Project Number: 800926

Project Acronym: HyPhOE

Project title: Hybrid Electronics based on Photosynthetic Organisms

DATA MANAGEMENT PLAN

Deliverable number:	D7.2
Delivery date:	28 February 2019 (M6)
WP number:	WP7
Author(s):	Sara Strömberg, Eleni Stavrinidou, Gianluca Farinola and Massimo Trotta
Lead Beneficiary:	UNIBA
Dissemination level:	Public

1. Project Participants

Full name	Short name	Contacts
LINKOPINGS UNIVERSITET	LiU	Eleni Stavrinidou (Coordinator), eleni.stavrinidou@liu.se
SVERIGES LANTBRUKSUNIVERSITET	SLU	Torgny Näsholm, torgny.naslund@slu.se Totte Nittylä, totte.niittyla@slu.se
INSTITUT POLYTECHNIQUE DE BORDEAUX	Bordeaux INP	Eleni Pavlopoulou, eleni.pavlopoulou@enscbp.fr
UNIVERSITA DEGLI STUDI DI BARI ALDO MORO	UNIBA	Gianluca Farinola, gianlucamaria.farinola@uniba.it
UNIVERSITE PARIS DIDEROT - PARIS 7	UPDiderot	Benoît Piro, piro@univ-paris-diderot.fr

2. Data Summary

This document constitutes version 1.0 of the Data Management Plan (DMP) for the HyPhOE project, funded by the EU's Horizon 2020 Programme under Grant Agreement number 800926. The DMP will present the procedure for the management of datasets created during the lifetime of the project and describes the key data management principles, in terms of data standards and metadata, sharing, archiving and preservation.

The HyPhOE project aims to develop a set of tools and methods for bi-directional electronic and chemical interfacing with photosynthetic organisms, which will pave the way for the following targeted applications: i. Energy systems based on electronically-functionalized plants and photosynthetic organisms. ii. Plant physiological control using bioelectronics systems. iii. Environmental monitoring using functionalized plants. Since the research of HyPhOE is highly interdisciplinary, including electronics, materials science, chemistry, plant biology and biotechnology, making research data from HyPhOE openly available, implies a great value to several research fields. The project expects to generate and collect only new data, constituting the following datasets, with an expected file size between 1 MB and 4 GB:

Table 1. Potential datasets

Data type	Format	Size
Experimental data with description (metadata)	.txt, ascii, .csv .tiff, .jpg, .tif, .eps, .bmp, .ai, .opj, .m, .mat, .xls, .exe, .mpr, .ocw, .fra .ND2	1 GB
Protocols	.doc, .pdf, .ppt	< 1 GB
Lab notes	.doc, .pdf, .ppt	< 1 GB
Reports	.doc, .pdf, .ppt	< 1 GB
Presentations	.ppt	< 1 GB
Crystallographic data	.cbf, .img	< 4 GB

3. FAIR data

3. 1. Making data findable, including provisions for metadata

Datasets and corresponding metadata will be identifiable using unique identifiers, DOIs (Digital Object Identifier) and a persistent link. No standards for metadata exist in the research discipline of HyPhOE, but a minimum content of the metadata collected will be documented and stored according to the Dublin Core Metadata Element Set (DCMES) with the following additions:

- Title of project
- Name of lead beneficiary responsible for the creation of the data set
- Year of publication of dataset
- Version of dataset used for the analysis
- Instrument name used for data acquisition (e.g. biologic potentiostat)
- Type of characterization (e.g. cyclic voltammetry)
- Name of repository/data archive
- Permanent identifier, DOI

Metadata content and format will be further developed in future versions of the DMP, and an example of a metadata file will be produced and amended to this plan, after the first scientific publication has been made available.

3.2. Making data openly accessible

According to Article 29.3 of the HyPhOE Grant Agreement, beneficiaries of HyPhOE must deposit research data including associated metadata, needed to validate the results presented in scientific publications as soon as possible unless a decision has been taken to protect the results.

At Linköping University, research data will be stored in the institutional repository DiVA (Digitala Vetenskapliga Arkivet) via the *Dataset* publication type. Each file may be up to 16GB and you can upload multiple files, even in compressed formats. Each dataset receives a unique identifier (Digital Object Identifier) and a persistent link. From the dataset it is possible to link to the publication. In addition, a document containing a description of research data e.g. an overview of the research results, various variables that are included in the data material and what they mean, how the material is to be interpreted, will be uploaded together with the dataset.

At SLU research data will be stored in the university's repository TILDA via the *Dataset* publication type. Each dataset receives a unique identifier (Digital Object Identifier) and a persistent link. From the dataset it is possible to link to the publication. In addition, a document containing a description of research data e.g. an overview of the research results, various variables that are included in the data material and what they mean, how the material is to be interpreted, will be uploaded together with the dataset.

