1 BACKGROUND, OBJECTIVES AND METHODOLOGY

ABSTRACT

Definition of biological recording;* Biological survey: need and network*; origins and objectives of CCBR; survey of relevant Government policy; objectives of the present review, methods used in the review; support received and responsibilities for the review.

1.1 BIOLOGICAL RECORDING: A DEFINITION

1.1.1 Biological recording, as agreed for the purposes of this Report, is defined as:

> **The collection, collation, storage, dissemination and interpretation of information, both in space and time, concerning kinds and numbers of wildlife, assemblages of organisms, and their biotopes, especially when the records are related to localized sites. It excludes comparable information concerning agricultural, horticultural and forestry crops, and stock, except in the context of general land use.**

Biological recording can operate at local, national and international levels. At the local level in the UK recording of geological information is undertaken often alongside biological recording, and the term *environmental recording* is widely used although without any clear definition of what is included.

1.2 HISTORICAL BACKGROUND

1.2.1 Observations on wildlife and landscape as an integral activity of travel, even for recreational purposes, account for the earliest biological observations such as those of Giraldus Cambrensis, William Turner or John Ray. Recording continued to be almost entirely the province of the interested amateur up to the late 18th century when a small but increasing number of professional botanists and zoologists began to make a variety of records for more explicitly scientific purposes and such professional recording has continued to increase. Early maps, especially Ordnance Survey and estate maps, often provide the earliest information about the location and area of major biotopes such as woodland and wetland. Nevertheless, much of the available information on both the taxonomy and on the distribution of the flora and fauna of Britain and Ireland has been, and continues to be, provided by amateurs working largely for their own interests.

1.2.2 Biological recording, undertaken formally but on a local scale, began in some areas in the late 19th century when many natural history and field clubs were established. There was a notable increase in biological recording which coincided with a spectacular increase in membership of the many general and specialist natural history societies in the years immediately after 1945. The reasons for this increase are not fully understood but in this period major mapping, and later, monitoring programmes were initiated. Coincidentally there was an upsurge of professional studies, in part through the rise of experimental taxonomy and, in part through the founding of a statutory Nature Conservancy in 1947 with its need for extensive distributional and monitoring data. The existence of the latter organisation also provided for the first time a limited source of funding for more complex or coordinated recording, monitoring and, to some extent, publication of such data.

1.2.3 A brief review of the state of biological recording in the UK in the mid 1980s - *Biological survey: need and network* - was published by a working party of the Linnean Society (Berry 1988). It noted that biological recording had become more effective, widespread and informative than previously, but recognized that the best use was not being made of existing information and even that some data were in danger of being lost. The report concluded that:
"Although considerable effort is expended on biological survey and surveillance in the United Kingdom by voluntary, professional and statutory bodies, no effective system exists for the overall co-ordination of recording and monitoring of wildlife and habitat resources"
It recommended the creation of a Coordinating Commission to plan and establish a national, computerized system and to investigate related issues such as the statutory and legal framework.

1.2.4 A meeting of interested organizations and institutions was arranged at the Royal Society by the Natural Environment Research Council (NERC) in February 1989. This meeting accepted, in large measure, the findings of the Linnean Society's working party and recommended their implementation through the establishment of a Coordinating Commission for Biological Recording (CCBR). CCBR is broadly representative of the recording community (Appendix 1), in particular, the use and management of biological records.

1.2.5 An independent chairman was appointed in February 1990 and the Commission published a Statement of Intent. After defining what was meant by *biological recording*, the Statement of Intent set out detailed objectives to meet the Commission's remit including a phased programme, of which the present investigation is the first part.

1.2.6 The first requirement to be identified was the need to obtain more detailed knowledge concerning the present position, status and legal aspects of biological recording than that described in the Linnean Society's report, and to define the probable future needs of a range of key users. The investigation was funded through contracts from the Department of the Environment (DOE) and the Joint Nature Conservation Committee (JNCC), with underpinning support from NERC. In addition, some initial financial help had come from the New Phytologist Trust. The CCBR representatives of principal funding bodies (DOE, JNCC and NERC) together with the independent Chairman and a representative of The Wildlife Trusts (formerly the Royal Society for Nature Conservation) formed a Board of Management for the project summarised in this report (see Appendix 1).

1.3 POLICY BACKGROUND

1.3.1 Although no *explicit* national policy concerning biological recording existed when the investigation began, the need to record and monitor wildlife is *strongly implicit* in much national and international legislation and international conventions adopted and ratified by the UK.

1.3.2 The earliest examples of such legislation are probably the late 19th century and early 20th century Acts to protect birds and seals, and the Local Government Act of 1888 which was used by local authorities to introduce by-laws to protect plants (Sheail 1976). Inevitably, this legislation was based on information about the occurrence of species and perceived threats to their survival. The Nature Conservancy was founded by Royal Charter in 1949 following a flurry of government activity in the immediately post-war years summarised in four Command Papers (6628, 7122, and 7235). Its role *"to provide scientific advice on the conservation and control of the natural flora and fauna of Great Britain; to establish, maintain and manage Nature Reserves in Great Britain, including the maintenance of physical features of scientific interest; and to organize and develop the research and scientific services related thereto"* led to pioneering activities such as the project to map the flora of the British Isles (Perring & Walters 1962) and, in 1964, the establishment of a national Biological Records Centre (Harding & Sheail 1992).

1.3.3 Subsequent legislation and conventions have built on this implicit need for data and the *de facto* supply of data though a variety of routes (see Chapter 2). In the White paper, *This Common Inheritance* (Cm.1200) and in *Action for the Countryside* (1992) the Government highlighted its broad policy to protect and enhance the beauty and diversity of the countryside and conserve its wildlife. The publication in October 1994 of the Planning Policy Guidance notes on nature conservation (PPG 9) (DOE 1994b) set out

the Government's policies on different aspects of planning (in England only) and refers explicitly to the need for *"adequate information about local species, habitats, geology and landforms"* (PPG 9, para 24).

1.3.4 Explicit recognition of the need for, and maintenance of, biological recording and monitoring has come during the preparation of this report as a consequence of the Government's becoming a party to the Rio Convention in 1992 and the ensuing publication of *Biodiversity, the UK Action Plan* (Cm.2428) early in 1994. The Biodiversity Action Plan Steering Group has set up a Sub-group on Data, which met for the first time in October 1994. The precise remit of this Sub-group has not been defined formally, but may cover the preparation of a catalogue of data sources, technical standards, the integration of data and legal issues.

1.3.5 The recent implementation of the EC Directive on the Freedom of Access to Information on the Environment (90/313/EEC) through the promulgation of the Environmental Information Regulation (SI 1992 No.3240) affects the public availability of biological records.

1.3.6 This report presents the results of the investigation undertaken by CCBR since 1992 and makes recommendations for future action. Some of the findings of this investigation have been incorporated already into *Biodiversity the UK Action Plan* and this report's final recommendations take cognisance of relevant government policy in the public domain up to 31st October 1994.

