



A joint Research Councils Programme
co-sponsored by Defra and SEERAD



Data resources for rural sustainability research: realising their combined potential

<http://reludata.csl.gov.uk>

A RELU Data Management Scoping Study

**Nigel Boatman, Alistair Murray, Helen McKay,
Naomi Jones, James Aegerter & Robert Stones**

Central Science Laboratory



Questionnaire

Response:

- 17% responded (114 of 646 sent)
- Non-respondents contacted, 60% 'not relevant', others mainly 'lack of time' (N=150)

Origin of respondents:

University/College	68%
Government Research Institute	14%
Charity/not-for profit institution	9%
Other	5%
Research Council Institute	4%

Research Interests

Environment	43%
Rural Development	31%
Farming and Food	27%
Economics	11%
Other	32%

Research Programme Funding Body

(26 past and present RCUK programmes were specified)

Joint RCs	11%
NERC	29%
ESRC	17%
BBSRC	13%
EPSRC	8%
UK Non-RC	68%
Overseas	7%

Data Access

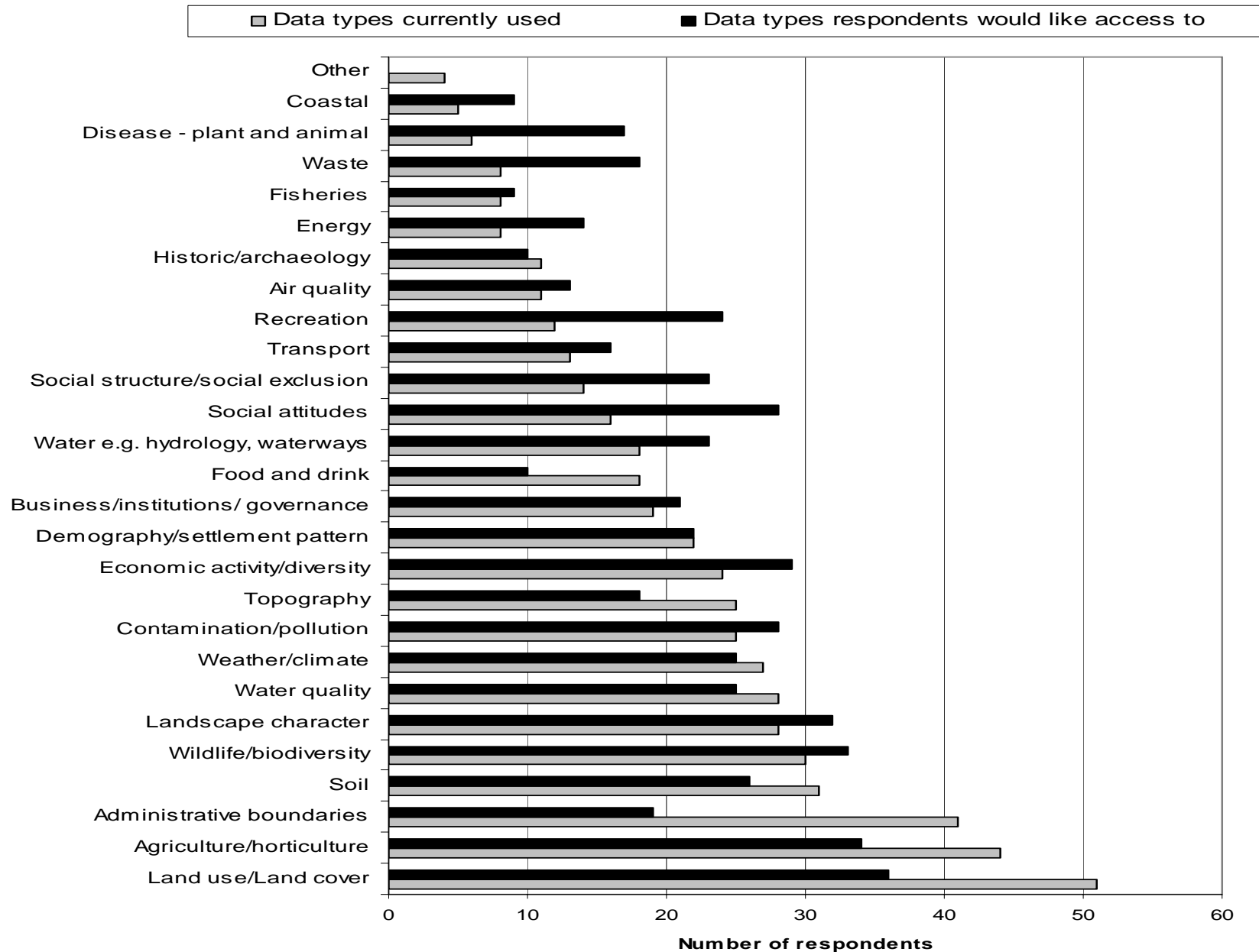
- 75% required access to external datasets

