

Department of Fisheries and Oceans

Policy for Scientific Data

Table of Contents

1. Effective Date	2
2. Application	2
3. Context	2
4. Definitions	2
5. Policy Statement	3
5.1 Objective.....	3
5.2 Expected Results.....	3
6. Policy Requirements	4
6.1 Principles	4
6.2 Data Nodes.....	5
6.3 Data Submission	5
6.4 Access	6
6.5 Inclusion of a Data Management Component in Science Projects.....	7
6.6 National Inventory.....	7
6.7 Acquisition of Data from Third Party Sources.....	7
6.8 Data Submitted under Regulations or Having Legal Aspects	7
6.9 Information Technology Support.....	7
6.10 Access to Information and Privacy Acts	8
6.11 Implementation	9
7. References	9
7.1 Relevant legislation	9
7.2 Related DFO Policies	9
7.3 Related Treasury Board Policies	9
7.4 Other Publications	9
8. Enquiries.....	10
Annex 1: Data Types.....	11

1. Effective Date

This policy effectively replaces the previous Management Policy for Scientific Data on September 1, 2013

2. Application

This policy applies to all scientific Data collected or generated by Fisheries and Oceans Canada.

3. Context

Fisheries and Oceans Canada, DFO, through its own programs and through exchanges with national and international organisations, has acquired over the years and continues to acquire, a large volume of Scientific Data and information. The Scientific Data policy is needed to preserve, enhance and facilitate the use of the valuable and irreplaceable Scientific Data resources of DFO.

This Policy supports the Departmental mission and strategic outcomes:

“Through sound science, forward-looking policy, and operational and service excellence, DFO employees work collaboratively toward the following strategic outcomes:

- *Economically Prosperous Maritime Sectors and Fisheries;*
- *Sustainable Aquatic Ecosystems; and*
- *Safe and Secure Waters.”*

by protecting the factual basis on which decision are made, policies are developed, and on which operations and services rest.

This Policy conforms to the Government of Canada Policy on Information Management, while allowing flexibility to accommodate the many Scientific Data sharing arrangements and obligations DFO has with external agencies in Canada and internationally.

This policy is to be evergreen, and will be linked to a strategic implementation plan. The policy and implementation plan will be revised and adjusted as imposed by resource constraints.

4. Definitions

Within this document the following definitions will prevail:

Data are qualitative or quantitative values that characterize either the environment, organisms living in the environment or human activities relating to environmental resources. Data can either be obtained through direct measurement or be derived from other Data for example numerical modelling output. It is understood that “Data” and “Scientific Data” in the context of this document are the same.

Policy for Scientific Data

Data Management comprises all the disciplines related to the development, execution and supervision of plans, policies, programs and practices that validate, control, protect, deliver and enhance the value of Data. This is a shared responsibility between DFO Science for Data collection, generation, assembly, transformation (content management responsibility) and stewardship, and organizations charged to provide the systems and services that enable management of Data and Metadata such as storage, applications, access and preservation in a commonly agreed organizational framework.

Data Node is taken to mean an entity within DFO comprising an organization and staff dedicated to the functions of Data content management and having access to the necessary infrastructures and services to perform these functions. There may be specialized Data Nodes that address a specific domain.

Managed Data Repository is a system under the care and control of a Data Node for storing and managing Data and Metadata of a given type.

Metadata is contextual information needed to identify, find, cite and properly understand the Data. It includes information on authors, the programs under which the Data was collected, contact for further information, the location where the Data can be found, description of how the Data is organized and formatted, the methods by which the Data was obtained as well as the units, precisions and accuracy of the Data

Science refers to the organization in DFO responsible for conducting scientific activities. It includes the Ecosystems and Oceans Science (EOS) sector, the organizations under the Regional Directors of Science, or any other bodies that hold similar responsibilities as the DFO organization evolves.

5. Policy Statement

5.1 Objective

The objective of this policy is to ensure the safeguard of scientific Data holdings and to maximise the usefulness of Data through interconnectivity, accessibility and adherence to standards.

5.2 Expected Results

The outcomes of this policy are:

- secure Data now and in future,
- interconnected and interoperable Data;
- useful Data discoverable and accessible through standard means; and
- cost effective Data management.