_____BOX_____

Examples of current national and international legislation, conventions and resolutions which imply needs for biological recording

National legislation
***Wildlife and Countryside Act 1981* - nature conservation policy and practice**
***Agriculture Act 1986* - monitoring of Environmentally Sensitive Areas (ESAs)**
***Environmental Protection Act 1990* - nature conservation policy and practice; formation of new statutory agencies**
***Town and Country Planning Act 1990* - survey of planning areas**
***Conservation (Natural Habitats, etc) Regulations 1994* - transposes Directive 92/43/EEC into national law**

International legislation
***European Community Council Directive 79/409/EEC* on the conservation of wild birds**
***European Community Council Directive 85/337/EEC* on the assessment of the effects of certain public and private projects on the environment**
***European Community Council Directive 92/43/EEC* on the conservation of natural and semi natural habitats of wild fauna and flora**

International Conventions (with date of UK ratification)
***Ramsar* (1976) - conservation of wetlands of international importance, especially as waterfowl habitat (Cm 6464)**
***Berne* (1983) - conservation of European wildlife natural habitats**
***Rio* (1994) - conservation and sustainable use of biological diversity**

Council of Europe Resolutions and Recommendations
***Resolution of the Committee of Ministers (R(76)17)* on the European network of biogenetic reserves**
***Recommendations of the Ministerial Committee (R(87)13)* on the creation of computerised data banks of biological information**

1.4 OBJECTIVES OF THE PROJECT

1.4.1 The objectives of this investigation were to make a detailed assessment of:
*** the kinds of biological records made and maintained;**

*** the resources deployed in such work;**
*** the purposes for which records were made and kept and the use made of them; and in the light of the findings;**
*** to propose recommendations for the establishment and operation of an integrated, computerised national system of biological records and recording.**

1.4.2 The topics to be investigated and assessed included:
*** the present situation concerning biological recording organizations, their holdings and activities;**
*** the principal current applications of biological recording;**
*** the legal aspects of holding such data and of making it available;**
*** future needs and the necessary actions to meet them, including technical specifications of appropriate hardware and software; the establishment of operational standards and appropriate operating policies.**

1.5 METHODS USED

1.5.1 A detailed questionnaire (Appendix 2) sought information about the present situation in the UK under the following subheads:
Details of Organisations: **contact; type; status; geographical coverage; scope and use of data; data exchange arrangements; operating policies; services provided and use made of data; resources[1];**
Data Holdings: **species data; habitat (biotope)-based and land type data; non-biological data;**
Computing details: **recording and storage media; computing experience; computing hardware used; database software and applications used for management of records; use of mapping systems, GIS and other special-purpose software.**

1.5.2 The list of recipients was drawn up after wide consultation with members of CCBR and the sponsoring agencies. They included national and local government departments, country conservation agencies, national parks, local records centres, wildlife trusts, natural history societies, scientific societies and various smaller groups known to be involved with biological records. The questionnaire was sent out to 600 organisations and backed up by means of telephone calls, visits and further discussions, or written submissions: CCBR received 355 responses (Appendix 3) of which about 200 can be regarded as complete.

1.5.3 The information obtained from the questionnaire was stored in a specially devised database, using *Advanced Revelation* and *Mapbase* software, which was used both for recording and, in part, analysing the information. The analysis also made use of *QuattroPro*, and *Graphics Works* was used to prepare tables and figures.

1.5.4 Literature relating to the topics of enquiry or subsequent recommendations was consulted, assessed and a full bibliography prepared (Appendix 5). Relevant national, EU and international legislation was examined.

1.5.5 Legal advice relevant to the owning and holding of individual biological records, and collections of them, was sought from solicitors and other academic sources (Appendix 4). Of particular interest were matters of intellectual property rights, especially relating to data compilations and computerised databases in the UK and EU. Public access to biological records was also considered.

1.6 CONTRACTS & RESPONSIBILITY

---

[1] **This information not to be disclosed except to CCBR and sponsors.**

1.6.1 Six tenders were received for the investigation in August/September 1991, but eventually CCBR agreed subcontracts with C J T Copp (Environmental Information Management) and the Institute of Terrestrial Ecology (ITE, NERC) (1st July 1992 - 30th June 1993, extended to allow for additional work, to 31st March 1994), together with Messrs Morrell, Peel and Gamlen of Oxford in connection with legal advice. Information from, and the views of, members of CCBR and the sponsors were received

throughout. The report was drafted initially by Sir John Burnett (Chairman CCBR) Charles Copp and Paul Harding (ITE), but has been seen by all members of CCBR and the Board of Management. It was edited finally by the authors in the light of the comments received and represents a report by CCBR to the sponsors of the project. A database was devised, developed and documented by C J T Copp to hold information obtained from the questionnaire. Copies of the database have been deposited with the principle sponsors (the Department of the Environment and the Joint Nature Conservation Committee) and at the Biological Records Centre, ITE, Monks Wood.

2 THE CURRENT STATE OF BIOLOGICAL RECORDING IN THE UK

ABSTRACT

Recent history of biological recording in the UK.
Organisations involved: size of the representative sample; kinds of organisations, location, staffing, funding, standards and policies, charging policies.
Data holdings: sources, numbers,temporal range, kinds - taxa, site, biotope, marine.
Survey, surveillance and monitoring: types and coverage.
Methods used: recording media employed, input methods, standards and validation - taxa, land cover, biotopes and vegetation types, spatial and geographical referencing - standards and validation.
Computerisation of biological records: extent used, data management, hardware and software in use, mapping programs, GIS, communications.
Data exchange and transfer: existing informal network, data transfer formats.
Non-biological data used.
Uses and users of data.

2.1 INTRODUCTION

2.1.1 The present state of all aspects of biological recording up to mid 1993 in the UK is reviewed in this chapter. The main source of the quantitative information cited is the questionnaire survey (see Appendix 2) which was conducted between September 1992 and August 1993. The 355 completed questionnaires included in the database, from which the analyses were prepared, are a self-selected sample from more than 2000 organisations potentially concerned with biological recording in the UK. The detailed data from the questionnaire survey are included in the CCBR database.

2.1.2 Biological recording in the 1990s is characterised by an array of surveys, methods and organisations which are dispersed and, in most cases, lack any form of coordinated overview. This situation is the inevitable consequence of responses by organisations and individuals to continuing changes in, for example, organisational policies, structures and funding, requirements for information, information technology (IT) and the capacity of modern IT systems to service these requirements. Although there have been significant moves towards integration of work by some key organisations in ornithology and marine recording, national coordination is generally poor. Almost the only component of biological recording which has remained relatively constant is the most important source of data, the volunteer specialists and biological societies, usually working in a recreational capacity. Nevertheless, even this

source has modified in response to external changes, particularly in the use computers to manage data and in the ways specialists and societies work with statutory bodies such as the conservation agencies.

