Plan Overview

A Data Management Plan created using DMPonline

Title: Extended areas of application of DESIREE by new instrumentation

Creator: Henning Schmidt

Principal Investigator: Henning Schmidt

Contributor: Henning Zettergren, Michael Gatchell

Affiliation: Stockholm University

Funder: Swedish Research Council

Template: SU-VR template

ORCID iD: 0000-0002-8209-5095

Project abstract:

This project serves to expand the DESIREE research area through new instrumentation through widening of the scope of experiments by introducing novel methods for ion-production, -cooling and -selection. This includes improved ion-production methods such as formation of cold cluster ions inside helium nanodroplets, an additional electrospray ionization system, and an effective source for metal cluster cations. A laser system is included for precision spectroscopy and manipulation of the quantum-state distribution of the stored ions for state-selective experiments. For studies of complex ions, detection of charged fragments with a wide range of energy-to-charge ratios is essential and we propose such improvements by introducing new detectors and DESIREE operation schemes. Finally, we ask for funding to further improve the already excellent vacuum in DESIREE. Further vacuum improvements are essential for the next generation of experiments.

ID: 93524

Start date: 01-01-2022

End date: 31-12-2026

Last modified: 16-03-2022

Grant number / URL: 2021-00155

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Extended areas of application of DESIREE by new instrumentation

0: Note on personal data!

Q1: I have read and understood the above declaration and hereby certify that this DMP contains no personal data except for information about project members such as PI and contact person.

- Yes

I: Description of data – reuse of existing data and/or production of new data

Q1: Data Quality Assurance measures (FAIR data): please check multiple options that will apply to assure quality and integrity of data collected, created or reused.

- 1. Non-proprietary file formats (e.g. .csv, .txt, .json, netCDF)
- 2. Sustainable file formats (e.g. .pdf; .csv; .txt)
- 3. Software specific file formats (e.g. Matlab - .mat; Stata - .dta)
- 4. Proprietary file formats (.doc ; .xls)

Options 5-7 are currently not implemented in all cases. This is something that we should aim to improve.

Q2: Dataset ID: at this initial planning stage, please find one main identifier (e.g. a DOI, Handle, URL, ...) for the entire dataset(s) in the project where possible, even if it comprises several data files of different types.

https://zenodo.org/communities/desiree-infrastructure/

Q3: Dataset Identifier Type: for your dataset ID above in Q2, please check the corresponding option in the list below!

- url

This project is primarily related to hardware development rather than a specific research project. General data will be made available at url.

Q4: Dataset Description (Abstract) - please describe the dataset(s) in the project! The description can be at a rather simple conceptual level, which does not have to point to individual data files.

The dataset should contain all of the information required to test and reproduce results that are published by users of DESIREE, as well as make new findings from existing data. Metadata regarding experimental conditions are vital, such as ionic species, energies, ion ring geometries, and timing information. Uncalibrated as well as calibrated data will be made available.

Q5: Title of dataset

Results from the DESIREE infrastructure

Q6: Are you re-using datasets that already have a definite distribution (identifier, access point or location, title ...)

- no

Q10: Type of dataset(s) / Resource type of the main dataset(s) of the project described by answers to Q2 / Q5 (thus, not primarily of re-used datasets).

- other

Primarily tables containing histograms (time-resolved data from neutral particle and fragment detectors) and correlated detection events from merged beam experiments. These will either be in ascii tables, e.g. csv-files, or in open hierarchical data structures.

Q11: Issue date (YYYY-MM-DD) of dataset in Q2 / Q5.

2022-01-01

II: Documentation and data quality

Q1: How will metadata be created for your dataset? If by*use of a repository* (*recommended*), please specify which, either from the given options, or - if Other - by giving a link(s) / URL(s) [if multiple separated by commas] as Additional Information below. Please*, do not write whole texts here*with line or paragraph breaks*, as this prevents automatic processing and evaluation of the DMP!*

- 6. README-file
- 4. Zenodo/StockholmUniversityLibrary

Metadata regarding experimental conditions should be clearly accessible either in a readme file or as part of the data structure. Additional information will be available in the database's metadata system.

Q2: Which metadata standards and vocabularies will you employ for general and domain specific metadata? (Multiple options possible. Some of them may overlap, then it is unnecessary to check all that hold a particular vocabulary specified in the text field.)

- 1. Metadata from II:Q1

Metadata will be available on the DESIREE Zenodo community.

