



# DATA MANAGEMENT PLAN

Deliverable 1.2

<b>Project number</b>	101177706
<b>Project acronym</b>	COLUMN
<b>Project title</b>	Colonial Legacies of Universities: Materialities and New Collaborations
<b>Start date of project</b>	1 April 2025
<b>Duration of the project</b>	48 months
<b>Date of submission of the report</b>	14 July 2025
<b>Project website address</b>	<a href="https://column-horizon.eu/">https://column-horizon.eu/</a>
<b>This report is based on</b>	Grant Agreement project 101177706 — COLUMN

Version	Notes	Date	Author(s)
1.0	Approved by consortium	27-5-2025	Frans de Liagre Böhl, Myrthe van Groningen, Gertjan Plets



UNIVERSITÄT GRAZ



UNIVERZITA  
KARLOVA



AARHUS UNIVERSITET



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

studio  
louter



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA



UNIVERSITÉ  
DE GENÈVE



Funded by  
the European Union

**About this document**

The research program COLUMN consists of multiple inter-related research projects executed by a consortium consisting of different research institutes and one industry partner. The Data Management Plan (DMP) is a living document that will be updated as the methodologies for each work package are developed. This document could be considered as a Data Management Framework.

As a framework, this document will provide for:

1. A number of principles which hold for the entire consortium, with regard to data capturing and -handling during the research projects as well as the long-term data dissemination and – retention after the finalization of these projects.
2. An overview of the structure of the main relevant (i.e./ involving research data) activities and responsibilities, who will be the primary responsible party for actual data management activities per (sub) project.

As the program is subdivided in Work Packages, these will be the units used in this data management framework.

## About COLUMN

Historically, universities have been foundational to promoting and shaping European culture, knowledge, and values including democracy, human dignity, freedom, human rights, and equality as a global project. They have also been leading institutions for promoting dialogue and building strategic alliances between the global north and the global south in the fields of science, heritage, and cultural and creative industries. In recent years, universities and museums in Europe have increasingly developed innovative approaches to address the colonial entanglements in their collecting practices, material legacies, understandings of national and transnational identities, and the production of knowledge. However, unfinished conversations with former colonies about colonial legacies in and beyond Europe remain barriers to cooperation and exchange as well as the promotion of European values.

COLUMN brings together partners from the creative industries, seven European universities from the Coimbra network and partner universities from the global south to rethink, reframe, and refresh approaches to European and postcolonial arts, culture, heritage, and societal values. Focusing on four pilot interventions that apply new collaborative and decolonial curatorial practices, COLUMN will lead the development of policies, recommendations, and good practices for new collaborations around European university heritage through strategic partnerships within and beyond the borders of the EU. Bridging knowledge production and creative arts practices, it builds an understanding of European cultural heritage as a vehicle to strengthen rights-based and equitable approaches to colonialism in dialogue with global south partners - thus paving the way for cultural cooperation in the true spirit of modern European values. Through its deliverables, the project redefines scholarly consensus, heritage practice, and policy.

## Consortium wide principles for data management, ethics and privacy

The following consortium wide principles and guidelines are formulated. These will primarily hold for the research partners, but may also be of interest, c.q. relevance for the participating commercial partners.

### A) Subsidiarity

The leading principle is that of subsidiarity: with all activities concerning data management, ethics and privacy concerning a specific Work Package, the rules and regulations of the researching partner responsible for that particular Work Package will be followed.

This subsidiarity has practical implications. With regard to data retention, data management and ethical assessment the rules of the leading research party will be followed. These rules and regulations can differ per consortium member.

### B) Joined processing responsibility

The second principle is that of joined processing responsibly for all the leading (i.e. primary responsible) research institutes. This agreement enables the data exchange between the research institutes and ought to be mentioned in the letters to the data subjects informing them about the research (sub)projects.