2.2 ORGANISATIONS

Numbers and types of organisations

2.2.1 Of more than 2000 organisations which have been identified by CCBR as being concerned potentially with biological recording (Table 2.2.1), only a minority were set up with this as their original objective or responsibility. A substantial proportion of this overall total (such as many local government planning departments and local natural history societies) are believed to lack any formalised mechanisms to collate or hold biological records. All planning departments have an interest in the use of site-based biological data, to enable them to fulfil their obligations under wildlife and planning legislation and to deliver balanced decisions on the use of the local landscape and natural environment. The routes by which many planning departments acquire this information are very varied although an increasing number (mainly in England and in Northern Ireland) provide some financial support for local data centres. Although some local societies often have a tradition of holding detailed records, most are likely to hold only those records which are in some way notable. A few local societies (e.g. Yorkshire Naturalists' Union (YNU), Bristol Natural History Society) hold very large and long runs of records, some of which may be of considerable historical importance. Nevertheless, for many of these 2000 or more organisations, the term *biological recording* is likely to mean very little: this is because they do not recognise that their activities relate to it or, in some cases, they are genuinely not involved with biological recording despite having a need for this type of data.

2.2.2 Many of the organisations covered in the survey fulfil several roles: for example 12% of wildlife trusts could be classified as local records centres, and 60% of local records centres could be classified as museums. In Table 2.2.1 each organisation is attributed once only to its presumed principal role, based on its title and parent organisation/address, and (for those from which a completed questionnaire was received) its own perception of its principal role.

2.2.3 It is impractical to provide a reliable estimate of the number of *active* field biologists and others involved with biological recording (especially species recording) in the UK. Estimated examples of some of the main functional groups involved with recording (particularly the collection of species data) are given in Table 2.2.2. These examples imply an overall figure of less than 50 000 people who could be regarded as a central core of active recorders. This figure is considerably smaller than the overall membership of wildlife and conservation organisations in the UK which has been estimated by May (1993) to be in the order of 1.25 million.

Table 2.2.1 **Summary of the organisational types covered by the CCBR survey in 1992/93**

Principal role of organisation	Estimated number in UK	Number of in CCBR inventory	Number of completed returns
Local records centre	82	82	47
County wildlife trust	47	47	26
Urban wildlife group	45	45	14
Badger group	77	77	43
Bat group	86	86	[1]
Bird group			
National society	3	3	2
County bird club	108	52	29
RSPB local group	175	0	0
Ringing group	120	0	0
County/regional planning department	65	65	13
Other local government planning dept.	>500	38	1 [2]
National natural history societies	115 [3]	15	5
Local natural history societies	1000 [4]	28	4
National voluntary conservation agencies	>30	30	7
Statutory nature conservation agencies	5	23 [5]	20 [5]
BSBI vice-county recorders (UK only)	118	107	85
Environmental consultancies	>100	16	9
Educational establishments	[6]	24	6
Museums with biological collections	396 [7]	42 [8]	33 [9]
National Parks	11	11	9
NRA regions/River Purification Boards	17	17	9
National biological recording schemes	>72	72	38
DoE and other government departments[10]		10	10 [5]
Research Council units and data banks	>30	30 [5]	6
Utility companies and agencies	>50	16	1

Notes

[1] Summary information was provided from the Bat Conservation Trust survey

[2] Other returns incorporated under local records centre

[3] Total from Meenan (1983)

[4] Total from Milner (1994)

[5] Number of individual datasets/databases separately identified and described in the CCBR survey database

[6] The extent to which all types of educational establishments are involved with biological recording was considered to be beyond the scope of the CCBR survey. Figures refer to Field Studies Council centres and some universities.

[7] Total from Garland (1989)

[8] Includes only some of the major national herbaria and zoological collections

[9] Includes overlap with local records centres

[10] Excludes the statutory nature conservation agencies, research councils and utility agencies.

Table 2.2.2 **Estimated numbers of active field biologists as examples of the potential numbers involved with biological recording (mainly species recording) in the UK**

Organisation/group	Number of members (approx.)
1) ***Associations of professionals***	
Institute of Ecology and Environmental Management	250
National Federation for Biological Recording	200
Association of Local Government Ecologists	110
Biological Recording in Scotland Campaign	100
2) ***Non-vocational societies/groups***	
Local bird groups [1]	10 500 [2]
British Trust for Ornithology	9 000 [2]
Local natural history societies[1]	6 500
Regular moth light-trap operators [3]	5 000
Botanical Society of the British Isles	2 700
BRC invertebrate recording schemes	2 500
JNCC/BRC Amphibian & Reptile Monitoring Scheme	1 200
Badger groups	1 000
Bat groups	1 000
JNCC/Butterfly Conservation - Rare moths network	270

Notes
[1] Estimated numbers of members active in recording (NB. Not total membership)
[2] Probable substantial overlap of active membership of local groups with the BTO membership
[3] P Waring (pers. comm.)

Organisational objectives in biological recording

2.2.4 Organisations with biological or environmental recording, records, data or monitoring in their title, such as the national BRC, national recording schemes and local records centres, can be expected to have a formalised role in biological recording. In the case of longer established organisations, this role is likely to have undergone significant changes in focus and emphasis during the period of their existence. For many, the formalisation of activities and development of policies has been retrospective, sometimes in response to a need to justify past and existing work with changing administrations and, in particular, to be accountable for the use of public or charitable funds. Some have moved to a further stage to implement closer control of data, for example through quality assurance procedures.

