Generic Place-based IDEA Data Policy

FAIR Island Project

Last edited: 2022-06-28 | V1.0.2

ABOUT THE MAJOR VERSION V1.0.0

V1.0.0 is the first official release of the Generic Place-based Data Policy. That includes significant community feedback on the V0.0.0 (https://doi.org/10.5281/zenodo.5781442) from February 2021 through June 2022.

V1.0.2 includes the correct DOI and citation. It also includes Hypothes.is for additional annotations.

FAIR Island Project

The stewardship of research data, material samples, and software underpins scientific reproducibility, synthesis, and reuse. Embedded in a multidisciplinary international research network, [Place] is the key use case for the **FAIR Island Project** (www.fairisland.org (http://www.fairisland.org)), which aims to demonstrate the benefits of global best-practice drawing on the FAIR data principles[1] and the CARE principles for Indigenous data governance[2], as well as guidance towards FAIR principles for research software[3]. The FAIR Island Project and [Host Organization][Place] Society are developing the [Place] Data trust as a core component of the [Place] IDEA.

[Host Organization] - Organization that controls access to the [Place], described in Preamble. Sponsors organizations that provide in-kind support, through donation, provision of people or a facility or instrumentation necessary for the development of the resource, etc. (DataCite Sponsor definition);

Cite this policy as:

Neil Davies, John Chodacki, Catherine Nancarrow, Maria Praetzellis, & Erin Robinson. (2022). FAIRIslandProject/Generic-Place-based-Data-Policy: V1.0.2 https://doi.org/10.5281/zenodo.6977454 (https://doi.org/10.5281/zenodo.6977454)

Tl;dr

In return for the support and access to [Place] (the physical island and data derived from it) provided by [Host Organization], participants in research projects agree to make their metadata and prepublication data available to each other through a "[Place] Data Trust" that facilitates data re-use according to mutually agreed terms and a code of conduct. The Trust is maintained by the [Host Organization] to support decision-making for wise and equitable stewardship of the [place]. It follows best-practice in research data stewardship. This policy does not impose legally-binding conditions nor replace or impinge on legal obligations and/or institutional policies that might apply.

This is a living document. Ideas or feedback can be contributed through commenting directly using Hypothes.is which will create issues in the Github repo or you can directly create an issue: https://github.com/FAIRIslandProject/Generic-Place-based-Data-Policy/issues (https://github.com/FAIRIslandProject/Generic-Place-based-Data-Policy/issues)

Preamble

The owners of [Place] have entrusted **[Host Organization]** (hereinafter "[Host Organization]")as the [Place]'s's environmental steward with authority to develop and implement the **[Place] Conservation & Sustainable Use Plan** (hereinafter the "CASUP")[4], and to manage access to the [Place] for research, education, or conservation purposes. All projects must be authorized by [Host Organization], which may seek advice from its [governing bodies like Scientific Advisory Board (SAB), Cultural Advisory Board (CAB), and/or CASUP Committee, as appropriate].

The handling of research data and other research outputs derived from [Place], including material samples and intellectual property, is a key consideration for [Host Organization]. The current document sets out its policy in this regard. As the environmental steward of [Place], [Host Organization] has an interest in the stewardship of data and material samples derived from the [Place] under its auspices. In addition to its role as environmental steward of 'physical' [Place], [Host Organization] is ready to act as the steward of the [Place]'s "digital twin" or "avatar", and to that end has established the **[Place] Data trust** in the context of a collaborative program:

[Place] IDEA

[Host Organization] is a collaborating node of the **Island Digital Ecosystem Avatars (IDEA) Consortium** which aims to build use-oriented 4-D simulations of entire social-ecological systems[5]. Island avatars leverage "digital twin" technology, and draw on all the data and knowledge about a place, to simulate future states of the social-ecological system under alternative scenarios of human activity and/or environmental change. **[Place] IDEA** enables potential actions (e.g., conservation management) to be assessed in silico prior to deploying them. It serves as a decision-support tool for the [Host organization] CASUP, facilitating scenario-based planning.
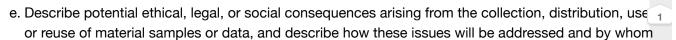
Purpose of this Document

The current document (hereinafter "the Policy") represents guidelines that define the general terms under which [Host Organization] grants Participants in scientific Research Projects (hereinafter "Participants") access to [Place]. The document is intended to be 'socially binding' with respect to the principles and common understanding described herein. The Policy does not preclude or replace any legal obligations or responsibilities that Participants or the institutions with which they are affiliated might incur, for example, under the regulations and laws of French Polynesia.

