How to Deal with Non-digital Data Guide

The benefits of digitising data

Non-Digital Data

Not all research data are digital. Most researchers keep handwritten laboratory notebooks, journals and other materials, examples of which may be surveys, paintings, fossils, minerals and tissue. However, non-digital data can be converted to a digital source in a variety of ways.

Why digitise data

- Increased efficiency: with a well-defined digitisation and document imaging plan, stakeholders can easily share, collaborate, compare, exchange and access documents, thus reducing waste of time and effort.
- Cost efficiency: the cost of printing and paperwork can be excessive.
- Ease of access: digitised objects can be easily accessed through the cloud or network using any device that has internet, anytime and anywhere.
- Security: file permissions and access rights to specific users or groups can be defined if needed. This will increase security and maintains the confidentiality of the document.
- Long term preservation: physical objects and information stored on, for example, paper are subject to degradation, and will degrade further every time they are handled manually.
- Data recovery: in the event of natural or manmade disasters, loss or damage of physical objects (data) can happen very easily. The digital form of these in the cloud or managed services would not be subject to the same threats, thus reducing risk.

This section is based on reference [1].

Digitising versus born-digital

There is a need to distinguish between those records that may have been digitised from a paper record or a real-world object *vis-à-vis* those that existed purely in digital form (born-digital), such as digital photos, CAD drawings, emails and so forth. [2] This guide relates to the former and the two classes may have different curation lifecycle models, since the physical object will also be subject to some curation, even if this may be its destruction.

Golden copy

The master version of any data is typically the golden copy and also implies a record that is of the highest quality. The golden copy of a record exists for all stages of its development. [3]

Retention period

Typically, basic research data and related material should be retained for a minimum of 10 years after the study has been completed. Clinical research data should be preserved for at least 20 years after the study has been completed. [4] PLease keep in mind that national legislation or codes of conduct may impose different periods.

Text digitising

Anything stored on paper can be scanned easily. For best results, scan your paper document as an image file, preferably as a TIFF image, or PDF. Using the latter method allows files to be searchable and Optical Character Recognition (OCR) is a process which allows conversion of text contained in images into editable text files.

If you are scanning sensitive data, we recommend scanning to an encrypted USB drive rather than to your network drive. However, you should transfer the data to a secured network storage space as soon as possible, then store the USB drive securely until you are ready to ask your IT support to perform secure data deletion. [5]

Video and audio digitising

Analogue video or audio recordings can be converted into digital sound or video files, and there are many hardware and software options available to do this. If the actual recordings are not required, and if, for example, only the words are needed, they can be transcribed and the video or audio data deleted.

Real-world objects

Those who carry out material structure studies, or restoration of museum pieces, often work with real objects - like biological material, fabrics, fossils and paintings. Digital technology allows them to observe these items in several ways. Cell and tissue specimens may exist as intact structures or may have been sectioned and adhered to glass slides. In the former case, there are numerous methods to 3D scan specimens such as various forms of tomography which can go from sub-nanometer to millimeter resolution. These will typically yield a 3D digital object which can then be virtually sectioned. In many cases, they are also used to discern physical characteristics or experimental signals that would not otherwise be possible. For sections, it is common practice to directly photograph these to create a digital record, and is the same for paintings. Other types of specimens such as fossils and minerals can also be

subjected to tomographic imaging methods to produce 3D virtual objects, while in other cases the experimental subject may be very large, such as a building or an entire landscape, in which case there are more expansive methods such as SONAR, RADAR and LIDAR.

If the real world object is not easy to scan, the only remaining option may be to take a digital photograph(s) at the highest resolution possible. The image should be checked to make sure it is of sufficient quality and provides enough detail, and is an accurate reflection of the real world object.

Advice for managing data in hard copy

This section is based on reference [6].

- Appoint an archive administrator and a deputy, and define clear key management duties.
- Make sure that only the archive administrator and their deputy have access to the physical part of the archive.
- Make sure that there is a clear procedure in place for archiving and reusing documents.
- A good archive is well organised. Use archive boxes and label them with the project name and number. Show information such as the name of the project leader and a serial number. For more information on documenting a paper archive, see the EAD standard for Encoded Archival Description.
- Create a corresponding digital database and log the information which is most relevant. Indicate the conditions according to which the contents of the archives can be accessed, as well as those governing their use. Keep a clear record of who has used the archive.
- Separate anonymised data from informed consent forms, for example, by keeping the forms in a locked cabinet in the archive.
- Keep the archive clean and up to date. Make sure that the contents of the physical archive match the records in the corresponding digital and hard-copy databases.
- Ensure the clarity of the procedures governing the removal or the destruction, or both, of archive materials.
- Records storage areas should be secure, clean, organised, safe, dry and accessible. If possible, the records storage should be a fireproof room. These are expensive, but very effective if set up properly with fire resistance for a minimum of four hours, a fire detection system, temperature and humidity control, dust-free conditions, and a secure locking device or some controlled access systems.

References

[1].https://www.aptaracorp.com/blog/10-advantages-digitization-and-data-capture-you-must-know

[2].http://www.nationalarchives.gov.uk/information-management/manage-information/digital-records-transfer/what-are-born-digital-records/

[3].https://www.ed.ac.uk/records-management/guidance/records/retention/golden-copy

[4].https://mrc.ukri.org/publications/browse/good-research-practice-principles-and-guidelines/

[5].https://library.bath.ac.uk/research-data/working-with-data/non-digital-data

[6]. https://www.ru.nl/rdm/archiving-data/archiving-non-digital-data/