BCOMING

D1.4 Data Management Plan

Project acronym: BCOMING

Project title: Biodiversity Conservation to Mitigate the risks of emerging infectious diseases

Call: HORIZON-CL6-2021-BIODIV-01

This project has received funding under the European Union's Horizon Europe research and innovation programme under grant agreement No 101059483.

Project no.	101059483
Project acronym:	BCOMING
Project title:	Biodiversity Conservation to Mitigate the risks of emerging infectious diseases
Call:	HORIZON-CL6-2021-BIODIV-01
Start date of project:	01.08.2022
Duration:	48 months
Deliverable title:	D1.4 Data Management Plan
Due date of deliverable:	31.01. 2023
Actual date of submission:	31.01. 2023
Deliverable Lead Partner:	**CIRAD**
Dissemination level:	Public

Author list

Name	Organization
Laurence Dedieu	CIRAD
Julien Cappelle	CIRAD
Bruno Rosset	CIRAD

Document History			
Version	**Date**	**Note**	**Revised by**
01	17.10.2022	Initial version by Laurence Dedieu	Julien Cappelle
02	17.01.2023	Draft report	WP Leaders
03	31.01.2023	Final report	Approved by GA

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While BCOMING is funded by the European Union, views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the European Research Executive Agency (REA) can be held responsible for them.

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the BCOMING consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the BCOMING Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the BCOMING Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Executive Summary

BCOMING data will follow the “FAIR” principles, meaning “Findable, Accessible, Interoperable and Re-usable”. The data will be made findable and accessible within the Consortium, and to the broader research community, stakeholders and policy makers. Also, data will be compliant with national and European ethic-legal frameworks, such as the General Data Protection Regulation (GDPR, Regulation (EU) 2016/679). The present data management plan (DMP) describes the data management life cycle for all data to be collected, processed and/or generated by the project. It includes information on the handling of research data both during and after the end of the project; the nature of the data, the methodology and standards applied, whether data will be shared or provided in open access, and how the data will be curated and preserved.

This initial version of the DMP will be updated during the project with the measures taken by the partners to ensure FAIR data. A final version of the DMP will be submitted as deliverable D1.5 at the end of the project (M48, July 2026).

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Table of Contents

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

List of tables

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Data Summary

Types, formats and size of data generated or re-used by the project

BCOMING will develop standardized data collection and analysis framework to support the co-construction of innovative biodiversity conservation strategies and zoonotic disease surveillance systems to reduce the risk of infectious disease emergence in three tropical biodiversity hotspots with different environmental and socio-cultural settings in Southeast Asia (Cambodia), West Africa (Guinea and Ivory Coast) and the Caribbean (Guadeloupe).

Additionally, previously collected datasets will be reused to conduct some activities of BCOMING. In that case, the partners will make sure that said data was initially collected in a lawful, fair, and transparent manner, and that these data are not covered by specific rights or licences and could thus be reused freely for research purpose. Appropriate procedures are described in Deliverable D9.3. POPD.

In particular four previous projects conducted by the consortium partners will provide datasets for BCOMING analyses:

- ZooCoV in Cambodia: ZooCoV is funded by ANR, coordinated by CIRAD and aims to prevent coronavirus transmission from wildlife to humans in Cambodia.

https://umr-astre.cirad.fr/en/research/projects/zoocov

- EBO-Sursy in Guinea: EBO-Sursy is funded by EU-DEVCO, coordinated by OIE, implemented notably by CIRAD and IRD and aims to strengthen early detection systems for wildlife in Africa (including Guinea) to prevent outbreaks of viral haemorrhagic fevers.

https://rr-africa.oie.int/en/projects/ebo-sursy-en

- BIODIV-AFREID in Ivory Coast: BIODIV-AFREID is funded by ERA-NET BIODIVERSA, coordinated by UANT and will investigate how biodiversity conditions (dis)favour transmissions of infectious agents from small mammals into human populations in African forests (including Ivory Coast).

https://www.uantwerpen.be/en/research-groups/eveco/research/main-ongoing-project/biodiv-afreid/

- INSULA in Guadeloupe: INSULA is funded by FEDER, coordinated by CIRAD and aims to assess the influence of ecosystem biodiversity and their anthropogenic changes on vector-borne diseases of humans, animals and plants in Guadeloupe.

