
DMP du projet "Ecology of the circulation of infectious agents in colonial vertebrate populations: surveillance, understanding and implications for biodiversity conservation in sub-Antarctic islands"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - DMP template (english)" fourni par Agence nationale de la recherche (ANR).

Plan Details

Plan title	DMP du projet "Ecology of the circulation of infectious agents in colonial vertebrate populations: surveillance, understanding and implications for biodiversity conservation in sub-Antarctic islands"
Deliverable	D1-DMP_v1
Version	First version
Plan purpose/scope	Data Management Plan of ANR project ECOPATHS, which will be up dated as necessary.
Fields of science and technology (from OECD classification)	Earth and related environmental sciences
Language	eng
Creation date	2022-03-29
Last modification date	2022-10-14
Identifier type	url
License	Creative Commons Attribution 4.0 International

Project Details

Project title	Ecology of the circulation of infectious agents in colonial vertebrate populations: surveillance, understanding and implications for biodiversity conservation in sub-Antarctic islands
Acronym	ECOPATHS

Abstract

Describing and understanding factors affecting the distribution and transmission of pathogenic agents in wild animal populations is important for basic and applied reasons, notably in the current context of global change and recurrent emergences of infectious diseases, which can be related to human activities. Various tools can be used to track, understand and respond to the emergence of infectious agents. Here, we propose to integrate biomedical and molecular epidemiology approaches with population ecology and anthropological approaches to address those issues. Building on our experience in these fields and on the unique situation of sub-Antarctic islands (Crozet, Kerguelen and Amsterdam), where dense but threatened populations of colonial marine vertebrates (albatrosses, penguins, seals) breed in highly structured but relatively simple communities that can be affected by infectious diseases, we propose to address critical hypotheses on the processes underlying eco-epidemiological dynamics and their impacts. A first axis of the work will focus on documenting how infectious agents are shared at a hierarchy of spatial scales among vertebrates host populations in the native vertebrate and introduced species communities of southern ocean islands. We will notably explore the potential usefulness of scavenger and predatory species, such as skuas and giant petrels, as epidemiological sentinels, and their potential roles as spreaders of pathogenic agents. A second axis will focus on the particular situation of Amsterdam Island, where recurrent epizooties of avian cholera are responsible for massive die-offs of nestlings of albatrosses and penguins, and where the test of the use of a vaccine approach is possibly hampered by the negative effects of introduced species like rats. That part of the project will uniquely benefit from the independently funded plan of eradication of introduced species from that island, to be implemented in 2022, and a collaboration with the stakeholder in charge of that eradication plan, the National Nature Reserve of the Terres Australes. A third axis will explore how modelling and social science approaches could be used to better apprehend the risks of emergence of infectious disease in such systems and ways to manage them. The project benefits from involving a consortium of partners and collaborators with a broad range of complementary skills. It will also draw upon the use of samples and access to field set ups provided by a multi-year field research program supported by the French Polar Institute. The results are expected to have broad basic and applied implications.

Funding

- French National Research Agency : ANR-21-CE35-0016

Start date

2022-01-01

End date

2024-12-31

Partners

- ANSES - Laboratoire de santé animale, sites de Maisons-Alfort et de Normandie (190123287X)
- Laboratoire d'anthropologie sociale (200112516E)

Produits de recherche :

1. Data related to all samples (Dataset)
2. Data related to the survival monitoring of the different species (Dataset)
3. Data related to all tracking devices (Dataset)
4. Samples collected on the fields (Collection)

5. Results from the analysis of the different samples (Dataset)
6. Collection of pictures and videos from camera trap on the field (Image)
7. Data resulting from social surveys (Text)

Contributeurs

Nom	Affiliation	Rôles
BOULINIER Thierry	Centre d'Ecologie Fonctionnelle et Evolutive	<ul style="list-style-type: none"> Project coordinator
Tornos Jérémy - https://orcid.org/ 0000-0002-8793-2518	Centre d'Ecologie Fonctionnelle et Evolutive	<ul style="list-style-type: none"> DMP manager Personne contact pour les données (Survival, Samples data, Camera trap, Tracking, Samples type, Samples analysis, Questionnaires)

DMP du projet "Ecology of the circulation of infectious agents in colonial vertebrate populations: surveillance, understanding and implications for biodiversity conservation in sub-Antarctic islands"

1. Data description and collection or re-use of existing data

1a. How will new data be collected or produced and/or how will existing data be re-used?

