This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **101006631**.

Supporting the Electric Vehicle REVOLUTION through maximising EV Range and End-of-Life Vehicle Recovery through optimisation of recycled plastics and advanced light materials

D7.2 Preliminary Data Management Plan
WP1 – Project Management

Deliverable Lead	Partner Name
Deliverable Due Date	30/06/2021
Actual Submission Date	30/06/2021
Version	2.0

Document Information

GRANT AGREEMENT NUMBER	101006631	ACRONYM		REVOLUTION
FULL TITLE	Supporting the Electric Vehicle REVOLUTION through maximising EV Range and End-of-Life Vehicle Recovery through optimisation of recycled plastics and advanced light materials.			
START DATE	01.01.2021	DURATION		36 months
PROJECT URL	REVOLUTION Website			
DELIVERABLE	D7.2 – Preliminary Data Management Plan			
WORK PACKAGE	WP7 – Project Management			
DATE OF DELIVERY	CONTRACTUAL	30.06.2021	ACTUAL	30.06.2021
TYPE	Report	DISSEMINATION LEVEL		Public
LEAD BENEFICIARY	Iconiq Innovation			
RESPONSIBLE AUTHOR	Stephen Ryley			
CONTRIBUTIONS FROM	Emre Elmas, Berk Gencer, Carlos Barreto, Santiago Moreira, Tugba Okay, Tommi Vuorinen			
TARGET AUDIENCE	☒ Public ☐ Restricted to other programme participants (including the Commission Services) ☐ Restricted to a group specified by the consortium (including the Commission Services) ☐ Confidential, only for members of the consortium (including the Commission Services) ☒ ORDP: Open Research Data Pilot			
DELIVERABLE CONTEXT/ DEPENDENCIES	This document is D7.2 Preliminary Data Management Plan. A Final Data Management Plan D7.2 will be produced at M36.			
EXTERNAL ANNEXES/ SUPPORTING DOCUMENTS	None			
READING NOTES	None			
ABSTRACT	The REVOLUTION Data Management Plan will be a live document throughout the project which defines how project data will be managed. It covers data types and sets, methodology and standards, security, open access and curation.			

Document History

VERSION	ISSUE DATE	STAGE	DESCRIPTION	CONTRIBUTOR
1.0	31/03/21	Draft	First draft	Stephen Ryley
1.1	26/06/21	Draft	Update following partner and WP leader feedback	Emre Elmas, Berk Gencer, WP leaders
1.2	29/6/21	Draft	Annotations removed	Stephen Ryley
2.0	30/06/2021	Ready to submit	Revision of design and minor changes	Emre Elmas
2.0	14.07.2021	Ready to resubmit	Removal of "business" classification and change EU emblem	Emre Elmas

Timetable for Updates

Deliverable Number	Version	Due Date
D7.2	Preliminary Data Management Plan	M6
n/a	Mid-point Data Management Plan	M18
D7.3	Final Data Management Plan	M36

REVIEW STATUS	☐ Draft ☐ WP leader accepted ☒ Coordinator accepted
ACTION REQUESTED	☐ to be revised by Partners ☐ for approval by the WP leader ☒ for approval by the Project Coordinator ☐ for acknowledgement by Partners

Disclaimer

Copyright message

CONTENTS

LIST of TABLES

Abbreviations

DMP	Data Management Plan
ORD Pilot	Open Research Data Pilot
DOI	Digital Object Identifier
XML	Extensible Markup Language
CSV	Comma-Seperated Values (file format)
TIFF	Tagged Image Format
WP	Work Package
CO	Confidential
PU	Public

1. Executive Summary

This document is the Preliminary Data Management Plan (DMP) for REVOLUTION project. It has been produced following the EC Open Research Data Pilot guidelines and answers the questions of the EC H2020 DMP template. DMP will be a live evolving document throughout the duration of the project updated when there are significant changes which at a minimum will be at beginning the (M6), mid-point (M18) and end (M34).

The document details how the REVOLUTION project data will be managed. This includes:

- details of data types that will be produced and handled during and after the project
- preliminary list of expected datasets
- description of methodology and standards
- how data will be exploited or made accessible for verification and re-use
- how it will be curated and preserved.

This will follow the FAIR data approach – Findable, Accessible, Interoperable, Re-usable.

1.1 Attainment of the objective

This Preliminary Data Management Plan is evidence of completion of REVOLUTION D7.2. The content and timing are as planned with no deviations.

