



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

Data resources for rural sustainability research: realising their combined potential.

Annex C

Review of data sources relevant to the RELU programme.

Helen M^cKay, Naomi Jones and Nicola Dennis,

**Central Science Laboratory, Sand Hutton, York YO41 1LZ, U.K.
h.mckay@csl.gov.uk**

29 June, 2005.



Contents

Executive Summary	1
Introduction.....	2
Portals, gateways and hubs	2
Environmental data	1
Social data.....	2
Geographical Information.....	3
Government (UK and EU).....	3
Data providers.....	4
Data initiatives.....	6
Environment Research Funders Forum	6
Defra Whole Farm Approach.....	6
Sustainable Development Research Network.....	6
SPIRE.....	6
INSPIRE	6
NERC DataGrid programme.....	6
SEEK.....	7
Great Britain Historical GIS project	7

Executive Summary

1. This review aims to provide users of land use and rural economy data with useful lists of data sources: portals, gateways, hubs, data providers and data initiatives. It is important to realize that the same dataset may be available from more than one source.
2. A portal is a network service that brings together information from diverse sources, and presents it to the user. Gateways and hubs are sites on the web that provide access to descriptions of and links to other sites on the web.
3. A number of portals, gateways and hubs are reviewed, ranging from EDINA (a JISC-funded national datacentre offering networked access to a library of data, free of charge for academic use) to INFOROUTE (gateway to information held by all UK Government departments).
4. Data can also be obtained directly from organisations collecting and/or holding them. A number of web sites are listed, with descriptions of the information available.
5. Finally, a brief description is given of a few relevant initiatives relating to data access or integration.

Introduction

This review aims to provide users of land use and rural economy data with lists of data sources that may be useful in (a) identifying whether certain data exist and (b) accessing data. These lists complement the results of a questionnaire survey of the RELU community (see reludata.csl.gov.uk web site), and are not exhaustive. Also, it is important to realize that the same dataset can be available from multiple sources. Portals, gateways, hubs, data providers and data initiatives are reviewed.

Portals, gateways and hubs

Technically, a portal is a network service that brings together content from diverse distributed resources using technologies such as cross searching, harvesting, and alerting, and collate this into an amalgamated form for presentation to the user. This presentation is usually via a web browser, though other means are also possible. For users, a portal is a, possibly personalised, common point of access where searching can be carried out across one or more than one resource and the amalgamated results viewed. Information may also be presented via other means, for example, alerting services and conference listings or links to e-prints and learning materials.

Gateways and hubs are sites on the web that provide access to descriptions of and links to other sites on the web. They carry out a similar role to portals insofar as they bring information together, but this information is only about the sources of content, not the actual content itself. Gateways and hubs allow you to discover where the content is, but they still require you to go to the site of the content to get it. Portals bring the content to you (JISC website: Portals Frequently Asked Questions http://www.jisc.ac.uk/index.cfm?name=ie_portalsfaq).

It is clear, however, that these definitions are not set in stone. In practice, the terms portals, gateways and hubs are all used for web sites that provide information on datasets, and sometimes direct access to the data.

Annex C. Review of data sources.

ENVIRONMENTAL DATA

EDINA	http://edina.ed.ac.uk/	EDINA, based at Edinburgh University Data Library, is a JISC-funded national datacentre. It offers the UK tertiary education and research community networked access to a library of data, information and research resources. All EDINA services are available free of charge to members of UK tertiary education institutions for academic use, although institutional subscription and end-user registration are required for most services.
Global Biodiversity Information Facility	http://www.gbif.org/	Web portal for access to data and information on biodiversity – access to museum specimen databases and also the 'what, where and when' of species distributions
Global Change Master Directory	http://gcmd.gsfc.nasa.gov/index.html	The goal is to enable users to locate and obtain access to Earth science data sets and services relevant to the global change and Earth science research. Web portal for access to data and services.
Global Terrestrial Observing System: Terrestrial Ecosystem Monitoring Sites	http://www.fao.org/gtos/terms/index.jsp	Actively collects, archives and/or distributes environmental data on terrestrial/coastal ecosystems. International web-accessible meta-database on monitoring networks, sites and the variables they measure
National Biodiversity Network	http://www.searchnbn.net/#	The NBN gateway provides access to biodiversity data for the UK. The site allows you to view distribution maps and wildlife data using a variety of interactive tools. It is completely dependent on the data providers. Organisations contributing to the NBN Gateway are listed, and include the Biological Records Centre, CEH, Local Councils and Wildlife Trusts. Use of the Gateway indicates acceptance of its Terms and Conditions. The user is permitted to view and download data for personal use (use for profit requires written permission from the data provider).
NERC catalogue	www.nmp.rl.ac.uk	Access to distributed meta-database on NERC environmental data holdings
The Knowledge Network for Biocomplexity	http://knb.ecoinformatics.org/home.html	Building and evaluating a metadata-based framework for integrating heterogeneous scientific data. Data access from a range of organisations in the US.
The Species 2000 Organisation	www.sp2000.org	Species 2000 is a "Federation" of database organisations working closely with users, taxonomists and sponsoring agencies. The thrust is to create an array of participant global species databases covering each of the major groups of organisms. Species 2000 will operate a dynamic Common Access System on the Internet through which users can locate a species by name across an array of on-line taxonomic databases.

