

Data policy of the Finnish Biodiversity Information Facility

1. Purpose of the data policy

The Finnish Biodiversity Information Facility (FinBIF) compiles Finnish biodiversity data (see Chapter 4 for details) in digital form into a data warehouse. Data contained in the FinBIF data warehouse not classified as sensitive [1] or, on other grounds, partially or entirely secret [2], is open and public information. Access to read and use data classified as secret on grounds stated above is always based on personalized access rights and requires a separate agreement with the owner of the data in question. Access rights to open data are based on user access licenses, which are described in the data transfer agreement and which may vary between datasets.

This data policy sets the principles for compiling, maintaining and using all digital datasets managed by the FinBIF.

The **FinBIF** is a virtual centre administered by the Finnish Museum of Natural History Luomus. This data policy is not intended to contradict Luomus' data policy used to direct the implementation of its statutory duties. In the event of conflicts, the data policy of FinBIF will be observed.

The data policy also acts as binding guidelines for data management at FinBIF. It applies to all FinBIF data and those working with the data, persons employed by Luomus and persons in any contractual relationships with the institution. In addition, all persons with access to data management at FinBIF based on user rights granted to them will be bound to observe this data policy.

2. Definitions of data policy terms

This data policy employs the following terms with the following meanings:

- **Data** refers to entities consisting of individual records stored in digital form. The number of records can vary from a handful to millions and they can include text, numbers, photographs, audio or comparable content, as well as related metadata.
- **FinBIF information system** refers to the information system maintained by the Information and Communication Technology (ICT) Team of Luomus. The system may be programmed either by the ICT team, by another current or previous member of Luomus staff, or it may be acquired from or provided to Luomus by an external party.
- **Information system** refers to a system based on one or more electronic databases where data is managed to serve a particular need.

- **Metadata** refers to information describing datasets, such as description of the content, method and time of creation, as well as the collector of the dataset, etc.
- **Primary dataset (primary copy)** refers to the original copy of the dataset that will be kept up to date. This version has metadata and an owner.
- **Secondary dataset (secondary copy)** refers to a complete or partial copy of the primary dataset, which will be kept up to date only by updating the primary dataset. Following each update of the primary dataset, the out-dated secondary copy is overridden by a new version.
- **Dataset formation** describes the manner or method in which datasets are compiled into target systems and the related processes.
- **Dataset updating** refers to amending the original data contained in the primary dataset, in other words, replacing information with new or corrected information either partially or completely. It also refers to replacing the secondary dataset with a new, updated version of the primary dataset.
- **Dataset maintenance** refers to keeping data up to date and available for use. For primary datasets, updates will be made immediately visible in the storage system. For secondary datasets, outdated data will be replaced with up-to-date data in new copies.
- **Dataset type** refers to datasets that are typically the result of a single process, including individual and descriptive metadata. This concept can be used to specify principles: a certain dataset type includes data covered by the same principles.
- **Openness of data** refers to content not classified as secret being open for anyone for any purpose, as long as access right principles, such as attributing sources in accordance with good scientific practice, are observed. Openness can be restricted by user licenses and confidentiality criteria (see Chapter 6, paragraph “Data usage”).

3. Acts, regulations and policies guiding the data policy

The data policy of the FinBIF is guided by the same acts and obligations as that of Luomus.

In its data policy, FinBIF complies with policies, strategies and programmes related to the open access to digital data agreed upon both at national and international levels including the EU, specifically the policy to increase the openness of data by the Finnish government.

The services of the FinBIF comply with the FAIR Data principles, originally published in 2016¹ (Findable, Accessible, Interoperable, Reusable). The European Council and the Finnish Ministry of

¹ <https://www.go-fair.org/fair-principles/>

Education and Culture is also committed to the FAIR Data Principles and obeying them is an official policy of both organisations.

The public nature of the Luomus' operations is governed by the Universities Act and the Act on the Openness of Government Activities, which has a direct impact on FinBIF's data policy as well. The data policies of the Luomus and, therefore, FinBIF are in line with the data policy of the University of Helsinki, the parent organisation of the Luomus.

Other regulations that have an effect on FinBIF's data policy are the following acts and regulations concerned with online services and government as well as related to the creation, use, distribution, utilisation and storage of digital data: the Nature Conservation Act, the Administrative Procedure Act, the Act on Information Management Governance in Public Administration, the Personal Data Act, the Act on Electronic Services and Communication in the Public Sector, the Act on Electronic Signatures, the Archives Act, the Act on Collecting and Preserving Cultural Material and copyright regulations.

The data policy of the FinBIF also takes into consideration the policies in Finnish legislation and strategy influenced by the EU's PSI (Public Sector Information) and the INSPIRE (Infrastructure for Spatial Information in Europe) directives. In accordance with the act (421/2009) and decree (725/2009) on spatial data infrastructure, Luomus has duties as regards the shared use of location data related to species distribution and biogeographical regions. Datasets must be maintained and shared in accordance with stipulations set in the act and the implementing rules of the INSPIRE directive.

