PAINLESS

Energy-autonomous portable access points for infrastructure-less networks

Grant agreement no. 812991
H2020-MSCA-ITN-2018

ITN - Innovative Training Network

D1.3– Data Management Plan

WP 1 – Project Management

Due date of deliverable:	Month 6	
Actual submission date:	02/ 04 / 2019	
Start date of project:	01/10/2018	Duration: 48 months
Lead beneficiary for this deliverable:	UCL	
Last editor:	UCL	
Contributors:	All beneficiaries	

Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

History table

Version	Date	Released by	Comments
V1.0	12.03.2019	Christos Masouros	n/a
V1.1	13.03.2019	David Lopez-Perez	n/a

Table of contents

Key word list

- Data management
- FAIR data
- Metadata
- Open access

Definitions and acronyms

Acronyms	Definitions
PAINLESS	Energy-autonomous portable access points for infrastructure-less networks
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-usable.
CA	Consortium Agreement
MB	Management Board

1. Introduction

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

The purpose of data collection in PANLESS is to aid the design of new telecommunications techniques and technologies, and to examine and verify their usefulness. The purpose of data generation will be again to verify the usefulness of the techniques, and the overall success of the project, as well as to disseminate and demonstrate the outcomes of the project.

Types and formats of data

PAINLESS may collect a) wireless network traffic data, to aid the development of telecommunication techniques, b) measurement data from our experiments, to evaluate and refine our techniques. There is no personal data to be collected throughout the progress of the program. Any data collection within the proposed research, may involve test data and measurements for the training and development of communication and energy management algorithms. Within the training program, data collection may involve attendance statistics and attendance sheets, cleared by the attendees as per the GDPR. Generated data may involve a) test/traffic data and results, b) software that implements the developed algorithms, c) measurement and experimental results, d) papers and reports of our outcomes. These may be in the formats of raw data sets, new software and codes, or documents.

Existing data re-use

Test, measurement or traffic data may be used to develop and refine our wireless communication and energy management techniques.

Origin of the data

Traffic and measurement data may already reside with the PAINLESS partners from previous research or be requested from external parties in the course of the project. All availability of existing data is subject to the IP regulations detailed in the CA, and for external partners, subject to IP agreements if needed.

Size of the data

Up to a few Gbits of data.

Data utility

The generated data will be useful to the research community for further development, to industry for commercialization and standardization, and indirectly to the public that will benefit from the PAINLESS technologies.

2. Activities carried out and results

2.1. Making data findable, including provisions for metadata

- The project publications will be published in IEEE, IET or other journals and conferences, all of which have unique identifiers such as Digital Object Identifiers. Other data such as measurement results and codes to be made accessible (subject to IP restrictions) will have an associated metadata document (stored as a .txt file) which describes key aspects of the data.
- Event listings are stored in a central spreadsheet and individual events are assigned a unique identifier of the format XXX_YYYYMMDD where XXX is the partner short name (as defined in the definitions and acronyms table) and YYYYMMDD is the start date of the event.
- Project deliverables are assigned a unique identifier PAINLESS-DX.X-YYYYMMDD. All files made publicly available should reference PAINLESS in their name, and we recommend the convention PAINLESS-xxxxxxxx where xxxxxxx is a meaningful short description. Photographs and audio/visual recordings should be named PAINLESS-XXX-YYYYMMDD-nnnnnnnn where XXX-YYYYMMDD is the event identifier and nnnnn is a brief description of the event/photograph content.
- An allowance has been made for this in the project metadata to optimize possibilities for re-use.
- Every Dataset will have an associated text document with its associated metadata.

2.2. Making data openly accessible

- The only data which de-facto will not be made openly accessible will be data which contains personally identifiable information (e.g. individual evaluation forms). These will be summarised, and any individual forms used for research publications (such as inclusion in 'user stories') will be redacted or anonymised before online storage. In addition, datasets, measurements, codes that are IP restricted as per the CA will not be made available in full, but the consortium will strive to make meaningful parts of these available for reproducibility. We will also strive to keep such restricted data to a minimum.
- During the project, a subset of summary data (e.g. event visitor statistics and feedback summaries) will be made accessible by one or more methods below:
  - Via newsletters, reports and other publications on the online knowledge sharing platform (togetherscience.eu) developed as part of WP3;
  - Via partner's local websites;
  - Via social media;
  - The PAINLESS website will provide open-access to the summer-schools proceedings ensuring a wide spread of the results and an increased awareness of the excellence of the PAINLESS network;

  - The project's journal/magazine articles will be made available to the wide public through open access and self-archiving, such as ArXiv, OpenAir, IEEE Open Access, and we will pursue open access publication venues.
- Detailed data will be available to all consortium partners via the project shared drive (with the exception of individual questionnaires which will be stored at each partner's premises). The access to this drive is restricted to project partners. Should other individuals wish to access the data for research purposes during the project, it will be openly shared on request. At the end of the project, data to be preserved will be stored in a suitable data repository. At this stage, we are using Microsoft Sharepoint.
- Data will be published using standard file formats (pdf, csv and others).
- With the exception of the knowledge sharing platform, all data will be accessed using standard tools. It is the responsibility of the Beneficiaries to provide appropriate documentation to make measurement results and software readily accessible and reusable.
- A relevant software is not seen as being a requirement, but should it be needed, we will provide the required open source to access and analyse the data, such as codes implementing our algorithmic solutions, or measurement/test results.
- For the duration of the project, any data and associated metadata and documentation will be stored on the shared drive, with no restrictions on use. At this stage, we are using Microsoft Sharepoint. Access conditions will be based on the FAIR principles. Internal or confidential data will only be accessible on a password-controlled central storage facility. For open data, we have not identified a need to identity the person accessing the data.

2.3. Making data interoperable

Data produced in the project are interoperable, therefore standard file formats and inter-disciplinary vocabularies will be used to facilitate data exchange and re-use. It is envisaged that every dataset will have metadata, aside of the project publications which will be open access and accessible as outlined in the previous section.

2.4. Increase data re-use

It is planned that Creative Commons Licenses will be used for all data to be preserved. Data will be made available in accordance with what specified in the Consortium Agreement Section 9, i.e.:

- Access Rights to Results and Background Needed for the performance of the own work of a Party under the Project are requested and granted on a royalty-free basis.
- Access Rights to Results if Needed for Exploitation of a Party's own Results shall be granted on Fair and Reasonable conditions, subject the Party requiring the grant of such Access making a written request to the Party from which it requires the Access Rights.

- Party and the Requesting Party shall not be otherwise deemed granted.
- Access Rights to Affiliated Entities will be granted on Fair and Reasonable conditions and upon written bilateral agreement.

All Personal Identifiable Information will be restricted to internal usage and not going to be shared with third parties. For shared information, standard format, open source software, and proper documentation will guarantee re-usability by third parties.
Data will remain re-usable for 10 + years, subject to EC policy changes.
Quality Assurance is the responsibility of the MB of the project.

2.5. Allocation of resources

An allowance of £2,440 has been made by the co-ordinator to cover the project web-site and archiving and storage requirements (including manpower to prepare and manage data as well as storage fees). Any additional costs will be covered by the project's common basket.
The MB is the ultimate responsible body for data management.

2.6. Data security

All envisaged including any personal data such as individual questionnaire responses will be stored in the project's share point, which will only be accessible on a password-controlled central storage facility. Personal data will be destroyed at the end of the project or as per GDPR regulations.