2. Introduction

2.1 Description of the Document

This document D7.2 is the first version of the Data Management Plan. A final version D7.3 will be submitted at M34. Additional version will be produced by the consortium at M18 and when there is a significant project change although there is no plan to submit these to the EC. This report has been created following the EC Guidelines for FAIR data Management in H2020 and follows the recommended template.

2.2 Purpose of DMP

DMP describes the data management life cycle for the data to be collected, processed and/or generated by the project. The process of collating the information in a report ensure that all partners are agreed and clear on how data will be managed during and after the project. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- the handling of research data during & after the end of the project
- what data will be collected, processed and/or generated
- which methodology & standards will be applied

- whether data will be shared/made open access and
- how data will be curated & preserved (including after the end of the project).

2.3 WP and Tasks Related with the Deliverable

This deliverable is one of a number that will be produced to fulfil WP7 Project Management Task 7.3 Quality Assurance, risk management and data management. In particular it is complementary with D7.1 Contingency Plan submitted at the same time (M6).

3. Data Summary

REVOLUTION project will generate and collect data during the planned work to develop and optimise production processes for recycled automotive plastic parts. Data collected will primarily be for the purposes of:

- developing machine learning and AI for the melt and injection moulding processes
- development of recycled plastic formulations
- validation and testing of parts
- economic, environmental, societal and sustainability analysis
- general project reporting and dissemination (images, graphics, reports)

The estimated data storage requirements for the project Work Packages are listed below. By far the largest data storage requirements will be required for WP2 predictive model control systems and demo data at approximately 500 GB. The total volume of all other project data will be approximately 20GB.

This is an acceptable and allowable volume for Iconiq REVOLUTION SharePoint.
Estimated data storage requirements for entire 3-year project duration:

WP1	3.5 GB	WP5	2 GB
WP2	10.5 GB + 500 GB	WP6	2.5 GB
WP3	0.5 GB	WP7	1 GB
WP4	1 GB		

The amount of data that will be uploaded to Zonodo open access repository will be dependent on agreements made between the partners on release of potentially commercially sensitive data but will be of the order of about 10 GB.

The types of data that will be generated are listed in Table 1 and a summary table of the datasets in appendix 1.

Long-term storage. Zenodo will be used for long term storage, a free open access FAIR repository supported by EC (OpenAIRE project) and hosted by CERN. Further details are given below.

Data standards. Choice of data formats has been based on industry standards and open data formats to ensure long-term useability. REVOLUTION will employ a metadata collection standard based on Dublin Core. Zonodo metadata is stored internally as JSON schema but can be exported in several standard formats including Dublin Core, MARCXML and DataCite Metadata Schema.

Type of data	Nature of data	Selected file format for sharing, re-use and preservation
Raw laboratory data	Laboratory experiments, protocols, results	Laboratory notebooks (including e notebooks) – lab notebook policies are in place within each partner organisation requiring meticulous lab book maintenance and regular sign off
Qualitative data	User requirement specifications, design specifications	eXtensible Mark-up Language (XML) text according to an appropriate Document Type Definition (DTD) or schema (.xml)
Quantitative tabular data with minimal metadata	Analytical performance data, laboratory experimental data	comma-separated values (CSV) file (.csv) or tab delimited file (.tab)
Quantitative tabular data with extensive metadata	Analytical performance data	SPSS portable (.por) file format
Digital image data	All images of project activities incl. consortium meetings, images of devices and device components, experimental data stills	TIFF version 6 uncompressed (.tif)
Digital video data Documentation	Demonstration of component and device operation; news and media items; project movie	MPEG-4 (.mp4)
Documentation	Draft and published manuscripts, project reports and deliverable documents.	Portable Document Format (PDF) (.pdf) Draft documents (.docx)

Table 1. Type, nature and format of data.

4. Fair Data

4.1 Making data findable, including provisions for metadata

Each data type will use an internationally accepted file format for its effective sharing, re-use and preservation. All data will be accompanied by appropriate data documentation which will include the names, labels, descriptions of variables, the explanations of any codes or classification schemes, as well as descriptions of derived data explaining how the data was processed, and what analyses or algorithms were used. REVOLUTION will also employ a metadata collection standard based on Dublin Core which will define the origin, creator, purpose, time reference, geographic location, access conditions and terms of use of the associated data. Metadata will facilitate effective searching and discovery of the data via the internet.