https://www.cirad.fr/en/worldwide/cirad-worldwide/projets/insula-project

The BCOMING project will standardise and complement the data collection initiated by the consortium members in the ongoing collaborative projects listed before. The data used for BCOMING analyses and models will originate from three biodiversity hotspots. In each region, study areas will be identified across an anthropisation gradient to represent 1) very low population density in pristine habitats; 2) low population density in fragmented habitats with rural agriculture; 3) high population density and intensive agriculture; and 4) very high population density in urban environments (Table 1). Two to four

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

sites, depending on the area size, will be selected for a similar sampling effort within each identified area. The project will particularly focus on the transition between pristine and fragmented forests where biodiversity loss is a major issue and biodiversity conservation an important stake.

Table 1. List of BCOMING study sites per region and type of habitat and main source of data collection

List of BCOMING study sites per region and type of habitat and main source of data collection					
Region	**Country**	**Pristine habitat**	**Fragmented habitat**	**Intensive Agriculture**	**Urban**
Southeast Asia	Cambodia	Mundolkiri forest[3,*]	Stung Treng Forest[1,3,*]	Battambang province[2,*]	Phnom Penh[2,*]
West Africa	Guinea	Ziama forest[4,*]	Past Ebola emergence sites in Forested Guinea province[4,*]	Kindia province1[,*]	Conakry [2,*]
Caribbean Caribbean	Ivory Coast	Tai Forest[1,*]	Tai Forest vicinity[1,*]		
	Guadeloupe	Tropical rainforests on Basse Terre [5,*]	Tropical rainforests on Basse Terre with human activities [5,*]		

[1]Biological samples will be collected longitudinally in BCOMING.[2]Biological samples will be collected cross-sectionally in BCOMING. Biological samples and associated ecological and socio-economic data from ZooCoV[3], Ebo-Sursy[4] and Insula[5] projects will be available for BCOMING. *Environmental samples will be collected in all sites using eDNA kits provided by Nature Metrics

Data collected in both terrestrial and aquatic ecosystems will be standardised in each sampling area to build datasets, including information about zoonotic pathogens in animals and humans, biodiversity levels and associated ecosystem services, socio-economic and environmental factors affecting the risk of transmission of infectious diseases.

Biological samples from hosts

The same biological samples will be collected across all BCOMING study sites to feed the standardised dataset (Table 2). The focus will be on bats and rodents since they are the most diverse orders of mammals and thus a major source of zoonotic pathogens. Furthermore, they are hosts to major zoonotic diseases in our study areas (SARS-CoV-2 in Southeast Asia and Ebola and Marburg in West Africa). Since they are the source of zoonotic diseases such as Buruli ulcer and trematodes infections, and because water can be used to monitor other infectious diseases such as COVID-19, the aquatic systems in the study areas will also be sampled in the effort to implement a global exosystemic approach. Thus, a mix of longitudinal and cross-sectional protocols will be designed (Task 2.1) (Table 1) depending on the current knowledge about the host-pathogen system under consideration and the objective (feeding WP3, WP4 and WP5 data analyses and models or testing WP2 and WP6 solutions).

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Table 2. List of samples included in the standardized dataset

List of samples included in the standardized dataset		
Sample sources	**Sampling methods**	**Sample types**
Wildlife - Bats	Mist nets, harp trap	Blood, oral and rectal swabs, feces, urine, wing punch, ectoparasites
Wildlife - Rodents	Traps and euthanasia	Blood, oral and rectal swabs, feces, heart, lung, liver, spleen, kidney, urine, ectoparasites
Wildlife - Fish	Fish nets and traps	Fin, clip, liver, spleen, kidneys
Wildlife - Other	Opportunistic sampling	Wild meat, carcasses, feces
Domestic animals	Veterinary surveys	Blood, feces
Humans	Public health surveys	Blood, feces
Environment: Water and soil	Standardized eDNA kits	Kit sample

Host (e.g. age, sex, and reproductive status) and site (e.g. geographic coordinates, habitat) metadata will be systematically recorded during pathogen sampling sessions. A standardised questionnaire will be used to record socio-epidemiological data as well as potential exposure factors for humans.

