This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776503
- CIRC4Life-H2020-IND-CE-2016-2017/CIRC-2017/TwoStage

A circular economy approach for lifecycles of products and services

Data Management Plan - #2

Deliverable 10.3.

PROJECT INFORMATION	
Type of Project	European Commission Horizon 2020
Topic	CIRC-01-2016-2017 Systemic, eco-innovative approaches for the circular economy: large-scale demonstration projects
Grant Agreement No.	776503
Project Duration	01/05/2018 – 30/04/2021 (36 months)
Project Coordinator	Nottingham Trent University (NTU)
Project Partners	Enviro Data (ENV), Jonathan Michael Smith (JS), Kosnic Lighting Limited (KOS), Centre of Research for Energy Resources and Consumption (CIR), European EPC Competence Center GmbH (EECC), The Institute for Ecology of Industrial Areas (IETU), SWEREA IVF AB (SWE), Make Mothers Matter (MMM), ONA PRODUCT (ONA), INDUMETAL Recycling (IND), GS1 Germany GMBH (GS1G), Laurea University of Applied Science (LAU), Center for European Policy Studies (CEPS), Institute of Communication and Computer Systems (ICCS), Recyclia (REC), S.A.T. Alia (ALIA)

DOCUMENT INFORMATION	
Title	Data Management Plan - #2
Version	V1.0
Release Date (dd.mm.yy)	30.10.19
Work Package	WP10
Dissemination Level	PU

DOCUMENT AUTHORS AND AUTHORISATION	
Document Responsible	NTU
Contributors	Bonnell, Jane (NTU); Wu, You (NTU)
Reviewed by	LAUREA, INDUMETAL, ICCS
Approved by	Professor Daizhong Su (NTU)

DOCUMENT HISTORY			
Version	Date (dd.mm.yy)	Description	Implemented by
V0.1	13.10.19	First draft	NTU
V0.2	21.10.19	Reviewers' comments are addressed.	NTU
V0.3	22.10.19	NTU Research Data Management team clarified the data security policies in SharePoint and the ICT repository.	NTU

Summary

The Horizon 2020 FAIR Data Management Plan (DMP) is designed to be applicable to any Horizon 2020 project that produces, collects or processes research data. CIRC4Life consortium has developed a single DMP for the project to cover its overall approach.

Deliverable 10.3, the CIRC4Life DMP#2, has been produced following a review of the CIRC4Life DMP#1 (D10.2). Deliverable 10.2 outlines the project's approach to the management of research data in accordance with Article 29.3 of the Grant Agreement. It details what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. Deliverable 10.3, the CIRC4Life DMP#2, therefore offers a status update as to the implementation of the plan. There are no significant revisions required to the project's research data management (RDM) strategy. At the point of writing Deliverable 10.3, there are no known changes or restrictions to data sharing as initially reported in Deliverable 10.2. Therefore, much of the plan remains unchanged. However, more information is now offered about the data that has been generated by the project under this reporting period (M6-M18).

A list of revisions is summarized as follows compared to DMP#1 (D10.2):

- Overview of the project data that are generated and used in this reporting period (Table 2.1; Section 2).
- Additional information is supplied in Section 3.1 to outline the techniques that will be employed to make data findable, including how NTU's institutional repository (IR) will be harnessed to increase the discoverability of project data (Section 3.1).
- Data security polices in ICT platform are reported (Section 5.1 – 5.3).
- Overview of the ethical related practices in this reporting period (Section 6).

Table of Contents

Acronyms and abbreviations

Abbreviation	Description
CEBM	Circular Economy Business Model
CSV	Comma-separated values
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
EPCIS	Electronic Product Code Information Services
EPC	Electronic Product Code
FAIR	Findable, Accessible, Interoperable and Re-usable
FNC	File Naming Convention
GDPR	General Data Protection Regulation
ICT	Information and Communications Technology
IR	Institutional Repository
IP	Intellectual Property
PC	Project Coordinator
RDM	Research Data Management
XML	Extensible Markup Language

1 Introduction

CIRC4Life project complies the FAIR data management concept to develop this DMP. FAIR data management requires the project 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.

