Deposition Instructions for EASY

May 2017

Data Archiving and Networked Services

DANS

Deposition Instructions for EASY

You are about to deposit data with DANS using the online archiving system EASY. This will ensure long-term archiving, access and reusability of your data.

You will deposit your data in EASY in the form of a dataset. A dataset consists of:
- A description in EASY, the metadata. A *Persistent Identifier* will be assigned by default: this is a unique identification with a permanent link that will always point to the dataset.
- The documentation files describing the research project and the data files.
- The data files.

This guide contains instructions for these three components. It answers questions such as:
- What are your choices within EASY? How do you enter data on-screen?
- What documentation does a researcher need to understand your dataset?
- Which data can and cannot be filed by DANS?

This guide also contains some specific directions for the following disciplines:
- Social and behavioural sciences
- History
- Life science and medicine
- Language and literature
- Archaeology

The guide, licences, terms of use and other information can be found on the DANS web page with data deposition guidelines.[1]

Your data will be archived according to the guidelines of the international Data Seal of Approval[2], the ICSU-WDS[3], and the NESTOR seal for Trustworthy Digital Archives.[4] Moreover, the metadata fields in EASY comply with the guidelines of the Dublin Core standard.[5]

---

1 https://dans.knaw.nl/en/deposit/information-about-depositing-data
2 http://www.datasealofapproval.org
3 https://www.icsu-wds.org/
4 http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html
5 http://dublincore.org

1. Preparing data

Please provide good metadata before proceeding with your deposition.

Data and documentation files

When preparing your deposition, you should select data and documentation files from your research which might be relevant for reuse. The data files must include complete and unambiguous variable and value labels. If you are depositing a large number of files, we would like to receive a file list. It should contain not just a list of file names, but also a description of the contents of the files and any interrelations between individual files.

When other researchers use your data, they should have information about the research design and the way in which the data have been collected and coded, i.e. an explanation of the variables, abbreviations and terminology used in the research project. In many cases, a code book and a publication about the research will be sufficient. Please make sure that your dataset contains this information.

When you deposit data in the discipline of History it is important to add to the dataset a description of the (archival) sources, the selection procedure used, and a description of the way in which the sources have been used. Please also indicate whether the information has been standardized, and which standards or classification systems have been used, e.g. HISCO for historical professional titles.

Data files in the Social and behavioural sciences should contain clear and complete *variable labels* and *value labels*. Essential documentation files are:
- Questionnaires or other research instruments
- The fieldwork report (if available)
- A code book, or a description of variables and information regarding (if applicable):
    - o The population
    - o Type of data (units of observation / analysis)
    - o The sample and the sampling procedure
    - o Response and non-response
    - o The data collection method
    - o Weighting variables
    - o Constructed and/or derived variables
    - o Information on anonymizing
- Publications based on the data (if available) or a bibliographical description of such publications.

Additional information can be found in 'Preparing data for sharing.'[6]

For Language and literature: CLARIN metadata

Have you received a CLARIN grant for your research? In that case, your dataset should preferably contain one or - if necessary - several metadata files in the so-called Component MetaData Infrastructure (CDMI) format (an XML format). If you have received a grant on the condition that you deliver your results in a CLARIN

---

[6] http://www.dans.knaw.nl/nl/over/organisatie-beleid/publicaties/DANSpreparingdataforsharing.pdf

compatible manner, the dataset must also contain CMDI metadata. For further information on CMDI, go to http://www.clarin.eu/cmdi .

Please select the appropriate (second) option on screen 1 in EASY to enable the availability of your CMDI metadata, e.g. for external search engines.

Anonymizing data

Before you upload the files, you should check whether they contain privacy-sensitive information within the meaning of the Dutch Personal Data Protecting Act.[7] If you give access to the data, both data and description must be completely anonymized. This includes personal contact information.

DANS offers the additional option of storing data in their original non-anonymized form on a server which is inaccessible to outsiders. Please contact DANS if you want to make use of this option.

In data for Social and behavioural sciences: Anonymized data files may not contain any variables which can be used to identify individual respondents. Examples of such variables include:

- Name of respondent
- Address of respondent
- Telephone number of respondent
- Respondent's BSN (citizen's service number) or SoFi (social and tax number)

The following variables should be re-encoded:
- Date of birth (re-encoded to year of birth)
- Postcode (re-encoded to four digits)

It is also recommended to re-encode exact occupations to a classification of occupations. You can either apply the Standard Classification of Occupations (SBC[8], in Dutch) used by Statistics Netherlands, or use the International Standard Classification of Occupations (ISCO)[9] compiled by the International Labour Organization (ILO).

