



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

# Guidelines for depositing datasets and software in AMS Acta

Document prepared by **AlmaDL**

AlmaDL – University's Digital Library Management and Development Unit  
Cultural Heritage Division (ARPAC)  
Alma Mater Studiorum – Università di Bologna

**Version 1, December 2025**

# Summary

<b>INTRODUCTION.....</b>	<b>1</b>
<b>PRELIMINARY ASSESSMENTS PRIOR TO DEPOSIT IN AMS ACTA .....</b>	<b>1</b>
CHOICE OF TYPE OF RESEARCH RESULT EXPLOITATION .....	1
OWNERSHIP OF RIGHTS FOR THE DEPOSIT .....	1
PERSONAL AND SENSITIVE DATA .....	2
CHOICE OF ACCESS LEVEL AND OF LICENCE OF USE FOR ITEMS .....	2
<b>PROCEDURE FOR DEPOSITING ITEMS TO AMS ACTA.....</b>	<b>3</b>
ACCESS TO AMS ACTA AND ACCEPTANCE OF THE TERMS OF USE OF THE SERVICE .....	3
MANAGEMENT OF ITEMS AND VALIDATION FLOW .....	3
<b>PREPARING ITEMS FOR THE DEPOSIT IN AMS ACTA.....</b>	<b>4</b>
METADATA REQUIRED FOR IDENTIFICATION OF ITEMS .....	4
<i>Title.....</i>	<i>5</i>
<i>Authorship responsibilities and affiliations.....</i>	<i>5</i>
<i>Contact person and contact email .....</i>	<i>6</i>
<i>Date .....</i>	<i>6</i>
<i>Abstract .....</i>	<i>6</i>
<i>Licence for use .....</i>	<i>7</i>
<i>Recommended citation and persistent identifiers (PID).....</i>	<i>8</i>
<i>Funding.....</i>	<i>8</i>
<i>Related resources .....</i>	<i>8</i>
SPECIFICATIONS FOR DATASETS AND SOFTWARE .....	8
<i>From data to datasets .....</i>	<i>9</i>
File formats .....	9
Naming files and folders and structuring of datasets .....	11
Organisation of data within files .....	12
Data & Code .....	12
Volume of the files of the dataset.....	13
‘Underlying data’ linked to a publication .....	13
Supporting documentation to be attached to the data (README file) .....	13
<i>Software .....</i>	<i>14</i>
Which Items can be deposited as 'software'? .....	14
Documentation to be deposited together with the Software .....	14
<b>CHECKLIST FOR DEPOSITING DATASETS .....</b>	<b>15</b>
<b>CHECKLIST FOR DEPOSITING SOFTWARE .....</b>	<b>16</b>
<b>FACSIMILE TEMPLATES &amp; GUIDE .....</b>	<b>17</b>
<b>DOCUMENT HISTORY .....</b>	<b>17</b>



This work is licensed under [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

## Introduction

[AMS Acta](#) is the institutional repository of the Alma Mater Studiorum – Università di Bologna for **the preservation, publication and dissemination of data, software and unpublished research outputs**, in line with FAIR and Open Science principles, in compliance with national and international standards.

The Users and document types admitted to the deposit service are defined in the [Service charter for the deposit of data, software and unpublished research outputs in AMS Acta](#) [Service Charter].

These Guidelines cover research data and software [Items] and in particular illustrate:

- aspects to be assessed before proceeding with the deposit;
- the AMS Acta login procedure and acceptance of the Terms of Use of the Service;
- the work area for the management of Items and validation procedures;
- the metadata required for identification of the Items;
- the specific requirements of the Items for admission to the repository;
- a checklist for the deposit;
- examples and facsimile templates (e.g. README files, organisation and names of files and folders)

## Preliminary assessments prior to deposit in AMS Acta

### Choice of type of research result exploitation

**Open access dissemination must be preceded by an assessment of the potential for economic, commercial or industrial exploitation of the research results.** Where protection [of intellectual property] or competitive exploitation may appear, publication in open access may take place only after the adoption of appropriate protection measures, guaranteeing the correct balance between the principles of transparency, knowledge sharing and safeguarding the strategic and institutional interests of the University of Bologna.

⇒ More information on potential competitive exploitation and protection measures can be found on the intranet information pages dedicated to [protection of intellectual property and exploitation of research results](#) or contact the **Knowledge Transfer Office (KTO) Process Unit** ([kto@unibo.it](mailto:kto@unibo.it)).