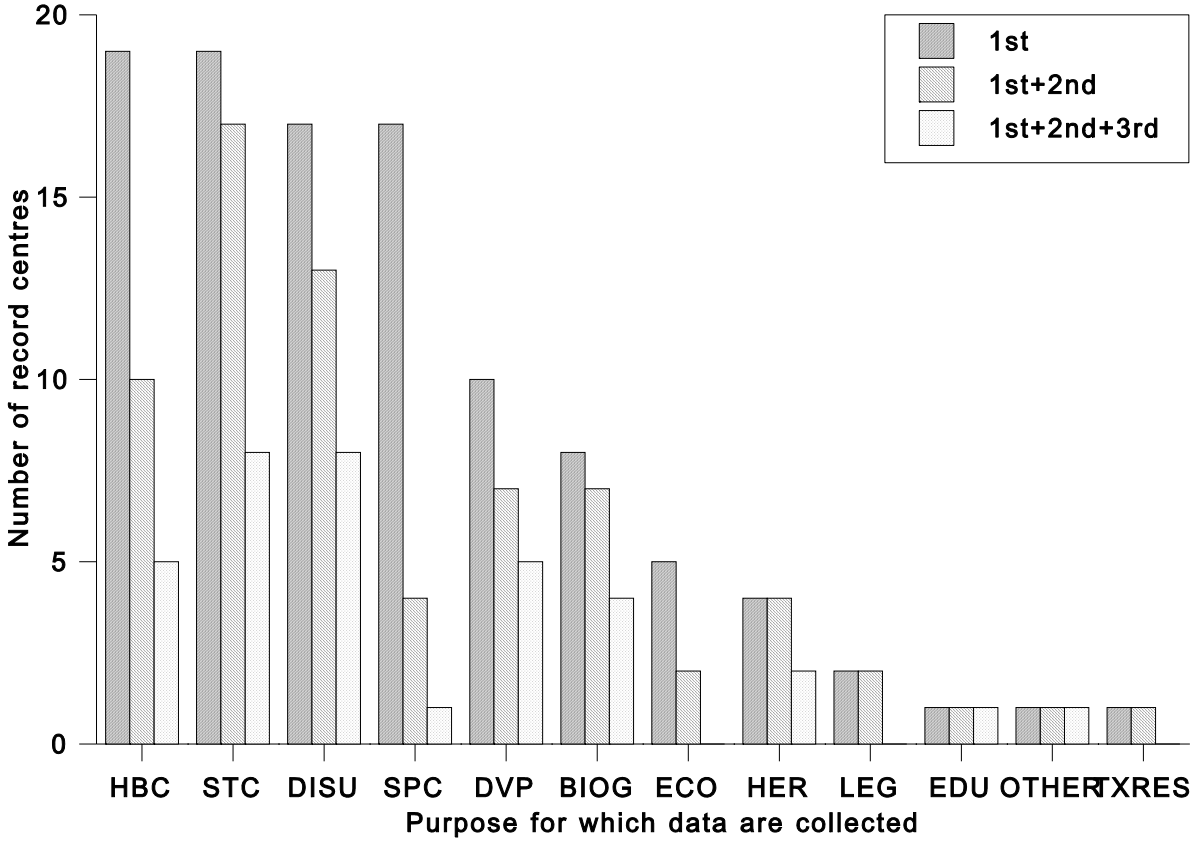
2.2.5 One of the main original objectives of BRC, national recording schemes and many of the older local records centres was to map the distribution of species, but now this is seen as only one of a range of objectives. In the 1990s their roles are concerned with providing national and regional overviews and site information for use in nature conservation and planning, and with biogeographic research, as much as with simple species distribution studies. For example, the objectives of BRC have evolved progressively since it was set up in 1964 (Harding & Sheail 1992). Local records centres have been established without the benefit of a guiding, coordinating or regulating body although the Biology Curators' Group (BCG) and subsequently NFBR have provided a technical forum for those involved with local centres. The only Handbook for Local Biological Records Centres (Flood & Perring 1978) dealt mainly with the practical aspects of operating centres. None of the wildlife trusts was set up with biological recording as part of its original mission, but subsequently most have taken on some form of recording and some now operate records centres (e.g. in Bedfordshire, Gloucestershire and Somerset).

2.2.6 Governmental organisations, such as JNCC and the other statutory conservation agencies, DOE, the National Rivers Authority (NRA), the Ministry of Agriculture, Fisheries and Food (MAFF) and the Forestry Authority have undergone radical changes in recent decades and continue so to do. As a consequence, their roles, as both producers and users of biological records, have undergone change and development. In particular, they now require far more data, for example to audit the effects of policies

and resultant legislation and to provide measurements of environmental changes. A notable example is the investment by MAFF in monitoring of Environmentally Sensitive Areas (ESA). Accountability for the use of scarce public funds is an important force in moulding objectives and priorities for biological recording in the government sector; for example, most of the survey and research undertaken by or for JNCC has been subject to external peer review since 1992, as has the work of BRC.

2.2.7 The overall objectives of biological recording, as defined in 1.1 apply in varying degrees to the wide range of organisations involved. The CCBR survey indicates that site, habitat and species conservation, development planning and biogeography are the main practical purposes for which data are collated and used (Figure 2.1). However, the main motivation for most species recording by volunteers is recreational, but usually with a focused interest in the distribution and conservation of species.

Figure 2.1 **Purposes for which data are collected and used. Figures are for a sample of 154 organisations**



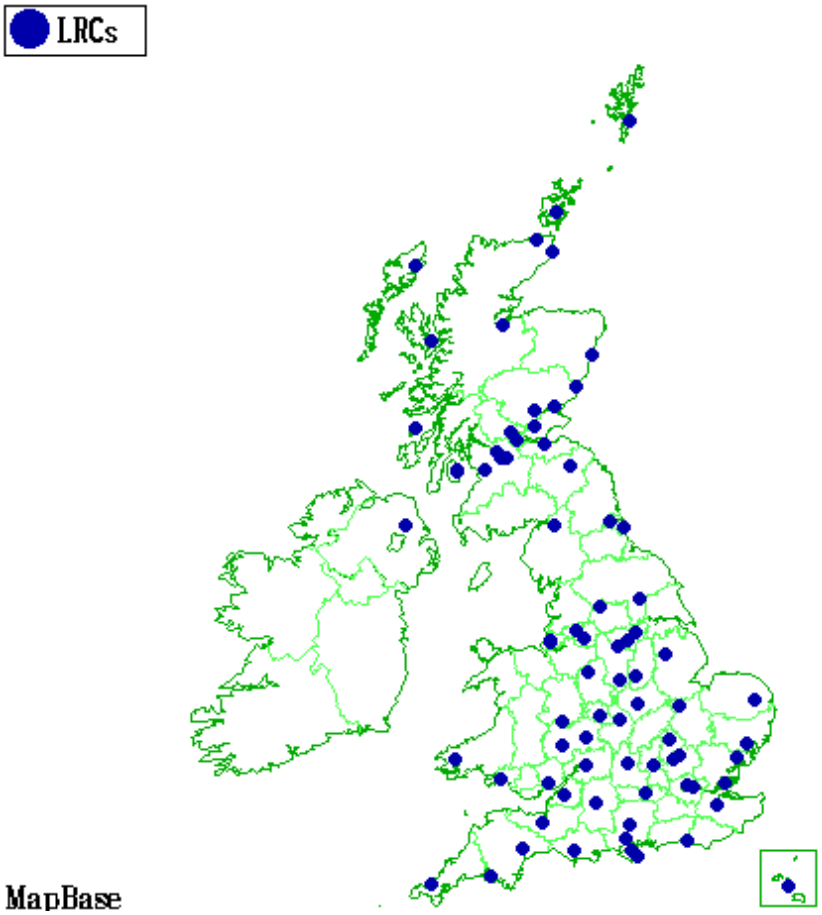
Key to abbreviations used in tables and figures

