Funded by the European Union

ACTPAC	
Project number:	101135289
Project name:	A Complete Transformation PAth for C-C backboned plastic wastes to high-value Chemicals and materials
Topic:	HORIZON-CL6-2023-ZEROPOLLUTION-01-5
Type of action:	HORIZON-IA
Starting date of action:	1 January 2024
Project duration:	48 months
Project end date:	31.12.2027
Deliverable number:	D13.9
Deliverable title:	Data management plan (initial)
Document version:	Ver1
WP number:	WP13
Lead beneficiary:	01-AU
Main author(s):	Zheng Guo (01-AU), Tina Kousted (01-AU)
Internal reviewers:	Celine Rabe (09-M&S), Damir Haskovic (09-M&S)
Nature of deliverable:	DMP
Dissemination level:	PU
Delivery date from Annex 1:	30-06-2024 (M6)
Actual delivery date:	30-06-2024 (M6)

Document history

Version	Date	Beneficiary	Description
0.0	15.05.2024	ALL PARTNERS	Filling questionary of the DMP of ACTPAC
0.1	05.06.2024	AU	First draft sent out for comments, AU WP leader
0.2	15.06.2024	M&S	Internal Review WP Co-Leader
0.3	20.06.2024	AU	Correction, Second version send out
0.4	25.06.2024	M&S	Second Internal Review WP Co-Leader
1.0	28.06.2024	AU	Submit

Executive Summary

This Data Management Plan (DMP) aims to set a general procedure and framework for the management of the data generated in the ACTPAC project. The DMP covers the actions from data acquisition to data curation and dissemination during the lifetime of the project and after the project is completed.

More specifically, the plan describes what kind of data is collected, generated and processed, what methodologies and standards are used, whether and how this data is shared and/or made publicly available, and how it is curated and preserved.

This DMP is in line with the agreed items and rules described in the Grant Agreement of the ACTPAC project (Grant agreement ID: 101135289) within Chapter 4 Grant Implementation, Section 2 Rules for carrying out the action, Article 15 — Data Protection, Article 16 — Intellectual Property Rights (IPR) — Background and Results — Access rights and rights of use and Article 17 — Communication, Dissemination And Visibility.

This DMP is in line with the FAIR principles as required by the European Commission and will be deposited in a public repository to allow full access.

The DMP describes how the generated data will be managed and made openly accessible throughout the lifetime of the project.

The ACTPAC DMP is a living document that will be continuously updated as it evolves throughout the lifetime of the project. A new version, updated with project outputs as the project evolves (new data sets, new publications, changes in data access or curation policies, etc.), will be submitted in M24 as D13.10 and M47 as D13.11.

Contents

Figures

Tables

Acronyms & Abbreviations

Term	Description
D	Deliverable
DMP	Data Management Plan
EC	European Commission
GA	General Assembly
GDPR	General Data Protection Regulation
PE	Polyethylene
WP	Work Package

1 Data Summary

This DMP is based on the template of "EU Grants: Data management plan (HE): V1.1 – 01.04.2022" [1]. It details what data the project will generate and manage, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. The signed Consortium Agreement by all beneficiaries of the ACTPAC project defines the ownership of key prior and foreground knowledge (IPR, data etc.) of all involved parties.

The ACTPAC project consists of 11 partners from 8 European or associated countries, including Denmark, Netherlands, Germany, France, Spain, Belgium, Austria and Turkey. Based on the nature and specificity of the research tasks, the ACTPAC project is split into 13 work packages (WPs), and the data management will be categorized according to the WPs and their connections.

The main objective is to establish a complete value-added industry-viable path to convert PE firstly into alkanes; then into high-value chemicals (monomers); and finally into PE-like but fully biodegradable polyesters. Except for the plan and documents associated with Environmental, social and economic studies (WP12), and Management, dissemination, exploitation and communication (WP13), most of the data are generated from lab experiments or pilot scale-up tests. The data belong to or are associated with new scientific findings, new processes, new technology, new catalysts or materials, new products and specifications, developed production lines and/or device prototypes.

The deliverable types within this project include R — Document, report, DEM — Demonstrator, pilot, prototype, DEC —Websites, patent filings, videos, etc, or DMP — Data Management Plan. The dissemination level is set at SEN - Sensitive or PU – Public.

The ACTPAC consortium will store all received data on the AU server, and measures will be taken to make it possible for third parties to access, mine, exploit, reproduce and disseminate the data including

- data gathered during the project, after the anonymization process and including associated metadata, as specified in the DMP;
- data produced during the project, including associated metadata, as specified in the Consortium Agreement and in the DMP.

A Microsoft Teams [2] account for the project has been created by AU to compile the ACTPAC data, including 1) public project reports and public deliverables; 2) all dissemination-related material; 3) consent forms signed by participants and 4) an updated live document of record keeping of the data collected.

The Teams account is administered by the AU team of the ACTPAC coordinator. The account is only accessible by invitation issued by the account administrators. All beneficiaries have been invited to join the Teams site and have been granted access to non-sensitive data or documents. A subset of the beneficiaries has been granted access to closed channels within the Teams account, which are intended for more sensitive data and documents, whose . privions will be in compliance with EU General Data Protection Regulation (GDPR). Anyone who gains access to these closed channels is fully controlled by the administrators of the Teams account and and the secure code is set and their access must be simultaneously approved by the controller, the coordinator and the data generator upon request. Concerning the metadata transfer to AU server, a safe online transfer such as encryption will be used.

The data will be presented in the form of spreadsheets (ODS, XLS or CSV), documentation (DOC, PDF, TXT or HTML), structure data (XML or JSON), image (JPG, TIFF, PDF or PNG) or video (WEBM, MP4, or MKZ etc).

The ACTPAC consortium will make every effort through the project's dissemination activities and media channels to make the data discoverable, accessible, intelligible, and usable by all interested stakeholders (including other EU related projects). A searchable index and an information classification system of all relevant datasets will be developed, applicable to all cases and data types, under the condition that all relevant IPR and copyright requirements pertaining to these datasets are protected or met.

1.1 Project briefing

The main objective is to establish a complete value-added industry-viable path to convert PE firstly into alkanes; then into high-value chemicals (monomers); and finally into PE-like but fully biodegradable polyesters.

Beyond the state-of-the art technologies, ACTPAC will design and deploy new catalysts and cross-metathesis modes for highly active and selective metathesis of PE into linear alkanes with a narrow distribution range (C6-C18, >90%). Two separate systems: a multi-enzyme machinery assembled in the recombinant cells, and a metabolic engineered yeast system, dedicated to the transformation of alkanes into monomers, will be developed. Monomers of diversified chain-lengths will be used for the synthesis of polyesters presenting different properties and polymer performances, assignable for various applications. A zero-waste solution to the plastic waste management is thus created to keep plastics out of the environment, and reclaim their values. The new properties and specific applications of the new polyester plastics produced from upcycling of PE waste will offer new business opportunities for SMEs by scalable, flexible and robust multi-product manufacturing processes for on-demand and small-volume output production.

To facilitate the data management for the ACTPAC project, the related information and key settings are listed in the following tables.

Table 1. Beneficiary list of the ACTPAC consortium

Participant No. *	Participant organization name	Country
1 (Coordinator)	Aarhus University, AU	Denmark
2	Utrecht University, UU	Netherlands
3	University of Münster, UM	Germany
4	Centre National de la Recherche Scientifique, CNRS	France
5	University of Groningen, UG	Netherlands
6	AIMPLAS	Spain
7	Biolynx	Belgium
8	Innovaplast	Turkiye
9	Minds & Sparks, M&S	Austria
10	B4PLASTICS	Belgium
11	CTCR	Spain

Table 2. ACTPACT WPs and Lead Beneficiaries

WP no.	Lead beneficiary
WP1+2	02-UU
WP3	01-AU
WP4+5	03-UM
WP6+7	04-CNRS
WP8+9	05-UG
WP10+11	01-AU
WP12	06-AIMPLAS
WP13	01-AU & 09-M&S

Table 3. List of the datasets in ACTPAC

| Dataset no.	Dataset title
Dataset 1	Catalytic metathesis of PE into C6-C18 alkanes
Dataset 2	Biological conversion of alkanes into diols/diacids
Dataset 3	Enzymatic conversion of alkanes into diols/diacids
Dataset 4	Chemical polymerisation for production of polyesters
Dataset 5	Chemo-Enzymatic polymerisation for production of polyesters
Dataset 6	Scale-up of catalytic metathesis of PE into C6-C18 alkanes
Dataset 7	Scale-up production of monomers
Dataset 8	Scale-up production of polymers
Dataset 9	Life cycle assessment and Techno-economic assessment
Dataset 10	Communication, dissemination and exploitation

Table 4. List of the data's relation to the project objectives of the ACTPAC project

Objective no.	Dataset title
Objective 1	Catalytic metathesis of PE into C6-C18 alkanes
Objective 2	Biological conversion of alkanes into diols/diacids
Objective 3	Enzymatic conversion of alkanes into diols/diacids
Objective 4	Chemical polymerisation for production of polyesters
Objective 5	Chemo-Enzymatic polymerisation for production of polyesters
Objective 6	Scale-up of catalytic metathesis of PE into C6-C18 alkanes
Objective 7	Scale-up production of monomers
Objective 8	Scale-up production of polymers
Objective 9	Life cycle assessment and Techno-economic assessment
Objective 10	Communication, dissemination and exploitation
Objective 11	Others (Pls specify)

Table 5. List of the data's relation with the result groups (RGs) of the ACTPAC project.

RG no.	Dataset title
RG 1	New catalysts and new chemical processing system of plastic wastes
RG 2	New biological transformation system
RG 3	New chemo/bio polymerisation catalysts and reaction system
RG 4	New industrial phenotype of monomer production
RG 5	New industrial phenotype of polymer production
RG 6	Science and citizen educational materials
RG 7	Societal/environmental impact and consequence
RG 8	Others (Pls specify)

1.2 Data reuse and generation
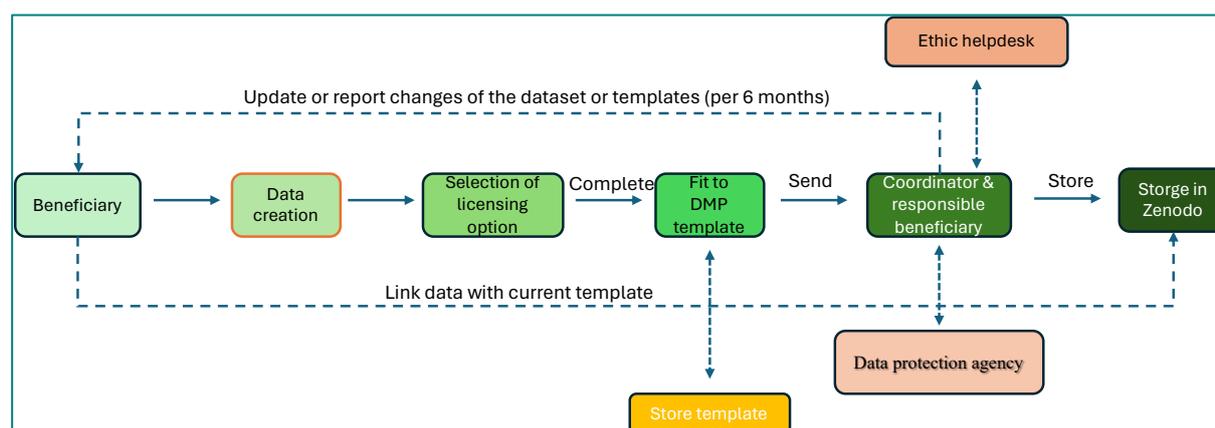
Figure 1. Procedures of data collection, updating, management and storage for the ACTPAC project

** **ZENODO** is a **CERN** Data Centre-backed research data repository for the long-tail of science, enabling researchers to preserve and share their research output from any science, regardless of the size and format.* [11]

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

The researchers of each beneficiary will decide what kind of data, materials or existing knowledge will be used, as the starting point of their research tasks described in the ACTPAC Grant Agreement. Most likely, the empirical and some approaches or methodology generated in the 5 research units, could be used for the technology development (WPs 1-9).

The existing methodology or approaches for Environmental, social and economic studies (WP12) and Management, dissemination, exploitation and communication (WP13) could be used as a basis to redesign and rebuild for the assigned new exploitable tasks in ACTPAC. The existing data as well as the knowledge of scale-up operation of previous studies from the industrial partners (WPs 10 & 11) may

be used for reactor redesign and optimization of variables. The relevant data has been agreed upon and listed as "background" in the appendix of the Consortium Agreement for each partner.

What types and formats of data will the project generate or re-use?

The deliverable types include R — Document, report, DEM — Demonstrator, pilot, prototype, DEC — Websites, patent filings, videos, etc, or DMP — Data Management Plan. The dissemination level is set at SEN - Sensitive or PU – Public.

The data will be presented in the form of either spreadsheets (ODS, XLS or CSV), documentation (DOC, PDF, TXT or HTML), structure data (XML or JSON), image (JPG, TIFF, PDF or PNG) or video (WEBM, MP4, or MKZ etc).

