CoreTrustSeal Application – Finnish Biodiversity Information Facility (FinBIF)

April 2024

R0. Background Information & Context

(1) Re3data Identifier

https://www.re3data.org/repository/r3d100013555

(2) Repository type

Specialist repository: biodiversity informatics.

(3) Overview

The Finnish Biodiversity Information Facility (FinBIF) serves as a comprehensive repository for biodiversity data. Our repository hosts a wide range of data types including a) verbal species descriptions complemented by multimedia, b) master taxonomy data, c) digitized scientific collection specimens, and opportunistic as well as systematic event-based nature observations (also known as biodiversity occurrence data/records), and d) metadata about occurrence datasets. The host institution of FinBIF is Finnish Museum of Natural History 'Luomus', an independent research institution under the University of Helsinki.

Currently FinBIF hosts information about c. 47.000 species and 50 million occurrence records divided into 540 datasets. Most of the data deposited to FinBIF are occurrence data. FinBIF handles two kinds of occurrence data: **a) primary data** which is maintained using FinBIF services, and **b) secondary data** which is maintained elsewhere but copied to FinBIF for publication and reuse.

FinBIF's services coalesce under a unified IT architecture, delivered via a multilingual online open data portal, Laji.fi, which is accessible in Finnish, Swedish, and English. We facilitate restricted-use data access for public authorities through a specialized portal. Our open data also integrates with international systems such as the Global Biodiversity Information Facility (GBIF).

FinBIF is designed to be both expansive and integrative, combining original data generation processes like digitization of collections and collecting new occurrence data with the harmonization, curation and dissemination of existing data provided by other parties. All our occurrence data is harmonized into shared data format so that it is easy to understand and use by a variety of data users (researchers, authorities, educators, amateurs etc.)

We are committed to a high level of integrity and trustworthiness in managing biodiversity data, ensuring its quality and accessibility for a broad range of users, including researchers, public authorities, amateur nature observers and the public. Our work has been documented by Schulman et al. (2021).

References:

- Laji.fi portal, https://laji.fi
- Schulman, L., Lahti, K., Piirainen, E. *et al.* The Finnish Biodiversity Information Facility as a best-practice model for biodiversity data infrastructures. *Sci Data* **8**, 137 (2021). https://doi.org/10.1038/s41597-021-00919-6
- Finnish Museum of Natural History Luomus, https://luomus.fi/en

(4) Designated Community

FinBIF serves a diverse and growing biodiversity informatics community, which is indicative of the repository's value and broad applicability. Our primary Designated Communities and their distinct needs and preferences are:

- **Researchers and Collection Curators**: They value comprehensive data types, advanced search capabilities, and data management tools. An R-package specifically designed for researchers is openly available.
- **Public Authorities**: Use FinBIF for data management related to threatened, protected and invasive species. They require secure, restricted-access portals and data integrity.
- **Education Sector**: Utilizes FinBIF's resources for both formal education and research projects, valuing ease of use and educational support materials.
- **Citizen Scientists**: Participate in systematic species surveys and opportunistic observations. They benefit from user-friendly interfaces for data recording, access and sharing.

While we emphasize development projects that come with resources and commitments, our project-based orientation does not compromise the needs of these varied priority user groups. Most development tasks are designed to benefit all users in a cross-cutting manner.

To meet the varied needs, we invest in good communication, documentation and instructions, and specific training through workshops and webinars.

We remain committed to serving our Designated Community through agile development, as it allows us to adapt and respond to their evolving needs.

References

- Publications and webinars (in Finnish): https://laji.fi/about/1952
- Instructions on Laji.fi (in Finnish): https://laji.fi/about/5699
- R-package: https://laji.fi/en/about/3123c1

(5) Levels of Curation.

FinBIF has a dual data management approach:

1. **Primary Data**: This constitutes about 80% of our repository and is collected or digitized using systems directly maintained by FinBIF and Luomus. The data is kept in its original format upon recording, ensuring authenticity and integrity, as well as harmonized.
2. **Secondary Data**: The remaining 20% of our collection comprises data sourced from systems outside of our direct maintenance, such as those managed by our partner organizations and researchers. This data undergoes a harmonization process, where either all, or only portion of the data most relevant to our designated community are ingested and stored.

All our data undergoes harmonization to a uniform format, tailored for our designated community. This practice ensures consistent data usability across varied sources. Furthermore, we retain custom data fields from most sources (e.g. research projects) and make them accessible, giving our users a comprehensive view of the information.

Levels of Curation Offered by FinBIF

- **A. Content Distributed as Deposited**: We have 0% data in this category, as our primary objective remains to be harmonization, ensuring data remains accessible and usable in the long run.
- **B. Basic Curation**: 0% of our data.
- **C. Enhanced Curation**: Over 99% of our data. It undergoes a harmonization process, which includes conversion into a shared format, and is supplemented with structured metadata. This format is tailored to suit our users, with consistent vocabularies. Each dataset undergoes a basic automated validation, the specifics of which might vary based on the source.
- **D. Data-Level Curation**: The remaining less than 1% has undergone data-level curation. Building upon the enhanced curation process, this data receives additional annotations. Authorized expert users conduct quality control, ensuring data integrity and correctness.

Feedback Mechanism and Continuous Improvement

- **User Feedback**: We operate as a user-focused repository. Feedback is actively sought from users, which informs our decisions to adjust vocabularies and improve data structures. The aim is to adapt to the changing requirements of our designated community.
- **User Annotations**: We provide users with the ability to annotate records. Users can add structured comments or quality classifications to the data, offering more detailed information for subsequent users.
- **Automated Validation Enhancements:** We are currently enhancing our automated validation processes by incorporating a rule-based system. This system is informed by expert knowledge to facilitate accurate data validation.

(6) Cooperation and outsourcing to third parties, partners and host organisations.

The host institution of FinBIF is Finnish Museum of Natural History 'Luomus', an independent research institution under the University of Helsinki, which has led its development, acquisition of funding, and managed national and international cooperation. However, a wide national collaborative network has been an integral part of FinBIF from the start, and many organisations from all sectors of society have participated in the development of FinBIF's services and in data mobilisation. Largest of these partners are the Finnish Environment Institute (Syke), Natural Resources Institute Finland (Luke) and Parks & Wildlife Finland (Metsähallitus Luontopalvelut).

There are several contractual collaborations with relevant stakeholders:

Service Providers

FinBIF's data and services are hosted by CSC - IT Center for Science and Helsinki University IT Center. Partnership with CSC is formalized through agreements. CSC has ISO 27001 information security certificate and has service level agreements for its services.

Services by the Helsinki University IT Center are hosted by our parent organization. FinBIF retains the ultimate responsibility for data preservation planning.

Public Authorities Collaboration

FinBIF's works in collaboration with public authorities, including two Finnish ministries (Ministry of Agriculture and Forestry, and Ministry of Environment) and related research institutes. This collaboration is

formalized through two signed contractual agreements, resulting in the launch of the Public Authority Portal in June 2020.

An agreement has been signed between the most important government biodiversity data managers and providers: Finnish Environment Institute (Syke), Natural Resources Institute (Luke) and Parks & Wildlife Finland (Metsähallitus Luontopalvelut). This agreement includes the principles of data management and sharing of biodiversity data with FinBIF.

