

- ▶ The checklist should help students work with the data according to the best research practices and prevent the risks of research misconduct and questionable research practices.

- ▶ Depending on the research field, your experience, and institutional policies and rules, the list might need to be adapted to your needs.

Plan & Preparation

Reasons

- The reasons for collecting new data are clearly stated.
- It is clear what the potential impact is on the subjects and the environment.
- It is clear how the data will benefit people.
- It is clear whether sensitive data will be processed.

Consistency & Responsibility

- The project is created according to the best international standards.
- The project takes account of broader technical integration and harmonisation.
- The data management plan is defined.
- The project identifies the leader, data controller, and other relevant roles.
- The project team reflects diverse opinions, backgrounds and various kinds of thought.
- All people involved have the necessary knowledge and skills.

Consent

- The research has ethical approval.
- Informed consent requirements have been determined.
- The consent form clearly explains what the users are consenting to.
- The consent form considers the rights of people unable to provide consent.
- Data subjects have explicitly provided consent (if humans are involved).
- Relevant authorities have explicitly provided consent if animals or the environment is involved.
- If the data come from a different project, the original consent is documented.

Acquisition

Primary data

- Data reliability is ensured (we know who collected the data and how).
- Only necessary and relevant data are collected from respondents.
- An efficient and valid data collection method was used.
- The sample of respondents corresponds to the target group.
- If necessary, the data are anonymised.

Secondary data

- Secondary data comes from a reliable source.
- We know where the data came from and who owns them.
- All sources of secondary data will be referenced transparently.

Storage & Protection

Storage

- It is clear where and how the data will be stored (what device and format).
- The data are stored only in a designated and protected repository and are not copied anywhere else.
- It is specified how and where the data will be backed up.
- There is a recovery plan in place to preserve the data for the future.

Protection

- There is a plan for how to protect and secure data with an appropriate level of security.
- Data comply with the necessary security requirements for the data security category.
- It is defined who has access to what type of data, or at what level of access.
- Only authorised persons have access to the data.
- Protection of the collected data during the transfer to the repository is ensured.
- Protection mechanisms such as data minimisation and anonymisation are in place.
- If sensitive personal data are involved, the impact on data protection will be assessed.
- Data encryption is used to protect data for privacy concerns.
- The identity of individuals is sufficiently protected.

Usage

Transparency

- Users are informed about what data are being provided and what is being done with them.
- The solutions developed by the project are open and can be used by others.
- All data modifications are fully traceable.

Consistency control

- The data are statistically consistent with the sample studied.
- It was stated how the technology could be attacked or abused.
- Possible sources of data bias are understood.

Processing

- The data are processed using appropriate methods according to their nature.
- The training data were tested to ensure that they are fair and representative.
- Correct functionality of the algorithms has been tested.

Visualisation, interpretation, & publishing

- Data modification has been notified.
- Data are visualised in a way that does not obscure their meaning.
- All potential limitations, biases, and conflicts of interest are clearly declared.
- The possibility of outliers and why and how they should not be omitted from the data set were considered.

Consequences

- It has been considered how the results might cause harm to an individual or group.
- Mechanisms have been set up for redressing those harmed by the results.

Archiving & reuse

Archiving

- It is clear where and how the data will be archived (what device and format).
- Local standards or regulations for archiving are taken into account.
- It is clarified whether the data will be open or private.
- Open data are publicly available and easy to find.
- It is defined who has access to what type of data, or at what level of access.
- Only authorised persons have access to private data.
- Cryptographic algorithms have been applied where necessary.
- Measures have been established to prevent the loss of archived data (physical and digital).

Reuse

- The data reuse has ethical approval.
- If data are shared with third parties, mechanisms for protection are in place.
- When data are shared, the associated metadata are also provided.
- It should be clearly stated whether the reused data are obsolete.

Destruction

- It has been verified that the data must be deleted in accordance with the specific rules.
- All copies of the data have been properly deleted.