TOPICS		ORGANISATIONS	
AGF	Agriculture and forestry	BADG	Local badger groups
BIOG	Biogeography	BIRD	Bird clubs and ornithological organisations
CAC	Countryside access	BRC	Biological Records Centre (Monks Wood)
DISU	Dissemination to users	BSBI	Botanical Society of the British Isles
DVP	Development and strategic planning	CNHS	County and local natural history societies
ECO	Ecological research	CONS	Consultants
EDU	Education	COU	County councils
HBC	Habitat conservation	DOE	Department of the Environment
HER	Heritage and education	LRC	Local records centres
LAT/LO	Latitude/longitude	L.AUTH	Local authorities (local government at various levels)
LEG	Wildlife and environmental legislation		
MC	Marketing campaigns	MDA	Museum Documentation Association
PCM	Pollution control and monitoring	MSC	Manpower Services Commission schemes
SPC	Species conservation	MUS	Museums
STC	Site conservation	NGO	Non-governmental organisations
STM	Site management	NPA	National Park authorities
TXRES	Taxonomic research	NRA	National Rivers Authority regions and Water Purification Boards
UTM	Universal Transverse Mercator grid		
WAT	Water resources	NRS	National recording schemes
OTHER	Other applications not listed	NVCA	National voluntary conservation organisations
		OGD	Other Government departments
		RC	Research councils
		RSPB	Royal Society for the Protection of Birds
		SNCA	Statutory nature conservation agencies
		UWG	Urban Wildlife trusts
		WLT	Wildlife trusts
		WWT	Wildfowl and Wetlands Trust
		OTHER	Other organisations

Distribution and coverage of organisations

2.2.8 The geographical location and geographical coverage of individual units concerned with biological recording are more important at the local level than at the national or country level. The main organisations currently holding data at a local level are local records centres, BSBI vice-county recorders, local specialist groups (e.g. bird, badger and bat groups), wildlife trusts, regional units of the country statutory conservation agencies and national park authorities. There is complete coverage of the UK by the statutory and voluntary conservation organisations, but coverage by local records centres is patchy, especially in Scotland and Wales (Figure 2.2).

Figure 2.2 **The distribution of local records centres**


LRCs
MapBase

_____BOX _____

Distribution and coverage - local biological recording units

Country statutory conservation agencies: Complete coverage of UK
Local records centres: Irregularly distributed (Figure 2.2), mainly for local government units
Wildlife trusts: Complete coverage of UK, by single counties, groups of counties, Scotland, Northern Ireland
National parks: All 11 national parks
BSBI vice-county recorders: Complete coverage of the UK
Bird groups: All counties, usually two or more groups in each county
Badger groups: } Almost complete coverage of the UK, but
Bat groups: } with sparse coverage in areas of low population and overlap or duplication in areas of high population

---

Staffing

2.2.9 Staffing levels at individual units are difficult to assess because biological recording frequently is only one of several official or perceived duties of the organisations which responded. The Survey provided information on staffing at 137 organisations, based on 1991/92 figures, which is summarised in Tables 2.2.3 and 2.2.4. Although these data give an impression of the staffing levels, based on 39% sample of the organisations from which data were received, it would be unwise to extrapolate from them to assess the overall numbers of staff involved in biological recording in the UK.

2.2.10 From these 137 returns, the salaried staff (full-time and part-time combined) in the six main areas of work are predominantly professionals with mixed duties (278) and field workers (158), with smaller numbers in managerial (64), clerical (29), data entry (24), computer support (19) and financial (5) roles. The totals for volunteers are distorted by three organisations which returned figures of 9000, 500 and 300 respectively for part-time volunteer field workers. Excluding these three, which almost certainly refer to society memberships, the totals for volunteers are comparable with those for permanent/contract staff. Not surprisingly there is bias towards mixed professional and managerial staff (combined 49%) in the salaried staff, compared with 10% in volunteers, but a bias towards field workers (80%) in the volunteers, compared with 27% in salaried staff. An overall impression can be gained from the CCBR survey that, for example, local records centres have an average of about 2 posts. The figures from the survey are only as complete as the information provided by respondents. For example, there is no assessment of the numbers of permanent and temporary nature reserve wardens, countryside rangers, heritage coast wardens, ESA monitoring staff, all of whom may undertake biological recording at some time as part of their duties, but whose data are often inaccessible.

Table 2.2.3 **Summary of staffing levels by main organisational types.**
The data are from returns provided by 137 organisations (1991/92 figures). They show the total number of staff posts, irrespective of status (e.g. permanent, contract, voluntary, full-time, part-time)

Type of organisation	**Number of staff posts**

(number of respondents)

	1	2	3	4	5	6	7
LRC (37)	**6.9**	**53.4**	**36.9**	**14.5**	**19.7**	**10.6**	**0.2**
Museums (27)	**3.5**	**33.8**	**7.7**	**6.4**	**21.9**	**1.1**	**0**
WLT (20)	**9.7**	**37.3**	**238.2**	**0**	**9.8**	**2.8**	**1.2**
NRS (16)	**3.0**	**10.0**	**433.0**	**2.0**	**3.0**	**3.0**	**0**
DOE/OGD (10)	**4.8**	**29.2**	**17.4**	**0.2**	**0**	**0**	**0**
NRA (8)	**12.0**	**87.0**	**17.0**	**2.0**	**1.0**	**0**	**0**
SNCA (8)	**3.0**	**21.5**	**29.0**	**4.0**	**3.0**	**3.0**	**0.5**
UWG (7)	**2.3**	**8.6**	**20.1**	**5.3**	**0.3**	**0**	**0**
NPA (7)	**3.0**	**34.3**	**6.0**	**1.2**	**0.3**	**1.0**	**4.0**
RC (6)	**12.0**	**9.0**	**92.0**	**1.0**	**4.0**	**1.0**	**0**
NVCA (5)	**6.0**	**10.0**	**540.0**	**4.5**	**5.0**	**1.0**	**0**
Ornithology (4)	**10.0**	**11.0**	**9113.0**	**10.0**	**10.5**	**2.0**	**0**

Key to staff posts

1 Managerial
2 Mixed duties
3 Field workers (NB The totals provided by respondents for the numbers of field workers clearly were the result of a variety of interpretation and cannot be used to estimate accurately the numbers of field workers associated with organisations.
4 Clerical
5 Data entry
6 Computer support
7 Financial

Key to abbreviations - see page 10.

Table 2.2.4 **Summary of the total numbers of staff posts involved with biological recording in 137 organisations (1991/92 figures). Shows the main work types and employment status.**

Work type	Employment status					
	Permanent		Contract		Voluntary	
	Full time	Part time	Full time	Part time	Full time	Part time
Managerial	**43.8**	**18.9**	**0**	**1.5**	**0**	**9.3**
Mixed duties	**193.9**	**37.6**	**28.5**	**18.3**	**13.0**	**22.3**
Field workers [1]	**51.5**	**30.5**	**28.7**	**47.1**	**0**	**10238.6**
Clerical	**10.8**	**7.4**	**9.0**	**2.1**	**0**	**13.7**
Data entry	**6.6**	**3.1**	**3.5**	**10.6**	**1.0**	**35.5**
Computer support	**8.3**	**2.5**	**2.0**	**6.0**	**0**	**5.3**
Financial	**4.7**	**0.3**	**0**	**0**	**0**	**1.0**

[1] The total for volunteer field workers is distorted by the returns from three organisations which included membership figures.