Annex C. Review of data sources.

SOCIAL DATA

Access to Archives	http://www.a2a.org.uk/	A2A allows the user to search and browse for information about archives in England, dating from the 900s to the present day. These archives are cared for in local record offices and libraries, universities, museums and national and specialist institutions across England, where they are made available to the public.
AHDS	http://ahds.ac.uk/	The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities. Presently, they cover five subject areas: archaeology, history, visual arts, literature, languages and linguistics and performing arts.
ESDS	http://www.esds.ac.uk/	The Economic and Social Data Service (ESDS) is a national data service that provides access and support for a range of key economic and social data, both quantitative and qualitative, spanning many disciplines and themes.
MIMAS	http://www.mimas.ac.uk/	Manchester Information & Associated Services is a JISC and ESRC-supported national data centre providing the UK with networked access to key data and information resources across a wide range of disciplines.
Social Science Information Gateway	www.sosig.ac.uk	The Social Science Information Gateway (SOSIG) is a freely available Internet service for students, academics, researchers and practitioners in the social sciences, business and law. It is part of the UK Resource Discovery Network
The Council of European Social Science Data Archives	http://www.nsd.uib.no/cessda/	CESSDA promotes the acquisition, archiving and distribution of electronic data for social science teaching and research in Europe. The web site allows easy access to the catalogues of member organisations.
UK Data Archive	www.data-archive.ac.uk	The UK Data Archive (UKDA) is a centre of expertise in data acquisition, preservation, dissemination and promotion. It is curator of the largest collection of digital data in the social sciences in the UK and houses a major collection of computerised historical material.
UK Resource Discovery Network	http://www.rdn.ac.uk/	The Resource Discovery Network is the UK's free national gateway to Internet resources for the learning, teaching and research community. The service currently links to more than 100,000 resources via a series of subject-based information gateways (or hubs). The RDN is primarily aimed at Internet users in UK further and higher education but is freely available to all.

Annex C. Review of data sources.

GEOGRAPHICAL INFORMATION

Gateway to Geographical Information	http://www.gigateway.org.uk/	Gigateway is a free web service aimed at increasing awareness of and access to geospatial information in the UK. The Data Locator can be used to find out what geographic datasets exist. The Data Directory will help locate organisations which supply geographic data, products and services.
United Nations GEO Data Portal	http://geodata.grid.unep.ch/	The GEO Data Portal is the authoritative source for data sets used by UNEP and its partners in the Global Environment Outlook (GEO) report and other integrated environment assessments. Its online database holds more than 450 different variables, as national, subregional, regional and global statistics or as geospatial data sets (maps), covering themes like Freshwater, Population, Forests, Emissions, Climate, Disasters, Health and GDP. Display them on-the-fly as maps, graphs, data tables or download the data in different formats.

GOVERNMENT (UK AND EU)

Defra Rural Evidence Hub		An initiative which aims to provide an evidence base of high quality information and knowledge of rural England. It is composed of statistical information and published documents in a searchable format and is due to go live in the summer of 2005.
European Environment Agency	http://www.eea.eu.int/main.htm	The EEA Data service provides access to data sets used in EEA periodical reports. The data sets contain aggregated data at national level. They cover at least 15 of the 25 EU Member States, but in many cases a pan-European coverage of 52 countries is available. Graphs and, in the future, maps can be generated from the data sets. Information about the source of each data set and its geographical and temporal coverage is attached.
EUROSTAT	http://epp.eurostat.cec.eu.int/	European Statistical Data Support provide a support network at national and European Level. Data is available free of charge.
Her Majesty's Stationary Office	http://www.inforoute.hmso.gov.uk/	Provides direct access to the Government's Information Asset Register (IAR) which lists information resources held by the UK Government, concentrating on unpublished resources. It enables users to identify the information held in a wide variety of government departments, agencies and other organisations.
Inforoute	http://www.opsi.gov.uk/cgi-bin/searchIAR.pl?DB=iar	Gateway to information held by all UK Government departments.
Magic	HTTP://WWW.MAGIC.GOV.UK/	A web-based interactive map to bring together geographic information on key environmental schemes and designations in one place.
OECD Statistics Portal	http://www.oecd.org/statistics	This provides free access to some OECD databases as well as extracts from other databases, classified by topic.