Table 6. Types and formats of data that ACTPAC will generate/collect

Data source	Data type	Data format [docx, txt, xls, etc.]	Data origin [WP]	Storage time [yr]
Direct measurement	R and DEC	Docx, pdf, CSV, xls, TIFF etc	1 & 2	During and 20 years after project
Direct measurement	R and DEC	Docx, pdf, CSV, xls, TIFF etc	3	During and 20 years after project
Direct measurement	R and DEC	Docx, pdf, CSV, xls, TIFF etc	4 & 5	During and 20 years after project
Direct measurement	R and DEC	Docx, pdf, CSV, xls, TIFF etc	6 & 7	During and 20 years after project
Direct measurement	R and DEC	Docx, pdf, CSV, xls, TIFF etc	8 & 9	During and 20 years after project
Direct measurement and simulation	R, DEM and DEC	Docx, pdf, CSV, xls, TIFF, and WEBM etc	10 & 11	During and 20 years after project
Model output and Simulation	R	Docx, pdf, xls, TIFF etc	12	During and 20 years after project
Survey and expert knowledge	R	Docx, pdf, xls, TIFF etc	13	During and 20 years after project

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

Aligned with the objectives of the ACTPAC project, the project will generate data concerning the new scientific findings, new technology and methodology, new catalysts, new value-added products and materials, new production lines and reaction/processing devices using waste PE as the starting feedstocks; thus create a completely new transformation path to upcycle PE waste.

The purpose of the data generation or re-use is to foster fast and extensive uptake of industrially relevant, cross-sectorial, cost-effective innovative technologies, as well as inspire relevant R&D, and innovation activities concerning upcycling and upvaluing of C-C backboned plastic wastes across

Europe. Thus, the positive impact of the ACTPAC outcomes on economical, industrial, and environmental goals can be maximized.

Not limited to the generation of new knowledge and technology, the new products and sideproducts from biological systems (WPs 3-5) and their unmatched properties, could be extended for other applications e.g., in the food, cosmetic or pharmaceutical sectors; speed up their uptake, and contribute to the green transition and environmental sustainability.

What is the expected size of the data that you intend to generate or re-use?

The expected size of data that will be generated or re-used is in a range of a few hundred megabytes to a few gigabytes.

What is the origin/provenance of the data, either generated or re-used?

Concerning the data origin/provenance, except for the plan and documents associated with Environmental, social and economic studies (Sustainability, circularity, TEA and LCA data) (WP12), and Management, dissemination, exploitation and communication (Business model, workshops and industry dialogue and demonstration, etc.) (WP13), most of the data are generated from lab experiments or pilot scale-up tests. The data belong to or are associated with new scientific findings, new processes, new technology, new catalysts or materials, new products and specifications, developed production lines and/or device prototypes.

To whom might your data be useful ('data utility'), outside your project?

The data is primarily targeted at all stakeholders related to recycling and upcycling of plastics wastes, including universities, research institutes, NGOs/NPOs, SMEs, civil society, policy makers and local administrations.

2   FAIR data

The Council of the European Union emphasises that "the opportunities for the optimal reuse of research data can only be realised if data are consistent with the FAIR principles (findable, accessible, interoperable and re-usable) within a secure and trustworthy environment"[3]. The FAIR principles are mentioned in the Communication "European Data Strategy (2020)" [4] by the European Commission as a way to implement interoperability.

Firstly, the ACTPAC consortium will store all data on the AU server. Later, the ACTPAC data will be deposited in the trusted European repository Zenodo, hosted by CERN, an intergovernmental organization. These measures will be taken to allow third parties to access, mine, exploit, reproduce and disseminate the following:

- Data gathered during the project, after the anonymization process and including associated metadata, as specified in the DMP.
- Data produced during the project, including associated metadata, as specified in the Consortium Agreement and in the DMP.

2.1 Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

A persistent identifier (PID) is a long-lasting reference to a digital resource and provides the information required to reliably identify, verify and locate your research data eliminating many misunderstandings. A PID may also be connected to a set of metadata which describes a digital resource.

For ACTPAC project, the scientific publications will use the Digital Object Identifier (DOI) for public dissemination.

Other than DOI, the ACTPAC consortium, will define its own ontology for data identification. The ontology is a formal naming and definition of the types, properties, and interrelationships of the entities (such as labs and WP to produce data).

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed?

ACTPAC will follow the FAIR principles [5], particularly in line with

- F2: Data are described with rich metadata [6]
- F3: Metadata clearly and explicitly include the identifier of the data they describe [7]
- F4: (Meta)data are registered or indexed in a searchable ressource [8]
- A1: (Meta)data are retrievable by their identifier using a standardised communication protocol [9]

The ACTPAC consortium will contact Zenodo. Data are described with rich metadata. Zenodo's metadata is compliant with DataCite's Metadata Schema [10] minimum and recommended terms, with a few additional enrichements.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Yes, keywords will be provided in the metadata for an easy discovery and potential reuse. The DOI is a top-level and a mandatory field in the metadata of each record.

Will metadata be offered in such a way that it can be harvested and indexed?

Yes. (Meta)data will be registered or indexed in a searchable ressource. Metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing.

Metadata of each record is sent to DataCite servers during DOI registration and indexed there.

2.2 Making data accessible

2.2.1 Repository

Will the data be deposited in a trusted repository?

The data will be deposited in the trusted European repository Zenodo.

"ZENODO is a CERN Data Centre-backed research data repository for the long-tail of science, enabling researchers to preserve and share their research output from any science, regardless of the size and format."[11] CERN's servers are stored in Europe, ensuring the respect of European laws, for instance with regards to GDPR (General Data Protection Regulation).

Zenodo hosts the data from the ACTPAC project, which allows to safely share and store the consortium's data and manage its visibility.

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

"ZENODO is an innovative and easy to use web-platform, which allows for upload, curation and sharing of the research data through an easy-to-use web interface and integration with other collaboration and data sharing services." [11]

For an effective data management, based on the data nature, WPs and sensitivity, appropriate arrangement with the identified repository will be made. The access rules will be set based on data sensitivity and granted aceesss by the ACTPAC data access committee. The coordinator AU will be responsible for account careation and uploading and moderation etc.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

"ZENODO ensures the discovery and citability of the research output by assigning a Digital Object Identifier (DOI) to every upload, as well as promotes software citation and preservation through one-click integration with GitHub." [11]

2.2.2 Data

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

Based on the dissemination level set, most of, but not all the data is openly available. Due to the sensitivity, IPR and commercial potential, access restrictions will be applied for some data pertaining to specific beneficiaries and their interests.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

The ACTPAC consortium will make research data accessible as soon as possible. However, if the patentable research data is filed as a pending patent, the consortium will follow the IPR legal restriction. The embargo periods should last no more than 2 years after the project is completed.

Will the data be accessible through a free and standardized access protocol?

Yes. The data will be assigned a globally unique and persistent identifier when it is deposited in Zenodo. The protocol is open, free, and universally implementable.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

For those data that have access restrictions, AU, as the coordinator of the ACTPAC project, will play the key role in ensuring secrecy from outside of the consortium or anyone who pose any conflict of interest. The coordinator then collects the data directly from the concerned parties, centralizes it and protects the files with a secret code, before distributing it to the targeted people. The process avoids direct communication between the partners and thus, any breach of confidentiality.

For possible access to the confidential data, the coordinator will be the first person to contact. The coordinator will contact the data owner and discuss the operability of the data access.

How will the identity of the person accessing the data be ascertained?

The protocol allows for an authentication and authorization procedure, when necessary.

Metadata are publicly accessible and licensed under a public domain. No authorization is ever necessary to retrieve it.

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

As per the Grant Agreement, ACTPAC scientific publications will be made fully accessible. However, the data marked as dissemination level Sensitive are excluded.

A data access committee to evaluate/ approve access requests to personal/sensitive data will be established by the end of the first project year after approval of the General Assembly (GA) of ACTPAC.

2.2.3 Metadata

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

All metadata in Zenodo may be freely used under the CC0 waiver. Zenodo registers DOIs for all uploads. Zenodo allows including information on alternate persistent identifiers, as well as linking to related persistent identifiers.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Metadata will remain accessible, even when the data are no longer available. Data and metadata 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.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

Zenodo sets its own Zenodo code, which is itself open source and is built on the foundation of the Invenio digital library which is also open source. The work-in-progress, open issues, and roadmap are shared openly in GitHub [12], and contributions to any aspect are welcomed from anyone. [13]

2.3 Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

Zenodo provides the following fucntions to make its stored data interoperable:

- **I1**: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
    - Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML.
- **I2**: (meta)data use vocabularies that follow FAIR principles
    - For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).
- **I3**: (meta)data include qualified references to other (meta)data
    - Each referenced external piece of metadata is qualified by a resolvable URL.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

In case this is needed, the ACTPAC consortium, will define its own ontology for data identification. The ontology is a formal naming and definition of the types, properties, and interrelationships of the entities (such as labs and WP to produce data).

Yes, the ACTPAC consortium may openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them.

Will your data include qualified references[1] to other data (e.g. other data from your project, or datasets from previous research)?

Yes, we will follow the Zenodo principle I3.

I3: (meta)data include qualified references to other (meta)data

- ▪ Each referenced external piece of metadata is qualified by a resolvable URL.

2.4 Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

To make the deposited data be reusable, Zenodo set convenient tools *to validate data analysis and facilitate data re-use.*[14]

R1: (meta)data are richly described with a plurality of accurate and relevant attributes

- ▪ Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Yes, Zenodo has set up [14]

R1.1: (meta)data are released with a clear and accessible data usage license

- ▪ License is one of the mandatory terms in Zenodo's metadata, and is referring to an <u>Open Definition</u> license.
- ▪ Data downloaded by the users is subject to the license specified in the metadata by the uploader.

Will the data produced in the project be useable by third parties, in particular after the end of the project?

The non-sensitive and non-confidential data produced in the project will be useable by third parties during and after the end of the project.

---

[1] *A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)*

Will the provenance of the data be thoroughly documented using the appropriate standards?

Yes, Zenodo set up the following to ensure the provenance of the data is documented using appropriate standards [14]:

R1.2: (meta)data are associated with detailed provenance

- All data and metadata uploaded is traceable to a registered Zenodo user.
- Metadata can optionally describe the original authors of the published work.

R1.3: (meta)data meet domain-relevant community standards

- Zenodo is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.

Describe all relevant data quality assurance processes.

A process for dissemination of the scientific publications, results, research outputs and any other data has been established in the ACTPAC consortium to ensure the quality and of information made accessible to the public. The process will secure the confidentiality status of the data and be presented in a comprehensive way to be easily re-usable by external parties. The process includes evaluating the confidentiality of the data to prevent any leak of sensitive information, then data will be reviewed by the coordinator and relevant beneficiaries to validate following criteria:

- A standard layout.
- Conformity with European Commission, Grant Agreement and Consortium Agreement rules.
- Inclusion of detailed data.
- High quality of delivered data.
- Readability and understandability of information.
- Inclusion of references, sources, etc.

In the case that beneficiaries do not meet these requirements, they will be asked to redo their datasets and re-submit the new version to the coordinator until the criteria are fulfilled. Following the quality check, the beneficiaries will have to abide by the publication policy in order to publish their content. The coordinator will perform a last quality check to ensure the set rules have been followed, which can lead to the open access publication of the data.
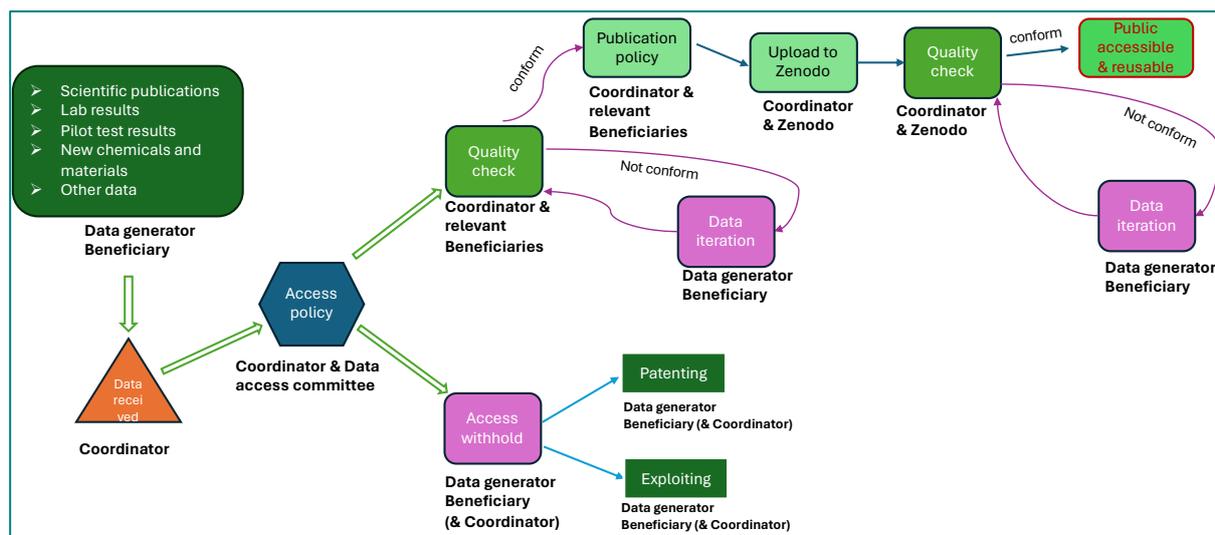
Figure 2. The procedure for the data quality checking and validation in ACTPAC

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

The management of other research outputs that will be planned and the related action will be described in the consoritium agreement of ACTPAC Digital Output Management Plan incorporated in the Dissemination & Communication Plan.

Other outputs such as algorithms, software etc, if any, will be decided by the inventors, together with the coordinator and Data Access Committee, for when and how to make them accessable publically.

The aspects related to the allocation of resources, data security and ethics will be carefully considered.