Article 1 – Convention

1. In return for the permission and support provided by **[Host Organization]** to access [Place] for research purposes, Participants in scientific Research Projects (hereinafter "Participants") agree to contribute to the conservation and sustainable use of the [Place] by participating in the **[Place] IDEA**, contributing to the **[Place] Data trust**, and abiding by the **IDEA Consortium Code of Conduct** (see Exhibit – IDEA Code of Conduct).

2. Participants who lead Projects (hereinafter "Principal Investigators" or "PIs") agree to submit a Project **Data Management Plan** (hereinafter the "DMP"), using the DMP Tool[6] or equivalent, with their initial Project application to [Host Organization]. Modifications of the DMP might be required prior to [Host Organization] granting approval to access the [Place]. [Host Organization] will provide a DMP template through the DMPTool that will:

  a. Provide ORCID IDs[7] for all Participants in the Project

  b. Describe intended outputs[8] (which might include datasets; original software; material samples)

  c. List any intended intellectual property (e.g., patents, copyright, etc.)

  d. List potential commercial uses of intended outputs

  e. Describe potential ethical, legal, or social consequences arising from the collection, distribution, use or reuse of material samples or data, and describe how these issues will be addressed and by whom

  f. Commit to using the Traditional Knowledge/Biocultural Notices[9] where appropriate to facilitate FAIR with CARE future data

3. Participants agree that [Host Organization] may publish a **Marker Paper** as an Open Access document (e.g., in Zenodo[10]) authored by the Principal Investigators and drawing on their Project Application and DMP. [Host Organization] may also make available a **Data Dashboard** as a public portal to the Project's metadata in the [Place] Data trust .

4. Participants agree to make all reasonable efforts to utilize and maintain identifiers and linkages generated during the Project, including the DOIs for Marker Papers and DMPs, and to connect them to future derived outputs, enabling dynamic updates to the [Place] Data trust .

5. If at some point during a Project, circumstances change and Participants can no longer contribute data under the terms described in its DMP and compliant with the terms described in this Policy (e.g., if commercial opportunities arise) the Principal Investigator shall request to make an arrangement with the [Host Organization].

Article 2 – [Place] Data Trust

1. Purpose: The [Place] Data Trust is maintained by [Host Organization] to help ensure that research data, material samples, and software relevant to [Place] are as Findable, Accessible, Interoperable, and Reusable (FAIR) as possible, in accordance with global best practice with respect to ethical, legal, social, and scientific considerations. The [Place] Data trust consists primarily of metadata and transaction records with links to metadata, data, or software that are maintained in suitable third-party repositories.

2. Stewardship: [Host Organization] agrees to provide stewardship of Project metadata and data in the [Place] Data Trust to the extent its resources allow, following the terms outlined in this Policy, and as stated in the Project's DMP or any agreed amendments thereof.

3. Project Participants agree that [Host Organization] may maintain copies of Project metadata, data, and software in the [Place] Data trust[11] that are contributed by the Project on a non-exclusive basis and with "no rights reserved" (under Creative Commons CC0). If justified in the original DMP or its subsequent amendments, more restrictive licenses may be applied (e.g., due to ethical, legal, social, or economic aspects - see Article 4)

4. [Host Organization] is under no legal obligation to maintain Project metadata, data, and software in the [Place] Data trust and Project participants are responsible for and free to archive its metadata, data, or other outputs elsewhere as appropriate.

5. By default, there is no embargo period on the release of metadata or data. If requested and justified in the DMP, however, an embargo period may be granted for specified data (though not for metadata) for up to 1 year after the stated end-date of the Project.

6. [Host Organization] personnel and other IDEA Consortium collaborators who have agreed to abide by the IDEA Code of Conduct are allowed pre-publication access and (re)use of data during any embargo period (see Exhibit 1).

7. Legal Structure: The [Place] Data trust is a voluntary non-legally binding relationship between the Project, its Participants, and [Host Organization] (which may have moral weight though not necessarily legal status).