Data from Medical Sciences are subject to legislation. Most importantly, patients must not be identifiable. Contact DANS with any questions that you may have.

Archaeologists make use of the e-Depot for Dutch Archaeology (EDNA). EDNA is a subdivision of DANS: EDNA's data and reports are filed in EASY. Specific components of EASY have been added for EDNA, e.g. specific fields for archaeology and a possibility to add metadata using the archaeological packing slip. Archaeological data and research reports must not contain personal contact data.

In principle, all documentation saved in digital form, such as maps, field drawings, photographs, tables and the publication, could be added to the archives. For that reason, all files that may be relevant to archaeological research in the future should be saved in the e-depot.
In addition to the description of the research project as a whole, a (meta)description

---

7 http://wetten.overheid.nl/BWBR0011468/2016-01-01

8 http://www.cbs.nl/nl-NL/menu/methoden/classificaties/overzicht/sbc/2010/default.htm

9 http://www.cbs.nl/nl-NL/menu/methoden/classificaties/overzicht/sbc/2010/default.htm

Deposit Instructions for EASY - May 2017

is necessary for each separate file that is deposited. For this reason, a documented file list (file metadata) and code books (attribute metadata) should be present in the dataset. More information about the metadata to be entered can be found in the document on archaeological metadata.[10]

File formats

DANS provides a list of preferred formats[11] for the files that will be deposited. DANS guarantees the long-term preservation and accessibility of the files that are deposited in a preferred format. To enable us to ensure sustainability we would like to receive data in a preferred format. Please contact a DANS data manager before depositing data when you plan to use a different file format.

File size

Files up to 100 MB can be easily uploaded to EASY. For larger datasets, it is recommended to make a ZIP file. Additional information can be found in 5: Uploading data files. We advise you to contact a DANS data manager before uploading (ZIP) files that are well over 100 MB.

2. Go to EASY

Go to https://easy.dans.knaw.nl and log in. New users will first have to create an account via *Register*.
EASY's user interface is in English to enable data access by non-Dutch researchers. The choice of documenting your data in Dutch or English is yours. See also the Language section under 4. Documentation and access level.
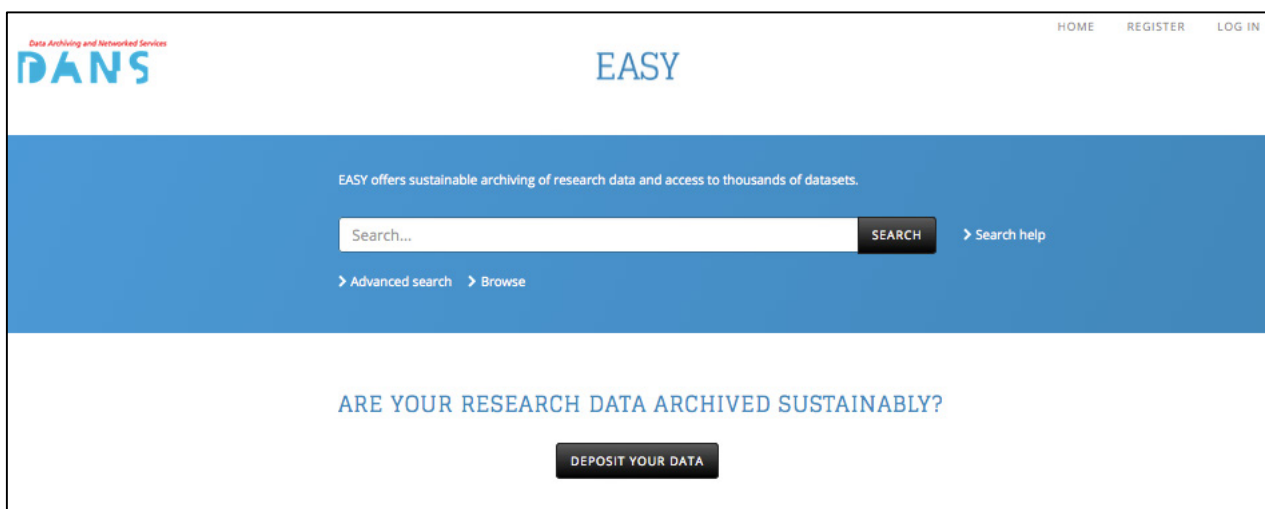

Figure 1: EASY website

---

10 https://dans.knaw.nl/en/deposit/information-about-depositing-data
11 https://dans.knaw.nl/en/deposit/information-about-depositing-data