Deliverable 10.3 is not intended as a strict technical implementation of the FAIR principles, it is inspired by FAIR as a general concept. The following documents have been referred in order to develop the CIRC4Life DMP:

- Guidelines on FAIR Data Management in Horizon 2020 (EC, no date)
- FAIR data principles (FORCE11, no date)
- FAIR principles (Wilkinson *et al.*, 2016)

The FAIR DMP template (EC, no date) is a set of questions that should be answered with a level of detail appropriate to the project. The DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, DMPs should have a clear version number and include a timetable for updates. As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

2 Data Summary

The CIRC4Life project involves an ICT (Information and Communications Technology) solution using a variety of technologies, which are important to provide the means to implement and assess the effectiveness of the Circular Economy Business Model (CEBM). These technologies will drive the data generation and collection processes and help inform the testing against the proposed methodologies of using an eco-point approach, sustainable product design and production, real-time tracking and monitoring technology based on EPCIS (Electronic Product Code Information Services), information logistics sharing infrastructure across the supply chain and associated data security and privacy.

Given that the ICT solution will use a number of tracking technologies, potentially including barcode, and EPCIS. The project process tasks will include full documentation on barcode code types, and EPCIS; and with regards the latter the EPC (Electronic Product Code) Core Business Vocabulary identifiers used. The documented descriptions will also contain an explanation on why particular identifiers or models were used for all the tracking technologies used, and how they related to the intended outcomes.

Example datasets with these descriptions will be supplied in an XML as well as Microsoft Word DOCX format for the descriptions, and example tracking data datasets are in CSV and Microsoft Excel XLSX format that will be compatible with as many software applications as possible.

Subsequent research will benefit from analysing the chosen identifiers used in attempting to improve the improved sustainability performance model. A clear description of the identifiers used, as well as the accurate documentation and example datasets of tracking technologies used, would be critical to this type of analysis.

With regards anticipating the expected size of data to be produced, the nature of the CIRC4Life project involving a number of collaborators makes an accurate guess difficult at this early stage, but it is proposed that the data generated will be in the order of Gigabytes (approx. 100), rather than Terabytes.

The importance of making these datasets FAIR (Findable, Accessible, Inoperable and Reusable) is recognised and planned for, but taking into account the demands of the 'living' nature of a data management plan as the project progresses, the final form of the datasets cannot yet be declared in detail, but the Project Coordinator (PC) can offer an outline of the data that has been collected.

To date, research activity has primarily focussed on intelligence gathering and project scoping. There now exists a subset of project data associated with the development of models, techniques and processes. These have been reported in the following project deliverables already submitted to the EU portal (See Table 2.1). Each of these deliverables and their associated data are contributing to the further project activities of the relevant WPs.

The project is also collecting data directly from consumers via online surveys and Living Labs in this reporting period. For example, consumer attitudes to reuse & recycling LED lighting products and the eco-points System have been measured using Survey Monkey, an online survey tool. A complete overview of the data collected, and an analysis of these results is reported in Deliverable 3.4 - Report on consumer satisfaction survey (M18).

Publicly available information about the project and some data are made available via the CIRC4Life Project Website, available at: https://www.circ4life.eu/

The CIRC4Life SharePoint Site stores all project data for the duration of the grant (see Section 5.1). Data is grouped and organised under each WP. Therefore, this site gives the Project Co-ordinator complete oversight of all data created by CIRC4Life is. The following table gives a summary of the data that has been collected to date; and, their relation to the WPs and corresponding deliverables. These datasets will be utilised as each WP progresses to inform the development and/or creation of future project outputs.