3   Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

The ACTPAC project is more experimental and delivers pilot-test results, thus workflows (flowchart or facility layout), protocols or models, and physical (products and materials) outcomes are foreseeable. A general procedure will be made by the executive body and proposed to the General Assembly. These kinds of outcomes will be managed by the coordinator and Data Access Committee.

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

In the ACTPAC consortium, each beneficiary will make their own judgement on the sensitivity and IPR necessity of the outcomes and their pertaining to FAIR data or not. They will cooperate with the coordinator and the data access committee to classify the data either as FAIR data moving for deposition in Zenodo or as the non-FAIR outcomes preserved in the ACTPAC dataset, but will be standardized and made them available for re-use on demand.

4 Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

Zenodo is partly funded by the European Commission and donations from other partners, therefore Zenodo's services are free of charge for ACTPAC. The fees for open access publication will be covered by each beneficiary.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

Other costs for the management of the non-FAIR data will be covered by the budget within the grant of ACTPAC project.

Who will be responsible for data management in your project?

The coordinator of ACTPAC, AU is responsible for the data management of the project, as part of WP13 focusing on project management, dissemination, communication and exploitation.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

For the data deposited in Zenodo, Zenodo supports long-term preservation of data deposits, as the repository is projected to be maintained for the lifetime of the host laboratory CERN, defined as at least the next twenty years (see Zenodo's Policies) [16].

For the non-FAIR data, the coordinator together with the beneficiaries will decide collectively and monitor the accessibility of data over time. After a certain time of exploitation, confidential results or data could be re-evaluated to make them publicly accessible. If costs arise, they will be discussed between the coordinator and the relevant beneficiaries to decide who will pay the costs or cost-splitting.

5 Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

According to Zenodo, they "take security very seriously and do [their] best to protect [the deposited] data [15].

- CERN Data Centre: Our data centre is located on CERN premises and all physical access is restricted to a limited number of staff with appropriate training and who have been granted access in line with their professional duties (e.g. Zenodo staff do not have physical access to the CERN Data Centre) .
- Servers: Our servers are managed according to the CERN Security Baseline for Servers, meaning e.g. remote access to our servers are restricted to Zenodo staff with appropriate training, and the operating system and installed applications are kept updated with latest security patches via our automatic configuration management system Puppet.
- Network: CERN Security Team runs both host and network based intrusion detection systems and monitors the traffic flow, pattern and contents into and out of CERN networks in order to detect attacks. All access to zenodo.org happens over HTTPS, except for static documentation pages which are hosted on GitHub Pages.
- Data: Zenodo stores user passwords using strong cryptographic password hashing algorithms (currently PBKDF2+SHA512). Users' access tokens to GitHub and ORCID are stored encrypted and can only be decrypted with the application's secret key.
- Application: We are employing a suite of techniques to protect your session from being stolen by an attacker when you are logged in and run vulnerability scans against the application.
- Staff: CERN staff with access to user data operate under <u>CERN Operational Circular no. 5</u>, meaning among other things that
  - staff should not exchange among themselves information acquired unless it is expressly required for the execution of their duties.
  - access to user data must always be consistent with the professional duties and only permitted for resolution of problems, detection of security issues, monitoring of resources and similar.
  - staff are liable for damage resulting from any infringement and can have access withdrawn and/or be subject to disciplinary or legal proceedings depending on seriousness of the infringement."

For the *sensitive data,* Zenodo allows closed access:

> "*Zenodo allows users to upload files under closed access. Closed access means that zenodo.org users will not be able to access the files you uploaded. The files are however stored unencrypted and **may** be viewed by Zenodo operational staff under specific conditions. This means that "closed access" on Zenodo **is not** suitable for secret or confidential data*."[15]

Will the data be safely stored in trusted repositories for long term preservation and curation?

Data will be safely stored in the world's largest general-purpose research repository, Zenodo hosted by CERN.

(Meta)data are accessible, and the metadata 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 at least the next 20 years.

6   Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

No ethical risk has been identified related to the ACTPAC's scope of actions. The coordinator, together with the executive board and the general assembly, will look out for any throughout the lifetime of the project to avoid ethical or legal issues.

The project will be carried out in compliance with Article 10, and Article 14.1, listed in the Grant Agreement of the ACTPAC: 1) ethical principles (including the highest standards of research integrity); 2) applicable EU, international and national law, including the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

Informed consent for data sharing and long-term preservation, following GDPR rules, will be included in questionnaires dealing with personal data.

7   Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

After the raw data is generated before transfer to AU, the beneficiary themselves will be responsible for the storage and security of the data.

The raw data will be firstly received by Aarhus University and stored on the AU server, before checking, validation, and finally being standardized and transferred to the deposition in Zenodo. AU will follow the Zenodo FAIR principles to assure the security and confidentiality of all received data.

References

[1]   Cohorts Coordination Board, 'Horizon Europe Data Management Plan Template', Cohort Coordination Board, Jan. 25, 2024. Available: https://cohortcoordinationboard.eu/toolkit/horizon-europe-data-management-plan-template/. [Accessed: Jun. 11, 2024]

[2]   Microsoft Teams. Available: https://www.microsoft.com/de-at/microsoft-teams/group-chat-software. [Accessed: Jun. 11, 2024]

[3]   Council of the European Union, 'Council conclusions on the transition towards an open science system'. May 27, 2016. Available: https://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf. [Accessed: Jun. 11, 2024]

[4]   European Commission, 'A European strategy for data', EUR-Lex.europa.eu, Feb. 19, 2020. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0066&from=EN. [Accessed: Jun. 11, 2024]

[5]   'FAIR Principles', GO FAIR. Available: https://www.go-fair.org/fair-principles/. [Accessed: Jun. 11, 2024]

[6]   'F2: Data are described with rich metadata', GO FAIR. Available: https://www.go-fair.org/fair-principles/f2-data-described-rich-metadata/. [Accessed: Jun. 11, 2024]

[7]   'F3: Metadata clearly and explicitly include the identifier of the data they describe', GO FAIR. Available: https://www.go-fair.org/fair-principles/f3-metadata-clearly-explicitly-include-identifier-data-describe/. [Accessed: Jun. 11, 2024]

[8]   'F4: (Meta)data are registered or indexed in a searchable resource', GO FAIR. Available: https://www.go-fair.org/fair-principles/f4-metadata-registered-indexed-searchable-resource/. [Accessed: Jun. 11, 2024]

[9]   'A1: (Meta)data are retrievable by their identifier using a standardised communication protocol', GO FAIR. Available: https://www.go-fair.org/fair-principles/metadata-retrievable-identifier-standardised-communication-protocol/. [Accessed: Jun. 11, 2024]

[10] 'DataCite Metadata Schema', DataCite Schema. Available: https://schema.datacite.org/. [Accessed: Jun. 11, 2024]

[11] K. Nowak, L. H. Nielsen, I. Pantopikos, and A. Themistoklis, 'Zenodo, a free and open platform for preserving and sharing research output.', May 24, 2016. doi: 10.5281/zenodo.51902. Available: https://zenodo.org/records/51902. [Accessed: Jun. 11, 2024]

[12] 'GitHub: Let's build from here', GitHub. Available: https://github.com/. [Accessed: Jun. 11, 2024]

[13] M. Potter and T. Smith, 'Making code citable with Zenodo and GitHub', Jul. 2015, doi: 10.5281/zenodo.45042. Available: https://zenodo.org/records/45042. [Accessed: Jun. 11, 2024]

[14] 'Principles', Zenodo | About. Available: https://about.zenodo.org/principles/. [Accessed: Jun. 11, 2024]

[15] 'Infrastructure', Zenodo | About. Available: https://about.zenodo.org/infrastructure/. [Accessed: Jun. 11, 2024]

[16] 'General Policies v1.0', Zenodo | About. Available: https://about.zenodo.org/policies/. [Accessed: Jun. 11, 2024]

[17] European Commission, 'Horizon Europe – Grants & reporting', European Commission. Available: https://rea.ec.europa.eu/horizon-europe-grants-reporting_en. [Accessed: Jun. 11, 2024]

[18] 'General Policies v1.0', Zenodo | About. Available: https://about.zenodo.org/policies/. [Accessed: Jun. 11, 2024]

[19] 'How to make your data FAIR', OpenAIRE, Jul. 26, 2018. Available: https://www.openaire.eu/how-to-make-your-data-fair/. [Accessed: Jun. 11, 2024]

Annex 1. Template dataset (Exampled by WP10)

ACTPAC Project_Data Management Questionnaire

This questionnaire is an essential component of the ACTPAC Data Management Plan deliverable, which is placed in the Annex of the report.

The purpose of the questionnaire is to categorise the nature and belongs of the specific WP dataset, which facilitates the data management and findability for the potential users of the data from the ACTPAC project.

The whole questionnaire consists of 11 sections. The questionnaire is designed as: short answer, description or multi-choice.

We hope the fulfilling of the questionnaire can be completed by May 15.

Email *

guo@bce.au.dk

Horizon Europe ACTPAC Project

List of the datasets
Catalytic metathesis of PE into C6-C18 alkanes
Biological conversion of alkanes into diols/diacids
Enzymatic conversion of alkanes into diols/diacids
Chemical polymerisation for production of polyesters
Chemo-Enzymatic polymerisation for production of polyesters
Scale-up of catalytic metathesis of PE into C6-C18 alkanes
Scale-up production of monomers
Scale-up production of polymers
Life cycle assessment and Techno-economic assessment
Communication, dissemination and exploitation

ACTPACT WP Data Summary

Work package No. *

10

Reference No.

-

Dataset selection *

Scale-up of catalytic metathesis of PE into C6-C18 alkanes
Scale-up production of monomers

Principal type of data contained in the data set *

- ☑ Quantitative
- ☑ Qualitative
- ☑ Numeric
- ☑ Text
- ☑ Images
- ☑ Audio
- ☑ Video
- ☑ Databases
- ☑ Non-structured data
- ☑ Source code
- ☑ Computational models
- ☐ Time series

Data description

Data will be both quantitative (parameters) and qualitative (product characterizations and property description). The documentation can be numbers including source code, condition settings etc, or text including operational journal, and protocols. The data can also be images such as reactor configuration and layout of the plant, or vedeo such as demonstration of operation procedure etc.

Data description *

Data will be both quantitative (parameters) and qualitative (product characterizations and property description). The documentation can be numbers including source code, condition settings etc, or text including operational journal, and protocols. The data can also be images such as reactor configuration and layout of the plant, or vedeo such as demonstration of operation procedure etc.

Methodology used to collect this data

Methodology used to collect this data *

The main methods are direct measurement. Some thermodynamic property of known molecules can also be used from other literature sources. Simulation data are generated from software or modellings etc. Videos are recorded in situ.

What is the source of the data? *

- [ ] Field work
- [x] Direct measurements
- [ ] Surveys
- [x] Simulations
- [x] Expert Knowledge
- [x] Model Output
- [ ] Other: ....

Relation to the data with the objectives of the project

Relation to the data with the objectives of the project *

- ☐ Catalytic metathesis of PE into C6-C18 alkanes
- ☐ Biological conversion of alkanes into diols/diacids
- ☐ Enzymatic conversion of alkanes into diols/diacids
- ☐ Chemical polymerisation for production of polyesters
- ☐ Chemo-Enzymatic polymerisation for production of polyesters
- ☑ Scale-up of catalytic metathesis of PE into C6-C18 alkanes
- ☑ Scale-up production of monomers
- ☐ Scale-up production of polymers
- ☐ Life cycle assessment and Techno-economic assessment
- ☐ Communication, dissemination and exploitation

Relation to the data with the results of the project *

- ☐ New catalysts and new chemical processing system of plastic wastes
- ☐ New biological transformation system
- ☑ New chemo/bio polymerisation catalysts and reaction system
- ☑ New industrial phenotype of monomer production
- ☐ New industrial phenotype of polymer production
- ☐ Science and citizen educational materials
- ☐ Societal/environmental impact and consequence

Why is this data collected?

The collection of this data will facilitate the fast access to the broad audience of the potential users and further commercial exploitation.

Please, provide a brief description of any external dataset used (For every external data set used please explain its origin, relevance and license.)

The existing data as well as the knowledge of scale-up operation of previous studies from the industrial partners (WPs 10 & 11) may be used for reactor redesign and optimization of variables. Most of the data in WP10 are generated from lab experiments or pilot scale-up tests. The data belong to or are associated with new scientific findings, new processes, new technology, new catalysts or materials, new products and specifications, developed production lines and/or device prototypes.