### Ownership of rights for the deposit

The responsibility for the Items deposited and disseminated through the repository rests with **the person carrying out the deposit, who must ensure that they have sufficient rights or have acquired the necessary permissions to authorise the University to preserve the Items and to disseminate them in open form.**

The dissemination of Items **must not infringe any third-party's interests**, such as, for example, intellectual property rights that may relate to content subject to copyright, confidential data, trade and

industrial secrets, inventions subject to patent applications, project background knowledge. Items must also not infringe any rights to the **protection of personal and sensitive data**.

## Personal and sensitive data

**AMS Acta may not be used to share Items that contain personal data and/or sensitive data** that can be traced back to the data subject without their express authorisation. The **information notices and informed consent forms** used for data collection must therefore expressly request the data subjects' permission to publish the data at least in an anonymous and/or aggregate form. It is also necessary to **remove any indirect element that allows identification from the anonymised and/or aggregated data** prior to the deposit.

⇒ For more information on current legislation and support in preparing informed consent forms, please refer to the relevant pages on the intranet dedicated to [privacy in research](#) and the [processing of personal data for scientific research purposes](#) or contact the 'Personal Data Protection' Professional Unit ([privacy@unibo.it](mailto:privacy@unibo.it)).

## Choice of access level and of licence of use for Items

Once the Items that may be deposited in the repository are identified, the most appropriate level of access and licence for use must be selected for each compressed file or folder, considering any constraints with third parties or obligations arising from University policies and research funding bodies. The access levels supported by AMS Acta are: **open access, open access with an embargo, access restricted to administrative staff only (AlmaDL)**.

It is recommended, whenever possible, that open access be selected so that deposited Items may be immediately accessible to any third-party user according to the principles of Open Science: *as open as possible as closed as necessary*.

In limited and justified cases, **a temporary restriction (embargo) may be applied to the files** composing the Items. For instance, if you deposit a dataset linked to an article that is in course of publication, you can set an embargo while waiting the official release of the publication.

Datasets with restricted access, and thus not public, are only permitted for upload provided that the explanatory documentation attached to the dataset (README file) contains a statement defining the conditions for the availability of the Items and the email address of a contact person.

**When published under restricted access or access with an embargo, Items are only accessible to the User who made the deposit and to AlmaDL staff.** In these cases, users of the repository that are interested in the research results can write to the relevant contact persons for the Items through an online form that can be viewed via the 'Request Copy' button on the public bibliographic record of the Item. Any requests for access should be handled directly by the contact persons indicated for the Items. **AMS Acta does not support restricted access modes for selected users.**

In addition to the level of access, **each file uploaded to AMS Acta must be given a licence for use** that defines the terms and conditions that users of the repository must comply with when re-using the Items. When choosing a licence for use, any restrictions determined by agreements with third parties and the obligations of research funding bodies should be taken into account. See the dedicated section '[Licence for Use](#)' for more information.

## Procedure for depositing Items to AMS Acta

### Access to AMS Acta and acceptance of the Terms of Use of the Service

Each User, in order to start the procedure for depositing an Item, must log in with their University credentials through the [AMS Acta homepage](#), by clicking on the link "**Login**".

When depositing a new Research Item for the first time, the [Terms of Use of the Service](#) must be read and accepted online. Acceptance becomes effective for all Items deposited in AMS Acta, including those that follow the first. **Failure to accept will not allow the upload and publication of Items through AMS Acta.**

⇒ For support, please contact AlmaDL staff by writing to [almadl@unibo.it](mailto:almadl@unibo.it).

### Management of Items and Validation flow

Once you have logged in for the first time and accepted the [Terms of Use of the Service](#), access is granted to your own work area which is named '**Manage documents**'. In this section, you can activate a new deposit procedure by clicking on the '**New document**' button, modify Items that have already been uploaded but not yet deposited, and view your previously published Items.

Each Item corresponds to an online bibliographic record [**Record**] containing the information needed to describe it uniquely [**metadata**] and the link to the files it comprises. Each Record is identified by a progressive number automatically assigned by the Repository [**ID**]. The Record is compiled through guided masks for entering metadata and uploading the files or compressed folders comprised in the Item [**metadata mask**].

In the work area, it is possible to check and manage the status of the Records created. Records can be assigned the following statuses:

- '**User inbox**'. The Item is in draft form, has yet to be formally deposited and is only accessible to the User.
- '**Being revised**'. The Item has been deposited by the User and is under revision [Validation] by AlmaDL. This status means the Item can no longer be modified by the User.
- '**Online archive**'. The Item has been validated and published by AlmaDL and its Record is visible and accessible from the web. In this case, the line relating to the Item will be highlighted in green.
- '**Deleted**'. The validated and published Item has been removed.

Items in the user's work area can be modified until they are deposited. Once they are deposited, they can only be modified by AlmaDL before they are published. The deposited Items are not automatically published in the online repository, but are transferred to a review area for formal checking [**Validation**] by AlmaDL.

