NEUROFIBRES

D 8.2 Data Management Plan

H2020 – FET PROACTIVE – 01 - 2016
Project Nº 732344

Project Title: **Biofunctionalised Electroconducting Microfibres for the Treatment of Spinal Cord Injury**

WP Leader: **SESCAM**

Due date of deliverable: **30/06/2016 (M6)**

Actual submission date: **Month 10**

Neurofibres consortium

- Servicio de Salud de Castilla La Mancha (SESCAM)

- The Chancellor, Masters and Scholars of the University of Cambridge (UCAM)

- Axon Cable (AXON' CABLE)

- Universita Degli Studi di Trento (UNITN)

- Kungliga Tekniska Hoegskolan (KTH)

- Universite D'aix Marseille (AMU)

- Universitat des Saarlandes (USAAR)

For more information on this document or NEUROFIBRES, please contact:

Dr. Jorge E. Collazos Castro – NEUROFIBRES Coordinator
jcollazos@sescam.org

Servicio de Salud de Castilla La Mancha
Hospital Nacional de Parapléjicos
Laboratory of Neural Repair and Biomaterials
Finca la Peraleda s/n
45071 – Toledo (Spain)

Dr. Beatriz Gallego Gómez – NEUROFIBRES Project Manager
bgallegog@externas.sescam.jccm.es

Fundación Hospital Nacional de Parapléjicos
Finca la Peraleda s/n
45071 – Toledo (Spain)
+34 925 396 812

Document control

Title	D 8.2Data Management Plan
Contributing work package	WP 8: Activities for the exploitation, dissemination and communication of results
Actual delivery date	19/10/2017
Dissemination level	PUBLIC
Contributors	All partners

Revision history

Version	Date	Authors	Description / Comments
V.0.0	25/09/2017	Dr. Beatriz Gallego (FHNP)	Creation
V.01	06/10/2017	Dr. Jorge Collazos (SESCAM), Dr. Beatriz Gallego (FHNP)	Revision
V.02	19/10/2017	All partners	Final version

Executive summary

1. Description of the deliverable content, objectives and purpose

The objective of WP 8 *Activities for the exploitation, dissemination and communication of results* is to ensure the appropriate measures to exploit, disseminate and communicate the main results derived from the project.

The deliverable 8.2 entitled as *Data Management Plan* (DMP) is a document which provides an analysis of the main elements of the data management policy that will be used by the members of the consortium with regard to the data generated through the life of the project.

The DMP will be released in compliance with the Horizon 2020 FAIR DMP template[1], provided by the European Commission in the Participant Portal, and updated over the course of the project in time with the periodic evaluation of the project (M12, M30 and M48).

---

[1]http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Acronyms

DMP: Data Management Plan

EC: European Commission

EU: European Union

FAIR: Findable, accessible, interoperable and re-usable

H2020: Horizon 2020

MFs: Microfibres

WP: Working Package

PU: Public

RE: Restricted

CO: Confidential

Table of contents

Deliverable report

Introduction

NEUROFIBRES is a Horizon 2020 project participating in the Open Research Data Pilot. This pilot is part of the Open Access to Scientific Publications and Research Data Programme in H2020[2]. The goal of the program is to foster access to data generated in H2020 projects.

Open Access refers to a practice of giving online access to all scholarly disciplines information that is free of charge to the end-user. In this way data becomes re-usable and the benefit of public investment in the research will be improved.

The EC provided a document with guidelines[3] for projects participants in the pilot. The guidelines address aspects like research data quality, sharing and security. According to the guidelines, projects participating will need to develop a Data Management Plan (DMP).

The purpose of the DMP is to provide an overview of the main elements of the data management policy that will be used by the Consortium with regard to the project research data. The DMP is not a fixed document but will evolve during the lifespan of the project.

The DMP covers the complete research data life cycle of the NEUROFIBRES project. It describes the types of research data that will be generated during the project, the strategies on research data preservation and the provision on access rights. The research data should be "FAIR", that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation solution.

The repository ZENODO has been chosen as the main repository to store, classify and provide Open Access to the stored data objects originated within the NEUROFIBRES project frame. Nevertheless, institutional data bases will be also considered to provide open access to specific data.

