rijksuniversiteit groningen

faculteit rechtsgeleerdheid

# Guidelines for scientific integrity and data protection in student research projects

**Barbara Brink**

Research Ethics Review Committee (CETOR)

version: January 2023

# Content

# 1. Scientific integrity

## 1.1. Code of conduct

When conducting research, students must adhere to the Netherlands Code of Conduct for Research Integrity (2018) of the Association of Universities in the Netherlands (VSNU). Important values included within this code are fairness, due diligence, transparency, independence and accountability. These can be regarded as the 'virtues' of a good researcher, guiding him or her to the right choices in all kinds of circumstances. This means the following for the researcher:

- Fairness includes ensuring that the research process is properly reported, that alternative views and counter-arguments are taken seriously, that there are margins of uncertainty, that no unsubstantiated claims are made, that no data or sources are fabricated or falsified, and that no results are presented as more favourable or less favourable than they actually are.
- Due diligence includes using scientific methods and utmost care in designing, conducting, reporting and disseminating research.
- Transparency includes making clear to others what data the research is based on, how the data were obtained, what and how the results were achieved, and what role external stakeholders played. If parts of the research or data are not disclosed, the researcher must provide a justification as to why this is not possible. At least to colleagues, it must be clear how the research was conducted and what the different phases of the research process were. At the very least, this means that the reasoning must be clear and the steps in the research process must be verifiable.
- Independence means, among other things, that the choice of method, the evaluation of data, the weight given to alternative propositions or the evaluation of other people's research or research proposals may not be based on non-scientific or non-political considerations, for example of a commercial or political nature. In this sense, independence also includes impartiality. Independence is required at all times in the design, implementation and reporting of research, but not necessarily in the choice of research subject and research question.
- Responsibility includes acknowledging that a researcher does not operate as an isolated individual and therefore - within reasonable limits - takes into account the legitimate interests of the participants, the clients, the funders and the environment. Responsibility also means conducting research that is scientifically and/or socially relevant.

The values are used to develop standards for good research practice.

## 1.2. Responsibility towards participants

When conducting research involving participants (e.g., interviews, surveys, observations), the student as a researcher is obliged to adhere to a number of ethical rules. These rules have been drawn up to protect the rights of participants. They are also part of the Netherlands Code of Conduct for Research Integrity. The main ethical rules for research involving human participants are:

- The participant gives written consent
- The researcher provides complete information about the research
- The researcher provides a debriefing and
- offers to keep the information confidential

Consent is one of the legal grounds for the processing of personal data for scientific research. The consent must meet the following conditions:

- Freely given: consent is not valid if the data subject has no real choice, feels compelled or pressured to give consent, or will be adversely affected by refusing or withdrawing consent.
- Specific: the consent must state for which purpose(s) the data will be processed.
- Informed: providing full information to participants is essential to their ability to make informed decisions.
- Unambiguous: the participants' wish to consent to the processing of their data resources must be clear.

The researcher needs permission from the participant to participate in the research project. Research participants should be informed in an accessible and transparent way about the purpose of your research and the benefits, risks and possible inconveniences. You should explicitly state that participation is voluntary, and that the participant has the right to withdraw his/her participation at any time without consequences. Furthermore, if you collect personal data, you must indicate why and how you are collecting personal data, guarantee that you will only use the data for your research, and indicate who has access to the data.

A *sample email* is attached to this guide , see Appendix 1. The email can be used to provide information about the research and explain what participation in the research means. By explicitly agreeing in a response email to this email, participants give informed consent to participate in the research.
This consent email will be kept separately from the interview data.

*Reliability and confidentiality of information*
The reliability of the interview reports is improved by giving participants the opportunity to read the report of interviews and opportunity to give their views on the report and point out inaccuracies. If you plan to publish the thesis, consult the respondents and supervisor first. It is wise to make arrangements with the respondents about this at an early stage, **even before the research project** has started.