Annex C. Review of data sources.

Data providers

The following table lists organisations and their web sites, which either provide direct access to data, or information on the existence of data.

British Geological Survey	www.bgs.ac.uk	Geoscience Data Index. Information on data available through licence.
Centre for Ecology and Hydrology	www.ceh.ac.uk	Has a large selection of nationally and internationally important hydrological, terrestrial and freshwater environmental data.
Centre for Environmental Data and Recording	www.habitas.org.uk/cedar	Information on the distribution of flora, fauna and geological sites within Northern Ireland and its coastal waters.
Countryside Agency	http://www.countryside.gov.uk/	Data tables associated with the National and Regional State of the Countryside Reports, as well as Rural Services Series data.
Defra	www.defra.gov.uk/	Defra's Information Asset Register covers information held by Defra, including databases, old sets of files, recent electronic files, collections of statistics, research etc. Direct access to the IAR is provided by the inforoute gateway.
English Nature	www.english-nature.org.uk	GIS Digital Boundary Datasets and National Map Series showing Designated Sites.
Environment Agency	www.environment-agency.gov.uk	Information on the existence of data but no access to datasets, although some summary data is available.
Environmental Change Network	www.ecn.ac.uk	The UK ECN is the UK's long-term, integrated environmental monitoring and research programme. ECN gathers information about the pressures on and responses to environmental change in physical, chemical and biological systems. Dynamic generation of graphs & tables and data download facility
Environmental Data Services	www.ends.co.uk	By subscribing one can connect to a database (updated daily) of current UK legislation, official reports, consultation papers and regulatory guidance.
Farm Scale Evaluation	http://www.farmscale.org.uk/index.htm	The FSE was set up to determine whether genetically modified herbicide-tolerant crops might have any significant effects on wildlife. The website contains a large number of free data sheets collected from beet, maize and oilseed rape.
Forestry Commission	www.forestry.gov.uk	As part of their on-line library they have reports and results from a number of surveys as well as research on topics such as woodland, planting, timber and visitor experience.
Joint Nature Conservation Council	www.jncc.gov.uk	This site contains the results of an analysis of the whole of the British flora, highlighting those that are threatened.
Met Office	www.met-office.gov.uk	The Hadley Centre data sets for climate are available here. These cover month temperatures since 1659 and rainfall since 1914.
National Soil Inventory (NSI) Data	www.silsoe.cranfield.ac.uk	This university site contains the Soil Data Gateway, however data sets cannot be viewed here but there are contact details should one wish to.
National Statistics (Office of National Statistics)	www.statistics.gov.uk	Contains time series data from National Statistics economic and socio-economic releases. These can be downloaded with the relevant software free of charge.
Royal Botanic Gardens, Kew	www.rbgekew.org.uk	Online databases

Annex C. Review of data sources.

Scottish Environmental Protection Agency	www.sepa.org.uk	SEPA holds a large range of data arising from its monitoring of the environment as well as operator returned data.
The Biological Records Centre	www.brc.ac.uk	BRC is the national custodian of data on the distribution of wildlife in the British Isles.
The Countryside Agency	www.countryside.gov.uk	Has a number of publications containing data sets on the state of the countryside in different regions, these can be viewed as a pdf or ordered, often free of charge, as a hard copy.
The Economic and Social Research Council	www.esrc.ac.uk	There is a wide range of research publications on a variety of issues available to download. It also contains a link to the Social Science Information Gateway and a fact and figures section containing a selection of different types of data.
The English Heritage Geophysical Survey Database	http://sdb2.eng-h.gov.uk	Index of geophysical surveys of archaeological sites
The Great Britain Historical Database	http://www.port.ac.uk/research/gbhgis/aboutthegbhistoricalgis/database/	GBHDB is a large integrated database of geographically-located historical statistics for Great Britain
The Natural History Museum	www.nhm.ac.uk	On-line databases.
The Royal Society for the Protection of Birds	www.rspb.org.uk	Contains a Data Zone with downloadable data on important bird areas and RSPB reserves.
UK Biodiversity Action Plan	www.ukbap.org.uk	This site contains an on-line library of downloadable data.
UK National Digital Archive of Datasets	www.ndad.ulcc.ac.uk	The National Digital Archive of Datasets (NDAD) is part of The National Archives. NDAD preserves and provides online access to archived digital data and documents from UK central government departments.
UKEDI	http://ukedi.ceh.ac.uk/	The UK Environmental Data Index currently contains metadata on 1245 datasets. It aims to address the problems of knowing what data exist, where they are located and how to gain access by providing users with a simple means for identifying sources of environmental data on air, water and land quality.

