

International
Institute of
Social Studies

Applying data management, privacy and open science principles in your research

Research project with non-European
partners and special category data

NOTE: This is a fictitious description of a research project but based on common factors which occur in quantitative research carried out at ISS. It has been developed at the request of participants of a [research data workshop](#) held at the ISS. The request was to provide a case study which showed how the principles discussed in the training could be applied to a 'real' case study. The research project discussed below aims to highlight the kind of issues which researchers need to consider when designing projects which collect personal data and to bring together relevant advice from privacy, data management, ethics and open science perspectives.

For specific guidance for your research project please contact:

Data Management: datasteward@iss.nl

Privacy: privacy@iss.nl

Ethics: researchethics@iss.nl

Table of contents

Background and ethics clearance

- 1 Who are the partners?
Advice on partnership agreements
What is the research question?
Who are the research participants?
Advice on consent
- 2 How were the participants recruited?
What did the research involve?
- 3 Did you need research permissions or ethics clearance?
Advice on ethics clearance

Data management and privacy

- 4 What data was collected?
Advice on data minimalism
Advice on special categories of data and sensitive data

- 5 Who, in which countries, had access to the raw data?
Advice on agreements in relation to personal data
- 6 How was data collected and shared?
Advice on data collection
- 7 Advice on data sharing
- 8 How was data stored during research?
Advice on data storage
- 9 How is data archived after research?
Advice on long term storage or archiving of raw data
- 10 When will you destroy data?
Advice on destroying data

Open science

- 11 What steps did you take in terms of Open Science?
Advice on pre-registration
Advice on open access data sharing
Advice on open access publishing

Who are the partners?

- **India:** an NGO, a university and three schools run by local government authorities
- **Netherlands:** a PI and two other researchers based at ISS
- **Canada:** an independent researcher



Advice: partnership agreements

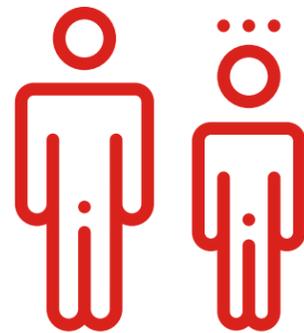
As well as the standard partnership agreements which would be used for these relationships you are very likely to require additional agreements related to data. Please talk to the Privacy Officer early on. S/he can help you work out what is needed. Please also see the section below on agreements in relation to personal data.

What is the research question?

The NGO provides a counselling service (mainly by phone) to families whose children are struggling at school (missing lessons; below average grades; signs of malnutrition). The service is funded by the local government. The research aims to evaluate the effectiveness of this counselling service and test ways of making it more effective, based on current research.

Who are the research participants?

- Children aged between 7 and 12 in three cities
- Their parents or legal guardians



How were the participants recruited?

The NGOs had participated in a previous study carried out by the Indian university. At the beginning of every school year the schools take basic height and weight measurements. They also keep a record of school results and school attendance on an ongoing basis. The schools contacted parents to ask whether they would be willing to take part in the evaluation.

Advice: Consent

Consent is viewed in two different ways. From the perspective of the General Data Protection Regulation (GDPR) there needs to be a legal basis for consent. For most research 'public interest' can be argued, however, from a research ethics perspective, fully informed consent is also a principle of all relevant ethics codes, including that of ISS. You can find ISS' ethics policy, guidelines and sample consent form template [here](#).

It is also possible to justify oral consent if written consent is inappropriate, however, you should then consider recording the consent in some other way, either by audio recording or with a witness present.

Your consent process must be very clear about the following aspects to comply with GDPR:

- Why you are collecting the data
- Who will have access to the data
- What you will do with the data
- When you will destroy the (personal) data
- What you will do with the data if the participant chooses to withdraw consent at any time

Consent/assent of minors/children

In the Netherlands, in order to carry out research with children under the age of 16 you must have consent of their parents/legal guardians. The age of consent varies in each country but in most ethics codes children will be seen as a group which is vulnerable and particular care needs to be taken when dealing with consent. As a safe benchmark, you should seek consent of parents or legal guardians if you are involving children or young people under the age of 18 years old. It is generally considered good practice to also request the assent of children and young people under 18 years old. As a researcher, you will need to adapt the way you ask for assent to the age and capacity of the child/young person, using appropriate language.

What did the research involve?

The research consisted of an evaluation including one control group and one treatment arm. The control group received a general information leaflet. The treatment arm received:

- Weekly counselling phone call with parents and children
- Tailored healthy eating and exercise advice from a trained nutritionist
- Support with setting realistic goals for schoolwork, sleep and exercise
- Follow up with SMS reminders

Did you need research permissions or seek ethics clearance?

