Horizon 2020

*flora***robotica**

Societies of Symbiotic Robot-Plant Bio-Hybrids as Social Architectural Artifacts

Deliverable D4.1

Data management plan (open research data pilot)

Date of preparation: 2015/09/30	**Revision: 1 (r81)**
Start date of project: 2015/04/01	**Duration: 48 months**
Project coordinator: UPB	**Classification: public**
Partners: *lead:* **UPB**	*contribution:* —
Project website:	http://florarobotica.eu/

H2020-FETPROACT-2014

DELIVERABLE SUMMARY SHEET

Grant agreement number:	640959
Project acronym:	*flora robotica*
Title:	Societies of Symbiotic Robot-Plant Bio-Hybrids as Social Architectural Artifacts

Deliverable N^o:	Deliverable D4.1
Due date:	M6
Delivery date:	2015/09/30

Name:	**Data management plan (open research data pilot)**
Description:	We report on identified options of data repositories, our data management approach, and individual data set types.

Partners owning:	**UPB**
Partners contributed:	—
Made available to:	public

Contents

1 Introduction

We have done a survey of available data repositories but we could not identify an appropriate repository that would fit the project's focus. This might be due to our nontraditional, innovative, and interdisciplinary approach of combining plant science, computer science, engineering, and architecture. For example, there are multiple data repositories that focus on plant science which are, however, very specific for genetic data, geographical data, etc. Therefore, we have decided to use a multipurpose data repository. We have chosen Zenodo[1] due to its assumed high reliability now and in the future.

In addition, we expect to generate big amounts of data. The overhead of storing and handling all of that data seems not justifiably. Hence, we have decided for a hybrid approach of selecting data that we store only for the lifetime of the project on our project's Cloud service hosted by OwnCube e.U.[2] and for which we share data based on individual requests. A more selective choice of data will be uploaded to Zenodo allowing for preservation beyond the project's lifetime.

Finally, we plan a strategy of storing almost all data by one service (Zenodo) for simplicity and as a holistic approach. At the same time, we allow to continue ongoing data sharing procedures of members of our project consortium, especially, concerning self-archiving practices of publications (conference papers and journal papers) and open-source availability of developed software via software repositories such as GitHub[3] and sourceforge[4].

---

[1]see http://www.zenodo.org, Zenodo is developed by CERN under the EU FP7 project OpenAIREplus (grant agreement no. 283595).

[2]OwnCube e.U., An der Liesing 2-34/7, 1230 Wien, Austria

[3]https://github.com/

[4]http://sourceforge.net/

2 Specifications for each data set

2.1 DATA1 – publications

2.1.1. Data set reference and name

DATA1 – publications (conference papers, journal papers, monographs)

2.1.2. Data set description

Scientific and peer-reviewed papers that are published by members of the project consortium. Groups with interest in this type of data set are the readerships of the respective conferences and journals.

2.1.3. Standards and metadata

We will use Zenodo[5]. Hence, Zenodo also defines and handles the metadata. Citing from their policies[6]: "Metadata is licensed under CC0, except for email addresses. All metadata is exported via OAI-PMH and can be harvested. [...] All metadata is stored internally in MARC [...]. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema according to OpenAIRE Guidelines." Version control is not necessary. Furthermore, we keep for each publication in this data set a BibTeX entry and its DOI[7].

2.1.4. Data sharing

The publications will be uploaded as PDF (Portable Document Format) on the servers of Zenodo. In addition, an incomplete selection of papers will be uploaded to http://arXiv.org and members of the project consortium will also continue their self-archiving practices.

2.1.5. Archiving and preservation (including storage and backup)

The data will be curated by Zenodo who state about their retention period: "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 the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories."[8]

---

[5]see http://www.zenodo.org, Zenodo is developed by CERN under the EU FP7 project OpenAIREplus (grant agreement no. 283595).
[6]https://zenodo.org/policies
[7]Digital object identifier, see http://www.doi.org/
[8]https://zenodo.org/policies

2.2 DATA2 – photo & video material of experiments

2.2.1. Data set reference and name

DATA2 – photo & video material of plant & robot experiments

2.2.2. Data set description

Generated data, photos and videos taken during experiments with natural plants and robots. Data is of interest for plant scientists and roboticists. It partially underpins scientific publications. The data is probably too specific to allow for integration.