1. Data Summary

1.1. Purpose of data collection and its relation with the project objectives

The purpose for data collection is to capture qualitative and quantitative evidence for their analysis to lead to the formulation of answers to the questions that concern the NEUROFIBRES project. Those data could be useful to scientists and industry with research and commercial interests in the fields of SCI, neuroprostheses, bio-electronic systems, electrical and mechanical engineering, biomaterials, affibodies, biotechnology, and electroconducting materials.

---

[2]http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

[3]http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

1.2. Types and formats of data

NEUROFIBRES will generate and collect multimodal measurements of the properties and behaviour of the MFs, the interconnected system and their mechanical interaction with the host tissue, numerical data on the features of the affibodies accounting for their electro-responsiveness and affinity for the target molecules, numerical and qualitative data regarding cell culture and animal experiments in which the novel bio-electronic tools will be tested, data statistics and numerous images related with the experiments.

The types and formats of data acquired within the NEUROFIBRES project frame include the following:

- Laboratory data: datasets (*.txt, *.doc, *.docx,*.xls, *.xlsx), multimodal measurements (*.txt, *.doc, *.docx,*.xls, *.xlsx), numerical data (*.XX), qualitative data (*.txt, *.doc, *.docx), data statistics (*.xls, *.xlsx), images (*.jpg, *.png, *.jpeg, *.tiff), videos (*.avi, *.mov, animated GIF).
- Research data: statistics (*.xls, *.xlsx), graphs (*.ogg, *xls, *.xlsx), bibliography (*.enl).
- Scientific texts: manuscripts and reports (*.doc, *.docx, *.pdf), publications (*.doc, *.docx, *.pdf), conference proceedings (*.doc, *.docx, *.pdf), conference presentations and posters (*.ppt, *.pptx, *.pdf), books and theses (*.doc, *.docx, *.pdf).
- Dissemination material: leaflets and fact-sheets (*.pdf), images (*.jpg, *.png, *.jpeg, *.tiff), animated images (*.gif), videos (*mp4), social network publications and website (*.html), presentations and templates (*.ppt, *.pptx, *.pdf).
- Management documents: deliverables (*.doc, *.docx, *.pdf), patents (*.doc, *.docx, *.pdf).

1.3. Re-using of data

Characterization of inflammation will be grounded on previous published and non published work aiming at the immunophenotypic identification of all fluorescent cell types in transgenic Thy1CFP//LysM-EGFP//Cd11C-EYFP mice as well as on methodological works providing the suitable parameters for intravital monitoring of cellular responses during the course of CNS pathology.

1.4. Origin of data

- **Laboratory and experimental data:** keysets of data obtained directly from the research setups. The sources will be identified in the deposited datasets.

- **Research data**: analyzed data from the laboratory which additional information such as charts, diagrams, statistics…etc.

- **Scientific texts** and **dissemination material**: prepared from the research data. Author-accepted manuscripts, media (images, video, gifs), leaflets etc. are included in this category.

- **Patents:** a potential product of the NEUROFIBRES research data.

- **Deliverables:** a set of reports for European Commission based on the research data derived from the NEUROFIBRES project.

1.5. Expected size of the data

The size of the data will be attempted to be as low as possible with the purpose of facilitating its storage and exchange. In any case, a single file should not exceed the limit of 50 GB, set by the repository ZENODO (see section 2.2.). The expected size of the data is listed below:

- Laboratory data: datasets (< 5MB each), images (1 – 10MB each), 5D images stacks and videos (up to 70GB per file).
- Research data: statistics, graphs, bibliography (< 1MB each).
- Scientific texts: manuscripts and reports (1 – 5MB), publications (100KB – 2MB), conference proceedings (100KB – 2MB) ,conference presentations and posters (100KB – 2MB), books and theses (30 – 50MB).
- Dissemination material: leaflets and fact-sheets (2MB), images (1 – 10MB each), animated images (300KB), videos (10MB), presentations and templates (100KB – 5 MB).
- Management documents: deliverables (500KB – 2MB), patents (500KB – 2MB).