FAIR DATA (DATA FINDABLE, ACCESSIBLE, INTEROPERABLE AND REUSE)

DATA FINDABLE

Tim's 5-star classificationof the dataset *

https://5stardata.info/en/

- ☑ Data is available under an open license
- ☑ Use a structured data (e.g., Excel instead of image scan of a table)
- ☑ Is available in a non-proprietary open format (e.g., CSV as well as of Excel)
- ☑ Use URIs to denote things
- ☐ Option 5

Metadata standards (Please cite the standard and format use for the date. If any this data set does not follow any standardized format, please provide a formal specification. For example, the Dublin Core Metadata Initiative, Inspire Initiative, ISO, etc.)
https://atlan.com/what-is-metadata/?ref=/metadata-standards/

Documentation stored in the data *

- [x] Information of the origin of the data
- [ ] Codebook
- [x] List of abbreviations
- [x] Description of variables
- [x] Technical information about files

Ontologies

- [x] FIWARE, Link regarding to fiware https://www.fiware.org/data-models
- [ ] Other external ontology, Link regarding to ontologies, https://www.w3.org/wiki/Lists_of_ontologies

Whom might it be useful? *

- ☑ Public Administrations
- ☑ Research groups
- ☑ Citizens
- ☑ Private sector
- ☐ Others, please specify

Channels to reach potential users *

- ☑ Personal/research group web page
- ☑ well-known specialist database
- ☑ Search Administration database
- ☑ Email of corresponding author
- ☑ Data access statement in published articles
- ☑ Personal networking
- ☑ Citation of data sets

DATA ACCESSIBLE

Accessibility *

- ◉ Public data
- ○ Confidential data

Obligation or intention to publish/share data *

- ◉ Yes
- ○ No

When will the data be published? *

- ☐ Immediately on collection
- ☑ Within sometime after the ends of the project (please specify):
- ☑ Within sometime after its collection (please specify):
- ☑ To coincide with publication of main results
- ☐ Other (please specify):

Expected difficulties file sharing

- ☑ Confidentiality
- ☐ Large file size
- ☑ Ownership/licensing
- ☑ Intended commercialisation
- ☐ Other (please specify):

DATA INTEROPERABLE

File format

Spreadsheet:

- ☑ ODS
- ☑ XLS
- ☑ CSV

Documentation

- ☑ DOC
- ☑ PDF
- ☑ TXT
- ☑ HTML

Structured data

- ☑ XML
- ☑ JSON

Image:

- ☑ JPG
- ☑ PDF
- ☑ TIFF
- ☑ PNG

Video:

- [x] WEBM
- [x] MP4
- [ ] MKZ
- [ ] Others (Please specify)

Methods or software tools needed to access the data (Please detail any necessary *
software to manipulate the information (if not standard).)

Zenodo sets its own Zenodo code, which is itself open source and is built on the foundation of the Invenio digital library which is also open source. The work-in-progress, open issues, and roadmap are shared openly in GitHub, and contributions to any aspect are welcomed from anyone.

DATA REUSE

License conditions and restrictions *

- [x] Copyright
- [x] Creative Commons (please specify)
- [x] Open Licence
- [ ] Other (please specify):

Please, list the owners of the copyright and intellectual property involved (please, *
list)

UU, AU, UM, Biolynx and Innovaplast

Access permissions and restrictions (List roles/individuals (internal & external) with   *
any limitations to access (e.g. scope, actions permitted),
including who has authority to grant additional access.)

The ACTPAC access committee and the IPR owners

DATA MANAGEMENT AND ALLOCATION OF RESOURCES

Partners *

	Collection	Curation	Preservation
AU	☑	☑	☑
UU	☑	☑	☑
UM	☑	☑	☑
CNRS	☐	☐	☑
UG	☐	☐	☑
AIMPLAS	☐	☐	☑
BIOLYNX	☑	☑	☑
Innovaplast	☑	☑	☑
M&S	☐	☐	☑
B4PLASTICS	☐	☐	☑
CTCR	☐	☐	☑

What are the costs for making data FAIR in your project?

Costs mainly from the data collection and curation, and the storage of the data is free.

How will these be covered?

Covered individually or split share

Primary storage medium and location *

☑ University shared or research storage

☐ Secure facility from a data provider

☐ Physical storage

☐ Cloud platforms

☑ Last resort platforms (e.g. Zenodo)

☐ Academic research network platforms (e.g. ResearchGate).

☐ Institutional open data repositories (e.g. AU based)

☐ Other (please specify):

Data curation processes (Please briefly describe the management of data throughout its *
life cycle.)

A standard procedure for DMP in the ACTPAC project has been made, for the data quality checking
and validation in ACTPAC.

How will long-term preservation and access be assured? (Please briefly describe how *
the data will be preserver after the end of the project.)

Metadata will remain accessible, even when the data are no longer available. Data and metadata 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.

Regularity of backups and data performed. Replicas in other different places (if any)

The data will first store in AU server and then deposit in Zenodo

File management versioning

- ☐ Unnecessary (i.e. overwrite original file)
- ☐ Control version software (e.g. Git please specify)
- ☑ Date/version number in filename/folder
- ☐ Other (please specify):

ETHICAL AND LEGAL ASPECTS

Ethical aspect (If know) No/Yes, please briefly describe:

Aspects regarding informed consent in data collection and information protection in data storage and access.

No, No ethical risk has been identified related to WP10.

Legal aspect (If know) No/Yes, please briefly describe:

Yes, The project will be carried out in compliance with Article 10, and Article 14.1, listed in the Grant Agreement of the ACTPAC: applicable EU, international and national law, including the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

Other aspects (Description optional)

No

Annex 2. Collection of all responses to the ACTPAC Project Data Management Questionnaire

DATA SUMMARY

9 responses (Missing response from UG, WPs 8&9)

Who has responded?

Response No.	Email
1	mllacer@aimplas.es
2	bekireser@bce.au.dk
3	peruch@enscbp.fr
4	cramail@enscbp.fr
5	celine.rabe@mindsandsparks.org
6	jochen.schmid@uni-muenster.de
7	guo@bce.au.dk
8	guo@bce.au.dk (1)
9	s.saedy@uu.nl
10	k.u.loos@rug.nl

ACTPAC WP Data Summary

- 9 responses

Response No.	Work package No.
1	12
2	3
3	6
4	7
5	13
6	4-5
7	10
8	11

9	1-2
10	8-9

Dataset selection

- 9 responses

Response No.	Selected data
1	Chemical catalysis for the synthesis of polyesters
2	Life cycle assessment and Techno-economic assessment
3	Enzymatic conversion of alkanes into diols/diacids
4	M&S Collection & Reporting
5	WP4_yeast
6	Scale-up of catalytic metathesis of PE into C6-C18 alkanes Scale-up production of monomers
7	Scale-up production of polymers
8	ACTPAC WP1 and WP2
9	Chemo-Enzymatic polymerisation for production of polyesters

Principle type of data contained in the data set

Principal type of data contained in the data set
10 responses

Type	Count (%)
Quantitative	9 (90%)
Qualitative	7 (70%)
Numeric	6 (60%)
Text	8 (80%)
Images	9 (90%)
Audio	1 (10%)
Video	2 (20%)
Databases	3 (30%)
Non-structured data	3 (30%)
Source code	2 (20%)
Computational models	3 (30%)
Time series	0 (0%)
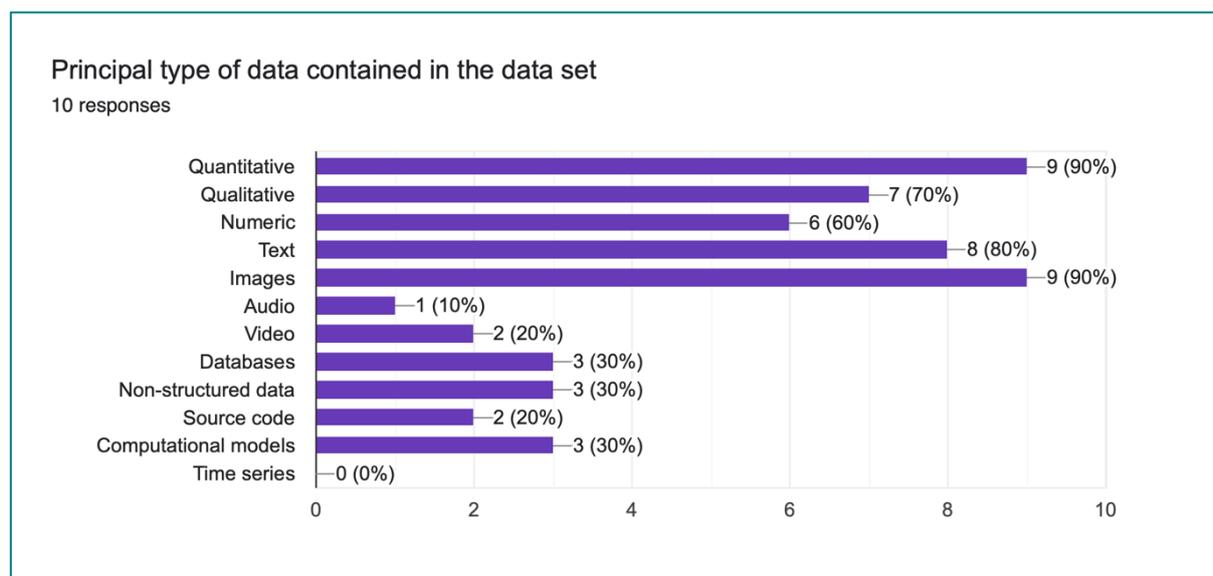
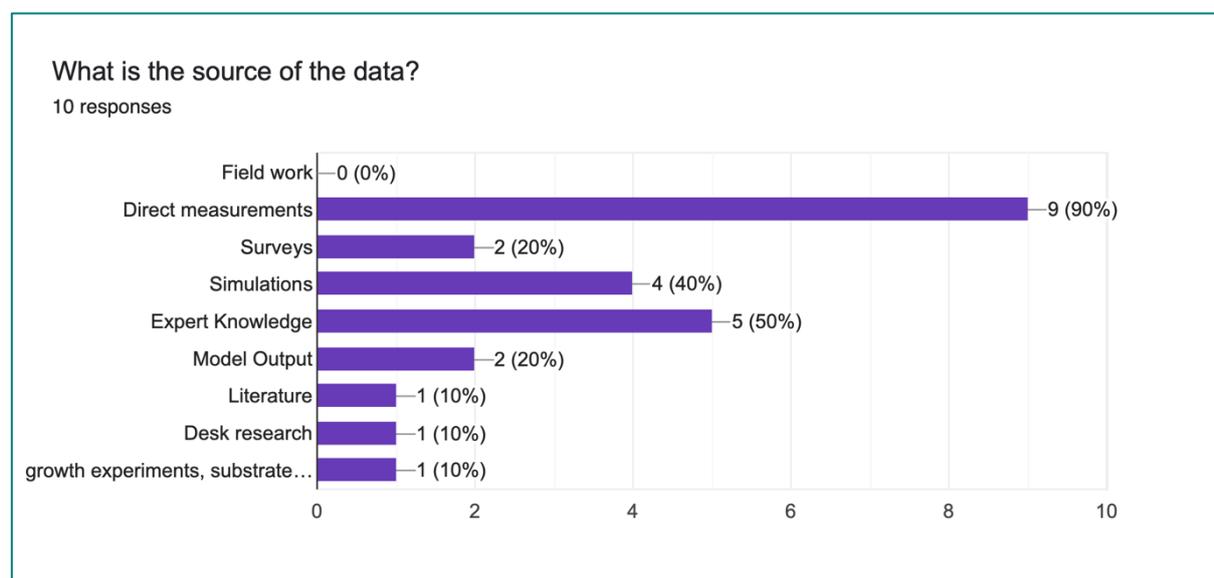## Data description

- 10 responses

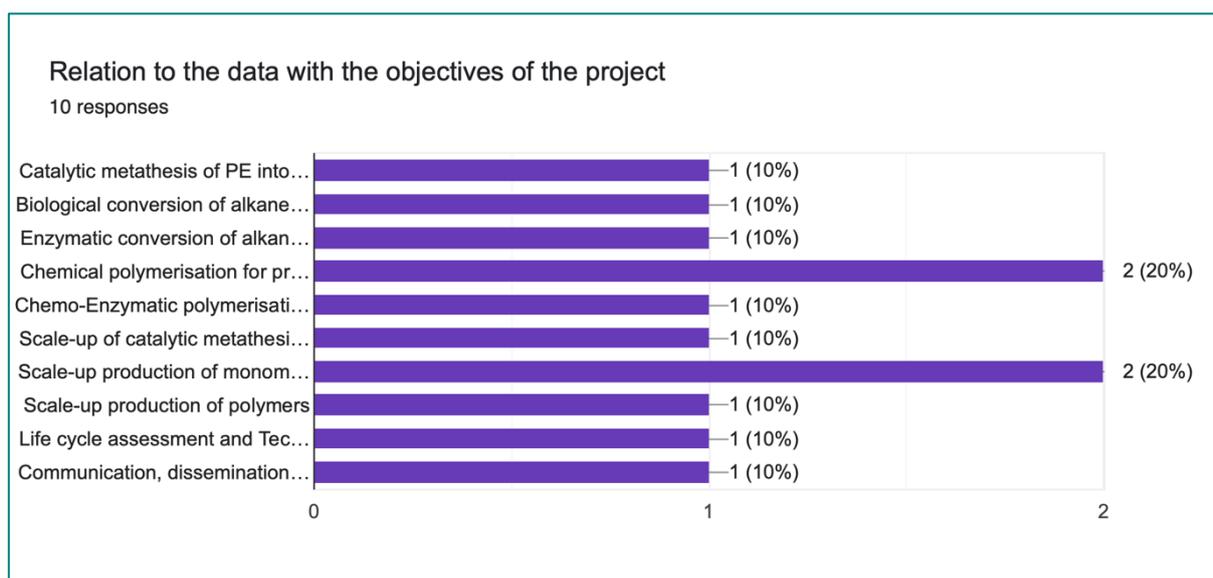| Response No.	Data description
1	Synthesis of polyesters
2	Input and description of the different processes involved in the production of the product (raw materials, energy and water consumption, waste, transport, working hours, machines used, processes...)
3	The data to be generated in WP3 includes; conditions for cell growth and protein expression, bacterial cell genotype, protein sequence, dna sequence, plasmid vector sequence, enzyme reaction kinetics data, optimum enzymatic reaction conditions, Gas chromatography data (with flame ionization detector and with mass detector), product yields from enzymatic reactions, substrate conversion levels, protein structure models (from computational analysis)
4	Collection of data needed for dissemination and exploitation management including publicly available stakeholder contacts, addresses etc.; texts and images for social media posts; information from internal questionnaires about project outputs, plans, contacts, activities, etc.
5	report
6	Data will be both quantitative (parameters) and qualitative (product characterizations and property description). The documentation can be numbers including source code, condition settings etc, or text including operational journal, and protocols. The data can also be images such as reactor configuration and layout of the plant, or vedeo such as demonstration of operation procedure etc.
7	In this WP, the main objective is to demonstrate scaling up of polymer (polyester) production. This WP will comprise of different technologies and methods to produce target polyesters from various chain length diols/diacids that were produced in WP6a, with the aim of reaching TRL 6-7. The data are related to the following deliverables: Deliverables D11.1 3 target polyesters produced in 300-liter reactor by chemical polymerization at TRL 6-7 with scalable protocol (M45) D11.2 Public report for scale-up production of 3 target polyesters by chemical polymerization at TRL 6-7 (M45) D11.3 Scale-up of 3 target polyester production by biocatalytic polymerization and demonstration of production at TRL 6 level (M46) D11.4 Public report for scale-up production of 3 target polyester by biocatalytic polymerization at TRL 6 level (M46)
8	ACTPAC WP1 and WP2 data
9	quantitative data from synthesis and characterization - qualitative data from synthesis and characterization - images from synthesis and characterization