Yes. Research clearance from the local department of education was required by law. The Indian University did not require additional ethics clearance once the government clearance was obtained. The research proposal was also submitted to the ISS Research Ethics Committee for approval. The Canada based researcher was not affiliated to an organization for this research so did not submit for any ethics approval, however his contract with ISS stipulated that he was expected to adhere to the ethical agreements of the project.

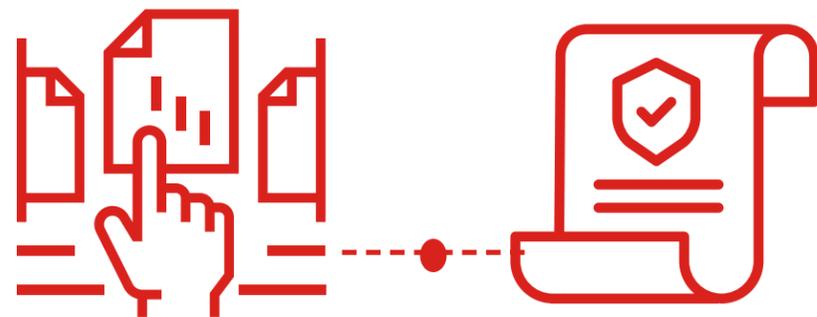
Advice: Ethics Clearance

ISS' policy is that if you conduct any research project (whether internally or externally funded) you must submit a request for approval to the Research Ethics Committee before starting data collection. All details and guidance can be found [here](#).

Although this document does not aim to go into detail on ethical issues, beyond those connected to data management, it is worth highlighting that because the project involves children, and possibly vulnerable children, the research team would need to provide adequate explanations of the different safeguards they would put in place.

Part of the guidance also refers to different options for gaining ethics clearance in the country where you collect data. In some countries, ethics approval will be required by law so you will need to investigate what the requirements are. Partners, particularly academic institutions with which you work, may also have their own policies. Local research partners are often the best source of knowledge. It is recommended that you obtain local ethics clearance wherever possible.

If you are not collecting data but asking organisations or individuals to do this on your behalf, then it is also your responsibility to make sure they are aware of the privacy and ethics standards you yourself are working to. In particular, you will need to draw up agreements with any organisation or individual who is working with, or in any way using, the data.



What data was collected?

Quantitative:

1. A list of parent and child participants with name, gender and email addresses. The list was provided by the NGO.
2. Administrative data containing school results, school attendance and anthropometric data. Provided by the schools.
3. Parental survey results with names, data concerning health, school results, school attendance and date of birth.
4. Child survey results with names, data concerning health, school results, school attendance and date of birth.

Qualitative:

1. Interviews were conducted with a sample of parents and children.



Advice: data minimalisation

One of the main principles of GDPR is data minimization. You need to consider what data is really necessary to answer your research question and not collect more data than is strictly necessary. For example, is it essential to record a specific birth date or would age brackets be sufficient? In this case age in months could be used rather than a specific date of birth. Data minimization also helps at the stage of de-identification of data – the less you have the easier it is to prepare it for eventual sharing and publication.

Advice: special categories of data and sensitive data

GDPR lists 8 categories of 'special category' data relating to: racial or ethnic origin; political opinions; religious or philosophical beliefs; trade union membership; genetic data; biometric data for the purpose of uniquely identifying a natural person; data concerning health; data concerning a natural person's sex life or sexual orientation

As you are collecting two areas of sensitive data you need to take special precautions when storing and sharing this data (see below). You must also be able to provide a very clear explanation as to why this data is essential for your research..

NB: Even if your data does not fall under the GDPR categories you may still consider the data sensitive because it could have negative consequences for the individual if it fell into the wrong hands (e.g. data about income level). If this is the case, you should also treat this data with the same level of care as special category data.

Who, in which countries, had access to the raw data?

- **India:** university researcher (collecting, cleaning, analysis); NGO researcher (collection for part of raw data)
- **Netherlands:** EUR researchers – lead PI and 2 other researchers (cleaning, analysis)
- **Canada:** independent researcher (cleaning, analysis)

Advice: agreements in relation to personal data

GDPR requires that standard agreements which relate specifically to data need to be put in place with all organisations handling data. These will vary depending on two things (1) the country in which the organization, or individual researcher (if unaffiliated) is based (2) who decides what in regard to the research data. In short:

1. Where is the organisation based?

- EU country or country with an adequacy decision by the European Commission: Regular agreement
- All other countries: Regular agreement combined with Standard Contractual Clauses.