All animal and human sampling will follow national and international guidelines and regulations (OIE Animal Health Code, WHO guidelines, CITES and the Nagoya Protocol). Approval from national and European ethics committees are detailed in the corresponding ethics deliverables (D9.1 for humans, D9.4 for Animals, and D9.5 for Non-European countries). Key staff members in the project from CERFIG, IPC, CIRAD, IRD, HZI and NM have long and documented experience in performing animal, human and environmental sampling. Biological samples will be stored in solutions allowing preservation of RNA and DNA at ambient temperature in the field for subsequent molecular analysis (Trizol, RNA later, NAP or ethanol), and in a virus transport medium for a subset of samples for virus isolation purposes. Where and when possible, samples will be placed in liquid nitrogen, transported to the laboratory, and immediately frozen at -80°C for future use. All biological samples collected during the project will be stored in the existing biobanks managed by the consortium partners. Associated to their related metadata and the BCOMING database, this will provide us with a unique standardised dataset to further study zoonotic risks across a large range of ecosystems.

The protocols for biological sampling are detailed in deliverable D9.1 for humans and D9.4 for animals with estimations of the number of samples that will be collected per site and protocol. Overall, thousands of biological samples will be collected during the project. Because most of them will come from the capture of wild animals, which can yield highly variable capture rates, an accurate estimation of the number of samples is impossible at this stage of the project. An update of the number of biological samples will be provided in the next versions of this deliverable.

Pathogen and epidemiological data

The pathogen detection and characterization methods developed during the project will focus on:

(i) three viral families known to comprise major zoonotic agents (coronaviruses, filoviruses and arenaviruses)

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

(ii) microbial communities (bacterial and viral) and parasites in both aquatic and terrestrial ecosystems, including potential new zoonotic pathogens.

Each biological sample will thus provide with a line of data regarding the presence or absence of the pathogens and microbial community of interest. Combined, these results from the detection methods implemented in the project will provide with classical epidemiological indicators such as prevalence and incidence rates that will allow to analyse the impact of biological, environmental and socio-economic factors on zoonotic risk (cf next section of this deliverable).

Additionally, genetic sequences data obtained from the pathogen characterization of positive samples will allow for phylogenetic analyses. Pathogen phylodynamics data will be established by analysing and comparing the newly generated sequences with those obtained by datamining the European and International public databases (ENA and NCBI) to extract all the available sequence datasets of bat coronaviruses, bat filoviruses and rodent arenaviruses in different geographic and environmental settings. This will allow to better understand the factors contributing to the spread and evolution of zoonotic pathogens of interest.

Biodiversity and Environmental data

At each study site, standardized diversity indicators, commonly used to assess biodiversity in ecology such as species richness, Shannon's diversity index or Simpson's diversity index will be estimated using species identified using the eDNA kits developed by Nature metrics and historical data provided by Fauna and Flora International (FFI). The eDNA kits, already successfully used in some study sites (Guinea and Cambodia) included in the project, will be used in all study sites to provide a standardized measure of biodiversity for water and soil samples. For each study site, data will be collected to describe the environment, the habitat type, the biodiversity, the climate.

Additionally, environmental indicators commonly used in ecology and epidemiology studies and known to impact the risk of zoonotic risk (such as forest cover, land use change or landscape fragmentation) will be produced from open satellite remote sensing images (Copernicus global land services). Potential drivers related to climates and the environmental suitability for vectors or hosts will also be extracted from global database. Data will preferably be extracted from the web (open-source database including Copernicus and other EU databases) and processed for correction of known errors (i.e., clouds and atmospheric interference) and should have as few gaps as possible. Processed imagery (timeseries and Transform Fourier Analysis (TFA)) will be provided at a 1 km resolution across the study extent. In order to fully incorporate existing and upcoming European infrastructure, we will work together with Terrascope to develop the different satellite processing chains. Existing and newly developed processing chains will be implemented on the Terrascope platform so that raw data doesn't need to be downloaded and that we use the latest state of the art imagery from the Sentinel platforms where possible.