3. Start the deposit procedure

In EASY, go to *DEPOSIT YOUR DATA* and click *Start deposit* to begin the deposit procedure.
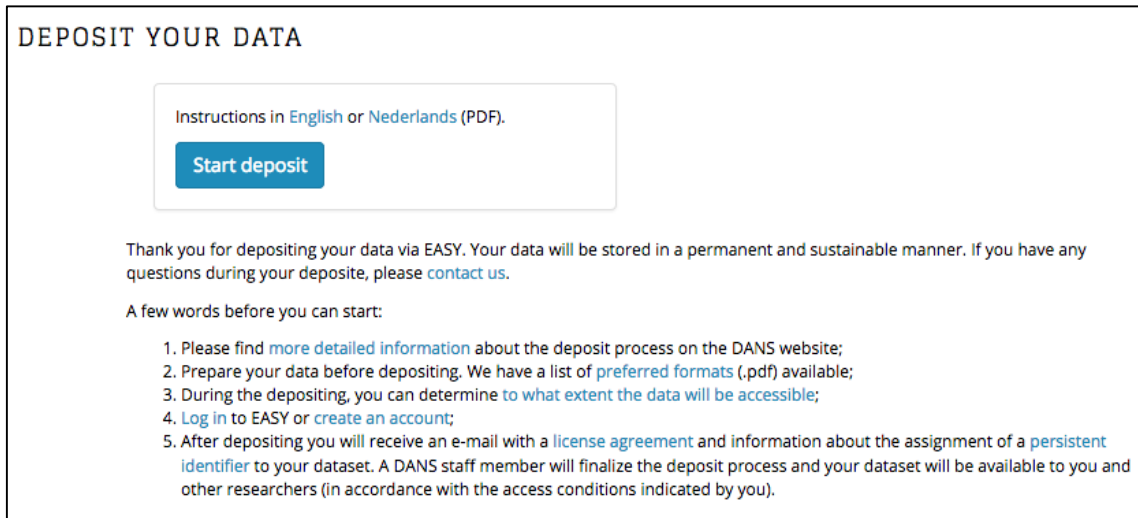

Figure 2: 'Deposit your data' with 'Start deposit'

4. Documentation and access rights

After clicking *Start deposit*, you will see a succession of six screens:

1. Import metadata – importing metadata. At the moment this is enabled only for Archaeology (archaeological packing slip) and Language and literature (CMDI metadata).
2. Primary information – primary information such as title, project leader, etc.
3. Content description – description of the content of your data.
4. Rights – rights, access category, date available, intellectual property
5. Upload files – the uploading procedure for data and documentation files.
6. Overview and submitting – overview of dataset and submitting of data.

These six screens contain eight required fields. Behind the field names of non-required fields, you will see *(optional).* Information in these non-required fields will add to the accessibility, usability and citability of your data.

Language
You can use any language in EASY, but you are advised to use the language that is used for the data. When this is not Dutch or English, e.g. German or French, please add a summary and keywords in Dutch and English to your metadata. For data from the Social and behavioural sciences the language of the *variable labels* and *value labels* is also important: when they are in English the dataset must be described in English. When in doubt contact DANS.
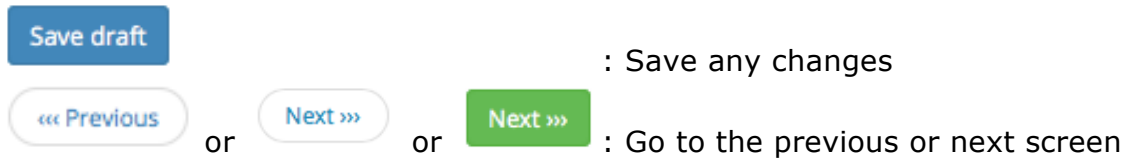
> By clicking the information button next to a metadata field, you open a help screen with additional information on that particular field.

Clicking the plus sign will enable you to add an extra field for the same item.

Save draft

At the bottom of each screen you will find these buttons:

Save draft : Save any changes

‹‹‹ Previous or **Next ›››** or **Next ›››** : Go to the previous or next screen

After each step in the procedure your input is saved. If you want to save your input in between, click *Save draft*. You can then log out of EASY, log in again later and resume the procedure where you left off. You will find your previously entered metadata set as 'draft' under the heading *My datasets* as soon as you return.

On the final screen, 6. Overview and submitting, you can submit the dataset by clicking **Submit** .