The data collected in this project will come from 4 annual sampling campaigns in the French Southern Lands (TAAF) between January 2022 and December 2025.

Samples and data will be collected from different sites, including sites on Crozet and Kerguelen archipelagos and Saint Paul and Amsterdam Islands.

Data will be used for further statistical analysis (R software), until the end of the ANR contract

Samples collected on the field will be brought to the laboratory for further lab analysis (PCR, serology, bacteriology).

Data and samples will be documented with all the details of the sampling campaign, in particular the names of the sites (geolocated), the animals ID, the protocols used, the measurements taken, the samples collected and the photos taken if necessary.

Data will be recorded on notebooks or electronic field notebooks.

Others data will be downloaded from devices (tracking data) and send to the lab via email or internet link.

All the laboratory experiments (protocols, follow-up of the experiment and results) and the stages of bioinformatics and statistical analyses will be recorded in a paper laboratory notebook.

1b. What data (for example the kind, formats, and volumes), will be collected or produced?

The data will be stored in different formats after collection:

- spreadsheet (Microsoft Office -> *.xlsx)
- database (PostgreSQL)

Data will then be formatted in standardized format for longterm storage and sharing :

- photos will be available in *.jpeg format
- genetic sequences will be available in *.txt format
- tracking (geolocalisation) data will be available in *.txt and .csv
- all other sampling data will be available in *.txt and .csv

Associated documents will be available -> *.txt or .pdf format

2. Documentation and data quality

2a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?

Metadata

- The database will be described using the EML (Ecological Metadata Language) standard.
- the EXIF (Exchangeable image file format) information of the photos will be saved

Organization and coding of directories

A specific directory named "Field procedure" will contain all the data collection protocols, the instructions for use and calibration of the measuring devices and the and calibration of the measuring devices and the images. It will also contain the organization and coding of directories and files. A specific directory named "Laboratory" will contain all the protocols, the notes of the apparatuses, the experimental results experimental results and bioinformatics and statistical analyses.

A "Field Data" directory will be created for each campaign. It will be organized according to the elements put in the document of organization and coding of directories and files.

All dataset will be organized according to the different sites and campaigns.

2b. What data quality control measures will be used?

During field campaigns, the quality and conformity of the data collection will be controlled by the person in charge of the campaign, assisted by the technician in charge of the measurements and the the data recording.

This control will mainly concern:

- correct measurement according to protocols
- correct sampling according to protocols, animal tissues, organs and other biological samples
- correct coding of the samples
- validation of the data collected in the field

All field data will be recorded on paper notebook or electronic notebook

During the lab experiments, the quality and the conformity of the data will be controlled by the field manager.

This control will concern:

- calibration and the good use of the laboratory equipment
- respect of the protocols
- traceability of samples

All lab data will be recorded in files according to the organizing document (1a and 1b)

3. Storage and backup during the research process

3a. How will data and metadata be stored and backed up during the research?

The data will be saved in differents media:

- Paper field notebook: saved in fireproof boxes at CEFE
- Electronic notebook : backup every evening on hard drive disk
- Spreadsheets: backup on the field laptop then transmission by e-mail of a copy to the CEFE team, then on storage server (automatic daily backups and control of the backups by the CEFE IT department)
- Data on server: backup according to the 3-2-1 principle (3 backups on 2 different media (SQL/binary) with 1 remote location (on the campus of the University of Montpellier))

The same attention will be given to the context data (sites, sampling method, campaign, persons,...)

The metadata (in EML format) will be saved on the servers of the CEFE EIS platform

Once data are safe and cleaned, they will be stored and shared on different hubs:

- OSU OREME platforms of data sharing for sampling data
 - Seabird Tracking database (Birdlife) for tracking data - <http://www.seabirdtracking.org/>
-

3b. How will data security and protection of sensitive data be taken care during the research

With the backup mechanisms in place, data security is assured.