1.6. Data utility

Making typical sample data openly accessible will be useful for validation of the research results, increasing citations and even facilitating other scientific breakthroughs. Original scientific data from all disciplines involved in the project are susceptible of being exploited by researchers or industry external to the project. On the other hand, there is great inconsistency in the results of experimental treatments for spinal cord injury (SCI), and much disparity exists in the methodology used for verification of the therapeutic effects. Searching for standardisation, a consensus on the minimum content of papers published on this matter was reached[4]. Our consortium will produce much more complete data and metadata than the proposed SCI standards, making an additional contribution to the field.

2. FAIR Data

In compliance with the European Commission guidelines, the data generated by the NEUROFIBRES project must be FAIR, that is findable, accessible, interoperable and re-usable. The decision to be taken by the project on how to publish its documents and data sets will come after the more general decision on whether to go for an academic publication directly or to seek first protection by registering the developed Intellectual Property.

---

[4]J Neurotrauma 2014;31:1354

2.1. Making data findable

All the records deposited in the ZENODO repository are indexed immediately in OpenAIRE, which is the Open Access Infrastructure for Research in Europe. OpenAIRE does this by aggregating European funded research output from nearly 1000 repositories from all over the world and makes them available via the OpenAIRE portal. Records indexed in OpenAIRE will be immediately available in the European Commission Participant Portal.

Metadata provision

The data generated under the NEUROFIBRES frame will be discoverable, identifiable and locatable by means of suitable metadata. Descriptive metadata refers to the information about the objects content. ZENODO repository offers the possibility to assign several tags (metadata) to all uploads, in order to make the content findable. The tags ZENODO offers are:

- Publication type (journal article, presentation, book, thesis…etc.).
- Title, authors, affiliation.
- Description of the content.
- Communities that the data belong to.
- Grants which have funded the research.
- Identifiers (DOI, ISSN, PubMed ID, URLs…etc.).
- Contributors.
- References.

ZENODO assigns other characteristics to uploads, in order to make the content interoperable, such as:

- Journal name, volume, issue and pages; in the case of a manuscript.
- Conference title, place, session…etc.; in the case of a conference proceeding.
- Publisher, place, ISBN, pages; in the case of parts of books and reports.

Clear versioning

The version of the data which is aimed to be deposited will be the final version. The digital object identifiers (DOIs) are automatically generated upon deposition on the repository. If necessary, posterior versions will be deposited; these posterior versions will be identified by their own DOI and identifiable by the date of deposition and file name.

Including semantic information such as the version number in a DOI will not be encouraged, because this information may change over time, while DOIs must remain persistent. Most importantly, version suffixes are not machine readable. ZENODO DOI versioning is linear, which means that the ZENODO version number may in fact not be the real version number of the resource.

The approach of NEUROFIBRES to the versioning is the one ZENODO provides: two versions (two DOIs) are semantically linked in the metadata of a DOI. This ensures that discovery systems have a machine readable way to discover that two DOIs are versions of the same resource.

Standard identification mechanism

All deposited data will be uniquely identifiable through the standard identifier DOI. Additionally, other identifiers, such as Handle, ARK, PURL, ISSN, ISBN, PubMed ID, ORCID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLsmay be used.

Data generated under the NEUROFIBRES frame will acknowledge the grant in the following way:

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732344".

and automatically be associated to the project via OpenAIRE portal.

Naming conventions

ZENODO DOI versioning is linear. Currently DOIs registered by Zenodo follows the pattern "10.5281/zenodo." where 10.5281 is the Zenodo DOI prefix and is a sequentially assigned integer. The word "zenodo" is semantic information, and as mentioned in the previous section, to include semantic information in DOIs it is not supported by NEUROFIBRES as it may change over time. The current practice was introduced when Zenodo was launched, and while it is not ideal ZENODO did not want to change the existing practice.

Keywords

All deposited data will have an associated group of keywords to facilitate identification.

2.2. Making data openly accessible

The main results of NEUROFIBRES are expected to be exploited industrially, and therefore some data cannot be made available for verification and re-use by persons and/or organisations external to the consortium. Making openly accessible some details on the fabrication or composition of the MFs, affibodies and interconnection system, or electrical parameters critical for succeeding in this challenging application, might jeopardise the achievement of the industrial and commercial impact of the project and therefore will be confidential information. Despite this, the consortium has a proactive, free-online access publication policy, provided that patents are applied before submitting the results to scientific journals. On the other hand, we aware that making data openly accessible will be useful for validation of the research results, increasing citations and even facilitating other scientific breakthroughs.

