

# The BiodiFAIRse GO FAIR Implementation Network

## Purpose of the Implementation Network

The global aims of the biodiversity field are to understand the underlying mechanisms of nature, document and capture the state and dynamics of ecosystems, and build predictive models for the future. This understanding is based on access to and use of data, models and analysis tools, produced in ever greater quantities, and used by diverse communities tackling different aspects of biodiversity from observations, collections, sampling and experimental data.

The analysis of biodiversity data is essential for ecosystem services, risk analysis, and human well-being. The impact goes well beyond provisioning for material welfare and livelihoods, to include food security, resiliency, social relations, health, and environmental indicators. Species loss has dramatically accelerated around the world and now poses an existential threat to some ecosystems and susceptible human societies. There is an urgent need to; 1) collect, preserve and share FAIR data on species and ecosystems before they are lost to the scientific record, and, 2) provide automatic workflows producing biodiversity indicators so researchers, planners or policy-makers have evidence-based models to understand the complex dynamics of biodiversity.

To accelerate progress, both in the completeness and coverage of data, and in the richness of available information, all relevant sources of data must be aggregated; including sample-based data sets, ecogenomics, molecular research, remote-sensing, literature records, local and regional checklists, and expert knowledge. These resources, records and diverse data types should be used not only as a source of occurrence information, but also as an effective discovery tool on species abundance, community compositions, and interrelated genetic data.

Towards these long term aims, the partners of the BiodiFAIRse IN plan to build a virtual research environment and tools, collectively bringing their expertise to FAIR compliance by adapting data exchange standards, promoting the use and mapping of controlled vocabularies and collaborating in the development of registries gathering FAIR research objects and processes, analysis tools, and scalable workflows.

## Key Challenges / Objectives

- Create a single platform aggregating all biodiversity data and species collections starting with the [PNDB](#) and related initiatives ([Recolnat](#), [INPN](#), [GBIF France](#)) in France, and spreading through partner networks and institutions in Europe, Africa, the French Pacific territories. This IN will initially have a national focus but will be open to the international community.

- Develop scalable FAIR workflows and services to aggregate, analyse and increase the level of FAIRness of this data.
- Increase the richness of metadata descriptions linked to PIDs used by other GO FAIR INs.
- Develop competences in semantic modelling, promote best practices and emerging standards within the diverse institutes and universities across the francophone community who contribute data and resources to this shared resource.
- Produce French language training materials, webinars, BYOD, and an e-learning platform to increase the awareness of FAIR and the Internet of FAIR Data and Services. Promote FAIR awareness and approaches through conferences, seminars and workshops.

## **Overarching Principle of Operation**

We commit to comply with the Rules of Engagement of GO FAIR Implementation Networks. We will work as a distributed network with regular GO FAIR meetings. We will apply FAIR approaches to research processes and resources for biodiversity building on existing services and shared standards. The BiodiFAIRse IN will, where possible, leverage existing activities, archives and resources to extend their level of FAIRness and wider use by the community.

## **Targeted Objectives for the IFDS**

### **Standards and mapping**

Currently GBIF France implements biodiversity standards developed by [GBIF](#) and [TDWG](#) at the international level (Darwin core, EML, ABCD...) and uses IPT software to publish data. PNDB and SINP use metadata standards based on EML. We will work on the harmonisation of mapping taking into account emerging developments and recommendations within GO FAIR. The lack of FAIRization of scientific data and related products is often rooted in difficulties for end users in describing and structuring data (raw as well as processed data) using rich metadata and controlled vocabularies. We propose to evaluate and improve existing tools for these purposes such as the IPT toolkit, Morpho, FME ETL or Openrefine. We will explore and build on existing Life sciences products (such as ISAtools) repurposed for ecology applications.

## **Data and metadata portals**

We propose to develop metadata and data portals based on FAIR utilities linked to PIDs and open licences, using the expertise already developed around the “Living Atlases” community (<https://living-atlases.gbif.org/>)

## **Collaborative and workflows oriented online platform for data access and analysis**

We envisage an Ecology software accessibility portal potentially based on Galaxy and capitalizing on the ongoing GEO BON EBV pilot as part of the EuroGEOSS “Biodiversity and ecosystem Action group”. This project is based on sharing ecological "analytics bricks" and workflows through the Galaxy-E initiative <https://github.com/65MO/Galaxy-E>

## **Training**

We have extensive experience in developing French language training programmes for Biodiversity (data curation, sharing, interoperability, and exploitation). We will develop and publish content on data processing at scale, metadata standards and workflows, and build on teaching resources developed by PNDB as a member of the Galaxy Training Network <https://galaxyproject.github.io/training-material/> (See examples on [phenology study](#) or [Species Distribution Modeling](#)). E-learning resources will be produced inspired by "software carpentry" approaches.

## **Promote Interoperability**

We will seek accords on standards and PID to link different research objects (data, workflows, researchers, etc.).

