Plan for data management solutions enabling to fill the known information gaps.

Deliverable 2.3

Project Title: Priodiversity LIFE

Project No: Project 101104217 — LIFE22-IPN-FI-Priodiversity LIFE

Work Package: WP2

Task 2.2: Conceptual Biodiversity Information Data Management System for Regional Biodiversity Action Planning (ReBAP)

Subtask 2.2.1: IT-development solutions to enable filling the identified gaps of biodiversity

Task Leader: Luomus (Finnish Museum of Natural History)

Contributing Partners: Syke (Finnish Environment Institute), MHJHT (Parks & Wildlife Finland), MHMT (Metsähallitus Forestry Ltd), SMK (The Finnish Forest Centre), ELY (KEHA-Centre aka Centre for Economic Development, Transport and the Environment)

Authors: Esko Piirainen (Luomus), Mikko Heikkinen (Luomus), Anne Koivunen (Luomus), Aleksi Lehikoinen (Luomus), Ulpu Leinonen (Syke), Päivi Sirkiä (Syke).

PRIODIVERSITY

NATURA 2000

Co-funded by the European Union

Summary

The Priodiversity LIFE project is improving Finland's biodiversity monitoring by updating data systems, harmonising practices, and making information more accessible. Monitoring has traditionally relied on multiple actors and systems, which have produced valuable datasets but also led to duplication, inconsistencies, and outdated workflows. Subtask 2.2.1 focuses on these challenges by defining the technical requirements — such as tools, applications, and technological choices — to address critical data gaps identified in the project, notably in birds of prey monitoring, moth monitoring, breeding bird monitoring in protected areas, and implementing the complete species lists. The work combines a guideline document with the development of practical tools, led by Luomus and Syke in collaboration with partners. It builds on existing infrastructures such as FinBIF and Finnish Nature Information Hub, Luontotieto.fi, ensuring that all outputs follow the FAIR principles — making biodiversity data findable, accessible, interoperable, and reusable. Following the guidelines outlined in this document (Deliverable 2.3), the next steps include developing new mobile and online tools and producing data services that support Regional Biodiversity Action Plans. Together, these efforts are making monitoring more coordinated and efficient.

1. Introduction

The Priodiversity LIFE project aims to strengthen Finland's biodiversity monitoring capacity by modernising data management systems, harmonising monitoring schemes, and improving the accessibility and usability of biodiversity information. Biodiversity monitoring in Finland involves a wide range of actors — including public authorities, research institutions, NGOs, and citizen scientists — often within scheme-specific arrangements that have developed over decades. Although this has enabled the collection of valuable long-term datasets, it has also led to fragmented workflows, inconsistent data formats, and multiple IT systems, many of which are nearing the end of their operational lifespan.

As a result, volunteers and professionals often need to enter the same observation into multiple systems, each using different formats and vocabularies. Data might be stored in systems with limited interoperability, or even in paper archives, making it more difficult to analyse trends, combine datasets, or respond swiftly to environmental changes. Data can also be difficult to find and utilise due to the scattered sources. At the same time, the availability of modern, centralised biodiversity data infrastructure — especially through the Finnish Biodiversity Information Facility (FinBIF) and Syke's Finnish Nature Information Hub (Luontotieto.fi) — presents a clear opportunity to simplify processes, cut down on duplication and increase the findability of biodiversity-related data.

PRIODIVERSITY NATURA 2000 Co-funded by the European Union

Within Priodiversity LIFE, Subtask 2.2.1 has focused on five concrete cases where these gaps are most evident and where solutions can deliver direct added value:

- **Birds of prey monitoring** — where multiple overlapping databases and reporting forms complicate data entry and create unnecessary duplication.

- **Moth monitoring** — where a long-term programme faces the risk of data loss unless legacy records are migrated into modern systems.

- **Breeding bird monitoring in protected areas** — where valuable datasets remain partly in paper archives and lack integration into national biodiversity infrastructures.

- **Complete species lists** — where an important methodology for collecting presence–absence data is still underused and requires better tools, uptake, and outreach.

- **Natural habitat types data collection, management and sharing** - where there are currently no uniform methods or solutions in use

Subtask 2.2.1 tackles these challenges by defining the technical requirements, such as tools, applications, and technology choices, needed to fill the most important biodiversity monitoring data gaps. As a first step, Deliverable 2.3 serves as a guideline document to help the technical development. The second phase involves developing concrete data management and monitoring tools. These tools are developed within the existing services and platforms of Luomus and Syke to make development efficient and effective. This subtask also includes the development of result services and data products. These new and improved existing data management tools and services will form Deliverable 2.24.

