BEXYL
BEYOND XYLELLA

Data Management Plan (DMP)

Deliverable 10.4

20 February 2023

Author(s): Ana Sánchez and Blanca B. Landa (CSIC), Antonio Vicent (IVIA)

Funded by
the European Union

DELIVERABLE INFORMATION

Start of the project:	01 Sep 2022
Duration:	48 months
Coordinator:	Blanca B. Landa
Deliverable title:	Data Management Plan (DMP)
Deliverable n°:	D10.4
Nature of the deliverable:	DMP
Dissemination level:	Public
WP contributing:	WP10
Lead beneficiary:	CSIC
Due date of deliverable:	28 Feb 2023
Actual submission date:	20 Feb 2023

Deliverable status

Ver.	Status	Date	Author(s)
1.0	Final	20 Feb 2023	Ana Sánchez and Blanca B. Landa (CSIC), Antonio Vicent (IVIA)

Funded by the European Union

TABLE OF CONTENTS

Funded by the European Union

Funded by the European Union

ABSTRACT

Data management plans (DMPs) are a cornerstone for responsible management of research outputs, notably of data used, obtained or developed by a research project, and which are mandatory in Horizon Europe for projects generating and/or reusing data.

DMPs are formal documents that outline from the start of the project all aspects of the research data lifecycle, which includes its organisation and curation, and adequate provisions for its access, preservation, sharing, and eventual deletion, both during and after a project. Writing a DMP is part of the methodology of the project, since good data management makes the work more efficient, saves time, contributes to safeguarding information and to increasing the value of the data among the beneficiaries themselves and others, during and after the research is finished. DMPs are thus a key means of support when planning and conducting a research project.

DMPs play a key role in helping researchers to adequately manage research outputs other than reports and publications, also in line with the guidelines on FAIR (Findable, Accessible, Interoperable and Reusable) Data Management principles from the European Commission. Such research outputs may be physical or digital, and include original software created during the project, workflows, protocols, new materials such as samples, host plants and vectors, DNA sequences, among many others. DMPs should reflect an adequate management strategy for such outputs as well.

This deliverable presents the DMP developed for BeXyl Project.

Funded by the European Union

1. INTRODUCTION

Research data management is mandatory in Horizon Europe for projects generating or reusing data. BeXyl project expects to generate and reuse data and other research outputs. So, this document provides a Data Management Plan (DMP) as part of the methodology of the project, since good data management makes the work more efficient, saves time, contributes to safeguard information and to increase the value of the data among the beneficiaries themselves and others.

DMP is the formal document that outlines from the start of the project all aspects of the research data lifecycle, which includes its organisation and curation, and adequate provision for its access, preservation, sharing, and eventual deletion, both during and after the project.

It also provides guidelines to BeXyl partners on how to make the project results accessible and easily findable, while simultaneously try to minimize the workload involved.

2. DATA SUMMARY

As practices in relation to data management, storage, and sharing differ widely across disciplines, the DMP should reflect common disciplinary practices. In addition to domain specificities, DMP across the board should address an overarching set of data-related requirements including those aspects related to making the data FAIR (Findable, Accessible, Interoperable and Reusable), following the principles of Horizon Europe Annotated Grant Agreement (AGA) in Annex 5 HE IPR RULES.

For this reason, the following guidelines are set out in this section of the DMP:

- State the purpose of the data collection/generation.
- Explain the relation to the objectives of the project.
- Specify the types and formats of data generated/collected.
- Specify if the existing data is being re-used (if any).
- Specify the origin of the data.
- State the expected size of the data (if known).
- Describe the data utility: to whom it will be useful.

2.1. Purpose of the data collection/generation and its relation to the objectives of the project

BeXyl has been built mainly on the foundations of the knowledge generated by the H2020 project XF-ACTORS to strengthen the EU research network tackling *Xylella fastidiosa* (*Xf*) outbreaks, to ensure the best use and exploitation of the research outputs gained so far.

For this reason, many data will be collected and new data will be generated to develop and implement tailored Integrated Pest Management (IPM) strategies to mitigate the impacts of current *Xf* outbreaks.