Screen 1: Import metadata

DEPOSIT DATASET - IMPORT METADATA (FOR SPECIFIC DISCIPLINES)

DEPOSIT:
1. *Import metadata (for specific disciplines)*
2. Primary information
3. Content description
4. Rights
5. Upload files
6. Overview and submitting

Instructions in English or Nederlands (PDF)

Next ›››

Start deposit

For all disciplines except 'Archaeology' and 'Language and Literature' go to next page.

Specific to Archaeology

Import pakbon (optional)   Browse…   No file selected.   Upload

Specific to the discipline of archaeology. Import (meta-)data from an "Archeologische Pakbon" file. If you do not have one, proceed to next step or page 2.

Specific to Language and Literature

Extra metadata (optional)
⦿ This dataset does not contain CLARIN metadata, i.e., no CMDI files.
○ This dataset contains CLARIN metadata, i.e., CMDI file(s).

Specific to the discipline of Language and Literature. The CLARIN research infrastructure requires data to be well documented. Therefore, if you have received CLARIN funding for the data that you are about to deposit, CMDI metadata (in XML format) should be part of the dataset.

Save draft   **Next ›››**

Next ›››

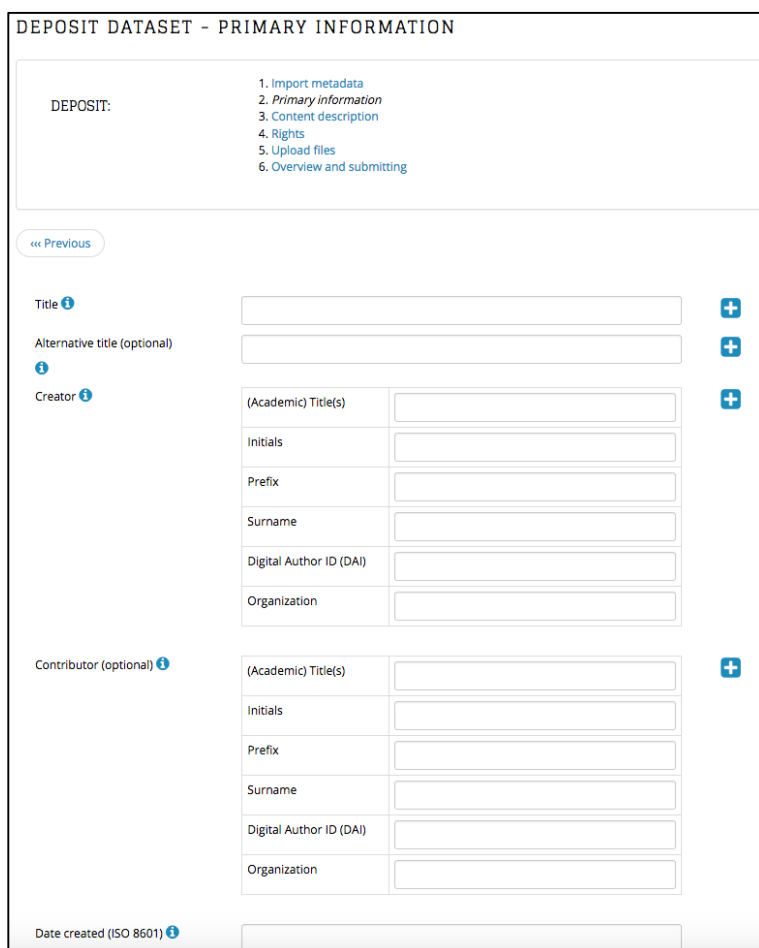*Figure 4: Screen 1 – Import metadata (for specific disciplines)*

Screen 1: Import metadata (for specific disciplines) is specifically designed for importing datasets from:

- Archaeology, with an archaeological packing slip
- Language and literature, with CLARIN metadata. EASY makes CLARIN metadata available to search engines compliant with internal CLARIN agreements.

Archaeologists generally use an archaeological packing slip (SIKB0102 standard).[12] This archaeological packing slip allows archaeologists who have completed their research project to indicate what is deposited where. Find data are delivered to several depositories (including DANS) through a standardized set of descriptions. The XML document can be uploaded here. It will enable EASY to save all metadata from the packing slip to the right field by default. This will minimize manual data entry. The packing slip itself will be added to the dataset as an additional file. More information about the packing slip can be found on the DANS website.[13]

For Language and Literature the field *Extra metadata* is available. Have you received a CLARIN Research Infrastructure grant for your research? In that case, you should preferably upload the CMDI metadata as an XML-file. Please tick the box when your dataset contains CLARIN metadata and upload the CMDI file together with your data files in screen 5: Files.