**The purpose of the Validation phase is to verify the accuracy and completeness of the descriptive information provided by the Users under their own responsibility during the deposit procedure and compliance with this Guide. AlmaDL reserves the right to request amendments and additions to the Items where necessary.** At the end of the Validation phase, AlmaDL notifies the User that the Item deposited in AMS Acta has been published.

Items that have been validated and published in the online repository can no longer be modified by the User. The User may, if necessary, create new versions or copies of the deposited Item by accessing the respective features ('**New version**' and '**Use as template**') via the option "**View the document**" which is available for each Item



and selecting then '**Change**'.

The 'Removal Request' feature is also available in the work area, however, **validated and published Items can only be removed under the conditions set out in the [Service Charter](#).**

## Preparing Items for the deposit in AMS Acta

### Metadata required for identification of Items

Each Item deposited in AMS Acta must contain within it, in a clear and unambiguous manner, specific minimum information allowing it to be identified without any ambiguity. The **key information to be included within each Item, and to be reported in the AMS Acta metadata mask**, is as follows:

- title
- authorship responsibilities and affiliations
- Contact person and contact email
- date
- abstract
- licence for use
- persistent identifier (PID) and recommended citation
- specification of any funding, research project, patronage or sponsorship
- references to any resources linked to the Item

## Title

**The title is a mandatory descriptive element.** The title should refer to the topic or subject concerned in the Item and enable its identification.

The title should be written in natural language and should possibly follow the conventions and punctuation of the language in which it is written<sup>1</sup> and of the scientific community to which the Item is addressed.

**The title must be unique, and must not be the same as that of any other resources linked to the Item.**

For example, if a dataset is associated with a scientific publication that presents and discusses the data contained in the dataset, the title and abstract of the dataset must be different from those of the publication.

⇒ In the AMS Acta metadata mask, remember **not** to insert the full stop at the end of the title.

## Authorship responsibilities and affiliations

**Authors** [Authors or Author] are the persons or institutions that had a primary and direct authorial role in the creation of the Item and are specified as such within it.

**Contributors** are persons or institutions that had a secondary authorial role in the creation of the Item, contributing to different parts or stages of its realisation, for example, in the collection, editing, production and storage of data. For an overview of the types of Contributor supported by AMS Acta, please refer to the [‘contributorType’ section of the Datacite guidelines](#).

For each Author or Contributor, the affiliation must be specified. The wording for the affiliated institutions, where possible, should be the official or most common wording used. For a list of standard wording, please consult the list of organisations in the [Research Organization Registry \(ROR\)](#). It is recommended to also indicate the [ORCID](#) identification code for each Author or Contributor.

**Authorship responsibilities are mandatory metadata fields** and must be present in the files composing the Item. It is not possible to include any references to authorship in the AMS Acta Record that are not expressly stated in the Item.

⇒ In the AMS Acta metadata mask, it is possible to specify the role of the Contributor by selecting it from a controlled list, based on the Datacite Metadata Schema.

In addition to the affiliation, it is also appropriate to include the ORCID identification code for each Author or Contributor.

---

<sup>1</sup> Titles in Italian are in lower case, the title and subtitle are separated by a dot. For example, *Opera aperta. Forma e indeterminazione nelle poetiche contemporanee*. In English, titles usually follow the 'title caps' convention, according to which all words in the title begin with a capital letter, except for articles and prepositions; the subtitle is separated from the title with a colon. For instance, *Orientalism: Western Conceptions of the Orient*. AMS Acta does not provide guidance on the style to be adopted, leaving it to the author to choose the most appropriate stylistic conventions to be used.

For each Author, it is also possible to specify the email address. It is advisable to specify the institutional university email address or, if the person does not have a permanent role in the University, a personal email address that is maintained over time. The emails specified are not displayed in the public interface of AMS Acta.

## Contact person and contact email

The email address of the person(s) to be contacted for any information or matters concerning the Item [**Contact Person**] must be specified inside the Item or in the accompanying explanatory documentation.

If you are depositing a dataset or software, please indicate one or more Contact Person(s) in the README file or other attached documentation.

For permanent University employees, we recommend using the institutional university email. For Authors or Contributors affiliated to the University of Bologna but not permanent staff (for example, research fellows, PhD students, etc.) it is recommended they also provide an email address that is maintained over time.

⇒ The Contact Persons indicated in the README file must be described as 'Contributor' with assigned the role type 'Contact Person' in the AMS Acta metadata mask.

In the 'contact email' field, only one email referring to one of the Contact Persons can be indicated. This email enables the 'Request a Copy' feature to be activated when access to files is temporarily or indefinitely restricted. The feature allows you to contact the email address specified as the contact via an online form. The email is not displayed in the form or in the public interface of AMS Acta.