Methodology used to collect this data

- 9 responses

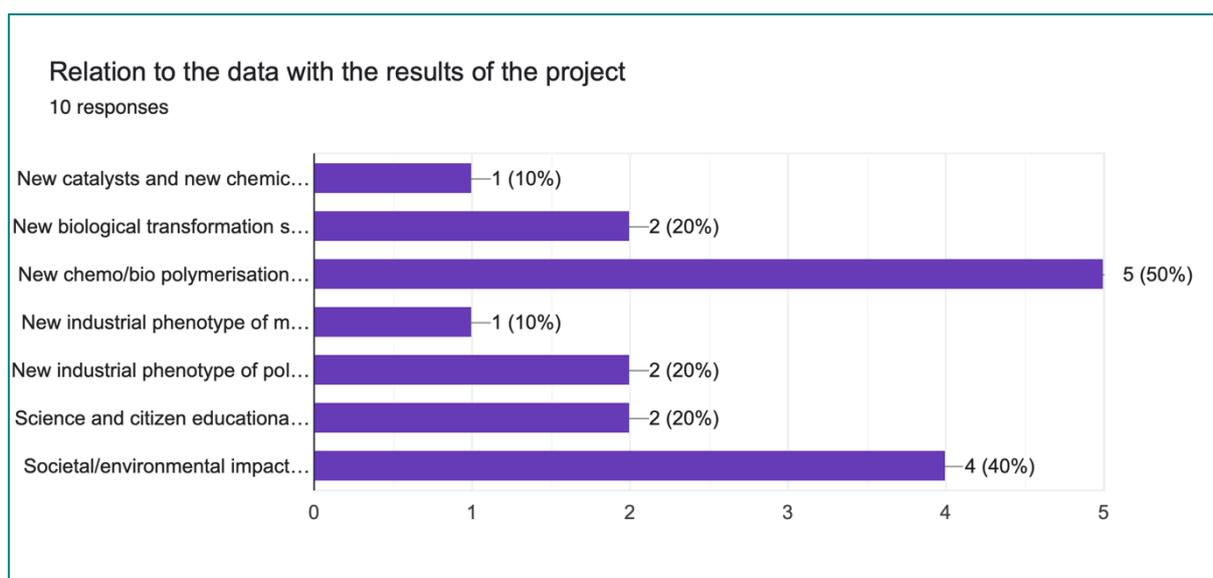| Response No.	Methodology used
1	experimental work
2	The main methods are direct measurement. Some thermodynamic property of known molecules can also be used from other literature sources. Simulation data are generated from software or modellings etc. Videos are recorded in situ.
3	Templates, excel or world
4	Experimental (lab experiments) and computational
5	Desk research, questionnaires
6	laboratory experiments, insilico design
7	experimental/imaging
8	lab journal and computer attached to equipment

What is the source of the data?

Relation to the data with the objectives of the project

Relation to the data with the objectives of the project
10 responses

Category	Value
Catalytic metathesis of PE into…	1 (10%)
Biological conversion of alkane…	1 (10%)
Enzymatic conversion of alkan…	1 (10%)
Chemical polymerisation for pr…	2 (20%)
Chemo-Enzymatic polymerisati…	1 (10%)
Scale-up of catalytic metathesi…	1 (10%)
Scale-up production of monom…	2 (20%)
Scale-up production of polymers	1 (10%)
Life cycle assessment and Tec…	1 (10%)
Communication, dissemination…	1 (10%)

Relation to the data with the results of the project

Relation to the data with the results of the project
10 responses

Category	Value
New catalysts and new chemic…	1 (10%)
New biological transformation s…	2 (20%)
New chemo/bio polymerisation…	5 (50%)
New industrial phenotype of m…	1 (10%)
New industrial phenotype of pol…	2 (20%)
Science and citizen educationa…	2 (20%)
Societal/environmental impact…	4 (40%)

Why is this data collected?

- 5 responses

Response No.	Reason for data collection
1	This data is collected so the studies from WP12 can take place.
2	This data is collected in order to find and optimize the best whole cell enzymatic system for the conversion of medium chain alkanes into diols and diacids. We aim desired selectivity as well as desired activity levels to enable scale-up of our biocatalytic system.
3	In order to facilitate the communication, dissemination and exploitation activities, to analyse the market landscape and to ensure the sustainability and impact of the project outputs.
4	base for strain engineering
5	The collection of this data will facilitate the fast access to the broad audience of the potential users and further commercial exploitation.
6	for research

Please, provide a brief description of any external dataset used

- For every external data set used please explain its origin, relevance and license
- 5 responses

| Response No.	Description of external datasets
1	Literature if more input are needed
2	Protein DNA and amino acid sequences from open databases like NCBI and UniProt Protein structures from open databases like Protein Data Bank (PDB)
3	-
4	The existing data as well as the knowledge of scale-up operation of previous studies from the industrial partners (WPs 10 & 11) may be used for reactor redesign and optimization of variables. Most of the data in WP10 are generated from lab experiments or pilot scale-up tests. The data belong to or are associated with new scientific findings, new processes, new technology, new catalysts or materials, new products and specifications, developed production lines and/or device prototypes.
5	The existing data as well as the knowledge of scale-up operation of previous studies from the industrial partners (WP 11) may be used for reactor redesign and optimization of variables. Most of the data in WP11 are generated from lab experiments or pilot scale-up tests. The data belong to or are associated with new scientific findings, new processes, new technology, new catalysts or materials, new products and specifications, developed production lines and/or device prototypes.

6	none

FAIR DATA (DATA FINDABLE, ACCESSIBLE, INTEROPERABLE AND REUSE)

DATA FINDABLE

Tim's 5-star classification of the dataset

https://5stardata.info/en/

Metadata standards
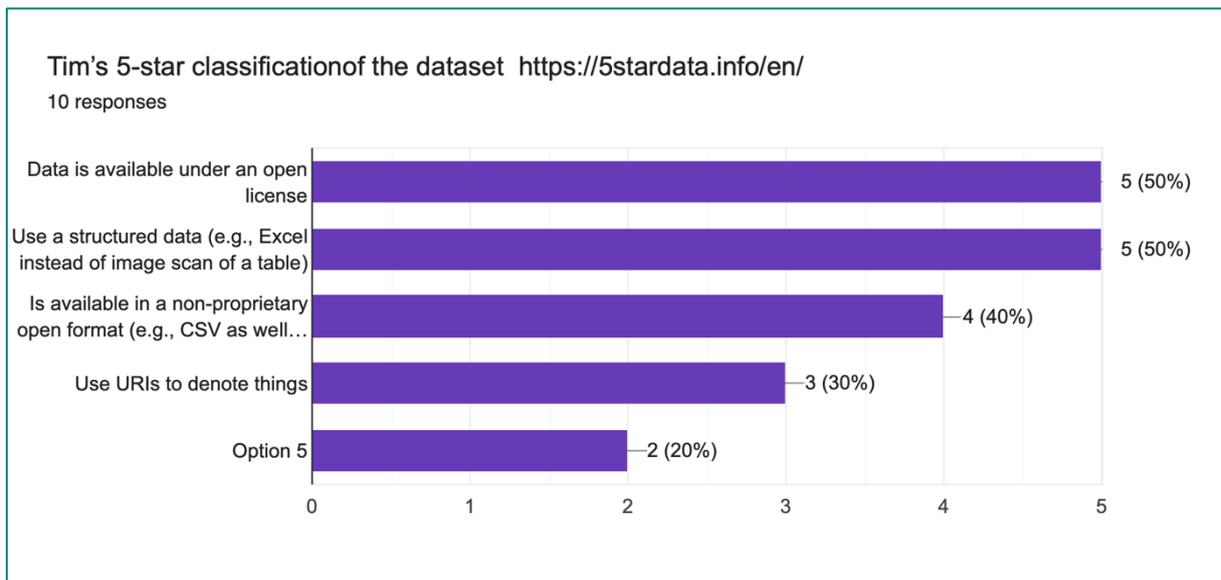
- ▪ (Please cite the standard and format use for the date. If any this data set does not follow any standardized format, please provide a formal specification. For example, the Dublin Core Metadata Initiative, Inspire Initiative, ISO, etc.) https://atlan.com/what-is-metadata/?ref=/metadata-standards/
- ▪ 3 responses

| Response No.	Metadata standards
1	The inventory will be based on ISO 14044 and ISO 14040 for the LCA studies.
2	XML
3	The metadata of the datasets published in Yoda complies with the Datacite V4 standard. In addition, the system facilitates validation by per-research-configurable metadata schemas based on JSONschema to support discipline specific metadata. The metadata is stored persistently along with the data and in the Yoda repository database. It is findable in Datacite upon publication of data. The metadata in the Yoda repository can also be retrieved via a service that complies with the Open Archives Initiatives Protocol for Metadata Harvesting (OAI-PMH).

