



OPERATIONAL SELECTION POLICY OSP30

GOVERNMENT AND PEOPLE: THE INTERACTION OF THE STATE WITH THE CITIZEN AS DOCUMENTED IN ELECTRONIC CASE FILES AND DATASETS

March 2005, Updated March 2006

Contents

1. Authority
2. Scope
3. Paper datasets and case files series
4. Electronic datasets and case files
5. The Selection Process
6. Summary Guide to the Selection Process
7. Repeated and On-going datasets
8. Dynamic datasets
9. Places of Deposit
10. Access

1 Authority

- 1.1 The National Archives' Acquisition Policy announced the Archives' intention of developing Operational Selection Policies across government. These would apply to the records created and stored by individual departments and agencies the collection themes described in the overall Acquisition Policy.
- 1.2 This Operational Selection Policy concerns the interaction of the government with the people, and is therefore intended to be applied across all central government departments where such interaction takes place. It is part of The National Archives' stated objective to ensure that our policies for records selection are socially inclusive¹. The Policy will detail a methodology for identifying and evaluating relevant material at the top level of state and citizen interaction. After publication and approval, specialised Policies addressing particular areas of state citizen interaction for particular subjects or departments will follow. (The datasets mentioned in the text are examples cited in order to facilitate the application of the policy).
- 1.3 Operational Selection Policies are intended to be working tools for those involved in the selection of public records. This policy may therefore be reviewed and revised in the light of comments from users of the records or from archive professionals, the experience of departments in using the policy, or as a result of newly discovered information or the development of new core datasets.
- 1.4 There is no formal cycle of review, but comments would be welcomed at any time. The extent of any review or revision exercise will be determined according to the nature of the comments received.
- 1.5 If you have any comments upon this policy, please email records-management@nationalarchives.gov.uk

Or write to:

Acquisition and Disposition Policy Manager
Records Management Department
The National Archives
Kew
Richmond
Surrey
TW9 4DU

2 Scope

¹ Work Plan 7 of the Corporate Plan 2004-05 to 2006-07 and Business Plan 2004-05 (three year objective)

In the development stage of this Operational Selection Policy various labels have been employed: it was originally concerned with documenting the Life Cycle of the Individual. This assumed it was possible to trace an individual's life through central government records. On the one hand, the main stages in an individual's life (birth, marriage and death) are already recorded and available in non-public record sources; on the other, many aspects of the individual's life are either the province of local government, (such as the compilation of electoral registers, or education), or are not captured as government records, (for example, employment, or moving house). The scope of this Policy is therefore more accurately defined as:

Information on the events in the lives of diverse individuals as occasioned by the interaction of the state and the citizen

This brings it into line with the relevant collection theme of the Acquisition Policy: *Interaction of the state with its citizens*. Furthermore, in order not to confuse potential users, it was felt that a title which focused on government records would be more appropriate, hence it has been re-titled "Government and people: the interaction of the state with the citizen".

Researchers must bear in mind that The National Archives can only select records that have been created by government departments and has no influence over the content of records, or the amount of detail about specific groups that may be in them.

3 Paper datasets and case file series

This Operational Selection Policy is primarily concerned with the appraisal of electronic records. Paper records documenting the interaction of the state and the citizen are by their nature voluminous, often difficult to sample and expensive to appraise, process and store. These problems have not fundamentally changed since the Public Record Office last reviewed its policy on the selection of paper case files and datasets² and the same considerations and criteria will continue to be applied to the appraisal of such paper records.

The selection of paper case files continues to be addressed by the appraisal methodology, the sampling Operational Selection Policy *OSP 0: The selection of case files: Sampling techniques*, (sampling of paper records is undertaken because of restrictions on space and the difficulties of managing large quantities of paper) and by the Acquisition Policy.

4 Electronic datasets and case files

² Sampling particular instance papers (RAD Occasional Paper No. 8, 1984), The National Archives Catalogue reference PRO 52/8

This Operational Selection Policy covers two kinds of electronic media: case files and datasets, defined for the purposes of this Policy as follows:-

Datasets are sets of records generally compiled from forms, either transcribed from paper, or increasingly, electronically generated, (often via the central government portals on the internet). The information so collected is then tabulated into a database. Typically, a dataset might contain the results of a survey, review, census or inventory.

Case files on individuals contain textual information. Although the information is much less structured than on a dataset, electronic case files are capable of being searched, in order to retrieve information on a particular individual, or on other criteria such as location.

The primary historical value of these record series resides in the aggregated use of the individual entries or returns: more or less sophisticated statistical methods may reveal social, economic or environmental trends. The information in just one return may have value to a specific individual or organisation. However, selection must primarily be conducted on the basis of the statistical use that may be made of the data.

Sampling. With electronic media, the need for sampling before transfer to The National Archives is removed. The information selected can be transferred to The National Archives in its entirety, ensuring the most comprehensive coverage of state and the citizen data. Researchers are then free to take their own samples, using their own criteria.

