ReCreate

Reusing precast concrete for a circular economy

Satu Huuhka et al.

Data Management Plan

Publisher / project	Reusing precast concrete for a circular economy
Project acronym	ReCreate
Deliverable	D10.4 Data Management Plan
Lead beneficiary	TAU
Author(s)	Satu Huuhka, Erik Stenberg, Rijk Blok, Kjartan Gudmundsson, Jukka Lahdensivu, Patrick Teuffel, Angelika Mettke, Viktoria Arnold, Leena Aarikka-Stenroos, Paul Jonker-Hoffrén and Mikko Knuutinen
Dissemination level	Public
Cite this document	Huuhka, S., Stenberg, E., Blok, R., Gudmundsson, K., Lahdensivu, J., Teuffeul, P., Mettke, A., Arnold, V., Aarikka-Stenroos, L., Jonker-Hoffrén, P. and Knuutinen, M. (2021). Data Management Plan. The ReCreate project.
Disclaimer	The content presented herein reflect the authors' views. The European Commission is not responsible for any use that may be made of the information this publication contains.

Funded by
the European Union

Contents

Funded by
the European Union

Abstract

This Data Management Plan (DMP) lays out the data management procedures and practices in the ReCreate project. The purpose of the plan is to ensure that the data developed in ReCreate will be FAIR – Findable, Accessible, Interoperable, and Reusable – and that resource allocation, data security and ethics are taken into account in the collection, handling, and storage of the data. The DMP is a living document which will be developed and updated throughout the project. The current document is the initial DMP devised in the beginning of the project.

List of abbreviations

3D	Three-dimensional
ASCII	American Standard Code for Information Interchange
BIM	Building Information Model(ling)
CC	Creative Commons
CO2	Carbon Dioxide
DfD	Design for Deconstruction
DMP	Data Management Plan
DOI	Digital Object Identifier
DPO	Data Protection Officer
EPD	Environmental Product Declaration
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
gbXML	Green Building Extensible Markup Language
GDPR	General Data Protection Regulation
IFC	Industry Foundation Classes
IP(R)	Intellectual Property (Rights)
LCA	Life Cycle Assessment
RMIS	Raw Materials Information System
URN	Uniform Resource Name
WP	Work Package

4	• Documenting the material and structural properties of the donor building's elements. • Detecting possible deterioration or harmful substance contamination in elements. • Determining which elements to deconstruct from the donor building on the basis of the residual service life. • Supplying condition data for WP 5 to act as the basis of reuse planning. • Creating a process to reliably ensure the relevant properties of salvaged elements in order to recertify them from waste into construction products. • Demonstrating to authorities that salvaged elements can be safely reused. Securing a building permit for the pilots in WP 5.
5	• Identifying and understanding the main technical aspects that influence reuse. • Determining necessary repair measures for the deconstructed elements based on condition data from WP 4. • Designing novel connectors that can be retrofitted to the elements to reconnect them for the cases where original connectors cannot be reused. Exploring the potential for design for deconstruction (DfD) in the new connectors. • Developing suitable design approaches for building design from salvaged elements. Understanding how reuse may influence roles of different experts and collaboration. • Securing building permits for the real-life pilots and erecting them. • Supplying data from the pilots for environmental and economic assessment in WP 6.
6	• Determining the environmental and economic impact of deconstruction and reuse. • Quantifying the environmental and economic impacts in comparison to (1) business-as-usual demolition and new construction from virgin materials and (2) demolition and new construction from recycled concrete. • Evaluating the potential environmental contribution of deconstruction and reuse on the EU level to the union's CO_2 reduction and circular economy targets. • Developing Environmental Product Declarations (EPDs) for deconstructed elements.
7	• Identifying the critical players in the deconstruction and reuse business ecosystem. • Understanding how the business is shaped by available technologies, and social and legal environments. • Identifying enablers and barriers of business-making from deconstructed elements. • Uncovering the operations and practices where it is the easiest to optimize value creation and value capture. • Planning the commercial and other exploitation of the project's results. Analyzing how the results can be transferred and replicated beyond the project. Managing the Intellectual Property Rights (IPR) arising from the project.
8	• Identifying regulatory possibilities and barriers for reusing precast elements in the EU, national and regional/local level. • Analyzing social acceptability of reusing elements among relevant stakeholder groups, covering in particular work and employment aspects. • Uncovering the mechanisms behind social acceptance of deconstruction and reuse at the intersection of 'business' and 'work'. • Proposing solutions to overcome the identified regulatory and social barriers.

