

The Lowland Catchment Research Programme

Data Policy

Roger V. Moore
C. Isabella Tindall

Centre for Ecology and Hydrology
Institute of Hydrology
Wallingford
Oxon
UK

Tel: +44 (0)1491 838800
Fax: +44 (0)1491 692424
E-mail: rym@ceh.ac.uk
E-mail: cit@ceh.ac.uk

The Catchment Hydrology and Sustainable Management Research Programme Data policy

Policy Statement

APPROVAL	Date	Name	Signature
-----------------	-------------	-------------	------------------

Author :

Checked :

DOCUMENT CONTROL

This document is under document control. Paper copies of this document are authorised if, and only if, they are signed in non-black ink.

Distribution List

LOCAR Steering Committee
LOCAR Data Managers Committee
LOCAR Data Centre
NERC Headquarters

Amendment History

Issue	Date	Author	Description
V1.R1.M1	28/06/01	Roger V. Moore C. Isabella Tindall	1 st Draft
V1.R1.M2	16/07/01	Roger V. Moore C. Isabella Tindall	2 nd Draft after comment by P.Heads, A.Gustard & A.McKenzie.

CONTENTS

1	PURPOSE AND READERSHIP	7
2	TERMS AND DEFINITIONS	7
3	BACKGROUND	8
4	GENERAL ISSUES RELATING TO DATA	9
4.1	The Nature of the Data Resource	9
4.2	Ownership and Custody of Data	10
4.3	Obligations of Those Holding Data	10
4.4	Obligations of LOCAR towards Scientists Collecting Data	12
4.5	Other Legal and Contractual Obligations	12
5	MANAGEMENT RESPONSIBILITIES FOR DATA WITHIN LOCAR	12
5.1	The LOCAR Data Centre	12
5.2	Long-term Maintenance and The Role of the Science & Technology Board	13
5.3	Detailed Responsibilities of the LOCAR Data Centre	13
6	PLANNING FOR THE MANAGEMENT OF DATA	14
6.1	Decisions to be made before Data are Collected	14
6.2	Minimum Standards of Stewardship for NERC ‘Corporate Data’	15
6.3	Technological Implications	16
7	ACCESS TO, AND CHARGES FOR, LOCAR’S DATA	16
7.1	The Public Interface to LOCAR’s Data Holdings	16
7.2	Granting Access to LOCAR’s Data	17
7.3	LOCAR’s Data Pricing Policy	17
7.4	Data Exchange, etc	18
8	THE IMPLICATIONS FOR SCIENTISTS HOLDING DATA	19
8.1	Responsibilities of <i>all</i> Data Holders	19

1 PURPOSE AND READERSHIP

The purpose of this document is to provide guidance on the implementation of the Natural Environment Research Council's data policies for the Lowland Catchment Research programme (LOCAR).

It has been prepared at the request of the LOCAR Steering Committee and NERC's headquarters for:

- The LOCAR Steering Committee
- The LOCAR Data Committee
- The LOCAR Data Centre
- Heads of other organisations which are providing data services
- Environmental scientists in academia, especially those working on LOCAR

The document is intended to provide the model on which the data policy statements for all LOCAR projects will be based.

The LOCAR Data Management Plan will describe how this policy will be implemented in operational terms.

2 TERMS AND DEFINITIONS

BGS refers to NERC's British Geological Survey.

CEH refers to the NERC's Centre for Ecology and Hydrology.

EA refers to the Environment Agency.

JIF is the Joint Infrastructure Fund established by the UK government and the Wellcome Foundation to fund research infrastructure.

LOCAR refers, for the purposes of this document, to the joint programme of lowland catchment research funded for five years by NERC and co-funded for the first two years by the NICHE/LOCAR JIF award.

NERC refers to the Natural Environment Research Council.

NICHE refers to the National Infrastructure for Catchment Hydrology Experiments.

NICHE/CHASM is the sister project of NICHE/LOCAR, where CHASM refers to the Catchment Hydrology and Sustainable Management project.

NICHE/LOCAR is a JIF award to fund the instrumentation of three lowland research catchments.

3 BACKGROUND

The LOCAR programme aims to improve the science required to support current and future management needs for permeable lowland catchments through an integrated and multidisciplinary experimental and modelling programme.

The scientific aims are:

- To develop an improved understanding of hydrological, hydrogeological, geomorphological and ecological interactions within permeable catchment systems and their associated aquatic habitats at different spatial and temporal scales and for different land uses;
- To develop improved modelling tools to inform and support the integrated management of lowland catchment systems.

