

Constructing a Preservation Policy: the case of the UK Data Archive

Matthew Woollard

Head of Digital Preservation
and Systems, UKDA

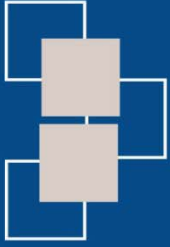
Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC



Requirements influencing policy

40th UKDA
1967-2007

- organisational and business need
- user requirements
- standards
- legislation
- technology and technical environment
- security

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

UKDA Mission Statement

“to support high quality research, teaching and learning in the social sciences and humanities, by acquiring, developing and managing data and related digital resources, and by promoting and disseminating these resources as widely and effectively as possible.”

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC



About data

- Finding
- Catalogue search
 - About catalogue
 - Help on searching
 - Browse by subject
 - Major studies
 - New releases
 - HASSET thesaurus
 - Other archives

Accessing

Creating and depositing

Preservation

Teaching and learning

SN 1 -Family, Work and Leisure in the London Region, 1970 : Main Study

[New Data Catalogue search](#)

[Download/Order](#)

[External View](#)

* Internal User *

[Documentation](#) [Keywords](#) [Publications](#) [Related Studies](#)

Title:

Family, Work and Leisure in the London Region, 1970 : Main Study

Subject Categories:

Community and urban studies - Society and culture
Family life and marriage - Social stratification and groupings
Leisure, tourism and sport - Society and culture

Depositor(s):

Willmott, P., Institute of Community Studies

Principal Investigator(s):

Willmott, P., Institute of Community Studies
Young, M., Institute of Community Studies

Sponsor(s):

Nuffield Foundation
Leverhulme Trust
Centre for Environmental Studies
Post Office
Sports Council

Abstract:

The purpose of this survey was to examine the relationship between work, home and leisure in order to explain changes in family patterns; to provide demographic, social/behavioural and economic data for the analysis of family life.

Main Topics:

Attitudinal/Behavioural Questions
Present and preferred occupation: type, hours, holidays, career structure, autonomy at work, whether physically or mentally tiring or boring. Journey to work.
Present and preferred residence: size, district, tenure, type, garden.

<u>Column</u>	<u>Title</u>
# 2/21	Pressed

Q.35
80

# 2/22	Physically tiring
-------------------	-------------------

Q.36
81

Codes

Not pressed	1 468
Sometimes pressed, sometimes not	2 352
Pressed	3 414
I/I	9 661
Not applicable (i.e. not working full or part time)	1928

Sometimes, sometimes not	0 148
Not tiring	1 528
Fairly tiring	2 380
Very tiring	3 213
I/I	9 659
Not applicable (i.e. not working full or part time)	1928

[illegible]

Pressed

		Frequency	Percent	Valid Percent
Valid	Not Pressed	468	36.7	36.7
	Sometimes pressed, sometimes not	382	29.9	29.9
	Pressed	417	32.7	32.7
	Incomplete information or N/A	9	.7	.7
	Total	1276	100.0	100.0

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Workplace

	Frequency	Percent	Valid Percent
Valid 0	2	.1	.1
Central area	211	10.9	10.9
Elsewhere Inner London (ILEA)	154	8.0	8.0
Outer GLC	366	19.0	19.0
Outer Metropolitan Area	385	20.0	20.0
Outside London Region	7	.4	.4
Works at home	27	1.4	1.4
No fixed place, travel around, "all over"	121	6.3	6.3
Insufficient Information / N/A (i.e., not working)	655	34.0	34.0
Total	1928	100.0	100.0

1/61

Workplace

Q. 18

56

Central Area (see special map) 1211
Elsewhere Inner London (ILEA) 2154
Outer GLC 336
Outer Metropolitan Area 4385
Outside London Region 5
I/I, Not applicable (i.e. not working) } 654
working) working at home, travels) 9
Works at home 627
No fixed place, travel around, "all over" 7121
1928

Note: Change from precodes 0 and 9

Supported by:

University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Standards used to inform

- Rights (Legal Protection of Databases, Re-use of PSI)
- BS ISO 14721 OAIS
- BS ISO 18492 Long term preservation of electronic information
- Information security standards (BS ISO 27001 and 27002)
- Records management standards (BS ISO 15489)
- Media/hardware storage and maintenance

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Standards used to inform

- Archival standards (BS 5454)
- Metadata standards
- File formats
- Also:
 - Other policies (TNA, English Heritage...)
 - Other guidelines (TNA, AHDS)
 - Planning tools

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Planning tools

- TRAC (Trustworthy Repositories Audit and Certification) Criteria and Checklist
- the NESSTOR catalogue of criteria
- the Digital Preservation Coalition's Handbook
- the DRAMBORA toolkit
- 'Core requirements for Digital Archives' (Center for Research Libraries)
- Data Seal of Approval published by DANS.

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Life-cycle approach

- Legal and regulatory framework
- Pre-ingest / ingest
- Archival storage
- Systems (hardware) / IT architecture
- Data Management / metadata
- Access

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Key changes in new policy

- Increased effort on authenticity, integrity and reliability;
- adherence to archival/records management standards;
- new concepts surrounding versioning and editions;
- definitions for archival packages;
- verification of AIPs;
- improved data management (i.e., non-resource discovery metadata);
- preservation of resource-discovery metadata;
- planning open implementation of preservation metadata.

Supported by:



University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

JISC

Σ S.S.R.C. SURVEY ARCHIVE

① Acquisition of data...

from students' p.h.d. theses, research organisations, market research organisations, academics, etc.



② Data arrives



in card or tape form with explanatory documents



③ Documents entered and checked

eg. codebook complete? questionnaire included? licence form signed?



④ Data entered and checked

eg. wild codes? pregnant males? too many cases? multi-punched?



⑤ Data converted to standard storage form

eg. data put on tape with back-up tapes created for security



⑥ Neat and tidy codebook compiled



⑦ Description written for inventory

the archive also... produces bulletins three times a year for distribution.

sends information to all archive representatives in every social science dept. in every university and major polytechnic.

arranges summer work-shops.

provides assistance for archive fellows.

arranges conferences.

distributes informative publications.



User query



Data query

eg. how do I set about getting a copy of the NOP political data?

① Data located from file



② Parameter cards prepared for run on data file



③ Produce copy of data and codebook

data are provided in a format compatible with the users' computer installation. the only costs incurred by the user are for production of a copy of codebook and for postage.

Text query

have you got any data on tourism?

① Check inventory

and obtain from other sources if necessary



② Locate relevant surveys

if data of interest application form sent to enquirer



③ Produce copy of codebook