### C) GDPR, Personal data, preferred legitimate ground, and ethics

Within this research program identifiable personal data will be processed. Certainly, in structured form and possibly in a more unstructured, qualitative form (open questions in surveys). However, all personal data will be collected in (visitor) surveys and is expected to be anonymized, after which GDPR-compliance is no longer an issue. Even when 'opinions' (with the implicit risk of asking for special categories of personal data) are involved, after anonymization these opinions can no longer be attached to a specific individual.

Given the deliberations in the previous paragraph, the preferred legitimate ground, demanded by the GDPR for the processing of personal data to be legal, will be *Public Interest*. This principle provides the researching partners with the most flexibility but warrant sufficient protection for the participating data subjects. Given the principle of subsidiarity, partners can deviate from this principle.

The central data management (UU) will offer the partners templates for information letters to be used informing the participants of the overall goal of the research program and the data sharing policy.

### D) FAIR data; anonymization, metadata, RAW data, DOI's and (long term) data retention

#### *FAIR data*

All WP's in which data will be created, c.q. captured, will offer these data sets in comprehensive data packages with adequate descriptions (metadata), a persistent identifier, user licenses and an accessibility corresponding the nature of the data, i.e. as openly accessible as possible.

#### *Open, Restricted and Closed Access*

The terms Open, Restricted and Closed Access, refer to the accessibility of archived data packages. While research is conducted (the 'research phase'), research data will be available only to the researcher involved in the COLUMN program.

All anonymous data packages will be published open accessible.

N.B. one should distinguish between data packages and metadata. Metadata (descriptions of datasets and -packages) should always be as open accessible as possible. This is taken care of by the repositories in which datasets are packed together in so called archival packages and stored for a longer period.

#### *RAW data & anonymization/pseudonymization*

In general, three types of datasets result from the research activities in COLUMN, sc.

- a) Annotated biographies, resulting from desktop studies
- b) Provenance reports, resulting from archival- and collection studies
- c) Sets of survey outcomes

Both a) and b) will not contain personal data, as the former contains holding locations of literature and the latter is defined as 'bibliography, inventory of depot numbers and objects and description.' (see *Grant Agreement, Descriptions of Deliverables*, 10.1 & 10.2, page 32). The processing of personal data, if at all will only occur in c) *Sets of survey outcomes*. In general, the anonymization of the structured data sets will take place by removing all structured elements (variables, columns) which contain personal data.

#### Handling administrative data

Administrative data, i.e. data which themselves will not be analysed but which is needed in the wider context of the research project, e.g. consent forms, key-files or participant administration, will both during and after research be processed on a location separate from the research data and accessible only on a need-to-know basis.

#### *Retention period*

The default long term retention (archival) period of research data is 10 years after the first publication of the research results or, if no publication is realized, after finalization of the research WP/sub-project.

#### *Cost long term retention*

The costs for long term retention of a data package is covered by the research partner creating that package for the duration of that retention.

#### *Metadata*

Each archival data package will be accompanied by a machine-readable set of metadata compatible with the DataCite metadata schema vs. 4.5 and will be uniquely referable (and findable) through a DOI.

For each archival data package resulting from this research program, the value of the first DataCite *Keywords* field shall be set to COLUMN.

Whenever possible data packages (e.g. an anonymized Open Access data package containing interview transcripts and a data package containing the RAW recordings of the interviews) are related, this relation will be indicated and described in the metadata. E.g. by a reference to one another through DataCite's relatedIdentifier.

A DataCite compatible metadata scheme is primarily for describing an archival package and less so for the content of the package. Hence, whenever needed c.q. relevant, datasets will be accompanied by codebooks describing the structure and semantics of a particular data file, e.g. a readme.txt describing the column-headers and contents of an excel file.

#### *Use Licenses*

Alle Open Access data packages will be published with a CC-BY-SA 4.0 user license. All data packages not offered to the public domain (*Closed* c.q. *Restricted Access*), will be offered with the same research-program broad customized user license. No data created in the context of this research program shall remain closed access for commercial reasons.

#### *Data ownership*

Though data ownership is a misnomer in relation to digital data, the first responsibilities for and control over the research data lays with the research partner creating these data.

