Plan Overview

A Data Management Plan created using DMPonline

Title: DECIDE. Deciding about Future Science: How Scientists Can Collectively Set Research Priorities in the Era of Big Science

Creator: Siska De Baerdemaeker

Principal Investigator: Siska De Baerdemaeker

Data Manager: Siska De Baerdemaeker

Affiliation: Stockholm University

Funder: European Research Council (ERC)

Template: ERC DMP

ORCID iD: 0000-0002-8642-5127

Project abstract:

Since World War II, we have entered the era of ‘big science’. Experiments in the natural sciences today can involve hundreds or thousands of scientists, they generate enormous amounts of empirical data and they require ever larger instruments. The associated cost of experiments has unsurprisingly also soared. Personal research funds or smaller research grants no longer suffice to fund certain experiments; instead, they require significant financial support from multiple government agencies over a long period of time. Because of this increase in costs, scientists have had to change how they decide what experiments to pursue. The central research question of this project is therefore: how can scientists make optimal decisions about what future experiments to pursue, given the scale of experiments and the limited available resources?

To answer this research question, the project will follow the approach of integrated history and philosophy of science. Through a careful study of the history of one of the most influential procedures for scientific funding decisions, the US Astronomy Decadal Survey, this project will: (i) trace the origins of current procedures for collective decision-making in science; (ii) investigate how these procedures have impacted the development of scientific research, specifically on the post-World War II history of astronomy, astrophysics and cosmology; (iii) formulate a novel philosophical account on what makes an experiment pursuitworthy; and (iv) apply current insights from social epistemology on collective knowledge-making and values and science to a new context, collective setting of research priorities.

Aside from its impact in history and philosophy of science, this project will have implications for the natural sciences by recommending revisions to current strategies for the setting of future research priorities. More broadly, this project will establish the nascent discipline of philosophy of astronomy and astrophysics in Europe.

ID: 209747

Start date: 01-07-2026

End date: 30-06-2031

Last modified: 18-09-2026

Grant number / URL: 101219395

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

DECIDE. Deciding about Future Science: How Scientists Can Collectively Set Research Priorities in the Era of Big Science

Summary

Project Acronym

DECIDE

Project Number

101219395

Provide a dataset summary

Publicly accessible data

The main data sources for the project are the Decadal Survey Reports, which are publicly available through the website of the US National Academies of Sciences, Engineering, and Medicine (https://www.nationalacademies.org/projects/consensusadvisory-activity). The project will also use the archives of the National Academies, which are accessible to researchers (https://www.nationalacademies.org/archives).

Newly generated data

Digital humanities research

Quantitative data about the published record in astronomy, astrophysics, and cosmology post-World War II will be generated using digital humanities methods and based on publicly available research databases like arXiv or astrophysics data system (ADS). This data set can be publicly available.

It is unlikely the data volume will exceed 1GB. It will likely be processed in a spreadsheet format, but the output will be a plain text file (e.g., .csv, .tsv) to ensure long-term accessibility.

It is currently unclear whether and what variables the data will have. Any variables will be documented in a readme file or code book.

Interview data

Depending on the evolution of the project, interviews with scientists may be conducted. This interview data will supplement the public record discussed above. Due to interviews likely containing personal data, this data set will be preserved but it will not be publicly available.

It is unlikely that the data volume will exceed 1GB, given that the likely format will be MP4. The interviews will be transcribed using accessible formats like .docx or .odt.

The metadata will contain the time and location of the interview, and the participants in the interview. It is currently unclear if the interviews will be highly structured. If so, the metadata will contain a list of questions.

FAIR data and resources

1. Making data findable

Publicly accessible data

This data is already publicly available and findable.

Newly generated data

Digital humanities data

The metadata will include information about how the data set was generated, including what research database was used, how the quantitative data was generated, and what format it is available in.

The data will be deposited in a public repository after the research is completed, e.g., FigShare (managed by Stockholm University), or the repository managed by the Swedish National Data Service (SND). At this point, the data set will be assigned a DOI and structured metadata in accordance with a widely used metadata standard by the repository.

Interviews

The metadata will include information about who the interview was conducted with, when it was conducted, and how it was conducted.

2. Making data openly accessible

Publicly accessible data

This data is already publicly available and findable.

Newly generated data

Digital humanities data

This data and the associated metadata will be made publicly available in a curated repository, e.g., FigShare (institutional instance managed by Stockholm University), or the repository managed by the Swedish National Data Service.

Interviews

Interview data will not be publicly available, since they will contain personal data and may concern sensitive personal data. Interviews will be made available to authorized researchers in accordance with GDPR and other relevant legislation. A metadata description of the interviews will be made publicly available in a curated repository.

3. Making data interoperable

Publicly accessible data

This data is already publicly available and findable.

Newly generated data

Digital humanities data

Standard methodology from digital humanities will be used. This primarily consists of quantitative analyses of the published record. This includes for example citation counts or key word searches.

The repository the data is published in will provide for a metadata standard, including categories and keywords in accordance with controlled vocabularies. The file formats will be reasonably open and sustainable (e.g., .cvs, .tsv).

Interviews

Quantitative interviews about the history of science will be conducted. The interviews will be recorded (audio and/or video) and transcribed.

The repository the interview metadata is published in will provide for a metadata standard. The file formats for the interviews will be reasonably open and sustainable (e.g., mp4 for audio, .docx or .odt for transcriptions).

4. Increase data reuse

Publicly accessible data

This data is already publicly available and findable.

Newly generated data

Digital humanities data

This data will be deposited in a data repository. Publications based on the data set will refer directly to the data repository. The published data set and metadata descriptions will be published with an open license (e.g., CC-BY 4.0). The published data are envisioned to remain public indefinitely.

Interviews

This data will not be publicly available since they will contain personal data. Participant consent will be acquired for using de-anonymized data in published articles. The published metadata descriptions will be published with an open license (e.g., CC-BY 4.0). The published metadata descriptions are envisioned to remain public indefinitely. The interview data will be preserved as long as the metadata is publicly available.

5. Allocation of resources and data security

Publicly accessible data

This data is already publicly available and findable. It is likely to have a K1 security classification ('completely non-sensitive').

Newly generated data

Digital humanities data

During the research, the data is stored on Stockholm University's secure cloud service 'Sunet Drive/Nextcloud', with servers within Swedish borders. Controlled access beyond the standards for Sunet Drive/Nextcloud is not required since the data contains no personal information. It is likely to have a K1 security classification ('completely non-sensitive').

Once the research is completed, the data will be deposited in a public repository, with relevant metadata.

Interviews

This data will not be publicly available due to privacy reasons. The interview data is stored on Stockholm University's secure cloud service 'Sunet Drive/Nextcloud', with servers within Swedish borders. This is appropriate for data of K3/K4 security classifications

('sensitive' and 'very sensitive' data, respectively), which is likely appropriate for the interview data. The data saved on the server is password protected and access is restricted to research group members of the DECIDE project. The data will be stored for at least 10 years, or for as long as the metadata is publicly available.

For the handling of personal data, Stockholm University will serve as the controller of the personal data management.

Accessibility and storage of datasets

Where will datasets, documentation and/or metadata be made *accessible*? (multiple options)

- Public data repository (please specify!)
- On direct request from authorized users
- SU Archive