Screen 2: Primary information

Figure 5: Screen 2 – Primary information

Screen 2: Primary information contains the essential information needed to understand the dataset. You are recommended to complete all fields in screens 2-5 (or as many fields as possible) to enable EASY users to find your dataset.

---

12 http://www.sikb.nl/datastandaarden/richtlijnen/sikb0102
13 https://dans.knaw.nl/en/deposit/information-about-depositing-data

Deposit Instructions for EASY - May 2017

When quoting from a publication, always use the name(s) of the author(s), title, and publication date as references. A similar principle applies to datasets. Enter the name of the primarily responsible researcher (*Creator*: project manager and/or lead author of the research publication) and the name by which the project is known - for archaeology this is the title of the report (*Title*) - as well as the date (*Date created*) on which the research and the accompanying data were completed, so that these data may be used to refer to the dataset. EASY will automatically add a Persistent Identifier to your dataset as soon as the deposit procedure is completed. These data will form the citation after the dataset has been published.

Bakker, Dr. H. (University X)(2017): *Title of the dataset*. DANS.
https://doi.org/10.5072/dans-xxx-xxxx

If desired, you can add several lines to a field using the + button. If you need to name more than one Creator in the description of your dataset – and the citation - you can add lines to the *Creator* field.

Beside the citation data, the *Description* field is also important. In this field the research project can be summarized in one paragraph of free text. If desired, you can extend this description with additional information on other persons and organizations (*Contributor*) that have contributed to the realization of the dataset. It is also possible to document an additional subtitle or alternative title in the *Alternative Title* field. Translating an English title into Dutch will help your dataset appear more often in the search results of other researchers.

Funding
When your project has received funding, the funding organization can be entered in the field *Contributor: Organisation*, for example:

Financer: The Netherlands Organisation for Scientific Research – NWO.

The funding registration number can be entered in the *Identifier* field, tab *3. Content description,* for example:

NWO-funding Veni, number [number]

Screen 3: Content description

Figure 6: Screen 3 – Content description

In Screen 3: Content description you can enter a description of the content of your research project. Assigning multiple keywords (under *Subject* and *Coverage*) is essential here.

In the *Coverage* part of the screen you can differentiate between the spatial location where the research project was carried out (*Spatial coverage*) and the period that is covered by the research project (*Temporal coverage*).

In principle, you can enter free text in the fields under Content description. Preferably use short descriptions and include concepts that characterize the dataset best.

For Archaeology, EASY contains the fields *Archis Zaakidentificatie / Archis Onderzoeksmeldingsnr., Subject (ABR Complex)* and *Temporal Coverage (ABR)*. It is possible to use a drop-down list to fill in several of the fields. At 'Subject' you can, for instance, use the ABR complex type (Archaeological Basis Register, Archis 1992) to describe the type of research site and at 'Temporal coverage' you will find the ABR abbreviations for the date. By using these terms, it will be possible for others to perform a more targeted search. You are not obliged to use the drop-down lists, you can still use free text to fill in the fields.

Geographical coordinates

By documenting the coordinates at Spatial coverage, it will be possible to show the research site or the archaeological site automatically on a distribution map (geographic search function). At 'Spatial coverage', the information can be entered as free text (to specify, for instance, the toponym, the city/town, municipality, and province), as a point (2 coordinates), and as a box (4 coordinates). When entering coordinates, the coordinate system to which they refer has to be selected from a drop-down list. Do not use diacritics when you use the *Rijksdriehoekstelsel* (RD). Diacritics can be used in Longitude-Latitude coordinates.

Identifier

EASY automatically assigns a Persistent Identifier to your dataset. To many projects other identifiers – e.g. ISBN, a project code or a report number - have already been assigned before they are deposited in EASY. These can be entered in the *Identifier* field. You can select the identifier scheme to which they refer from a drop-down list. Use '-*no scheme*-' when none of the identifier schemes in the list is applicable to your identifier.

When you have received funding, you can enter the funding registration number as an Identifier. Always enter both the registration number and the funding organisation. *NWO-funding Veni, number [number]*

Relation

In the Relation field you can refer to a related dataset, publication or journal article. It is advised to mention the DOI (Digital Object Identifier) here. If a paper has not yet been published, add *(forthcoming)* after the *Title* and mention in the *Remarks* field where and when the paper is expected to be published.