### 1.3. Data protection

Research projects involve data, such as interview results, questionnaires, observations or personal documents. In those cases, projects most likely involve data that can directly or indirectly identify individuals. Such data must be handled and stored with the utmost care.

Researchers are guided in data management by a set of legal instruments and policies that ensure that they collect and process data securely and ethically. To assist the researcher in this, the Faculty board has established a Faculty data management protocol.
A specific template for a data management plan (DMP) has been created for students, see Appendix 2.
Central to this are the aforementioned VSNU Code of Conduct, but also the GDPR. The GDPR regulates data protection and privacy for all EU citizens.

Good Research Data Management (RDM) is one of the most important responsibilities of a professional research organization and of individual researchers.
It is strongly recommended that students prepare a DMP, regardless of what data are collected and discuss this DMP **- prior to the study**- with their supervisor.

In the DMP, after a brief description of the project, the researcher indicates what kind of data will be collected, whether this involves (special) personal data, where the data will be stored during and after the project, and other relevant information. Completing a DMP is a good way to think about the data collection process in a structured way. It can also be useful to indicate whether a further ethical review is needed.

For the data management requirements, the Faculty draws upon the GDPR, but also the university research data policy from 2015 and the principles of FAIR, in which case FAIR stands for findable, accessible, interoperable and reusable data management.

The following section provides a checklist for the proper handling of personal data and responsible data management and helps determine the best way to store a research data securely.

### 1.4. Data protection checklist

This checklist will help to ensure that research data are handled in a safe and compliant manner. It focuses on different phases of a research project. It can be used to consult before, during and after the completion of a project.

***Before starting a project***

- Draw up a data management plan. This plan is discussed with the supervisor.
- Think about what data is needed for the project, knowing that you must use as little data as possible (data minimization principle).
- Think about the kind of data needed for the project. Consider the following:

*Type of data*

Personal data
When using personal data, consider anonymization (removing identifying information) whenever possible or feasible. If anonymization is not possible or feasible, apply pseudonymization (replacing identifying information with numbers, for example) of the data. Read more about these techniques on the DCC website (link).

Special personal data
Special personal data, such as data about someone's health, religion or political views, should only be collected when strictly necessary for the project. In such cases, discuss this with the supervisor first.

Confidential data
Even if no personal data is shared within a project, confidential information may be processed, for example sensitive (financial) company data. In that case, it is important to clearly explain how such confidentiality will be safeguarded.

*Collaboration*

In case data are processed in collaboration with other parties, a data sharing agreement from the external organization may be required. In that case, please contact the P&S coordinator for research (Maarten Goldberg, email: m.goldberg@rug.nl).

***Collection of primary data***

- Draw up an informed consent form (see appendix 1).
- When considering using data from social media, i.e.data scraping, discuss this with the supervisor first. Data on social media (personal or not) are controlled by the social media provider, so it is necessary to check the terms and conditions used by the media platform, in advance.
- If data collection tools such as Qualtrics are used, make sure that the tool is only used to collect data and that the data are downloaded as soon as possible to the recommended storage facility for processing and archiving (see section 1.5). Do not leave data in the tool.

***During the execution of the project***

- Only collect data strictly necessary for the project, based on the principle of data minimization.
- Obtain, if applicable, informed consent from each data subject and store the forms adequately; preferably use the Y-drive or the Brightspace group drive.
- Whenever possible, anonymize the personal data you process. If not possible, pseudonymize them. This is especially true for aggregated data.
  Store the research data securely, using the recommended storage, preferably the Y-drive. see 1.5.
- Do not give anyone other than the supervisor access to the personal research data.
- Make regular backups of the work and data sets to the recommended storage.
- Do not use unreliable public Wi-Fi networks, e.g., in a café, hotel or train, when processing personal data such as when forwarding them: connect to secure networks such as “Eduroam” and use a VPN service.
- Use UWP so that a secure connection (VPN) can be established.
- If you use your own device for research, make sure you use legitimate software, such as the operating system, antivirus software and VPN.
- It is preferable not to use an external storage device to store personal data; if you absolutely must, make sure you USB stick or other device is encrypted.
- Hide printed documents containing personal data when you don't need them and don't just leave them on your desk or elsewhere.
- If you walk away from your computer or other device containing personal data, lock the device, by pressing "Windows+L" in Microsoft Windows.