## **Promote Reuse**

We will promote the use of dynamic Data Management Plans, standardised machine readable licences for data and metadata, well described portable FAIR workflows, as well as data products, tool parameters and versions. We will advocate to broaden the use and promotion of FAIR [Ecology standards](#) within the IN and the broader community.

## Develop “small scale” pilot initiatives for high demand applications (killer Apps)

In addition to these higher level objectives, which contribute to existing international initiatives, tools and solutions, some tasks for this IN will be oriented towards very specific “niche” domains with a high take-up potential. One example; **knowledge based descriptive data management and interactive identification**: Xper3 (<http://www.xper3.fr/>), enables publishing and collaborative editing of specimen and taxon descriptions but will evolve to become compliant with the FAIR principles. We aim to make the connexion between Xper3 / Recolnat and Ontologies to standardize morphological data for specimens of the French Natural History Museum collections.

## Membership

We consider this Manifesto to be one way by which the undersigned stakeholders can **speak with one voice** on a number of critical issues that are of generic importance to the objectives of FAIR, and on which we feel we have reached consensus. Although this collaboration is initially foreseen as a national IN, each of the partners has extensive international collaborations, and our philosophy is to be open to any individual or organisation that would benefit from joining the IN.

*The initial list of signatories include :*

**Yvan Le Bras**, MNHN/UMS PatriNat (AFB-CNRS-MNHN) : Responsable Pôle National Données Biodiversité **PNDDB** (AFB, CNRS, FRB, IFREMER, INRA, IRD, MNHN, BRGM, CIRAD, INERIS, IRSTEA)

ORCID [0000-0002-8504-068X](https://orcid.org/0000-0002-8504-068X)

**Anne-Sophie Archambeau**, IRD/UMS PatriNat (AFB-CNRS-MNHN) : Node manager **GBIF France**. (AFB, UMS PatriNat, MNHN, IRD, UPMC)

ORCID [0000-0001-6902-1465](https://orcid.org/0000-0001-6902-1465)

**Eric Chenin**, IRD/UMR UMMISCO (IRD, Sorbonne Université, autres tutelles universitaires): Head of Delegation GBIF France

**Jean-Christophe Desconnet**; IRD/Directeur de la Mission Infrastructures et données numériques

**Philippe Grandcolas**, CNRS/ Directeur UMR ISYEB (Sorbonne Université, MNHN, CNRS, EPHE)

**Michel Guiraud**, MNHN/Directeur **DGD Collections: Recolnat**, (MNHN, Université Montpellier, Université Clermont Auvergne & associés, université de Bourgogne, IRD, INRA, CNAM, CNRS, Agorologie, Tela Botanica, GBIF France)

**Frédéric Huynh** IR Système Terre

**Marc Pignal**, MNHN/Directeur **RI ReColnat** (MNHN, Université Montpellier, Université Clermont Auvergne & associés, université de Bourgogne, IRD, INRA, CNAM, CNRS, Agorologie, Tela Botanica, GBIF France)

**Jean-Denis Vigne** MNHN/Directeur **DGD REVE**

**Régine Vignes-Lebbe**, UMR **ISYEB** (Sorbonne Université, MNHN, CNRS, EPHE) :  
**Xper3** team

*Further collaborators and potential future partners :*

National initiatives with potential interest:

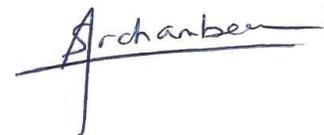
- National project dedicated to structure data from national citizen science initiatives: "[65 Millions of Observers](#)" through Romain Julliard
- French node of the Europe "**Analysis & Experimentation on Ecosystems**" network [AnaEE](#) through Christian Pichot
- French node [BOLD](#) & [iBOL](#) through Rodolphe Rougerie
- [INIST](#) as french **Datacite** node for support concerning DMP through [OPIDoR](#) (French DMPtools) and DOI provision
- [BBEES](#) "Databases on Biodiversity, Ecology, Environments & Societies" Service Unit through Cécile Callou
- Link with agronomic research, Odile Hologne, head of department of scientific information of **INRA**, french research institute in agriculture, environment and food.

International initiatives with potential interest:

- **European Galaxy node** through Björn Grüning (representing links for Elixir + Bioconda + Biocontainer) and Bérénice Batut (representing Galaxy Training Network and related infrastructure as link towards Software carpentry)

**Signed by :**

**Anne-sophie Archambeau**



**Jean-Christophe Desconnet**

**Eric Chenin**

**Philippe Grandcolas**

**Michel Guiraud**

**Frédéric Huynh**

**Yvan Le Bras**

A handwritten signature in black ink, appearing to be 'Yvan Le Bras', with a long horizontal stroke extending to the right.

**Marc Pignal**

**Jean-Denis Vigne**

**Régine Vignes-Lebbe**