Human socio-epidemiological and health data

Human socio-epidemiological and health data will be collected by field surveys, using a standardized questionnaire, or by reusing previously collected datasets to evaluate the exposure of study participants to emerging infectious diseases. The questions and data that will be collected from study participants are detailed in Deliverable 9.1. They will include identification and localisation data, some socio-economic data and data relating to health and also personal and farming practices of the study participants. No genetic nor biometric data will be collected.

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

To complement the questionnaire, qualitative data will be collected from a Human-Centered Design approach to document the existing human needs, attitudes, preferences and behaviours, including wildlife trade (hunting, consumption and trading behaviours (e.g. bushmeat) of people).

Collection and treatment of these data will strictly follow the European and national legislations relating to protection of personal data as detailed in Deliverable 9.3 POPD. Therefore, collection of data will be conducted in strict compliance with the local legal and ethical framework when this legislation exists such as in Guinea, Ivory Coast and France or, for countries without such legislation like Cambodia, with strict compliance with the European legislation (Deliverable D9.1 Humans and D9.5 NEC). The procedures to collect the free and informed consent of each study participant are fully described in the Deliverable D9.3. Only individuals who have signed the consent form may then see their personal data collected and processed as part of the BCOMING activities.

To ensure the security and confidentiality of these data, including protection against unauthorised or unlawful access, accidental loss, destruction or damage, all partners will implement appropriate measures as described in the Deliverable D9.3. Data will be pseudonymised or anonymised (e.g. aggregation) before processing, sharing or transfer. No personal data will be transferred from a non-EU country to the EU prior to anonymisation of the data. Only the partner in charge of the activity will have access to the non-anonymised datasets and to the corresponding consent forms. Anonymised datasets will then be shared with the consortium as part of the central project database.

Data utility

The raw data collected in BCOMING will be analysed with different methods to improve our understanding of the emergence of zoonotic diseases and to develop tools to prevent future emergences:

Spatio-temporal models for integrated risk assessment and risk mapping:

The main objective of this spatio-temporal statistical modelling is to assess both the risk of disease spread and the emergence from a set of known risk factors across a range of fragmented habitats and landscape. In the BCOMING project, we aim to embed the effect of covariates into process-based representations of social networks, agent-based dynamics, as well as evolutionary and propagation dynamics. We propose to assess and map the risk to support WP4 analysis in terms of species communities; further study the pathogens spillover and risks based on host contact patterns, the spread of infections and inference of model parameters and hidden processes (inference of the role of reservoir or hidden components); the connection between household behaviour with climate, biophysical, institutional economic and social factors; and future adaptive responses to disease exposure and/or proposed remediation/control efforts. A key innovation will be to draw realistic spillover, emergence and spread scenarios by coupling, at the scale of hosts-and-pathogens communities and in explicit geographic contexts, the dynamics of adaptation and populations on the one hand, and the effects of landscape and environment on the other hand. Thanks to the range of biodiversity levels in the selected study sites, comparison of the modelling results will be provided to study the impact of biodiversity loss and other risk factors on zoonotic pathogen circulation levels.

Microbiome analysis, impact of biodiversity levels on microbiomes composition and structure and the presence of zoonotic pathogens:

The main objectives of the microbiome analysis in the BCOMING project is 1) to characterise the influence of biodiversity on microbial communities, providing additional microbial diversity metrics and

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

2) to identify potential interactions between bacteria and/or microbiome structure and zoonotic pathogens prevalence. To go beyond microbiome description and provide a preliminary assessment of the impact of biodiversity on microbiome structure and of zoonotic risks, we will analyse the gut microbiota of a subset of 2000 individuals from different mammal groups (bats, rodents, domestic animals and humans) along the urbanisation gradient using metagenomics approaches, and identify potential correlations between biodiversity metrics and bacterial richness and diversity. Using bacterial association networks analysis, we will investigate the relationship between microbiomes composition and structure and zoonotic pathogen prevalence in wildlife, and assess whether bacterial consortia or network structures associated with zoonotic pathogens can be found in farm animals and humans. Unveiling how biodiversity at multiple levels influences the risk of zoonotic emergence will ultimately inspire mitigation strategies (tested in WP5) based on evidence linking the environment and health.