Funding

2.2.11 The majority of organisations responding to the CCBR survey were unable to give details of their funding. For those that did, as with staffing levels, the multiplicity of duties undertaken by many units made it difficult to disentangle the funding allocations specifically for biological recording activities from overall departmental budgets. In particular, the perception of units as to what constituted the costs of their operation varied from those that gave the full economic cost (including all institutional overheads), others that gave direct staff costs and consumables only, to what (at £475/year) could only be an annual budget for consumables. Local records centres which responded with details of their funding rely mainly on core funding from their parent or host institution (Table 2.2.5). Even one of the few commercially based local records centres gave a return indicating that 23% of its budget came from core funding, but with almost all the rest (71%) from service agreements and data sales. The estimated average level of core funding of local records centre, with 2 salaried posts, is about £58 000 per year (1992/93 prices).

Table 2.2.5 **Summary of sources and amounts of income in 1992/93 at local records centres**

Source of income (number of respondents)	Range of income £
Core funding from parent/host institution (11)	475 to 367 000
Contract work (7)	50 to 21 000
External grant income (5)	1 200 to 20 000
Provision of services (4)	50 to 52 600

Organisational Policies and standards

2.2.12 The role of published policies and standards in biological recording is to promote confidence in the quality of records and their management. The need for standards has been a recurring topic in meetings and publications of the National Federation for Biological Recording (e.g. authors in Copp & Harding 1985, Harding & Roberts 1986, Stansfield & Harding 1990) and is seen by many as central to the future introduction of accreditation for records centres. However, little has been achieved so far in the way of published examples of policies and standards. Many respondents to the Survey said that policies were being considered or actively being developed but few actually had examples in place. A draft general Code of Practice for the collection and care of biological records has been published by Ely (1994) and Paine (1992), but none of the respondents to the Survey had yet used either as models.

2.2.13 The Biological Recording in Scotland Campaign (BRISC) has established an accreditation system for collectors and collators of biological records. The BRISC accreditation has five grades, to suit different levels of records management and services; from full function museum record centres to local natural history groups. Progression to each grade is achieved by meeting criteria such as the number of animal and plant groups covered, standard of records management and services provided, as defined by the BRISC Committee. At present, written policies or documented protocols are not part of the criteria for BRISC accreditation although this is being considered.

2.2.14 There are no specific legal requirements relating to the management or provision of biological records that require organisations to maintain certain standards or publish policies, although a number of more general regulations do affect the operating policies of organisations. The care and quality of records are included under the provisions of the Data Protection Act where data are computerised and contain

reference to identifiable people (e.g. recorders). Access to, and charging for, data is covered by the EC Directive on freedom of access to environmental information. This Directive applies in particular to government funded organisations and local records centres funded by local authorities. NRA regions are statutorily obliged to publish data on water quality in a public register, but not to do the same for biological survey data. However, several NRA respondents recorded that they had policies to make survey information available.

2.2.15 In the case of local records centres, the lack of readily available policy statements, particularly on data quality and access to data, has the potential to undermine confidence in centres by both the suppliers of data and the prospective users of data. A recurrent criticism, made by those environmental consultancies, statutory bodies and utility companies which were contacted during the course of the Survey, concerned the absence of comparability between local records centres, particularly in relation to access to data and charging policies. At a local level, cooperation and confidence between organisations (e.g. planning departments, local records centres and wildlife trusts) is good, being based on personal contact and close working relationships. However, the absence of formalised objectives and operating procedures can hamper the development of new funding arrangements and discourage wider use of the information held by local centres.

2.2.16 The effort expended in preparing policy statements, establishing quality control procedures and developing standards is repaid by better understanding by staff of their roles within organisations, which results in improved reliability and professionalism as perceived by potential users. This has been recognised by a number of local records centres and wildlife trusts anxious to shed their earlier *amateur* image in the process of building new relationships with sponsors and business partners. Examples of detailed business development plans containing policy statements and objectives were received from one wildlife trust, one urban wildlife trust and one local records centre. At least two wildlife trusts are considering the application of BS5750 (ISO 9001) Quality Assurance to their work and one wildlife trust has contracted an external consultant to prepare operating policies and quality assurance protocols.

2.2.17 The results from the Survey relating to policies are summarised in Table 2.2.6.

Table 2.2.6 **Policies held in relation to biological recording in a sample of 198 organisations that provided information to the CCBR Survey**

Policy Statement	Number of replies	Percentage of sample (N=198)
Statement of aims	85	43
Data collection and collation	35	18
Charging for data and services	28	14
Data access	26	13
Code of practice for data management	22	11
Data validation	18	9
Data security	13	7
Publication	13	7
Other policies	9	5
Backup and archiving of data	2	1

2.2.18 Information on policies was supplied by 198 respondents to the Survey (56% of the total returns), of which 85 organisations had a written constitution or statement of aims. This was across the

range of organisation types from small badger groups to large centrally funded bodies including the statutory nature conservation agencies. Within the larger organisations the publication of mission statements, strategies, attainment targets and operating policies in relation to the whole organisation is becoming common, but examples of specific policies and standards relating to biological recording are few and generally informal. A number of organisations, including BTO, have formal mission statements which encompass recording and the National Park authorities publish detailed Park Plans which contain their objectives and strategies for conservation and survey.

2.2.19 Of the 35 organisations which had formal policies on the scope of data collection and collation, only six were local records centres (13% of local records centres that responded), although most records centres operate informal policies or follow locally established practices. Numbers for all other types of policy are much lower: of the 198 respondents, 18 had written statements on data validation (including only two local records centres), 13 had a policy on data security and two on data backup and archiving.

2.2.20 Of the 26 organisations which had written policies on data access, eight were local records centres. Policy statements on access tended to be vague and individualistic although generally of the form: 'free and complete access to *bona fide* wildlife conservation bodies/students otherwise each case is decided on its merits': there being no guidelines on what constitutes a *bona fide* body or what is considered meritable. Another feature of existing data access policies is reference to withholding information on 'confidential' or 'rare' species, again with no explanation of the criteria by which these are judged although one policy mentioned Red Data Book species.

Charging policies

2.2.21 Charging for biological information is perhaps one of the most contentious issues facing organisations and one for which clear statements of policies are probably essential. Charging policies must take account of the issues relating to the ownership of data and copyright, which are covered in Chapter 3. The Survey shows that the majority of wildlife trusts, local records centres, local societies, recording schemes and smaller organisations have only informal arrangements for charging and that most transactions are either free of charge or are dealt with on an *ad hoc* basis, often based on what the supplier thinks the potential customer will be prepared to pay. This approach leads to many anomalies, particularly in instances where data are exchanged between wildlife trusts, local records centres and planning departments, and an enquirer may approach each organisation in turn, looking for the best price. The lack of coordination and the variation in charging policies between organisations with apparently similar functions in the supply of data creates problems for regionally based users. For example, a potential user may be given data free of charge by one supplier but be asked to pay considerable sums for comparable data by another supplier, even in the same county. Such anomalies are frequently related to the extent to which the supplier of data has secured sources of funding and is not wholly or substantially dependent on contract earnings to maintain its core activities.

2.2.22 Only 28 organisations declared written charging policies in the Survey. A small number of organisations in the public sector have fixed and published scales of charges. The National Parks, for example, acting as planning authorities, publish full scales of fees (e.g. for submitting building plans) and charges for services including, in some cases, the supply of data. At the other extreme, one national museum has a charging policy which is confidential.