Table 2. Data types, formats, sharing, and potential users.

Generated data (WP)	Raw data	Raw data format	Generated data format	Sharing of generated data outside consortium	Potential users of generated data
Precast concrete taxonomy (WP1)	Precast system handbooks, technical specifications, interviews	Literature, archival documents, transcriptions of interviews, drawings by interviewees	Online database or publication, possibly integrated to the RMIS	Fully Open Access	Researchers, practitioners, and policy-makers interested in reuse. Can also benefit the aforementioned groups in the context of building maintenance, repair, retrofit, and renovation.
Point clouds data (WP2)	3D coordinates of point cloud data	ASCII format or other binary data codes into the point cloud library	The data format and library will be adjusted for input to BIM (open IFC format)	Used method: Fully Open Access.	Point clouds data (WP2)
BIM-based pre-decon-struction audit (WP 2,3)	Building data, 3D model and additional (element) information	BIM model and databases data (e.g. quality check data) depending on exchange format such as Revit based or otherwise needed for inter-operability	BIM object in IFC/ Revit / gbXML formats	Audit Model: Fully Open Access	BIM-based pre-deconstruction audit (WP 2,3)
Quality check data (WP4)	Measurements and tests of deconstructed elements and their materials	Visual observations and numerical data from laboratory equipment	Condition investigation and laboratory test reports	Fully Open Access	Researchers and practitioners engaging in other reuse projects. Policy-makers that aim to encourage other reuse projects.
Redesign and reassembly process (WP5)	Drawings and calculations models	Drawings and calculations models	Reports, publications and construction document-tation	Reports, publications: Fully Open Access; Construction document-tation to building owners, contractors and authorities	Researchers and practitioners engaging in other reuse projects. Policy-makers that aim to encourage other reuse projects.

:ReCreate

Huuhka, S. et al. (2021)
Data Management Plan.

Environmental Product Declarations (EPDs) for reuse of precast concrete components (WP6)	Energy consumption, carbon emissions and other environmental impacts of reuse from the pilot projects	Life Cycle Assessment (LCA) data according to the template, which is the same for each participating country	(Inter)-National EPDs	Fully Open Access	Environmental Product Declarations (EPDs) for reuse of precast concrete components (WP6)
Operational costs data (WP7)	Interviews, media-sourced and company sourced documents and reports, observation data	Transcriptions of interviews, fieldnotes, archival documents	Reports, publications	Conditional access to analytical data on project or operation level. No access to detailed data to protect trade secrets.	Researchers and practitioners engaging in other reuse projects. Policy-makers that aim to encourage reuse.
Data on ecosystem actors (WP7)	Interviews, media-sourced and company sourced documents and reports, observation data, ethnographic data	Transcriptions of interviews, fieldnotes, archival documents	Reports, publications, ecosystem maps	Conditional access to pseudonymised data to protect participants	Researchers and practitioners engaging in other reuse projects. Policy-makers that aim to encourage reuse.
Changes in work processes (WP8)	Interviews, ethnographic data	Transcriptions of interviews, fieldnotes	Value-chain based professional skills inventory	Conditional access to pseudonymised data to protect participants	Changes in work processes (WP8)
Attitudes to reuse of precast concrete components (WP8)	Survey among stakeholders	Survey data	Publication	Fully Open Access after anonymization to protect participants	Researchers in other reuse projects. Policy-makers that aim to encourage reuse.
Legal environment of pilot projects (WP8)	EU- and national laws, local regulations	Legal texts	Synopsis and abstraction of legal norms, published as a guide	Fully Open Access	Researchers and practitioners engaging in other reuse projects. Policy-makers that aim to encourage reuse.