These aims will be achieved through the study of:

- The surface and near-surface environment – runoff, recharge and material transport;
- Groundwater processes in lowland catchments;
- Physical, chemical and biological processes within the valley floor;
- In- stream, riparian, hyporheic and wetland habitats and their dependence on flow regimes;
- The impacts of society on the functioning of lowland permeable catchments.

The broader programme objectives are:

- To provide an underpinning science base to meet the requirements of the UK and CEC for environmental protection in lowland permeable catchments within a framework of sustainable and integrated river basin management;
- To establish flagship sites as a basis for long-term monitoring, to provide the necessary data to define natural variability and response to environmental change;
- To build strong inter-disciplinary links in order to harness the science skill base more effectively to tackle problems of national importance;
- To promote research collaboration between universities, research institutes and user groups such as regulators and water utilities;
- To equip the UK research community in order to maximise opportunities for international collaboration and to contribute to international initiatives;
- To provide training in the scientific expertise required for integrated catchment management;
- To disseminate and exploit the research outputs effectively through close collaboration with users, workshops, seminars and the WWW, as well as through publications and conferences;
- To provide suitable data stewardship.

As the first step in achieving these objectives, the LOCAR programme will instrument a number of catchments and assemble the resulting data in a single Data Centre. From here they

and all other LOCAR data will be made available to the science programme that will follow. This document sets out the high level vision and rules that will guide this data gathering exercise. It is based on the NERC Data Policy Handbook, Version 2.1 (March 1999).

Scientists often consider the end point of their research to be the production of **publications**, whether in open literature or in the form of reports to customers. While these form one output from scientific activity, science in general but particularly environmental science such as LOCAR involves the **collection of data**, and the subsequent **management of these data**. While data will indeed be manipulated by the researcher to provide material for publication, **data are a resource in their own right**. Properly managed and preserved, they can potentially be used and re-used by future researchers, and exploited commercially or educationally. Such additional uses, often never envisaged in the first instance, will make a further contribution to the advancement of science and for the benefit of society.

Environmental data are often irreplaceable; they are always unique, if only in the timing of collection. They can also be extremely expensive to collect. For these reasons LOCAR attaches great importance to ensuring that maximum benefits are derived from data once acquired. LOCAR has therefore adopted and adapted NERC's formal policies relating to data. The LOCAR policy is set out in Annex 1. Their full implementation must inevitably be an evolutionary process since some aspects of policy can only be achieved when concomitant funding is available. As a first step it is essential that the policies and their implications are understood at a management level throughout the LOCAR community. Managers will then need to ensure that appropriate guidance is passed down to all who need it within their areas of the programme.

4 GENERAL ISSUES RELATING TO DATA

4.1 THE NATURE OF THE DATA RESOURCE

'Scientific data' may be held in either analogue or digital form and be stored either on paper or a variety of computer-compatible media. There is a spectrum running from 'raw data' through 'processed data' to 'information' and ultimately 'knowledge'. The material held by libraries, and the physical specimens in curated collections, might be regarded as 'data', although they are outside the usual sense of the word. It would be idle to attempt a rigid definition of the term 'data'.

These policy guidelines have generally been designed with **digital, computer-based data sets** in mind, although the **principles** can often be extended to cover a wider field. 'Data' so defined still exhibit considerable diversity. For example, there are: major computer-based databases which are continuously managed and extended by the validation and addition of incoming data from multiple sources; the output files from computer-based predictive models, which can be regarded as 'environmental data' even though they do not constitute measurements of the environment; digital data sets deposited as archives after research projects have ended; and data sets held by individual scientists on which they are actively working as part of their research. Some of these last will be ephemeral and never justify long-term preservation. To give an extreme example, the transitory contents of a scientist's calculator might be regarded as 'scientific data'! **But many data sets are potentially valuable**, whether scientifically, educationally or commercially, even when (or indeed because) they were collected a long time ago.

4.2 OWNERSHIP AND CUSTODY OF DATA

If the term ‘data’ is difficult to define precisely, then so too is the concept of ‘LOCAR’s data’. Data sets will be collected or generated by NERC scientists, Higher Education Institutions (HEIs) and, may be, public and commercial partners. Data sets will also be placed into the custody of LOCAR as a result of voluntary deposits, negotiated exchanges or purchase, with concomitant obligations on LOCAR.