## Date

**The date is mandatory information.** It is possible to specify whether this is a date of creation or draft date or a release or publication date.

Within the Item or in the accompanying documentation, it is possible to specify all the different types of date; in the AMS Acta metadata mask, you must instead indicate the one you consider most significant.

The date can be in the formats YYYY, MM/YYYY, DD/MM/YYYY. For example '2025', 'July 2025', or '11 July 2025'.

⇒ Please note that **the date indicated in the metadata mask will be reported by AMS Acta in the bibliographic citation of the Item.**

## Abstract

**The abstract is a mandatory metadata field** and represents a brief and structured summary of the Item, presenting its essential contents in a clear and immediate manner. Its function is to quickly inform the



reader about the objectives, context, methodology, main results and impact of the research. It must be independently understandable, concise and accurate, facilitating indexing and retrieval of the Item in databases and bibliographic databases.

When depositing data that form the basis of a publication (so-called 'underlying data'), it is recommended to include an abstract focusing on the description of the content of the dataset without reproducing that of the publication.

## Licence for use

Each file or compressed folder of an Item **must be assigned a licence for use** that clarifies the terms and conditions of use of the content of the Item by third parties.

As mentioned above, **when choosing a licence, Authors must take into account the nature of the data or content and any constraints to which it is subject**, for example, third-party rights such as intellectual property rights or privacy, contractual clauses signed with funding bodies, confidentiality agreements, agreements requiring the adoption of restrictive licences.

For example, the European research funding programme, Horizon Europe, requires that data from funded projects be released under Creative Commons licences [CC BY](#) and [CCO](#) (or equivalent), unless there are special constraints or other restrictions that must be duly documented.

However, as far as software is concerned, the libraries, proprietary or non-proprietary, used for development may place constraints on the choice of licence. It is essential to make sure that there are no conflicts regarding the licence chosen.

It is recommended to use standard licences such as the Creative Commons also supported by AMS Acta.

It is necessary to **indicate the chosen licence within the Item files by means of a declaration**. For example, in the README file of the dataset, the Creative Commons licences may be cited as follows: *'this dataset is distributed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, <https://creativecommons.org/licenses/by/4.0/>'*.

⇒ In the metadata mask, **a licence must be specified for each compressed file or folder uploaded to AMS Acta**. AMS Acta allows you to choose different types of licence, from the more open licences, such as Creative Commons CC0 and CC BY<sup>2</sup>, to the more restrictive licences, which only allow use for personal, study and teaching purposes.

In the case of deposit of a Software, if the licence you wish to associate with the Item is not implemented by the repository and therefore cannot be found in the metadata mask, please contact AlmaDL to assess the possibility of inclusion.

---

<sup>2</sup> More information on Creative Commons licences can be found at <https://creativecommons.org/share-your-work/cclicenses/>.

## Recommended citation and persistent identifiers (PID)

It is advisable to provide within the Item a suggestion for citing the Item itself using a codified citation style used in the relevant disciplinary area<sup>3</sup>.

At least one persistent identifier such as a DOI must be specified in the files constituting the Item. If the Item being deposited does not already have an attributed DOI, AMS Acta can automatically associate one in the form **10.6092/unibo/amsacta/ID**, where ID is the identification number assigned when the Record was created.

## Funding

**If the Item is the result of funded research, a reference to the funding programme and the research project must be included in the Item files**, in the format required by the funding body, including, where necessary, any symbols or logos, declarations and the link to the project website. For example:

*'This research was funded by European Union's Horizon Europe Programme, grant number 1234567890'*

*'Project info: MADFORWATER (DevelopMent AnD application of integrated technological and management solutions FOR wasteWATER treatment and efficient reuse in agriculture tailored to the needs of Mediterranean African Countries), funded by European Union, Horizon 2020 Programme. Grant Agreement no. 688320, DOI <https://doi.org/10.3030/688320>; <https://www.madforwater.eu>'* (in case of datasets or software)

⇒ In the AMS Acta metadata mask, it is possible to capture or manually enter information concerning research funding.

## Related resources

**It is recommended to include in the README file of a dataset or software any references to related information resources.** Citing academic publications, preregistrations, code repositories or supplementary materials complete with their persistent identifiers such as a DOI or HANDLE allows for better contextualisation of the Items, facilitating their understanding, reusability and proper citation. This enhances the transparency of research, promotes the traceability of results and increases the overall impact of the work, facilitating collaboration between scientific and interdisciplinary communities.

