP3	Mean-field approximations using classical ODE systems and novel universal differential equations for scientific machine learning.
Evolutionary neurobiology-inspired federated networks insufficiently developed: Medium	WP3	Spiking neural network models as alternatives to evolutionary neural biology-inspired algorithms in federated networks.
Cooperative forecasting mixing eco-evolutionary dynamics and neural nets in large scale federated networks insufficiently developed: Medium	WP3	Mix eco-evolutionary dynamics models with less alternative neural nets models working a smaller spatiotemporal scales.

Tables 3.1a-c). IFISC-CSIC team focuses on multilayer network modularity, community detection and decentralization metrics for pattern detection in data knowledge discovery (D1.2, Tables 3.1a-c). All teams in WP1 will join efforts to merge evolutionary semantic algorithms, multilayer network metrics, automation, reproducibility and visualization to produce the data knowledge discovery graph for the sustainability of the Seas case study (D1.6, Tables 3.1a-c). ROBHOOT v.2.0's team composed by EAWAG, and TARTU ULIKOOL and will merge eco-evolutionary biology-inspired networks to deep learning networks, the "Evolutionary biology-inspired AI algorithms" approach (D2.1 and D2.2, Box 1, Table 3.1a-c and Figure 1, green). The overall goal of this milestone is to connect evolutionary biology mechanisms to deep learning networks to generate a causal knowledge discovery technology to make patterns interpretable (Deliverable D2.2, Section 3.1.2, Table 3.2.a-c and Figure 3). The team for this milestone add inter-module complementarity expertise to ROBHOOT v.1.0's team: Now the skills focus on data-scientists trained in deep learning networks and evolutionary biologists with expertise in evolutionary ecology theory and evolutionary-inspired networks (section 3.1.2 and Figure 1, green). Milestone two generates a causal knowledge discovery for the sustainability of the Seas containing 9 million entries, 1612 species using around 11 sampling methods and more than 15 countries (D2.6, Figures 1, green). Interdisciplinarity in ROBHOOT is achieved not only at the intra-module development stage, but also at the inter-module stage where causal knowledge discovery and evolutionary biology-inspired AI algorithms might form the basis for the interdisciplinarity breakthrough ideas reflected in the highly complementarity skills of the consortium. The first two modules in ROBHOOT contain researchers from Estonia, Spain and Switzerland.

The ROBHOOT consortium wants to advance the rapidly evolving digital ecosystem by making cooperative discovery a fundamental feature of it. For this purpose, a science-enabled data and causal knowledge discovery technology is not enough if they stay isolated from a discovery technology embedded in large-scale networks. To discover novel scenarios for ecosystem sustainability, Discovery in federated networks should learn to learn from heterogeneous data-sources in the context of evolutionary neural biology-inspired algorithms. To achieve scalability for the discovery in federated networks, eco-evolutionary dyamics and neural-inspired protocols in federated networks is the excellency feature of ROBHOOT v.3.0 (section 3.1.3). ROBHOOT v.3.0's team composed by SRC and UNIGRAZ, develop eco-evolutionary dynamics scenarios for ecosystem sustainability and neural biology-inspired federated networks, respectively (Box 2). The team forming ROBHOOT v.3.0 also requires contrasting skills: First, theoreticians working in eco-evolutionary dynamics guarantee scalable implementation of evolutionary processes in federated networks. Second, neurobiologists in collaboration to developers aiming to explore the role of evolving neural biology-inspired solutions accounting for heterogeneity and dimensionality in federated networks. ROBHOOT v.3.0 is a fundamental stepping-stone for developing "Cooperative Forecasting": it first guarantees proper eco-evolutionary dynamics along species-rich ecosystems is implemented. Then these species-rich ecosystems represent the basis for discovery of novel paths that increase sustainability goals. And these novel paths are searched along many nodes of a network replicating eco-evolutionary dynamics scenarios that interact and learn from each other to find better forecasting scenarios at a global scale. ROBHOOT v.3.0's implements heterogeneous groups of cooperating and competing neurons in federated networks for making cooperative forecasting a standard global property of ROBHOOT (Deliverable D3.2, Tables 3.1a-c). Milestone three generates discovery in federated networks for the sustainability of the Seas to provide populations of scenarios satisfying biodiversity and sustainability maintenance while guaranteeing commercial interest of many interacting groups and stakeholders within and among countries (Deliverable D3.6, Figure 3, blue). ROBHOOT v.3.0 contain researchers from Sweden and Austria. ROBHOOT architecture aims to guarantee strong reproducibility, automation, and visualization-communication along its whole life cycle and development. The team formed by the SDSC (D1.4, D2.4 and 3.4), ICREA (D1.3, d2.3 and D3.3, and SME (D2.5, D3.5 and D4.5), will implement reproducibility, automation, and visualization and reporting, respectively, features crossing all ROBHOOT milestones to secure dissemination along its life cycle (Figure 1 and Gantt chart).