By means of the *Relation* field you can link related datasets. In this field you can refer, for instance, to a related dataset or website. Please fill in the title and the full web address (including *http://*) under URL.
You can specify the type of relation by means of a dropdown list. For more information on this, please consult the help text in EASY.
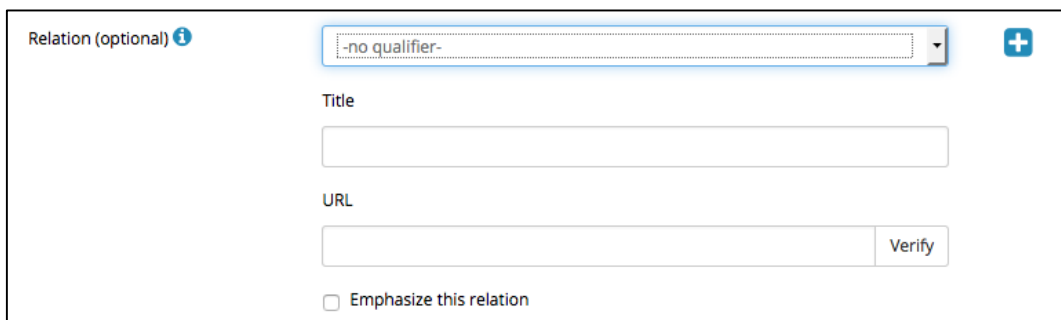

Figure 7: Relation field

By checking *Emphasize this relation* this *Relation* will be visible on the first page of the dataset (the *Overview* tab) and attract extra attention.

The screen *Content description* can also be used to provide information on the deposited digital files. This may be of a technical (*Type* and *Format*) or of a

substantive nature (*Language, Source, Date*). For example, it is possible to enter the period during which the research project was carried out, or its completion date. You don't have to enter the deposition date (*Date submitted*) here, this date is added by default during the deposition process.

Remarks
The metadata field *Remarks* can be used for additional information on the dataset that cannot be entered in any other field.
You can also use this field to communicate data use suggestions to potential users, or to send requests to the DANS data manager. This is particularly useful when depositing a dataset with both *Open access* and *Restricted access*: request permission files. In such cases, you use the *Remarks* field to indicate which files should be made available according to which access category.