Deliverable 2.3 is co-authored by Luomus and Syke, in consultation with other WP partners, including SMK, MHJHT, MHMT, and ELY, ensuring that technical solutions are developed collaboratively and aligned with user needs and policy requirements.

The following sections offer in-depth case studies for each of the five key monitoring areas. They showcase the current progress, discuss the challenges faced, and outline the development needs to be addressed within Priodiversity LIFE.

2. Information gaps in biodiversity data

2.1. Recognised data needs

2.1.1. Birds of prey monitoring data

Current status

From the perspective of a volunteer participating in birds of prey monitoring projects, participants need to fill in data in several formats and locations for the same observation. For example, if nestlings are ringed at a nest that is located on a birds-of-prey monitoring grid, the participant needs to fill in the data to: 1. the ringing database, 2. a nest card, 3. in species monitoring grid card, and 4. also include the numbers in the submission to the so-called summary study. While it is strictly necessary to submit data only to the ringing database, many participants engage in all the different monitoring schemes. The user needs to use two different IT systems, and the monitoring grid and summary scheme reporting is done using a paper form. Depending on the nesting species, the user must also report the data using different fields and vocabularies, since there are four different nest monitoring schemes: 1. Pandion haliaetus, 2. Haliaeetus albicilla, 3. other birds of prey, and 4. other non-predatory birds.

From an IT-point of view, there are five different databases: the ringing database ("Tipu"), osprey monitoring database ("Pandion"), white-tailed sea eagle monitoring database ("Haliaeetus"), other-bird-of-prey monitoring database ("Petopesä"), and a fifth is for the nest card system for non-bird-of-preys ("Pesäkortti"). The structure of the last four databases is roughly the same, but each nest monitoring scheme has its own fields and vocabularies.

The bird ringing data management is handled using three IT systems: the main database (Tipu https://bitbucket.org/luomus/tipu/), an API (Tipu-API https://bitbucket.org/luomus/tipu-api), and "Sulka" (https://bitbucket.org/luomus/sulka), an Excel-like data entry system for the bird ringers.

Data management for the five nest monitoring schemes is done using two IT-systems: all the five databases are administered using the same framework (PeSe https://bitbucket.org/luomus/pese). Participants fill in their data using a multipurpose form service Kirjekyyhky (https://bitbucket.org/luomus/kirjekyyhky/).

Finally, all services utilise the same authentication and authorisation portal, Lintuvaara (https://bitbucket.org/luomus/lintuvaara/), which combines all services to one location for data administrators and participants.

There are no modern IT systems for schemes handled using paper forms: the grid monitoring scheme and the summary monitoring scheme. Data is stored in text files and analysed using scripts.

Description of the development need

The Finnish Biodiversity Information Center (https://laji.fi) has developed multipurpose IT systems that are used currently to implement 18 different monitoring schemes. Many of the participants in the bird-of-prey monitoring scheme are already familiar with FinBIF services and utilise them to participate in various bird monitoring schemes, such as line transect, point count and winter bird monitoring. To make participation in monitoring schemes easier, there are three main courses of action:

1. Combine the different nest-card monitoring schemes into one monitoring scheme under FinBIF (same fields and vocabularies for all the different studies)
2. Remove the need to fill in separate grid and summary forms
3. Automate data flow from nest-monitoring schemes to bird ringing database entry

These steps make fieldwork easier (using the same fields and vocabularies for all bird species) and reduce the need to report data on several different forms. A bird ringer would first fill in a nest card, and an automatic data entry row would be generated in Sulka for them to fill in additional data to the ringing database. (Combining ringing database and monitoring schemes is not logically feasible, since bird ringing has many aspects that a regular monitoring scheme does not have, such as handling distribution of rings, etc.). Grid studies would be replaced by filling in nest cards and calculating the numbers from the nest card data. The same goes for the summary form: data can be calculated from nest cards, and in the case of field-nestlings (where a full nest card is not appropriate), a lightweight nestling entry.

These development needs would help eliminate multiple IT systems, freeing development resources from maintenance to improving services.

Using the data would be easier, as they would be in a more harmonised format (with the same fields and vocabularies). The needs of public authority users can be taken into account when designing the common data model to make bird-of-prey nest data easier to use in land-use planning and nature management solutions. Nest grid and summary data would be findable, accessible, and preserved in a well backed-up and administrated database (instead of text files).