Annex C. Review of data sources.

Data initiatives.

This section provides a brief description of some initiatives relating to data access or integration.

ENVIRONMENT RESEARCH FUNDERS FORUM	http://www.erff.org.uk/tors.asp	The Environment Research Funders' Forum is a new focus group bringing together the UK's major public sector sponsors of environmental science. The Forum will allow funders to identify and take strategic action on any gaps in environmental research and training. It will also shape future science direction through horizon scanning and other activities and add value to UK environmental interests internationally. To do this members will share their strategies, priorities, activities and other information. The structure for ERFF working groups diagram on the web site mentions a database project board, which suggests that data will be managed.
DEFRA WHOLE FARM APPROACH	http://www.defra.gov.uk/farm/sustainable/wfa/	The WFA is an IT framework which will provide a single access method for farmers to Defra and other participating national and local government departments. Ultimately this would give access to Government information currently held in different databases: animal records, Integrated Administration and Control (IACS) information, habitat maps, water catchment area data, land designations, archaeological sites, etc.
SUSTAINABLE DEVELOPMENT RESEARCH NETWORK	http://www.sustainabledevelopmentresearch.org.uk/sdrguide/	The Sustainable Development Research Network (sponsored by Defra) produce a guide to Research Centres and Evidence Providers. The report maps out - provides an overview and guide to - the main UK academic departments, institutes and research groups undertaking cross-cutting sustainable development research and their relevant areas of expertise. Details of key sustainable development research organisations from outside of the academic sector (such as voluntary sector organisations and consultancies) are also included where appropriate. Data would presumably be obtained directly from the organisations listed.
SPIRE	No URL.	SPatial Information REpository is a Defra initiative which will create a geographical information repository. The initial phase of the project involved consultation with users, in order to plan the inclusion of specific layers in SPIRE. The current phase is addressing data standards, and a future phase will identify gaps in data.
INSPIRE	http://www.ec-gis.org/inspire/	The INfrastructure for SPatial InfoRmation in Europe initiative aims at making available relevant, harmonised and quality geographic information for the purpose of formulation, implementation, monitoring and evaluation of Community policy-making. INSPIRE principles state that: <i>It should be possible to combine seamless spatial information from different sources across Europe and share it between many users and application.</i> The project is currently in progress, with the Preparatory Phase extending from 2005 to 2006.
NERC DATAGRID PROGRAMME	http://ndg.badc.rl.ac.uk/	It is intended that this programme will build a grid which will make data discovery, delivery and use much easier than it is now, facilitating better use of the existing investment in the curation and maintenance of quality data archives. Further it is intended to make the connection between data held in managed archives and data held by individual research groups seamless in such a way that the same tools can be used to compare and

Annex C. Review of data sources.

		manipulate data from both sources. The user will be able to compare and contrast data from a range of (US, European, UK, NERC) datasets from within one specific context. The project will initially support specific datasets within the meteorological and oceanographic community, however it will be fully functional for defined data types. The technology will be extensible and will be designed to allow appropriate data held across all the NERC disciplines to be available via the NERC DataGrid interface.
SEEK	http://seek.ecoinformatics.org/	The Science Environment for Ecological Knowledge (SEEK) is a system designed to facilitate not only data acquisition and archiving, but integrating, transforming, analyzing, and synthesizing ecological and biodiversity data that was previously intractable. The goals of SEEK are to build cyberinfrastructure to address the many challenges associated with data accessibility and integration of large-scale biocomplexity data in the ecological sciences. SEEK will enable ecologists to efficiently capture, organize and search for data and analytical processes from their desk-top in a user friendly interface -- ultimately providing resources typically out of reach for many ecologists.
GREAT BRITAIN HISTORICAL GIS PROJECT	http://www.port.ac.uk/research/gbhgis/	The Great Britain Historical Geographical Information System is a digital collection of information about Britain's localities as they have changed over time. Information comes from census reports, historical gazetteers, travellers' tales and historic maps. The system uses ArcInfo, Oracle database and open source MapServer software.
EU JRC	http://www.jrc.cec.eu.int/	The EU Joint Research Centre provides customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.