⇒ In the AMS Acta metadata mask, it is possible to enter the persistent identifiers of linked resources by specifying the nature of the relationship. For an overview of the types of 'relations' supported by AMS Acta, see the ['relationType' section of the Datacite guidelines](#).

## Specifications for datasets and software

The following section provides guidance on how to structure datasets and software to be deposited.

---

<sup>3</sup> Codified citation styles. Example dataset citation in APA format: Amurri, A., Giachino, E., Misuraca, E., Peroni, S. (2025). *UNIBO IRIS Bibliographic Data Dump, Dated 30 May 2025, Updated on 3 July 2025* [Dataset]. University of Bologna. <https://doi.org/10.6092/unibo/amsacta/8427>

## From data to datasets

Sharing research data, whether or not they support scientific articles, is considered a foundation for the realisation of Open Science principles to ensure transparency, reproducibility and integrity in research practices and scientific communication. Some research funding bodies, such as the European Commission or the Wellcome Foundation, require recipients of funding to share research data in the form of datasets.

**A dataset is a collection of related data, collected within a well-defined structure. A dataset may contain data of any type (raw, processed, etc.) or nature (numerical, textual, images, etc.) and the criteria for aggregating them may be multiple:** same research purpose; same techniques or methodologies used to produce them; same geographical area of reference; same subject or conceptual class; same relevant research phenomenon; same author or group of authors, etc.

Data are often, but not necessarily, encoded in numerical format. Annotated bibliographies, interview data, image or audiovisual collections or other types of evidence may also be collected, deposited and disseminated in the form of datasets.

**The datasets must include supporting documentation** (for example, README files, guides, codebooks, questionnaire templates, experimental protocols, etc.) **and any code prepared *ad hoc* to process and/or analyse the data** (for example, scripts in Python, R or other programming languages designed to merge data from different sources, 'clean' and/or normalise the data, perform statistical analyses, visualise the data, etc.).

⇒ For more guidance on best practices in research data management, see the [University Guidelines for Research Data Management](#)<sup>4</sup>; for support in the planning phase for research data management and the drafting of the Data Management Plan contact the [Data Stewards](#) of the University of Bologna ([aric.datasteward@unibo.it](mailto:aric.datasteward@unibo.it)).

### *File formats*

The files that make up the Item must be saved in formats that allow for accessibility, re-use and interoperability of the data in the long term.

**Formats based on well-documented and commonly used standards, possibly open, unencrypted and uncompressed** are therefore to be preferred, for example, simple text files encoded in standard formats (UTF-8, Windows-1252, etc.) and .csv files for tables.

---

<sup>4</sup> Alma Mater Studiorum - Università di Bologna (2024) University guidelines for research data management. Bologna: University of Bologna, DOI <https://10.6092/unibo/amsacta/7835>.

⇒ For an overview of the most recommended formats for preservation see for instance the pages of the UK Data Service<sup>5</sup>, the Library of Congress<sup>6</sup>, the Dutch national repository of research data (DANS - Data Archiving and Networked Services)<sup>7</sup> and the Agency for Digital Italy (AgID)<sup>8</sup>.

**If it is not possible to use standard open formats, it is good practice to use proprietary formats that are widespread and known in the community.** For instance, if open formats such as .odt/.txt or .ods/.csv cannot be used, .docx and .xlsx formats can be used, even if they are proprietary, as they are widely used.

If, on the other hand, the proprietary format is not widespread, the **file must be converted to an open standard format or, at the very least, to another proprietary but widely used standard format.** Only in the event of conversion problems can the original format of the data be retained, attaching the documentation relating to the applications necessary for reading and using the file.

Data produced by instruments and stored in proprietary, instrument-specific formats must be converted to standard (open) formats or to widely known proprietary formats. The software of the instruments often has functions for exporting data in multiple formats, and it is advisable to check whether the exportable formats include those suitable for storing the type of data being deposited. Alternatively, **it must be specified in the accompanying documentation whether files in proprietary format can nevertheless be opened with free or open source software. Any code developed ad hoc for using the data must be included in the dataset.**

It is possible to provide data within the same dataset in several formats, for example, in a proprietary format (.xlsx), while also providing a conversion to a standard open format (.csv).

**Data must also be preserved in the format most appropriate to their nature.**

For example, **numerical tabular data** should not be provided as a table in a text document such as .odt or .docx, but in a format more suitable for tabular data (.csv, .ods or .xlsx).

Similarly, for **extended bibliographies**, structured formats are preferable to lists of citations in text documents. For instance, through the use of tables (.csv, .ods or .xlsx), in which each column corresponds to information on the individual publication cited (e.g. 'authors' column, 'title' column, 'publisher' column, etc.), or in specific formats for bibliographic reference management (e.g. .bib and .xml).