5 The Selection Process

5.1 Macro Appraisal

Selection will be done by following the precepts of Macro-appraisal (as outlined in The National Archives' Appraisal Policy): assessing the value of records at a government, departmental or unit level rather than at an individual document or file level:

2.4.3.5 Macro-appraisal will be used as an aid to:

- identify datasets and case files of potential archival value
- identify areas of overlap between departments, so avoiding duplication in material selected for the archive

5.2 Selection Criteria

In order to apply this Operational Selection Policy, government departments, (having, where appropriate, conducted the census of case-files and datasets as outlined in section 2 of the Appraisal Policy Background paper) will evaluate their datasets and case files against the following selection criteria:

- 1) Is the subject of the record an individual? Or, occasionally, a social collection of individuals such as a family or a household, all of whom may or may not be separately identified in the record?
- 2) If so, does the data *together with the circumstances of its compilation* provide information on the individual in addition to that in the core datasets?
- 3) Does it record key functions of the department?
- 4) If not, does it concern under-represented or disadvantaged groups?
- 5) Does it have 'archival value': does it provide information on the individual likely to be of research value either to academics or to a broad section of the public?

Looking at these criteria in more detail:

1) **Is the subject of the record an individual, or a social collection of individuals?**

The definition of an individual within the scope of this Operational Selection Policy would exclude employees of the state (civil servants and military personnel, as these are being employed by, and not interacting with, the state) but include those who had been conscripted to work for the state. It could include: those who had not been awarded formal citizenship under the Nationality, Immigration and Asylum Act 2002; taxpayers; benefit claimants; students or prisoners.

If the dataset does not fulfill these citizen criteria it is not subject to this Policy.

2) **If so, does the data *together with the circumstances of its compilation* provide information on the individual in addition to that on the core datasets?**

Core datasets are those that provide the maximum amount of elemental information on citizens. It is important that the selection process does not merely duplicate information which is already being preserved at The National Archives. However, although the data may appear to be a duplication (a list of names and

addresses say), when the reason for the collection of those names and addresses is taken into account, (for example, a set of basic data about individuals or families resident in caravans would proportionately be weighted to cover travelers), it can be demonstrated that additional information, reflecting the life experiences and diversity of the UK population, is being supplied on those individuals. It is not sufficient to assess the contents of the dataset in isolation but to consider the circumstances of its compilation.

The core datasets, which may change over time, are discussed at Annex A.

If the dataset does not pass this test it will not be selected.

3) **Is it recording key functions of the department?**

In order to prevent the unnecessary duplication of information from different government departments, only those records that relate to a key function of the department should be selected. For example, the Immigration Appellate Authority's IRIS Database concerning case management of immigration appeals, or the Department of Social Security's Earnings Top-up Scheme.

However, this would exclude trivial or ephemeral transactions, for example, a database listing individuals who had enquired about receiving a service on-line would not be selected; whereas one documenting on-line completion of forms to apply for a benefit, or to make a tax return, would be.

More detailed advice on the selection of datasets covered by key functions will be addressed in future Operational Selection Policies.

If a key function has been identified, proceed to question 5.

4) **Social Inclusion: does the record concern minority or disadvantaged groups?**

The National Archives' Social Inclusion Policy

Social Inclusion is a short-hand term for tackling the problems of people who have become isolated from opportunities because of multiple disadvantage – poverty, disability, discrimination, unemployment, issues of health, housing, education and child care. It is about offering appropriate services.

Work is ongoing at The National Archives to address the problems encountered by those who were socially excluded from accessing the archive material already deposited.

What is also needed are measures to ensure that records of particular value for the study of minority history, whether that be ethnic, religious, or other minority groups at risk of social exclusion (such as people with disabilities, the elderly,

lesbians and gay men, unemployed people, people with mental health disorders), will be preserved. In order to be certain that such records are not overlooked, those that fail to record key functions at question 3 should be considered for selection if they concern under-represented or disadvantaged groups.

5) **Does it have ‘archival value’: does it provide information on the individual likely to be of research value either to academics or to a broad section of the public?**

All datasets considered for selection must pass this final test.

The government policy that occasioned the collection of the data should be referred to as an aid in making this decision. However it should be born in mind that such information can be of use to researchers for purposes quite other than the policies which initiated them. For example, taxation records could be used to study the economic condition of the state; criminal datasets could be used to study law and order and sentencing policy; unemployment benefit records could be used to study industrial history; all of the aforementioned might, one day, prove of interest to genealogists.

6) **If the dataset is not to be selected, is it illustrative of a department’s way of working?**

If this is case, although the entire dataset would not be selected, one snap shot could be taken.