Despite behaviour that might suggest the contrary, **data sets frequently do not belong to those who have collected them**. They generally belong to the employers of such data collectors (eg NERC or the HEIs) or to those who have paid for the data collection (eg bodies commissioning research or the Ordnance Survey). Ownership can become more complex when data collection has been jointly funded. Ownership of data sets implies the right to **exploit** them and, if continued maintenance becomes uneconomic, the right to destroy them. If legal ownership is in question, or is excessively fragmented, it may be impossible for **anybody** to exploit data. The important point is that all parties who may have a claim to ownership or part ownership of a data set should agree at the outset as to how it is to be exploited and how the benefits are to be shared.

Examples of LOCAR data sets and the ownership situations that might arise are:

Data type	Funding	Ownership
Site exploration data – pre-bid, e.g. test boreholes.	NERC and JIF Consortium	NERC and JIF Consortium
Baseline survey data, e.g. land use	NERC and JIF Consortium	NERC and JIF Consortium
Catchment data collected from LOCAR and EA instruments by the Catchment Service Teams and EA personnel	Mixed funding where it will probably be difficult to identify accurately who contributed what.	NERC, JIF Consortium and EA
Annual surveys of e.g. ecology	Mostly funded by NERC	NERC
Data provided by e.g. the EA	EA	EA

4.3 OBLIGATIONS OF THOSE HOLDING DATA

It follows that scientists in the LOCAR community will often be holding data owned in whole or in part by others. There are consequent **obligations on the holders (and their management) to look after the data responsibly**, so that these interests are not compromised, and indeed to be **aware of the issues that are important in this context**.

For example, researchers will often exchange data sets in a spirit of scientific co-operation and it is part of the normal process of science that they should do so. However, if a data set has been distributed to third parties without any restrictions on its subsequent re-use or redistribution, then any later commercial exploitation of that data set is likely to have been undermined, so losing income that might have been spent on further science. At times, therefore, data transfers need to be controlled by formal licence agreements. To attempt a rigid definition of the circumstances when this should apply would be as futile as defining the term ‘data’ in the first place. The environmental science community needs to develop a culture where scientists consider such issues, use their own judgement and **seek guidance when appropriate**.

Data sets can easily be destroyed unintentionally, or effectively lost, by failure to take adequate and continuing precautions to safeguard them; they may be held on vulnerable computer media without adequate back-up, such media may become obsolete over time, and the data formats may be undocumented. Unique paper records are vulnerable to fire or physical deterioration. Researchers and their management must be aware of these dangers.

Scientists will frequently process the data they have collected selectively, or with specific application packages, in order to prepare material for publication in the scientific literature. But the full value of the data collected may only be realised if the entire data set is subjected to generic processing (eg to ensure calibration and adequate quality control) and is sufficiently documented to allow others to re-use it at a later date. The original collector may be the only person in a position to undertake such work and so to unlock the full potential of the data. Those holding data collected under LOCAR funding will be expected to co-operate in validating and publishing them in their entirety - **when this can be justified in terms of their scientific value** - rather than merely creaming off a subset for immediate publication in the literature. As a logical consequence, NERC will accord due recognition and support to the 'publication' of data in this way.

It is self-evident that data sets can only be fully exploited if potential users are aware of their existence. Scientists will be expected to co-operate with the LOCAR Data Centre in providing information about them.

LOCAR fund holders will be required to:

- Prepare and agree a data management plan with the LOCAR Data Centre
- Prepare and agree quality assurance procedures with the LOCAR Data Centre
- Depending on the agreement with the Data Centre either:
 - deposit with the LOCAR Data Centre a copy of the data resulting from the supported research as soon as practicable together with documentation/metadata describing these data.
 - or
 - deposit with the LOCAR Data Centre appropriate metadata

The Data Centre will then be in a position to make the data available to others (under suitable constraints) for further bona fide research only. The Intellectual Property Rights to the data need not be transferred.

4.4 OBLIGATIONS OF LOCAR TOWARDS SCIENTISTS COLLECTING DATA

Individual scientists, principal investigator teams and programmes will be permitted a reasonable period of prior right¹ to publish results from data sets which they have collected. Moreover, **NERC will give due recognition to staff for their production of publishable, application-independent data sets alongside that of papers in the scientific literature**, and Grants Committees recognise that the activity is a legitimate call on NERC funding. The long-term professional maintenance of such data sets, and their dissemination as a service to those who need them, is an activity often quite distinct from the data acquisition and subsequent research of the originator. This long-term data stewardship will be undertaken by the LOCAR Data Centre, described below, which, **subject to funding**, will be established on a permanent basis independent of the finite life of the LOCAR project.