2.2.3. Standards and metadata

Data is labeled with experiment ID. Photos are numbered in chronological order or alternatively labeled directly with the time in the experiment (e.g., minutes after start of experiment). Videos are labeled with the experiment ID and numbered chronologically if there are multiple parts. Version control is not necessary. For each experiment dataset there is a short textual description of the experiment itself, when it was started, when it was ended, and who supervised it. For the data that is uploaded to Zenodo, Zenodo defines and handles the metadata (see above).

2.2.4. Data sharing

Given the amount of this kind of data that will be generated throughout the project, we do not plan a central data deposit for all of it. Individual photos and videos that are representable for the respective experiment and/or relevant for publications will be stored in the project's Cloud service (OwnCube e.U., An der Liesing 2-34/7, 1230 Wien, Austria) for the duration of the project, shared among project partners, and data requests for that data from outside the project will be handled individually. A smaller selection of that data will be uploaded to the servers of Zenodo and hence allow for easy sharing.

2.2.5. Archiving and preservation (including storage and backup)

For the data stored on our cloud service we cannot guarantee preservation beyond the project's lifetime but that is taken care of by the smaller selection of data that will be uploaded to Zenodo (lifetime of the host laboratory CERN is 20+ years, see above).

2.3 DATA3 – developed software

2.3.1. Data set reference and name

DATA3 – developed software

2.3.2. Data set description

A variety of different software packages are going to be developed in this project including software to run microprocessors (e.g., Raspberry Pi[9], sensors, actuators, to run complete robot platforms, and also simulation software frameworks. The developed software is mostly of interest for roboticists while some of the simulation software might also be of interest for plant scientists. The software partially underpins scientific publications and is probably too specific to allow for integration.

2.3.3. Standards and metadata

For service such as GitHub and sourceforge the metadata is very specific and depends on the respective hosting service. For the data that is uploaded to Zenodo, Zenodo defines and handles the metadata (see above).

2.3.4. Data sharing

For software we plan a hybrid approach. For software that is under development (i.e., within the project's lifetime) we are going to use a project-internal software versioning and revision control tool based on Apache Subversion (SVN) [10] which is hosted by the "Zentrum für Informations- und Medientechnologien (IMT)" of the University of Paderborn (UPB, project coordinator). For software releases we are going to use standard and well accepted hosting services for scientific open-source software such as GitHub[11] and sourceforge[12] but also to provide the software via Zenodo[13].

2.3.5. Archiving and preservation (including storage and backup)

For the software stored on our project-internal software versioning system we cannot guarantee preservation beyond the project's lifetime but that is taken care of by the released software packages that will be uploaded to Zenodo (lifetime of the host laboratory CERN is 20+ years, see above).

---

[9]https://www.raspberrypi.org
[10]https://subversion.apache.org/
[11]https://github.com/
[12]http://sourceforge.net/
[13]see http://www.zenodo.org

2.4 DATA4 – sensor data of experiments

2.4.1. Data set reference and name

DATA4 – sensor data of plant & robot experiments

2.4.2. Data set description

Generated sensor data acquired during experiments with natural plants and robots. This includes data from sensors monitoring the plants (temperature, humidity, etc.) and from sensors of the robots (proximity sensors, force sensors, etc.). Data is of interest for plant scientists and roboticists. It partially underpins scientific publications. The data is probably too specific to allow for integration.

2.4.3. Standards and metadata

Data is labeled with experiment ID, numbered in chronological order or alternatively labeled directly with the time in the experiment. For each experiment dataset there is a short textual description of the experiment itself, when it was started, when it was ended, and who supervised it. For the data that is uploaded to Zenodo, Zenodo defines and handles the metadata (see above). Version control is not necessary.