Another agreement is signed by all public authorities that are eligible to use the Public Authority Portal (PAP) for an open access to all data governed by FinBIF including sensitive and restricted-use data. This includes for example Finnish municipalities.

Collection Management Collaboration

FinBIF has also developed Kotka, a national multi-purpose collection management system used by natural history museums and botanic gardens in Finland. This collaboration is also based on identical agreements signed between the institutes and Luomus.

Citizen Science and NGO's

FinBIF collaborates closely with NGO nature societies and other organisations that provide valuable information through citizen science. This is done mainly by campaigns (e.g., Atlas of Finnish Fungi funded by a private foundation). NGO's may also collaborate with FinBIF by using custom-made data entry forms at the observation services of FinBIF.

iNaturalist Network

FinBIF is a member of the iNaturalist Network through contractual agreement and supports a localized portal, iNaturalist Finland. Data sourced through this portal are integrated into FinBIF's total data mass.

Global Biodiversity Information Facility (GBIF)

FinBIF is the official national node of the Global Biodiversity Information Facility (GBIF), which is the organization responsible for sharing biodiversity data internationally. FinBIF shares its biodiversity data to GBIF using an automated API. Data is also read from GBIF into FinBIF's data warehouse and updated on a regular basis (e.g., eBird occurrence data recorded from Finland).

References

- Introduction to FinBIF Public Authority Portal: https://laji.fi/en/about/5633
- CSC ISO 27001 certification: https://www.csc.fi/en/-/iso-27001-sertifikaatti-nyt-myos-csc-n-ict-alustoille
- SLA for CSC Rahti Container Cloud Service: https://rahti.csc.fi/agreements/sla/
- SLA for CSC cPouta Community Cloud Service: https://research.csc.fi/documents/48467/0/cPouta+SLA.pdf/144785b7-4571-4ab0-92a1-832e01ae1fae?t=1587646455277
- iNaturalist Finland: https://inaturalist.laji.fi
- GBIF: https://www.gbif.org/what-is-gbif

Organisational Infrastructure

Mission & Scope (R01)
Self-Assessed Compliance Level: Implemented

Mission: The mission of the Finnish Biodiversity Information Facility (FinBIF) is to actively compile, manage, and share biodiversity data, particularly biodiversity occurrence data. FinBIF is committed to long-term preservation and accessibility of this data, aligning with FAIR principles of data management—Findable, Accessible, Interoperable, Reusable. We place an emphasis on data quality, incorporating methods to ascertain and improve accuracy.

Scope: FinBIF focuses on setting national standards for biodiversity data management and taxonomy, thereby enhancing the findability and accessibility of biodiversity data. Our data repository accommodates both open-access and restricted-use data, ensuring responsible data sharing. Our focus is on data from Finland and data collected by Finnish institutions.

Approval and Mandate

1. Finland's National Biodiversity Strategy and Action Plan: Aligns with and supports our mission.
2. State Funding from the Ministry of Environment: Provided initial funds of approx. 2.4 million euros (2015-2018) to establish FinBIF.
3. FinBIF FIRI -projects: Three construction cycles of FinBIF as a research infrastructure, with funding from the Academy of Finland.
4. Roadmap for Finnish Research Infrastructures 2021–2024: FinBIF is accepted on the national research infrastructure roadmap.
5. Public Authority Agreements: Authorization for sharing restricted-use data through a secure portal. Alltogether more than 150 agreements with different public authority organizations, like municipalities and national agencies.
6. Nature Conservation Act (NCA) Legislation: Recognizes FinBIF (Suomen Lajtietokeskus) as an integral part of the national nature conservation IT-infrastructure.
7. Govermental bill on the new Nature Conservation Act defines the scope of FinBIFs operations.
8. University of Helsinki Strategic Plan 2021–2030: FinBIF's implementation aligns with the strategic plan, particularly emphasizing open science, FAIR principles, and accessible research infrastructures.

References

- FinBIF Mission Statement: https://laji.fi/en/about/2954
- National Biodiversity Strategy and Action Plan: https://www.biodiversity.fi/actionplan/action-plan/summary-of-actions
  - Archived at https://web.archive.org/web/20210615162955/https://www.biodiversity.fi/actionplan/action-plan/summary-of-actions
- Roadmap for Finnish Research Infrastructures 2021–2024: https://www.aka.fi/globalassets/1-tutkimusrahoitus/4-ohjelmat-ja-muut-rahoitusmuodot/4-tutkimusinfrastruktuurit/roadmap-for-finnish-research-infrastructures-20212024.pdf
- Nature Conservation Act Legislation 9/2023 (in Finnish): https://www.finlex.fi/fi/laki/alkup/2023/20230009

- Govermental bill on the new Nature Conservation Act (in Finnish): https://www.eduskunta.fi/FI/vaski/HallituksenEsitys/Sivut/HE_76+2022.aspx
- University of Helsinki Strategic Plan 2021–2030: https://www.helsinki.fi/en/about-us/strategy-economy-and-quality/strategic-plan-2021-2030/strategic-plan-of-the-university-of-helsinki

Rights Management (R02)

Self-Assessed Compliance Level: Implemented

FinBIF service manages and shares all data and supporting documentation in accordance with the FAIR principle "as open as possible, as closed as necessary". FinBIF employs a systematic approach to rights management that ensures all data, media and metadata rights are respected. Data deposition to FinBIF it is done with an agreement with the data owners, governing the terms of storage, handling, and sharing, aligned with the FinBIF data policy. These agreements are flexible to accommodate exceptions mutually agreed upon by FinBIF and the data providers.

Licenses for Data and Metadata

FinBIF defaults to the Creative Commons Attribution 4.0 International license for all biodiversity occurrence data to promote open access while ensuring proper attribution to the data providers. Alternative licensing arrangements, including different Creative Commons licenses or "All rights reserved," are also available based on the depositor's preference.

Metadata is licensed under Creative Commons Zero to enhance usability and facilitate integration into various research initiatives.

Creators of media files linked to occurrence data can independently select from a range of Creative Commons licenses or opt for "all rights reserved," ensuring a choice in the level of rights management.

Depositor Agreements

Agreements with data depositors' outline:

- Dataset metadata.
- Intellectual property rights, retained by the original data owner.
- Licensing terms for data usage.
- Citation instructions.
- Responsibilities for maintaining the dataset's primary copy.
- Procedures for dataset updates.
- Restrictions on data use, such as research embargoes.
- Access protocols for restricted or sensitive data.
- Sharing data with other biodiversity repositories.

Specific Agreements

- **Major Biodiversity Data Owners:** Written agreements with key Finnish biodiversity institutions delineate the terms of data sharing, usage restrictions, and further dissemination rights.

- **Public Authorities:** Agreements permit Finnish public authorities to utilize all FinBIF data for official purposes, subject to stringent use and protection conditions. Access is limited to authorized personnel only, and usage is systematically logged.
- **Natural History Museums:** The Kotka collection management system's utilization is formalized with agreements outlining data ownership, access, and openness principles.

Terms of Use for Individual Contributors

The Notebook observation service, utilized by individuals for submitting observations, is governed by its terms of use, asserting that all public data is subject to the Creative Commons Attribution 4.0 International license. Users can opt to anonymize personal information and restrict access to sensitive data. FinBIF reserves the right to selectively disseminate such data to support e.g. biodiversity research and conservation.