Table 2.1 Overview of CIRC4Life collected data in this reporting period (M6-M18)

Tasks/Tools/Activities related to data	Format	Work Package	Related deliverables in this reporting period (M6-M18)	Responsible Partner(s)
Recycle bin cards, User Profiles	JSON	WP2	D3.1 - Development of eco-shopping and eco-account tools	NTU, EECC, ICCS
Recycle bin traceability data	EPCIS, JSON	WP2	D2.3 - Development of the ICT system for reuse/ recycling	NTU, EECC, ICCS
Data entry Tool	JSON	WP4	D4.3 - Report on development of adaptor systems for eco-point, LCA and EPCIS event data interoperability	ENV, ICCS
CIRC4Life Eco Account	JSON	WP4	D3.1 - Development of eco-shopping and eco-account tools	NTU, ICCS
CIRC4Life Master Product Data	JSON-LD	WP4	D4.3 - Report on development of adaptor systems for eco-point, LCA and EPCIS event data interoperability	ENV, ICCS
CIRC4Life Live Product Data	JSON-LD	WP5	D5.2 - Development report and documentation for traceability components and tools	EECC, ICCS
LCA Module	.xlsx	WP1	D1.2 - Report on sustainable (environmental, social and economic) impact analysis	NTU, JS, KOS, ONA, ALIA
Brokerage Tool	XML	WP4	D7.3 - Report on the stakeholder involvement along the supply	GS1, ICCS
Workshop data from Living Labs [1]	.doc	WP7	D7.4 - Experience and recommendations of end-user engagement across circular business model development	LAU
Stakeholder contact list	.csv	WP9	D9.1 - Communication plan	MMM, NTU
Consumer surveys	.csv	WP3	D3.4 - Report on consumer satisfaction survey	MMM

[1] For all data, since actual data acquisition will occur during demonstration activities, it is implicitly noted that WP7 activities are relevant besides the technical WP that is referred to.

3 FAIR Data

3.1 Making Data Findable, Including Provisions for Metadata

Through the active phase of the project, data management will include simple organisational measures such as following a file naming convention (FNC) and will take the form of CIRC4Life_ DocumentName_ ResponsiblePartner _ YYYYMMDD_ Version.docx. Document filenames will be kept short to avoid unnecessarily long paths, always include the last person to edit the document, and a version indicator.

For example, CIRC4Life_ProjectManagementPlan_NTU_20181009_V0.1.docx

All project documentation will be stored in a Microsoft project SharePoint site (as part of an Office 365 for Business cloud service instance) that will enable full control over editing permissions of project participators. SharePoint as a collaboration and document management platform also offers good functionality for platform-wide metadata control of content.

For additional knowledge management, an Excel file will be included and promoted that contains a list of keywords, and the accepted definitions to be used by all the participating project partners. It will also act as a controlled vocabulary to ensure consistent knowledge organisation that will aid subsequent retrieval and re-use. This document will be easily accessible for project collaborators using the project SharePoint site and included on the project website.

Project data will be made discoverable through the inclusion of a detailed descriptive record that will be added to NTU's institutional repository. This service is indexed in Google and Google Scholar; therefore, the records will be retrieved when anybody searches for the keywords associated with the CIRC4Life project. Further information about this is provided in the section below. A data availability statement will also be included in the project's published outputs to direct readers to a full overview of project data, as well as the terms and conditions of accessing and using publicly available data. This method will help increase the visibility of the data and make it easier for people to locate and access them.

3.2 Making Data Openly Accessible

The final datasets to support published outcomes will be deposited in Arkivum, the NTU instance of the Archive-as-a-Service data safeguarding and long-term lifecycle management solution. A Digital Object Identifier (DOI) will be minted using DataCite service, a global provider of DOIs for research data, and included in the description (with additional metadata) in the NTU IR to facilitate persistent identification and discovery. The metadata included in the IR record will identify the file formats the datasets are available in, which will indicate to potential users the software application required to effectively use the datasets. Any software code will also be described in the IR and archived, but intellectual property (IP) implications may restrict access. The development of the ICT solution will be accompanied by software documentation that will provide instructions for reuse, and these will be archived and made available with the proviso again of potential IP implications.