Naming convention: REVOLUTION_Data_WP#_<serial number dataset>_<dataset title>
Zonodo uses Digital Object Identifiers (DOI) making data and metadata citable and easily discoverable. DOI is resolvable, persistent and globally unique. Data files are versioned, each version has its own unique DOI. Data is also linked to EC project topic identifier. On creation of record, metadata is indexed and searchable on Zonodo's engine and is sent to and indexed on DataCite servers.

4.2 Making data openly accessible

REVOLUTION will use the following open repositories: (a) Zenodo; (b) Project Website (public deliverables and dissemination materials).

Zonodo has been specifically designed through the OpenAIRE project to allow open access to EC project data. All metadata will be openly available under CC0 licence. Other licences can be used if required. All open content is openly accessible through open APIs meaning easy integration with other Apps and thus also searchable from these Apps. All metadata and data files can be harvested through standard protocols such as HTTP and OAI-PMH. Source code is available as Open Source (GPL).

4.3 Making data interoperable

Standard vocabularies used by Zonodo will be used to ensure interoperability of data. For (meta)data Zonodo uses a formal, shared and broadly applicable language for knowledge representation. JSON Schema is used for internal representation of metadata. Metadata can be exported in other formats such as Dublin Core and MARCXML. Vocabularies used follow FAIR principals and when required Zonodo can use external vocabularies, e.g., license (Open Definition), funders (FundRef) and grants

(OpenAIRE). (Meta)data will include qualified reference to other data with each external reference qualified by a resolvable URL.

4.4 Increase data re-use (through clarifying licenses)

Zonodo items are retained for the lifetime of the repository. Zonodo host CERN currently has an experimental programme defined for at least the next 20 years. A succession plan is in place to transfer data to other repositories if the CERN repository closes.

Use and re-use is subject to the license under which the data objects were deposited. CC0 Licence will be used unless a specific need is identified for a different licence type. For Zonodo (meta) data are richly described, each record contains a minimum of DataCites manadatory terms and optionally DataCite's recommended terms and Zenodo's enrichments. (Meta)data are released with a clear and accessible data usage licence, data downloaded is subject to the licence specified by the uploader. All (meta)data is traceable to Zonodo user and can describe the original authors.

Zonodo Restricted Access: Users may deposit restricted files with the ability to share access with others if certain requirements are met. These files will not be made publicly available, and sharing will be made possible only by the approval of depositor of the original file. At present REVOLUTION do not plan to use this restricted access approach but may consider if a specific need is identified.

5. Allocation of Resources

The only cost, other than effort, associated with data management for REVOLUTION will be for Open Access publication. Size of datasets is small enough to be stored on internal partners standard servers and REVOLUTION SharePoint. Largest amount of data will be produced at Farplas who already have specific facilities in place so no extra costs will be incurred. In addition, Zonodo repository will be used for long term open access storage free of charge.

Effort for data management activities - particularly for WP2 and to a lesser extent WP4, WP5, WP6 – has been planned for and incorporated with other activities.

- Effort budget allocated to Data Management in WP7 – approx. 2 person months.
- Budget allocated for Open Access publication – 5000 €.
- Cost of long-term storage – zero.
- Data Manager is Berk Gencer, Farplas.

6. Data Security

6.1 Project internal communication platform and file store

Secure storage. Project data is stored on Iconiq Innovation's Microsoft365 SharePoint. Statements on how SharePoint safeguards data on the cloud can be found on the Microsoft365 site. Iconiq's IT services are supplied by Nubis365 who are certified and approved to IT security certification schemes and standards ISO 27001:2013, Cyber Essentials, PCI, Security Standards Council, Silver Microsoft Partner Network.

In addition, data may be stored on partners internal systems before transferring to Iconiq SharePoint. Each partner has their own data security policies and procedures, all have provided statements and evidence of compliance with ISO 27001:2013 or equivalent security standards.

In particular Farplas in-house system for capturing the data from their injection moulding machines which typically captures about 50MB of data per machine per month. They have a Kiban system in place so that approved users from other partners can access the data directly if required. Farplas have a robust Information Security Management system with TiSAX certification and are currently completing paperwork for ISO 27001 certification (expected to be issued September 2021).

Data recovery. SharePoint Site Collection recycle bin and Second Stage Recycle Bin store content for up to 93 days during which time files can be restored if required. Microsoft backup SharePoint every 12 hours and requests can be made to restore data through Iconiq Global Admin. Office365 files have version history so previous versions can be restored for up to 30 days.
Transfer of sensitive data. Data is encrypted in transit and at rest using Microsoft365 OneDrive for Business and SharePoint.