Data Publication, Citation and Protection

Metadata and all individual data records incorporate the license information in both human-readable and machine-readable formats. Additionally, citable downloads offer guidelines on data citation.

Examples of how data is presented:

- Link to search:
  https://laji.fi/en/observation/map?target=MX.36235&finnishMunicipalityId=ML.365&time=2023-01-01%2F&loadedSameOrBefore=2023-10-18
- Link to citable download landing page with basic information, citation guidelines and data:
  http://tun.fi/HBF.80116?locale=en
- Link to metadata of one of the datasets (collections) within the search:
  http://tun.fi/HR.4471
- Link to one individual data record of the search download:
  http://tun.fi/JX.1639909#23

Sensitive data is safeguarded according to guidelines established in collaboration with Finnish authorities and biodiversity specialists.

Restricted Data Usage Requests

Anyone can request access to restricted data e.g. for research purposes. These requests are logged, documenting the applicant's details and intended use. The decision to grant access to sensitive data resides with the data owner and is considered on a case-by-case basis. Redistribution of such restricted data by third parties is expressly forbidden.

Monitoring and Compliance

FinBIF's policy does not include active monitoring of data use post-access; responsibility for addressing instances of non-compliance lies with the data owners and users.

Ethical Considerations and Privacy Practices

FinBIF adheres to ethical guidelines in data management, particularly regarding the dissemination of sensitive data and the privacy of users. Data ownership is respected, and the transfer of data to FinBIF does not affect ownership rights. Practices around anonymization and privacy protection of data contributors are integral, ensuring that personal information is handled with care, especially in the context of sensitive data.

Additionally, FinBIF staff receive training on managing sensitive species data. These measures ensure a comprehensive approach to rights management, balancing legal, ethical, and privacy considerations.

References

- FinBIF Data Policy: https://cms.laji.fi/wp-content/uploads/2020/01/Data-policy-of-the-Finnish-Biodiversity-Information-Facility_Final-2020.pdf
- Principles of handling sensitive data: https://laji.fi/en/about/875
- Terms of Use for Notebook Service (in Finnish): https://laji.fi/about/5809
- Specific agreements are confidential and available upon request, detailing the terms of data sharing and usage with major Finnish biodiversity stakeholders and public authorities.

Continuity of Service (R03)

Self-Assessed Compliance Level: In Progress

FinBIF and the Finnish Museum of Natural History Luomus, stands on a robust framework for biodiversity data preservation, strongly supported by both legislative mandates and a wide societal endorsement. Anchored by the Universities Act, Section 72, Luomus is entrusted with national collections and is pivotal to Finland's biodiversity informatics landscape. The inception of FinBIF, mandated at a seminal 2012 cross-sectoral assembly and underscored in Finland's Biodiversity Strategy and Action Plan (action 40), reflects a concerted national effort towards shared data utility.

A pivotal aspect of FinBIF's structural stability and operational continuity is its foundational support that spans the administrative domains of three key Finnish ministries: the Ministry of Education and Culture, the Ministry of the Environment, and the Ministry of Agriculture and Forestry.

The evolution of FinBIF, has been marked by strategic development and funding procurement, within national and international cooperation. This cooperative ethos is exemplified by the establishment of seven advisory and co-creation groups, ensuring a participatory development process. The diverse funding streams, ranging from EU infrastructure funds to private-sector partnerships, attest to FinBIF's dynamic funding model and societal value. This multifaceted funding approach facilitates ongoing development and serves as a buffer against financial contingencies.

The Reformed Finnish Nature Conservation Act (2023; paragraph 117 §) which defines FinBIF as part of Finland's nature data management system, forms the legal foundation of Luomus's and FinBIF. The repository's services are thus sustained by a stable legislative and collaborative structure.

The solid societal and institutional foundations of FinBIF, coupled with its expansive support network (see partner organizations on section R0.6), mitigate the risks associated with unforeseen circumstances, safeguarding the continuity of its vital services to depositors and users.

Ensuring continuity

FinBIF acknowledges the necessity of sustained vigilance in risk management. Our IT systems and associated risks are documented in a system provided by the University of Helsinki. We are creating a new methods to assess technical risks, which will work alongside our current automated system checks. Maintenance responsibilities are allocated on a daily rotation. This ensures constant oversight and immediate response if needed.

The agile software development methodology adopted by FinBIF enhances its responsiveness to rapidly evolving technical or organizational landscapes. This agility ensures that the services can adapt effectively to unforeseen changes while maintaining data integrity and access.

Functions and services offered

FinBIF handles two kinds of data: **a) primary data** which is maintained using FinBIF services, and **b) secondary data** which is maintained elsewhere but copied to FinBIF for publication and reuse.

FinBIF operates a "Data Bank" for deposition, storage, management, sharing and long-term preservation of primary biodiversity occurrence data. The "Kotka" collection management system complements this by managing natural history museum collections' data. "Notebook" observation system provides tools for Citizen Science data and projects. Other primary data are stored within various systems, e.g. the taxonomy database for taxonomical data and names or organisms.

Secondary data management is facilitated through the "Data Bank" and custom application interfaces (API's) for data sharing via the FinBIF data portal, Laji.fi. While responsibility for long-term preservation of this data rests with the original owners, FinBIF ensures access and availability.

Luomus, fortified by its mission and institutional mandate, guarantees the continuity of service, irrespective of external funding scenarios. The stewardship of biodiversity data is ingrained in its operational ethos. While there is no formal succession plan or written agreement with a potential successor, the robust nature of Luomus's institutional framework and its foundational role in national biodiversity informatics render the prospect of cessation remote. This assurance is based on Luomus's statutory responsibilities, diversified funding structure, and the integral nature of its services to the Finnish biodiversity community.

In conclusion, while the formalization of a succession plan is recognized as an area marked "In Progress," the existing measures and the strong institutional commitment of Luomus to its mission substantially mitigate the risks associated with service continuity. In a crisis situation FinBIF responsibilities could already be transferred to our partners.

The current strategic approach is to bolster the existing strong foundations with our partners while continuing to develop a formalized continuity and succession plan and risk assessment plan within the next three years that will elevate the repository's compliance status to "Implemented."

References:

- Finnish Museum of Natural History Luomus: https://luomus.fi/en
- Universities Act, Section 72: https://www.finlex.fi/fi/laki/kaannokset/2009/en20090558_20160644.pdf
- Finland's Biodiversity Strategy and Action Plan action 40: https://www.biodiversity.fi/actionplan/action-by-category/cross-cutting-issues/monitoring-and-research/joint-use-of-data
  - Archived at https://web.archive.org/web/20210615162707/https://www.biodiversity.fi/actionplan/action-by-category/cross-cutting-issues/monitoring-and-research/joint-use-of-data
- Advisory and co-creation groups of FinBIF: https://www.nature.com/articles/s41597-021-00919-6/tables/4
- Reformed Nature Conservation Act: https://ym.fi/en/reform-of-the-nature-conservation-act & https://www.finlex.fi/fi/laki/alkup/2023/20230009

Legal & Ethical (R04)

Self-Assessed Compliance Level: Implemented

FinBIF, as a national service, with numerous data providers, has been under thorough scrutiny by the partners to create the trust. The service has been built from the start to handle all kinds of biodiversity data including sensitive, restricted-use and embargoed data. There are a number of technical and policy-level measures in place to guarantee the confidentiality of information stored in FinBIF.