The use of open science practices in the BeXyl project (foreseen in the Horizon Europe programme) increases the impact of research and accelerates progress in relation to

Funded by the European Union

improving capacity building for the prevention of new *Xf* introductions and the management of current outbreak areas in Europe.

2.2. Types and formats of data generated/collected

Public access to research data and published material from the project BeXyl will be guaranteed according to these categories, not excluding some new that can appear during project development:

- **Training materials and publications:** lists of host plants and vectors, photo descriptions of symptoms, scientific and extension publications, etc.
  Molecular and genetic data: DNA sequences and genomes, RNA, MLST, HTS, barcoding, microsatellites, etc.
  Protocols, code and raw data: geospatial datasets, airborne image data, modelling codes, maps, eco-physiological and meteorological data, protocols for pathogenicity tests will be described in all scientific publications. In addition, protocols, program code and raw data will be included in institutional repositories.
  Physical access to biological materials: Access to insect specimens, bacterial strains, etc. will be granted by depositing representative materials in public culture collections, observing the pertinent biosecurity provisions.

2.3. Existing data re-use and data utility

The data/datasets collected or generated in the context of BeXyl will be available to external users under specified license terms and subject to ethics considerations. A full description of the available licenses can be found in section 3.4.1.

The potential users considered for research data management in BeXyl is a broad group, encompassing from those experts interested in specific basic research to the stakeholder's community. BeXyl envisages that the community of contributors will be composed of the following groups:

- Farmers/foresters and Nurserymen
- Agronomy/forestry consultants and technical staff
- Farmers' organizations and associations
- Diagnostic laboratories and phytosanitary inspectors
- Official agencies, plant health services and policy makers
- Research community (including the Advisory Board)
- Agri-food industry
- Scientific and daily press.

A distinction can be made between research datasets supporting peer reviewed academic publications and datasets forming the basis on which the project deliverables are elaborated. The research data underlying project deliverables requires a big investment in terms of human resources, therefore the main aim of making this data public is for others to use and build on this data, preventing duplication of work and improving efficiency, quality and speed

Funded by the European Union

of research. This also holds true for scientific publications, but for these validation and reproducibility are key considerations as well.

2.4. Origin of the data

The origin of the data comes from a wide variety of sources:

- Training materials as lists of host and vectors, descriptions of symptoms.
- The results obtained by the project partners in different types of experiments.
- Maps and images acquired or processed and developed by project partners.
- Field and greenhouse data on *Xf* response of diverse crops and cultivars, collected in previous projects (e.g. POnTE and XF-ACTORS) and during BeXyl development.
- Airborne hyperspectral imagery from the aircrafts used in XF-ACTORS and new ones obtained in BeXyl and commercial satellite systems.
- Modelling codes, program code and raw data from previous research.

2.5. The expected size of the data

The data collected/ generated by BeXyl is very heterogeneous in size ranging from few megabytes per item (e.g. rates of disease data) to many gigabytes (e.g. satellite images or genome sequences).

The estimated amount of data to storage at the time of writing this DMP is:

.

- *Experiment data*: An estimation is difficult to forecast, because each partner uses different data and procedures, but we expected around 1 Gb per partner, no more than 30 Gb in total.
- *Models*: The sum of all datafiles containing alphanumeric data will not exceed 1 Gb.
- *Satellite and remote sensing imagery*: 1 Tb per location of study and year including hyper, thermal and satellite imagery. It is estimated a total of 10 Tb.
- *Publications, dissemination material and additional data*: We expected around 0,5 Gb per partner, no more than 15 Gb in total.
- *Molecular and genetic data*: We expect around 1Gb per *X. fastidiosa* genome including raw Illumina and Oxford Nanopore reads, and assembled genomes. A total of 500 genomes are expected to be sequenced yielding less than 1 Tb. RNASeq data from genome expression experiments will represent around 1-2 Gb per experiment, representing less than 1 Tb in total.

2.6. Data Utility

Although it is difficult to predict how and by whom the specific data will be used in the future, this DMP is written with the understanding that the data generated by BeXyl will be useful at least to:

- Researchers working in prevention and practical solutions for disease control.