The personal data contained in the datasets will be treated in accordance with the General Data Protection Regulation² (GDPR) applied in all EU countries as of 25th of May 2018.

4. Data policy scope

The data policy covers all data managed by FinBIF. Datasets are received for management, storage and distribution on the basis of agreements.

The following types of datasets are included:

1. Data on collection specimens; collection specimens include complete organisms, macro- or microscopic parts of organisms, tissue samples or DNA, and fossils
2. Data on collections that contain strains of living plants, fungi or microbes; strains may consist of one or more individual specimen or growth, or several specimens or growths produced by a single seed or spore lot, or a vegetative reproduction or propagation event. These may be in various states of their lifecycle (for example, spore, seed, seedling, full-grown)
3. Observation data without collection specimens

4. Material documenting collection specimens or observations, such as photographs, audio recordings, drawings, etc.
5. Monitoring and mapping data: such as satellite tracking data, and other data compiled, combined and processed in order to produce maps or statistics, or to implement monitoring
6. Nomenclatures and taxonomic hierarchies
7. Results of scientific measurement and classification
8. Various analysis results: sample dating results, isotope ratio and element composition assays, DNA sequences and genotype assays, etc., and related process knowledge
9. Metadata related to the dataset types listed above
10. Spatial and location data related to dataset types listed above, such as collection and observation locations or regions, location within collections, etc.
11. Personal data related to dataset types listed above, such as name and organisation of the data producer, contact details, etc.

5. Data policy goal

The general goal of the data policy is to promote the usability and use of data managed by FinBIF in support of research, education, administration, nature conservation, business life and society by following the FAIR Principles.

In order to reach this goal, the data policy strives to

- create consistent practices for storing, publishing and distributing data;
- ensure that primary copies, or updatable versions of data, are available for all data in electronic form with metadata and owner details attached;
- enable the distribution of biodiversity datasets for open use as quickly, efficiently and exhaustively as possible;
- to increase cooperation between data collectors, creators and users.

The data policy steers Luomus and its contracting partner staff in managing and handling data received and shared by FinBIF. It also contributes to different parties observing consistent principles when submitting and using data. By defining principles of data management and distribution, the policy indirectly improves the quality of data.

6. Principles of data creation, maintenance and use

Data creation

Secondary datasets are stored (copied) into the data warehouse of the FinBIF from primary dataset sources. Copying is based on written agreements or other documented agreements (e.g. acceptance of the terms and conditions of the observation recording system) made between source material owners or managers and Luomus. Existing data is transferred to FinBIF automatically, through data transfer interfaces or by utilising other efficient methods. New data is

accumulated into FinBIF either from the primary dataset storage systems of its contracting partners or the primary dataset storage systems that are currently in use as well as those that will be established as part of FinBIF's services. Secondary copies will be copied to the FinBIF data warehouse at scheduled intervals, with a delay or in real time, depending on the dataset in question.

Data is owned by the organisation identified as the creator; or, when the material is created as a private activity, an individual. Transferring data to FinBIF for storage and distribution does not affect the ownership of the data. Data ownership refers to the right to make decisions on data use and transfer to the FinBIF data warehouse. This is arranged with a mutual agreement or with some other kind of documentation.

Dataset maintenance and storage

Primary dataset storage systems are a part of the services provided by FinBIF, in addition to which new systems will be built. Datasets within these systems are maintained and updated by Luomus and data owners in a manner agreed separately. Contracting partners of FinBIF are responsible for maintaining and updating data stored in the primary dataset storage systems owned or managed by them. Irrespective of ownership, control, or update responsibility, only the secondary copy (i.e., the copy that is updated through the update of the primary dataset) of the primary datasets will be stored in the FinBIF data warehouse. Data management and storage will be conducted in accordance with decisions and instructions that comply with the acts, regulations and policies listed in chapter 3 of the data policy.

Copies of biodiversity data transferred to the FinBIF data warehouse will be standardised with the FinBIF core database primarily before storage. Among other things, the core database includes the national nomenclature and taxonomy.

A metadata listing distributed through public websites will be maintained for datasets. Information on new datasets will be added to the list no later than after the data is completed, at the time of data transfer or, in the case of datasets compiled continuously, without delay after data transfer to the FinBIF systems has begun.

Data usage

As a rule, digital data managed by the FinBIF is public and free to use [4] (open data). As a rule, data is shared in compliance with the principles set by the Creative Commons (CC) licenses [4]. Selection and implementation of CC licenses [5] is conducted in accordance with agreements made with data owners. Data publication will be conducted in accordance with decisions and instructions that comply with the acts, regulations and policies listed in Chapter 3.

