

1. Purpose of Data Archiving and Networked Services (DANS):

DANS is an online platform that provides an opportunity to share and reuse of research data. Data collected for a certain research purpose contains the answers to research questions from very different research in the same or in another research discipline. The availability of research data facilitates the reproduction of research, which is an important condition for science. The researchers get DOI (digital object identifier) for their uploaded data, which provide a persistent link to the location of data on the internet and the uploaded data can be cited. See point 4.1 for example of the reference of the uploaded data.

The process of uploading data can be divided into two major parts: (a) preparation of data and metadata; and (b) real time process of uploading data to the DANS server.

2. Preparation of data and metadata

Following points should be considered for the preparation of data (points 2.1 to 2.5) and metadata (point 2.6):

- 2.1 The data gathered during whole PhD project may be difficult to organize in one single project. Therefore, data can be divided into the main chapters/publications of PhD dissertation and the data of each chapter/publication can be uploaded separately to DANS server. However, based on the diversity of the data, PhD students can adopt either they want to upload the data for each chapter or the data for whole PhD research, but keep in mind the point 2.3 below.
- 2.2 The data to be uploaded must contain the initial input, the intermediate input and output, and the final output of the research. But, it is very important to decide “which data to upload” so that the uploaded data describes the whole research, but in a concise way. See point 4.2 for example.
- 2.3 It should also be noted that the maximum size of the uploaded data should not exceed 100 MB. For larger size, one needs to contact the administrator of DANS before uploading.
- 2.4 The DANS document “[DANSpreferredformatsUK.pdf](#)” describes the accepted format of the data files. The formats of data files are very important so that DANS could offer the long-term guarantees in terms of usability, accessibility, and sustainability. Please read the description of the format carefully in the section 2.1 of the DANS document.
- 2.5 Sometimes, providing data as an excel file has its own advantage so that the important calculations on data could be shown. Therefore, if the data is provided as an excel file, the csv version of the excel file should also be provided for long-term preservation purposes.
- 2.6 The next step of the preparation is the documentation of well-defined metadata. Metadata should be provided as a PDF file. The metadata should start with the uploader name,

affiliation, data, and maybe version number. Metadata file should be divided into the two sections “Introduction” and “Description of uploaded files”. The following information should be provided:

- 2.6.1 The first section “Introduction” describes the general introduction of the research including brief methodology and the information of the published paper, if available, related to the uploaded data. See point 4.3 for example.
If the open source software are used in the research and these software can’t be distributed by the user due to the copyright policy, please provide the link to download the software used in the research in this section.
- 2.6.2 The next section “Description of uploaded files” describes the uploaded file list (in the format suggested for archaeological metadata) and the description of variables and other information in each data file. Description of each data file should be provided as a table. At least, following information should be provided in the table:

file_name	File name of the data file					Part 1
file_content	Description of the contents					
Software	Software and version used					
purpose	Specific purpose of the data file					
Information of variables in the data						Part 2
Variable	Symbol	Unit	Sensor used for measurement	Frequency	Remark	

Part 1 of the table gives the data file name (it is good choice to highlight the name of each uploaded file in red or in another colour), file content, software used to generate data file, and most important the purpose of the data file in the research. Part 2 provides the information of variables in the data file. A note can be written, if required, anywhere in the table to give extra information. See point 4.3 for example of this table.

It should be noted that part 2 can also be written as a separate file in the form of code book (preferably as PDF file) if there are lot of information to provide. In this case, part 2 should simply look like:

othmat_codebook	exact file name of the related codebook (samplelist_code.txt)
-----------------	---

The code book may contain any information of data file such as the name of tables in a database, the sheets in a spreadsheet, meanings of the variable names in the table, spreadsheet or GIS-table. There is no specific format of code book, however the contents of the code book should be clearly described. For each data file, a separated code book, if required, should be provided.

Another important point of the metadata is that the order of the tables of the data files should be in-line with the logical steps of the research.