Q3: Which of the following data quality documentation and safeguard measures, if any, would you consider applying to your dataset?
(Multiple options possible. For options 4. *Pre-registration,* 6. *Supplementary documentation* or 9. *Other*, please specify to the extent possible in the comment area, e.g. by giving the URLs of particular services you intend to use for certain data quality measures.)

- 3. Integrity check of data files
- 2. File format and software description
- 5. Repeated measurements
- 7. Validation of data input

III: Storage and backup

Q1: Where will you store and backup your data during the project? (Multiple options possible.)

- 6. Repository in II:Q1
- 3. Own harddrive
- 5. "Cloud server" (please specify)

In addition to Zenodo page, data will be stored locally on lab computers and the Physics Department's data storage system (https://nextcloud.fysik.su.se).

Q2: What volume (X) of data will you need to store and backup?

- 1 TB < X < 10 TB

Q3: What security measures will you need to employ to protect your data during the research process? (Multiple options can be selected).

- 1. Password protection
- 4. Private links
- 6. Other (please specify)

Neither internal nor external users have the privilege to modify or delete raw data stored in the local data storage system.

IV: Legal and ethical aspects

Q1: Will the creation, collection or reuse of dataset(s) in your project entail processing of *personal data*, i.e. any information relating to an identified or identifiable natural person (a '*data subject*', that is a *living* person)?

- no

Q2: Will the creation, collection or reuse of dataset(s) in your project entail any of the following:

- **(a) processing of *special categories of personal data* according to the General Data Protection Regulation (EU 2016/679), i.e. personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation**
- **(b) processing of personal data regarding violations of law that include crimes, judgments in criminal cases, penal law sanctions, or administrative deprivation of liberty**
- **(c) physical interventions on research subjects or deceased persons**
- **(d) methods with the purpose of affecting a research person physically or mentally, or which includes an apparent risk of injuring the research subject either physically or mentally**
- **(e) studies of biological material that has been taken from a living or deceased person, and can be traced to that person**

Or, further, will the creation, collection or reuse of dataset(s) in your project include:

- **(F) Data from animal research**
- **(G) Data on genetic resources and/or traditional knowledge associated with genetic resources**
- **(H) Data that can be used for military purposes or concerning products that can be used for military purposes**
- **(I) Data that are sensitive in some other respect**

(Possible ethical review documentation applying to any of points (a-e) will be asked for separately.) If you answer yes to any of the points (F-I) and there is already relevant documentation or applications, please provide (a) reference(s) to any application(s)/ approval(s)/decision(s)/document(s), if possible by URL(s) / PID(s) such as DOI(s) giving direct access, or registration no. (Swe. *diarienr.*) in the text field below.

- no

Not expected to be relevant for the measurements done at DESIREE.

Q5: Intellectual Property Rights -License(s) of data. Please select the usage license(s) for dataset(s) and/ or software

***produced in your project*. (Multiple options possible. If *Other*, please specify by a URI or other file location for each of the additional license(s), separated by commas, in the comment area.)**
[For *previously existing datasets* that you will be *re-using* , corresponding license(s) are entered in section *I:Q9 Description of data - re-use ..*.]
If you choose the option '*No license*', being aware that it *might* make your dataset less FAIR, please state the reason for this choice. [There are *fully legitimate reasons*, concerning e.g. personal data and sensitive data, that cannot be shared.]

- 01. CC-BY-4.0

Published data will be made accessible with a Creative Commons 4.0 license. No personal or other sensitive information is present in the collected data.

V: Accessibility and long-term storage

Q1: Where will datasets, documentation and/or metadata be made accessible? (Means or location of *distribution*).

- 1. Repository in answer to II:Q1
- 3. SU Archive
- 2. Supplement to journal article / publication

All data should be made available on repository. Data may also be made available as supplemental material in publications, depending on the journals requirements.

Q2: What will be made directly accessible (e.g. via repository in Q1, or as supplement to online journal)?

- Metadata and some datafiles

Any data files relevant to a given research study will be made available at time of publication. Test measurements that do not impact the primary measurements may be excluded, as will data that cannot be readily transferred from proprietary internal formats to openly accessible formats. The latter will not have an impact on the usefulness of the data, but are available upon request.

Q3: When will data files and/ or metadata and documentation be made accessible?

- Only after publication of journal article / paper

Data should be made available in connection to publication. Ideally the data will be made available when a manuscript is submitted and contain a reference to the data repository.

Q4: How will you ensure that all data files, documentation and metadata are transferred to SU digital archive for long-term preservation?

- Automatic harvest & transfer from repository

Automatic harvesting from Zenodo.

