Tetiaroa Research Data Policy

Tetiaroa Society, French Polynesia

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

Feedback form: https://forms.gle/7aGfkZndD5bmkGrw7

Preamble

The owners of Tetiaroa have entrusted **Tetiaroa Society** (hereinafter "TS") as the private atoll's environmental steward with authority to develop and implement the **Tetiaroa Conservation & Sustainable Use Plan** (hereinafter the "CASUP")[1], and to manage access to the atoll for research, education, or conservation purposes. All projects must be authorized by Tetiaroa Society, which may seek advice from its 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 Tetiaroa, including material samples and intellectual property, is a key consideration for Tetiaroa Society. The current document sets out its policy in this regard. As the environmental steward of Tetiaroa, TS has an interest in the stewardship of data and material samples derived from the atoll under its auspices. In addition to its role as environmental steward of 'physical' Tetiaroa, TS also serves as the steward of the atoll's digital representation ("digital twin" or "digital avatar") and to that end has established the **Tetiaroa Data Trust** in the context of two collaborative programs:

Tetiaroa IDEA

Tetiaroa Society is a collaborating node of the **Island Digital Ecosystem Avatars (IDEA) Consortium** that aims to build use-oriented 4-D simulations of entire social-ecological systems[2]. 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. **Tetiaroa 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 Tetiaroa CASUP, facilitating scenario-based planning.

[1] https://www.tetiaroasociety.org/sites/default/files/research-docs/Tetiaroa-CASUP.pdf

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

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, Tetiaroa is the key use case for the **FAIR Island Project** (www.fairisland.org), which aims to demonstrate the benefits of global best-practice drawing on the FAIR data principles[3] and the CARE principles for Indigenous data governance[4], as well as guidance towards FAIR principles for research software[5]. The FAIR Island Project and Tetiaroa Society are developing the Tetiaroa Data Trust as a core component of the Tetiaroa IDEA.

Purpose of this Document

The current document (hereinafter "the Policy") represents guidelines that define the general terms under which Tetiaroa Society grants participants in scientific research projects (hereinafter "Participants") access to Tetiaroa. 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 **Tetiaroa Society** to access Tetiaroa for research purposes, Participants in scientific Projects (hereinafter "Projects") agree to contribute to the conservation and sustainable use of the atoll by participating in the **Tetiaroa IDEA** (see Preamble), contributing to the **Tetiaroa Data Trust** (see Article 2), and abiding by the **IDEA Consortium Code of Conduct** (see Exhibit 1).

2. Participants who lead Projects (hereinafter "Principal Investigators" or "PIs") related to Tetiaroa agree to submit a Project application to TS. A key component of this Application is a **Data Management Plan** (hereinafter the "DMP"), that may be developed. Modifications of the DMP might be required prior to TS granting approval to access the atoll. TS will provide a DMP template, using the DMP Tool[6] or equivalent, that will require:

    a. ORCID IDs[7] for all Participants in the Project
    b. Description of intended outputs[8] (which might include datasets; original software; material samples)
    c. List of any intended intellectual property (e.g., patents, copyright, design rights and confidential know-how)

---

[3] Making data Findable, Accessible, Interoperable, and Reusable, particularly for machines https://doi.org/10.1038/sdata.2016.18

[4] Collective benefit, Authority to control, Responsibility, and Ethics - CARE principles for Indigenous data governance. *GIDA* https://www.Gida-Global.Org/care

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

[6] https://dmptool.org/

[7] https://orcid.org/

[8] https://wellcome.ac.uk/funding/managing-CASUP/developing-outputs-management-plan

   d. Description of potential ethical, legal, or social consequences (including potential commercial uses) arising from the collection, distribution, use or reuse of material samples or data; how these issues will be addressed and by whom.
   e. Commitment to using the Traditional Knowledge/Biocultural Notices[9] where appropriate.

3. Participants agree that TS 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 including its DMP. TS may also make available a **Data Dashboard** as a public portal to the Project's metadata.

4. Participants agree to make all reasonable efforts to utilize and maintain persistent identifiers and linkages generated during the Project and to connect them to future derived outputs, enabling dynamic updates to the Tetiaroa 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, the Principal Investigator shall request to make an agreement with TS, and any other relevant parties, to cover the Project going forward from that point.

Article 2 – Tetiaroa Data Trust

The Tetiaroa Data Trust consists of transaction records and other metadata, data, and other research products from Projects carried out on Tetiaroa under the terms of this Policy. These digital objects ("Project Data") may be stored in multiple repositories, which should be described in the Project's Data Management Plan.

