



VETBIONET

Veterinary Biocontained facility Network for excellence in animal infectiology research and experimentation

Terms from “VetBioNet Data Management Plan” to be respected by applicants to the “TNA Call for Proposals to Advance Research on Epizootic and Zoonotic Diseases”

Table of contents

1. Introduction to Open Data.....	2
2. How to share FAIR data.....	3
3. Estimation of resources needed for ensuring open access to research data.....	6
4. Data security.....	7
5. Ethical aspects.....	7



1. Introduction to Open Data

1.1 Purpose of data collection/generation and public sharing

The H2020 topic under which VetBioNet has been funded (INFRAIA-01-2016-2017) participates per default in the open access to research data pilot which aims to improve and maximize access to and re-use of research data generated by H2020 projects for the benefit of the research community.

VetBioNet will collect and generate data to advance research on epizootic and zoonotic diseases, with the final objective of strengthening the present European capacity and competence to meet the challenges of (re)emerging animal infectious diseases.

Particularly, VetBioNet will collect data generated from its Joint Research Activities (JRA), which are designed to improve the scientific and technological standards of the integrated services provided by the network infrastructures. For this, the [FAIR principles](#) (Findability, Accessibility, Interoperability, and Reusability) and [ARRIVE guidelines](#) for scientific reporting will be adopted.

Moreover, VetBioNet will also collect data from Transnational Access (TNA) research projects on a voluntary basis, namely when the owners of the data at stake agree to make them available and with a possible delay of 2 years after the end of their TNA projects. **The VetBioNet consortium encourages you, as possible beneficiary of a TNA funding - which in turn is provided by the European Commission - to grant public access (i.e.: open access) to the final research data that could arise from your project.**

1.2 Data generated/collected through TNA projects: types and formats

VetBioNet TNA projects will probably generate data of several types, including phenotypic, genotypic and sequencing data from pathogens and hosts. The final research data produced by TNA projects will be curated in a harmonized way and must be stored in publicly available standard data repositories (e.g. Gene Expression Omnibus, Sequence Read Archive, etc.) or in a repository chosen by the TNA project owner, facilitating sharing of information with the wider community of scientists and end-users.

Research data (phenotypic, genotypic, sequencing and other kinds of data) **generated by researchers outside the VetBioNet consortium through TNA Activities will be covered by data management on a case-by-case basis, as these data may be required to be protected by confidentiality terms specific to each researcher selected for conducting**

a TNA project within VetBioNet. Specific agreements will be sought with researchers conducting TNA projects - allowing them to openly share only a part of their research data and with a possible delay of 2 years after the end of their TNA projects.

1.3 Who will find data useful

The data that TNA projects will make openly available will be useful for the veterinary field, particularly for the scientific community interested in animal infectious diseases, characterizing models and livestock production. The data will also be useful to policy makers, funders and industry who have an interest in biosecurity for animal diseases and zoonoses through sustainable and safe production of livestock species for the public good.

2. How to share FAIR data

2.1. Making foreground data findable, including provisions for metadata

It is essential that final research data arising from TNA projects are searchable with metadata. A standard identification mechanism will be used to describe the data.

Particularly, the naming convention followed will be:

[work package].[task].[(TNA id).[text description].[version].[format]

Search keywords must be provided to optimize possibilities for re-use of data and raw data will be curated accordingly. The keywords to be chosen shall relate to animal species and pathogens analyzed by the specific TNA project and followed methodologies.

Clear version numbers must be provided.

Metadata must be produced to describe how the data were analyzed and summarize important portions of data.

Researchers have to use the taxonomic international accepted names for pathogens, the common pathogen titration methodologies expressing doses such as TCID₅₀, pfu, etc.

The VetBioNet consortium will provide a simple form to all TNA project owners who will develop experimental models for animal infectious diseases. The form will include the following fields: Animal species, Pathogen, Doses, Route of inoculation, Biosecurity

requirements (following OIE classification), Mortality percentage, Clinical signs, Pathogenesis, Immunity and Reference to published data. The form could show in summary the main project achievements, and could facilitate the search for information about animal models related to a particular pathogen, animal species, etc.

2.2. Making foreground data openly accessible

Data sharing with third parties will be subject to a data-sharing agreement established by the VetBioNet IPUDC (Intellectual Property Use and Dissemination Committee). The agreement will indicate the conditions of use, criteria for access, and acknowledgements. TNA project participants who wish to withhold patentable or proprietary data can do so, and advice on this point will be given by the IPUDC.