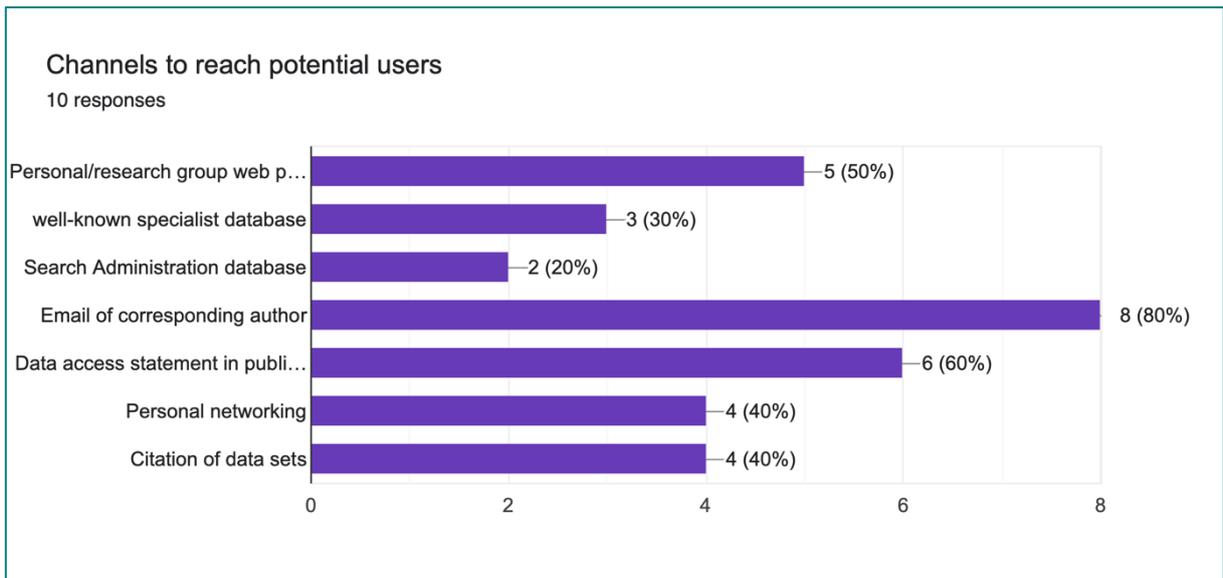| 4	ISO

Documentation stored in the data
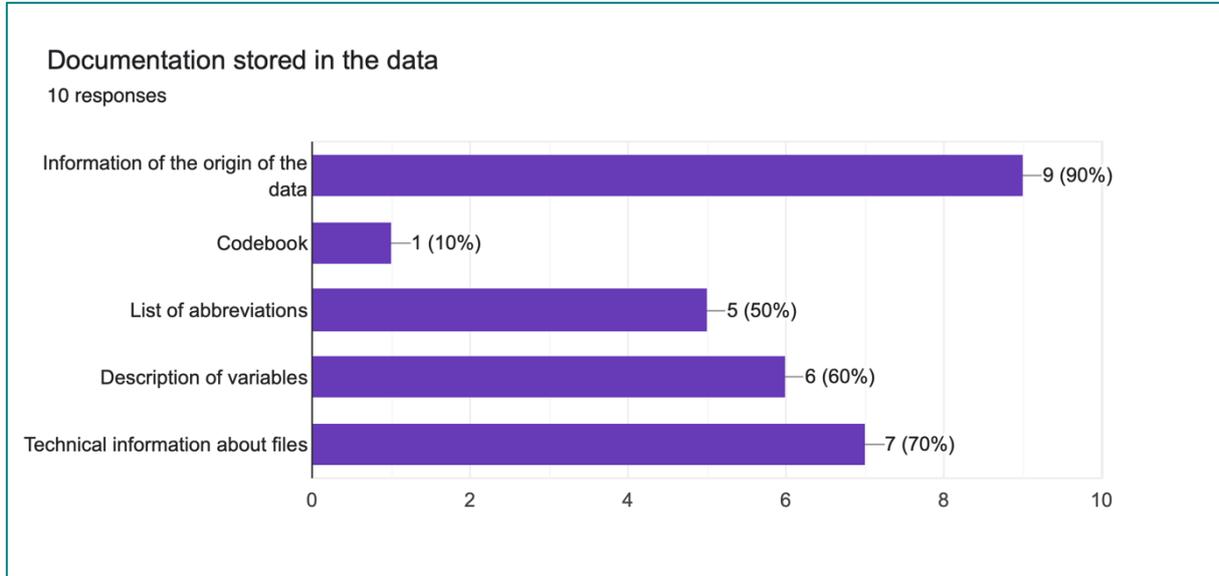
Ontologies
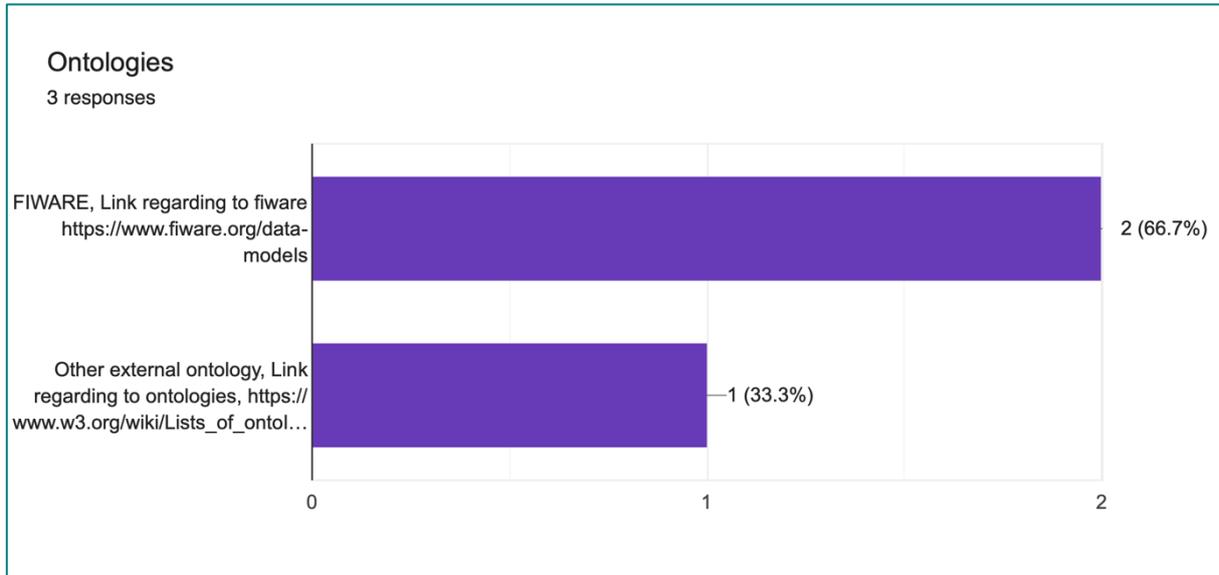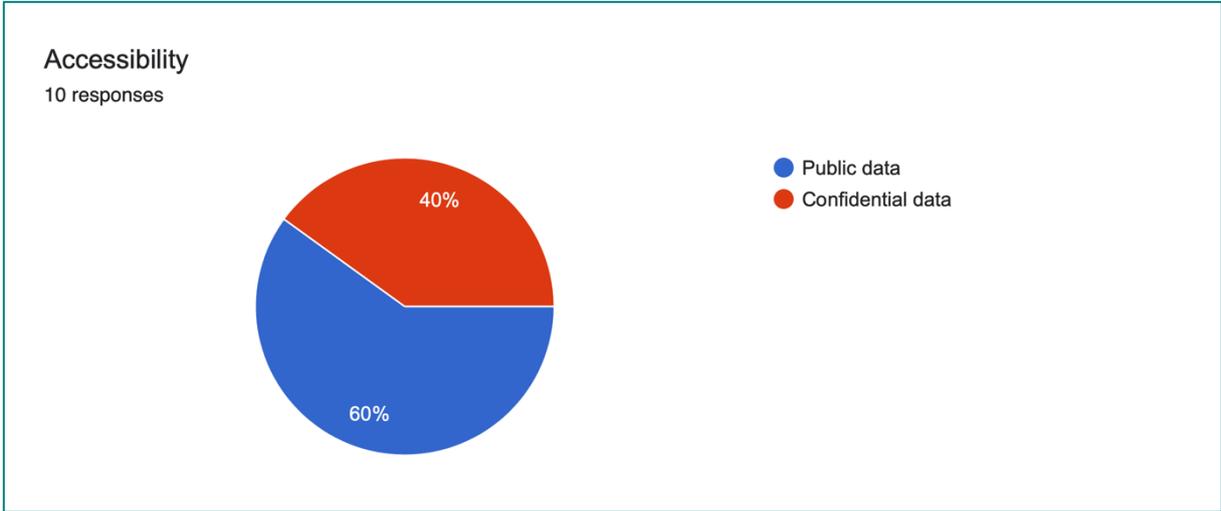
For whom might the data be useful
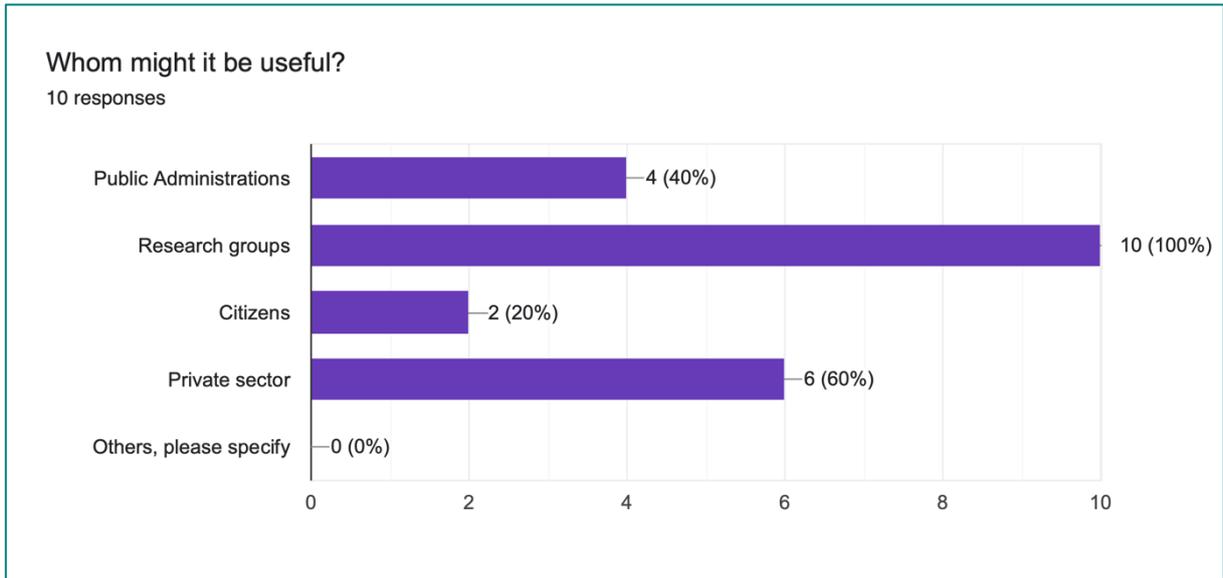
Channels to reach potential users