In the case of **graphs, histograms**, etc., the data to be preserved is not the image but the numerical data presented graphically within them.

---

<sup>5</sup> UK Data Service Recommended formats, <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>

<sup>6</sup> Library of Congress Recommended Formats Statement 2024-2025, <https://www.loc.gov/preservation/resources/rfs/TOC.html>

<sup>7</sup> DANS file format page, <https://dans.knaw.nl/en/file-formats/>

<sup>8</sup> AgID. Open Data Guidelines. Annex B - Relevant standards and Open Formats, <https://docs.italia.it/AgID/documenti-in-consultazione/Ig-opendata-docs/it/bozza/allegato-b-standard-di-riferimento-e-formati-aperti.html#allegato-b-standard-di-riferimento-e-formati-aperti>

### *Naming files and folders and structuring of datasets*

It is requested that **uniform file and folder names be adopted** that are in line with internationally promoted best practices<sup>9</sup> and **a consistent structure be used for organising files into folders and subfolders within the dataset**. The conventions and organisation adopted should be **documented as much as possible in the README file** to facilitate data sharing and reuse.

⇒ If the data are related to research projects involving the formulation of a **Data Management Plan**, it is important to follow any naming and structuring rules in said document.

In general, for file and folder names it is recommended to:

- choose short names<sup>10</sup> but with significant content-related elements;
- avoid spaces that separate the elements making up the file name and use, for example, dashes ( - ) and/or underscores ( \_ );
- do not use special characters;
- when numbering, insert zeros at the beginning of the sequence (for example, use '001, 002, ..., 100' instead of '1, 2, ..., 100');
- indicate dates in YYYYMMDD or YYMMDD format, etc.

**Several compressed files or folders may be deposited for each dataset.** For the repository to maintain usability, it is advisable to group large numbers of files into an appropriate number of compressed folders, provided that the volume of the compressed folders is in line with the indications in the '[Volume of the files of the dataset](#)' section.

The internal structure of the compressed folders can consist of various levels of subfolders ('nested' structure). **It is recommended not to adopt a deeply nested structure** (more than 4-5 levels of subfolders) and to avoid creating compressed folders within other compressed folders.

⇒ When including files in compressed folders, it is best to select the individual files together and give the command to compress them instead of putting them all in one folder and then compressing that folder. Otherwise, a redundant intermediate folder is created which increases the nesting level and unnecessarily lengthens the file path.

---

<sup>9</sup>Some examples of internationally promoted guidelines:

UK Data Service (<https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/organising/>);  
CESSDA Training (<https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/File-naming-and-folder-structure>);

Stanford Libraries training material (<https://guides.library.stanford.edu/data-best-practices/name-files>);

Harvard file naming guidelines (<https://datamanagement.hms.harvard.edu/plan-design/file-naming-conventions>)

<sup>10</sup> Please keep in mind that in major operating systems, the maximum allowed path length of a file (including the length of the file name) is 255-260 characters, and that the more the dataset structure is nested in folders and subfolders, the longer the path to the files in the subfolders will be. Accordingly, it is advisable to keep the length of the file names to within around 30-40 characters whenever possible.

Within zipped folders generated in Mac operating systems there are often residual system folders and files named '.\_\_MACOSX' and '.DS\_Store', which are hidden on Mac systems but can be viewed on Windows systems. It is advisable not to include such folders in the deposited material. In this regard, see for example the [instructions on the Stack Overflow forum](#). Similarly, it is a good idea not to include in the zipped folders deposited the hidden files of the type 'Thumbs.db' generated on Windows systems.

AlmaDL provides a short guide for file and folder naming and structuring into folders and subfolders (link in the “[Facsimile templates & Guide](#)” session).

### *Organisation of data within files*

The data within the files must be clearly displayed and understandable.

**In the case of numerical data, it is necessary to specify**, directly in the file itself or at least in the [README file](#), **the units of measurement and any abbreviations and acronyms they are associated to**. If data are presented in spreadsheets (for example, .ods or .xlsx), the use of **colours or other text formatting** is not recommended. Any groupings or selection parameters must be added in a new column.

It is not advisable to submit spreadsheets (.ods, .xlsx) with **active formulas**, as they may not be compatible between different software or between different versions of the same software. In such cases, in the spreadsheet it is necessary to present only the numerical values of the data that is calculated using the formulas and to provide a detailed description separately of the formulas applied to obtain these values<sup>11</sup>.

It is also not advisable to submit spreadsheets containing **macro codes**<sup>11, 12</sup>, due to compatibility and security issues. In such cases, it is advisable to keep the commented macro code in a separate text file.