Screen 4: Rights

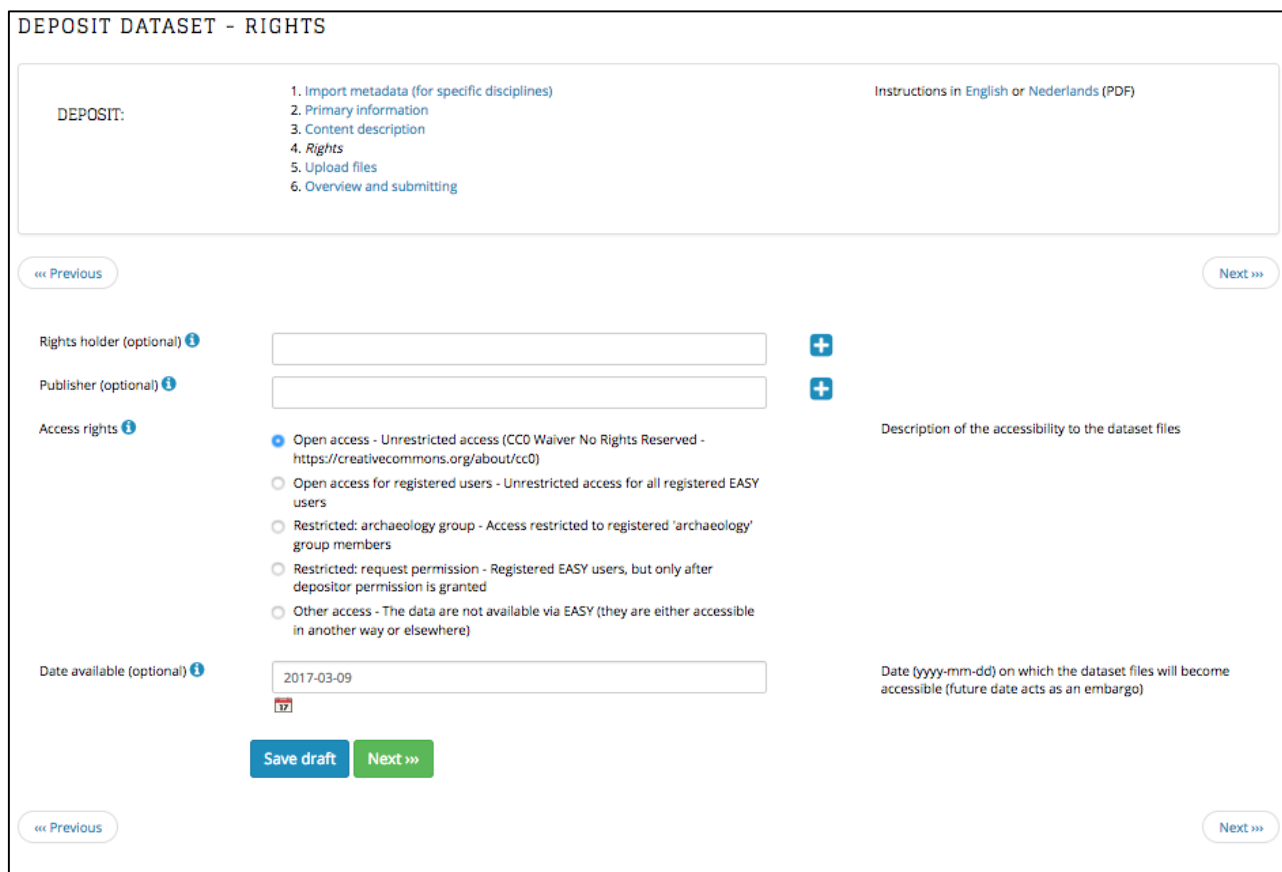*Figure 8: Screen 4 – Rights*

In Screen 4: Rights you can indicate who is the rights holder for your dataset (*Rights holder*) and who is the publisher of the data, or in archaeology, the report. For data published by Wageningen University 'Wageningen UR' is entered in the *Publisher* field.

The field *Access Rights* is used to indicate the conditions on which your data will be accessible.

Open access – CC0 Waiver (accessible to everyone without logging in to EASY, in accordance with the Creative Commons Zero Waiver licence.[14]) One of the conditions of the CC0 licence is that you waive your rights to the data.

The other options are:
Open access for registered users (accessible to logged in EASY users in accordance with the DANS licence agreement).[15]

Restricted access: archaeology group (accessible to logged in professional archaeologists and archaeology students in accordance with the DANS licence agreement).[16] This category does not apply to non-archaeological data.

Restricted access: request (users can access and download the data only after you have given them permission.) This is in accordance with the DANS licence agreement.[17]

Other access (the data are accessible in another repository than EASY). The category *Other access* can only be used when you have a contract with DANS.
You can access additional information on these categories by checking the information button 🛈 on the screen. At your request a data manager can assign the *Open access* category to selected files within a *Restricted: request permission* dataset. In such cases, you use the *Remarks* field to indicate which files should be made available according to which access category. Before publishing the dataset, the data manager will adjust the access category.

It is standard DANS practice to assign the *Open access for registered users* category to the documentation files belonging to a dataset with the *Restricted: request permission* access category. This enables interested EASY users to judge from the documentation if the dataset is relevant to their own research project before requesting access to the data files.

The selected access category relates only to the possibility of downloading the data files belonging to a dataset. The metadata belonging to a dataset are not subject to rights and accessible without restriction as *Open access CC0 Waiver*.

Date available
The *Date available* field is used to set a temporary embargo on the data for a maximum of two years. As soon as the embargo period has ended the dataset will be accessible in accordance with the access category under *Access rights* that you selected when you deposited the dataset. Checking the information button on the screen will give access to additional information. At your request DANS archive managers can prolong the embargo period.
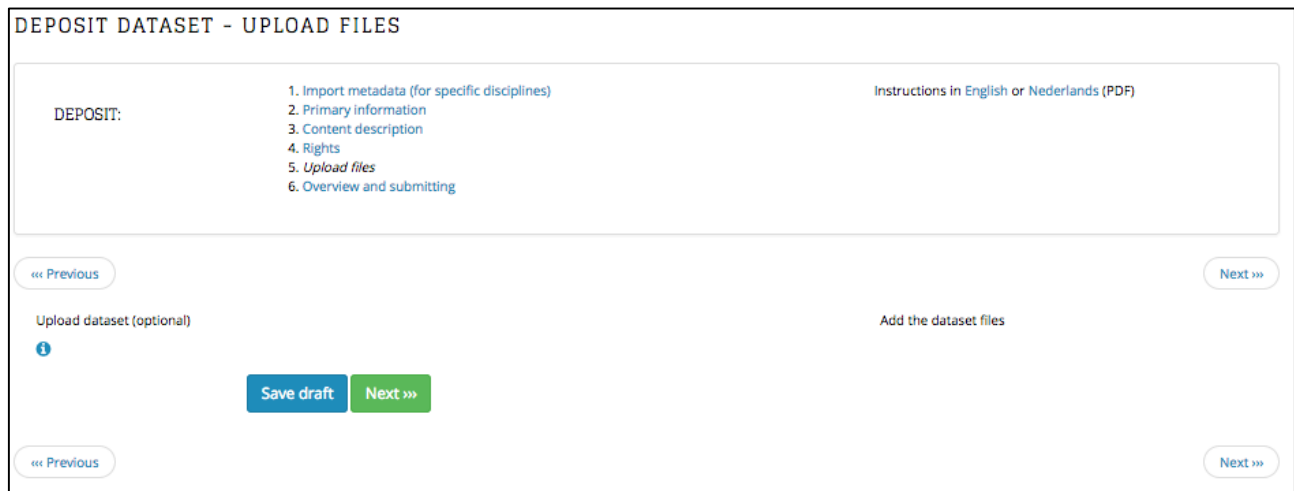
---