Host community models: Joint Species Distribution Model at pathogen level:

The main objectives of using a Joint Species Distribution Model at the pathogen level is (1) to understand what factors determine the spatiotemporal assembly of the pathogen community detected in WP2, (2) use these key determinants to guide decision about SIR model architecture and (3) predict the effect of host community changes on parasite composition. In the BCOMING project, parameterisation of these models with empirical data from various locations will allow to analyse how the pathogen communities are associated with ecological covariates (environmental and climatic covariates, host species diversity and host abundance), reservoir covariates (individual host species, age, sex, weight and microbiome structure), and temporal variables (season and year). Variance partitioning over these groups will enable us to filter out the important associations. In addition, the model permits us to investigate how interspecific variation in covariate response can be attributed to functional traits; for example, the influence of biodiversity loss on the transmission mode. This innovative modelling approach will lead to predictions on how parasite composition alters under different biodiversity change scenarios.

Mechanistic models:

The main objectives of mechanistic modelling in the BCOMING Project are to (1) make predictions about the dynamics of the circulating pathogens in a wildlife population; and (2) design cost-effective surveillance strategies. In the BCOMING project, three SIR models will be developed based on data previously collected in the study areas of the project: coronaviruses in insectivorous bats in Cambodia, filoviruses in frugivorous bats in West Africa and West Nile virus in Culex mosquitoes in Guadeloupe. To go beyond the state of the art, a key step will be the quantification of uncertainties using Bayesian techniques (Markov chain Monte Carlo), thereby permitting uncertainty to be fully integrated in the prediction steps. Another key innovation of the SIR models developed in the project will be to incorporate the logistical and economic constraints of surveillance in the study areas in conjunction with Task 6.4. Therefore, cost-effective sampling strategies maximising the probability of the detection of the pathogen will be tailored to each case-study.

Results from these analyses will be of interest to a large range of actors:

In academia, researchers working on the surveillance of emerging zoonotic diseases but also scientists in social sciences, climate change adaptation, biodiversity, landscape planning & economics interested by the risk of emerging zoonotic diseases will benefit from the scientific results and publications from the project to fuel their own research.

Local actors such as representatives of wildlife associations, farmers, hunters, consumers of wildlife, value chain stakeholders (retailers) that live at the interface with wildlife will benefit from the project in

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

several ways. The risk of emergence of zoonotic diseases from wildlife will be reduced thanks to an improved detection capacity of zoonotic pathogens and a better understanding of the factors leading to emergence.

National authorities as well as the European Commission and its agencies, international bodies and government representatives will benefit from the development of an integrated management of zoonotic exposure, biodiversity and livelihoods based on the One Health participatory systems approach.

FAIR data

At this stage of the project no data has been collected and an overview of the historic data is being prepared to summarize the main datasets coming from previous projects that will be available for BCOMING analyses (Deliverable 3.1). Nevertheless, partners of the project already have various means to make their data FAIR and we can already anticipate the following measures to ensure it for BCOMING. A finalized data management plan detailing how BCOMING data have been and will be made FAIR will be submitted at the end of the project as Deliverable D1.5.

Making data findable, including provisions for metadata

Data from the BCOMING project will be centralized in a database hosted by CIRAD, facilitating findability and the access of the data. A public interface, similar to Ebo-Sursy's portal of scientific data (https://database.ebo-sursy.eu/public-interface/) will facilitate the findability of the data. Additionally, these data will be made available in public repository upon publication.

At minima, partners will be compelled to provide a basic metadata record that complies with the OpenAIRE application of the DataCite Metadata Schema to describe the datasets in the BCOMING databas. To ensure findability, interoperability and reuse, datasets will be described and documented as much as possible using domain-specific metadata and controlled vocabularies. Examples of such standards are Mesh or UMLS (Unified Medical Language System) for Health data, PDBx/mmCIF for proteomic, MIBBI for biological data, GCMD (Global Change Master Directory) or CF (Climate forecast) for co-variates, ISO 19115 for spatial data, DDI for survey data, etc. When necessary, other appropriate standards will be searched in the Research Data Alliance Metadata Standards Directory (https://rdamsc.bath.ac.uk/).