It is also not advisable to include **graphical representations of the data (for example, graphs and histograms)** in the spreadsheet. If it is deemed necessary to support the data with this type of visual information in order to increase its ease of understanding, usability and re-use, it should be attached in a separate file within the dataset folder in image format (e.g. .png, .jpeg, etc.).

If data are presented within a publication and are shown in certain figures and tables, **do not use article references e.g. 'figure 1, 2, ...' in the dataset**, but rather describe the type of data and the units in which they are expressed. As a matter of fact, it is essential that the content of the dataset is understandable even without viewing an associated publication.

### *Data & Code*

When the data is the result of processing by means of a code, **the dataset should contain the source code of the scripts used for processing, the input data and the processed data**, especially if the processing is difficult to reproduce or computationally intensive.

---

<sup>11</sup> Artefactual Systems and the Digital Preservation Coalition (2021) *Preserving Spreadsheets. Data Types Series*. Digital Preservation Coalition. DOI: <http://doi.org/10.7207/twgn21-09>.

<sup>12</sup> Macro code. Code generally written in Visual Basic or Javascript that allows repetitive actions to be performed within the spreadsheet.

If the input data already exist and are freely accessible in repositories or other external platforms that ensure their preservation over time, they should be cited within the README file and in the metadata field 'Related Resources' of AMS Acta, specifying the appropriate relationship.

#### *Volume of the files of the dataset*

AMS Acta imposes no limits on the number or size of compressed files and folders to be uploaded per Record. **With regard to the size of the individual file or compressed folder, it is generally suggested not to exceed 3GB.**

However, if the total volume of data to be submitted is very large, for example, it exceeds 30 GB, then **compressed folders of larger volumes, up to 10 GB, may be deposited.** In these situations, you are requested to contact AlmaDL prior to depositing.

In the case of compressed folders, it is good practice to specify the algorithm and level of compression used in the documentation attached to the dataset.

#### *'Underlying data' linked to a publication*

If data are associated with and presented in a publication (so-called '**underlying data**'), the title and abstract of the dataset and publication must differ. The data and the publication, although related, are in fact distinct 'objects'. The abstract must focus on describing the content of the dataset and the title must reflect the content of the dataset.

The 'underlying data' may have authorship responsibilities that are different to those of the publication.

It is recommended to submit the 'underlying data' before the official release of the publication they refer to in order to include the citation of the dataset complete with the DOI attributed by AMS Acta in the publication itself.

Furthermore, **adding the DOI of the deposited dataset to the 'FAIR data' section present in the bibliographic record in IRIS of the research publication is recommended**, in order to link the publication described in IRIS with the relevant data.

#### *Supporting documentation to be attached to the data (README file)*

Each dataset deposited in AMS Acta must be supported by **documentation files that enable the user to understand its content and the methodologies used to collect and produce the data, without having to ask the authors for explanations.**

AlmaDL requires at least one README file to be attached to the dataset which, in addition to the basic metadata (title, authorship responsibilities, date, licence, information on possible funding and related resources), contains information on:

- content of the dataset (files making up the dataset, their name, format and possible structuring into folders and subfolders);
- detailed description of the dataset
  - abstract;



- description of the content of individual files;
- instructions/technical information on file formats and software. For example, indication of the software needed to open files or to convert them;
- information on data sources;
- methodology by which the data were collected and generated, if necessary for re-use;
- meaning of abbreviations and acronyms, if present

AlmaDL provides a README file template to be completed and attached to the data (see the “[Facsimile templates & Guide](#)” session).

⇒ **Some of the information to be indicated in the README file coincides with the information to be entered in the metadata mask in AMS Acta at the during the deposit and must therefore match.**

**The supporting documentation to be attached to the data is not limited to the README file**, but may also include other types of documents such as guides, codebooks, questionnaire templates, experimental protocols, scripts, etc.

## Software

### *Which Items can be deposited as 'software'?*

In AMS Acta, software can be deposited both in the form of **executable files (.exe)** and in the form of **source code**.

⇒ If the Item to be deposited consists of one or more scripts for processing/analysing data and needs to be deposited together with input and output data, it is suggested that this Item be deposited as a dataset and not as software.

The following guidelines for depositing software in AMS Acta are in line with internationally recognised and adopted best practices<sup>13</sup>.

### *Documentation to be deposited together with the Software*

In AMS Acta, the deposited software must be accompanied by at least two basic files: README file and LICENSE file.