	Discipline			Institution		All
	Social scientists	Environ. scientists	Economists	Government	Education	
N	26	63	22	16	79	114
% Yes	69	78	68	88	70	75
% No	27	17	27	6	25	22
% Do not know	4	5	5	6	5	4

Data access (types of data)

- The most frequent data types accessed were:
 - land use/cover
 - agriculture/horticulture
 - administrative boundaries.
- There was an apparent difference between data types currently used and future needs, e.g.
 - An increased future need to access datasets on: social attitudes, social structure/social exclusion, recreation, energy and waste

Data access (types of data)



Datasets and access arrangements

- Most commonly accessed datasets:
 - Digimap (21%)
 - Countryside Survey (20%)
 - Population Census (15%)
 - Agricultural Census (15%)
- Access arrangements for commonly used datasets:
 - 11 free access
 - 6 licence (Digimap, Soil)
 - 4 access rights
 - 2 self-owned

How respondents discovered the datasets they used in their research.

Method of discovery	N	%
Colleagues	69	65.7
Internet: owning organisation website	55	52.4
Internet: data portal	40	38.1
Library	31	29.5
Catalogue	10	9.5
Other	6	5.7

Difficulties with access

	Discipline			Institution		
	Social scientists	Environ. scientists	Economists	Government	Education	All
N	26	66	22	16	78	114
%Yes	23	41	41	63	31	38
%No	50	44	55	38	56	49
%Do not know or N/A	23	11	5	0	13	13

Nature of difficulties with access

	Current (n = 80)	Anticipated (n = 86)
Cost	44%	43%
Confidentiality	21%	26%
Ownership issues	17%	19%
Expertise/data structure	12%	13%
Other	5%	

- Current difficulties:
 - Agricultural Census
 - confidentiality 4
 - expertise/data structure 3
 - cost 2
 - NSRI Soils and NATMAP
 - cost 8
 - expertise/data structure 1
- Anticipated difficulties:
 - Soil
 - cost 6
 - availability 1
 - Schemes (uptake)
 - confidentiality 5

Data integration

- 25% currently integrate data or use integrated datasets

	Discipline			Institution		
	Social scientists	Environmental scientists	Economists	Government	Education	All
N	26	66	22	16	78	114
%Yes	12	33	14	31	21	25
%No	81	42	64	50	62	57
%Do not know	4	20	23	19	18	18

Types of data integrated (from examples given):

	Current (n = 43)	Future (n = 37)
Environmental/Environmental	64%	72%
Environmental/Socio-Economic	29%	12%
Socio-Economic/Socio-Economic	7%	16%

- Most integration was into landscape units e.g. km², field (67%) rather than a political units e.g. county, output area (29%)

Tools for data management and integration

	Discipline			Institution		
	Social scientists	Environmental scientists	Economists	Government	Education	All
N	26	66	22	16	78	113
% Spreadsheets	73	77	64	69	74	75
% Mapping/GIS	31	53	59	44	50	53
% Statistical software	23	39	36	50	36	50
% Database software	19	32	0	44	21	36
% Graphical	42	61	32	63	47	24
% CIS	12	9	9	19	9	15
% Other or N/A	23	12	14	25	14	10

Further work

- Consultation
 - Social scientists
 - GIS/statisticians
 - Data managers
 - Data providers
 - Computer scientists

Consultation

- **Social scientists**
 - What were the difficulties with the questionnaire?
 - Explore data integration issues
 - Explore QA, archiving and DM issues

Consultation

- **GIS / Statisticians**
 - Methods used for data integration
 - 3 most important issues
- **Some preliminary results:**
 - Methods: common framework e.g OS mastermap, LCM2000
 - Issues: scale, licensing and availability of data, quality, availability of metadata, how to map qualitative data,

Consultation

- **Data managers**
 - What worked and what didn't?
 - Explore data interchange/sharing/communication between researchers
 - Explore archiving
- **Some preliminary results:**
 - Personal approach
 - Timeliness
 - Enforcement
 - Resources

Consultation

- **Data providers**
 - Resources
 - Problems
 - Technological developments and trends
- **Some preliminary results:**
 - Resource demanding
 - Internet
 - Quality issues

Consultation

- **Computer scientists**
 - New technologies
 - Breadth of access
- **Some preliminary results:**
 - Quantity of data
 - Grid technology
 - E-social science
 - Scanners

Further work

- Workshop on data integration

19 May, King's Manor, York

- Report