DATA ACCESSIBLE

Data accessibility

Accessibility
10 responses

- Public data
- Confidential data

40%

60%

Obligation or intention to share data

Obligation or intention to publish/share data
10 responses

- Yes
- No

20%

80%
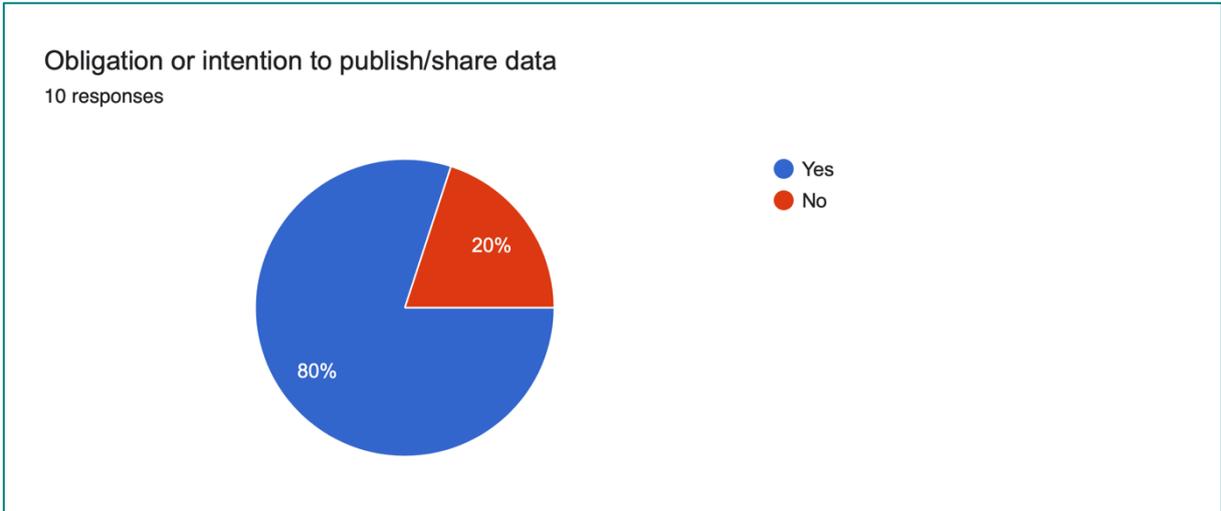
When will the data be published?
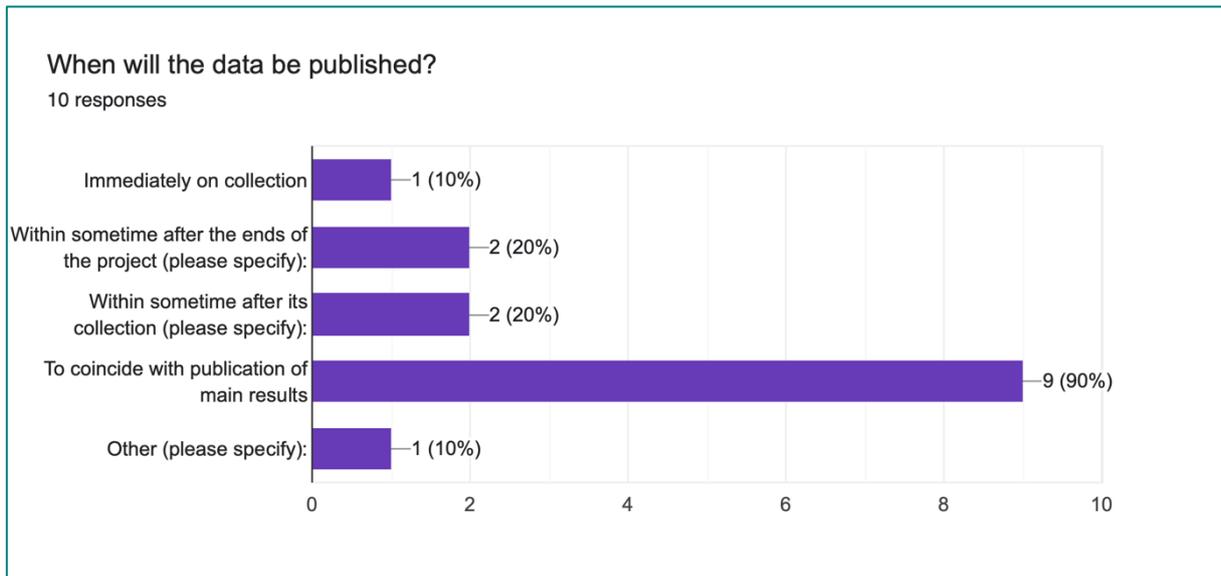
Expected difficulties with regard to file sharing
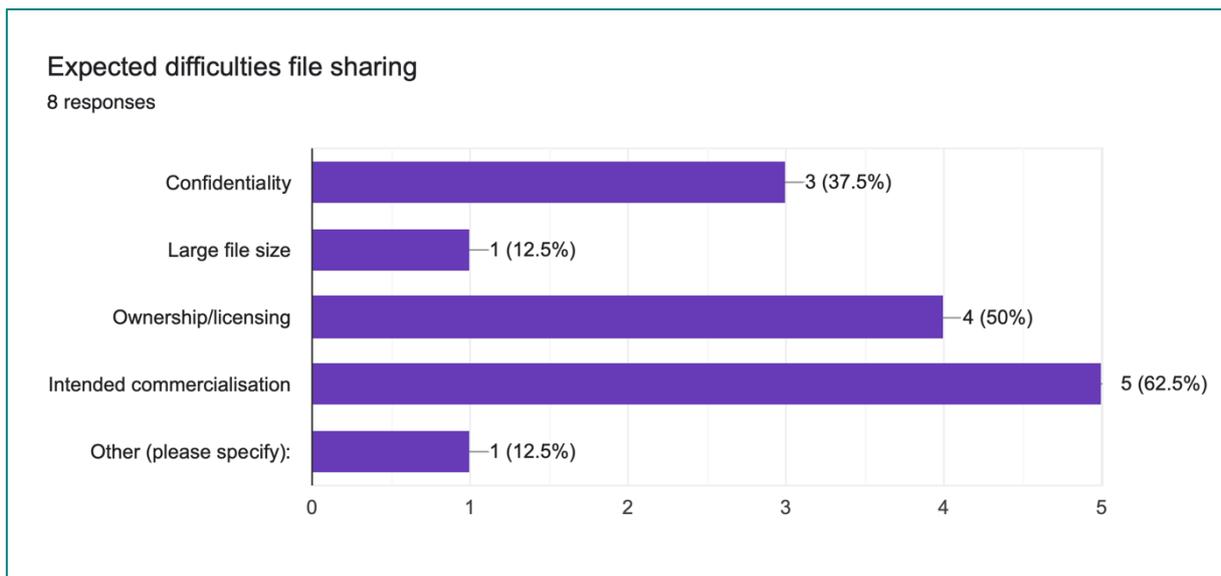
DATA INTEROPERABLE
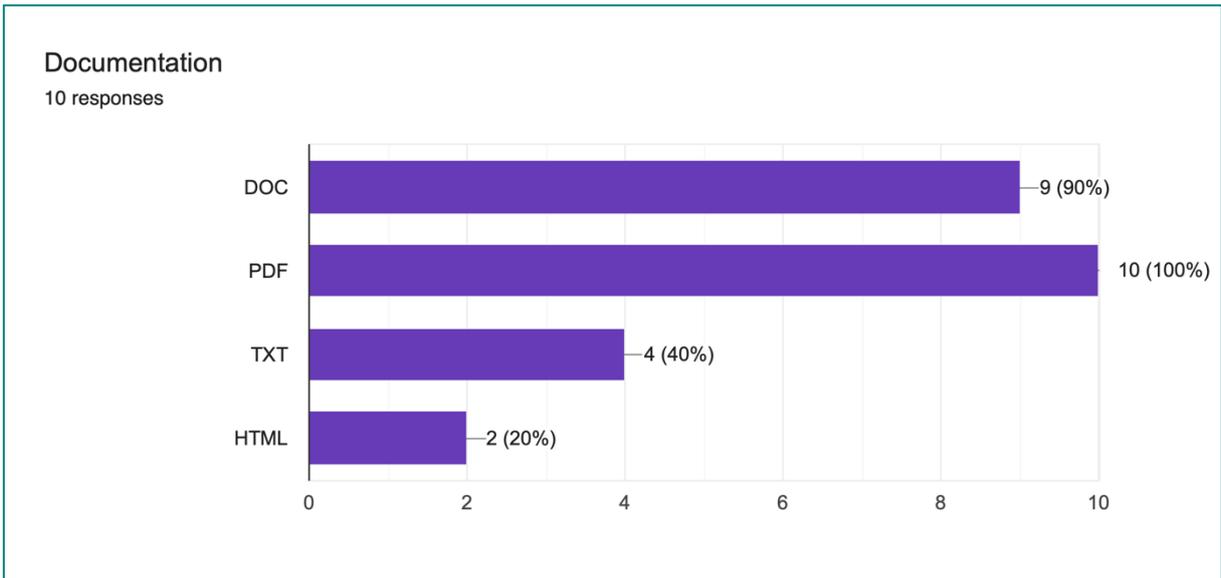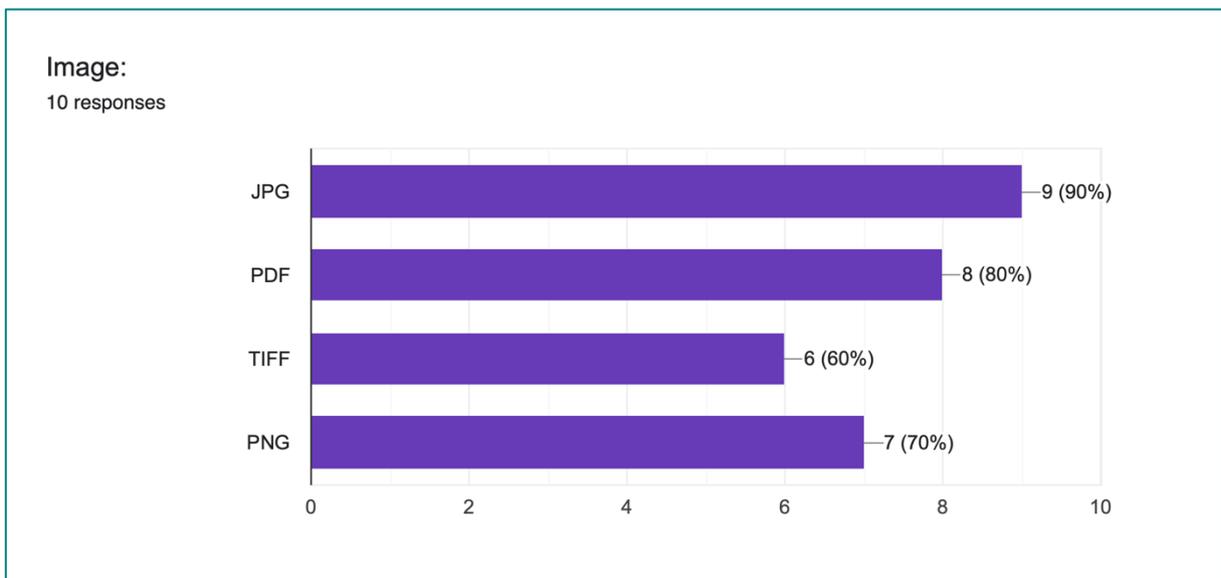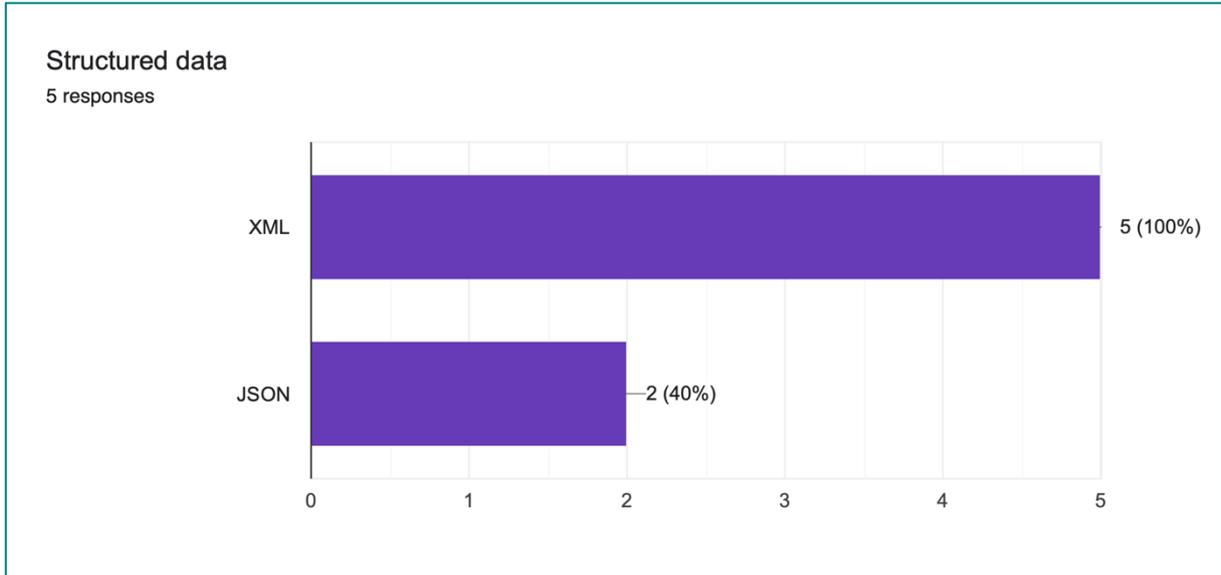
File formats

File format
10 responses

ODS	3 (30%)
XLS	8 (80%)
CSV	6 (60%)

Documentation
10 responses

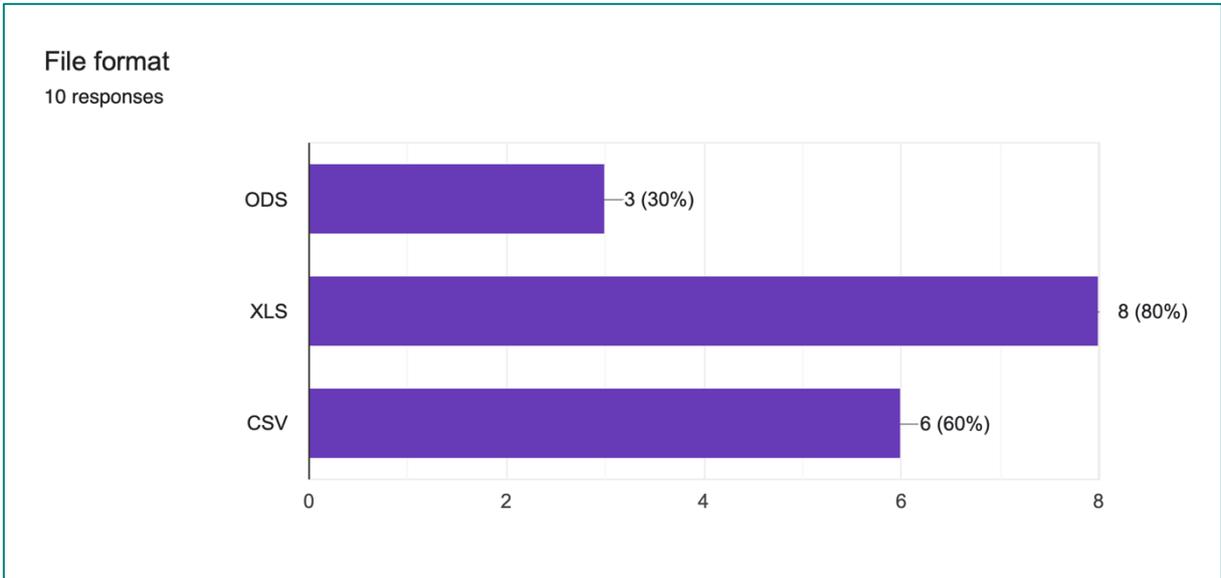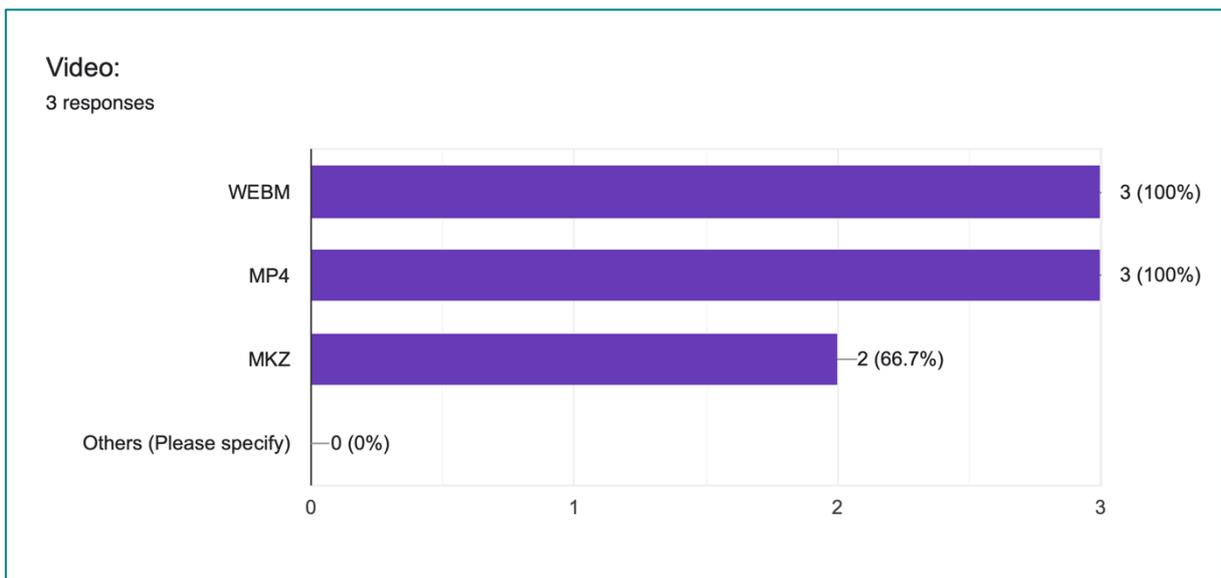|	
DOC	9 (90%)
PDF	10 (100%)
TXT	4 (40%)
HTML	2 (20%)

Structured data
5 responses

XML	5 (100%)
JSON	2 (40%)

Image:
10 responses

JPG	9 (90%)
PDF	8 (80%)
TIFF	6 (60%)
PNG	7 (70%)

Video:
3 responses

WEBM	3 (100%)
MP4	3 (100%)
MKZ	2 (66.7%)
Others (Please specify)	0 (0%)

Methods or software tools needed to access the data

- Please detail any necessary software to manipulate the information (if not standard).
- 9 responses