6.2 Certified long-term repositories

Zenodo will be used for long-term data storage and open access. Data can be loaded and safely stored on Zenodo in a closed and restricted manner. Data can then be made open at a later date, e.g., after journal paper has been published or an embargo period elapsed. System is built on Invenio Software battle tested by CERN for 12 years. All data files 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. Data stored on Zenodo is not encrypted and closed access data could be viewed by Zenodo staff, so this repository won't be used for confidential information.

7. Ethical Aspects

Personal (contact) data will only be stored for the project partners on Iconiq SharePoint in a manner compliant with European and national law, GDPR.

No personal data will be collected for project purposes and therefore no personal data will be stored on Zenodo long term open access repository.

8. Other

None anticipated at present.

9. Conclusions

Preliminary Data Management Plan has been created and agreed between the consortium partners. Document will be kept live and updated when significant changes occur. The Zenodo repository will be used for long-term open access storage. As this was designed for the purpose of sharing EC project data, through the Open AIRE project, is well suited to the project requirements and FAIR data principals.

10. Bibliography

- EC Guidelines on FAIR Data Management in H2020, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Zonodo open data repository, https://zenodo.org/
- Zonodo FAIR principles, https://about.zenodo.org/principles/
- OpenAIRE, https://www.openaire.eu/
- How SharePoint and OneDrive safeguard your data in the cloud, 05/25/2018, https://docs.microsoft.com/en-us/sharepoint/safeguarding-your-data
- Data Encryption in OneDrive for Business and SharePoint Online, 07/02/2018, https://docs.microsoft.com/en-us/microsoft-365/compliance/data-encryption-in-odb-and-spo?view=o365-worldwide

11. Annex I – (Expected) Datasets

Dataset Name	WP	Description	Format	Lead	Class	Size (MB)*	Comments	Repository
D1.1 Specifications Material	WP1	Specifications for material design	PDF	Norner	CO	20		Deliverable Report at Project Sharepoint.
D1.2 Virgin SRPP Panel	WP1	Data from the development of SRPP for the REAR SEAT BACK PANEL from virgin PP	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
D1.3 SRPP PCR% Panel	WP1	Data from the development of SRPP with PCR content for the REAR SEAT BACK PANEL	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
D1.4 SRPO for IM	WP1	Data from the development of general purpose SRPO for injection molding	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
D1.5 Virgin PMMA B-P Cover	WP1	Data from the development of virgin PMMA for the B Pillar Cover	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data

D1.6 PMMA PCR% B-P Cover	WP1	Data from the development of virgin PMMA for the B Pillar Cover	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
D1.7 TPC Crash Box	WP1	Thermoplastic composite for the Crash Box	PDF, Other	Norner	CO	40+500	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
D1.8 PP PCR% Bumper	WP1	PP with PCR content for the Lower Rear Bumper	PDF, Other	Norner	CO	40	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint Associated datasets at IT system of the partner generating the data
Predictive process model	WP2	Pre-trained ANN, test results and metrics	Python object/library JSON, CSV	IDENER	CO	1500	The deliverable is confidential but selected data sets and result may be made public on eventual Technical, Scientific Publications or Patent Applications	Deliverable Report at Project SharePoint Associated datasets at IT system of the partner generating the data
Predictive model control algorithm	WP2	Configuration data and parameters	CSV, JSON	IDENER	CO	100	The deliverable is confidential but selected data sets and result may be made public on eventual Technical, Scientific Publications or Patent Applications	Deliverable Report at Project SharePoint. Associated datasets at IT system of the partner generating the data
Predictive model control system demonstration data	WP2	Input/Output data for training and validation, test and validation results, metrics	Python object/library; CSV, JSON	Farplas	CO	~~500 GB	The deliverable is confidential but selected data sets and result may be made public on eventual Technical, Scientific Publications or Patent Applications	Deliverable Report at Project SharePoint. Associated datasets at IT system of the partner generating the data