In the **README file**, in addition to the metadata necessary for the identification of the Items (title, authorship responsibilities and affiliation, contact information, date, abstract, licence for use, information on possible funding, recommended citation and persistent identifiers, linked resources), information must also be provided on:

- software version and any notes relating to it (e.g. change logs and bugfixes);

---

<sup>13</sup> As an example, refer to the guide for using Software Heritage to archive research software: Roberto di Cosmo (2020) 'How to use Software Heritage for archiving and referencing your source code: guidelines and walk-through', <https://annex.softwareheritage.org/public/guidelines/archive-research-software.pdf>



- runtime environment (the programming language and its version, the runtime platform, the operating systems on which the software can be installed, any dependencies on libraries or other software);
- platform on which the software was developed, e.g. GitHub, GitLab, etc. (if any);
- links to the website and documentation of the software or project that led to its development (if any);
- any description of the main functionalities of the software;
- instructions for compiling/installing the software (or a reference for where to find a file containing this information, typically named INSTALL);
- instructions for using the software (or a reference to documentation files for its use).

The file **LICENSE.txt** must contain the terms of the project licence in full. In the case of open source software licences, it is recommended to use the standard names and licence texts found in the [SPDX License List](#). More information on open source software licences and how to apply them can be found on the [Open Source Initiative](#) website and on [GitHub's licence selection tool](#).

⇒ If the software was developed using **proprietary libraries**, it is essential to ensure that the licence to which you wish to associate it does not conflict with the terms of use to which these libraries are subject.

In line with emerging software preservation practices<sup>14</sup>, which suggest providing citation metadata and other metadata describing the Item in machine-readable format, it is also possible to encode the recommended software citation in Citation File Format ([CFF](#)) and the descriptive metadata in JSON-LD format according to the [CodeMeta](#) schema. The [CFFINIT](#) and [CodeMeta generator](#) tools allow the creation of CITATION.cff and codemeta.json files that can be submitted together with the Item.

## Checklist for depositing datasets

- **Do I have the necessary authorisations to deposit this dataset?**  
see section [Preliminary assessment prior to deposit in AMS Acta](#)
- **Does the dataset contain personal data or data indirectly allowing the identity of the persons involved to be traced?**  
see section [Personal and sensitive data](#)
- **Have I established an access level and selected a licence for use?**  
see section [Choice of access level and of licence of use for Items](#)

---

<sup>14</sup> guide for using Software Heritage to archive research software: Roberto di Cosmo (2020) 'How to use Software Heritage for archiving and referencing your source code: guidelines and walk-through', <https://annex.softwareheritage.org/public/guidelines/archive-research-software.pdf>

- **Is the level of access and the licence in line with the policies of my funding body?**  
see sections [Choice of access level and of licence of use for Items](#) and [Licence for use](#)
- **Are the files in a standard, open and non-proprietary format?**  
see section [File formats](#)
- **If the data is not in an open format, have I attached a conversion of the data in an open format?  
Have I provided guidance on the software to be used to open files?**  
see section [File formats](#)
- **Do file and folder names contain no special characters, follow a convention and are organised in a documented structure?**  
see sections [Naming files and folders and structuring of datasets](#) and [Organisation of data within files](#)
- **If the dataset accompanies a publication, have I attributed the required descriptive metadata and am I able to insert the citation in the publication?**  
see sections [Recommended citation and persistent identifiers \(PID\)](#), [Related resources](#), and ['Underlying data' linked to a publication](#)
- **Have I prepared and attached at least one documentation README file with the metadata and information required by the repository?**  
see section [Supporting documentation to be attached to the data \(README file\)](#)

## Checklist for depositing software

- **Is the Item I have to deposit a software?**  
see sections [Data & Code](#) and [Which Items can be deposited as 'software'?](#)
- **Do I have the necessary authorisations to deposit this software?**  
see section [Preliminary assessment prior to deposit in AMS Acta](#)
- **Is the level of access and the licence in line with the policies of my funding body?**  
see sections [Choice of access level and of licence of use for Items](#) and [Licence for use](#)
- **Has the software been developed using proprietary libraries?**  
see sections [Licence for use](#) and [Documentation to be deposited together with the Software](#)
- **Have I complied with the terms of use of any open source modules and libraries I have used?**  
see section [Licence for use](#)

- Have I prepared and attached at least a documentation README file with the metadata and information required by the repository and a LICENSE file, which contains the full text of the licence?

see section [Documentation to be deposited together with the Software](#)

## Facsimile templates & Guide

Facsimile templates and other guide (e.g. README file, short guide for file and folder naming and structuring into folders and subfolders) are available on [the page describing these deposit Guidelines on the University Library System website](#).

## Document History

20/10/2025	AlmaDL	first draft
30/10/2025	Gruppo di Lavoro Open Science (GLOS)	revisions
06/11/2025	Comitato di Indirizzo del Sistema Bibliotecario di Ateneo	approval
11/12/2025	AlmaDL	pubblication