NTU uses DataCite Metadata Schema to describe the data and is actively updated and promoted by interactively coordinating with community standards, like ORCID (https://schema.datacite.org/). Access to datasets will be open and currently available as a request/mediated service via the Library Research Team at NTU, which will necessarily identify the individual requiring access to the datasets and provide statistical information on dataset usage. However, availability will always be determined by licence conditions, and that will allow for some granularity of access (including restrictions) to be potentially specified by different project collaborators. DataCite also has a reciprocal dataset registry service to aid discoverability and potential reuse of datasets that have been determined to be openly available.

The data will be preserved for a minimum of ten years in the Arkivum service, or in any other subsequent solution used.

Access to project documentation and data is only available to those who have access to the SharePoint site, which is determined by the PC and implemented by NTU research staff participating in the project. Additional backup and data restore procedures will be agreed using this solution full functionality. A further backup will be performed daily to an on-premise storage solution hosted by NTU.

As stated NTU has a long-term storage solution (i.e. Arkivum), available to enable preservation and curation. It is an Archiving-as-a-Service solution that combines a hybrid local and cloud storage synchronised combination that guarantees data integrity and redundancy with full security. This ensures that funder, institutional or publisher retention compliance is satisfied, as is the authenticity of the original data for open data requirements or post research review if necessary.

3.3 Making Data Interoperable

NTU uses DataCite metadata schema as DataCite looks to community practices that provide data citation guidance. The PC will foresee that SharePoint project site will include appropriate mandatory DataCite metadata elements for project files. Such as (Creator, Title, Description, Access to the dataset, Data Collection Method, Data Processing/Preparation Activities), and are included as mandatory metadata classification fields for file inclusion on the SharePoint site.

Using standard vocabularies for all data types is not initially anticipated as the data definition is integral to the tracking technologies of barcodes and the EPCIS Standard already. It will be clearly stated what barcodes code type used, and EPCIS Standard used.

However, the PC will ensure that any datasets that are included as CSV will have a description definition for all the data elements included, and these are mapped to the industry standard description of the identifiers used for the different tracking technologies.

3.4 Increase Data Re-use (Through Clarifying Licences)

All openly available data will be offered on a CC-BY SA share-alike basis, so attribution will be required and that any repurpose, or re-use will be shared on the same basis. IP requirements may determine that some datasets are not open and will remain closed but archived in Arkivum. The description of the individual datasets in the NTU IR will make the conditions of the licence clear.

Archival in Arkivum is for ten years, and will always be available, allowing for licence conditions, for external viewing during that period.

During the active phase of the project, data quality assurance process will be organised by the PC (responsible individual for data management) to discover inconsistencies and other anomalies in the data, as well as performing data cleansing activities (e.g. removing outliers, missing data interpolation) to improve the data quality. It is foreseen that this will involve sampling datasets initially, with a thorough assessment of particular grouped datasets if significant inconsistencies are discovered.

4 Allocation of Resources

The costs of using Arkivum Archiving-as-a-Service solution available at NTU at the end of the project is £0.50 per GB of archived data. As the projected final amount of data is expected to be in the order of 100 (estimation) Gigabytes, this will cost approximately £50 per year, or £500 for ten years.

NTU will support the archiving of funded research datasets to promote the open access and data agenda.

The responsibility of data management on the CIRC4Life project is the PC of the leading collaborating institution (NTU). Everyday workflow tasks will be delegated, but the PC will ensure that consistent data management is performed for the duration of the project and will conduct six monthly reviews on the use of controlled vocabulary, file naming and versioning conventions and that the organisational logic of the SharePoint site is adequate.

The PC will be responsible for selecting the datasets to be archived, for a period not less than ten years. Preservation will support publication outcomes, research deemed of long-term value, as well as project communication channels dissemination output and literature. Should personal details be included in the data preserved, then anonymity need to be maintained, but traceable should there be a need for source data verification.

Any data deemed to be not worth saving in the active and archival stage needs to be destroyed in accordance with NTU Information Systems data destruction policy.