Funded by the European Union

- Agronomic experts to prevent, monitor, and find adequate responses to EU plant pests.
- Researchers and policy experts evaluating the socioeconomic implications for farmers affected by the plant pest.
- Researchers and policy experts interested in relevant information on resistant/immune or more resilient host plants used to restore current *Xf* outbreak areas and facilitate sustainable disease management.
- Companies commercialising resistant cultivars to *Xf.*
- Undergraduate and graduate students in crop protection.
- Farmers and technical staff in farms.
- General public.

3. FAIR DATA

The purpose of this data management plan is to issue clear guidelines to make sure that the data generated under the project that can be made public are both open and findable, accessible, interoperable and reusable (FAIR) for both WWW search engines and individuals.

3.1. Making data findable, including provisions for metadata

According to the Guidelines on FAIR Data Management of European Commission, this section of the Data Management Plan is expected to give answers to the following issues:

- Outline the discoverability of data (metadata provision)
- Outline the identifiability of data and refer to standard identification mechanism.
- Outline naming conventions used.
- Outline the approach for clear versioning.
- Outline the approach towards search keyword.
- Outline the approach for clear versioning.
- Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how.

3.1.1. The discoverability of data

BeXyl project plans to make discoverable the dataset and documents (articles, technical reports, ...) generated within the project uploading them as open data in institutional repositories which will require metadata and will provide by default Digital Object Identifiers (DOI). To simplify operations, BeXyl will differentiate between the following categories:

Training material and publications
Because of mandatory open science practices in Horizon Europe research outputs will be published in open access journals. Where possible, the Open Research Europe (ORE) open access publishing platform of the European Commission will be chosen. In this system, names of authors and reviewers are public, the review reports are open access published alongside the article, allowing reciprocal discussion between authors and reviewers.

Funded by the European Union

When appropriate, preprint versions of a manuscripts will be either submitted to a recognized preprint repository (e.g., bioRxiv or agriRxiv) or directly deposited in institutional repositories (e.g., at P1-CSIC: DIGITAL.CSIC; P2-IVIA: ReDivia; P20-EPPO: Euphresco DROP).

Nevertheless, as a default all documents generated in the BeXyl project can be uploaded to the DIGITAL.CSIC, as this is the coordinating institution of the project.

Once uploaded, the authors will inform to the Coordinator Team (CT) to include the publication's DOI in the list of project results at the project's website (https://www.bexylproject.org).

All this process is explained in details at D10.2 BeXyl Operational Manual (Point 3.5 Management of publications).

Molecular and genetic data
Molecular and genetic data will be deposited in NCBI/GenBank/SRA and/or ZENODO, whenever needed. Accessions to these data will be conveniently included in all published material (e.g., scientific and technical articles, etc.).

Protocols, code and raw data
Protocols, program code and raw data will be included in ZENODO, GitHub or Bitbucket repositories. For statistical analyses, open-source statistical software such as R and Python will be used. For relevant studies, model code will be externally checked for reproducibility using the code-check platform.

3.1.2. Naming conventions used and clear versioning

Naming conventions are useful when there is the need to manage a large number of documents on similar topics, as in BeXyl. The naming conventions used for BeXyl Project are the following (defined as codification in D10.2 BeXyl Operational Manual):

- Identification: BeXyl
- Filename: Description of data (Minutes, Report etc.).
- Version: 01 to 99.
- Extension: File extension.

Example: BeXyl_DX.X_01.doc

3.1.3. Research keywords provision

As a rule, the keywords will provide the necessary information on the document subject, but it is recommended to use one or two that make a clear link to BeXyl. If possible, it is recommended to used one or several of these: @bexylproject, #BeyondXylella, #BeXylinfo, #Xylellafastidiosa

Funded by the European Union

3.1.4. Type of metadata to create

Meta data are extremely important to ensure that our results and findable in the WWW for search machines or individuals. To consolidate this the Research Data Alliance (https://rd-alliance.org/) offers a Directory of Metadata Standards. Data repositories may also provide guidance about appropriate metadata standards. In this case we have followed DIGITAL.CSIC repository to define the fields.
It is always useful to consider what other documentation is needed to enable reuse. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data and software used to collect and/or process the data.