#### *Embargo on data*

During the research phase, i.e. as long as data are being analyzed, the data will only be accessible to the research team.

### Data Reuse

In principle only in case of restricted access data is reuse an issue. In principle, the creator of the data archive will be the first party to decide over requests for reuse and will do so according to the rules and regulations of that particular research institute. During research, data sets are only reusable by research partners in the consortium. It is to be expected that all datasets resulting from this research project can be offered for reuse.

### Data Access Management

During research each research partner in the lead of a WP will be responsible for adequate access management. After research, i.e. when archived data packages are involved, the responsibility for correct access management lays in the hands of the repository processing the Archival Data Packages.

Access to research data while the data is still analyzed is only for the researchers of the project. It is expected that alle archival data packages can be published as *Open Access*.

### Costs and resources for data management

Each research institute which is in the lead for a WP will organize data management for that WP and will cover the costs.

### Data security

During research data sets will be stored on the data infrastructures of the research partners, each of which provide for differentiated role-based access, adequate data recovery and incident management procedures.

## 1. DATA SUMMARY

The data is collected to answer research question on how reframe, and refresh approaches to European and postcolonial arts, culture, heritage, and societal values. In general, the research project will yield three types of dataset which are created to answer the underlying research question:

- a) Annotated biographies, resulting from desktop studies
- b) Provenance reports, resulting from archival- and collection studies
- c) sets of survey outcomes

It should be noted that a number of events will be organized in the context of this research project which might be registered through audio/video recording. However, these events do so much not 'yield data' used to answer the research questions and will hence not be regarded as data which needs to be covered by this plan.

### Data list

In the COLUMN there is a number of WP's/Activities in which datasets will be created. These are listed in the table down here.

**Table of tasks with datasets to be collected:**

Deliverable No	Deliverable Name	WorkPackage	GDPR? (1)	Format	IP Owner
D4.1	Annotated bibliography colonial history	WP4	No	multiple (2)	CU
D5.1	Annotated bibliography heritage	WP5	No	multiple (2)	UP
D6.1	Provenance reports Collections UNIBO South Africa	WP6	No	multiple (2)	UNIBO
D8.1	Report archival research campus	WP8	No	multiple (2)	AU
D10.1	Provenance reports gardens Bologna	WP10	No	multiple (2)	UNIBO
D10.2	Provenance report gardens Utrecht	WP10	No	multiple (2)	UvS
D15.1	Visitor survey: collections intervention	WP15	Yes	Tabular	UP
D15.2	Visitor survey: campus intervention	WP15	Yes	Tabular	AU
D15.3	Visitor survey: gardens intervention	WP15	Yes	Tabular	UU
D15.4	Visitor survey: intangible intervention	WP15	Yes	Tabular	UNIGE

(1) GDPR? – a Yes in this column indicates that the set should, at least until anonymization, must be managed in compliance with the GDPR.

(2) The reports themselves will be in text format; the underlying sources can be multiple. Each provenance report will in fact be a text with an extensive reference apparatus to the holding locations of the objects described in the report.

The expected amounts of data will not be such that issue with regards to either transport or storage costs will pose a problem. As a first estimate, not more than 50GB in total will be expected.

## 2. FAIR DATA

For general principles, see above section '**Consortium wide principles for data management, ethics and privacy; D) FAIR data**'.

During research, i.e. as long as data sources are being analyzed transformed etc. the data will be closed and only available for the research team. Only when a publication is due and the underlying data need to be deposited, all relevant datasets will be packed into a so-called archival data package, containing the research data, a set of machine-readable metadata and a use license, a file describing the conditions for reuse (if any).

All archival data packages resulting from this research projects will have a DOI assigned, thus making these packages uniquely identifiable and referable.

Digital Object Identifiers are handed out by the organization DataCite and can only be assigned after a minimal set of metadata has been provided. Data packages from the COLUMN program will be metadated using the DataCite 4.5 scheme or higher.

For each archival data package resulting from this research program, the value of the first DataCite *Keywords* field shall be set to COLUMN.