The data will be accessible to campaign managers and their collaborators during the research process (shared on a secure CORE).

They will be modifiable only by the person in charge of the data (field manager, data manager)

The data will be available in open access only after quality control. An embargo period might be possible also for analysis and publication.

Data issued from questionnaires, including personal informations, will be protected and not accessible. A partial and degraded dataset might be shared online but with a full respect of RGPD (no personal and sensitive data accessible).

4. Legal and ethical requirements, code of conduct

4a. If personal data are processed, how will compliance with legislation on personal data and on security be ensured?

The personal data of those involved in the field campaigns (field manager, scientists...) will be subject to informed consent for their collection during the project. These data will never be disseminated.

This information, which is useful for the organization of the campaigns, will only be accessible to the campaign managers and organisations (IPEV, TAAF and others) in charge of the campaigns.

Medical examinations (if required for field campaigns) will be kept confidential and personal except for medical staff on sites.

Data issued from questionnaires, including personal informations of people not involved in the research program (other civilian interviewed), will be protected and not accessible. A partial and degraded dataset might be shared online but with a full respect of RGPD (no personal and sensitive data accessible). Data will be stored with full respect of sex, gender, ethnicity.

4b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

The holder of the intellectual property of the data is the CEFE, CNRS to which the project leader belongs.

The observation data, processed and completed, will be freely accessible (CC-BY-4.0 license free to use on the condition that it is attributed to the author by citing his name).

4c. What ethical issues and codes of conduct are there, and how will they be taken into account?

The project leader (CEFE) will apply the institutional code of ethics of the institution.

The study is in conformity with the prefectural approval of the establishment as far as the animal species collected are concerned.

Ethics committees (CNPN, CEP,...) have validated the scientific necessity of using the species studied. All living individual manipulated will be released.

Samples and data are and will be acquired under permits from the Ministry of Research regarding animal experimentation. Import permits of samples from animal are also obtained as requires.

The committee found that the principles of humane treatment are respected and that the conditions of use of the animals are optimized (3R principle "Replace, Reduce and Refine") taking into account the experimental needs. Campaign managers and responsible for the campaigns and technicians in charge of handling animals and taking the tissue samples are qualified in animal experimentation.

5. Data sharing and long-term preservation

5a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

The data collected in the framework of this ANR will be available in open access (CC-BY-4.0 license) after the final publication embargo of the processed data.

- The data will be made available through the OREME data sharing platform (<https://data.oreme.org/>)
- Genetic data will be shared on Genbank (<https://www.ncbi.nlm.nih.gov/genbank/>)
- Tracking data on Seabird Tracking database (Birdlife) (<http://www.seabirdtracking.org/>)

5b. How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?

All data will be open except for personal data which will be destroyed.

The data of planning and organization of the campaigns, observations and measurements and laboratory analyses will be saved on the CEFE servers for the duration of the project and for 3 additional years.

At the end of this period, only the paper field notebooks, the paper laboratory notebooks and the database will be kept.

The database will be maintained as part of the CEFE's EIS platform.

The data deposited in Genbank, OREME and Birdlife will also be preserved as long as possible as a long term storage.

5c. What methods or software tools are needed to access and use data?

The data will be made available through the different platforms (Genbank, OREME and Birdlife).

Data will be stored in open access format and available and readable by open access software. No particular tool (needing license) is necessary, data will be interoperable with open access tools.

The formats used will be in .txt, .csv, .jpeg, .pdf according to the type of data.

5d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

The deposit in the OREME hub will automatically generate a DOI for the deposited datasets and will be accessible upon request.

Genetic data deposited in Genbank will be identifiable by a unique international accession number.

6. Data management responsibilities and resources

6a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

The management of the data will be ensured by the project leader, namely Thierry Boulinier, with the help of CEFE data management specialists: Jeremy Tornos and Marie-Claude Quidoz

6b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

All the steps taken throughout this project ensure that the data will be FAIR. The writing of a 'datapaper' will contribute to the FAIRisation process of the data resulting from this project.

In view of the few material resources required, all equipment expenses are covered by the EIS platform platform of the CEFE and the OSU-OREME.