2. Depending on who decides what with regard to the research:

- **Joint controller agreement:** where two or more parties are responsible for making decisions about how the same data is processed and are ultimately in charge of and responsible for the processing
- **Separate controller agreement:** parties are individually responsible for their own research – this is usually the case with separate researchers
- **Controller-processor agreement:** one party (controller) is responsible for the research, but another party (processor) uses the personal data, under the responsibility of the controller and for the goals of the controller (e.g. a company that simply stores the research data, but does not use it for their own purposes)

It might sound complicated but help is on hand! Please reach out to privacy@iss.nl who can help you get the right agreements in place.

How was data collected and shared?

1. List of parent and child participants with name, gender, e-mail: the list was provided in an excel file sent by email to the Indian university researcher from the NGO. This was forwarded to the PI by email.
2. Administrative data containing school results, school attendance and anthropometric data: provided in excel files sent by email to the Indian university researcher by the schools. The files were amalgamated and then sent as one file to the PI.
3. Parental survey results with ID number allotted by researcher, data concerning health, school results, school attendance and date of birth: surveys completed on paper and transcribed into Excel by local researcher then emailed to PI.
4. Child survey results with ID number, data concerning health, school results, school attendance and date of birth: surveys completed on paper and transcribed into Excel by local researcher then emailed to EUR PI.
5. Key files linking anonymous IDs to research participants: stored by Indian and Dutch university researcher.
6. Interviews were conducted with a sample of parents and children: handwritten notes were taken by the Indian university researcher, transcribed and emailed to the lead PI. No names were used.
7. The PI uploaded all files to EUR Document Vault (Blackberry Workspace) and did not share further.



Advice: data collection

When collecting personal data and particularly special category and sensitive data you can often design your collection methods to maximise safety and security. For example, in this case, the surveys of children and parents did not include individual names but a number which was linked to a key held only by the Indian university researcher. A rule of thumb is to keep as much data as possible de-identified and to transfer (or, if possible, directly collect) data into a secure environment as soon as possible. For example, interview notes could be immediately scanned and uploaded to a secure online environment and paper copies destroyed.

Advice: data sharing

This data set includes both personal data, special category and sensitive data. Therefore, you need to take particular care sharing the data, both in terms of the tools you use to share data but also in terms of how you connect to the internet and the device that you use. EUR Document Vault is the recommended storage service and an account is available on [request](#). However, if you are dealing with less sensitive data, you could also consider using the EUR licensed versions of OneDrive or SURFdrive or when not sensitive at all EUR Dropbox.

Transfer of files	Recommendation	Rationale
Sending files by e-mail (when files contain identifiable personal data)	Encrypt data with 7-Zip before sending and share password separately.	When a small number of files only needs to be transferred once.
Transferring large files	Use SURF filesender	When files are shared occasionally and worked on individually.
Uploading files to shared storage (when files contain identifiable personal data)	Use EUR Document Vault (Blackberry Workspace)	When files have to be shared regularly with a group of people (and access by the group is needed after sending).

Connecting to internet	Recommended option	Rationale
Wifi	Use Eduroam if available. On any other wifi network use eduVPN .	Eduroam provides a secure connection to internet. On any other (public) network use eduVPN to secure your connection.
Wired	Use only on personal device. Do NOT store or work with (including sending) unencrypted data on shared devices.	When files are shared occasionally and worked on individually.
.../3G/4G (mobile data connection)	Use eduVPN for best security.	Copies of files on shared devices might be accessible to other users of the device (even when deleted).

How was data stored during research?

Raw data set (India) was stored on EUR Document Vault by EUR PI as it needed cleaning. However, it was not shared through EUR Document Vault.



Advice: data storage

Your data set includes both personal data and sensitive category data. Therefore, you need to take particular care storing the data. The next page has some examples of the storage tools you could use. However, it is quite common, as in this case, that all the data sources are merged into one file and/or need to be accessed simultaneously whilst they are being cleaned and analysed. In this case, you should choose the option which fits the most sensitive data. If in doubt, EUR Document Vault provides a very secure storage space.