6. Policy Requirements

6.1 Principles

Scientific Data sets are valuable national resources that have been acquired through decades of investment, enabling DFO to maintain world leadership in aquatic sciences. These Data are irreplaceable, and must be protected and managed to ensure long-term availability.

To be useful scientific Data must be accessible, usable and of known quality. Because of the complex and often unique nature of scientific Data as well as the fact that scientist are themselves the largest users of Data, it is essential that Science at DFO be responsible for managing the quality, flow, organization and dissemination of Data to ensure its meaning and usefulness are preserved.

Scientific Data retain their value indefinitely and therefore must be preserved forever. It is paramount that Data be managed and archived so that they remain intelligible in the long term. To ensure proper management and protection, all scientific Data collected by DFO must be submitted to a Managed Data Repository immediately after the measurements have been processed or derived Data created.

For decision making, the most recent Data are typically the most critical. To obtain maximum benefit to DFO and to the user community at large, scientific Data must be made available as soon as possible.

The Government of Canada's Open Data Policy as well as DFOs philosophy promote full and open access to Data whenever possible and subject to particular departmental, national and international obligations regarding sensitivity and confidentiality.

Canada's aquatic environment is influenced by and influences a coupled water and atmospheric system that extends globally and cannot be effectively monitored by one nation alone. Scientific Data regarding the aquatic environment are therefore more useful when pooled and shared with the international community. To obtain access to all of the Data and information that are pertinent to Canadians, Canada must be able to exchange Data with other world Data centres, while keeping in mind departmental, national and international obligations regarding sensitivity and confidentiality.

Derived Data, for example analysis products or numerical model output, are not necessarily subject to all the safeguards justified for observational Data, as they can in principle be reproduced. In certain cases, however, long term preservation may still be required for legal reasons or to document the basis for decisions and policies. The needs for discovery and accessibility of derived Data are nevertheless the same as for observational Data. A strategy must therefore be developed on a case by case basis for derived Data sets, stating the rationale for any departure from the present Policy.

6.2 Data Nodes

All DFO scientific Data must be managed as part of an integrated system accessible through regional and national Data Nodes. Data Nodes will have the responsibility to:

- Respond to internal and external requests, in accordance with 'Access' Section below.
- Maintain inventories and documentation for all Data holdings for which they have designated responsibility, including Metadata and links to where the Data are stored.
- Provide basic Data retrieval, integration and summarization to satisfy common needs of DFO programs and in conformance with governmental and departmental policies.
- Coordinate sharing arrangements with other organizations.
- Perform, in concert with the Data providers, Data quality control, verification and removal of duplicates and other value added processing.
- Ensure through partnership with relevant service organizations that Data and Metadata are protected against loss, remain accessible in the long term and survive in the event of organizational changes, shifts in responsibility, retirements, etc...

6.3 Data Submission

It is the responsibility of Science program managers to ensure that Data with the accompanying Metadata are submitted to the appropriate Data Node in a timely fashion whether collected internally, under contract to or in partnership with other agencies. This is important to ensure that Data are quickly migrated into an environment where the risk of accidental or circumstantial loss are minimized, and where the supporting Metadata is integrated with the Data to ensure that contextual information is preserved for the long-term for appropriate interpretation and use of the Data.

Timely fashion is taken to mean that:

- (a) Data sets will be submitted immediately after the measurements are processed or Data is derived;
- (b) submission will not be delayed while Data analysis, statistical treatment, interpretation and publication occur, and
- (c) submission will include Metadata prepared by the Data collector, in accordance with Treasury Board Secretariat guidelines, to document the methodologies and other details needed so that others may cite the Data and be aware of any potential limitations or features of the Data.

Exceptions to this policy are possible if:

- (a) the responsible manager and the responsible Data Node have agreed that the Data in question are not appropriate for submission, or

- (b) it can be demonstrated that there is a legal imperative that categorically prohibits submission of the excluded Data, or
- (c) an extension or exemption from the policy is sought for other reasons and granted in writing by the Regional Director of Science.

Data submission to the responsible Data Node does not mean that the Data will be openly accessible. Thus concerns about access shall not be a valid reason for not submitting Data. It is the responsibility of the Regional Director of Science to designate specific Data as classified for preventing its open sharing.