Identification and Management of Legal and Ethical Standards

FinBIF has identified five major legal and ethical issues in data management with the following guidelines in place:

1) Transparency and Openness in Public Agency Operations

We adhere to the principle of "as open as possible, as closed as necessary," aligning with the FAIR data principles. While prioritizing data findability and accessibility, we respect legal and contractual limitations, providing a structured service for access requests to restricted information (Restricted Data Request Service). Our openness policy and methods are detailed in the article "As Open as Possible, as Closed as Necessary – Managing legal and owner-defined restrictions to openness of biodiversity data". Data download procedures include clear guidelines on disclosure and use, ensuring informed and responsible data handling by users.

2) Protection of Endangered Species Data

In compliance with the Act on the Openness of Government Activities (section 24, item 14), we take special care in handling data on endangered species and habitats. We have an automated sensitive data protection process defined together with a multi-stakeholder working group of Finnish authorities and biodiversity institutions. This ensures the protection of species and habitats while allowing data accessibility under controlled conditions. Additionally, data owners can obscure their observations for various reasons, including research embargoes.

3) Privacy and Personal Data Protection

We prohibit the unauthorized entry of personal details and ensure compliance with Finnish law and GDPR. Our privacy policy governs the handling of personal data, giving data owners the ability to obscure personal information in their data.

4) Data Quality and Accessibility

FinBIF acknowledges the importance of every observation record. We provide tools to access data based on relevance and reliability, including standardized metadata. Quality classifications, annotations, and evidence-based identifications enhance the reliability of our data. Our commitment to data quality and fitness-for-use is exemplified by our accessible services, including our API and R-package.

5) Case-Specific Data Use Agreements

We respect the wishes of data owners regarding data openness. Transferring or sharing data to FinBIF does not affect data ownership: data are owned by the organization or individual who has shared it for FinBIF.

The sharing rights for restricted datasets are negotiated individually, with the aim of ensuring authority access for official duties, and data use is governed by mutual agreements or relevant documentation.

Management of Data with Disclosure Risks

The measures in place at FinBIF demonstrate a comprehensive approach to managing data with disclosure risks. Our technical and policy-level safeguards are tailored to ensure the confidentiality and ethical use of sensitive biodiversity data.

1) Data Segregation in FinBIF Data Warehouse

Our data storage strategy is designed to mitigate risks associated with sensitive data. The FinBIF Data Warehouse segregates data into two distinct databases: one for publicly accessible data and another for all data, including sensitive, restricted-use, and embargoed data. This separation ensures that sensitive data is adequately protected and not inadvertently disclosed.

2) Access Control and Usage Agreements

Access to sensitive data is strictly limited to a select number of employees from Finnish public authorities. These authorities are bound by agreements that dictate responsible data usage, focusing on protecting both species and personal privacy. The agreements explicitly prohibit the sharing of sensitive data.

3) Case-by-Case Data Release Mechanism

Data owners or their appointed representatives have the discretion to release data on a case-by-case basis, often for specific purposes such as research or land-use planning, using the Restricted Data Request Service. Users who access this data are subject to strict terms of use. These terms mandate responsible use, prohibit unauthorized sharing or publishing, and restrict usage to the purposes specified in the data request.

Training and Guidance on Responsible Data Handling

FinBIF and Luomus staff are trained in the risks of releasing information about sensitive species as a part of our induction process.

References

- Restricted Data Request Service (in Finnish): https://laji.fi/about/1890
- 'As Open as Possible, as Closed as Necessary' – Managing legal and owner-defined restrictions to openness of biodiversity data: https://biss.pensoft.net/article/37395/
- Act on the Openness of Government Activities, section 24, item 14: https://www.finlex.fi/en/laki/kaannokset/1999/en19990621_20150907.pdf
- Sensitive data protection implementation process: https://laji.fi/en/about/875
- Quality control and annotation process description: https://laji.fi/en/about/877
- FinBIF API: https://api.laji.fi
- FinBIF R-package: https://laji.fi/en/about/3123c1
- FinBIF Privacy policy: https://laji.fi/en/about/848

Governance & Resources (R05)

Self-Assessed Compliance Level: Implemented

FinBIF, housed within the Finnish Museum of Natural History Luomus, part of the University of Helsinki, demonstrates a robust governance and resource allocation framework, essential for its mission in biodiversity data management. Establishment of FinBIF is also emphasised in Finland's Biodiversity Strategy and Action Plan (action 40).

Governance Structure

FinBIF's governance structure is clearly delineated, reflecting a hierarchical organization. The governance includes:

- Rector of the University of Helsinki.
- Dean of Faculty of Biological and Environmental Sciences.
- Director of Luomus as an independent Research Institution.
- Advisory steering group for FinBIF was formed in March 2024, led by the Finnish Prime Minister's Office.
- Unit Director of Biodiversity Informatics Unit, i.e. Director of FinBIF.
- Team Leaders for the ICT, Species Information, and Digitisation teams.
- Autonomous governance in FinBIF partner organisations

This structure ensures a streamlined decision-making process and accountability at each level, as detailed in the organizational diagram (https://laji.fi/en/about/2982, the new steering group missing here)

Staffing and Expertise

FinBIF employs 23 full-time staff (including data digitizers and researchers), with approximately half under permanent contracts, indicating a stable workforce. The staff is divided into specialized teams, ensuring focused expertise in each operational area. The up-to-date list of employees is available on the Luomus website, providing transparency in staffing.

Funding and Financial Sustainability

Funding for permanent staff salaries is sourced from the core funds of Luomus (University of Helsinki), signifying long-term financial support. Temporary staff funding is secured through various internal and external projects, showcasing the ability to leverage additional resources.

References:

- The organization of FinBIF: https://laji.fi/en/about/2982
- Biodiversity informatics Unit staff: https://www.luomus.fi/en/biodiversity-informatics-unit
- Finland's Biodiversity Strategy and Action Plan (action 40): https://www.biodiversity.fi/actionplan/action-by-category/cross-cutting-issues/monitoring-and-research/joint-use-of-data
  - Archived at https://web.archive.org/web/20210615162707/https://www.biodiversity.fi/actionplan/action-by-category/cross-cutting-issues/monitoring-and-research/joint-use-of-data

Expertise & Guidance (R06)

Self-Assessed Compliance Level: Implemented

FinBIF has a robust approach to maintaining and enhancing expertise and guidance. The repository has implemented several mechanisms to ensure that it possesses the necessary skills and knowledge, both internally and through external engagement, aligning with the needs of its Designated Community in biodiversity informatics.

In-House Development and User Interaction

FinBIF's in-house ICT development team (developers and computer scientists) together with species information team (biologists and a GIS specialist) engages in agile development, maintaining continuous interaction with service users. This is facilitated through regular webinars and training sessions, enabling the team to understand and respond to user challenges effectively.

Expert Advisory Groups

From its inception, FinBIF established seven expert advisory groups for co-creation of content, services, and policies. These groups, such as the Public Authority Expert Group, meet regularly or on demand, ensuring that the repository's services and policies are informed by up-to-date expert advice.

National and International Collaboration

FinBIF actively participates in national working groups and international collaborations like DiSSCo European RI (DiSSCo) and Global Biodiversity Information Facility (GBIF) governing boards. This broad engagement ensures that FinBIF remains aligned with global standards and practices in biodiversity informatics.