To ensure a traceable route for quality control and high visibility, all datasets and documents submitted to the database or a repository from BeXyl must contain the information defined in this metadata template as a minimum.

Table 1: Template for metadata

Dataset/Document Content, Use and Provenance	
Authors	Who generated the data
Title	Description of the dataset
Date	Publication date
Publisher	if applicable
Citation	Refers institutional publication
Reference	Refers papers (if applicable)
ISSN	From paper
Document Type	Xls, jpg…
Spatial situation data	For datasets
Temporal situation data	For datasets
Matadata format and standards characteristics	
Key words	#...
Summary	Summary of the document/dataset
Sponsors	If applicable
Funding Agency	EC in this case
Description	Description of the information uploaded
URI	Link to the URI
DOI	Link to the DOI
Grant Agreement	BeXyl GA nº 101060593
Data access and Sharing	
Data access policy	Open/restricted/embargo
Licenses	Link to license
Embargo date	If applicable
version	Preprint/Postprint
peer revised	YES/NO
Dissemination/reuse procedure	Link to repository
Responsible	Responsible for uploads

Funded by the European Union

3.2. Making data openly accessible

According to the Guidelines on FAIR Data Management of European Commission this section of a Data and Publication Management Plan gives answers to the following issues:

- Specify which data will be made openly available. If some data is kept closed provide rationale for doing so.
- Specify how the data will be made available.
- Specify what methods or software tools are needed to access the data. Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open-source code)
- Specify where the data and associated metadata, documentation and code are deposited.
- Specify how access will be provided in case there are any restrictions.

3.2.1. Openly available data

Open Research Data (ORD) procedure does not necessarily mean opening up all your research data. Rather, the ORD follows the principle "as open as possible, as closed as necessary" and focuses on encouraging sound data management as an essential part of research best practice.

Being *Xf* a priority quarantine pathogen in the EU, BeXyl will also adopt the open science practices proposed in the code of ethics for plant health emergencies by the International Society of Plant Pathology (ISPP) Including:

- Early and open sharing, the project complies with the FAIR (Findable, Accessible, Interoperable, and Reusable) principles and data policy;
- Open access and reproducibility of research outputs;
- Open peer review: using Open Research Europe (ORE);
- Citizen, civil society and end-user engagement:
    - Co-design activities involving workshops and focus groups with stakeholders, NPPOs and EPPO participating in outbreak response programs for *Xf* as well as farmers implementing crop restoration and IPM programs in current outbreak areas in the EU.
    - Co-creation activities in BeXyl are fully implemented through involving citizen science, that is also considered to improve early detection of *Xf* with passive surveillance by citizens.
    - Co-assessment activities will be implemented through the Advisory and Stakeholder boards, to monitor and evaluate activities progress and check that proper interactions with the civil society are made.

3.2.2. Accessibility of the data

BeXyl data will be accessible through institutional repositories. Organizations with their own repository will use their own one. For authors without access to an institutional repository, and when they cannot be uploaded in the repositories of other co-authors, the document will be uploaded as a first option or at DIGITAL.CSIC as this organization is the coordinator of the project. As an alternative to DIGITAL.CSIC BeXyl considers the ZENODO repository in the

Funded by the European Union

community OpenAIRE (European Commission Funded Research) and the Euphresco DROP that is managed by P20-EPPO.

3.2.3. Methods or software tools that need to access the data

No specific software tool is needed to access BeXyl data. The data in the repositories are accessible at least individually through a set of Application Programming Interfaces (APIs). This means that a user can query datasets one by one and/or request several datasets if they have been published together (e.g. in a .zip file).