Whenever possible data packages (e.g. an anonymized Open Access data package containing interview transcripts and a data package containing the RAW recordings of the interviews) are related, this relation will be indicated and described in the metadata. E.g. by a reference to one another through DataCite's relatedIdentifier.

NB: this is about the description (metadata) of archival data packages, i.e. a description of an archival package, which itself can contain multiple datasets (e.g. the outcome of different surveys).

Where necessary, individual datasets will be stored with readme-files describing the set. E.g. a tabular set of survey answers might be accompanied by a readme-file explaining abbreviated column-headers, the semantical standards used (F=Female, M=Male), etc.).

### 2.2. Making data openly accessible

For general principles, see above section '**Consortium wide principles for data management, ethics and privacy; D) FAIR data**'.

Research data produced in the project is owned by the partner that generated it. Each project partner has a duty to disseminate its results unless it can show legitimate reason for not doing so and must give the other project partners advance notice that it intends to do so. Project partners must also share sufficient information on the results they plan to disseminate, giving at least 45 days' notice.

Each of the partners can decide in what repository they store the data packages that came out of their specific research activities, as long as the packages have DOI assigned.

The data packages can be searched for a found in regular data catalogues like GoogleScholar, DataCite Commons and OpenAir.

### 2.3. Making data interoperable

For general principles, see above section '**Consortium wide principles for data management, ethics and privacy; D) FAIR data**'.

Given the nature of the expected data resulting from this research project, data interoperability is not to be expected.

### 2.4. Making data re-useable

For general principles, see above section '**Consortium wide principles for data management, ethics and privacy; D) FAIR data**'.

All archival data packages resulting from this research project will in principle be offered under Open Access. Exceptions will be any set of raw survey data. If these are being processed in archival storage that this will be as Closed or restricted Access Datapackages. However, it is expected that anonymized versions will be available for reuse.

There are no immediate costs anticipated to make the datapackages produced FAIR. The data will be deposited in the repository for at least 5 years after the conclusion of the project. Any unforeseen costs related to open access to research data in Horizon Europe are eligible for reimbursement during the duration of the project under the conditions defined in the Grant Agreement (see Article 14 and Annex 5).

Associate Professor Gertjan Plets (project leader and PI UU) is responsible for data management within the project, specifically for creation of the data management plan and updating the data management plan and ensuring the datasets are recorded. The PI of each partner will have overall responsibility for implementing the data management plan.

Each partner should respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with European Commission and local legislation. Dataset validation and registration of metadata and backing up data for sharing through repositories is the responsibility of the partner that generates the data in the WP.

The datasets in the repository will be preserved in line with the European Commission Data Deposit Policy. The data will be preserved indefinitely (minimum of 5 years).

#### **4. DATA SECURITY**

For general principles, see above section '**Consortium wide principles for data management, ethics and privacy; D) FAIR data**'.

Survey data will be anonymized as soon as possible.

For the duration of the project, data will be stored on the institutional storage infrastructures, using the institutional security settings. It is the responsibility of the partners to ensure that these are on a level which can be expected from a modern research institute.

Following completion of the project, all responsibilities for data recovery and secure storage will go to the repository storing the dataset, Data will be archived and preserved in the Yoda repository. This provides options for making data openly available and other data restricted access as required.

#### **5. ETHICAL ASPECTS**

The project partners will comply with the ethical and data protection principles outlined in Articles 14 of the Grant Agreement. All activities must be carried out in compliance with:

- Ethical principles, and the highest standards of research integrity, as set out, for instance, in the European Code of Conduct for Research Integrity (European Science Foundation, 2011), including avoiding fabrication, falsification, plagiarism or other research misconduct);
- applicable international, EU and national laws; and
- ensuring that informed consent is obtained for data sharing and long-term preservation of datasets acquired by questionnaires dealing with personal data.

#### **6. OTHER**

Project partners must adhere to their own institutional policies and procedures for data management in addition to those policies stipulated by the European Commission.