Q5: Will specific systems, software, source code or other types of services be necessary in order to understand, partake of or re-use / analyse data in the long term?

- no

Data should be available in a clear structure that is agnostic to the analysis software.

Q6: Will the software you will use to collect, create, handle, transform, refine or analyse data also be needed to replicate or rerun experiments, partake of your datasets or open datafiles?

- no

Q7: Will the software / code you will use to collect, create, handle, transform, refine or analyse data be ... (multiple options possible)

- 1. Non-proprietary/Open Source (e.g. Python, R, XSLT)
- 2. Proprietary/Commercial (e.g. Matlab, Stata)

Data collection is primarily handled with LabView modules. Data analysis primarily with Python scripts, though this may vary with the user.

Q8: Will you be using Software in the "cloud" / Software-as-a-Service (SaaS) to create, handle, transform, refine or analyse data ?

- No

Typically, no. But this may vary with the user.

VI: Responsibility and resources

Q1: Who is responsible for data management and (possibly) supports the work with this while the research project is in progress?

- 1. PI
- 5. Others (please specify!)

Users of the facility will be required to make data available at time of publication. The contract PI will be responsible for enforcing this requirement.

Q2: What resources will be required for data management to ensure that data fulfil the FAIR principles? (Multiple options possible.)

- 3. Repository account(s)

Q3: Please estimate total extra costs (C) for data management, that is not covered by grant funding (or regular SU services, such as RDM-team support).

- < 10000 SEK

We are not sure at the moment and will have to see how this progresses.

VII: Funding requirement fulfilled for initial version

Q1: I hereby certify that the prefect / responsible head of department or institute has (re)viewed this initial DMP as fulfilling the requirements for funding. I am aware that answering Yes will send this Initial version of the DMP to Archive for long-term preservation, and that future editing will then be in Phase 2, the final version.

- Yes

VIII: DMP administrative information

Q1: Please give an Identifier *of the Contact Person* designated in Project details (even if same as PI), *-not* the *name* of the contact person, but only the identifier-string (that is within the " " in the examples below).

0000-0002-8209-5095

Q2: Please select Type of Identifier given as answer to Q1 above.

- orcid

Q3: Affiliation (Department / Institution) of Contact Person. Please select main Department / Institution affiliation *from drop-down menu* (ordered after faculties as in this list), or else choose "Other" and specify in comment area below!

- Physics

Q4: Language used for this DMP. Please select!

- eng: English

Q5: Funder(s). Multiple choice possible. If Other, please specify funder name(s) in the Additional Information text field, if more than one separated by commas.

- VR - Swedish Research Council

Q6: Grant ID. Please specify, if possible as a URL. (The Grant ID can often be the same as the PROJECT-ID in SweCris, e.g. https://www.vr.se/swecris#/project/2010-00383_VR)

VR 2021-00155

Q7: Funding status. Please choose one from the dropdown menu.

- granted

IX: Full DMP - additional Datasets and identifiers, Reference list and Project end

Q1: Additional dataset(s)
Please fill in the table below in accordance with the given example by replacing *None* in *Title, Identifier and Type* with *real values* for your dataset(s) after the *T1:, Id1: , Type1:* etc. You can add / delete rows if needed, but make sure the *new entries* are still in *italics* and leaving the last row without real values with *None* (as this will help us process your

DMP data output for review.)
Identifier type: select from the same list as in section I-Q3: *ark, doi, handle, url, other.*
If *other*, please specify the type of dataset ID below the table as e.g. *"Local filename"* or *"Project-ID".*
The Description, Type of dataset (software, images, text, spreadsheets, sound, video, other) and Issue date for these additional datasets will as default be the same as for the main dataset described in section I: Q4, Q10 and Q11, so you might have to adjust these answers to fit for all datasets, or specify these new values for each additional dataset (below the table).

Short Title of dataset	**Identifier of dataset**	**Identifier Type**
T1: None	*Id1: None*	*Type1: None*
T2: None	*Id2: None*	*Type2: None*
T3: None	*Id3: None*	*Type3: None*
T4: None	*Id4: None*	*Type4: None*
T5: None	*Id5: None*	*Type5: None*

Q2: List of References / Sources / Publications (other than reused datasets, in sec. I:Q6-9).
To be updated during all research project.

References: authors (year): *titles*	**Identifiers / Locations**

Q3: Research project ended?
Please indicate if the research project described by this DMP is completed, so the full and final version of this DMP can be sent to long-term archive.

- No