2.   FAIR data

2.1. Making data findable, including provisions for metadata

The data that can be made openly accessible without compromising the protection of personal data and the commercial exploitation of the project's results will also be made findable by depositing it a data repository , where it is assigned with a Digital Object Identifier (DOI) or a Uniform Resource Name (URN) and metadata with keywords to describe the content of the dataset. The datasets will be named with a descriptive name that best communicates their content; the connection to the ReCreate project and a specific WP and task within it will be provided in the metadata. In addition to this, the types of metadata to include will be follow the selected data repository's guidance and standard.

It is expected that data is made open and findable only in its 'final' state so that version numbers are not needed. Should a data versioning scheme be needed for internal project purposes, it is devised. In case open data is updated during the project, the new data will be assigned different identifiers that distinguish it from already published data.

Any data that will not be published due to commercial exploitation reasons will remain at the custody of the exploiting partner(s) according to a separate agreement about the assignment of Intellectual Property Rights (IPR). As such data cannot be shared outside the ReCreate consortium, it will not be assigned a DOI. Should the partner(s) in question revoke the decision to commercially exploit the data, the possibility to make the data open and findable will be reconsidered.

2.2. Making data openly accessible

In line with the Horizon 2020 Open Research Data Pilot, ReCreate is committed to open data principles and the sharing of the data when it does not compromise the commercial exploitation of project results or the personal data or consent of human participants. The data to be made publicly available in specified in Table 2.

As a rule, data is made openly accessible through Open Access data repositories apart for special cases, such as the Precast Concrete Database, the integration of which to the EU's Raw Materials Information System (RMIS) will be explored. The deposited data will encompass metadata to describe the dataset and a machine-readable license specifying the terms of data reuse. The repositories to be used will be specified in due course. Zenodo, IDA Research Data Storage Service and/or the Finnish Social Science Data Archive have so far been considered.

Generally, to access the numerical and/or textual data, no special software will be required. However, for Building Information Models (BIM), BIM software is required. The aim is to open the BIM data in an Industry Foundation Classes (IFC) format, which is intended to make BIM models interoperable in different software.

Business-sensitive data is only opened in an aggregated or synthetic format to protect commercial exploitation of results by the project beneficiaries. Business-sensitive results will only be published with an IPR protection clearance. An Innovation Manager will be named by the end of 2021 to coordinate IPR management. An IP protection roadmap will be devised to avoid conflicts with IP protection.

Data from social sciences research is only opened conditionally in an anonymized format to protect participants' identities and to ensure the ethical standards of research on humans can be met in data reuse. The conditional access is granted by the researchers themselves that gathered the dataset. The procedures relating to data from research on humans and personal data protection are provided in the deliverables D11.1 'H – Requirement No. 2' and D11.2 'POPD – Requirement No 3'.

The restricted datasets are only accessible by researchers that gather or process the data. These datasets are stored and processed in environments only accessible over multi-factor authentication by named individuals conducting the research. Please see Chapter 4 'Data security' for details.

For any data that cannot be made openly accessible, Tampere University's repository can nevertheless be used for storing the metadata.

2.3. Making data interoperable

The aim is to make as much of the Open Access data interoperable as possible. Datasets in numeric and/or textual format are stored in standard formats which enable easy reuse and re-combination with datasets from other origins. The aim is to make data requiring special software applications, such as Building Information Models (BIM), accessible in the open Industry Foundation Classes (IFC) format. However, in some cases, using the IFC format can mean that some functionalities are stripped from the model. For any BIM models that require such functionalities to be useful, the original software-specific format will be retained.

ReCreate is a multi- and interdisciplinary project. Construction is the project's overarching discipline, and so, standard language of the construction sector is the vocabulary that will be used in the majority of the project's data. However, some of its WPs operate from the premises of different disciplines, namely business, legal research, and social sciences. It is crucial for the success of the research in these WPs that discipline-specific standard vocabularies can be used to express ideas and theories and as analytical tools to identify

phenomena in the data. So, depending on the nature of the data, the vocabularies of the data will be a mix of standard construction and other discipline-specific language.

2.4. Increasing data reuse

The datasets to be published in Open Access will be assigned with appropriate Creative Commons (CC) licenses. Since the data to be published in openly will not contain commercial exploitation interests by the project partners, it is expected that types of CC licenses that allow commercial use as long as the data source is credited, such as CC-BY, can be employed. In most cases, it is not necessary to impose further conditions on the data. Since the project's aim is to facilitate business, the Non-Commercial (-NC) condition is not meaningful requirement. To enable commercial exploitation by non-project partners, the Share Alike (-SA) condition should, as a rule, not be required either. However, the aforementioned conditions can be applicable to certain datasets and the need for them should be determined on a case-to-case basis.