2.4.4. Data sharing

Given the amount of this kind of data that will be generated throughout the project, we do not plan a central data deposit for all of it. Individual photos and videos that are representable for the respective experiment and/or relevant for publications will be stored in the project's Cloud service (OwnCube e.U., An der Liesing 2-34/7, 1230 Wien, Austria) for the duration of the project, shared among project partners, and data requests for that data from outside the project will be handled individually. A smaller selection of that data will be uploaded to the servers of Zenodo and hence allow for easy sharing.

2.4.5. Archiving and preservation (including storage and backup)

For the data stored on our cloud service we cannot guarantee preservation beyond the project's lifetime but that is taken care of by the smaller selection of data that will be uploaded to Zenodo (lifetime of the host laboratory CERN is 20+ years, see above).

2.5 DATA5 – logging & tracking data of experiments

2.5.1. Data set reference and name

DATA5 – logging & tracking data of plant & robot experiments

2.5.2. Data set description

Generated data acquired during experiments with natural plants and robots including data generated by image processing of photos and logging data of our hardware systems. Data is of interest for plant scientists and roboticists. It partially underpins scientific publications. The data is probably too specific to allow for integration.

2.5.3. Standards and metadata

Data is labeled with experiment ID, numbered in chronological order or alternatively labeled directly with the time in the experiment. For each experiment dataset there is a short textual description of the experiment itself, when it was started, when it was ended, and who supervised it. For the data that is uploaded to Zenodo, Zenodo defines and handles the metadata (see above). Version control is not necessary.

2.5.4. Data sharing

Given the amount of this kind of data that will be generated throughout the project, we do not plan a central data deposit for all of it. Individual photos and videos that are representable for the respective experiment and/or relevant for publications will be stored in the project's Cloud service (OwnCube e.U., An der Liesing 2-34/7, 1230 Wien, Austria) for the duration of the project, shared among project partners, and data requests for that data from outside the project will be handled individually. A smaller selection of that data will be uploaded to the servers of Zenodo and hence allow for easy sharing.

2.5.5. Archiving and preservation (including storage and backup)

For the data stored on our cloud service we cannot guarantee preservation beyond the project's lifetime but that is taken care of by the smaller selection of data that will be uploaded to Zenodo (lifetime of the host laboratory CERN is 20+ years, see above).

2.6 DATA6 – logging data & images of simulations

2.6.1. Data set reference and name

DATA6 – logging data & images of simulations

2.6.2. Data set description

Generated data acquired during simulations of natural plants and robots including logging data of simulations, screenshots, and other visual representations of data. Data is of interest for plant scientists and roboticists. It partially underpins scientific publications. The data is probably too specific to allow for integration.

2.6.3. Standards and metadata

Data is labeled with experiment ID, numbered in chronological order or alternatively labeled directly with the time in the experiment. For each experiment dataset there is a short textual description of the experiment itself, when it was started, when it was ended, and who supervised it. For the data that is uploaded to Zenodo, Zenodo defines and handles the metadata (see above). Version control is not necessary.

2.6.4. Data sharing

Given the amount of this kind of data that will be generated throughout the project, we do not plan a central data deposit for all of it. Individual photos and videos that are representable for the respective experiment and/or relevant for publications will be stored in the project's Cloud service (OwnCube e.U., An der Liesing 2-34/7, 1230 Wien, Austria) for the duration of the project, shared among project partners, and data requests for that data from outside the project will be handled individually. A smaller selection of that data will be uploaded to the servers of Zenodo and hence allow for easy sharing.

2.6.5. Archiving and preservation (including storage and backup)

For the data stored on our cloud service we cannot guarantee preservation beyond the project's lifetime but that is taken care of by the smaller selection of data that will be uploaded to Zenodo (lifetime of the host laboratory CERN is 20+ years, see above).

3 Conclusion

A data management plan is not a fixed document, instead it might be refined later during the project. The feasibility and utility of this data management plan will be monitored during the next months and years of the project. Also the number of data requests from outside the consortium will be recorded. Depending on that experience we will consider appropriate actions and changes of this plan.