Started from "Letter about scientific integrity" by Prof. M Verbeek in 2012 (MV/tv 0012.003840), applied by S Eriksson to Swedish research. PhD students in SE's research ethics courses at Uppsala University have continuously revised them. They promotes re-use of data, generation of better data sets, visibility, credibility & more. Data management takes time and effort to implement, but it also facilitates easy access to data for others as well as for ourselves.

General recommendations for the handling of research data (version 2.2)

1. *Raw data*: If possible, maintain the original, raw data (including laboratory notebooks, samples, specimens, photographs, scanned papers, etc.); or, if not possible, document them such that the researcher and/or data collecting facility are able to convincingly demonstrate that the original research data has not yet undergone any selection, purification or transformation steps.

2. The *data collection* or *generation* process should be clearly described in the research records through meta-data. This includes aim, equipment (logs), software (use Git to track changes), corrections made, primary reference results, etc., but also the dates, names and roles of the researchers involved and organizations providing data (i.e. research agencies or collaborators). Make descriptions detailed so that it is possible to trace the collection process or to reproduce it.

3. *Analysis*: Document data input, analysis procedure and troubleshooting in detail (SOP), so that the analysis/simulation gets replicable. Try analysis blinded. Provide any code written especially for the analysis. Use freely available software with version number noted. Store identifiable & fully described data sets for each crucial data compilation, purification or transformation step.*

4. All raw data and the documentation of the data collection, input and analysis process should be *stored* safely, electronic data on a central server with backup or in duplicate to prevent accidental deletion. No original data should be removed from the research body without explicit permission from authorized officials. Only use authorized cloud solutions.

5. Data that are e.g. sensitive or protected by secrecy should be *securely* handled (by measures such as 'pseudonymization', logging access, encryption or password protection).

6. Data should be preserved according to national regulations on *archiving*, and, if possible, kept in a standardized format that facilitates the aggregation and re-use of data. See examples here.

7. To ensure *public access* to data: describe the data set, and order and list data as well as the accompanying research records in such a way that a layperson can understand what is archived.

8. *Open data*: It is good practice to register or publish data sets in open repositories (such as Dryad, GDC or dbGap) when or after reporting results, or after filing a patent application.

9. Original data and research records should only ever be *deleted* for justified reasons and after results have been made public, and reasonable time has passed for verification of research results (often at least 10 years after publication is believed to be reasonable).

10. It is also good practice to use resolvable and persistent identifiers such as e.g. *digital object identifiers* (DOIs) and make use of them e.g. in citations or on a CV, in order to distinguish the data and make them traceable (www.doi.org).

* Crucial steps transform data such that it is impossible to revert to the rawer data when only the transformed data is available.