



The National Archives

Datasets preserved by The National Digital Archive of Datasets

Request for information about datasets preserved by The National Archives, via The National Digital Archive of Datasets (NDAD)

Questions asked and information provided:

[What does the National Digital Archive of Datasets preserve? Are the datasets actual records, or a finding aid providing the location of actual records?](#)

The National Digital Archive of Datasets preserves actual electronic records and it contains original data generated by HM Government. This often takes the form of surveys and reports. Users can query and analyse the data according to their requirements.

[Who decides what should be preserved on the National Digital Archive of Datasets, and in which format? Do you have any criteria, or a policy regarding what should be preserved or discarded?](#)

Material for the National Digital Archive of Datasets is selected in accordance with The National Archives' Operational Selection Policy (OSP), which can be accessed via this link:

<http://www.nationalarchives.gov.uk/recordsmanagement/selection/osp-subject.htm>.

The format for storage is relational tables based on extraction through, in general, comma separated variable tables.

[How do you guarantee the authenticity or fixity of the material and ensure no one can tamper with it?](#)

The datasets are stored securely under archive conditions. Fixity metadata is held in a database along with other object metadata. An automated process checks all database entries against the file system, and produces a report if any file cannot be accessed or if the metadata differs from that held in the database. Access to the datasets is provided by online presentation. Users are not able to change the data presented online.

[Who is responsible for the validity of the URLs of your archive?](#)

Both The National Archives and the University of London's Computer Centre, as service provider, are jointly responsible for the website and its contents.

How do you deal with the provenance issue? Can an end user know the origin of the material he finds?

The National Archives' Records Management and Cataloguing Department source the data and metadata in the originating Government Department. The data and metadata are then collected and checked for completeness and integrity, with the assistance of the University of London's Computer Centre. Information about provenance appears on the National Digital Archive of Datasets' online catalogue, in the entries for 'Source of acquisition' and 'Digital processing and conversion'.

Who is responsible for the technical infrastructure and for ensuring that the digital content will always interact with its digital surroundings?

The University of London's Computer Centre is responsible for managing and operating the National Digital Archive of Datasets on behalf of The National Archives. The University of London's Computer Centre is therefore responsible for ensuring that the service continues to operate effectively. This includes migrating to new hardware devices, software packages, or media if any of the current operating systems become obsolete.

Who is responsible for the whole archive? Is it the government with the support of the University of London, or are there any other sectors involved?

The data remains in the legal custody of The National Archives, but is managed by University of London's Computer Centre, as our service provider.

How do you deal with technological obsolescence?

There are two aspects to technical obsolescence. The first relates to the media on which the information is stored; the second to the format in which the information was written.

Media: The National Archives has the facility to read some obsolete media (such as floppy discs) and to transfer, or migrate, it to the latest media. The data can then be processed on in-house servers. Much of the information The National Archives receives currently is stored on up-to-date media, so this is less of a problem.

Format: With regard to format, The National Archives has the expertise to export data from proprietary software, even if it has been transferred to us in an earlier version (eg DBMS). The data can then be migrated to the appropriate format for preservation.

Date of disclosure: 12 April 2007