***After completion of the project***

Preferably save the data using the Y drive and discuss with your supervisor how long the data should be kept there.
If desired, also discuss the possibilities of accessing or using the data after completion of the research.

## 1.5. Secure data storage

Choosing an appropriate, safe method to store the research data is a fundamental way to preserve the data and to ensure the privacy of research subjects. Moreover, adequate storage prevents serious situations such as a data breach or loss of data.

If the project contains confidential or personal data, make sure that no one but the researcher and the supervisor has access to it. Do not use Dropbox, iCloud, a personal (private) email or drive to store and share personal data. Instead, use the faculty's storage services, preferably the Y drive or Unishare. For groupwork you can use Brightspace, however Brightspace is not suitable for archiving. Please contact the faculty privacy and security coordinator (Maarten Goldberg, e-mail M.Goldberg@rug.nl) if you would like to use the Y drive for the project.

The table below helps to find the most suitable storage solution, depending on the type of project.

| | University Y drive | Google drive RUG | Personal Google drive | Locked drawer | Personal computer / laptop | USB | Drop-box | Brightspace |
|---|---|---|---|---|---|---|---|---|
| In my project personal data are processed | Yes | Yes | No | Recom-mended | With encryption & only temporary during data collection. | Not recom-mended.[1] | No | For groupwork, part of a course, no archiving |
| In my project special personal data are processed | Yes | No | No | Yes | With encryption & only temporary during data collection· | Not recom-mended. | No | No |
| In my project no personal data are used | Yes | Yes | Not recom-mended | Yes | Yes | Not recom-mended. | Yes | For groupwork, part of a course, no archiving |

[1] If the USB stick is lost, it is not only unfortunate for the research project since data is lost, it is also considered a data breach, which will have to be reported to the Faculty P&S Officer or the UG Data Protection Officer. The advice is therefore only to use this temporarily, for example whilst doing data collection and to make sure the USB stick is encrypted.

# Appendices

## Appendix 1. Sample e-mail informed consent for research participants

Dear Sir/Madam ....

- intro: how to find/approach the respondent

- intro: research

    o Title of the research

    o Brief introduction, including purpose and method

    o Duration of the research

    o Participant contribution (including possible advantages and disadvantages for the participant)

If you have any questions about this, please let me know. You can reach me at...

Through this mail I request your cooperation for this research. Your contribution will include an interview lasting ... minutes.

This interview will be conducted by means of .... I will record this, but only for the purpose to make a report. Once I have completed the report, I will delete the recording.

I understand how my personal information will be obtained, handled and protected. This includes ... [mentioning data protection measures here].

Please be aware that if you consent, you are free to refrain from answering questions and that you can still withdraw until the completion of the research. Please be aware that late notification may mean that the data collected cannot be deleted due to the integrity of the research.

Should you have any complaints during the course of the research, please contact my supervisor, ...

If you have no further questions and are willing to participate, I would like to ask you to indicate your voluntary agreement to participate in this research by replying'' to this email.

## Appendix 2. Template Data Management Plan

This template is based on the Faculty's Research Data Protocol and the Dutch Code of Conduct for Scientific Integrity

| | |
|---|---|
| **1. General** | |
| **1.1 Name & supervisor**<br>State your name, the name of your supervisor and the date this form was discussed with him/ her. | |
| **1.2 Organization**<br>Provide details on the organization where the research takes place | |
| **1.2 Description of the research project**<br>Provide a short description of the project and the central research question | |