3.2.4. Data and associated metadata, documentation and code deposit

The different institutional repositories where data are planned to be deposited based on the partners availability are:

- Preprint repository:
  - bioRxiv: https://www.biorxiv.org/
  - agriRxiv: https://www.cabi.org/publishing-products/agrirxiv/
- Post print repository:
  - CSIC, DIGITAL CSIC: https://digital.csic.es/
  - IVIA, ReDivia: https://redivia.gva.es/
  - UNITO: IRIS: https://iris.unito.it/
  - ZENODO repository: https://zenodo.org
  - Euphresco DROP: https://drop.euphresco.net/
- Program code and raw data:
  - GitHub: https://github.com/
  - Bitbucket: https://bitbucket.org/product/
  - ZENODO repository: https://zenodo.org

Additionally, the BeXyl website (https://www.bexylproject.org/) will contain some key documents and a comprehensive and clear list of the project data and result with the link to the repositories from which they can be accessed and downloaded.

3.2.5. Provision of data access in case of restrictions

As far as possible, final verified datasets produced by the project will be made available as open access via institutional repositories. Initial versions of the datasets will have restricted access to consortium partners for specified periods as indicated in deliverable descriptions. In the interests of personal data protection, restrictions will be considered for certain datasets (e.g. socioeconomic analysis), in which case anonymised and/or summarised findings will be made available via an openly-accessible report. For dataset that for any reason could not be made open at the end of the project for any reason (e.g. used within the project but coming from private companies or project that restrict their availability) contact to the owner of the data will be provided to allow individual requests for permission to use the data.

Funded by the European Union

3.3. Making data interoperable

According to the Guidelines on FAIR Data Management of European Commission this section of a Data and Publication Management Plan is expected to give answers to the following issues:

- Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.
- Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

3.3.1. Interoperability of data in order data to be exchanged and re-used

BeXyl works towards interoperability of our data through the use of standardize templates for all the generated information as described in section 3.1.4.

3.3.2. Specification of data and metadata vocabularies, standards or methodologies

In order for data to be FAIR the consortium will strive to comply or reuse existing standards whenever possible BeXyl recommends the use of AGROVOC (Caracciolo et al. 2013) developed for FAO and available in http://agrovoc.uniroma2.it/agrovoc/agrovoc/en/index/

3.4. Increase data re-use

According to the Guidelines on FAIR Data Management of European Commission this section of a Data and Publication Management Plan is expected to give answers to the following issues:

- Specify how the data will be licenced to permit the widest re-use possible.
- Specify when the data will be made available for re-use.
- Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project. If the re-use of some data is restricted, explain why.
- Describe data quality assurance processes.
- Specify the length of time for which the data will remain re-usable.

3.4.1. Data license to permit the widest re-use possible

To ensure that the data gathered in the BeXyl project can be made available for re-use, the datasets will be released under a license that allows datasets to be re-used and new work to be derived from them. Partners are being advised about appropriate licensing schemes and license attributions to facilitate data re-use and ability to derive new work from the data and make these works open access.

Funded by the European Union

Scientific publications

Creative Commons licences (https://creativecommons.org/licenses/) are used for every published result. The default licence is Attribution (CC BY), and this will be the encouraged by BeXyl, however the level of restriction will need to be determined on a case-by-case basis. The chosen licence needs to be recorded in the metadata information before to upload the publication.

Possible Creative Commons licences:

- Attribution (CC BY)

  This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

- Attribution-ShareAlike (CC BY-SA)

  This license, which is BeXyl´s encouraged license, lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to "copyright" free and open-source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.

- Attribution-NoDerivs (CC BY-ND)

  This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

- Attribution-NonCommercial (CC BY-NC)

  This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

- Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)

  This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.

- Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

  This license is the most restrictive of the six main Creative Commons licenses, only allowing others to download your works and share them with others as long as they credit you, but they cannot change them in any way or use them commercially.

Funded by the European Union

- Public Domain Dedication (CC 0)

  This license is a public dedication tool, which allows creators to give up their copyright and put their works into the worldwide public domain. CC0 allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, with no conditions.

Metadata

Metadata of deposited publications must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles and provide information about the information as described in section 3.1.4. Where applicable, the metadata must include persistent identifiers for any research output or any other tools and instruments needed to validate the conclusions of the publication.