Which data will be made openly accessible

Any dissemination data linked to exploitable results will not be put into the open domain if they compromise its commercialization prospects or have inadequate protection.

Categories of outputs that NEUROFIBRES will give Open Access (free of charge) include:

- Scientific publications (author accepted manuscripts, supplementary files, and conference proceedings).

- Research Data (key datasets accompanying publications that are needed to validate the results).

- Other research data that may be of interest to scientists and/or industry

- Deliverables (public).

We will provide restricted access to the members of the Consortium only for templates (e.g. deliverables templates) and documents concerning internal meetings (e.g. minutes of meeting).

Dissemination and outreach material will be openly available via the NEUROFIBRES website and the network social sites.

How the data will be made available

The data will be available via ZENODO repository. The data which is owned by the Consortium will be deposited as soon as possible, in the repository with open access rights. In the case the data needs to be protected (for example, in the case of a future publication or patent), the data will be deposited with embargoed access. The embargo will be lifted after the data has been disclosed.

The author accepted manuscripts will be deposited as soon as possible, and at the latest after 6 months after publication. The **preferred** route to deposit the manuscripts will be the **'Green' model**: the author or representative deposits the published article or final peer-reviewed manuscript in an online repository. Whenever the 'Green' route cannot be assured, the Consortium will provide open access following the 'Gold' route: an article is immediately released in Open Access mode by the scientific publisher upon publication; the payment of publication costs is shifted away from subscribing readers.

All data deposited by the NEUROFIBRES project in the ZENODO repository will be available via the OpenAIRE portal:

https://www.openaire.eu/search/project?projectId=corda__h2020::719cd36fe3741c6 0b7bdc234b8867fe9

Methods or software tools needed to access the data

The software tools necessary to access the data will be standard software tools and free of charge, such as Open Office, or Adobe Acrobat Reader. Other standard programs could be also used, such as Microsoft Office package. These well-used standard software are likely to be readable in the future.

In the case of research data files that cannot be opened with standard programs, reliable information on the instruments and tools needed to validate the results will accompany the data setsand will be indicated in the section "Additional notes", upon deposition in the ZENODO repository.

Where the data documentation and code are deposited

The consortium has chosen ZENODO as the central scientific publication and data repository for the project outcomes. The online repository has been created through the European Commission's OpenAIREplus project and is hosted at CERN.

The ZENODO community NEUROFIBRES has been specifically created to gather all data under the frame of the NEUROFIBRES project:

https://zenodo.org/communities/neurofibres/?page=1&size=20

Additionally, the public data will be included in the European Commission Funded Research (OpenAIRE) ZENODO community (https://zenodo.org/communities/ecfunded/), which is curated by ZENODO.

Restrictions

The deposited data is accessible under four types of access **rights**:

- Open Access.

- Embargoed Access.

- Restricted Access.

- Closed Access.

The data which is owned by the Consortium will be deposited as soon as possible, in the repository with **open access** rights. In case there is any **embargo** period on the deposited data, the access to the data will be granted after the embargo period and at the latest after 6 months after publication. In the case of author-accepted manuscripts with an embargo period longer than 6 months, the Gold Open Access route will be followed.

However, in some cases, data will be **restricted** due to scientific purposes, i.e. validation of the results, statistical analyses on research results, background information for new projects, etc. This also applies to some documents such as templates (e.g. deliverables templates) and documents concerning internal meetings (e.g. minutes of meetings), which are not public.