Where LOCAR-funded academics deposit data with LOCAR, the Data Centre will disseminate these data to other bona fide researchers as a service to the environmental science community.

4.5 OTHER LEGAL AND CONTRACTUAL OBLIGATIONS

LOCAR, NERC and individual members of the LOCAR scientific community, are bound by legal and contractual obligations which relate to data. In particular

- Some of LOCAR's data sets may constitute 'environmental information' within the meaning of the Environmental Information Regulations 1992, which gives effect to the EC Directive on the Freedom of Access to Environmental Information.
- 'Personal data' held in computers are subject to the Data Protection Act 1984. While most environmental data sets will not contain information about individuals that would be covered by the Act, any that do must be registered with the Data Protection Registrar.
- Data may be collected under the terms of a contract or Memorandum of Understanding with another body, which may specify the rights and obligations of the contracting parties with respect to data sets and their dissemination.
- Data sets needed by a researcher may not belong to NERC, but have been obtained under licence from another supplier; the terms and conditions of the licence must be observed.

5 MANAGEMENT RESPONSIBILITIES FOR DATA WITHIN LOCAR

5.1 THE LOCAR DATA CENTRE

It will be clear from the foregoing that LOCAR's interests in ensuring that data are managed as a coherent resource may be wider than the immediate concerns of the researcher who originates or holds individual data sets. NERC believes that issues relating to scientific data are best handled, and related services best provided, by those with a background in the underlying discipline. **The LOCAR Steering Committee will therefore delegate responsibility for its data and implementation of its data policies to the Centre for Ecology and Hydrology and the British Geological Survey which will establish a LOCAR Data Centre**, as part of the National Water Archive and which will be responsible for all LOCAR data.

¹ Typically, but not necessarily, two years. Note, however, that a prior right to publish does not mean exclusive access to the data.

It is important to distinguish the Data Centre's **responsibility** for data from actual **data custody** itself. In some cases data will be physically transferred to the Data Centre, for example, the results of the field programme, while in others, the Data Centre will keep records of where data are held.

The responsibilities of the Data Centre are listed at 5.3. Contact details are given in Annex 2.

5.2 LONG-TERM MAINTENANCE AND THE ROLE OF THE SCIENCE & TECHNOLOGY BOARD

Long-term data stewardship is an expensive activity and the long-term costs arising from the LOCAR project will be born by NERC. The justification for funding it must lie in the contribution that it will make to science, the creation of wealth, or to improvements in the quality of life. In any area of science, the scale of data stewardship, whether undertaken by NERC's Designated Data Centres or elsewhere within the NERC community, should reflect the anticipated overall value (scientific or otherwise) of the data over the long term. This value will often be much greater than any immediate financial return obtainable from commercially licensing access to the data in the short term. Often, data will be of value to users beyond the immediate NERC environmental science research community. In managing data a balance must be struck between the scale of the commitment needed (across the entire lifetime of the activity) and the anticipated benefits, both scientific and financial. Judgements on the balance will be affected by the characteristics of the related scientific discipline, and are therefore the province of NERC's Science & Technology (S&T) Board and Expert Groups. The data stewardship activities in each science area should be responsive to the overall science strategy. The S&T Board will set clear objectives to the Data Centres as to the nature of the return they are expecting from their investment in data stewardship and monitor Data Centre activities and costs. It will ensure that science activities within their disciplines are paying due attention to data issues and in recommending the allocation of resources they will satisfy themselves that there is appropriate funding in each of the following areas.

- Core funding via NERC Centres/Surveys to meet the data requirements of their science strategy and commercial work. This area is the responsibility of the Centre/Survey Boards.
- Specific funds for each Thematic Programme so as to ensure the proper management and stewardship of data generated by that programme.
- Funding to provide a data service to HEIs, other than as above.
- Funding to allow Data Centres to adopt lead roles at national or international level where appropriate.
- Funding to grasp opportunities relating to data stewardship - eg to acquire and preserve data at risk.

In some cases, direct receipts from commercial licensing of data will contribute toward the above funding lines, so subsidising data management and scientific research.