The content of the table is not limited as given in the above above. The full list of table content is given below. For more details, read the description of the table contents in the DANS document "[DANSarchaeologicalmetadataUK.pdf](#)".

file_name	File name of the data file (samplelist.xls)
file_content	Description of the contents (list of soil, seed and wood samples)
data_format	General technical description (relational database)
software	Software and version used (MS-Word, version 2000)
hardware	Original system (Intel-PC)
original_OS	Original operating system (DOS)
data_collector	Specific individual responsible for the contents of the file (hired)
purpose	Specific goal for the collection of data in this file (random sample)
collection_mode	method of collection (height measurements in a grid of 5x5 metres)
analytic_units	analysis/storage of features, observations or records (description of individual features)
data_appreciation	evaluation of data quality, re-use value or limitations (5 to 10cm error margin in height measurements)
geog_cover	spatial coverage area (500 x 250m around central point)
geog_unit	features where spatial coordinates have been used (in kilometres)
mapprojection	name of the card projection or local measurement system (RD)
local_georef	minimum of 2 reference points for the conversion of local to national coordinates
source_document	(analogue) source of digital data (digitalised field measurements 1:50)
othmat_citation	name of the files that contain related, supplementary information
othmat_codebook	exact file name of the related codebook (samplelist_code.txt)
notes	supplementary and specific instructions for (re-)use

2.6.3 The final step is to combine all data files and metadata description file in one zip file. This zip file should be uploaded to the DANS server. All the contents of the zip file will be extracted automatically during the process of uploading.

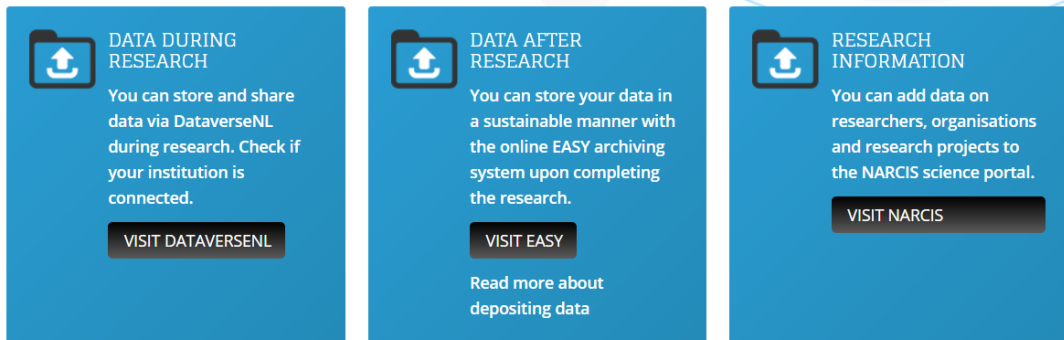
3. Process of uploading data to the DANS server

The zip file, which contains data files and metadata file, is uploaded on the DANS server after creating an account. The process of uploading takes 20-30 minutes given the data files are well-prepared and all information required (read carefully all the points below) during the process of uploading are ready. All account access rights are assigned to the account holder. The account holder is only responsible for uploading data and further modification in the data files if requested by the DANS data manager. After uploading the zip file, DANS data manager checks the uploaded contents and they suggest, if required, some corrections by email. Finally, data is made available on the DANS server and DOI of the uploaded data is provided. The whole process may take one or two weeks.

3.1 Create an account on the DANS server

3.1.1 There are three options to upload the data: (a) DATA DURING RESEARCH; (b) DATA AFTER RESEARCH; and (c) ADDING DATA TO NARCIS SCIENCE PORTAL. This document explains only the option “b” to create an account and uploading the data. Data after the research are uploaded through EASY archiving system. See the screenshot and link below.

Use the link <https://dans.knaw.nl/en/deposit> to create the account.



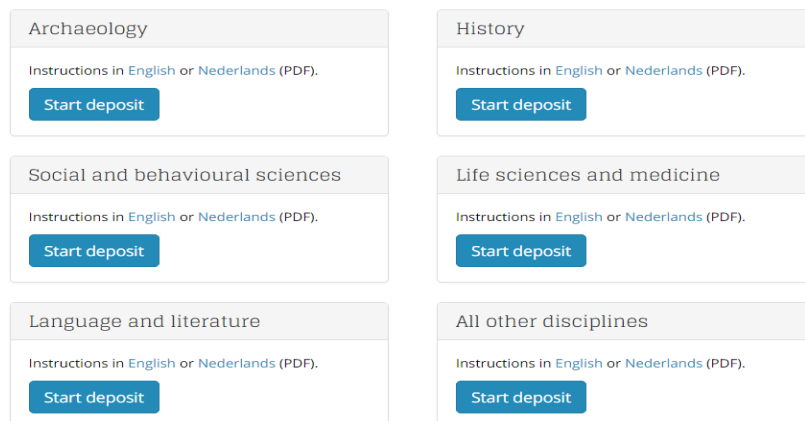
3.1.2 Click on the option “VIST EASY”. The following options will be displayed.

DEPOSIT YOUR DATA

To start depositing your data you will need to:

1. Prepare your data;
2. Log in to EASY or create an account;
3. Choose the relevant discipline below to start the deposit procedure.

More information about the deposit process can be found on the DANS website.