Technical solution
The Lintuvaara authentication service would be replaced with the already existing FinBIF authentication service. The already existing FinBIF Notebook service can replace the five different nest-monitoring databases. The Notebook would require several new features to accommodate the needs of these new monitoring schemes, but the underlying technologies and database solutions are well-suited for the task. The Kirjekyyhky from service can be shut down when all forms are submitted using the FinBIF Notebook.

Timetable
The technical solutions described in this section are highly complex, and it may not be possible to implement all of them within the scope of this project. The primary goal is to prepare precise and detailed plans and, where necessary, seek external funding for their future implementation.

Responsible: Luomus

2.1.2. Moth monitoring data

The national moth monitoring programme (Valtakunnallinen yöperhosseuranta) is coordinated by the Finnish Environment Institute (Syke) and has been running continuously since 1993. Using a network of light traps deployed across Finland, it records the occurrences and individual abundances of moth species with weekly accuracy from spring to autumn. In addition to species and abundance data, the programme collects ancillary information such as trap type, habitat type, and sampling effort metadata.

Moths are ecologically important and serve as sensitive indicators of environmental change, especially in nocturnal ecosystems that are not covered by most other insect monitoring programmes. The long-term dataset supports biodiversity trend analysis and helps detect changes in species distributions and abundances over time, contributing to national and international biodiversity assessments.

Current status
Until recently, moth monitoring data were stored in the legacy YÖPETI system, which is approaching the end of its life. Without a replacement, there would be no viable long-term platform for collecting, managing, and sharing these data. Within the Priodiversity LIFE project, a new moth monitoring data management solution has been developed on the FinBIF Notebook platform. The system has been in pilot use since the 2024 field season and moved into full production use in 2025, incorporating improvements based on feedback from monitoring coordinators and volunteers responsible for gathering and identifying specimens.

Description of the development need

The main data gap to be addressed is the absence of a sustainable, long-term system for storing and managing moth monitoring data. While the new Notebook-based system now supports current and future monitoring activities, historical records—totalling approximately 1.3 million occurrences collected since 1993—remain in the YÖPETI system. These data must be migrated to the new platform to ensure long-term preservation and availability. Migration will include data conversion to the Notebook schema, quality control and validation, and correction of any detected errors.

Once integrated into the new system, all data will be openly accessible via FinBIF's Laji.fi portal and shared with the Global Biodiversity Information Facility (GBIF) in accordance with open biodiversity data practices. The open publication of the complete dataset will allow researchers, conservation practitioners, and policymakers to combine moth monitoring data with other biodiversity datasets, supporting cross-taxon and ecosystem-wide analyses.

Technical solution

The new moth monitoring data system is implemented within FinBIF's Notebook system, ensuring compatibility with FinBIF's broader biodiversity data infrastructure. The Notebook form system is available as open source at https://github.com/luomus/laji-form

The system is accessible at https://laji.fi/project/MHL.1136/about

The data is accessible at https://laji.fi/observation/statistics?collectionId=HR.4511
Historical data from 1993-2019 is available partially and as a secondary copy, pending migration of the primary data to the new system.

Timetable

- Pilot use: 2024 (completed)
- Production use: 2025 (completed)
- Migration of historical data from YÖPETI: 2026 and beyond

Responsible: Luomus, Syke

2.1.3. Breeding bird monitoring in protected areas

Current status

The breeding bird monitoring of birds in the protected areas has been conducted using line transect method. The primary survey results and the maps of the survey sites have been stored in paper format since the 1980s by the Finnish Park and Wildlife Services

(Metsähallitus). Approximately half of this data has been digitised to FinBIF Notebook system (Laji.fi) through the collaboration of Luomus and the Finnish Environmental Institute (FEO project). Metsähallitus has been collecting the data into its own Excel sheets. Furthermore, small part of the annual surveys of the common bird monitoring coordinated by Luomus have also been occurring in protected areas. These survey results have been manually extracted from the Luomus database for Metsähallitus and the two data sets have not been communicating automatically. The aim has been to centralise the data entry and allow data sharing to all potential users (e.g. authorities, researchers, managers).

Description of the development need

The rest of the primary data in the paper forms will be digitised to the Notebook system. Furthermore, the data from the new survey including repeated surveys from the old survey sites will be directly entered to the Notebook system. The system will be developed so that the whole dataset from the protected area network can be downloaded through the system in uniform format.

Technical solution

Each bird record and a segment of the line transect will be automatically classified if this is situated inside or outside protected areas through GIS information.