5 Data Security

As described in Section 3.2, the long-term security and preservation of project data will be managed by NTU using the Arkivum appliance, known publicly as The NTU Data Archive. This section details the security arrangements for project data during the grant period. Data will be moved from the Circ4Life SharePoint site to Arkivum at the end of the project, or before if data is released simultaneously with scientific publications.

5.1 Data Security Policies in CIRC4Life ICT Platform

All CIRC4Life data will be collected, stored, protected, shared, retained and destroyed, upholding state-of-the-art security measures and in full compliance relevant EU legislation, bearing in mind the demands of crowdsourcing and flood modelling contexts. As a general rule, this data will be stored on paper and/or computer files protected by state-of-the-art physical and logical security measures: the archives containing the paper folders are locked; the computer files are stored in computers and hard disks, accessible only by authorized personnel (within the relevant CIRC4LIfe partners) through password. This data will not be in any case shared with or disclosed to anyone outside the research team until data has been finalised for publication and approved for release by the Consortium.

A data minimisation policy is adopted by CIRC4LIfe, which means that only data strictly necessary for running the participatory and demonstration activities is collected and processed. Personal data, if any, collected and stored within CIRC4Life and for the purpose of the project aforementioned activities will be permanently and irrevocably erased on the project completion. Nevertheless, only if an individual participant has provided his/her free, specific and informed consent, name, age, professional occupation and professional views will this data be included in project outputs. If such a consent is not provided by the individual participant, only information that may be processed in a way that inhibits tracing his/her opinions back to him/her (anonymised information) will be a part of the activities.

In particular and with respect to access control and data protection, CIRC4LIfe data (including data from all the input sources) will be collected by the CIRC4Life ICT platform (hosted in ICCS premises) which will store

them in the relevant databases. This platform and the associated data repository system provides the means for deploying access control policies as described in Task 4.3. These are flexible and fine-grained means to assign permissions to roles and users in such a manner that access to resources can be controlled sufficiently for all eventualities. Moreover, the system will provide secure access to data repository by using security protocols such as OAuth2 for authorization and secure encrypted HTTPS calls. More specifically an open source software, KeyCloak, is used for the access to the ICT platform and the authentication for all the applications.

5.2 CIRC4Life End Users Personal Data

CIRC4Life ICT platform data will be stored in the ICCS data repositories. These are secure servers with limited access to Internet (with established filtering rules allowing access only for specific web requests or for authorized personnel through a VPN network). Security provisions are also taken for the physical infrastructure (rooms) where these servers reside. Furthermore, a specific encryption with the usage of KeyCloak will provide an additional level of security to the personal data.

5.3 Data Anonymization

Data anonymization refers to the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

However, the explicit introduction of anonymization is not sufficient by itself to preclude any other measures of data protection. Therefore, security policies for data protections should always be enforced as strictly as possible. The controllers of the CIRCC4Life ICT Platform should require that data anonymization is enforced before any dataset is uploaded to the ICT Platform.

6 Ethical Aspects

Any personal data gathering within the CIRC4Life project will conform to informed consent expectations that are expected with regards current Data Protection legislation, and the EU General Data Protection Regulation (GDPR) that started to implement on 25th May 2018.

The project team has developed the CIRC4Life Ethical Clearance Checklist. This must be completed by each team whose responsible task involve with ethic issues. This checklist addresses aspects associated with data use and data retention. It requests confirmation that the team is familiar with GDPR and that the activity has been designed through close consideration of the issues surrounding Data Protection. The archive of signed checklists from each team (see Table 6.1) are placed in the CIRC4Life SharePoint Site to provide administrative oversight to the project management of all ethical aspects of RDM. All participant information sheets explain how data will be used during the project. This is provided in language of the participants and is written in clear and straightforward language. The CIRC4Life Survey Privacy Notice is an example of this. These are available here: https://www.circ4life.eu/survey-privacy-policy.

An ‘Ethics requirements’ work package has been developed, as Deliverable D11.2, D11.3 and D11.7 of Work Package 11 has been submitted in this reporting period, in order to addresses the ethics requirements of the CIRC4Life project.