| | |
|---|---|
| **2 Data collection – the creation of data** | |
| Do you conduct "classic" legal research (doctrinal research)?<br>Data used consists solely of publicly accessible documents only, such as statutes, rulings, annotations, scientific articles and books. If this is the case in your research, please tick yes. You can then skip the rest of the questions and go directly to the final question and submit the form. You do not need a RDMP. Should you, in a later phase of your research, decide to collect data, you have yet to answer the rest of the questions.<br>Please note: as soon as you use and edit data from the above-mentioned sources , for example to create tables, this is considered creating new data! In that case you will need a data management plan ( RDMP).<br>If you do create data in your research (primary and/or secondary), please complete the other questions below. | Yes/no |

| | |
|---|---|
| **3. Ethics, storage, sharing and archiving** | |
| **3.1 Human subjects and research ethics** | *Explanation* |

| | |
|---|---|
| Does your research involve human participation or the collection (processing) of personal data?<br>If your answer is yes, then you must follow the ethical rules for research involving human subjects.<br>Use the checklist and informed consent template provided in the manual and complete Appendix 3: Personal Data.<br>Obtain consent from your supervisor before beginning your research.<br>*For more information and assistance about research with participants, your supervisor can contact the P&S Research Coordinator at any time.* | 'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person (Art. 4 GDPR).<br>https://www.gdpreu.org/the-regulation/key-concepts/personal-data/<br>NB if you process personal data, please also complete Appendix 3 |
| **3.2 Methods of data collection**<br>What method(s) do you use for the collection of data. Please select all choices that apply | ● Structured individual interviews<br>● Semi-structured individual interviews<br>● Structured group interviews<br>● Semi-structured group interviews<br>● Observations<br>● Survey(s)<br>● Experiment(s) in real life (interventions)<br>● Secondary analyses on existing datasets<br>● Other (explain) |
| **3.3 Where will the data be stored *during* the research?**<br>Where will the (raw) data be stored during the research?<br>Please select all choices that apply | ● On the X drive of the UG network<br>● On the Y drive of the UG network<br>● On a (shared) UG Google drive<br>● Other (explain) |
| **3.4 Will the data be stored after the research?**<br>Are the (raw) data stored after the research? If so, where? | ● Data will be destroyed after the research, it will not be stored<br>● On the X drive of the UG network<br>● On the Y drive of the UG network<br>● On a (shared) UG google drive<br>● Other (explain) |
| **3.6 Sharing of data**<br>Do you share research data with anyone other than your supervisor?<br>If yes please indicate with whom. | |

| |
|---|
| **4 – Final comments** |

| | |
|---|---|
| Do you have any other information that was not addressed in this template that you think might be useful to include? | |

## Appendix 3: Personal data

| Answer these questions when you collect personal data | |
|---|---|
| **A. Categories of human subjects**<br>Categories of human subjects, i.e., persons whose personal data will be processed<br>Explanation: human subjects may be `vulnerable persons’, depending on the circumstances relating to the project and or the research subject.<br>People are vulnerable in the context of the GDPR when:<br>- they cannot freely give their consent or freely oppose participation in the research, i.e., students who have to participate in a research project in order to get credits;<br>- they are not able to form an opinion about the processing of their personal data, e.g., subjects with a mental disability<br>- the processing of personal data about them could be harmful to them, i.e., processing of political opinions of people in countries with oppressive governments, see Appendix 4 | - Adults (not vulnerable) > 18 years<br>- Minors < 16 years<br>- Minors < 16 years<br>- Minors < 18 years<br>- Patients<br>- Vulnerable persons (others) (please provide an explanation) |
| **B. Consent of human subjects**<br>For the processing of personal data in the research, the consent (informed consent) of the research subjects is requested.<br>Explanation: consent in the context of the GDPR is freely given, specified, the data subject must be informed and there must be no ambiguity (the data subject must have no doubt whatsoever about providing their consent)<br>Consent is given by a clear affirmative act of the data subject. | - Yes, informed consent is required and will be obtained<br>- No, informed consent is NOT required: The research subject’s permission (informed consent) is not requested for the processing of personal data in the research. |
| **C. Informing participant**<br>Have you informed the research subjects about the data collection procedure and about any information that may be relevant to their decision to participate or not to participate in this research, such as debriefing procedures? You can answer `no’ to this question if this is a natural field experiment.<br>“A natural field experiment is conducted in a different environment than the lab where the subjects naturally perform these tasks and where | Yes / No<br>*Explanation*<br>You can answer `no" if this project is a natural field experiment in Harrison and List's taxonomy (JEL, 2004)<br>`A natural field experiment is conducted in a different environment from the lab where research subjects naturally perform these tasks and where research subjects do not know they are in an experiment.’ |