The aim is to make the data that is to be published in Open Access available for reuse as soon as possible, without delay. When this is, depends on the specific data in question. For some data items, the publication for reuse can be made as soon as the data has been processed into its final format and a platform for its sharing has been determined. For datasets associated with planned peer-reviewed publications, the data can be made openly available at the time of the results' publication. Once the Open Access data is deposited into a repository, it will be usable by third parties and will remain so after the end of the project for as long as the repository is operational.

Data that is related to commercial exploitation interests by the project partners will not be made available for reuse.

3. Allocation of resources

Making data FAIR is considered an integral part of the ReCreate project and any costs that may incur will be covered with the project funding as a part of the work in the Work Packages. Each Work Package and individual researcher is responsible for their own data management.

The long-term preservation of data will be discussed in the Core Group of the project, consisting of the Coordinator, the Work Package Leaders, and the Country Cluster Leaders, during the final year of the project. As much of the data as possible, namely the open data, will be deposited in repositories that will maintain them over the long term, in which case the allocation of resources for the long-term preservation is also made by the repository in question. The potential long-term preservation or destruction of the outstanding, access-restricted datasets, will remain the duty of the research team(s) that created them.

4. **Data security**

During the project, the collectors and holders of the data are responsible for storing the data safely and ensuring that backups are run regularly. The sensitivity of the data influences how the data must be stored and transferred between project partners. Raw and generated data is classified into three data sensitivity classes as follows:

1. Open-source data
2. Semi-confidential data
3. Confidential data

Open-source data encompasses raw data from public sources as well as project-generated data for Open Access publication. Apart for the requirement to back up regularly, there are no other requirements for the storage and transfer of open-source data.

Semi-confidential data consists of project data that is neither open-source nor confidential. Most of the project data belong to this class. It is project-generated data that needs processing (such as naming, classification, etc.) before it is meaningful to publish and becomes understandable for external users. This type of data should be stored in and transferred through password-protected environments. The processed data, once in its final format and ready for Open Access publication, becomes open-source data.

Confidential data encompasses personal data as well as business-sensitive data. This type of raw data will not be published for the use of others. However, the researchers generating and processing the data may store it for their own future research use, as set out in the data protection sheet of the study in question, over the long term. This type of data must be stored in and transferred through environments protected by multi-factor authentication. The access to the raw data must be limited only to the persons who need the access to conduct the research. Access may also be granted for service providers necessary to complete the research, such as companies that provide interview transcriptions from audio or video files.

Confidential data may be publishable in a processed, e.g. anonymized or aggregated, format. The researchers will assess the possibility to curate the data so that no sensitive information is revealed throughout the research process. Once the data is curated this way, it becomes open-source data. In some cases, the curation necessary to anonymize the data may go to such lengths that the data ceases to contain any meaningful information, capable to reveal noteworthy matters about the studied phenomenon. In such cases, it is not sensible to use the project's resources for data anonymization and consequently, this type of data will not be published.

Any data that is prepared for Open Access publication will be stored in and shared through a free data repository, such as Zenodo, IDA Research Data Storage Service or Finnish Social

Science Data Archive, which will take care of the long-term preservation of the deposited data.

5.   Ethical aspects

The data protection sheet provided for the interviewees before they sign the informed consent form covers the issues of personal data anonymization, long-term preservation, and open data sharing. No data with personal information will be shared openly, and personal data is destroyed when it is no longer needed. The data procedures relating to research on humans and personal data protection are detailed in the project deliverables D11.1 'H – Requirement No. 2' and D11.2 'POPD – Requirement No 3'.

6.   Other issues

Data management experts of Tampere University will support ReCreate researchers in all phases of data management when needed. A dedicated Data Protection Officer (DPO) and GDPR Legal Counsel in Tampere University provide regulatory guidance and instructions for the project personnel and monitor compliance with the GDPR and the university guidelines. Horizon 2020 DMP templates are capitalized on in the development of this plan.