User Feedback Mechanisms

Users can provide feedback and suggestions via a feedback form (accessible from Laji.fi), email (helpdesk@laji.fi) or the Laji.fi Forum (https://foorumi.laji.fi/).

Annotation System

FinBIF's annotation system allows any user to comment on data, with a pool of experts available for expert annotations. This interactive feature ensures data quality and user engagement.

Continuous Professional Development

Staff development is prioritized, with IT staff receiving training from the University of Helsinki and external sources. Participation in international conferences and projects, such as Biodiversity Information Standards (TDWG), Biodiversity Next and ICEDIG, underscores a commitment to professional growth and staying current with sector developments.

References

- FinBIF webinars (in Finnish): https://www.helsinki.fi/en/news/unitube?search=lajitietokeskus
- FinBIF expert advisory groups: https://www.nature.com/articles/s41597-021-00919-6/tables/4
- FinBIF annotation system (in Finnish): https://laji.fi/about/3806

Digital Object Management

Provenance and authenticity (R07)

Self-Assessed Compliance Level: Implemented

The FinBIF repository actively maintains the authenticity and provenance of its digital objects. This is achieved through a combination of data and metadata management systems, user authentication processes, and detailed versioning of all primary data.

Data and Metadata Management

FinBIF employs diverse procedures for managing digital objects, tailored to each data type's unique requirements. For primary occurrence data records and dataset metadata, FinBIF maintains a comprehensive version history. This history includes information about user identity for each change, except for system admin modifications (rarely done) where user identity is not recorded. Secondary copies of data from partner databases and media objects (images, audio) do not have a version history.

Authenticity through User Authentication

User authentication and authorization are central to FinBIF's operations. Users are provided accounts with specific roles and permissions, ensuring that only authorized changes to digital objects are made. While the identity of citizen science service users is not always known, email verification ensures a basic level of authenticity.

Data Integrity and Provenance

For data integrity, FinBIF utilizes checksums for media objects and relies on the inherent reliability of ACID-compliant databases. Our IT infrastructure, built on API access-token systems, further secures data transfers between systems. Each digital object is assigned a unique URI identifier, ensuring traceability and provenance. In cases where partners restrict data sharing, FinBIF provides a mechanism to indicate limited data availability, maintaining transparency about the data's scope and origin.

References

- FinBIF Enterprise Architecture Document (in Finnish): https://laji.fi/about/1357

Deposit & Appraisal (R08)

Self-Assessed Compliance Level: Implemented

FinBIF Data Policy describes what kind of data we manage: species occurrence data from Finland or collected by Finnish organizations, and information about Finnish species. FinBIF adheres to a structured deposit process, ensuring that all data follows our data model and metadata standards. Deposition of data in other proprietary formats is not possible. This ensures all data is interoperable and reusable.

FinBIF employs various methods for data collection. These methods include:

- ETL (extract-transform-load) processes data imports from large databases via an API, created in collaboration with data owners.
- Tools for submitting occurrence records, for
    - citizen science participants (Notebook observation system)
    - professionals (FinBIF Data Bank)
    - Museum and gardens (Kotka collection management system)

By controlling these tools, FinBIF ensures adherence to our standards. Depositors cannot bypass these methods, upholding our deposit integrity. Data not fitting the required standards or formats are not accepted. For datasets of significant value but in obsolete formats, FinBIF offers tailored migration solutions. However, the responsibility for resource allocation for such migrations lies with the data owner.

The provision of structured metadata is mandatory prior to data upload. Data is compatible with domain standards (see R09, R10 & R13).

The transfer of data to the FinBIF repository is governed by formal agreements, established between the data provider and Luomus. These agreements depend on the data exchange method. For ongoing data transfers, such as API uploads, we typically establish a bilateral agreement, clearly outlining the principles and methods of data transfer. Conversely, for one-time data uploads, the agreement may be more flexible, ranging from detailed terms to basic agreements conducted via email. This structured approach ensures clarity and mutual understanding between FinBIF and data providers, aligning with our commitment to maintaining data integrity and trust.

References

- FinBIF Data Policy: https://laji.fi/en/about/2982
- Instructions on deposition of data to FinBIF (in Finnish): https://laji.fi/about/1289

Preservation plan (R09)
Self-Assessed Compliance Level: Implemented

FinBIF is committed to the long-term preservation of species information and biodiversity occurrence data as a core aspect of its mission. The FinBIF data policy, reviewed regularly, delineates the responsibilities and principles guiding data management, with a focus on long-term preservation. Notably, occurrence data ownership remains with the original creator, and data management responsibilities are defined in agreements or relevant documentation with depositors. This approach ensures clarity in roles and responsibilities for preservation.

Our preservation strategy involves maintaining biodiversity occurrence data in the FinBIF data format, aligned with the international ABCD (Access to Biological Collections Data) standard. This ensures long-term data understandability, interoperability, and accessibility. Metadata adheres to the EML (Ecological Metadata Language) standard, promoting compatibility and standardized data curation. In instances of format changes, FinBIF oversees necessary data migrations.

Primary data responsibility resides with FinBIF, while data owners retain responsibility for secondary copies. Data owners, however, retain control over their data, including the right to request removal. In such cases, citable downloads are preserved perpetually in text format to maintain persistent citations. Identifiers, once assigned, are kept persistent and are never reassigned.

Data is currently hosted and backed up by the University of Helsinki IT services. We are in the process of extending collaboration with CSC's (Finnish IT Center for Science) Finnish National Digital Preservation Service.

References

- FinBIF Data Policy: https://laji.fi/en/about/2982
- CSC Finnish National Digital Preservation Service: https://research.csc.fi/-/digital-preservation-service
- ABCD Standard: https://www.tdwg.org/standards/abcd/
- EML Standard: https://eml.ecoinformatics.org/

Quality Assurance (R10)

Self-Assessed Compliance Level: Implemented

The FinBIF upholds a robust approach to technical quality and standards compliance, ensuring that our users can effectively evaluate the quality of the data they access. This section outlines our methods and standards for quality assurance, focusing on occurrence data and metadata quality during the curation process.

Data and Metadata Quality Assurance

FinBIF enforces a systematic approach to data quality, which varies based on the source and nature of the data. Our quality assurance process includes:

1. Predefined Formats and Automatic Validation: Data must conform to one of several predefined formats for ingestion. We employ automatic validation checks to ensure data completeness and format compliance. Data lacking mandatory fields or in an incorrect format is rejected.
2. Custom Validation Rules: We utilize custom rules for different data types (e.g., bird or butterfly records) to identify and flag geographic inconsistencies.
3. Dataset Metadata: Each dataset includes metadata detailing its quality level (professional, specialist, citizen science), collection methods, and a quality description of any pre-ingestion assurance measures. This metadata, co-developed with data owners and FinBIF staff, aids users in assessing data relevance and quality.

Expert Evaluation

Over 100 volunteer experts across various fields evaluate and annotate individual records, assigning quality ratings (verified, neutral, uncertain, erroneous) or flagging them for verification. This process leverages the expertise of a diverse group of specialists to enhance data reliability.

Importantly, the original data remains untouched during this process. Instead, annotations are attached to the original data, allowing data users to view both the original data and the accompanying validation results and annotations.