It is common for DFO to participate in cooperative science programs with other partners and for these programs to set up their own data storage and management systems. In such cases data should still be submitted to a Data Node in order to ensure it can be inventoried as a DFO data asset. The Data Node will then submit data to the program and if appropriate use the program facility as a Managed Data Repository, thus avoiding duplication. Any such data sharing of data and data management function must be approved in at the Regional Director of Science or Director General levels.

It is understood that increasing use of real time data systems enable submission, validation, quality control and loading into Managed Data Repositories for quasi-immediate access. This is encouraged and nothing in this policy should be interpreted as an impediment to real time data systems.

Data encompassed by this policy include all scientific Data that may be collected, derived or otherwise acquired by DFO. Annex 1, which may be updated from time to time, lists several Data types held at DFO.

6.4 Access

DFO scientific Data are a public resource and subject to full and open access within two years of being acquired or generated. In cases where, in the opinion of the Regional Director of Science, there is a danger of improper or incorrect interpretation of the Data, steps shall be taken to ensure that potential users are fully apprised of this possibility and a contact shall be identified to provide assistance in proper use and interpretation.

Exceptions will be made to this policy in the event that one or more of the conditions below are met:

- DFO investigators have written approval from the Regional Director of Science to delay access to the Data; in such cases, the letter of approval will include the rationale for the delay, and an agreed-upon date for the release of the Data;
- There are third party agreements, privacy concerns, or legal restrictions that prevent open Data access;
- The Data are of commercial benefit to DFO, in which case they will be managed according to Departmental intellectual property management regimes and prevailing policy;
- The Data would be protected under s.18 of the Access to Information Act.

Where there is uncertainty or dispute over whether a Data set meets the last three conditions, legal advice shall be sought and followed.

Third party agreements for the provision or exchange of Data affect Data management in DFO and must therefore be approved by the National Science Directors Committee (NSDC) or equivalent body as determined by the ADM responsible for Science, to ensure consistency with this Policy.

6.5 Inclusion of a Data Management Component in Science Projects

All science project proposals and plans must demonstrate the existence of a comprehensive Data management plan to address the management of scientific Data collected or generated during the life of the project.

6.6 National Inventory

A national inventory of DFO scientific Data holdings will be maintained. Each of the Data Nodes will maintain the inventory for the Managed Data Repositories under their control, in conformance to the Treasury Board's policy suite including standards on Metadata. DFO Science with its IT partners will make Science Data and Metadata accessible and interoperable with other government systems such as the Government of Canada Open Data Portal.

6.7 Acquisition of Data from Third Party Sources

DFO Science acquires relevant scientific Data from other national and international sources where these Data contribute to the goals of DFO. This must be done in an open and transparent manner with DFO's rights and duties clearly agreed upon by all concerned parties. Any such agreement must be approved by NSDC.

6.8 Data Submitted under Regulations or Having Legal Aspects

Scientific Data that have legal aspects constraining their distribution, whether collected by DFO or submitted by third parties will be kept in their original form and appropriately secured. If confidential Data are submitted by third parties, a letter of agreement from the third party will be obtained indicating that the Data are provided on a confidential basis. Such agreements are subject to approval as stated in section 6.4.

6.9 Information Technology Support

Data management is a shared responsibility between Science and IT organisations such as DFO's Information Management & Technology Services (IM&TS) and the government's Shared Services Canada (SSC).

Science is principally concerned with knowledge, in other words data content. In addition to being the collector, generator and principal user of scientific Data, Science takes on the stewardship for Data. Data stewards represent the interest for

data producers and consumers and hold the ultimate responsibility for Data assets, their quality and their use. To fulfill these roles and responsibilities Science relies on technology infrastructure and services.

IT organizations are concerned with the representation of data or knowledge in electronic media. The representation of the data can be number, images, videos, text, etc. IT is the custodians for Data, ensuring that they are organised, preserved, protected and accessible.

The following are some of the specific responsibilities of Science as domain expert:

- Data collection, generation and assembly;
- Data submission;
- Data distribution and sharing in the scientific community;
- Content manipulation including the production of derived data, validation and quality control of content (i.e. insuring that data reliably captures the properties of nature or human activity in nature);
- The specification of IT infrastructure and service needed for Science Data management.