Injection moulding process data	WP2	Data from Injection machines and sensors (Energy analyser and water collector)	CSV Parquet	Farplas	CO	50(per Injection machine / month) +500(EA) + 500(WC)	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Sharepoint and Kibana interface
D3.1 Final report detailing process specification	WP3	A review of WP1: are all the specifications in accordance with D1.1 A report on all activities from WP1 to WP3	PPT (initial) PDF (final)	Fraunhofer IAP	CO	15	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project Sharepoint (Final)
D3.2 Training of neural network for material models	WP3	A material model relating compositions of newly developed materials (compounded) to their properties (but will only add data for those compositions for which you could also provide the data from D3.4)	Other	Fraunhofer IAP	CO	50	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Associated datasets at IT system of the partner generating the data OR A report
D3.3 Optimized material models	WP3	A report on recyclability and environmental parameters	PDF	Fraunhofer IAP	CO	25 (or "0" in case of no improvement to model curve)	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Deliverable Report at Project (Sharepoint)
D3.4 optimized Component design	WP3	A presentation of the outcome of topology optimization on various automotive components	PPT	Fraunhofer IAP	CO	15	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	A Presentation at Project Sharepoint

D3.5 optimization of production strategies	WP3	A presentation on the manufacturing simulation of various developed automotive components	PPT	Fraunhofer IAP	CO	15	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	A Presentation at Project Sharepoint
D3.7 model and simulation report	WP3	A report detailing the optimization procedure and simulation results of the finalized component	PDF	Fraunhofer IAP	CO	25	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	A report describing all necessary steps of D3.5 and 3.6 at sharepoint
D3.6 Scale up specification targes	WP3	Details of the trial / prototype mould to be used in validation of aesthetic, paintability, thermal stability and scratch resistance by farplas (in T1.4.2). + Details of the actual geometry of the components being optimized	PDF	Fraunhofer IAP	CO	30	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Cad-files / PDF drawings (on Project - sharepoint)
Sensor integration platform	WP4	Datasheets of electronic components, electronic schematics and designs, electronic measurements, list of fabrication steps, report	PDF, XLSX, DXF, GERBER FILES. PPTX	IMEC	CO	1000	The information is confidential. Some information will be made public in scientific publications, patents	Project sharepoint and partners IT system
Test & Validation of demo parts	WP4	A report related material characterizations and part validations	PDF	TOFAŞ	CO	30	The deliverables are confidential but specific selected datasets may be public in paper, patent applications, eventual Technical	PDF in Revolution Project Share Point

							and Scientific Publications	
EV range analysis	WP4	Virtual range analysis report focusing on potential range extension with respect to extend of light weighting to additional components besides crash box.	PDF	TOFAŞ	CO	10	The deliverables are confidential but specific selected datasets may be public in paper, patent applications, eventual Technical and Scientific Publications	PDF in Revolution Project Share Point
Environmental data	WP5	Mass and energy balances and auxiliaries of the selected systems and their corresponding alternatives. Environmental and climate-related indicators	CSV, Excel, PDF	IDENER	PU	2000	The final deliverable is public. Selected data sets may be made public on eventual Technical, Scientific Publications or Patent Applications	Data provided by partners. Intermediate results will be presented to the consortium. Deliverable Report at Project SharePoint (Final)
Social & societal data	WP5	Indicators and associated parameters	CSV, Excel, PDF	IDENER	PU	100	The final deliverable is public. Selected data sets may be made public on eventual Technical, Scientific publications or patent applications	Intermediate results will be presented to the consortium. Deliverable Report at Project SharePoint
TEA data	WP6	Mass and Energy balances and auxiliaries of selected systems and corresponding alternatives, costs, economic and financial indicators	CSV, Excel, PDF	IDENER	CO	2000	The deliverable is confidential but selected datasets may be made public on eventual Technical, Scientific Publications or Patent Applications.	Data provided by partners. Intermediate results will be presented to the consortium. Deliverable Report at Project SharePoint (Final)
Network Patent Analysis	WP6	Patents citation analysis, clusters of patents, valuation of foreground IP	Excel, PDF	IDENER	CO	300		Deliverable Report at Project SharePoint (Final)
Dissemination materials	WP6	Video, graphics, flyers / information leaflets, public general	TIFF, MPEG, PDF	IDENER	PU	500		Project SharePoint, project website.

		interest reports, marketing material						
Social media	WP6	Post and articles in social media such as LinkedIn or Twitter	PPTX, DOCX	IDENER	CO	100		Continuously shared with consortium for approval before final publication.
General project reporting – CO	(WP 7)	Deliverables, periodic reports	PDF	Farplas, Iconiq	CO	500		Project SharePoint.
General project reporting - PU	(WP 7)	Public deliverables, dissemination records	PDF, Excel	Farplas, Iconiq	PU	500		Project SharePoint, Project Website.