8. Beneficiaries of the [Place] Data trust include all stakeholders in [Place] including those who provide data, those who are provided with prepublication access to the data (such as the IDEA Consortium), and those who benefit from what is created from the data (most importantly the [Place] of Tetiaroa and its stewardship through the CASUP).

Article 3 – Material Samples

1. Each project's DMP shall include a description of the physical collection and removal (extraction) of material samples from [Place] - whether archeological, geological/chemical, biological, or social (built environment). Projects shall describe in their DMP how they plan to register these sampling events, apply unique identifiers, implement relevant metadata standards, and track derived material (sub)samples, (meta)data, and other derived outputs.

2. [OPTIONAL, if you have specific policies with respect to geo, bio, or anthropological samples, insert here for example: In the case of biological samples, the [Host Organization] has established the **[Place] Biocode & Genomic Observatory**, a scientific research program of the [Institution/s] with the following objectives:]

3. In accordance with the Convention on Biological Diversity (Nagoya Protocol) and associated regulations of the [Country] government, all biosamples collected on [Place] contribute to the option[[Place] Biocode & Genomic Observatory] and shall comply with the [Template ABS Agreement see sample: "Tetiaroa Access & Benefit Sharing (ABS) Agreement (https://drive.google.com/open?id=1W0SK1YihP2eHkt7WhOwJBp5mhy7o1HBx)" .]

Article 4 – Ethical, Legal, and Social Aspects

1. Certain data types or samples require special consideration for ethical, legal, social (including privacy, commercial confidentiality) and scientific reasons, such as, inter alia, data or samples associated with traditional knowledge, archeological artifacts, endangered species, medical data, and human subject research. In some cases, government permits might be required. In other cases, there might be commercial considerations. These aspects should be considered and addressed in Project DMPs.

2. Requests to deviate from the default terms described in this document should be submitted to [Host Organization] with justification in the Project application and DMP, or subsequently if new issues arise. In addition to potential review by [Host Organization] Advisory Boards, other Institutional Review Boards or government agencies might also be involved as appropriate.

3. Projects are encouraged to consider the use of Traditional Knowledge (TK) and Biocultural (BC) Labels, developed to: "support Native, First Nations, Aboriginal, and Indigenous communities in the management of their intellectual property and cultural heritage specifically within the digital environment"[12], and Biocultural (BC) Labels, which extends the TK Label initiative to genetic resources[13].

Article 5 – Commercial Projects

1. Projects with commercial intent require authorization from [Host Organization] (and possibly other parties under separate project-specific, written agreements. Commercial use is defined here as:

   a. Commercial product development; the sale, lease, or license of the Outputs; using the Outputs to produce or manufacture products for general sale, not including publications and/or copyrightable works; transferring tangible rights by sale or license resulting in payment beyond cost; conducting market research; seeking pre-market approval.

Article 6 – Dispute Resolution

1. Disputes concerning the use of data or samples in ways that breach the good-faith engagements made under this data policy are adjudicated by [Host Organization] with input from its Advisory Boards, which will hear from both sides and make recommendations. Sanctions for failing to comply with the IDEA Code of Conduct or other aspects of the [Place] Data Policy may include the rescinding of rights of future access to [Place].

2. This dispute resolution mechanism does not cover breaches of applicable laws or regulations, which shall be dealt with by the competent authorities.

Article 7 – Financial Aspects

1. Nothing in this document commits [Host Organization] or any other parties financially.

Article 8 – Compliance Monitoring

1. The [Host Organization] will identify a delegate to check compliance to the Policy

Exhibit – IDEA Code of Conduct

The following is adapted from the Toronto Statement on Prepublication International Data Sharing[14].

Participants in the IDEA Consortium agree to abide by the following principles:

Respect the Privacy, Ethical, and Legal rights of anyone, including Indigenous peoples[15], who might be adversely affected by the collection, distribution, use or reuse of data, samples or derived information and knowledge. Apply consent, confidentiality, anonymization and other ethical considerations, where appropriate.