Option 1: “Prepare your data”

The DANS has specific requirement for the preparation of data for different disciplines that can be read through the option “Prepare your data”. Most of the research in ITC will fall under “All other disciplines”. There is no specific requirements for the format of metadata for all other disciplines. However, a well-structured format of metadata as explained in section 2 (which is motivated by the guidelines for the preparation of data and metadata for Archeology) should be followed during the preparation stage.

Option 2: Log in to EASY or create an account

To create an account, click on the option “create an account”. The registration page will be displayed. There are some mandatory and optional fields. The information to be provided in each field is easy to understand. The fields of the option “organization” are however optional, but the researchers of ITC should provide the name, department, and address of ITC. There is one another optional field “DAI (Digital author identifier)”. The Digital Author Identifier (DAI) is a unique national number assigned to individual researchers employed at Dutch universities and research institutes. The DAI is the only independent (non-commercial) identification number for Dutch researchers. If the information of DAI of the researcher is available then this should be provided. The screenshot of the “Registration” page is given below.

REGISTRATION

Please fill out the details below to create an EASY account, with which you can download or deposit research data. Downloading Open Access data does not require a login.

Account	
Username	<input type="text"/> A short name you will use to log in (minimal 5 characters, only letters and numbers).
Password	<input type="password"/> A secret phrase you will use to log in (minimal 6 characters).
Confirm password	<input type="password"/> Repeat the password to make sure you typed it correctly.
E-mail address	<input type="text"/> For future correspondence, specify an address which is as permanent as possible.
Person	
Title (optional)	<input type="text"/> Your academic title (e.g. "Prof. Dr.").
Initials	<input type="text"/> (e.g. "R.M.")
Prefix (optional)	<input type="text"/> (e.g. "de" or "van de")
Surname	<input type="text"/> Your last name.
Function (optional)	<input type="text"/> Your role (e.g. as on your business card).
Telephone (optional)	<input type="text"/> Your telephone number.
Primary discipline (optional)	<input type="text" value="Choose One"/> The foremost discipline in which you are occupied/interested.
Secondary discipline (optional)	<input type="text" value="Choose One"/> An additional discipline in which you are occupied/interested.
Tertiary discipline (optional)	<input type="text" value="Choose One"/> An additional discipline in which you are occupied/interested.
DAI (optional)	<input type="text"/> Your Digital Author Identifier.
Organisation	
Organisation (optional)	<input type="text"/> The institute or company at which you are occupied (e.g. "Utrecht University").
Department (optional)	<input type="text"/> Part of an organisation (e.g. "Department of social sciences")
Address	<input type="text"/> (e.g. "Kerkweg 1")
Postal code	<input type="text"/> Postal or zip code (e.g. "3500 AA")
City	<input type="text"/> (e.g. "Utrecht")
Country (optional)	<input type="text"/> (e.g. "The Netherlands")

Option 3: Start deposit

After creating an account, now click on the option “Start deposit” under “All other disciplines”. The login page will appear. The username and password created in Option 2 will be used to login the EASY archiving system.

3.2 Upload the zip file containing data files and metadata files

There are five steps to fill up during the process of uploading. There is an “i” mark associated with each field in each step that gives the general explanation of the information that has to be filled in that field.

3.2.1 Required elements: In this step, information about the creator, title and short description of the research, date of creation of research project, and the access rights should be provided. If the information on DAI is available, this should be provided in the field “Digital Author ID (DAI)” under the option “Creator”. Otherwise the field “DAI” can be left blank. The field “Date created (ISO 8601)” under the option “Date created” should be filled with the date of completion of the research project. However, PhD students, if they have already submitted their thesis, can also fill the date of final thesis submission. The field “Access right” under the option “Access” has four types of accessibility of the uploaded data: open access for unregistered users, open access for registered users, restricted (request permission), and other access (data available not by EASY). Before choosing the type of access right, an agreement should be made with all contributors/collaborators of the research data/project. The field “Data available (optional)” under the option “Access right” can either be changed or left with the default date of uploading the data files. The screenshot of the “Required element” page is given below.