	Data set	Recommended storage	Rationale
1	List of child and parent participants	Local copy on encrypted drive (internal or external) Cloud copy in EUR Document Vault (no need for encryption) or File encrypted with 7-Zip on SURFdrive or EUR Dropbox.	directly identifiable, containing personal data, small size
2	Administrative data	If not necessary to share outside EUR: a (shared) network folder with restricted access or same as above (nr. 1.)	directly identifiable, containing personal and special category data, small size
3 & 4	Child and parent survey results	Same as above (nr. 1.)	Directly identifiable, special category personal data, medium size
5	Key file	Local copy on encrypted drive (internal or external)	

6	Interview notes – if no personal or other identifying data included	Minimum one local copy and one cloud copy (no security measures required)	Non-identifiable, non-personal data, small size
	Interview notes – if personal or other identifying data included	If not necessary to share outside EUR: a (shared) network folder with restricted access or same as above (nr. 1.)	Directly or indirectly identifiable personal data, small size

NB. Wherever you are working with local copies of data, e.g. if you download your data to work on it, it's really important to create local back-ups.

How is data archived after research?

The research team was not yet sure about long term storage and archiving of the raw data.

Advice: long term storage or archiving of raw data

Before long-term storage, and assuming that personal data may need to be accessed again, data should be pseudonymised. Your data set includes both personal data and sensitive category data. Therefore, you need to take particular care storing the data.

Below are some examples of the storage tools you could use. However, it is quite common, as in this case, that all the data sources are merged into one file and/or need to be accessed simultaneously whilst they are being cleaned and analysed. In this case, you should choose the option which fits the most sensitive data. If in doubt, EUR Document Vault provides a very secure storage space.

Dataset	Recommended storage	Rationale
List of child and parent participants	After pseudonymisation: DANS Easy . Key file at EUR Document Vault .	Secondary data
Administrative data	After pseudonymisation: DANS Easy . Key file at EUR Document Vault .	Secondary data

Child and parent survey results	After pseudonymisation: DANS Easy . Key file at EUR Document Vault .	Primary data
Key file(s)	EUR Document Vault	Administrative data
Interview notes	After de-identification: DANS Easy	Primary data

When will you destroy the data?

The research team has not yet made plans regarding any data destruction.

Advice: destroying data

There is no definitive answer. As a rule of thumb you should not keep personal data for longer than is necessary in relation to the reasons you have collected it. However, data collected for research purposes can be kept for longer periods if anonymised. In short, whilst you continue to process the data, and as long as it is in line with the reasons for which you collected the data, you can continue to keep it.

If research is funded through a grant, the funding organisation will often have its own rules regarding data retention and destruction and you will need to adhere to those. In many instances EUR uses the previous VSNU guidelines of 10 years as a reasonable period before destroying data. Whatever you choose, it is important that you include this information on the consent form.



What steps did you take in terms of Open Science?

Proposed trial was registered at the American Economic Association's registry for randomized control trials ([AEA RCT Registry](#)).

Advice: pre-registration

Open Science incorporates a broad range of issues, not all of which will be appropriate or feasible for all research, however there are some key 'good practice' steps which researchers can consider as they begin to explore the principles of Open Science. The Open Science Framework (OSF) platform includes a wide range of information, support and resources for researchers. One place where researchers can start is pre-registration of their research, meaning that researchers publicly report their research hypothesis which is date-stamped. There are various templates for pre-registration available [here](#). Subsequent changes to the research can also be registered as well as final reports and links to resulting publications and data.

Data will be uploaded to a trusted data repository like [EUR Data Repository](#) (Figshare) or a repository recommended by the specific discipline ([re3data.org](#))

Advice: sharing open access data

Sharing de-identified data is another important step in promoting open science. It means that the data you have collected can be re-used by other researchers to answer other research questions than your own and that your study can be replicated. Sharing data promotes confidence in scientific research and can lead to new collaborations. However, take care that at the point of uploading data, all your data is de-identified. As well as sharing the data, it's a very useful step to share analysis scripts and procedures in order to facilitate computational reproducibility. If point-and-click software was used then providing a list of actions to reproduce the results is invaluable in enabling other researchers to replicate the study.

Advice: open access publishing

Publishing research results in Open Access is becoming the norm in many countries. The [Taverne amendment](#) means that any publication associated with a Dutch university can be made public after 6 months. In addition EUR already has a large number of no-charge open access agreements in place with journals and you can check the publication you are interested in [here](#). If the publisher requests a copy of your research data, you may prefer to share a link to your chosen data storage service.

Even if you choose not to publish in an open access journal, you can make a copy of the author's or pre-print version publicly available through PURE. The Library will check any publisher embargo rules. The Library will check any publisher embargo rules. Finally, EUR has an Open Access Fund and you can find more about this and other options [here](#).