Rapid prepublication data release to the IDEA Consortium (e.g., through one of its regional chapters such as the [Place] IDEA), which is a program with the following characteristics:

- Large scale (requiring significant resources over time) and broad utility

- Creating reference data sets

- Associated with community buy-in

Coordinating nodes (e.g., [Host Organization] for the [Place] IDEA) should facilitate the specification of data-release policies and compliance with agreed DMPs by:

- Explicitly informing applicants of data-release requirements, especially mandatory prepublication data release

- Ensuring that evaluation of data release plans is part of the peer-review process

- Proactively establishing analysis plans and timelines for projects releasing data prepublication

- Fostering investigator-initiated prepublication data release

- Helping to develop appropriate consent, security, access and governance mechanisms that protect research Participants while encouraging prepublication data release

- Providing long-term support of databases

Project leaders should state their intentions and enable analyses of their data or material samples by:

- Informing data/sample users (e.g., through a 'marker paper' at initiation of the project) about the data/samples being generated, data standards and quality, planned analyses, timelines, and relevant contact information

- Providing relevant metadata and documentation that will assist other researchers in reproducing and/or independently analyzing the data, while protecting privacy and other ethical, legal, and social aspects concerning individuals and communities

- Ensuring that all Participants in their Project are informed that their data will be shared with other scientists in the IDEA Consortium and the associated terms

- Publishing their initial global analyses in a timely fashion

- Providing a copy of their data, protocols, software and other outputs, in an easily retrievable form and with appropriate metadata and documentation that facilitates usage of both pre-processed and processed data (e.g., following FAIR data principles)

- Ensuring the long-term maintenance of these resources (e.g., through deposition in appropriate repositories and collections).

Data analysts/users should freely analyze released prepublication data or samples and act responsibly in publishing their analyses by:

- Respecting the scientific etiquette that allows data producers to publish the first global analyses of their data set

- Reading the citable document(s) associated with the Project(e.g., 'marker paper')

- Accurately and completely citing the source of prepublication data, including the version of the data set

- Being aware that released prepublication data may be associated with quality issues that will be later rectified by the data producers

- Contacting the data producers to discuss publication plans in the case of overlap between planned analyses

- Ensuring that use of data does not harm Participants and is in conformity with ethical, legal, and social aspects and approvals

Footnotes

1. Making data Findable, Accessible, Interoperable, and Reusable, particularly for machines https://doi.org/10.1038/sdata.2016.18 (https://doi.org/10.1038/sdata.2016.18) ↵

2. Collective benefit, Authority to control, Responsibility, and Ethics - CARE principles for Indigenous data governance. GIDA https://www.Gida-Global.Org/care (https://www.Gida-Global.Org/care)↵

3. Lamprecht, Anna-Lena et al. (2020) 'Towards FAIR Principles for Research Software'. DOI: 10.3233/DS-190026↵

4. Sample CASUP https://www.tetiaroasociety.org/sites/default/files/research-docs/Tetiaroa-CASUP.pdf (https://www.tetiaroasociety.org/sites/default/files/research-docs/Tetiaroa-CASUP.pdf)↵

5. Simulating social-ecological systems: the Island Digital Ecosystem Avatars (IDEA) consortium https://doi.org/10.1186/s13742-016-0118-5 (https://doi.org/10.1186/s13742-016-0118-5)↵

6. https://dmptool.org/ (https://dmptool.org/)↵

7. https://orcid.org/ (https://orcid.org/)↵

8. https://wellcome.ac.uk/funding/managing-CASUP/developing-outputs-management-plan (https://wellcome.ac.uk/funding/managing-CASUP/developing-outputs-management-plan)↵

9. https://www.enrich-hub.org/ (https://www.enrich-hub.org/)↵

10. https://zenodo.org/ (https://zenodo.org/)↵

11. https://theodi.org/article/what-is-a-data-trust/ (https://theodi.org/article/what-is-a-data-trust/)↵

12. https://localcontexts.org/ (https://localcontexts.org/)↵

13. https://www.enrich-hub.org/ (https://www.enrich-hub.org/)↵

14. **Prepublication data sharing** (https://www.nature.com/nature/journal/v461/n7261/full/461168a.html) Toronto International Data Release Workshop Authors Nature **461**, 168-170 (10 September 2009) doi:10.1038/461168a (doi:10.1038/461168a)↵

15. Collective benefit, Authority to control, Responsibility, and Ethics - CARE principles for Indigenous data governance. GIDA https://www.Gida-Global.Org/care (https://www.gida-global.org/care)↵