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 and ecosystems. Evolving networks in nature with ever changing traits and connectivity patterns can inspire a new discovery computation for a global-sustainable knowledge-inspired society. Many studies have shown sustainability could be achieved by strengthening transparency, communication, and rapid access to discovery technologies. Sustainability goals, however, 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 discovery computation technology for a knowledge-inspired society. We introduce evolutionary biology-inspired and artificial intelligence solutions for sustainability in natural ecosystems. We validate our approach with a sustainability of the Seas case study in federated networks, where many distinct groups of species, humans and technologies coexist 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 computation 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 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 biology-inspired federated networks ROBHOOT

1 Excellence

1.1 Radical vision of a science-enabled technology

Ecosystems collapse around the globe in the absence of technologies to discover novel ways of sustainable exploitation of complex ecosystems. In this regard, rapid, real time, heterogeneous- and cooperation-based, discovery computation is currently a major issue revolving around data-driven intelligent machines and knowledge inspired societies facing global sustainability challenges. 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 computation yet, despite rapid changes in trait and interaction have been observed in experimental and theoretical systems [16, 18]. 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 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 [22] but often avoids evolutionary changes for exploring new computing capabilities [27]. The same situation occurs for artificial neural networks that also make limited use of novel computing capabilities as a consequence of evolutionary changes in interactions and traits [29]. **The goal of this project is to implement eco-evolutionary-biology 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 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 diversifying information processing for discovery computation? With connections and traits 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 evolutionary processes among many distinct traits and agents to decipher new solutions for harvesting renewable resources. This allows representing real-time solutions for spatiotemporal ecosystems with evolving renewable resources, which is a key problem in many digital and natural ecosystems.

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

1. To extend existing theories of eco-evolutionary biology-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.
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 computation principles for forecasting in federated networks when evolutionary changes in interactions and traits occur in a large and diverse 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: Most studies of data discovery focus on advanced analytics functions to reveal insights, ignoring the discovery of data-sources 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, 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 accounting for many data-sources are still to be developed [5, 7].

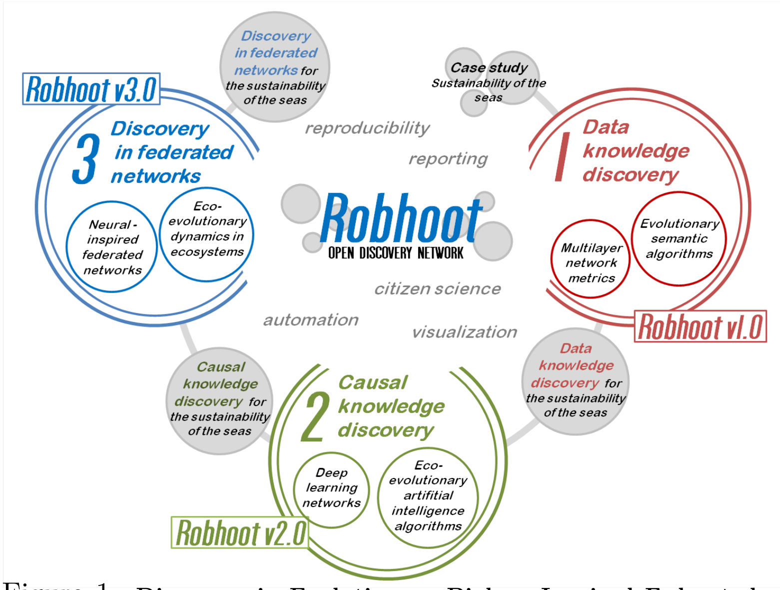

Figure 1: **Discovery in Evolutionary Biology-Inspired Federated Networks**. ROBHOOT target knowledge discovery when heterogeneous groups of species, humans and technologies share ecosystem resources for a sustainable knowledge-inspired society: It introduces three science-enabled technologies: Evolutionary biology-inspired semantic algorithms for ROBHOOT v1.0 (data knowledge discovery, red), 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). ROBHOOT uses the sustainability of the Seas case study in federated networks to offer a compact open-source technology with full reproducibility, automation, visualization and reporting for an open citizen science.

Going beyond ROBHOOT will go beyond state-of-the-art to implement evolutionary-biology inspired semantic algorithms. We will explore insertions and deletions, different types of recombination and crossover, and other evolutionary-based processes to find datatype properties from ontologies and raw-data from non-semantic databases. ROBHOOT will contrast many searching algorithms to gain understanding of the replicability of 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: Causal, explainable or interpretable discovery from observable data has been extensively studied (refs ++). Many of these studies have used symbolic reconstruction of equations by symbolic regressions or evolutionary methods (refs +++). However, a common gap throughout much of the literature is that of reconstruction from partially known models where the parameters represent eco-evolutionary processes from where explainability of the data can be done efficiently. The classical view on biology-inspired information processing technologies is to consider plasticity without structural changes, or without co-evolution among many interacting components (refs +++). Recent experimental evolution studies shows that rapid trait changes with new information processing capabilities 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 and adaptation is key to understand the emergence of new traits and information processing abilities to elaborate new discovery computation strategies in ecosystems.