Standards Compliance

FinBIF data is compatible and adheres to internationally recognized standards for biodiversity data (Darwin Core, ABCD and EML). Our commitment to standards ensures interoperability and enhances the utility of our data within the global biodiversity informatics community. As technical standards evolve, FinBIF is committed to updating its practices.

Issue Resolution

Our approach to resolving quality issues includes returning problematic digital objects to depositors for rectification before data is ingested, fixing issues internally, applying quality flags, and providing detailed explanations in metadata. This comprehensive strategy ensures transparent and responsible data curation.

Links to Related Data

We provide links to related digital objects (e.g. species and other taxa), publications (DOI's), and utilize controlled vocabularies and ontologies to enhance data discoverability and usability.

References

- Description of FinBIF quality control and expert evaluation process: https://laji.fi/en/about/877

Workflows (R11)
Self-Assessed Compliance Level: In Progress

FinBIF manages occurrence and metadata records with a set of integrated IT systems. These systems ensure adherence to predefined workflows, thereby standardizing the handling of all data objects. This process is automated for efficiency and consistency.

Incorporation of new data sources into FinBIF follows a standardized, internally documented workflow. This process includes:

- Initial contact with the data owner.
- Discussion on data format, rights, and agreements.
- Metadata entry.
- Data conversion to FinBIF format.
- Data entry, including data cleaning (if necessary), based on automatic validation.
- Automatic data publication.
- Post-publication quality control, as elaborated in Requirement R10.

The digitization of old natural history museum collections is governed by its own set of internally documented workflows, ensuring that historic data is accurately and consistently integrated into the repository.

Workflows at FinBIF are flexible and can be adjusted to cater to different data types and provider needs. Decisions within these workflows are made by qualified staff, ensuring expert judgment in workflow adaptation and exception handling.

While the application does not currently detail change management procedures for workflows or mechanisms for tracking workflow execution, this is an area identified for development. A plan is underway to implement a system for further documenting workflows, tracking workflow execution and handling exceptions, and to formalize change management processes for workflow adjustments. This should be completed within the next three years.

Different workflows for varying levels of data security, as mentioned in R04 (Legal and Ethical), are not currently differentiated, because sensitive data is secured automatically. FinBIF will explore the need for and potentially develop distinct workflows for data with different security requirements within the next three years.

More detailed model including the data source visualisation of FinBIF, available as Fig 4 of the referenced article.
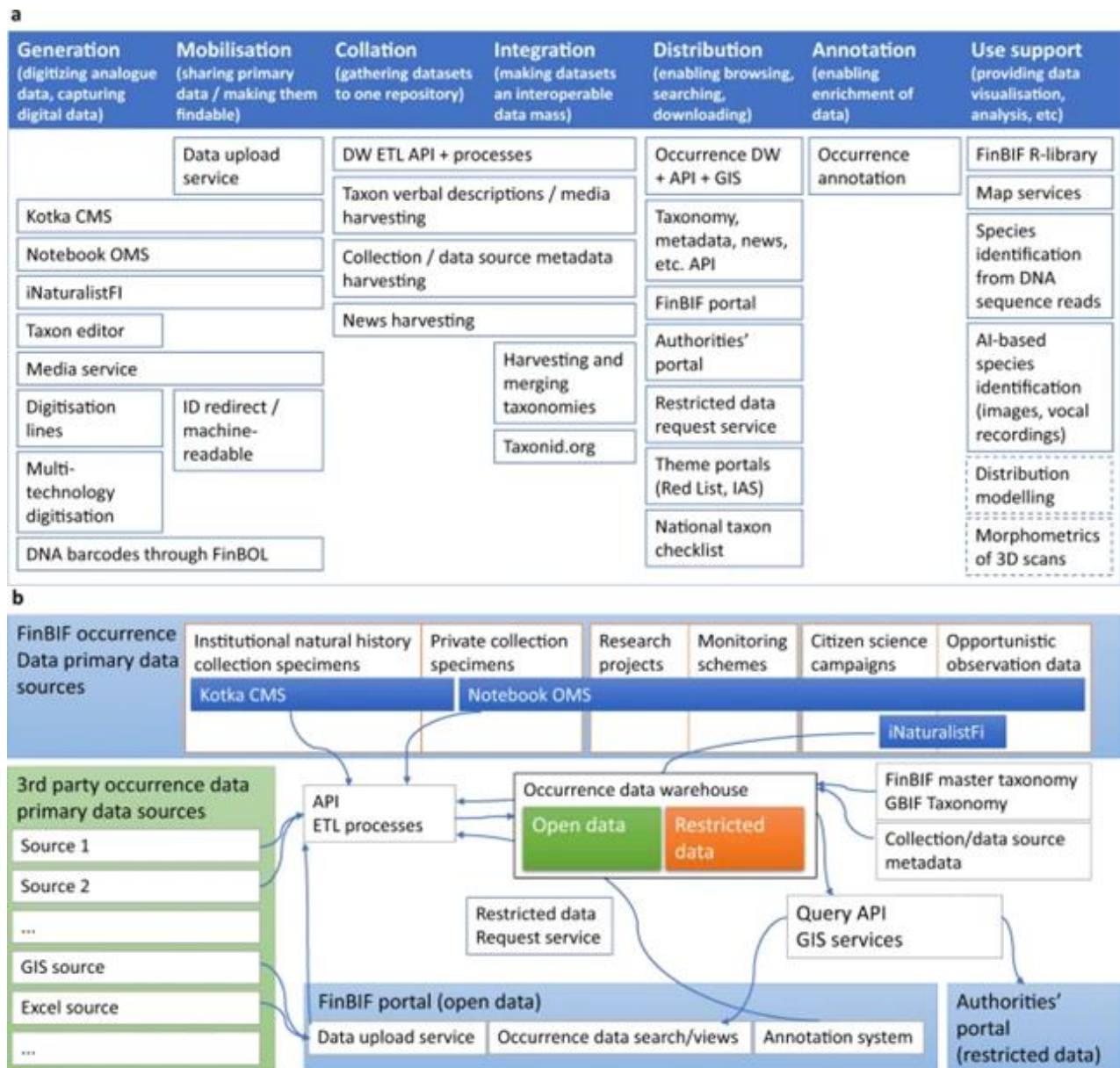
References:

- The Finnish Biodiversity Information Facility as a best-practice model for biodiversity data infrastructures: https://www.nature.com/articles/s41597-021-00919-6

Discovery and Identification (R12)
Self-Assessed Compliance Level: Implemented

FinBIF ensures that users can discover and cite digital objects through a robust infrastructure, utilizing persistent HTTP-URI identifiers, metadata standards, and inclusion in key registries. The repository's approach encompasses a comprehensive search facility, adherence to international metadata standards, unique and persistent identification of objects, machine harvesting of metadata, and the facilitation of recommended data citations.