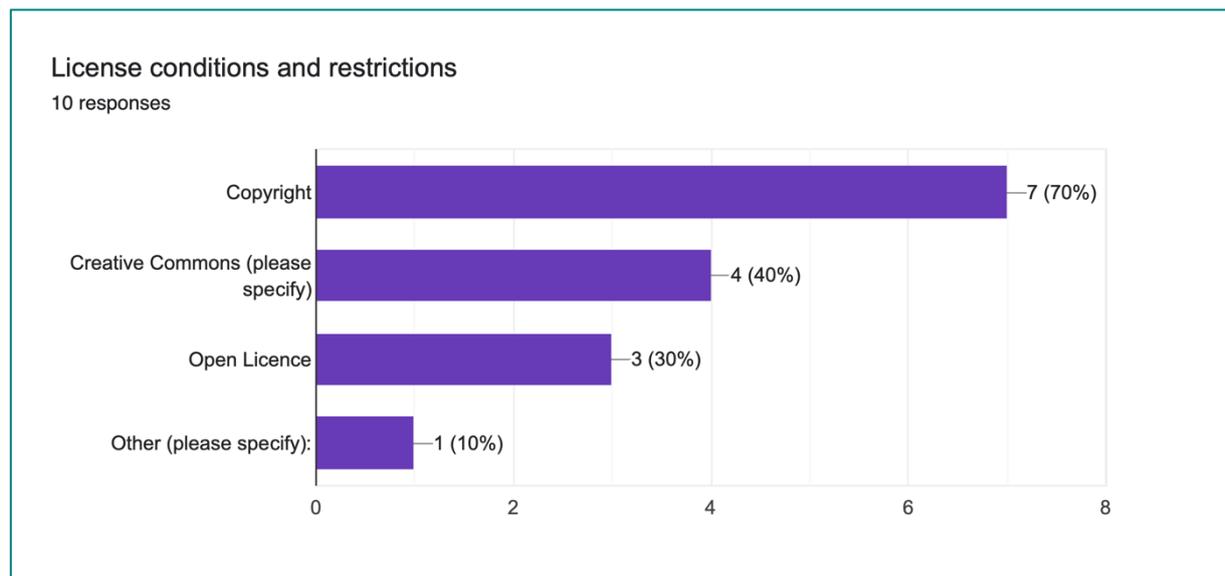| Response No.	Methods or software tools needed to access the data
1	Excel, Word, Adobe Reader
2	Zenodo sets its own Zenodo code, which is itself open source and is built on the foundation of the Invenio digital library which is also open source. The work-in-progress, open issues, and roadmap are shared openly in GitHub, and contributions to any aspect are welcomed from anyone.
3	The data will be collected by excel or word files
4	Standard
5	-
6	standard
7	All data will be shared in open and recommended formats. Where specific disciplinary software is needed to read or process the data, said software will be described in the documentation (README file) accompanying the datasets.
8	data is ASCII

DATA REUSE

License conditions and restrictions

License conditions and restrictions
10 responses

- Copyright — 7 (70%)
- Creative Commons (please specify) — 4 (40%)
- Open Licence — 3 (30%)
- Other (please specify): — 1 (10%)

Please, list the owners of the copyright and intellectual property involved

- 9 responses

Response No.	Owners of the copyright and intellectual property
1	CNRS, University of Bordeaux, Bordeaux INP
2	Partners of each process
3	publishers of the journals
4	-
5	not known yet
6	UU, AU, UM, Biolynx and Innovaplast
7	UG, B4PLASTICS, CNRS, CTCR
8	Utrecht University
9	UG

Access permissions and restrictions

- List roles/individuals (internal & external) with any limitations to access (e.g. scope, actions permitted), including who has authority to grant additional access.
- 10 responses

Response No.	List of access rights
1	Head of CNRS, Head of University of Bordeaux, Head of Bordeaux INP
2	The ACTPAC access committee and the IPR owners
3	Partners information, about each processes and formulations they are involved in, will be used, so it will be necessary to check with each partners how the information can be share.
4	Bekir Engin Eser
5	The consortium has access to the data. Additional access can be provided by M&S.
6	not known yet
7	Prof. Bert Weckhuysen; he is the WP leader has the authority to grant access to data.
8	decided per dataset

DATA MANAGEMENT AND ALLOCATION OF RESOURCES
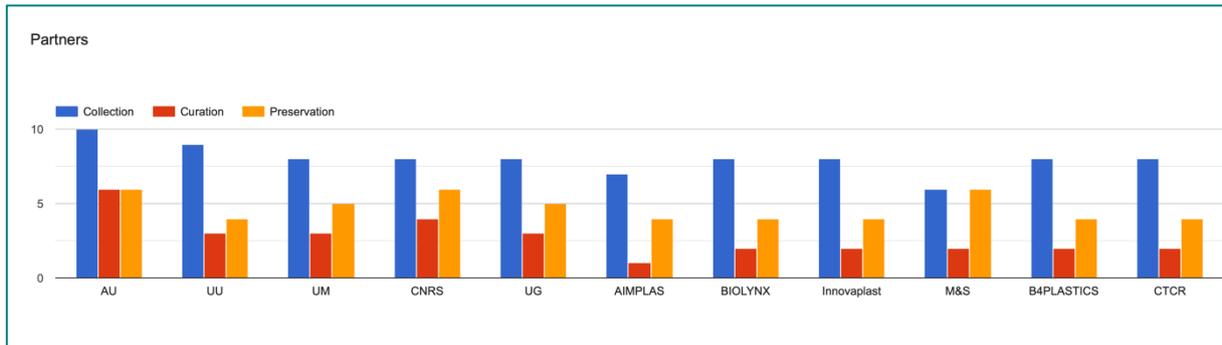
What are the costs for making data FAIR in your project?

▪ 4 responses

Response No.	Costs for FAIR data
1	Costs mainly from the data collection and curation, and the storage of the data is free.
2	N/A
3	not known yet
4	our university currently covers this

How will these be covered?

▪ 4 responses

Response No.	Coverage of costs
1	Covered individually or split share
2	N/A
3	not known yet
4	our university currently covers this

Primary storage medium and location
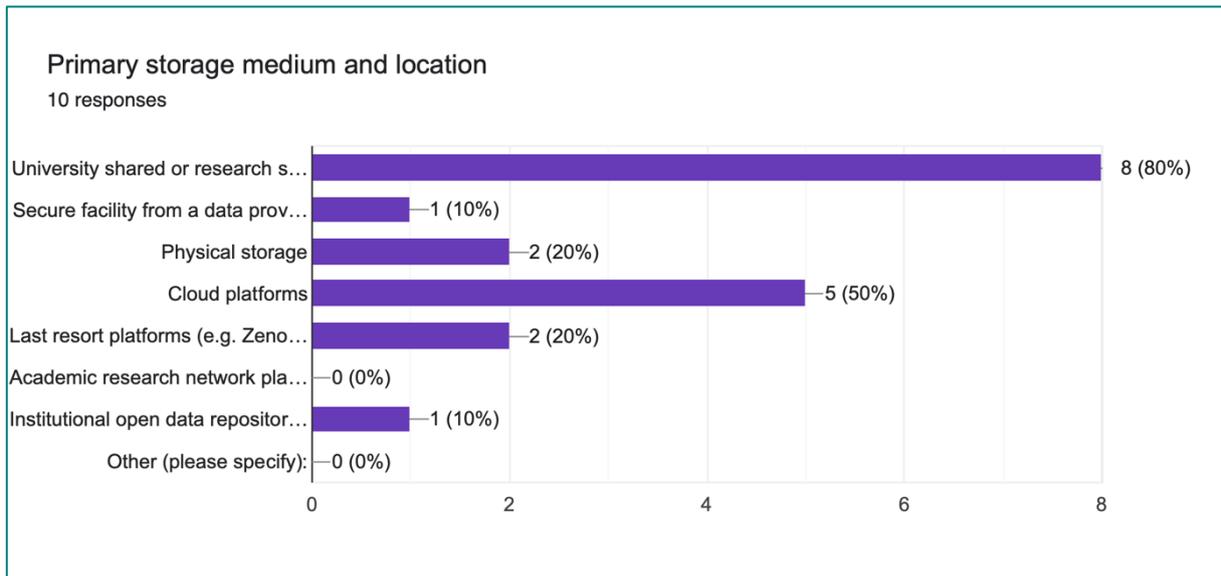
Primary storage medium and location
10 responses

University shared or research s...	8 (80%)
Secure facility from a data prov...	1 (10%)
Physical storage	2 (20%)
Cloud platforms	5 (50%)
Last resort platforms (e.g. Zeno...	2 (20%)
Academic research network pla...	0 (0%)
Institutional open data repositor...	1 (10%)
Other (please specify):	0 (0%)

Data curation processes

- Please briefly describe the management of data throughout its life cycle.
- 10 responses

Response No.	Description of the data curation process
1	Cloud platform of the Bordeaux INP
2	A standard procedure for DMP in the ACTPAC project has been made, for the data quality checking and validation in ACTPAC.
3	Data will be collected via email with word and excel files. Then they will be save at AIMPLAS cloud platform (one drive) and also at the project share point if it is necessary (teams). Then this data will be worked to make the sustainable studies of WP12.
4	Data will be curated for publications and for communicating results to other partners
5	Data will be organised, integrated and presented by M&S. The data will be processed for project deliverables, market analysis, and communication activities (e.g., articles, social media posts, factsheets, etc.)
6	raw data will be stored, data will be analysed and used frof graphical display, data will be used to extract next experimental design
7	all the data is stored locally, and after publication, will be uploaded in Yoda (the UU)
8	done by the respective researcher

How will long-term preservation and access be assured?

- Please briefly describe how the data will be preserver after the end of the project.
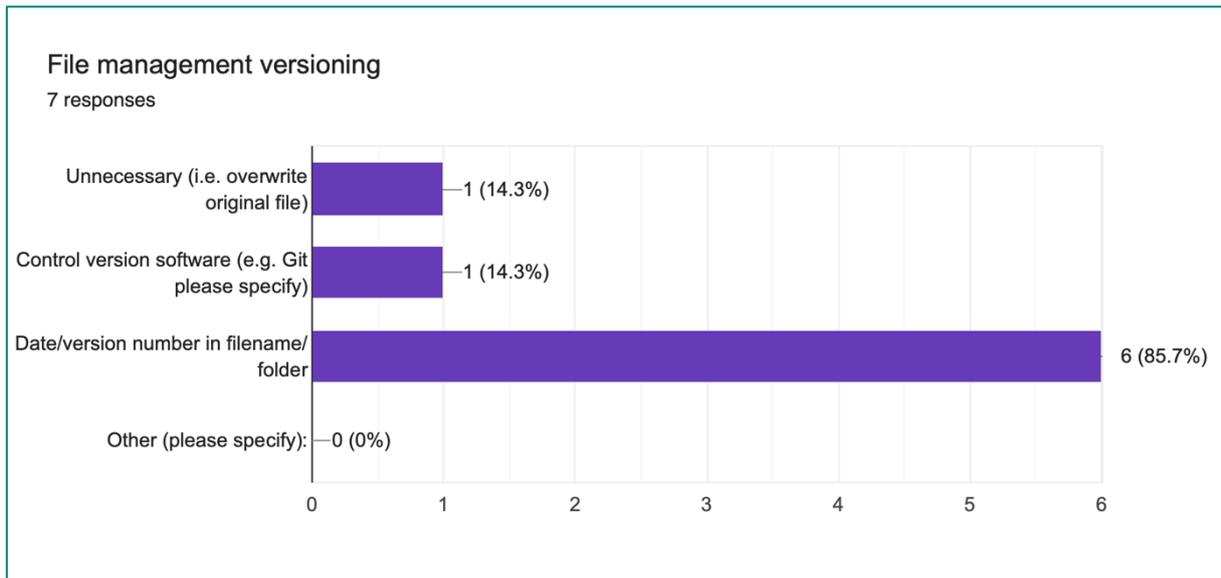- 10 responses

Response No.	Preservation and access times
1	5 years after the end of the project
2	Metadata will remain accessible, even when the data are no longer available. Data and metadata 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.
3	This project will have a code in our server (one drive) and all the information produced during this project will be preserved there.
4	Aarhus University's data storage capabilities will be used for long term storage
5	The data will be stored on the M&S Google Drive account until at least one year after the end of the project.
6	long term storage at university server, long time storage in project cloud, publication with open access
7	10 years
8	server of the UG

Regularity of backups and data performed. Replicas in other different places (if any)

- 6 responses

Response No.	Backups and replicas
1	I do not know if the company has copies of the project folder
2	Mainly cloud/network storage
3	regular back up of university servers (mirrowing)
4	The data will first store in AU server and then deposit in Zenodo
5	Every 6 months will be backuped. Metadata will remain accessible, even when the data are no longer available. Data and metadata 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.
6	server of the UG

File management versioning

ETHICAL AND LEGAL ASPECTS

Ethical aspect (If know)

- No/Yes, please briefly describe
- Aspects regarding informed consent in data collection and information protection in data storage and access.
- 6 responses

Response No.	Ethical aspects
1	No
2	No, only publicly available data will be collected.
3	No
4	No, No ethical risk has been identified related to WP10.
5	No, No ethical risk has been identified related to WP11.
6	No

Legal aspect (If know)

- No/Yes, please briefly describe:
- 6 responses

Response No.	Legal aspects
1	Yes, the project will be carried out in compliance with Article 10, and Article 14.1, listed in the Grant Agreement of the ACTPAC: applicable EU, international and national law, including the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.
2	No
3	No
4	Yes, might be use for IP

OTHER ASPECTS

- Description optional
- 5 responses

Response No.	Other aspects
1	No
2	No
3	No