3.4.2. Data availability for re-use.

The main requirement essentially calls for researchers to provide free-of-charge scientific information (unless it is protected; e.g. patenting) in the form of peer-reviewed scientific publications (Article 29.2 of the Model Grant Agreement) and Research data (Article 29.3), and to ensure continuous online access to it. The full instructions are available in the official Horizon Europe Open Access document.

Peer-reviewed scientific publications

This type of information usually refers to articles and final manuscripts that were assessed by other scholars, typically organized by the journal or the publisher.

These are subjected to Open Access following two steps:

Step one: Deposition in Institutional, subject-based or centralized repositories as described in section 3.2.4.

Step two: Ensuring Open Access to the publications through one of the available options: “Green” or “Gold”. Choosing the right option for you largely depends on your overall dissemination, communication, & exploitation strategy, and availability of resources.

- Green OA → an avenue referring to self-archiving, depending on an embargo period (if any). This option generally consists of providing access to scientific data through one of the following: the Horizon Europe research project’s direct website, the author’s website, the host institution’s website or an independent central open repository.
- Gold OA → refers to paid archiving and the process of depositing scientific publications on websites and platforms that are not your own. Here the Open Access is immediate upon publication. Since the main premise of Open Access is that it is free, the payment is therefore handled by the author and not the reader. Within the execution phase of any Horizon Europe project, the associated costs of such publications are eligible for reimbursement as part of the Horizon Europe grant.

Funded by the European Union

Choosing the right option for you largely depends on your overall dissemination, communication, & exploitation strategy.

The difference between the two types of Open Access uncovers several pros and cons for choosing one type over the other:

1. Namely, self-archiving can potentially require a lot more resources and management on the author's end. But, if you already have a website for your Horizon Europe project, this then seems like a logical step to take. Since a project website can receive less exposure and traffic than a professional publisher, it is important to have a strategic marketing strategy for the website to ensure the scientific data and information extends and reaches its target audience.
2. That being said, depositing information in dedicated professional publishing platforms can help ensure the information is better circulated and reaches a wider audience. Clearly, this option requires setting aside budget for these types of activities, and forming your information to the style and expectations of such platforms.

While scientific articles are the most widely accepted forms of sharing research results, there is a strong urge to also provide Open Access in Horizon Europe to other forms of scientific publications such as books, conference proceedings, and more.

Research data

This type of information includes all kinds of data (e.g. numerical and visual) that were collected during research to be further examined. Dissemination of these data complies with the ORD Pilot and currently is included by default in the grant agreement. However, considering that not all data can be shared, the commission follows the "as open as possible, as closed as necessary" approach, thus allowing opting out either prior signing or retroactively upon legitimate reasons. Among these is potential commercial implementation of the data, security issues and protection of personal data considerations.

Similar to scientific publications, Open Access to research data is satisfied through two steps:

Step one: Deposition in designated research data online archives. (When choosing an archive, you may refer to the following listings of optional repositories: Registry of Research Data Repositories and Databib)

Step two: Promoting availability and free of charge usage of the data first by granting copyright permissions (as Creative Commons Licenses). And then, enable access to this data via your Horizon Europe project's website or any other website that you use. This data sharing activity will be part of BeXyl dissemination strategy.

In conclusion, the Gold Open Access will be considered for high impact papers or results if the BeXyl partners involved in the research consider it valuable for maximizing its impact, and budget had been allocated for this purpose in the proposal.

Funded by the European Union

3.4.3. Data utility by third parties, in particular after the end of the project

As already mentions in section 2.6., BeXyl´s data can be of interest to a broad range of experts, farmers and general public. By its nature, these data will remain of interest after the life of the project and this DMP is designed to ensure that they will be available indefinitely.

Beneficiaries should also make sure to provide repository address and basic access instructions as part of any dissemination related to the datasets.

3.4.4. Data quality assurance processes

Quality control of the data and information included in the ORD for BeXyl is based in these principles:

1- The experimental data will be subjected to the fundamental principle of research integrity as set out in the European Code of Conduct for Research Integrity. This implies compliance with reliability of research quality, reflected in the design, the methodology, the analysis and the use of resources.
2- The scientific papers generated by BeXyl will be subjected to an external peer-review as part of their editorial process. Beneficiaries have the possibility to publish at no costs in Open Research Europe, the European Commission open access publishing platform.
3- The technical materials and key summary documents (such as policy briefs) and dissemination materials of high interest will be subjected to an internal peer-review within BeXyl before been made openly available.