Search Facilities and Metadata Standards

- All occurrence and taxonomic data are available at Laji.fi portal, with comprehensive search and browsing functionality.
  - Example: Passerine bird observations from Finland in January 2023: https://laji.fi/en/observation/list?target=MX.30559&countryId=ML.206&time=2023-01-01%2F2023-01-31
- Since FinBIF manages the national Kotka collection management system, also the scientific collection data is accessible to all relevant collection holding institutes as the data is originally recorded. The open data of the collections is also available and accessible through the Laji.fi-portal by e.g. selecting Kotka CMS as a data system. (https://laji.fi/en/observation/list?sourceId=KE.3)
- FinBIF is listed in the Registry of Research Data Repositories, Global Biodiversity Information Facility (GBIF) and the national Open Data Portal, enhancing its discoverability.
- Metadata descriptions are shared e.g. to GBIF using Ecological Metadata Language (EML) standard. Metadata is expanded to match new requirements.
- FinBIF provides access through Species.fi and third-party services like GBIF, ensuring data is findable and accessible, with metadata linked to each data record.

Persistent Identifiers

- FinBIF uses HTTP-URI identifiers recommended by the World Wide Web Consortium for all objects, ensuring global uniqueness and persistence. All objects we manage (e.g. datasets, occurrence records, users, taxa, localities) are given a unique and persistent identifier.
  - Examples:
    - Identifier for accessing metadata for 4th Finnish Bird Atlas dataset: http://tun.fi/HR.4471
    - Identifier for accessing an observation: http://tun.fi/JX.1672309
    - Identifier for Chaffinch species: http://tun.fi/MX.36237
- A redirection service is in place for these identifiers, guiding users to relevant information in both human-readable and machine-readable formats, ensuring long-term resolution and accessibility.
- The addition of DOI identifiers is underway for certain data types, such as dataset metadata and natural history collection data, with implementation expected by the end of 2024. This move signifies a commitment to enhancing the persistence and citation integrity of critical data sets.

Machine Harvesting and Interoperability

- Metadata and data are machine-actionable through both FinBIF API (https://api.laji.fi) and machine-readable through HTTP-URI identifiers. Machine-readable format can be accessed by either content negotiation or providing a parameter.

- Example: http://tun.fi/JX.1672309?format=json
- FinBIF offers data compatible with CETAF Specimen Preview Profile (CSPP) and Minimum Information about a Digital Specimen (MIDS) standards in both RDF+XML and JSON-LD formats using CETAF compliant vocabulary This ensures that data is not only discoverable but also interoperable across biodiversity data ecosystems.
- To help the researches' work, an R-library has been developed by us and is freely available here: https://luomus.github.io/finbif/

Registry Inclusion and Data Citation

- Inclusion in re3data.org and linkage to GBIF.org facilitates broader visibility within the research community.
- FinBIF provides clear instructions for citing datasets, whether in part or in whole, ensuring that data use is properly attributed. Data downloads are provided with a persistent HTTP-URI identifier, which leads to the downloaded dataset, metadata and citation instructions.
  - Example: http://tun.fi/HBF.83012

References

- FinBIF on GBIF: https://www.gbif.org/publisher/04fd2e13-6881-4e5c-9dd1-8fdd9ab993c1
- FinBIF on re3data: https://www.re3data.org/repository/r3d100013555
- FinBIF on Finnish open data portal: https://www.avoindata.fi/data/en_GB/dataset/lajitietokeskus
- EML standard: https://eml.ecoinformatics.org
- MIDS, Minimum Information about a Digital Specimen: https://www.tdwg.org/community/cd/mids/
- CSPP, CETAF Specimen Preview Profile: https://cetafidentifiers.biowikifarm.net/wiki/CETAF_Specimen_Preview_Profile_(CSPP)
  - CETAF is the Consortium of European Taxonomic Facilities, of which Luomus is a member. https://cetaf.org
- R Interface for the Finnish Biodiversity Information Facility API: https://zenodo.org/records/10496151

Reuse (R13)

Self-Assessed Compliance Level: Implemented

FinBIF is committed to enabling the effective reuse of its biodiversity occurrence data and associated digital objects over time, thereby ensuring their long-term value to the Designated Community. This commitment is realized through a strategy that encompasses engagement with our designated community, the adoption of widely accepted data formats and standards, and the provision of metadata and documentation.

Engagement with Designated Community

FinBIF actively engages with its designated community, which includes researchers, public authorities, amateur naturalists, and partner organizations such as the Global Biodiversity Information Facility (GBIF) and the Consortium of European Taxonomic Facilities (CETAF). This engagement is facilitated through project and meetings as well as various working groups that focus on research, observation data, and public authority issues. These groups provide valuable input that informs the development and enhancement of FinBIF's services, ensuring they meet the evolving needs of the community.

We are also monitoring and potentially adopting new data formats and standards as they emerge in biodiversity informatics. For example, we are now (3/2024) working on including machine-generated occurrence data and related metadata (e.g. automatic audio recordings).

Data Formats and Standards

To support reuse, FinBIF utilizes a range of data formats and standards that are recognized and valued within the biodiversity informatics community. These include:

- Database Records: Migrated to current RDBMS technologies, ensuring long-term accessibility.
- Media Objects: Stored in TIFF and JPG for images, and WAV, FLAC and MP3 for audio, with considerations for future migrations to open-source formats if necessary.
- Data Accessibility: Offered through UTF-8 TSV text files, Excel, OpenDocument Spreadsheet formats, and a comprehensive API delivering data in JSON, JSON-LD, and XML formats for broader use cases including GIS services.

Metadata and Documentation

To enhance understandability and reuse, each digital object in FinBIF is accompanied by detailed metadata and is accessible via a persistent identifier (see section R12), ensuring stable referencing over time. The FinBIF schema provides descriptions and definitions of data fields, supported by extensive documentation available on our website. This includes guidelines on data formats, such as occurrence data field documentation, ensuring users have the necessary information to effectively utilize the data.

Licensing and Standards Compliance

All digital objects are assigned clear licenses, detailed on our website, promoting transparency and informing users of reuse conditions. FinBIF adheres to global biodiversity informatics standards like the Darwin Core (DwC) and Ecological Metadata Language (EML) for data sharing and interoperability.

Continuous Improvement and Change Management

FinBIF is dedicated to the continuous improvement of its data management practices, including the management of changes to data, metadata, and documentation. This ensures that the repository's offerings remain relevant and valuable to the designated community, addressing their needs for data reuse effectively.

References

- FinBIF Schema: https://schema.laji.fi/
- Occurrence data fields documentation: https://laji.fi/about/2883
- Licensing information: https://laji.fi/en/about/2986

Information Technology & Security

Storage & Integrity (R14)

Self-Assessed Compliance Level: Implemented

FinBIF approach storage and integrity is based on documented processes, strategies for maintaining multiple data copies, risk management techniques, monitoring of storage media, and integrity checks.

Documented Processes

FinBIF stores and harvest various types of data. Overview of different data flows can be found at https://wiki.helsinki.fi/xwiki/bin/view/FinBIF/FinBIF/1%20Overview/. A technical overview of storage locations is documented at https://wiki.helsinki.fi/xwiki/bin/view/FinBIF/FinBIF/1%20Overview/Architecture/. Technical aspects of the different implementations are discussed in chapter R15.

FinBIF staff are guided by internal documentation that describes the management of all storage locations. This documentation covers the operational procedures and responsibilities for ensuring data integrity and accessibility. The primary storage infrastructures are managed in cooperation with the University of Helsinki IT Department and CSC – IT Center for Science, both of which provide professional data center services with data redundancy solutions and disaster plans.

CSC services, cPouta and Rahti provide service-level agreements (SLAs). University of Helsinki does not have a formal SLA and documentation about their services is available in Finnish only and is not publicly accessible.