DEPOSIT DATASET - REQUIRED ELEMENTS

DEPOSIT:

1. Required elements
2. Upload files
3. Recommended elements
4. Additional elements
5. Overview and submitting

Date created

Date created (ISO 8601) ⓘ

Date created (optional) ⓘ

Creator ⓘ

(Academic) Title(s)	<input type="text"/>
Initials	<input type="text"/>
Prefix	<input type="text"/>
Surname	<input type="text"/>
Digital Author ID (DAI)	<input type="text"/>
Organization	<input type="text"/>

Access

Access rights ⓘ

- Open access - Unrestricted access (CC0 Waiver No Rights Reserved - <https://creativecommons.org/about/cc0>)
- Open access for registered users - Unrestricted access for all registered EASY users
- Restricted: request permission - Registered EASY users, but only after depositor permission is granted
- Other access - The data are not available via EASY (they are either accessible in another way or elsewhere)

Date available (optional) ⓘ

Audience ⓘ

3.2.2 Upload files: In this step, zip file containing all data files and metadata file is uploaded. The option “Choose file” is used to add the zip file from the computer and the option “Upload” is used to upload the zip file. The content of the zip file will be extracted automatically. The screenshot of the “Upload file” page is given below.

DEPOSIT DATASET - UPLOAD FILES

DEPOSIT:

- 1. Required elements
- 2. Upload files
- 3. Recommended elements
- 4. Additional elements
- 5. Overview and submitting

[<< Previous](#)

Upload dataset (optional) No file chosen

3.2.3 Recommended elements: This can be observed that all the field in this steps are optional. However, information in the following fields must be provided:

- a) Contributor (optional) – Here the information about the person or organization that contributed to the data is provided. Any number of contributors can be added here using the “+” sign. PhD student should add the name of promoter and supervisors. However, there may be other contributors from other organizations, their names should also be added. For each contributor, DAI, if available, can be provided.
- b) Right holder (optional) - In this field, provide the following information: “Faculty of GeoInformation Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands”

The information of the other field can be provided based on the data. Please use “i” sign associated with each field to get information about that field. The screenshot of the “Recommended elements” page is given below.

DEPOSIT DATASET - RECOMMENDED ELEMENTS

DEPOSIT:

- 1. Required elements
- 2. Upload files
- 3. Recommended elements
- 4. Additional elements
- 5. Overview and submitting

[<< Previous](#)

Contributor (optional) i	(Academic) Title(s) <input type="text"/>
	Initials <input type="text"/>
	Prefix <input type="text"/>
	Surname <input type="text"/>
	Digital Author ID (DAI) <input type="text"/>
	Organization <input type="text"/>
Rights holder (optional) i	<input type="text"/>
Subject (optional) i	<input type="text"/>
Spatial coverage (optional) i	<input type="text"/>
Temporal coverage (optional) i	<input type="text"/>
Source (optional) i	<input type="text"/>
Identifier (optional) i	<input type="text"/>

3.2.4 Additional elements: In this step, the format of each data file, and the ‘relation’ to the published articles are provided. The title and DOI of the published article can be provided in the field “Title” and “URL” respectively. Additional remarks can be added in the field “Remarks” to provide an extra information about the data that can help users in to

understand the dataset/instruments/coding used in the project. In this field, however, the following information must be provided:
“A file list and description of data (i.e., metadata) in each file are provided in the uploaded pdf file "NAME_METADATA_FILE.pdf". This pdf file also provides a brief description of the methods adopted in this research.”

The screenshot of the “Additional elements” page is given below.

DEPOSIT DATASET - ADDITIONAL ELEMENTS

DEPOSIT:

1. Required elements
2. Upload files
3. Recommended elements
4. *Additional elements*
5. Overview and submitting

« Previous

Format (optional) ⓘ

Relation (optional) ⓘ

Title

URL

Emphasize this relation

Language

Language (ISO 639) (optional) ⓘ

Language (optional) ⓘ

Remarks (optional) ⓘ

3.2.5 Overview and submitting: In this step, all information provided in the previous steps together with the uploaded data files can be viewed. After the acceptance of the licence agreement, the data can be finally submitted.

Warning: After submitting it will not be possible to make any changes to the data (but you can always contact DANS to alter metadata or files).

4. Example of data and metadata

4.1 DANS provides a reference to the uploaded data as:

Raj, R. (Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands) (2016): Variance-based sensitivity analysis of BIOME-BGC for gross and net primary production. DANS. <http://dx.doi.org/10.17026/dans-z8b-s53t>.

4.2 In Raj (2016) [see the reference in point 4.1], the data belongs to the sensitivity analysis of process-based simulator BIOME-BGC. This experiment generated a large sample space of input parameters (47000 vectors of parameters) of BIOME-BGC simulator that finally resulted into very large (>10GB) size of daily simulated output files, which were the intermediate output, but in any case, these were not feasible to upload. Therefore, only the annual mean of simulated output (for which the sensitivity indices were calculated) generated with each input parameter vector was uploaded as a text file.