3.4.5. The length of time for which the data will remain re-useable

The data will remain reusable as long as the Institutional repositories will be operational, which at the time of writing this project are indefinitely.

4. ALLOCATION OF RESOURCES

According to the Guidelines on FAIR Data Management of European Commission, this section of a Data and Publication Management Plan is expected to give answers to the following issues:

- Estimate the costs for making your data FAIR. Describe how you intend to cover these costs.
- Clearly identify responsibilities for data management in your project.
- Describe costs and potential value of long-term preservation.

4.1. The costs for making data fair

The BeXyl consortium has the knowledge and tools to develop an ORD based on FAIR principles. This is one of the project objectives and appropriate resources were allocated by each partner to cover costs for data harmonisation, integration and publication. The costs for

Funded by the European Union

making the data accessible is mostly due to time required to provide the standard procedures for providing them with a common language for metadata in this DMP, the additional information for metadata and to deposit the documents or data in the repositories.

Public project deliverables and datasets are published on the BeXyl website and several institutional repositories. For this reason, there is no additional cost for publication on the website because this is included in the cost of operating the website as a whole, included the public and private parts. In addition, ZENODO and other institutional repositories are free of charge for Horizon Europe projects. No financial resources for storage, cloud, hosting, IT infrastructures etc. are then required from the project regarding the provision of data in the public domain.

Open access publications in peer-reviewed journals are part of BeXyl dissemination activities. In line with the BeXyl Communication and Dissemination Strategy, part of the project funding will be made available for gold access – particularly for any major results or breakthroughs –, which is preferable as they avoid embargo periods, enabling faster dissemination and subsequently increased visibility of the research. Along with gold access, green access will be used for disseminating projects' outcomes, which, though relevant, are not considered major results or breakthroughs.

The costs of Open access publications through Gold access could be an eligible cost in Horizon Europe and it is also possible to use Open Research Europe, the European Commission open access publishing platform. For these reasons, Open Access publications do not increase the costs for making data FAIR.

4.2. Responsibilities for data management in your project

Different consortium members will be tasked with carrying out certain aspects of the DMP. BeXyl data management responsibilities are shared between data producer (i.e. those that are producing the data), work package and tasks leaders (i.e. overseeing the project tasks) and Consortium Team (i.e. ensuring compliance with the Open Access).

The leaders of the task(s) generating each dataset have the ultimate responsibility of the correct implementation of the provisions in the DMP.

4.3. Cost and potential value of long-term preservation

Given the variety and dynamicity of the datasets generated by the BeXyl project, at present it is difficult to determine to what extent data will be accessed in the long term, and therefore calculate the value of long-term preservation. Nevertheless, the strategy pursued by this DMP, which minimizes cost of preparation and preservation of data using public, high volume and general-purpose repositories (e.g. DIGITAL.CSIC) while maximizing its openness and long-term availability, optimizes the return of the investment in scientific and technical knowledge generated by BeXyl.

Funded by the European Union

5. DATA SECURITY

The data and documents included in this DMP will benefit of the standard security for integrity (recovery) and unauthorized access provided by default by the Institutional Repositories.

With regards to privacy and sensitive data management, it is confirmed that these activities will be rigorously implemented in compliance to the privacy and data collection rules and regulations as they are applied nationally and, in the EU, as well as with the HE rules.

It should be stated that the protection of the privacy of participants is a responsibility of all persons involved in research with human participants. Privacy means, that the participant can control the access to personal information and is able to decide who has access to the collected data in the future. Due to the principle of autonomy, the participants will be asked for their agreement before private and personal information is collected. It will be ensured that all persons involved in the project activities understand and respect the requirement for confidentiality. The participants will be informed about the confidentiality policy that is used in this research project.