5.3 DETAILED RESPONSIBILITIES OF THE LOCAR DATA CENTRE

The LOCAR Data Centre is responsible for the following:

1. specifying the procedures, formats and media in which data will be received from the field and disseminated to users.
2. receiving, storing and making available LOCAR data and meta data.
3. ensuring the adequate physical custody, validation, dissemination and review/purging of data for the LOCAR project (In many cases the Data Centre will undertake this stewardship themselves; when it does not, it will at least have contact with those who do).
4. maintaining standards of data stewardship for LOCAR.
5. pro-actively seeking out data within LOCAR which would merit more active stewardship; encouraging the deposit of data sets; and advising LOCAR-funded academics on whether and where their data should be deposited on completion of their projects.
6. promoting the case for investment where necessary to facilitate the above.
7. promoting the use of LOCAR data by devising and promulgating catalogues, directories, leaflets and brochures.
8. formally arranging licences to control the release of data sets to non-LOCAR/NERC recipients, the uses to which the data sets may be put and their further dissemination; and to protect NERC from legal liability. Pricing data in accordance with NERC guidelines in order to derive revenue and/or other scientific benefits from LOCAR's data resource.
9. advising on the licensing/purchase of non-NERC data required by researchers. In some cases they will undertake this on behalf of the scientist, to ensure the most favourable terms and conditions, and best value for money.
10. handling all requests for LOCAR's data with the specific invocation of the Environmental Information Regulations.
11. maintaining up-to-date information on data holdings in their discipline on the World Wide Web.
12. acting as a gateway to other NERC custodians of data.
13. reporting to the LOCAR Steering Committee via the LOCAR Data Committee (a sub-committee of the Steering Committee).
- 14.

6 PLANNING FOR THE MANAGEMENT OF DATA

6.1 DECISIONS TO BE MADE BEFORE DATA ARE COLLECTED

It is a requirement of the NERC, and therefore the LOCAR, data policy that any proposal

to undertake science, which will involve the acquisition of data sets, should include at the outset consideration of what is to be done with them once acquired. Valuable scientific or commercial opportunities may be lost if this fundamental principle is neglected.

Before **any** project at any scale is started, the following issues should be addressed and resolved with the LOCAR Data Centre.

- What existing data sets will be needed by the project and what, if anything, will it cost to acquire them?
- What data sets will be produced by the project and who will be responsible for their initial management?
- What are the requirements for inter-operability between such data sets and how are they to be met in practice? What data (and associated software) standards and quality assurance arrangements should be set in place?
- Is the implication that participants will require a project-specific data *service*; if so how will it be provided and what will it cost?
- How can the data to be assembled be best exploited, whether scientifically or commercially?
- Who will take the lead in this exploitation and how will any commercial benefits be shared?
- Should specific data products be published or otherwise promulgated as a direct output from the project?
- Which, if any, of the data sets should be accorded long-term stewardship and become part of NERC's environmental data resource?
- Which NERC Data Centre will undertake this long-term stewardship?
- Should any specific data services set up during the project be perpetuated after it has finished? If so, will continuing funding be required for them, and can a source be identified?
- What are the overall resource implications of the above plans?
- What are the technological implications of the above plans?

All LOCAR projects are required to prepare a written data management plan addressing the above issues and establish quality assurance procedures. Both should be agreed with the Data Centre at the start of the project.

6.2 MINIMUM STANDARDS OF STEWARDSHIP FOR NERC 'CORPORATE DATA'

If a data set is to form part of LOCAR's and hence NERC's enduring data resource, then the following minimum standards are required.

- The ownership and Intellectual Property Rights to the data set must be established and LOCAR/NERC's policy towards exploiting and making it available to third parties agreed.
- The data set must be catalogued to the level of detail required by the Data Centre, so that it can be mentioned, if appropriate, in LOCAR and NERC catalogues on the World Wide Web.
- Formal responsibility for the custody of the data set must be agreed.
- The data must be fully 'worked up' (ie calibrated, quality controlled, etc) with sufficient

qualifying documentation to be of use to third parties without reference to the original collector.

- The technical details of how the data are to be stored, managed and accessed must be agreed and suitably documented.
- The technological implications must be established.
- The resources needed to carry out these intentions over the planned life of the data, in terms of staff (whether in project teams or the Data Centre) and Information Technology (IT) equipment/ infrastructure, must be estimated and sources identified.
- A review mechanism must exist to reconsider periodically the cost benefits of continuing to maintain the data. The intention to destroy or put at risk data should be publicised in advance, allowing time for a response by interested parties.