| | |
|---|---|
| the subjects do not know they are in an experiment.” (JEL, 2004) | |
| **D. Categories of personal data that are processed**<br>Select all choices that apply | **0 General personal data** such as:<br>- Name and address<br>- Telephone number<br>- IP-address<br>- E-mail address<br>-<br>**0 Special personal data**:<br>– Nationality<br>– BSN (Citizen's Service Number) or V-number (registration number)<br>– Information revealing racial or ethnic origin<br>– Information revealing a person's political opinions<br>– Information about a person's physical health<br>– Information on a person's mental health<br>– Information about a person's sex life or sexual orientation<br>– Information revealing religious or philosophical convictions<br>– Information revealing membership of a trade union<br>– Biometric information<br>– Genetic information<br>– Criminal record<br>– Other (please explain below) |
| **E. Providers of personal data**<br>Who will provide the personal data to be used in the research?<br>If the personal data will be provided by an external party (not being the student or the UG), please describe the party/parties and indicate for each party whether an agreement has been signed. | - Data supplied by the researcher<br>- Data supplied by the University of Groningen<br>- Data have been supplied by an external party (please explain below)<br>*Explanation:*<br>These are agreements which contain provisions on the handling of personal data used in your research. In the case of interviews or surveys, the researcher provides the data himself. |
| **F. Technical / organizational measures**<br>Select which of the following security measures are used to protect personal data.<br>See the DCC website for more explanation (link) | - Pseudonymization<br>- Anonymization<br>- Encryption of storage<br>- Encryption of transport<br>- None of the above<br>- Other (describe below) |
| | |

| | |
|---|---|
| **G. Personal data transferred outside the EU**<br>Is personal data transferred to countries outside the European Economic Area (EU, Norway, Iceland and Liechtenstein)?<br>Please note: After 1 May 2021 the UK is also considered a country outside the EU. | |
| If you would like any assistance in answering these questions, please contact your supervisor. Your supervisor can contact the P&S Research Coordinator at any time for more information or advice.<br>P&S Research Coordinator contact information:<br>Maarten Goldberg<br>E: m.goldberg@rug.nl<br>T: 06 319 83 053 | |

## Appendix 4. Additional information on processing personal data

*What is meant by vulnerable participants?*
Vulnerable participants are defined as those who have (or think they have) a dependency relationship with the sponsor or implementer of the research. This dependency may be psychological, social, economic, political, or other. A dependency relationship can result in the research subjects' feeling compelled to give consent or, on the contrary, not daring to refuse the processing of their data. Examples include employees or students of the UG, children vis-à-vis their parents/carers/school, but also other vulnerable persons, such as patients vis-à-vis their health care provider/insurer, people entitled to benefits vis-à-vis the municipality or UWV, prisoners vis-à-vis the judicial authorities, and so on.

*What is general personal data?*
European law provides a legal basis to support the fundamental rights to privacy and data protection of legal and natural persons {individuals, companies and institutions}. The General Data Protection Regulation (GDPR) ensures the protection of individuals in relation to the processing and sharing of personal data. To understand which rules apply it is necessary to distinguish between the various kinds of data.

The GDPR stipulates that `a natural person (data subject) is identifiable when it is possible to identify him or her, directly or indirectly, by using certain identifiers. Examples of personal data are dates of birth, addresses, telephone numbers, e-mail addresses, social media pictures, or other background information about a specific person.