Timetable

- Digitising the old data 2025-2027
- Technical development of the databases 2026

Responsible: Luomus, MH

2.1.4. Complete species lists

Complete species lists are a method where observers record all species of a selected group encountered during an excursion, and explicitly indicate which species were not observed. This generates presence–absence data, a dataset type that has historically been missing in Finland. Such data is particularly valuable because it provides a more complete picture of biodiversity than simple presence-only records, enabling more accurate analyses of species distribution, abundance trends, and habitat quality.

While presence-only data can suggest where a species occurs, absence information allows for more robust modelling and inference — for example, identifying range contractions, assessing the effectiveness of conservation measures, or prioritising sites for protection. Globally, complete list recording is well established among birdwatchers and some other

taxa, but in Finland, there has been neither a strong tradition nor widely available tools for it. This has left significant methodological gaps in biodiversity monitoring, particularly for groups such as fungi, lichens, and certain insect taxa, where systematic coverage has been sparse.

The Priodiversity LIFE project addresses this by introducing and scaling up complete list recording in Finland, aiming both to improve the methodological scope of biodiversity data collection and to expand monitoring coverage for underrepresented taxonomic groups. In addition to producing valuable datasets for research and conservation, this approach also encourages a more systematic observation culture among both citizen scientists and professionals.

Current status

The FinBIF (Laji.fi) Notebook observation system has been extended to include complete list forms. Currently, there are 10 active forms covering taxonomic groups such as butterflies, flowering plants, bumblebees, and edible mushrooms. Data entry is possible via both the web interface and the FinBIF Mobile Notebook app (Android and iOS).

So far, approximately 370,000 complete list observations have been entered, including a bird atlas dataset which covers most of this data. All data are available as open FAIR data through the FinBIF portal (https://laji.fi/observation/list?collectionId=HR.5615,HR.4471) and are also shared with the Global Biodiversity Information Facility (GBIF).

Description of the development need

The complete list forms are in active production use, but user feedback has identified opportunities for improvement, particularly to:

- Make the forms more appealing and intuitive to increase participation across different taxa and user groups.
- Reduce data entry errors through improved guidance and validation.
- Promote adoption beyond the current base of experienced amateur naturalists (e.g. birdwatchers) to include professional surveyors, nature consultants, and conservation authorities.

Communication and outreach will be essential. Collaboration with ELY Centres and relevant networks will help embed complete list methods into standard monitoring practices.

The mobile app will require continuous development to remain compatible with evolving smartphone platforms and to meet changing user expectations. The current approach keeps the tools simple to maximise participation, but incremental improvements will ensure accessibility, stability, and ease of use.

Technical solution
The complete list functionality is implemented within FinBIF's open-source Notebook system (https://github.com/luomus/laji-form), ensuring full integration with Finland's national biodiversity data infrastructure. Forms are accessible via https://laji.fi/save-observations. The system design allows for future expansion to additional taxa and languages.

Timetable

- Production use: 2024
- Incremental improvements: 2025 and beyond

Responsible: Luomus

2.1.5 Natural Habitat Types Data Collection, Management and Sharing

Current Status
Natural habitat types have become a key indicator for biodiversity monitoring and an essential tool for conservation planning in the 2020s. Reliable data on their extent and ecological condition is critical not only for statutory conservation tasks but also for broader societal decision-making. Despite its importance, habitat type data in Finland remains fragmented, with no centralized repository or coherent information architecture to support efficient access and use.

The Finnish Environment Institute (Syke) plays a leading role in producing and utilizing habitat data, particularly for national assessments and EU-level reporting. Metsähallitus Parks & Wildlife Finland (MHLP) manages data from state-owned protected areas. Other contributors include the Natural Resources Institute Finland (Luke), ELY Centres, the Finnish Forest Centre, the Geological Survey of Finland (GTK), and municipalities, which collect data through land use planning and ecological surveys.

Development Needs
Other projects (e.g. LYSEKAPP, LYSEKTIETO) have already initiated the development of practical tools for national, interoperable field data collection, including mobile applications for recording monitoring data and systems for managing and distributing that data. The Priodiversity LIFE project can build on these foundations by further developing the solutions based on identified needs, such as expanding the user base, improving data sharing, and enhancing the ability to input data into the system.

Technical Solution
With the basic technical infrastructure in place, Priodiversity LIFE can focus on refining and scaling these tools. A key priority will be enhancing interoperability by working collaboratively across organisations to define shared vocabulary and variables for use in field monitoring and

other relevant applications. This will help ensure that data is consistent, reusable, and accessible to a wide range of users.

Timetable
Technical and interoperability development in 2026 and beyond.

