



Guidelines for management of data at the Norwegian Polar Institute

1. Purpose

These guidelines are intended to secure the best possible data for research, monitoring, management and environmental information. An open data policy serves to maximise returns to society on the Institute's investments in data collection, as well as ensuring public accessibility of environmental information, and the credibility, transparency and verifiability of the Institute's research and environmental consulting. These guidelines set forth principles concerning data ownership and copyright, and benchmark good practices for documentation, dissemination and long-term management of the data sets.

This data policy will ensure that the Norwegian Polar Institute fulfils both ICSU World Data Systems criteria for permanent data archives, and the data policy standards that are demanded for full participation in Norwegian and international data collaborations and networks that are important within the Institute's fields.

2. Definitions

"Data" are defined as any set of observations or measurements that are registered in written or digital form, are intended for use in research, environmental monitoring, or mapping, and can be described by metadata. The term can also include scientific samples and specimens.

Reference data, monitoring data, environmental data, and map data are data collected through documented methods for purposes that are wholly or partly related to management, counselling, information, documentation, mapping, or other purposes that fall within the Institute's mission to society.

Research data are data that are collected for a specific scientific project, and are unique to that project. Digital topographic maps are exempt from these guidelines.

3. Ownership and responsibilities

The Norwegian Polar Institute owns all data and scientific material collected by its employees in the course of their duties. This also applies to externally funded projects unless the funding entails specific exemptions with regard to data ownership.

Department Directors are responsible for implementation of these guidelines. Project managers are responsible for ensuring that the data from a project are properly documented and permanently stored.

4. Basic principle: open data

All data owned by the Norwegian Polar Institute shall be open and accessible free of charge for further use, in accordance with government policy for release of public and publicly funded data¹. When data are (re)used

¹ See Norwegian regulations concerning freedom of information, environmental information, geodata, copyright, national implementation of OECD principles, and the Institute's letter of appropriations. The Freedom of Information Act (offentlighetsloven) and the Environmental Information Act (miljøinformasjonsloven) can be expected to heighten the need for transparency concerning monitoring data and other data that form the basis for consultancy and information to the public on the state of the environment.

by external entities, the data source must be cited in accordance with Institute guidelines as published at the time.

Specific datasets can nonetheless be protected from access if this is necessary to maintain personal integrity or if public access to the data could pose a significant risk of harm to vulnerable species or natural resources. The protection should be limited to the sensitive portions of the dataset².

Access to research data can be restricted for a specified and limited period of time to protect the researcher's right to first scientific publication of the dataset.

5. Licenses

To guarantee the rights of both the Institute and the individual who generated the data to be credited when the data are (re)used, and to protect them against liability associated with such use, data from the Norwegian Polar Institute are to be published under open data licenses such as the Norwegian license for public data (NLOD) or compatible international licenses such as Creative Commons attribution (CC BY)³.

6. External collaboration

When the Institute collaborates with outside agencies in projects that require exchange of data, the collaboration should implement the same licenses to ensure access to data and protect intellectual rights during open data exchange. When data are collected by the Institute as part of external programmes with their own data policies⁴, those policies should be followed, but in a manner that does not restrict the Institute's open publication its own data.

7. Publication rights (rights management)

Employees at NPI have a right to be first to publish research data they themselves have generated. To safeguard this right, research data may be kept confidential until the dataset has been published for the first time, albeit only for a specified period of time. The protection period shall run to a specific date, and should not normally extend for more than two years after data collection is completed. Within the contractual conditions and external requirements that apply to the project, the protection period can be prolonged if necessary to safeguard the right of first publication. The Section Leader is responsible for setting the release date and can extend the protection period to up to four years after data collection is completed. A data protection period exceeding four years shall be granted only in exceptional cases and must be approved by the Director of the Department.

When the protection period is over, or if no such period has been defined, data that have been processed, documented, and stored at the Institute are to be openly accessible at data.npolar.no.

8. Data lifecycle management

Reference data, monitoring data, and environmental data are managed by the Institute's Data Centre. Research data are also managed by the Centre once they have been processed, documented, and released for archiving. For new projects that entail collection of data, the project description shall include a plan for long-term management of the data. This plan is to be formulated in consultation with the Institute Data Centre, and should also include a brief description of which datasets will be collected (metadata); plans for data processing, analysis, and quality control; transformation to finished data product; storage site and storage

² For example, geographic coordinates of denning or breeding sites of species that are vulnerable to human activity.

³ CC BY 4.0 NO (Norwegian version) when it becomes available.

⁴ Such as SCAR, CliC, IPY, GBIF, IODE, etc.

format; protocols and services that will be used to make the data accessible; documentation for future reuse; and date for publication.

Data integrity⁵ must be ensured throughout the lifecycle of the dataset, through general data security, version control and, if necessary, access control. The Institute Data Centre is responsible for transferring data to new storage media when required to ensure data integrity or to stay abreast of technological developments. Projects involving data collection are not considered completed until all the data have been archived and documented. This is to be checked by the Director of the Department prior to approval of the final report.

To ensure future verifiability, data that form the basis for scientific publications, reports, and analyses must be unequivocally identifiable. Data should be archived with unique, persistent digital identifiers that can be linked to a version of the dataset as it stood on a given date⁶.

9. Data storage

All data are to be stored in original form, at the Institute, or at remote data centres with which the Institute has signed contracts. The data should be stored according to best practices for long-term data management, in well-documented standard formats, and backups should be stored separately from the original data. Digital data should whenever possible be stored in generic format, independent of system, in standard open-code formats and – if applicable – be readable with open source software.

Where applicable, data should be collected and documented according to standard practice for the data type or discipline in question.

10. Documentation

All datasets must be registered in the Institute's database of metadata. The meta-database will make it possible to retrieve, maintain, and publish datasets. When a dataset is updated or supplemented with new data, the metadata must also be updated. Alternatively, a new and separate edition of the dataset can be published.

All datasets archived at NPI are to have adequate documentation for independent use in the future. Metadata must be searchable at data.npolar.no and must be possible to share with external data directories. Storage information must always be kept up to date in the meta-database.

Future use of the dataset will often require documentation that goes beyond what is recorded in the meta-database. Such information must be compiled in a data report from fieldwork / data collection. The data report must contain all the information required to make use of the data set, including:

- * Field log
- * Documentation of the data collection method, quality assurance and validation
- * Documentation of software that has been used and/or is required
- * Description of instruments, sources of error, measurement resolution, etc.

The data report must be available electronically, and its URI registered in the meta-database.

⁵ This means that the data set is maintained intact, without any form of deterioration or inadvertent or unauthorised change.

⁶ The most likely types of identifiers are UUID and DOI. Both link to a URI that brings up the actual dataset.

11. Accessibility

Metadata and archived datasets are to be made searchable and accessible over the internet through digital services based on standard protocols⁷; they will have unique and persistent identifiers, and be accessible through services for searching, downloading and transformation.

Data dissemination over the internet requires that both metadata and all preconditions for dissemination of the datasets are machine-readable. Metadata, and if possible dataset itself, should therefore contain formatted information on ownership, copyright, data licensing, citation, publication date, version, any restrictions on use and other rights. Metadata are to be associated with the data through unique, persistent digital identifiers.

12. Personal integrity

All data should in principle be free from personal information other than title, name, place of employment, and working address and phone number. If a dataset includes any other personal information, NPI must follow the regulations set forth by the Data Inspectorate (Datatilsynet) and the Data Protection Act (personopplysingsloven).

⁷ HTTP, OPeNDAP, WMS, WFS and others