Bordeaux INP will use the open data repository HAL (*Hyper Articles en Ligne*, https://hal.archives-ouvertes.fr/). HAL is an open archive implemented by the *Centre pour la Communication Scientifique Directe* (CSDD) of the *French National Centre for Scientific Research* (CNRS). It is a platform made in the respect of open access principles, for archiving and disseminating scientific publications and data. HAL is compatible with the European OpenAIRE project. In case of uploading a publication, HAL can automatically retrieve the metadata to complete the filing, either directly from the pdf file (Grobid – GeneRation Of BIbliographic Data), or, from the DOI number, using the CrossRef database. In case of uploading unpublished data, each dataset will prior receive a unique identifier (Digital Object Identifier number) and then will be uploaded to HAL. In addition, a document containing a description of research data e.g. an overview of the research results, various variables that are included in the data material and what they mean, how the material is to be interpreted, will be uploaded together with the dataset.

UNIBA will use the data repository IRIS – Institutional Research Information System (https://ricerca.uniba.it/). IRIS is password protected archive designed for the collecting and handling research products, including publications. Similarly CNR uses the data repository PEOPLE (https://www.cnr.it/people/) which is also a password protected archive designed for the collecting and handling research products, including publications. Data deposited in these databank can be retrieved via specific fields, including Author, DOI and funding grant.

UPDiderot will use the French HAL open data repository (https://doc.archives-ouvertes.fr/), also used by Bordeaux INP. The multidisciplinary open archive HAL exists since 2000. HAL is based on the principle of self-archiving, the direct deposit with the authors of the full text of publications of any kind, but also thesis, research report, data with metadata, etc.

3.3. Making data interoperable

The beneficiaries of HyPhOE project aims to collect and document data in a standardised way to enable exchange and re-use between researchers of HyPhOE. Even though the HyPhOE consortium is highly interdisciplinary, standard protocols and methods will be used to generate research data experimentally.

Thus, the formats for research data will be similar across the consortium. Generated data will be preserved on institutional platforms minimum until the end of the HyPhOE project and maximum according to national guidelines and legislation for archiving.

3.4. Increase data re-use (through clarifying licences)

To enable reuse of the HyPhOE datasets and make the data available to the widest audience possible, licensing will be made through Creative Commons. At this stage of the project, all 7 Creative Commons licenses can be possible and considerations for choosing license will be made continuously throughout the project.

4. Allocation of resources

The costs foreseen to make HyPhOE datasets openly available are primarily personnel costs related to HyPhOE beneficiaries managing and storing datasets according to this plan.

Prof. Gianluca Farinola and Prof. Massimo Trotta at Universita Degli Studi di Bari Aldo Moro (UNIBA) are responsible for data management within HyPhOE. The implementation of the data management plan is the responsibility of the PIs of the HyPhOE beneficiaries.

5. Data security

No sensitive data, such as personal data will be generated within HyPhOE.

During the course of HyPhOE, datasets will be stored locally by the responsible beneficiary, detailed in the table below.

Table 2. Data Storage

Short name	Data Storage
LiU	Data generated at LiU will be stored at internal unit servers at Laboratory of Organic Electronics, LiU. All data is backed up automatically by the central IT department at LiU (network file storage at LiU). Backups are regularly taken, which means that if data is lost for any reason, it can in most cases be recreated.
SLU	Data generated at SLU will be stored at internal unit servers at Umea Plant Science Centre. Backups are regularly taken, which means that if data is lost for any reason, it can in most cases be recreated.
Bordeaux INP	Data generated at Bordeaux INP will be stored at internal unit servers at the Laboratoire de Chimie des Polymères Organiques (LCPO). A network-attached storage (NAS) server is used for this purpose, which is a file-level computer data storage server. The NAS installed at the LCPO contains four storage drives and is connected to

	the internal computer network of the LCPO. All data is backed up automatically by the NAS, ensuring that if data is lost for any reason, it can in most cases be recreated.
UNIBA	Data generated at UNIBA/CNR will be stored at internal unit servers at Chemistry Department - Uniba. All data are backed up automatically. Backups are regularly taken, which means that if data is lost for any reason, it can in most cases be retrieved.
UPDiderot	Data generated at UPDiderot will be stored at internal computers and at an internal unit server at the ITODYS laboratory. All data are backed up automatically each day by the server, which means that if data are lost by the users at their working computer, they can be restored within a short time. This server cannot be accessed from outside the laboratory network.

6. Ethical aspects

No ethical or legal issues can be foreseen within the scope of the proposed project. All partners will comply with Article 34 of the Grant Agreement and the HyPhOE consortium will use techniques and methodologies (including for data collection and management) that are appropriate for the field.

7. Other issues

None of the HyPhOE beneficiaries currently have any official policy for managing and storing research data.

8. Timetable for updates of the DMP

Updates of the HyPhOE DMP will be made over the course of the project in conjunction with significant changes such as generation of new data (not previously covered in the DMP), changes in consortium policies (such as new IPR strategies) or changes in the consortium composition. There will also be updates in conjunction with the periodic reporting