6.3 TECHNOLOGICAL IMPLICATIONS

Data stewardship implies the need for an underlying infrastructure of IT equipment and support. This will include adequate communication links to the sites holding data, computing hardware and software (eg database- and file- servers with adequate backup and secure off-line media storage, and geographical information systems) and operational/technical support for such facilities.

The rapid evolution of IT presents both opportunities and threats to those responsible for data: opportunities in that they may be able to adopt more efficient and cost effective techniques for data management (subject to appropriate investment); and threats in that changes in technology render former media obsolete, forcing the need for investment in alternatives and in the potentially crippling costs of transcription. Moreover, IT developments affect instrument technology, making it feasible to collect ever larger quantities of data, and offsetting the improvements in storage capacity.

The membership of NERC's Data Strategy Group includes IT expertise so that the Data Centres such as LOCAR's will be aware of these opportunities and threats. In response, the Data Centres can ensure that project data plans are realistic and incorporate adequate IT provision; and prepare cases for IT investment to evolve their data services in the light of their customers' needs.

7 ACCESS TO, AND CHARGES FOR, LOCAR'S DATA

7.1 THE PUBLIC INTERFACE TO LOCAR'S DATA HOLDINGS

The LOCAR Steering Committee intends that enquiries concerning LOCAR's data holdings, or requests for access to them, should normally be addressed in the first instance to the Data Centre. In addition to its detailed knowledge relating to LOCAR's data, the LOCAR Data Centre will have some knowledge of the holdings of others Data Centres within and without NERC, and will refer enquiries onward to them if necessary.

7.2 GRANTING ACCESS TO LOCAR'S DATA

It is important that all holders of data in the LOCAR community who receive requests from third parties for access to the data sets they hold should first establish, before they supply them, that

- they have a right to transfer the data requested (eg the copyright/IPR is not owned by some other party who has not authorised disclosure).
- the transfer is not compromising LOCAR's or NERC's interests (eg the use of the data, and NERC's liability, are suitably constrained by licence agreements, and that the pricing, or zero-pricing, is appropriate).

The advice of the Data Centre should be sought if necessary.

Many of NERC's and possibly some of LOCAR's data sets may constitute 'environmental information' within the meaning of the Environmental Information Regulations 1992. These provide a **right of access** in a timely manner to such information, though it is important to note that 'freedom of access' need not imply 'free access'. There are circumstances under which such requests either **may** or **must** be refused (eg if the data contain personal information). **Holders of LOCAR or NERC data who receive requests for access to them which specifically invoke the Environmental Information Regulations should, in all cases, seek the advice of the LOCAR or appropriate NERC Designated Data Centre.**

Furthermore, **if such holders of LOCAR or NERC data receive any requests (whether invoking the regulations or not) which they intend to refuse, they should seek guidance from the Data Centre as to the lawfulness of their intention and discuss their grounds for doing so.**

7.3 LOCAR'S DATA PRICING POLICY

NERC and therefore LOCAR will not normally sell data or the Intellectual Property Rights to its data. Where a charge is levied, NERC will generally sell the **right to use data** and provide access to the data so that the right can be exercised.

In making data sets available (when it has the right to do so) NERC Council seeks to realise their value in the advancement of environmental science and/or the creation of wealth. It is therefore Council policy to charge for the provision of data at a rate that is dependent on the use to which the data will be put and to specify formally any consequent restrictions on the use of the data in formal licensing agreements. This policy enables Council to reflect the value of the data and the cost it has incurred in data acquisition, custody and retrieval, when satisfying requests for wholly commercial applications. By contrast, it is NERC policy to enable access to data by *bona fide* scientific researchers at no more than the direct costs involved. Depending on circumstances, this will be achieved by providing data either: wholly free of charge (eg when provision of data to a specific community has already been funded centrally); for a nominal handling charge; or at an appropriately discounted rate.

'Bona fide researchers' are those conducting academic research solely to advance the state of knowledge and not for commercial gain; where appropriate, NERC will peer review the

application. Note that access to data may be permitted even when NERC is not itself funding the bona fide research and in this case represents 'support in kind'.

Such provision of data for *bona fide* research must inevitably be based on mutual trust. Council must be satisfied that data supplied free or at reduced cost for research will not be re-used for commercial purposes in breach of the licensing arrangements, and may revoke academic privileges in cases of abuse.

The results of *bona fide* research should be published in the public domain and unless agreed otherwise, processed data sets derived from NERC's data supplied under such terms should be offered to NERC. Even where commercial charges are made for NERC's data, it may be appropriate to agree within the licence that processed data will be returned to NERC under suitable terms.