Table 6.1 Archive of CIRC4Life Ethical Clearance Checklist in this reporting period (M6-M18)

WP/Task name	Signed by	Signed time	Brief description on the activities and objectives
Task 7.2 - Implementation living labs (M10-M30)	KOS, LAU	23/05/2019	Leasing service workshop in Lighting Industry Association (LIA), Telford, UK on 28 May 2019.
Task 7.2 - Implementation living labs (M10-M30)	ALIA	29/03/2019	To organize at least three workshops in relation to Task 7.2. Living Lab Implementation Activities, in order to plan DEMO4. Each workshop will last around 2 h 30 minutes and will involve both, citizenship and public administration.
Task 8.4 - Policy alignment (M2-M33)	CEPS	03/04/2019	The interviews for the first sub-task (analysis of policies and regulations) were conducted between October and December 2018. The team collected information through interviews with companies involved in the demonstration of the project's circular business models but also with other companies involved in the electronics and food value chains in order to collect additional perspectives and information.
Task 2.6 - End-user awareness for reuse/recycling (M10-M16)	IND	27/03/2019	Three different primary schools were selected in order to carry out a training process with students and professors, focused on end-users awareness for recycling and reuse. During this training, besides the educative process itself, the intelligent container will be placed in the school in order to give the opportunity to the students to dispose their WEEE and put in practice the acquired knowledge. The participants are the students and the professors. Neither photos nor videos will be taken without authorization and under no circumstances personal data will be collected or used.
WP7 - Stakeholder Interaction and End-user Involvement (M1-M30)	LAU	19/10/2018	First Innovation Camp event 2018 included 70 participants, out of which 40 external participants. External participants were selected based on the procedures and selection criteria described in D11.1 section 2.2. Participants were classified based on the Quadruple Helix type (academia, business consumer, policy) and selected to ensure transparent and fair participation of QH types, as well as gender and geographical balance.
Task 3.6 - Consumer (satisfaction) survey (M13-M18)	MMM	08/07/2019	Conduct online surveys using Survey Monkey in order to easily process and analyse the data; Complement these surveys which tend to attract only certain type of consumer with a medium/high education with physical interviews using questionnaire in key spots such as recycling/collecting points, zero waste/organic shops, and trade fairs in several countries.
Task 6.3. - Demonstration of CEBM with tablets (M19-M33)	REC	07/05/2019	A pilot test will be conducted in the Basque Country (Spain), specifically in Getxo municipality. For that, Getxo's inhabitants will employ the APP developed in the project and the intelligent containers placed in Getxo for disposing their devices. The participants of this test pilot will do it voluntarily and their number is unknown. Any personal data will be linked to the APP and will comply the GDPR.

7 Outlook and Conclusions

With regards, conducting research project NTU expects staff and students involved in research to adhere to the policies pertaining to 'Code of Practice for Research', 'Research Ethics Policy' and the 'Research Data Management Policy'. NTU has also sought to follow the requirements and recommendation of Horizon 2020 EU funding as described in this document. Lastly, as a UK HE organisation, NTU has provided support for research active members to be aware of national guidelines on open data, and to follow the principles wherever possible.

The DMP will be the guide document for project's data treatment and management. As has been seen, the DMP describes which and how data is collected, processed or generated, but also outlining the methodology and standards used. Furthermore, the DMP explains whether and how this data is shared and/or made open, and how it is curated and preserved. Finally, it should be taken into account that the DMP evolves during the lifespan of the project. Thus, this second version will be updated in M36 to reflect the status of the CIRC4Life project with respect to its data management.

8 References

EC (no date) *Data management - H2020 Online Manual*. Available at: http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm (Accessed: 10 October 2018).

FORCE11 (no date) *The FAIR Data Principles*. Available at: https://www.force11.org/group/fairgroup/fairprinciples (Accessed: 10 October 2018).

Wilkinson, M. D. *et al.* (2016) 'The FAIR Guiding Principles for Scientific Data Management and Stewardship', *Scientific Data*, 3, p. 160018. doi: 10.1038/sdata.2016.18.