Making your data secure

Certain categories of data produced or collected for research require greater security, with particular attention being paid to the conditions in which data is stored and shared.

To ensure the physical security of your data and prevent loss, you should comply with good storage practice.

Although the GDRP (General Data Protection Regulation) allows work on personal and sensitive data for scientific research (renvoi vers démarche 22), it nevertheless requires you to take specific precautions.

In addition to using secure storage, it is recommended that you **encrypt** this data during the project. Encryption guarantees the confidentiality of data by only allowing legitimate people to access data and content. Files are encrypted using a key: this key is then required to decipher the documents, be it by humans or machines. Without the key, the data is unreadable. Several software packages allow you to simply encrypt data and/or file directories yourself.

During the project it is also recommended that you **pseudonymise** data, that is, replace any directly identifying data (family name, first name etc.) in a dataset with indirectly identifying data (aliases, sequential numbers, etc.).

Pseudonymisation also provides a way of processing individuals' data without being able to directly identify them. But insofar as it is frequently possible to recover their identity using third-party data, making the operation reversible, the data concerned retains its personal nature.

In order to enable broader exploitation of this data and its reuse by third parties, another option is available: **anonymisation**. This removes the identifying nature from a dataset, meaning the data can be disseminated without infringing the privacy of the people concerned. It is based on a set of techniques for randomisation and generalisation.

Encryption and the anonymisation of data are irreversible steps. The decision to choose one or the other depends on the uses envisioned for your data, and needs to be anticipated as early as possible in a research project: anonymisation entails a necessary loss in data quality, but enables broader dissemination of datasets; encryption preserves all the information, but drastically limits possibilities for accessing data and its subsequent reuse.

Contact
For questions about research data
guichet-ardoise@groupes.renater.fr
Fiona Edmond
Research data manager - University Library
fiona.edmond@univ-rennes2.fr
Morgane Mignon
Coordinator of the Digital Humanities platform - MSHB
morgane.mignon@mshb.fr
Further information

- Personal Data : definition (CNIL)
- English-French glossary on data protection (CNIL)
- Anonymisation of personal data (CNIL) (French)