Disciplinary standards will be used for data and metadata. The data generated by the project will arise from a number of different interrelated fields. Therefore, no single metadata standard will apply to all the cases. MIxS genomic standards will be used, but we will check data generators to identify suitable standards from the Research Data Alliance Metadata Standards Directory, which will include PDBx/mmCIF and MIBBI or use the Fairsharing registry to find the appropriate metadata standards. The choice of the repository may also depend on the data repositories recommended by the journal chosen for publication.

Making data accessible

The data created by the partner institutions will be accessible following their open access/science policies. Furthermore, data deposited by the partners in the BCOMING database will be available upon request to the partner in charge of the dataset. In case the datasets would be used for

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

commercial purposes, partner institutions would sign an agreement with the third party based on fair and reasonable conditions. Access modalities will follow the Horizon Europe's Programme Guide's recommendations related to open science.

Additionally to the BCOMING data base, to increase accessibility to the data, relevant datasets will be shared by deposition into relevant open access and trustworthy data repositories with attribution of an open licence such as the Creative Commons CC-BY. When disciplinary-specialized repositories are available, such as for genetic sequences, they will be used to share the datasets. For example, data will be deposited in GenBank, Dataverse and/or GBIF for biodiversity data. Otherwise, repositories will be selected from the EU recommended registry (http://www.re3data.org/) during the course of the project and according to the type of data and to the public targeted. Besides, institutional repositories such as the Cirad Dataverse (https://dataverse.cirad.fr/) or the European repository Zenodo (https://zenodo.org/) may also be selected in the absence of specialized repositories. Therefore, for each dataset produced and deposited in a data repository, a unique ID will be associated that will be included in the articles published by the project to improve data visibility, access, sharing, and reuse.

As described above, a combination of domain-specific-metadata standards will be used to describe the BCOMING datasets. When depositing the datasets in appropriate domain-specific repositories, depositors will make their best effort to document a maximum of scientific metadata available in the repository.

Making data interoperable

To ensure interoperability of data and research outputs, BCOMING will follow pre-existing standards recommended by scientific communities (such as: norm ISO 8601 for the dates, ISO 19115/32018 for geolocalisation, EML for ecological data, Norm DDI for survey and social data, those of the ECDC for epidemiological data, as well as WHO, OIE and RDA recommendations). In principle, data and metadata will be requested, stored and transferred (across partners) in a comma-separated values (CSV) format. MS Excel-compatible files, including .xls(x) format will be also accepted to facilitate exchange. Other data formats that will be used in BCOMING include .fasta, .png, .gml, and .nexus. For statistical purposes, other formats include .sas7bdat (SAS), .RData (R), .SAV (SPSS) and .mat (matlab). As far as possible, the consortium will reuse existing controlled vocabularies for providing metadata to resources.

Increase data re-use

To facilitate reuse, data will be deposited in relevant repositories under open formats and open licences such as the Creative Commons CC-BY 4.0, or equivalent, not excluding the use of the CC-BY-NC licence for some specific datasets. Core project partners will be encouraged to openly deposit their data using a Creative Commons version BY 4.0 licence, or equivalent.

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

Other Research outputs

Several models will be developed during BCOMING

- SIR model for surveillance optimisation
- Agent Based Models for the support of participatory workshops

These models will be shared through publications in scientific journals and through the publication of the codes in repository such as Git Hub (https://github.com).

Several diagnostic methods for pathogens will be developed during BCOMING

- PCR protocols for the detection of coronaviruses and other bat-borne and rodent-borne pathogens
- Luminex assays for the detection of anti-coronavirus and anti-ebolavirus antibodies in bats and humans.

The assays for these detection tools will be published in scientifc journals and the protocols may be detailed in platform such as https://www.protocols.io/ with a reference in the scientific paper.

Allocation of resources

Data management coordination:

For the preparation of the initial version and the update of the DMP during the project, 2 staffs from CIRAD are planned to work 6 PM during the project.

Additionally, a data Manager from CIRAD is planned to work 4 PM for BCOMING for the development and management of the BCOMING database.

All partners will also contribute to the management of data and to their publication.

CIRAD long-term infrastructure

The BCOMING database will be located on a server hosted by CIRAD during and after the project. All costs for maintenance of the database will be covered by the IT department of CIRAD during and after the project.

The parameters for calculating the maintenance cost are: Memory size, Storage space and Saved space.