6. ETHICAL ASPECTS

Participation in Horizon Europe requires compliance with the highest standards of research ethics and integrity, and with EU, national and international law. Horizon Europe regulation 2021/695 (Articles 18 and 19) & Grant Agreement (Art 14, Annex 5): guiding principles and general obligations.

Regarded to ethics requirements, a specific deliverable (D11.1) will be developed in BeXyl project considering the following dimensions:

- In the specific case of Personal Data:

The social scientists will map stakeholder networks and explore the attitudes, experiences, values and behaviours of different stakeholders regarding existing/future IPM technologies, their deployment and barriers to uptake. We will gain an understanding of barriers, opportunities and costs for Xf preparedness and management through interviews, surveys and focus groups.

As is standard in social research, data will be recorded in a manner which prevents identification of the participants and, except for case studies, research sites. This involves avoiding the use of names on Dictaphone recordings and anonymity in transcriptions.

- About Data Protection:

Data collection will be done in compliance with the General Data Protection Regulation (GDPR). In addition, the collection of data will be conducted in compliance with data protection acts, legislation, and directives, both at the European and the national level. The coordinator as a data Controller, will ensure that data will be stored securely during the period of the project, and provisions will be made to ensure that the data will be secured safely beyond the lifetime of this project. Confidential information will therefore be securely stored to prevent breaches of confidentiality.

Funded by the European Union

- About Use of Animals:

BeXyl will explore the use of canine olfactory for *Xf* detection. The experiments and research programs will consider the Directives 2010/63/UE which aims to improve the welfare of animals used in scientific procedures.

The Italian National Canine Association (ENCI) will be in charge of training the dogs to recognize *Xf*-infected plants of olives and other plant species. Training activities will be initially performed in a selected location in the southern part of the Apulia (Italy) where the pathogen is endemic, and then in nurseries and in sites where plants lots for markets are located.

- About Non-EU Countries:

The consortium BeXyl has foreseen the active involvement of research Centers located in non-EU Countries affected by *Xf*, as they can represent an important added value to the research activities.

Unless it will be considered extremely necessary during the project implementation, no exchange of infectious materials (living bacterial cells in axenic cultures or *Xf*-infected plant materials, or living insect specimens) from non-EU Countries to EU, or viceversa, has been foreseen in BeXyl. As alternative, the exchange of non-viable (non-infectious) DNA-preparations and insect vectors stored in alcohol will be done, as such these samples (not containing living bacterial cells) are not covered by the EU regulations on the possession and handling quarantine plant pathogens (Regulation (EU) 2016/2031 and Commission Delegated Regulation (EU) 2019/829).

- About Plant Health Rules:

All the activities foreseen in the project with viable stages of *Xf* will follow the provisions of Regulation (EU) 2016/2031 on protective measures against plant pests. Particularly, those specified in Article 8 on Union quarantine pests used for official testing, scientific or educational purposes, trials, varietal selections or breeding. An official authorisation will be mandatory in each case where viable stages of *Xf* will be used, to ensure that the activities will be properly recorded and carried out in a quarantine station or a confinement facility designated by the competent authority. Likewise, activities will be conducted only by authorized personnel with appropriate scientific and technical competence. Furthermore, specific measures against *Xf* laid down by Commission Implementing Regulation (EU) 2020/1201 will be observed in demarcated (outbreak) areas.

Funded by the European Union

7. REFERENCES

Horizon Europe Annotated Model Grant Agreement (AGA). Annex 5 HE IPR RULES. Available in: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga_en.pdf

Guidelines on FAIR Data Management in Horizon 2020. Available in: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Eglen & Nüst. 2019. CODECHECK. Septentrio Conference Series (1). https://doi.org/10.7557/5.4910.

Caracciolo, C., Stellato, A., Morshed, A., Johannsen, G., Rajbhandari, S., Jaques, Y., Keizer, J. 2013. The AGROVOC linked dataset. Semant. Web J. 4(3), 341–348 (2013)

Registry of Research Data Repositories: https://www.re3data.org/

Horizon Europe regulation 2021/695 (Articles 18 and 19). Available in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R0695

Model Grant Agreement (MGA). Available in: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf

Funded by the European Union