4.3 In Raj (2016) [see the reference in point 4.1], the first section of the metadata file is written (highlighted in italics) as:

Introduction

This dataset contains the underlying data for the study:

Raj, R., Hamm, N.A.S., van der Tol, C., Stein, A., 2014. Variance-based sensitivity analysis of BIOME-BGC for gross and net primary production. Ecological Modelling 292, 26-36, DOI: <http://dx.doi.org/10.1016/j.ecolmodel.2014.08.012> .

The above journal paper can also be found as "RajEtAl2014.pdf" in the uploaded files.

This research conducted variance-based sensitivity analysis (VBSA) of the process-based simulator BIOME-BGC for gross and net primary production (GPP and NPP) output in a Douglas-fir stand at the Speulderbos forest site, The Netherlands. Such an analysis allowed identification of the most influential input parameters of BIOME-BGC and provided insight into the simulator function. A VBSA produced the sensitivity indices (first and second order) whose value indicated the degree with which simulated GPP and NPP were sensitive to the input parameters. A VBSA was implemented using SIMLAB version 2.2. For this implementation, a sample space of the input parameters was generated in SIMLAB using defined probability distribution functions. The result was then exported from SIMLAB and reformatted to allow sequential simulator executions of BIOME-BGC version 4.2 on a Windows 7 system. For each input parameter vector, the BIOME-BGC was run for four years (2007-2010) to simulate daily value of GPP and NPP. The annual mean GPP and NPP for each year was then calculated by the mean of daily simulated GPP and NPP respectively. These outputs were then arranged in a suitable format to read back into SIMLAB to calculate the first and total order sensitivity indices. Each simulated output resulted in four values of each sensitivity index for each input parameter. The large number of BIOME-BGC parameters makes VBSA computationally expensive. Therefore, before applying VBSA, a computationally cheap Morris method in SIMLAB was conducted to screen first the most influential parameters and then considering only screened parameters in VBSA by fixing other input parameters in advance. The details of methodology can be found in Raj et al. (2014) (RajEtAl2014.pdf).

Due to the copyright policy, SIMLAB and BIOME-BGC can't be provided by the uploader. However, these software are open source and can be downloaded from the following links:

- I. SIMLAB version 2.2: <https://ec.europa.eu/jrc/en/samo/simlab>
- II. BIOME-BGC version 4.2: <http://www.ntsq.umd.edu/project/biome-bgc>

The details of each uploaded files are provided in the section 2.

A table for one of the uploaded data files "MetVar_Speuld_2007_2010.xlsx" is provided in the second section (highlighted in italics) of the metadata file as:

Description of uploaded files:

<i>file_name</i>	MetVar_Speuld_2007_2010.xlsx (see also Note 1 below)				
<i>file_content</i>	<p>a. Half-hourly measurements (2007-2010) of meteorological variables (Precipitation, global radiation, air temperature, relative humidity, and vapour pressure deficit) acquired from the Speulderbos flux tower, the Netherlands.</p> <p>b. Daily values (2007-2010) of meteorological variables calculated from half-hourly measurements</p>				
<i>Software</i>	MS-office, version 2013				
<i>purpose</i>	<p>The purpose of this file is to prepare the meteorological input file (daily values of meteorological variables) for BIOME-BGC from half-hourly measurements.</p> <p>The file "MetVar_Speuld_2007_2010.xlsx" is not the direct input for BIOME-BGC. A specific format is required for this (see Note 2 below).</p>				
<i>Geographic location</i>	Speulderbos flux tower is located at 52°15'08.1" N, 05°41'25.8" E				
Information of variables used in the excel file					
<i>Variable</i>	<i>Symbol</i>	<i>Unit</i>	<i>Sensor used for measurement</i>	<i>Frequency</i>	<i>Remark</i>
Precipitation	Prpc_h	mm	Tipping bucket (Vaisala, Finland)	Half-hour	48 half-hourly measurements are acquired on each day from 2007 to 2010
Global radiation	Rd	W m ⁻²	Net radiometer CNR1 (Campbell Scientific, INC.)	Half-hour	
Air temperature	T	°C	Weather sensor WXT510 (Vaisala, Finland)	Half-hour	
Relative humidity	RH	%	Weather sensor WXT510 (Vaisala, Finland)	Half-hour	

Note 1:

Half-hourly and daily meteorological variables are also presented separately as comma delimited files ("**Half-hourlyMetVar_Speuld_2007_2010.csv**" and "**DailyMetVar_Speuld_2007_2010.csv**" respectively) for preservation purposes.

Note 2:

The content of *DailyMetVar_Speuld_2007_2010.csv* is also presented as a specific format of meteorological input file ("*Speuld0710.mtc41*") required by BIOME-BGC.