Data curation concerns the maintaining, preservation and the adding of value to research data throughout its life cycle. Data needs to be actively managed so that it can be discoverable, understandable and reusable. Due to how relatively cheap it is to create huge amounts of data, there is a significant amount of data to be managed. The effort spent on curation should be balanced with the value of the data. The skills of archivists and records managers can be put to use in the management of research data, in particular their understanding of provenance and context, and also in determining how long different research data should be retained and how to preserve it over long periods of time.

The management of research data will potentially require working with research data managers who assist researchers in the management of their data. Data managers will often support researchers in a number of areas:

- **Planning** – Assisting researchers before they start their research in helping them plan how the researcher will manage their data during the course of their research.
- **Providing the infrastructure** where data can be stored securely and accessed and shared appropriately.
- **Stewardship** – Adding value to the data by linking the research to the bigger picture by describing datasets and ensuring the information can be found.

Data managers will most likely be found in universities. It may be the case that there is a lack of awareness of the skills archivists and records managers can bring to the management of data but the professions will need to work together. There may also be challenges at determining who holds which responsibilities in regards to managing the data. It is important that the electronic data is managed consistently with other information across the organisation.

When managing research data, it is important to establish what it is you are trying to preserve. Many researchers want to reuse data and to be able to do so they require the contextual information surrounding the data including: the workings; the notes; and the processes of the research; so that the research can be reconstructed. It is therefore important to preserve accompanying records and not just the data itself. Data can exist in many different forms and have many different uses at the same time and decisions will need to be made on how much of the data to keep – will a sample be sufficient? Sometimes data is edited/cleaned to correct errors and decisions will also be needed on whether all versions of the data should be maintained. Often research data may come to the archives/records centre without sufficient contextual information. In this case it will be necessary to work with others including those in the relevant research field and if possible the researchers themselves to make sense of the data and accompanying records.

When promoting collections, it may be necessary to encourage the use of datasets and explain to researchers how they can be used. Like with other types of information, research datasets may
contain sensitive/personal information and the data will need to be managed appropriately and in line with relevant legislation such as the Data Protection Act.

Some interesting case studies in which research data has been reused were discussed during the day:

- Ian Dreary, Professor of Psychology at the University of Edinburgh, discussed how the reuse of Lothian Birth cohort study data (intelligence tests given to all 11 years old in Scotland in 1932 and 1947) has been essential to his research. As a sample of a whole section of the population has been preserved, the data has been able to be linked with other datasets (leading to at least 300 papers).
- Victoria Cranna, Archivist and Records Manager at the London School of Hygiene and Tropical Medicine, discussed how records from Ebola outbreaks in the 1970s were used following the Ebola outbreak in 2014 to see what lessons could be learned from the data.

In summary, research data can be a valuable record and like other records it is important to ensure that the accompanying records and contextual information are also preserved. When dealing with research data:

- you will have to determine what precisely it is that you are trying to preserve
- who you will need to work with to ensure that the data is managed throughout its lifecycle and what everybody’s responsibilities will be
- make decisions based on retention similarly to other records, also taking into account future use of the information for research and funder requirements
- how to preserve the data which may be held in complex file formats over long periods of time
- consider any governance issues surrounding the management of the data, for example in regards to access

Useful resources:

The Digital Curation Centre http://www.dcc.ac.uk/

The Research Data Alliance (RDA) has a group for archivists and records managers who are not currently well represented and need to work on how to communicate with the rest of the research community. Further information can be found at the following url: https://rda-alliance.org/groups/archives-records-professionals-for-research-data.html

The University of Edinburgh have information on their website regarding their data management program: http://www.ed.ac.uk/information-services/research-support/data-management

Some information on how The National Archives have archived datasets can be found at: http://www.nationalarchives.gov.uk/webarchive/archiving-datasets.htm