Where academics supported by NERC grants have deposited resulting data with NERC, the Intellectual Property Rights (IPR) to those data need not be transferred to NERC. NERC will make the data available to other *bona fide* researchers either free of charge or for no more than a handling charge. If requests for such data are received from *commercial* users, NERC will refer them to the owners of the data unless other arrangements have been negotiated.

Since disseminating information about data sets can only increase their utilisation for science, education or wealth creation, it is NERC policy that access to its basic catalogues and indexes should be encouraged and be 'free'; ie incur no more than any direct costs involved. Sophisticated software based catalogues supplied in machine readable form may be marketed as data in their own right, with inexpensive availability to bona fide researchers as above.

The LOCAR Data Centre will be in a position to provide further guidance on licensing.

7.4 DATA EXCHANGE, ETC

The LOCAR Steering Committee intends its policy to be sufficiently flexible to encourage collaboration with other bodies and to enable the value of its data assets to be realised on occasions by arrangements such as data exchange, rather than by charging a price for the licence to use them.

Partners in data acquisition will not be charged for unprocessed data but where LOCAR has analysed, interpreted or generally added value to the data, a charge for the resultant value-added information may be made.

LOCAR's data may be used as a potentially tradeable asset for reciprocal exchange agreements without prejudice to current contractual arrangements. Data will be exchanged for research purposes where it is clear that the research will lead to a contribution to knowledge within LOCAR's remit, or to benefits in kind to LOCAR. Such negotiations will normally be undertaken by the LOCAR Data Centre.

8 THE IMPLICATIONS FOR SCIENTISTS HOLDING DATA

8.1 RESPONSIBILITIES OF ALL DATA HOLDERS

All scientists in the LOCAR community who are involved with the collection and/or maintenance of data should adhere to the following guidelines.

- They should be acquainted with the principles contained in this document, follow them when planning the collection of any new data, and where practicable attempt retrospectively to manage to the same standards data already collected
- In particular, they should be aware of the role of the LOCAR Data Centre
- They should liaise with the Centre if
 - i they require data which will need to be licensed/purchased at significant expense. Through its knowledge of other potential requirements for the same data the Data Centre may be able to arrange a more favourable deal than an individual scientist. In some cases (e.g. satellite imagery) it is standard practice for the Data Centres to effect the transaction on behalf of the scientist; the cost of such data licences are not allowable on research grants.
 - iii they are providing a service disseminating data to users outside their immediate team. This will enable them to ensure that their activities, eg in relation to making data available and charging for access, are consistent with LOCAR and NERC policies.
 - iii they have data sets they believe to be of 'publishable quality' that would justify publicising and disseminating more widely to users, whether scientific or commercial. In particular, recipients of NERC grants or LOCAR funding should liaise with the appropriate Data Centre, and offer to deposit with it a copy of data resulting from the research supported.
 - iv they are holding data sets which would require the expenditure of resources if they were to be rendered publishable, but which might justify such investment in terms of the subsequent scientific or commercial return.
 - v they are approached with requests for data invoking the Environmental Information Regulations.

ANNEX 1: LOCAR DATA POLICY STATEMENT

A Data Acquisition

With regard to the acquisition of data, LOCAR:

- i regards data sets as a valuable resource in their own right;
- ii will ensure that the maximum benefits are derived from data acquired by LOCAR;
- iii requires that due consideration be given to the 'post-project' stewardship of data prior to approval being given for a 'project';
- iv will establish, at the outset, how the data acquired as a result of such projects will be exploited (whether commercially or scientifically); who will be responsible for this exploitation; and how the benefits will be shared;
- v requires that recipients of LOCAR related funding offer to deposit with the LOCAR Data Centre a copy of data sets resulting from the research supported. This offer will not prejudice the intellectual property rights;
- vi requires that processed data sets derived from LOCAR/NERC's data, which were provided on 'academic' terms, be offered to the Data Centre.