14 http://creativecommons.org/publicdomain/zero/1.0
15 https://dans.knaw.nl/en/about/organisation-and-policy/legal-information

16 https://dans.knaw.nl/en/about/organisation-and-policy/legal-information
17 https://dans.knaw.nl/en/about/organisation-and-policy/legal-information

Deposit Instructions for EASY - May 2017

5. Upload files

Screen 5: Upload files is used to upload documentation and data files.

Figure 9: Screen 5 - Upload files

Use the *Browse*-button in the metadata field *Upload dataset* to find the file. Select the correct file and check *Upload* to start the data transfer.

ZIP files in EASY
We advise to bundle all files of a voluminous dataset in one ZIP file before you upload them. EASY will unpack the ZIP file during the uploading process and restore the original files and directories.

EASY will only process and unpack ZIP files with the .zip extension, it does not support other ZIP formats, such as 7zip (.7z) or Winrar (.rar). DANS is regularly requested to store ZIP files in EASY, and to make them accessible to users without unpacking them. We advise against this because we can't foresee if in the long run it will remain possible to open ZIP files. For this reason, we always ask for a set of unpacked files to be delivered to us together with the ZIP file.

Very voluminous datasets (and ZIP files with a volume well over 100 MB) can't be uploaded by EASY users. Please enter all project metadata and mention in the *Remarks* field that the files are too big to upload. You can then submit the dataset without the files. The data manager who processes your dataset will contact you to discuss alternative methods of depositing and accessing the files.

(Open Access) journals
Have you written a journal article for an Open Access journal such as *PlosOne*, and are you required to deposit your research data before the article is published? In that case, you can deposit the data and the documentation file in EASY, but don't forget to mention the name of the journal and the expected publication date. Amongst other information, the documentation file contains a description of the dataset structure, with a description of variables (if applicable) and the connection between the files.

When your article has been published, we would like to receive the Persistent Identifier (DOI, Handle, etc.). This will enable us to establish a link between the dataset and the article. When the article has been published on paper we would like to receive the bibliographical data. When the article is available online without a Persistent Identifier, we would like to receive the bibliographical data including the internet address (URL). Added Persistent Identifiers will help to draw more attention to your article.

6. Submitting files

Figure 11: Screen 6 – Overview and submitting

In Screen 6: Overview and submitting you will see an overview of all the metadata you entered. Check the data one last time and accept the terms of the DANS Licence Agreement by checking the box. After this you can click *Submit* to complete the deposit. The dataset will now be submitted to DANS where a data manager will review and publish it.

Deposit Instructions for EASY - May 2017

After clicking *Submit* you can no longer modify or update the dataset. If you still want to apply changes, you should contact a DANS data manager (info@dans.knaw.nl or for archaeology archeologie@dans.knaw.nl).

If the data submission fails, you may have neglected to fill in a required field. This field will be indicated in an EASY error message. Once you have filled in the field in question, click 'Submit' again.

After having submitted the dataset successfully, you will receive a confirmation e-mail with the licence and the Persistent Identifier of the dataset. This Persistent Identifier will be activated once DANS has published your dataset.

7. Publication by DANS

After you have deposited the data, a DANS data manager will process your dataset according to an established protocol.[18] They can ask you to adjust the dataset. If all the conditions are met, they will make the dataset publicly available with the access level and any embargos you may have set.

Your data will now remain sustainably archived to be found and reused by others.

More information

Information about depositing research materials can be found on the DANS website.[19]

If you have any questions or issues, please contact a DANS data manager at info@dans.knaw.nl or, for archaeology, archeologie@dans.knaw.nl, or +31 70 34 94 450.

---

18 https://dans.knaw.nl/en/deposit/information-about-depositing-data
19 https://dans.knaw.nl/en/deposit/information-about-depositing-data

Deposit Instructions for EASY - May 2017