In research, personal data is collected by researchers and students, mainly through surveys, interviews, questionnaires, data collection software, etc.

Please note: as combinations of data could be used to identify a person, this data – which may at first not look like personal data – must be counted as such!

For example:

| | |
|---|---|
| the combination 'name X'<br>'man'<br>'owns a Toyota'<br>'in the Netherlands' | = directly identifiable (person X) |
| the combination 'man'<br>'owns a Toyota'<br>'address'<br>'in the Netherlands' | = indirectly identifiable (address leads to person X) |
| the combination 'man'<br>'owns a Toyota'<br>'in the Netherlands' | = truly anonymous |
| the combination 'man'<br>'owns a Toyota'<br>'in Groningen' | = given the number of Toyota owners, probably still truly anonymous |
| the combination 'man'<br>'owns a Lada'<br>'in the Netherlands' | = unclear, because how many people own a Lada in the Netherlands? |

| | |
|---|---|
| the combination 'man'<br>'owns a Lada'<br>'in Groningen' | = probably indirectly identifiable (the combination leads to person Y) |

As you can see, data that does not directly identify a person', but must still be considered to be personal data if it is reasonable to assume that the data could be combined, either now or in the future, with other data to identify a person .

This includes datasets that use unique identification numbers, even if you as the researcher really do not know who the numbers refer to, or datasets with unique combinations. This is pseudonymized, and therefore personal data.

*What is special (or sensitive) personal data?*
The GDPR brings a special focus on sensitive data, and stipulates that we must be extra careful with that data. According to the GDPR (art. 9), the following personal data should be considered as special category (sensitive) personal data:

- personal data that reveals racial or ethnic origin;
- personal data that reveals political opinions;
- personal data that reveals religious or philosophical beliefs;
- personal data that reveals trade union membership;
- data about a person's health;
- data concerning sexual behaviour or sexual orientation;
- genetic data (data that provides unique information about physiology or health and/or the health of family members);
- biometric data that allows the unique identification of a person (data that provides unique information on physical, physiological or behavioural characteristics).

Special category (sensitive) personal data explained by the "Autoriteit Persoonsgegevens" (in Dutch)

Special category (sensitive) personal data by the "UK Information Commissioner's Office" (in English)

*Conducting research with special personal data*
To do scientific research with special personal data explicit permission must be requested from the research subject for the processing of these data. Of course, in addition, all GDPR conditions apply to be able to process ordinary personal data.

See also the DCC website:
www.rug.nl/research/research-data-management/data_protection-gdpr/consent

Note: If there is a high risk of misuse of personal data you should first perform a "Data Protection Impact Assessment (DPIA)".  Please check, using the "DPIA scan" in the Appendix of the Readers Guide, whether this may be the case.

*Why is the data provider asked?*
If you do not collect the personal data yourself but receive it from another party, or if you have personal data that you have collected yourself is processed by a party outside the UG, an agreement such as a 'data sharing agreement' or 'data processing agreement' must be drawn up for this purpose in order to comply with the requirements of the AVG.

*What are technical and organizational measures?*
Technical measures are, for example, pseudonymization and encryption. Organizational measures include not storing data on devices or keeping consent forms and transcripts separate.

More information on the website of the DCC:
www.rug.nl/research/research-data-management/data_protection-gdpr/measures/

www.rug.nl/research/research-data-management/data_protection-gdpr/consent

*Why is it asked if personal data is transferred to countries outside the European Economic Area?*
The GDPR imposes restrictions on the transfer of personal data outside the European Union, to third countries or international organizations, to ensure that the level of protection of individuals provided by the GDPR is not undermined.
In this case, it must be demonstrated by the researcher that it is necessary for the research to transfer these data, for example because there is a research partner in those countries and what (additional) measures have been taken to achieve the same level of protection. For example, by working together on the data in a virtual workspace (Virtual Research Workspace).