2.2.23 Two aspects are common to almost all the charging policies supplied to the Survey or reviewed during interview.

- **All organisations have classes of user to whom charges do not apply.**
  These may include sponsoring organisations, organisations with whom data are exchanged, *bona fide* naturalists, conservation organisations, natural history societies, educational users and the national Biological Records Centre.
- **Most organisations have identifiable customers who pay for data and services.**
  These may include environmental consultants, private companies, NRA regions and utility companies. (However, in some cases NRA regions and some utility companies may be exempted from charges because they are partners in the exchange of data or are sponsors).

2.2.24 The position of individual members of the public is not always clear: most museum-based local records centres regard the provision of information to the public as being central to their activities, although some will introduce charges where requests for data are regarded as excessive, or will refuse access to data altogether if the request is suspect (e.g. from a known collector). A small number of local records centres do not provide services to the public but, in these cases, there are no formal policies for charging and staff judge requests 'on their merits'.

2.2.25 The possibility of contravening access to information regulations and possible dispute over the ownership of data leads all those that supplied copies of charging policies to state that charges are for the labour and resources of extracting and copying data, not for the actual data. Only one organisation had a specific statement that data is a valuable resource and charges should reflect this (but even this organisation had a long list of users which were exempt from charges). Two organisations, which use mainly site-based data from in-house surveys, charge to reflect the original costs of collecting data. Typically charges were calculated according to the number of hours labour involved in extracting and preparing data (examples given were between £25 and £30 per hour), but some holders of predominantly site-based data charged a rate per site (e.g. about £50). Most declared a minimum charge (about £50) and a few had more complex charging arrangements based on the format of the data supplied (e.g. publication, map, floppy disk, photocopies) and its original storage medium (computer or paper files). One local records centre based in a county council mentioned recovery of the full economic costs in its charging policy, but the list of exemptions to this charge was long.

2.2.26 The perception that environmental information has an inherent financial value is not uncommon, but none of the data centres covered by the Survey is wholly or primarily funded through the provision of information to users on a wholly commercial basis. At the present time, almost every local records centre is being funded either directly from the core funding of its parent organisation, through grant aid, or service agreements with local authorities. Where other income is made, it is derived from commissioned survey or the provision of expert advice (such as interpreted data). In responses to the Survey, only 20 organisations listed services, such as the provision of data (other than to core funding organisations or under contract), as a source of income; in most cases earnings were only small sums ranging from £20 to £2500 per year (average about £500 in 1992). Even in the USA, The Nature Conservancy, which is a private organisation, has been able to raise only 5% of its income from charging for data.

2.3 BIOLOGICAL DATA HOLDINGS

2.3.1 For the first time, the national resources of data on species, biotopes and land cover have been partially assessed in the CCBR Survey. The Survey was targeted in such a way that information on most of the principal collating and managing agencies could be compiled in the database. Regrettably, a few important agencies failed to respond to the survey and others supplied incomplete information on their data holdings. Summaries of the data resources are given in the following sections: sources of data (2.3.2 to 2.3.10), the temporal range of data (2.3.11 to 2.3.20), data on taxa (2.3.21 to 2.3.32), data on biotopes and land types (2.3.33 to 2.3.36) and data on the marine environment (2.3.37 to 2.3.41). Some information from the CCBR's review of biological data holdings and information sources has been used already in the preparation of Chapter 9 of *Biodiversity, the UK Action Plan* (Cm.2428).

Sources of biological records

2.3.2 This section reviews the sources of data in the sense of the collectors and suppliers, such as volunteers, in-house staff and contractors, and the types of organisations which hold data and are, therefore, sources of data to the user community. It does not consider the methods for sourcing data, such as field survey, air photography and satellite imagery, which are dealt with elsewhere (Section 2.4).

2.3.3 The Survey received quantified returns from 951 respondents which included information on sources of data and datasets which are summarised in Table 2.3.1. The numbers of records in these datasets are expressed as percentages of the national totals of records contained in the datasets (Table

2.3.2). When considering the actual numbers of records in these datasets, the strong national and local bias towards records of birds should be noted (Table 2.3.2) and a further analysis has been made (Table 2.3.3) which excludes all ornithological data.

2.3.4 The direct contribution of data by volunteers accounts for about 29% of the total number of all datasets, but for over 70% of all taxa records (not including national recording schemes and collections) and 36% of all biotope/land type records. This figure is even higher if the hidden contributions of data, in the form of records extracted from collections, publications, and copies of records from other sources are included, many of which would have originated from volunteers. These figures greatly exceed the next major sources: surveys carried out by the in-house staff of organisations, contractors, Manpower Services Commission (MSC) schemes and the statutory conservation agencies. There are important differences in the sources used by the major types of organisation identified in the Survey.

2.3.5 Volunteers are a particularly important source of data for bird clubs and the major ornithological organisations. This is especially significant because ornithological datasets account for 60% of the national total of taxa records! County wildlife trusts and urban wildlife groups also rely heavily on volunteers for species data on taxa, as do local and regional natural history societies and national recording schemes.

2.3.6 Government-funded organisations and some local government departments are dependent mainly on surveys by in-house staff and commissioned surveys. The majority of taxa data held by the statutory nature conservation agencies were collected by their own staff or by contractors, although in many cases volunteers have made significant contributions to some datasets, most notably the Invertebrate Site Register and ornithology data supplied by BTO, the Wildfowl and Wetland Trust (WWT) and RSPB. Contracted surveys are the main source of data for other government departments and local authorities planning departments, although information on taxa is not an important part of the data holdings of either type overall. Local planning authority ecologists apparently contribute less than 10% of the taxa datasets held by their organisations, but their work is expected to be concerned mainly with commissioning surveys from others. The NRA regions rely almost entirely on data from in-house sources, largely because water quality survey data which relate to taxa are not available from other sources. The National Park authorities collect more than 40% of their own data, but other information comes from a wide range of sources.

2.3.7 The entry for the Biological Records Centre is anomalous in that it implies that no use is made of amateurs whereas the figures include 24% from national recording schemes (and therefore mainly from volunteers) and from collections and publications, which also originate mainly from volunteers. Almost all of these national schemes are coordinated by BRC.

2.3.8 Various government employment schemes (such as MSC) were a feature of biological recording at the local level during the 1970s and 1980s. Data collected by surveys as part of these schemes contribute only about 5% to the national total of datasets and less than 10% each to the main users of these schemes - wildlife trusts, local records centres, museums and National Parks authorities. However, the schemes made more significant contributions to surveys of biotopes and land types at the local level (Table 2.2.4).

2.3.9 It is essential to distinguish between the types of 'habitat' information associated with taxa data and the more rigorous and demanding characterisation of biotopes and land cover, for example in Phase 1 Habitat surveys and the National Vegetation Classification (NVC). The sources of data on biotopes and land types are similar to those for taxa datasets, but there are some important differences (Table 2.3.4). The most important difference is the role of volunteers, who make less overall contribution to these datasets than to taxa datasets in almost all organisations, although they do make significant contributions to biotope and land type datasets in museums and wildlife trusts, for example in Phase 1 Habitat surveys.