Multiple Copies Strategy

To mitigate risks of data loss, FinBIF maintains multiple backup copies of data across the two principal infrastructures. This strategy is supported by automated backup processes, including incremental and full backups of primary data and harvested data. The use of RAID (Redundant Array of Independent Disks) technology further enhances data redundancy.

Images and other media data are stored primarily to CSC infrastructure and backed up to the CSC Ida service for safe research data storage as a frozen dataset.

Risk Management Techniques

The reliance on two principal infrastructures with their own disaster plans and the maintenance of test and production environments minimizes risks. Continuous communication with infrastructure providers and annual needs assessment meetings with the University of Helsinki IT department are part of this strategy.

Monitoring Storage Media and Data Deletion Procedures

Procedures for monitoring the health of storage media and for the approved deletion of data and metadata are implied through the use of professional data centres and formal backup processes.

Integrity Checks

FinBIF employs automated backup processes and utilizes technologies that inherently support data integrity, such as RAID. The repository's infrastructure facilitates recovery in the event of data loss, and integrity checks, such as fixity checks, are implied within these processes.

References

- CSC cPouta: https://research.csc.fi/-/cpouta
- CSC Rahti: https://research.csc.fi/-/rahti
- CSC Ida: https://www.fairdata.fi/en/ida/

Technical Infrastructure (R15)
Self-Assessed Compliance Level: Implemented

FinBIF's technical infrastructure is built on a combination of well-supported operating systems, cloud technologies, and open-source software, aligning with the needs of our designated community. We operate within a dual-infrastructure framework provided by the University of Helsinki and CSC – IT Center for Science, ensuring high availability, redundancy, and scalability for managing biodiversity occurrence data.

Core Infrastructure and Software

FinBIF operates on a dual-infrastructure model, utilizing OpenStack-based virtual machines and OpenShift Kubernetes platforms provided by the University of Helsinki IT Department and CSC – IT Center for Science. These infrastructures are designed to offer high availability, redundancy, and scalability, which are critical for managing the extensive data and metadata associated with biodiversity occurrence records. We use community-supported, mostly open-source software for repository management, such as the Elasticsearch search engine for data querying and Vertica RDBMS for data warehousing. Most of our infrastructure is based on open-source software, which, coupled with comprehensive documentation, facilitates community engagement and continuous improvement of the repository's technical capabilities.

FinBIF technical storage infrastructure consist of the following:

- JSON document storage - Data storage for primary occurrence data. Used to store complex records, like occurrence documents and museum specimen records. Based on Oracle RDBMS.
- Triplestore storage - Data storage for primary data. Used to store simpler objects like datasets and media metadata, taxonomy, schema properties and users. Based on Oracle RDBMS.
- Several services like user authentication and data request service that have their own small relational databases.
- Media publication platform A primary data storage. Contains media objects as physical files.
- Media archival service - A backup data storage. Contains large original images and backups of media on the publication platform.
- Data warehouse - A secondary data storage. Vertica RDBMS/big data search engine to which occurrence data from various primary data sources are synchronised for searching.
- Search engine - A secondary data storage. Data from JSON document storage, triplestore storage and various geographical features are loaded to an Elasticsearch search engine for faster searching.

Implementation and its features are widely documented in Schulman et al. 2021.

IT Service Management and Standards Compliance

FinBIF adheres to international and community technical standards, including Darwin Core for biodiversity data and MIDS for specimen description, to ensure interoperability and data quality. The repository's technical infrastructure development is informed by these standards, and compliance is actively monitored through participation in global initiatives such as DiSSCo and Synthesys+. Version control for repository-generated software is managed through publicly accessible code repositories, promoting software integrity and facilitating community contributions.

Availability and Bandwidth

The repository's technical infrastructure is designed to meet the demands of its Designated Community, ensuring sufficient availability and bandwidth. Continuous communication with infrastructure providers and annual service assessments help FinBIF maintain and adjust its infrastructure to meet evolving needs.

Monitoring Technical Change

Regular reviews of infrastructure capabilities, user feedback, and technological advancements guide the repository in adapting its technical infrastructure. This proactive approach ensures that FinBIF remains at the forefront of biodiversity data infrastructure.

References

- Schulman, L., Lahti, K., Piirainen, E. *et al.* The Finnish Biodiversity Information Facility as a best-practice model for biodiversity data infrastructures. *Sci Data* **8**, 137 (2021). https://doi.org/10.1038/s41597-021-00919-6
- FinBIF Technical Architecture: https://wiki.helsinki.fi/xwiki/bin/view/FinBIF/FinBIF/1%20Overview/Architecture/
- DiSSCo: https://www.dissco.eu/
- Synthesys+: https://www.synthesys.info/
- MIDS, Minimum Information about a Digital Specimen: https://www.tdwg.org/community/cd/mids/
- FinBIF repositories:
    - Github: https://github.com/luomus
    - Bitbucket: https://bitbucket.org/luomus/workspace/repositories/

Security (R16)

Self-Assessed Compliance Level: Implemented

FinBIF acknowledges the significance of a robust security system, especially given the sensitivity of biodiversity data and its users. Considering this, a comprehensive analysis of potential threats has been undertaken, followed by risk assessment and consistent security measures.

Levels of Security and Data Classification

At FinBIF, data on biodiversity occurrence and personal information are classified into two categories based on security requirements: public data and restricted-use data. These are stored in two separate databases, each tailored to its security needs. This bifurcation ensures that restricted-use data receives specialized protection while maintaining accessibility for public data.

Processes and Risk Analysis

In 2020, FinBIF services underwent a third-party security assessment by Silverskin Information Security Ltd. The outcome of the assessment was rated "Good". A confidential report of this assessment is available upon request. FinBIF continuously evaluates potential threats from various sources, be it human error, malicious intent, or technical failure. Based on these evaluations, risk levels are ascertained, and necessary countermeasures are put in place. This risk management process is dynamic, adapting to emerging threats and the evolving landscape of biodiversity informatics.

Procedural Safeguards and Collaboration

Recognizing the broader implications of data security, especially with personal information, extensive legal consultations have been undertaken. Collaborative efforts with the Ministry of Environment, Ministry of Agriculture and Forestry, Helsinki University, Finnish Environment Institute (Syke) and Natural Resources Institute Finland (Luke) ensure that our data management procedures align with intellectual property rights, as well as legal and ethical standards, including best practices for publishing biodiversity data while protecting sensitive information.

Human Resources and Training:

FinBIF has a dedicated IT-admin with the ISC2 Systems Security Certified Practitioner certification. This role will lead ongoing security evaluations, risk assessments, training, and countermeasure planning. Through regular workshops and training modules, all personnel are kept updated on best practices, ensuring that security remains a collective responsibility.

Facility and Physical Security

The University of Helsinki and CSC, our hosting infrastructures, prioritize restricted access and uphold stringent security standards for their server rooms. Internal best practices, such as secure storage of work equipment and confidential information, further support our security measures. CSC is certified according to the ISO 27001 standard for information security.

Authentication and Authorization Procedures

To ensure secure and user-friendly access, FinBIF implements a tiered authentication and authorization system, integrating various authentication methods tailored to specific user groups.

References

- Information about security at CSC: http://urn.fi/urn:nbn:fi:lb-201710264