Responsible: Syke

3. Methodology

3.1. FinBIF

FinBIF (the Finnish Biodiversity Information Facility) serves as Finland's unified, open-access platform for collecting, managing, and distributing data about the nation's biodiversity. It brings together data from digitised natural history collections, species checklists, DNA barcoding, systematic monitoring programmes, citizen science, and more, all under a single IT framework and portal.

FinBIF is notable for handling every stage of the data lifecycle:

- Digitisation: converting physical museum specimens into digital records.
- Data generation: capturing observations via structured surveys as well as opportunistically through citizen contributors.
- Collation and Integration: combining datasets from multiple sources into a coherent, harmonised data mass through FinBIF's Data Warehouse.
- Annotations by taxon experts enrich the data quality and feedback loops help improve the source data.
- Publishing: FinBIF shares data according to the FAIR principles.

At the core of FinBIF lies a national taxonomic backbone—a dynamic checklist of Finnish species (currently 44 thousand), supporting consistency in naming and data integration. Yearly snapshots are published under open licenses for citation and reference.

IT Architecture & Services

All data flows and services operate under a shared IT architecture, including:

- A data warehouse where harmonised records are stored.
- APIs powering the public portal (Laji.fi / Species.fi) and enabling third-party access.

- Support for both open data and restricted-use data, the latter accessible only to authorities or via request—ensuring sensitive species are protected while still enabling use.

FinBIF's portal offers:

- Browsing and downloading data on species, their occurrences, images, and descriptions.
- Tools for entering new observations (including via iNaturalist Finland).
- Visualisations, statistics, and metadata to aid both lay users and professionals.

Managed by the Finnish Museum of Natural History (Luomus), FinBIF harnesses cooperative governance and strategic planning, and is certified as a trustworthy data repository (CoreTrustSeal), reflecting its commitment to high standards in data stewardship.

3.2. Description of the Notebook system

Notebook is FinBIF's system for managing biodiversity observation data. It is a web-based tool for recording both *opportunistic* observations (e.g., an individual spotting a rare bird) and *structured monitoring data* from organised surveys and monitoring schemes. It serves a wide range of users — from national monitoring programmes and citizen science projects to communities of species enthusiasts.

At its core, Notebook uses a form engine called LajiForm. This engine turns a predefined data structure (a "schema") into an interactive online form. What makes LajiForm special is that it can adapt the way questions are asked to match the needs of different audiences and contexts. For example, instead of showing a long list of technical fields, the form can guide the user step-by-step: if a location is marked as a potential flying squirrel nest, it can prompt relevant follow-up questions and automatically record certain details in the background. This makes the system more intuitive and helps collect better quality data.

Forms can be highly customised — from simple versions for schoolchildren to complex, compact forms for expert recorders. This flexibility is possible because Notebook separates the *data structure* from the *user interface*. Designers can control which questions appear, when they appear, and how they respond to previous answers.

Technically, LajiForm is built using modern web technologies (React) and is based on an open-source form library (RJSF) originally developed by Mozilla. FinBIF has significantly extended this library and contributed improvements back to the open-source community.

Monitoring scheme reporting sites (such as routes and nests) are managed by so-called NamedPlaces. These places can be public or visible only to approved participants or individual persons. In some schemes they can be reserved for the year using an in-built reservation system. Additional information, such as maps, hints for the participants can be included for a place.

3.3 Finnish Nature Information Hub (Luontotieto.fi)

Luontotieto.fi, the Finnish Nature Information Hub, is an online service designed to improve the discoverability, usability, and impact of nature data. It brings together existing datasets, guides users to their original sources, and provides practical information on how to access and apply the data.

The Hub's data search function allows users to easily browse metadata on nature datasets produced by various Finnish organisations. Currently, metadata is harvested from the Finnish Environment Institute (Syke), the Finnish Biodiversity Information Facility (FinBIF), and the Natural Resources Institute Finland (Luke). The long-term goal is to provide a comprehensive search interface for all nature-related datasets in Finland.

Each metadata entry includes key details such as the dataset name, description, production and update dates, the responsible organisation, and contact information. Where available, links are provided to the original metadata page and to view or download the dataset directly from its source.

In addition to data discovery, the website offers insights into nature monitoring programs, methods, and the role of nature data in decision-making and environmental assessments. Luontotieto.fi also promotes collaboration among data providers to improve the FAIR principles of data—Findability, Accessibility, Interoperability, and Reusability—ensuring that nature data becomes more efficiently available and usable across various sectors.

The Priodiversity LIFE project can contribute to the further development of Luontotieto.fi by identifying and addressing evolving user needs. This may include enhancing functionalities, expanding content, and ensuring that data produced within the project is made discoverable through the service.

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.