B Data Management

In managing its data holdings, LOCAR will:

- i have defined points of contact associated with its Data Centre, with which agreements must be reached as part of the planning for any new activities which generate new data sets, so that the full implications (including any commitments on the Designated Data Centre or requirements for IT infrastructure) can be established at the outset;
- ii use these defined points of contact to publicise its data, providing a means whereby the environmental science community can find out what information it can obtain, and on what terms;
- iii facilitate access by customers to LOCAR data holdings;
- iv publicise its intention to put at risk or destroy data, before doing so, when it is considered that the cost of keeping data sets outweighs the apparent benefits;

C Data use

With regard to the use of data, LOCAR will:

- i require that LOCAR data are made available to all parties involved in the LOCAR programme and, after two years, to the whole academic research community for non-commercial research;
- ii require that the users of data sets inform the originators of their use and the purpose to which they will be put;
- iii require that the users of data acknowledge the originators, for example, by co-authorship or citation;
- iv ensure that individual scientists, principal investigator teams and participants in programmes will be permitted a reasonable period during which they have a prior right to

- v publish² the results derived from the data collected by such individuals and teams³;
exchange data for research purposes where it is clear that the research will lead to a contribution to knowledge within LOCAR's remit, or to benefits in kind;
- vi regard data as a potentially tradeable asset for reciprocal exchange agreements, without prejudice to current contractual arrangements; and
- vii specify formally any restrictions on the use of data in formal licensing agreements.

D Charging for Access to LOCAR's Data

Recognising the value of data as a resource, but mindful of the manner of its acquisition, LOCAR will:

- i not charge partners in data acquisition for unprocessed data;
- ii ensure that access to the results by any other persons or for any commercial or other activities not referred to above shall only be given with the consent of the party to which the results belong and upon such terms and conditions and in return for such payments as the party concerned shall require;
- iii consider whether or not a charge should be made where LOCAR has analysed, interpreted, or generally added value to the data;
- iv charge* for the provision of data at a rate that is dependent on the use to which the data will be put;
- v specify formally any consequent restrictions on the use of data in formal licensing agreements;
- vi enable access to data by bona fide scientific researchers at no more than the direct costs involved (nominal handling charge) or at an appropriately discounted rate;
- vii encourage access to basic catalogues and indexes of LOCAR data, for which only direct costs will be charged; but
- viii market software-based catalogues, supplied in machine readable form, in their own right.

* or accept 'payment in kind' e.g. exchange

² Typically, but not necessarily, two years. Note that a prior right to publish does not imply exclusive access to the data.

³ Detailed arrangements will be set out in the Data Management Plan.

ANNEX 2: NERC DESIGNATED DATA CENTRES

Further Information about NERC's Designated Data Centres can be obtained from the NERC World Wide Web home page URL <http://www.nerc.ac.uk/>

NERC Council has delegated responsibility for data, and the implementation of its data policies, to seven Designated Data Centres as follows:

National Water Archive

[responsible for the LOCAR Data Centre, NERC's hydrological data, and the Government's National River Flow Archive]

Centre for Ecology and Hydrology
Wallingford
Oxon OX10 8BB

tel 01491-838800
fax 01491-692424
e-mail nwamail@ceh.ac.uk

Antarctic Environmental Data Centre

[responsible for all NERC's data from the Antarctic, irrespective of discipline]

British Antarctic Survey
High Cross
Madingley Road
Cambridge CB3 0ET

tel 01223-221400
fax 01223-362616
e-mail aedc@bas.ac.uk

British Atmospheric Data Centre

[responsible for atmospheric sciences data]

Rutherford Appleton Laboratory
Chilton
Didcot
Oxon OX11 0QX

tel 01235-446432
fax 01235-445848
e-mail badc@rl.ac.uk

British Oceanographic Data Centre

[responsible for marine data]

Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead
Merseyside L43 7RA

tel 0151-653-8633
fax 0151-652-3950
e-mail bodcmail@pol.ac.uk

National Geosciences Information Service

[responsible for geosciences data, but NERC's science-based archaeology community should liaise directly with the Archaeology Data Service, Dept of Archaeology, University of York, Kings Manor, York YO1 2EP tel 01904-433954 fax 01904-433939 e-mail info@ads.ahds.ac.uk]

British Geological Survey
Keyworth
Nottingham NG12 5GG

tel 0115-9363109
fax 0115-9363276
e-mail ngis@bgs.ac.uk

Environmental Information Centre

[responsible for all other NERC terrestrial and freshwater data]

Institute of Terrestrial Ecology
Monks Wood
Abbots Ripton
Huntingdon
Cambs. PE17 2LS

tel 01487-773381
fax 01487-773467
e-mail eic@ite.ac.uk

NERC Earth Observation Data Centre

[responsible for NERC's holdings of Earth observation data]

Rutherford Appleton Laboratory
Chilton
Didcot
Oxon OX11 0QX

tel 01235-446168
fax 01235-445848
e-mail neodc@rl.ac.uk