6 Summary guide to the Selection Process

Questions		Possible answers	
1	Is subject of the record an individual, or a social collection of individuals?	If yes – next question	If no - not subject to this Policy , proceed to question 6
2	Does the data <i>together with the circumstances of its compilation</i> provide information on the individual in addition to that on the core datasets?	If yes – next question	If no - proceed to question 6
3	Is it recording key functions of the department?	If no – next question	If yes - proceed to question 5
4	Does it concern under-represented or disadvantaged groups?	If yes – next question	If no - proceed to question 6
5	Does it have ‘archival value’: does it provide information on the individual likely to be of research value either to	If yes – Select	If no - proceed to question 6

	academics or to a broad section of the public?		
6	Is it illustrative of a department's way of working?	If yes – Take a single snapshot	If no - not subject to this Policy

7 Repeating and On-going datasets

In order to track changes in society over time datasets that are repeated (for example, a set of surveys conducted over a number of years), will be selected for transfer at agreed intervals, to be decided in conjunction with The National Archives' client manager. Those that are continually added to will need to be archived through regular snapshots, at a frequency to be decided by the department and The National Archives.

8 Dynamic datasets

Datasets, as well as acquiring information on new individuals can also be added to in ways that modify the information on existing individuals (dynamic datasets). There is a risk that some information may be not be accessible. Work is in progress to address this problem and explore methods of preserving all the data captured throughout the dataset's lifespan.

9 Places of Deposit

Local and Regional records, falling under 2.2.1 and 2.2.2 of the Disposition Policy, may be selected for local/regional deposit, (under s. 4 (1) of the Public Records Act 1958), but held at The National Archives until the chosen repository would be able to offer suitable conditions for the long term storage of e-records. The same selection criteria would be applied to local or regional records as to national ones, except that the additional category under procedure 4, (social inclusion), would be considered; does it record individuals in a particular region or locality (a minority in the context of the nation as a whole), and contain information not otherwise accessible or available?

10 Access

These records are likely to contain personal details of citizens who are still alive; such details would not be immediately available to researchers. Access conditions will not influence the selection of any records. The appropriate period of closure would be determined for each record collection following consideration of appropriate exemptions under the Freedom of Information Act 2000, and the relevant principles under the Data Protection Act 1998. As a general rule, personal information will be exempt from disclosure under the Freedom of Information Act for 100 years from the subject's date of birth, assuming, if this is

unknown, an age of 16 for an adult and 7 for a child. In some cases where disclosure would result in an actionable breach of confidence, or would cause extreme mental damage to the subject's family, the information will remain exempt after the subject's death.

11 Links

The National Archives:

Acquisition Policy

<http://www.nationalarchives.gov.uk/recordsmanagement/selection/acquisition.htm>

Disposition Policy

<http://www.nationalarchives.gov.uk/recordsmanagement/disposition/policy.htm>

Appraisal Policy

<http://www.nationalarchives.gov.uk/recordsmanagement/selection/appraisal.htm>

OSP 0: The Selection of Case Files: Sampling Techniques

<http://www.nationalarchives.gov.uk/recordsmanagement/selection/pdf/osp0.pdf>

Digital Preservation

<http://www.nationalarchives.gov.uk/preservation/digitalarchive/>

(The National Archives' long-term preservation strategies for electronic data)

National Digital Archive of Datasets (NDAD)

<http://www.ndad.nationalarchives.gov.uk/>

External Digital Data repositories:

United Kingdom Data Archive (UKDA),

<http://www.data-archive.ac.uk/University of Essex>

Annex A

Government and People: The interaction of the state with the citizen as documented in core records

The Core Datasets

The decennial census of population contains the most comprehensive details of the UK population and is the present core dataset from the point of view of this policy. The first question when appraising datasets should be: does this information merely duplicate that which is on the census? (N.B. the census data itself is not at present held in the format of a digital dataset, although it is likely that future census returns will be digitized on release to the public, (as was the 1901 Census)). Nonetheless, the information on the census returns, irrespective of format remains the starting point. If the records under consideration do contain further information, then it may be appropriate to preserve them in addition to the core datasets.

Example forms detailing the questions asked on the 1971 – 2001 censuses are provided below to assist reviewers:

1971 Census example form

<http://www.celsius.lshtm.ac.uk/modules/forms/census1971.pdf>

1981 Census example form

<http://www.celsius.lshtm.ac.uk/modules/forms/census1981.html>

1991 Census example form

<http://www.celsius.lshtm.ac.uk/modules/forms/census1991.pdf>

2001 Census example form

<http://www.celsius.lshtm.ac.uk/modules/forms/census2001.pdf>

Proposed comprehensive datasets

At the time of writing, there are two proposed large scale projects whereby the state will capture comprehensive data on individuals: the National Identity Register which would contain up-to-date personal details, and the National Health Service National Program for IT.

At such time as these datasets should become operational they will be selected for permanent preservation, and would, along with the decennial census of population, form the core citizen datasets, against which the information on all other datasets and case files should be evaluated.

Records of registrations of births, marriages and deaths in the UK are core datasets but are not public records. Their preservation is safeguarded by other relevant legislation. For The National Archives' policy on the selection of public records of registrations of births, marriages and deaths overseas, see Operational Selection Policy *OSP18 The Registration of Births, Deaths and Marriages of United Kingdom Citizens Overseas*.