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

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

### 1.2 Science-to-technology breakthrough that addresses this vision

**Data knowledge discovery (WP1)**

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

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

**Causal knowledge discovery (WP2)**

***Eco-evolutionary diversification-inspired AI algorithms (EEDA)***: Causal discovery from observable data has been extensively studied [10]. Many of these studies have used symbolic reconstruction of equations by symbolic regressions or evolutionary methods [11]. A common gap in much of the literature is one where parameters represent eco-evolutionary diversification processes, and thus, discovery can be explored broadly. The classical view on biology-inspired information processing technologies is to consider plasticity without structural changes, or without di-

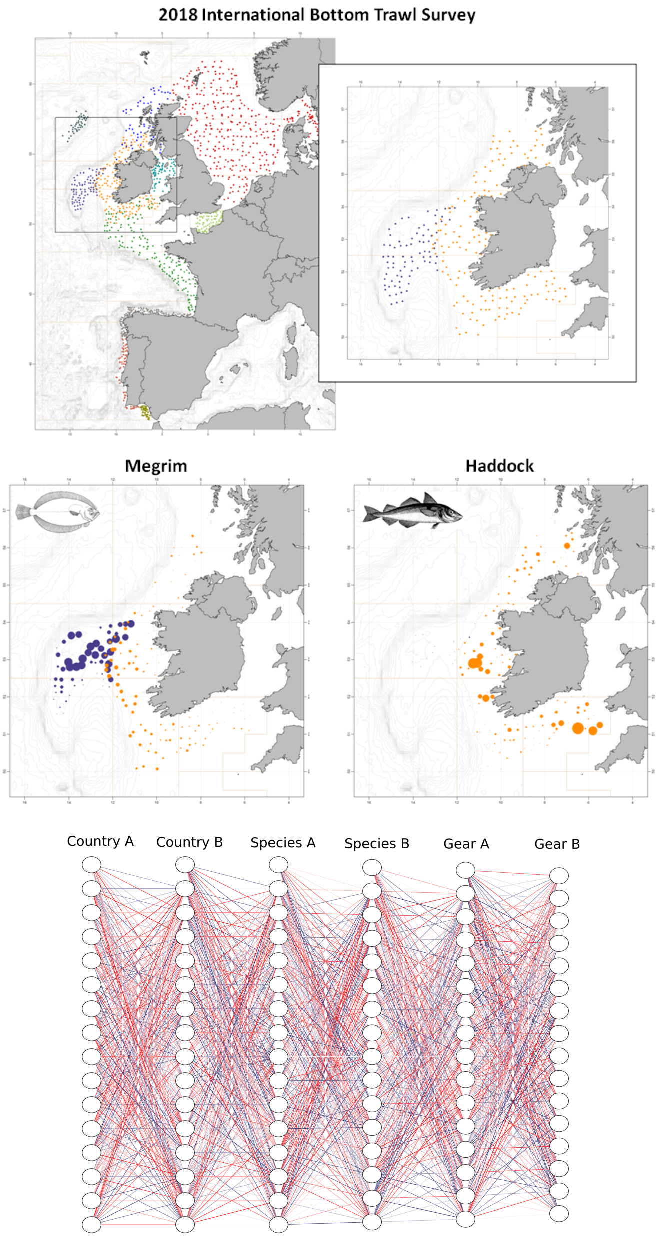


Figure 2: **Discovery for Sustainable Ecosystems**. ROBHOOT target discovery computation for new sustainability paths in complex ecosystems. The sustainability of the Oceans case study [7] will be enriched to validate the technology when many species, human groups and technologies exploit resources (**Top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