3.4 Resources to be committed

Bibliography

[1]

[2] *Blue-cloud.*

[3] *Elixir open data.*

[4] *HOT.*

[5] *Knowledge Graph COVID-19.*

[6] *Openstreet.*

[7] International bottom trawl survey working group (ibtswg). *ICES Scientific Reports*, 2019.

[8] Elli Androulaki, Artem Barger, Vita Bortnikov, Srinivasan Muralidharan, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart, Chet Murthy, Christopher Ferris, Gennady Laventman, Yacov Manevich, Binh Nguyen, Manish Sethi, Gari Singh, Keith Smith, Alessandro Sorniotti, Chrysoula Stathakopoulou, Marko Vukolić, Sharon Weed Cocco, and Jason Yellick. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains. *Proceedings of the 13th EuroSys Conference, EuroSys 2018*, 2018-Janua, 2018.

[9] BigchainDB GmbH. BigchainDB: The blockchain database. *BigchainDB. The blockchain database.*, (May):1–14, 2018.

[10] Emanuele Cozzo and Yamir Moreno. Dimensionality reduction and spectral properties of multilayer networks. 2018.

[11] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. pages 1–10, 2019.

[12] Manlio De Domenico, Albert Solé-Ribalta, Elisa Omodei, Sergio Gómez, and Alex Arenas. Ranking in interconnected multilayer networks reveals versatile nodes. *Nature Communications*, 6:1–6, 2015.

[13] Johnny Dilley, Andrew Poelstra, Jonathan Wilkins, Marta Piekarska, Ben Gorlick, and Mark Friedenbach. Strong Federations: An Interoperable Blockchain Solution to Centralized Third-Party Risks. 2016.

[14] Nikolai Durov. Telegram Open Network. pages 1–132, 2017.

[15] Fussmann G. F., Loreau M., and Abrams P. A. Eco-evolutionary dynamics of communities and ecosystems. *Functional Ecology*, 21:465–477, 2007.

[16] Giuseppe Futia and Antonio Vetrò. On the integration of knowledge graphs into deep learning models for a more comprehensible AI-Three challenges for future research. *Information (Switzerland)*, 11(2), 2020.

[17] Hairston N. G., Ellner S. P., Geber M. A., Yoshida T., and Fox J. A. Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*, 8:1114–1127, 2005.

[18] Yolanda Gil and Bart Selman. A 20-Year Community Roadmap for Artificial Intelligence Research in the US. aug 2019.

[19] Golem. The Golem Project Crowdfunding Whitepaper. *Golem.Network*, (November):1–28, 2016.

[20] Roger Guimerà, Ignasi Reichardt, Antoni Aguilar-Mogas, Francesco A. Massucci, Manuel Miranda, Jordi Pallarès, and Marta Sales-Pardo. A Bayesian machine scientist to aid in the solution of challenging scientific problems. *Science Advances*, 6(5):eaav6971, 2020.

[21] Vlad Günther and Alexandru Chirita. " Scienceroot " Whitepaper. 2018.

[22] Raban Iten, Tony Metger, Henrik Wilming, Lídia Del Rio, and Renato Renner. Discovering Physical Concepts with Neural Networks. *Physical Review Letters*, 124(1):1–18, 2020.

[23] Wolfgang Maass. Noise as a resource for computation and learning in networks of spiking neurons. *Proceedings of the IEEE*, 102(5):860–880, 2014.

[24] Wolfgang Maass. To Spike or Not to Spike: That Is the Question. *Proceedings of the IEEE*, 103(12):2219–2224, 2015.

[25] Ocean Protocol Foundation, BigchainDB GmbH, and DEX Pte. Ltd. Ocean Protocol: A Decentralized Substrate for AI Data & Services Technical Whitepaper. pages 1–51, 2018.

[26] Anthony O'Hare. Inference in High-Dimensional Parameter Space. *Journal of Computational Biology*, 22(11):997–1004, 2015.

[27] Esteban Real, Chen Liang, David R. So, and Quoc V. Le. AutoML-Zero: Evolving Machine Learning Algorithms From Scratch. 2020.

[28] Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and & Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature*.

[29] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, jan 2015.

[30] Christian Steinruecken, Emma Smith, David Janz, James Lloyd, and Zoubin Ghahramani. The Automatic Statistician. pages 161–173. 2019.

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 all management aspect.

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 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 andassociate 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 experienced in modelling 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 modelling, 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 funcitoning 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.

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).