Final research data can be made openly available by TNA project owners after one of the 3 following criteria are met:

- relevant scientific publications based on the data at stake have been accepted;
- a patent application has been published;
- 2 years after the project end.

Biosecurity reasons will be considered before making data openly available, as well as possible IP protection measures that will allow further exploitation of produced data.

Finally, TNA project participants are encouraged to give public access to the raw data that will not be subject to a patent application, at the latest 2 years after the end of their projects. TNA project participants will autonomously store the raw data used to generate scientific papers in a repository of their choice.

Documentation about the software needed to access the raw and final research data has to be provided by the TNA project owner upon request.

All freeware programs needed to access the data must be provided in the data repositories used by the TNA project owner.

Where possible, open data and associated metadata, documentation and codes should be deposited in certified repositories which support open access.

The TNA project owner can choose to provide access to open data stored in institutional repositories only to registered users (name, affiliation and email can be requested) who have accepted the terms of use.

Data that could raise societal concerns (such as biosecurity risks) must not be shared in open access mode.

2.3. Making data interoperable

To ensure interoperability of data across different projects and where applicable (e.g., data on gene sequences), TNA project participants should upload basic datasets in standardized forms in a primary database as required by the journals in which they publish their results. Where adequate, data generated in TNA projects can be defined according to the Animal Trait Ontology for Livestock (ATOL: <http://www.atol-ontology.com/index.php/en/>).

The VetBioNet consortium will provide common terminology, key words, units and formats that the TNA project owners must apply to the data and other documents to be publicly shared. Standard vocabularies for all data types present in data sets of TNA projects will be employed, to allow inter-disciplinary interoperability.

2.4. Increase data re-use

FAIR principles must be applied also for optimal formatting of data.

Final research data shall be made openly available through publicly available standard data repositories or on an open database chosen by the TNA project owner after one of the 3 following criteria are met:

- relevant scientific publications based on the data at stake have been accepted;
- a patent application has been published;
- 2 years after the project end.

Possible restrictions could be put in place due to the need to finalize patenting processes or publish scientific works based on project data.

Both raw and final research data stored in open repositories shall remain re-usable for a minimum of 10 years.

Each animal experiment conducted by TNA project participants will require approval from an ethical committee. In order to ensure data quality, guidelines will be provided by VetBioNet

coordinator and shared among all TNA project participants to establish common procedures for acquiring, storing and amending data. Finally, [ARRIVE guidelines](#) must be adopted for ensuring metadata quality.

3. Estimation of resources needed for ensuring open access to research data

All data produced during the TNA project can be initially recorded in numbered workbooks, signed and dated and retained at each respective hosting institution, or stored on network attached file systems that will be regularly archived and automatically backed up. Data will be checked for quality and accuracy and all protocols recorded and adapted to clear guidelines. Where necessary, data will be archived in each institution's archive system. Data will be made fully available to the public in a timely manner after passing each hosting institution's quality control criteria and after Intellectual Property considerations have been taken into account. To enable long-term accessibility and validation, data will be stored in formats that are open, non-proprietary, and in common use by the research community.

Each TNA project beneficiary will be responsible for uploading its own final research data on publicly available standard data repositories or on an open database of his/her choice. Management of final research data publicly shared is the responsibility of the TNA project beneficiary, who has to take into account the data preservation costs while preparing the budget of his/her project proposal. A 10-year term would be an adequate preservation period for the data publicly shared by the TNA project owner.

Raw data can also be stored and preserved by the TNA project beneficiaries who generated/collected them. Raw data shall be retained and remain accessible for at least 10 years after completion of the project. Data storage facilities shall be maintained in accordance with the manufacturer's warranty and guidelines and data backed up at regular intervals and stored safely and securely.

Moreover, TNA project participants are encouraged to publish research data as supporting materials together with their project publications, to facilitate preservation of data for future re-use by other projects or research initiatives.

4. Data security

Final research data arising from TNA projects shall preferably be stored in publicly available standard data repositories (e.g. Gene Expression Omnibus, Sequence Read Archive, etc.). Alternatively, data can be stored in a repository chosen by the TNA project owner, if data backups and data security are adequately ensured.

5. Ethical aspects

Ethical issues that can have an impact on data sharing will be analyzed on a case-by-case basis by the TNA project beneficiary and the hosting research institution.

Informed consent for data sharing and long-term preservation must be guaranteed if the TNA project needs to launch questionnaires dealing with personal data.