The cost of memory is 88,583 €/Go/year

The cost of the storage space is 1,002 €/Go/year

The cost of the backup is 1,289 €/Go/year

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

So for the BCOMING server which will have 4 GB of Ram and 50 GB of disk space the annual cost will be 4 * 88,583 + 50 * 1,002 + 50 * 1,289 => 468,882 €/year => for 5 years : 2344,41 €.

s.

Open Science budget by partner:

Some BCOMING partners have included in their Budget specific ressources to support their open science policy in BCOMING:

CIRAD 6,000 EUR

IRD 10,500 EUR

IPC 6,000 EUR

INRAE 3,000 EUR

CERFIG 5,000 EUR

Data security

Data security will be organised at two levels in BCOMING, the partners collecting the data will ensure the security of the data upon collection before anonymising them and uploading them into the central database developed and hosted by CIRAD.

The first level of data storage will thus be provided by the BCOMING partners collecting and producing raw data: CIRAD, CERFIG, FFI, HZI, iDE, IPC, IRD, UL. All these partners have proper IT systems for data storage on high security levels. For example, for the BCOMING database hosted by CIRAD, the full server is saved daily by the IT department and is available for 7 days. Additionnally local saves are made by the data manager monthly and a copy of each dataset is sent to the relevant partner monthly.

To ensure the security and confidentiality of processed personal data, specific measures are implemented such as physical and organisational access restrictions, authentication and authorisation processes, workstation security, logging and monitoring, incident management, production of backup and copies, data pseudonymisation and anonymisation so as to limit the risk of any type of unlawful or accidental data breaches, including destruction, loss, alteration, or unauthorised disclosure of the datasets.

The second level of data storage will concern the data to be shared and centralized among partners to fuel the analyses planned in the project. Once processed, the raw datasets will be shared within the project, in an anonymised version, constituting a central project database managed by CIRAD as part of WP1 “Coordination and management”. This central database will be secured through the CIRAD’s institutional information system security measures as described in the Deliverable D9.3:

Organisational measures compiled in various institutional documents:

A yearly reviewed information system security policy,

A best practice framework communicated to scientists, server managers in the scientific units and IT specialists,

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

An IT charter.

A Risk and Compliance Management in place to document and manage significant events and the risks they involve.

A security management platform, combining SIEM and vulnerability scanning functions, measuring the security level of the servers.

A source code management and static security/quality analysis platform regularly measuring the security level of application codes.

Moreover, the database will be protected by specific security measures:

The database will be accessible only if the user has created an account validated by the project administrators. Access to individual accounts will be conditioned to a valid email address and a password including at least 8 characters, a special character and at least one lower case and one upper case letter.

Users can only access their team's data. A 4-digit code is provided to the team so that their agents can validate the registration (in the registration form, the agent chooses the name of his team in a drop-down menu and the 4-digit code provided must correspond to their team).

Ethics

The data collected in BCOMING are raising several issues due to their nature (personal data, human samples, animal samples) and the location of their collection (natural environment in Non-European countries).

The consortium is well aware of all the ethics requirements associated with the collection, management and use of such data as all partners have collaborated in the past on this topcis.

All these Ethics requirements are covered in details in the following deliverables:

D9.1 – Humans – Submitted January 2023

This deliverable is notably covering the informed consent procedures for the collection of personal data and details the sampling procedures.

D9.3 – Personal data – Submitted August 2022

This deliverable is detailing the collection, management and use of personal data in the project. All items are strongly linked with the data management plan.

D9.4 – Animals – Submitted January 2023

This deliverable is notably covering the procedures of capture, sampling and data collection from animals, including the authorizations required to conduct such studies.

D9.5 – Non-European Countries – Submitted January 2023

This deliverable is notably covering the procedures for the transfer of biological material between partners and the fair benefit sharing of project data and results. In particular, the research carried out in the project is in strict compliance with the rules and legislation EU Regulation (EU) N°

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.

511/2014 and Commission Implementing Regulation (EU) N° 2015/1866 relating to the Nagoya Protocol on Access and Benefit Sharing (ABS).

Other Issues

No other specific issues have been identified at time of submission of this initial version of the DMP.

This project has received funding from the European Union's research and innovation programme under grant agreement No 101059483.