1. Definition: "A Data Trust can be defined as: "a legal structure that provides independent stewardship of data." [11] While not necessarily a legal structure, the Tetiaroa Data Trust is a set of mutually agreed terms that should be considered socially-binding, and may in some cases be legally-binding. In the Tetiaroa Data Trust, Principal Investigators (Trustors) assign a stewardship role to Tetiaroa Society (Trustee) under the terms described in their Project's Data Management Plan (DMP) to share with Tetiaroa and its stakeholders (Beneficiaries) fair and equitable benefits arising from the Project's data.

2. Purpose: The Tetiaroa Data Trust is maintained by TS to help ensure that research data, material samples, and software relevant to Tetiaroa 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.

3. Stewardship: TS agrees to provide stewardship of Project metadata and data in the Tetiaroa 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.

[9] https://www.enrich-hub.org/
[10] https://zenodo.org/
[11] https://theodi.org/article/defining-a-data-trust/

4. The Project agrees that Tetiaroa Society may maintain copies of Project metadata, data, and software in the Tetiaroa Data Trust 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)

5. TS is under no legal obligation to maintain Project metadata, data, and software in the Tetiaroa Data Trust and the Project is responsible for and free to archive its metadata, data, or other outputs elsewhere as appropriate.

6. Prepublication data sharing: Metadata and copies of data should be submitted to the Tetiaroa Data Trust as soon as possible, where they are available to TS personnel and other IDEA Consortium collaborators for reuse subject to mutually agreed terms and the IDEA Code of Conduct (see Exhibit 1).

7. Public release of data: TS may release metadata to the public immediately. By default, there is no embargo period on the public release of data, however, if requested and justified in the DMP, 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.

8. Legal Structure: The Tetiaroa Data Trust is a voluntary non-legally binding relationship between the Project, its Participants, and TS (which may have moral weight though not necessarily legal status).

9. Beneficiaries of the Tetiaroa Data Trust include all stakeholders in Tetiaroa including those who provide data, and those who benefit from what is created from the data - most importantly the atoll of Tetiaroa itself (and its stewardship through the CASUP).

Article 3 – 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 TS with justification in the Project application and DMP, or subsequently if new issues arise. In addition to potential review by TS 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].

4. Tetiaroa Society has developed approaches for certain data types and products and these will be amended to this Policy as Exhibits. For example biological samples subject to Nagoya Protocol), (see Exhibit 2).

Article 4 – Commercial Projects

1. Projects with commercial intent require authorization from Tetiaroa Society (and possibly other parties) under separate project-specific 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 5 – 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 Tetiaroa Society 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 Tetiaroa Data Policy may include the rescinding of rights of future access to Tetiaroa.

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

Article 6 – Financial Aspects

1. Nothing in this document commits TS or any other parties financially.

Article 7 – Compliance Monitoring

1. The Tetiaroa Society will identify a delegate to check compliance to the Policy

[12] https://localcontexts.org/
[13] https://www.enrich-hub.org/

Exhibit 1 – 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 release of prepublication data to the IDEA Consortium (e.g., through one of its participating place-based programs, such as the Tetiaroa IDEA). IDEA programs have 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., Tetiaroa Society for the Tetiaroa IDEA) should facilitate the specification of data-release policies and compliance with agreed Data Management Plans 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)

---

[14] **Prepublication data sharing** Toronto International Data Release Workshop Authors *Nature* **461**, 168-170 (10 September 2009) 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

- 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 and metadata, 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
- *In case of major data project using substantial, significant or otherwise potentially sensitive prepublication data, users are encouraged to submit a Project application to TS for access to the data (digital Tetiaroa) to facilitate review and collaboration with relevant stakeholders*

Exhibit 2 – Material Samples

1. Each project's DMP shall include a description of the physical collection and removal (extraction) of material samples from Tetiaroa - 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. In the case of biological samples, the Tetiaroa Society has established the **Tetiaroa Biocode & Genomic Observatory**, a scientific research program with the following objectives:

    a. to comprehensively inventory the biodiversity of Tetiaroa, including all species of fauna and flora – plants, animals, algae, fungi, and microbial groups;
    b. to generate time-series biodiversity observations at the bio-molecular level, including nucleotide sequencing and other "multi-omic" approaches[16].
    c. to test new technological and scientific approaches for the analysis of biodiversity patterns and ecological processes in general; and to make material samples and information available to the research community.

[16] Davies et al. Founding Charter of the Genomic Observatories Network. GigaScience 2014, http://www.gigasciencejournal.com/content/3/1/2

d. In accordance with the Convention on Biological Diversity (Nagoya Protocol) and associated regulations of the French Polynesian government, all biosamples collected on Tetiaroa contribute to the Tetiaroa Biocode & Genomic Observatory and shall comply with the "Tetiaroa Access & Benefit Sharing (ABS) Agreement"