Data managed by FinBIF may include data classified as sensitive on the basis of either general legislation [6] or decisions made by the Finnish environmental and natural resources administration [7]. Availability of and access to sensitive material is usually restricted. Under the leadership of FinBIF, a national and binding decision in principle [8] containing general policies for availability and access restrictions, as well as an attached list of sensitive species in Finland, has been drafted.

Openness can be restricted in several ways, the most common of which is making location data coarser. Time dependent restrictions to the access of data can also be used to restrict otherwise open data. In addition, FinBIF may have datasets whose openness and availability are restricted at the request of data owners or managers as agreed. The rights related to sharing this type of data are also defined in accordance with the user access licenses stated above on a case-by-case basis. These also include embargos on the use of scientific research data, which sets a date before which a given dataset is not open. Access-restricted research material may be made available with the consent of the owner.

Data will be stored in the FinBIF data warehouse in the form they were received from the data owners. The restrictions stated above will only come into effect when data is transferred from the data warehouse to the public service of the Laji.fi website and made available for use. In accordance with the goals promoting the open use of FinBIF data, FinBIF strives to ensure that parties classified as authorities have unlimited access to the data for official duties. However, this is based on a separate voluntariness agreement between the owners of the material and FinBIF.

Data disclosed voluntarily through FinBIF for use by authorities is protected under section 24, subsection 1, item 16, and section 26, subsection 1, item 2 of the [Act on the Openness of Government Activities](#). In practice, data disclosed for official use by authorities cannot be further disclosed without the consent of the data owner. This is ensured by mutual agreements between FinBIF and the authorities when agreeing on how authorities will have access to all FinBIF data.

Data openness must not violate data protection. Restrictions to the openness of data and their justification are described in the metadata. Open data is made available to the academic community, public administration, organisations, businesses and private citizens for free use through the FinBIF portal (Laji.fi), managed by Luomus. Other public distribution channels for open data may be established.

The open and public nature of FinBIF data is promoted by sharing data and metadata from the data warehouse for publication in the Global Biodiversity Information Facility portal (gbif.org) and potentially in other similar services.

7. Sharing responsibilities

Drafting, updating and administering the data policy

FinBIF's managers are responsible for the drafting, updating and managing of the data policy. The director of Luomus will ratify the data policy.

The Biodiversity Informatics Unit (BIU) of Luomus is responsible for maintaining this data policy, and the policy is managed as a part of the enterprise architecture of FinBIF [9]. The BIU will annually review the need to update the policy. During update reviews, the Unit may direct enquiries regarding the policy update needs to relevant parties. In addition, the Unit will evaluate the compatibility of the policy with corresponding partner policies.

Implementing the data policy

Responsibilities for implementing the data policy of FinBIF are shared as follows:

- The head of BIU of Luomus supervises the implementation of the data policy as regards datasets managed by FinBIF.
- BIU is responsible for implementing the data policy in FinBIF data systems for all dataset types.
- For their part, all members of the Luomus' staff and other persons in any contractual relationship with the institution are responsible for implementing the data policy when acting on behalf of FinBIF and producing, distributing or using datasets covered by the policy.

REFERENCES

[1] Sensitive data refers to data that the public is restricted from. Access to this data is limited to certain individuals, groups or organisations only. Sensitive classification is based on the assumption that publishing the data in question may harm that which it concerns.

[2] In addition to sensitivity, datasets may be restricted on other grounds that are usually related to the ownership and management of the data, such as exclusivity based on research or business development.

[3] JHS public administration recommendation 189

[4] <https://creativecommons.org/licenses/by/4.0/deed.en>

[5] <https://creativecommons.org/licenses/?lang=en>

[6] Act on the Openness of Government Activities (1999/621); Nature Conservation Act (1996/1096), Chapter 6, Protection of species

[7] The Finnish environmental administration consists of the [Ministry of the environment](#) and the areas of responsibility for the environment and natural resources administered by the regional Centres for Economic Development, Transport and the Environment under the purview of the [Ministry of Economic Affairs and Employment](#), and the areas of responsibility for environmental permits administered by the Regional State Administrative Agencies, as well as the [Finnish Environment Institute](#), and the Housing Finance and Development Centre of Finland. Responsibility for natural resources administration is shared by the Ministry of the Environment, the Ministry of Agriculture and Forestry and the Ministry of Economic Affairs and Employment.

[8] As a result of an inclusive process, a document "[The Finnish Biodiversity Information Facility – list of sensitive species data \(in Finnish\)](#)" was completed and distributed to all participating parties on 29 July 2016. A preparatory meeting was arranged on 29 September 2015 to reach a decision in principle, with the Ministry of the Environment, the Ministry of Agriculture and Forestry, the Finnish Environment Institute and the Finnish Museum of National History in attendance.

[9] Enterprise architecture refers to the modelling, describing and planning of operations, data requirements, data systems and technological solutions in accordance with a consistent model. In the planning of its enterprise architecture, FinBIF observes the JHS public administration recommendation 179 when applicable.