2.3. Making data interoperable

The following abbreviations will be used in the data:

aa: Amino acid

Au: Gold

BBB: Blood Brain Barrier

CARS: Coherent Anti Stokes Raman Spectroscopy

CD: Circular dichroism

CFP: Cyan Fluorescent Protein

CNS: Central Nervous System

7

DRG neuron: Dorsal Root Ganglion Neuron

EGFP: Enhanced Green Fluorescent Protein

EYFP: Enhanced Yellow Fluorescent Protein

FEM: Finite Element Method

HFBM: Hierarchical Fibre Bundle Model

HPLC: High performance liquid chromatography

K_D: Equilibrium dissociation constant

MALDI-TOF: Matrix-assisted laser desorption ionization time of flight

MF: Microfiber

MS: Mass spectrometry

Myeline: insultating and nutritive layer around axons

OPO: Optical Parametric Oscillator

Pt: Platinum

QD655: Quantum Dots emitting at 655 nm

QFM: Quantized Fracture Mechanics

RU: Response units

SCI: Spinal Cord Injury

Secondary neurodegeneration: delayed loss of neurons with regard to impact

SPPS: Solid phase peptide synthesis

SPR: Surface plasmon resonance

UDHS: Unilateral Dorsal Hemi Section

Video: Time lapse acquisition

2P: Two-Photon Microscopy

5D images: Multiparametric images (xyz, color, time):

2.4. Data re-use

Data re-use is subject to the **license** under which the data objects are deposited.

Licence to the data

The ZENODO repository offers five types of licences for the files under Open Access right. The license chosen by NEUROFIBRES will be a Creative Commons Attribution Non-Commercial 4.0. The characteristics of this license are:

- The licenser allows to copy, distribute and communicate publicly the work, as well as create and disseminate works derived from the former.

- The licenser allows the reproduction, dissemination and public communication of the work only for non-commercial purposes.

By uploading content, no change of ownership is implied and no property rights are transferred to CERN. All uploaded content remains the property of the parties prior to submission.

When the data will be made available for re-use

The data will be immediately available for reuse upon deposition. As we have stated in Section2.2, the data deposited will be conferred Open Access rights on deposition and at the latest after 6 months.

Third parties and re-usability

The data produced may be used by third parties, since will be openly available during the lifetime of the ZENODO repository.

Data quality assurance processes

The cornerstone of digital preservation is data integrity: data is complete and unaltered as it was when it was originally recorded. All data files are stored in ZENODO along with a MD5 checksum of the file content. Files are regularly checked against their checksums to assure that file content remains constant.

Length of time for which the data will remain re-useable

All data files deposited in ZENODO are stored in CERN Data Centres, primarily Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis. Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least. In case of closure of ZENODO repository, the best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.

The NEUROFIBRES data is expected to be available after the end of the project duration (31.12.2020), for at least, the lifetime of the ZENODO repository.

Records can be retracted from public view; however, the data files and records are preserved and can be no longer changed or retrieved. The record's metadata can be modified in ZENODO repository.

3. Allocation of resources

Estimated costs for making the data FAIR

The estimated cost of the article processing charges is in average 3.500 € per publication[5]. Considering an estimate of 3 publications per year per partner from the second to the fourth year of the project, the total cost of making the data openly accessible for the NEUROFIBRES project is about 200.000 €.

The associated costs are covered by the author and/or co-authors of the publication as agreed in the NEUROFIBRES grant agreement (eligible costs in Horizon 2020 projects).

Responsibilities for data management

Any member of the Consortium can upload content in the repository. The content will be approved by the coordinator of NEUROFIBRES, SESCAM. All approved items cannot be deleted. New versions of the content can be uploaded together with previous versions; all versions are simultaneously available.

Value of long term preservation

The value of long-term preservation is on ensuring and facilitating the accessibility and usability of the preserved data. It involves planning, resource allocation and application of preservation methods that have been described in Section 2. The goal is the accurate reordering of authenticated content over time, so it remains usable as technological advances render original software obsolete.

4. Data security

In this section, data recovery, secure storage and transfer of sensitive data is addressed.

All data files are stored in CERN data centres. CERN has considerable knowledge and experience in building and operating large scale digital repositories and a commitment to maintain this data centre to collect and store 100s of PBs of LHC data as it grows over the next 20 years. In the highly unlikely event that ZENODO will have to close operations, CERN guarantees the migration of all content to other suitable repositories, and since all uploads have DOIs, all citations and links to ZENODO resources will not be affected.

---

[5] R. van Noorden, Open access: The true cost of science publishing. *Nature* **495**, 426 (2013).