Going beyond ROBHOOT will, for the first time, employ eco-evolutionary biology-inspired solutions to implement AI process-based methods to represent spatiotemporal causal inference in systems containing large heterogeneity and dimensionality (Figure 2). Eco-evolutionary biology-inspired models will be extended to deep process-based learning networks including trait and interactions as evolutionary changes to understand diversification patterns in these systems. The search for causal knowledge discovery will be applied to the data knowledge discovery for the sustainability of the Seas started in 1965 and currently containing 9 million entries, 1612 species (i.e., 50 variables and traits per species), around 20 countries and 11 sampling methods (Figure 2). 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 [21, 30].

Discovery in federated networks (WP3)
Evolutionary neural biology-inspired discovery in federated networks: Technologies in digital ecosystems around federated networks are scarce and mostly focus on decentralization, scalability and security fronts [9, 10, 14, 15, 20, 25]. 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 ([21, 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 sustainability of the Seas federated network

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 [14], 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 exploitation of ecosystems. 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].

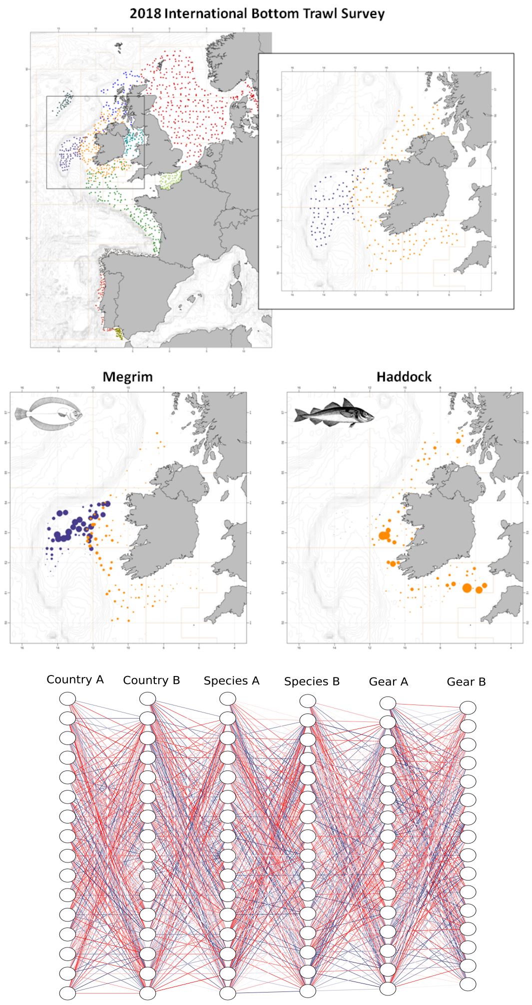

Figure 2: **Discovery for Sustainable Ecosystems**. ROBHOOT target discovery computation for new sustainability paths in complex ecosystems. The sustainability of the Seas case study [8] will be enriched to validate the technology when many species, human groups and technologies exploit resources (**Top left**, data-color points represent 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 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 by 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.

ROBHOOT v.3.0 connects 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.

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, i.e., evolutionary biology, ecology, computational neuroscience, computer science, data science, complex systems and experts in communication and biodiversity sampling methods. Data knowledge discovery gained by the analysis of the computation discovery capabilities of evolutionary-inspired semantic algorithms will be developed by the evolutionary biology, computer science and complex system members of the consortium (EBD-CSIC, IFISC-CSIC, SDSC). Data discovery can be transferred to the causal domain addressed 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 supported by ICREA, SDSC, and our company-partner SCITE, 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 cross-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 in ecosystems. 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. We will start with evolutionary biology-inspired semantic algorithms for data discovery applying it to 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. 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 paths of sustainability in our 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 open-source technology. We will develop the modeling in fast computing languages to develop low-level Agent Based Models (ABM) along all the theoretical development of the proposal (i.e., Julia, C++). We will contrast the ABM with differential/difference equations methods when a large number of agents, traits, 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 1.4a, Critical risks for the research approach). Our implementation activities are all complemented by numerical investigations contrasted for speed, replicability, and robustness with the sustainability of the Seas case study (Figure 2). The success of ROBHOOT would represent a breakthrough in the current discovery computation with direct application to sustainability of ecosystems and beyond. 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 open-source languages will lead to fast implementations of the demonstrators with high flexibility that will permit a rapid transit to the public (Impact section).

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.
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	WP2	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 +++).