



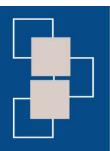
Constructing a Preservation Policy: the case of the UK Data Archive

Matthew Woollard

Head of Digital Preservation and Systems, UKDA





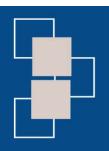




Requirements influencing policy

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







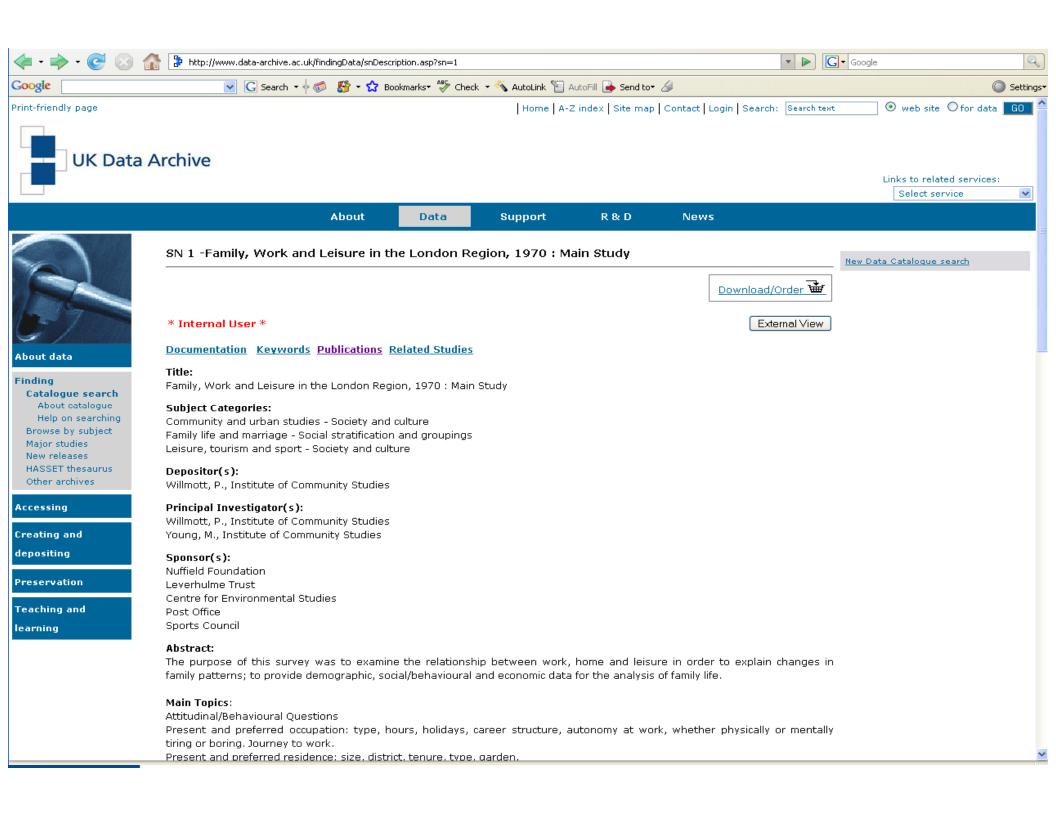
UKDA Mission Statement

"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



Column Title

#= 2/2 | Pressed

Q.35

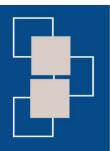
2/52/22 Physically tiring Q.36

Codes

Not pressed	1468
Sometimes pressed, sometimes not_	2 352
Pressed	3.414
1/1	19 661
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	19 659
Not applicable (i.e. not working	1928
full or part time)	**

<u> File</u>	<u>E</u> dit	<u>S</u> earch	<u>P</u> roject	<u>V</u> iew	Forma <u>t</u>	Column	<u>M</u> acro	<u>A</u> dvanced	<u>W</u> indow	<u>H</u> elp
11004	• 151	516445	20412	60930	000099	999999	99990	99999999	929206	51065412653370000370004801¶
11004	214	199913	09911	2123:	112042	001500	00000	0002500:	1500000	00899999999999999200100¶
11004	302	200000	000000	0304:	100001	020000	00000	00000000	0000010	03060406040606060406060405¶
1										03818810009903999990202452¶
11004	•514	417901	.00000	14114	431690					190419999573188208010·····¶
11004	-									·····я
1										41094212252370000370005701¶
										314999099999999999210211¶
1										00161626161606160616061606¶
1										12821221883225105210206421¶
I				_						130610009273058248010·····¶
	_									···············17601····················
						_				510652122424230007201042019
1										008999999999999999220011¶
										00040404040404040504040404¶ 04884810009906999990202461¶
				•						190999999273098208320····¶
										···············17601·················
	_			50730						7,001 1,001 1,001 1,001
										999999999999999999999999999999999999
										00040404060404040404060404
										94884811882004006220004432¶
										110311208172128200320·····¶
										··653811212317601·············
11009	• 151	716345	20321	20530	000031	945328	89990	99999999	9027094	43999999999999999999999999
11009	299	999999	999999	99999	99999	999999	99999	99999999	999999	P0000256666666666666666666
11009	302	200000	000000	04040	000000	000000	00000	00000000	000000	00050506040405040404040505¶
11009	404	040000	01000	00020	000002	020000	00000	0081301:	1099999	91881810009215213090103461¶
11009	•523	617000	99010	14213	343810	146009	91120	41999268	381132	110311253373131108440·····¶
11009	6					233	12999	92999142	2222199	99······¶
11010	141	726155	10822	00810	000009	999999	99990	99999999	909310	5499999999999999999999999
11010	299	999999	999999	99999	99999	999999	99999	99999999	999999	999999999999999200000¶
1										00040406060406060404040606¶
										94881810009005999990000252¶
1										190419999271128908440····¶
										··616123221117601···········
										32123179999370000370099999¶
										99999999999999999999999999999999999999
										99040426040406040405041604¶
										02881821881993999990000372¶
										290129999275128248440····¶ ·····
	· 6/8		(71739)	7371	11/255	/ N# · · ·				u





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:











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

1/51

Supported by:





JISC

Workplace

Central Area (see special map)

Elsewhere Inner London (ILEA)

Outer GLC

Outer Metropolitan Area

Outside London Region

I/I, Not applicable (i.e. not

working) working at home, travels

Works at tone

No Sixed plant, Travel Grand,

Note: Change from precodes O and

all ores 7121

928

121)

2.154

.3**3**LL

4385 5. 🎽

65 ¥

6 27





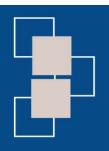
Standards used to inform

- 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











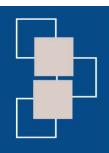
Standards used to inform

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











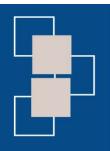
Planning tools

- 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.











Life-cycle approach

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







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







①Acquisition of data...

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



② Data arrives



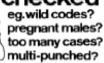
in card or tape form with explanatory documents



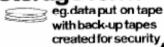


Documents entered and checked

eg.codebook complete? questionnaire included? licence form signed? Data entered and checked



⑤ Data converted to standard storage form



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 workshops.

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?





② Parameter cards prepared for run on Cards file



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

and obtain from other sources if necessary



if data of interest application form sent to enquirer