All of the above tasks required domain specific knowledge.

The following are some of the specific responsibilities of IT organizations with respects to scientific Data:

- The provision of storage for Data including related services such as backup and quality controls to ensure content does not become corrupt;
- The organisation of storage for Data, which includes data base structures and data models;
- The accessibility to the stored data, which includes networking and access interfaces;
- The provision of data entry, manipulation and access tools (hardware and software), either through procurement or development;
- The evolution of IT data systems and services with technology and in particular the assurance that data content is preserved and unaltered through migrations.

None of these tasks require domain specific knowledge; however, close collaboration with Science is required to ensure that these systems and services fulfill its needs.

The distinction between content and container as well as the requirement for subject domain knowledge are the principles for separation of duties in this data management partnership.

6.10 Access to Information and Privacy Acts

DFO Science and Oceans sectors will manage their Data in a manner consistent with good Data management practices and with the Access to Information and Privacy Acts. When scientific Data are requested under the Act, ISDM officials or the responsible Regional Director of Science shall provide the Data to the ATIP Secretariat and inform them as to whether the Data can be disclosed or not, with supporting justification.

6.11 Implementation

It will be the responsibility of the Regional Directors of Science and the ADM responsible for Science to implement and ensure adherence to this policy. Inter-regional and inter-sectorial issues and concerns will be addressed by the ADM responsible for Science. A strategy and a plan for implementing the present policy will be developed by NSDMC and coordinated with the science strategic plan. The strategic implementation plan shall be inclusive of DFO science Data partners at all levels and shall strive to include Data management within the life cycle of scientific Data as seamlessly as possible.

Given the potential resource requirements, implementation may be phased over multi-year period. Where new or upgraded infrastructure, databases or applications are required these requirements will be passed on for implementation by relevant IT service organizations according to agreed upon plans.

7. References

7.1 Relevant legislation

- [Access to Information Act](#)
- [Canada Oil and Gas Operations Act](#) (Offshore Physical Environmental Guidelines and Guidelines Respecting Physical Environment Programs during Petroleum Drilling and Production Activities on Frontier Lands)
- [Canada Shipping Act](#)
- [Department of Fisheries and Oceans Act](#)
- [Fisheries Act](#)
- [Oceans Act](#)
- [Species at Risk Act](#)

7.2 Related DFO Policies

- [DFO Intellectual Property Management Policy](#)
- [DFO/CCG Data Policy](#)

7.3 Related Treasury Board Policies

- [Policy on Information Management](#) (and associated directives and standards)
- [Policy on Access to Information](#)

7.4 Other Publications

- [DFO National Science Data Management Strategy](#) (2006)
- [Implementation of the National Data Management Strategy](#) (2009)
- [Intergovernmental Oceanographic Commission of UNESCO's Oceanographic Data Exchange Policy](#)

- [GCPedia DAMA Guide to the Data Management Body of Knowledge](#) (2009)

8. Enquiries

For further information on this policy or on accessing scientific Data please contact:
Director, Integrated Science Data Management Service
W082, 12th Floor, 200 Kent St.
Ottawa, Ontario, Canada, K1A 0E6
(Tel) 613-990-0265
(FAX) 613-993-4658
e-mail: isdm-gdsi@dfo-mpo.gc.ca

Annex 1: Data Types

Some Data Types Covered Under the Management Policy of DFO Scientific Data

- A. Physical oceanographic Data
- B. Hydrological Data (e.g. flow volumes of streams and rivers)
- C. Meteorological Data
- D. Biological oceanographic Data
- E. Marine chemistry Data
- F. Contaminants Data
- G. Fisheries Data
- H. Biological Data (from commercial catch sampling, trawl and acoustic surveys, sentinel fisheries and industry surveys, science logbooks, etc.)
- I. Field and lab Data in support of stock assessments
- J. Fish health Data
- K. Freshwater and marine habitat Data
- L. Freshwater biological Data
- M. Experimental Lakes Area (ELA) Data
- N. Data collected by the Canadian Hydrographic Service, subject to CHS agreements and operational practices.