

## ***Data integrity***

The role and purpose of 'data' in research is so important that understanding the various aspects of data handling within a project cannot be underestimated. Whilst many researchers are familiar with handling data within their own type of research work, many are not as familiar with basic standards and definitions about what constitutes data and is therefore subject to regulation and duties relevant to maintaining integrity.

### Who owns the data?

If your research is self-funded or supported or is an independent piece of work, then it is most likely that you as the researcher own the data that you have produced or collected. However, in many circumstances, the University may have rights over data if the research was granted resource by the institution or if you have used University resources or equipment substantially in the research. If your work has been funded by a third party, a funding body or industry, then it is most likely that the funder or sponsor of the research owns the data and may have rights to their use which go beyond those of the researcher. You must, as a researcher, be clear about ownership of the data and cannot automatically presume that you own the data. This is particularly important if, for instance, you change institutions. You should also be aware that if the data are owned by an institution which is classified as a 'public' body, as a university is, then data might be subject to a disclosure request under the Freedom of Information Act, even though certain exemptions from disclosure are in force for research data. Such exemptions depend on a clear path of research output into the public domain, so are not unlimited.

## Defining 'data'

Data' is defined for the purposes of research integrity as anything which is gathered, collected, stored, and analysed and/or processed with the aim of obtaining a result or outcome forming or contributing to new knowledge. Data is only 'data' in this sense when it is framed and contextualised for the purposes of analysis. However, anything can count as data if pursued for the purposes of research. Data is the 'information' that you generate, collect and then store in some form during a project and all forms of data, regardless of the stage of research or the form it takes, have implications for the integrity of the research.

It is important to bear in mind that objects and other media can be data. It is important to bear in mind whether an object is data itself or whether it is merely instrumental to the data of interest. For instance, if an object is being measured, does the object itself constitute data, or are the measurements only considered as data? A scientist studying rocks may consider the rocks themselves as data, as well as any measurements obtained from the rocks; whereas, a scientist studying animals is less likely to treat the animals themselves as data, but only any measurements obtained from them. Likewise, working with texts provides some difficult areas. Where research data is derived from writings of authors, there is a difference between the writer (*not* data); texts written by that writer (*possibly* data, depending on what use is made of them as a whole); the words and sentences that constitute those texts (most likely *instrumental* to the data); the ideas that are communicated by those words and sentences in the texts (*data* when subjected to analysis). Audio or video recordings constitute data if subjected to analysis and especially so if the recordings have been made by the researcher, in which case everything captured on the recording is considered as data. Web-based materials become data when the subject of analysis. Other areas of difficulty are observational materials, such as observations of behaviours exhibited by members of the public: these become data when subject to analysis in research.