Table 2.3.4 **Sources of records in biotope and land type datasets**

No. datasets	WLT 99	COU 50	SNCA 43	RC 2	OGD 26	NVCA 23	NRS 10	NPA 38	NRA 21	MUS 57	LRC 117	ALL 410
Amateurs	31.5	6.8	0.5	0.0	1.3	75.5	50.2	9.4	0.0	22.5	16.0	19.4
Copy Records	0.3	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	4.3	1.4
Contractors	13.5	18.4	40.8	0.0	68.5	0.6	6.0	38.0	7.5	11.1	12.4	18.4
In-house	31.2	40.1	21.2	100.0	25.0	18.0	17.0	26.9	87.1	19.7	17.0	29.3
MSC	8.9	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.2	25.7	10.1
NGO	5.1	8.0	0.7	0.0	1.3	0.8	13.1	7.4	4.8	11.6	7.6	5.8
Other	0.1	1.2	0.0	0.0	1.9	4.6	1.5	2.0	0.0	1.9	0.9	1.2
Publications	0.5	4.5	1.3	0.0	0.4	0.0	4.0	0.0	0.1	4.7	1.2	1.3
SNCA	9.1	9.4	35.0	0.01.5	1.5	4.2	13.5	0.4	15.4	15.8	12.8	

Key to abbreviations see page 10

2.3.10 Development planning and much of the focus of practical nature conservation measures are directed towards defined sites. Consequently, much of the work of professional ecologists is site-based and is concerned mainly with biotopes and land types rather than taxa (except rare species). In-house and contracted surveys are the main sources of biotope and land type data for all organisations other than bird clubs and national taxa-based recording schemes. Some 45% of the biotope and land type datasets held by the local records centres, local wildlife trusts and county planning departments that responded to the questionnaire, have been gathered in this way, particularly through schemes funded through MSC or grant-aided by the Nature Conservancy Council (NCC).

Temporal ranges of biological data

2.3.11 Biological records have neither been made nor accumulated at a constant rate. In particular, there has been a notable increase in biological recording since 1945. Data obtained from the questionnaire have been grouped into five unequal periods. The first two, pre-1900 and 1900-1939 are somewhat arbitrary but the remaining three, 1940-69,1970-1979 and post-1980, reflect the main phases of recent recording. The Survey data are summarised in Tables 2.3.5, 2.3.6, 2.3.7 and 2.3.8. They illustrate differences between the periods for taxon-based records, biotope and land-type records, and between the data holdings of the main organisation types.

2.3.12 The temporal ranges of 1085 taxa datasets in Table 2.3.5 demonstrate that over 60% in the sample have been collected since 1980 and just over 85% since 1970. There are slight variations between taxonomic groups. Over 60% of the records of popular groups (e.g. birds, mammals, amphibians, reptiles and Lepidoptera) fall into the post-1980 range. Over 50% of records of the more specialist groups (e.g. flies and other insects) date from earlier periods, for although these groups show peaks in the 1900-39 period, this probably is an underestimate of the true levels of recording. The interpretation of these differences is not obvious but some of the differences reflect factors other than popularity. For example, the large numbers of records of fungi and molluscs post-1980 reflect the more directed activities of specialist recorders. A significant underestimate of early records remains because the survey did not include the major museum and herbarium records, which are not readily accessible (see 2.3.11 & 2.3.32). Their inclusion would have greatly enhanced the number of records prior to 1940.

Table 2.3.5 **Average percentage of records in different date ranges for a sample of 1085 taxon-based datasets**

	Pre-1900	1900-39	1940-69	1970-80	Post-1980	N=1085
Fungi	**4.012.6**	**5.4**	**10.4**	**66.8**	**27**	
Lower Plants	**5.210.8**	**12.4**	**17.8**	**54.2**	**80**	
Higher Plants	**4.54.2**	**11.8**	**22.3**	**58.5**	**150**	
Invertebrates*	**2.312.0**	**4.4**	**25.7**	**55.6**	**194**	
Molluscs	**3.95.5**	**11.5**	**17.3**	**62.4**	**33**	
Arachnids	**1.110.1**	**5.9**	**24.3**	**57.7**	**37**	
Insects	**4.410.6**	**9.5**	**28.0**	**47.3**	**176**	
Flies	**3.715.2**	**8.1**	**22.7**	**47.2**	**31**	
Beetles	**5.816.1**	**8.2**	**17.4**	**51.1**	**36**	
Lepidoptera	**4.58.5**	**7.1**	**17.3**	**62.0**	**48**	
Vertebrates*	**2.93.4**	**1.0**	**18.7**	**74.0**	**12**	
Fish	**4.34.5**	**6.6**	**29.4**	**50.6**	**25**	
Herptiles	**4.24.8**	**6.0**	**15.9**	**67.7**	**50**	
Birds	**3.94.6**	**7.5**	**19.8**	**63.6**	**117**	
Mammals	**1.82.7**	**9.5**	**17.2**	**68.5**	**69**	
All Plants	**4.69.2**	**9.9**	**16.8**	**59.8**	**257**	
All insects	**4.612.6**	**8.2**	**21.3**	**51.9**	**291**	
Other inverts	**2.49.2**	**7.2**	**22.5**	**58.6**	**264**	
Birds	**4.55.5**	**7.3**	**18.1**	**62.8**	**117**	
Other verts	**3.33.9**	**5.7**	**20.3**	**65.2**	**156**	
All Data4.0	**9.08.2**	**21.7**	**63.4**	**1085**		

*** undifferentiated**

2.3.13 Some organisations (Table 2.3.6), concerned mainly with development planning and conservation, possess little or no pre-1980 taxon-based records, *e.g.* country conservation agencies, government departments, local authorities, national parks. The period of their records reflects new kinds of user-demand, although their total contribution to the national records is small (about 5%), of which the statutory conservation agencies account for about 3%. In some cases these recent records reflect intensive monitoring activity, *e.g.* in national parks, 97.6% of whose records are post-1980. In others, notably government departments such as DOE, they reflect the development of new projects such as the Countryside Information System (CIS) and an increasing concern to be informed statistically about the current state of the countryside.

Table 2.3.6 **Average percentage of taxon-based datasets collected in selected date periods, arranged by organisation types**

	Pre 1900	1900-39	1940-69	1970-80	Post 1980	
WLT/UWG	**0.0**	**0.2**	**0.8**	**25.1**	**73.8**	**Museums**
19.8	**23.1**	**8.6**	**13.1**	**35.2**		
LRC	**3.1**	**11.8**	**9.3**	**25.4**	**49.7**	
NRS	**4.2**	**8.2**	**10.7**	**21.0**	**53.3**	
BSBI	**5.2**	**5.7**	**17.4**	**29.8**	**45.3**	
BIRD	**0.1**	**0.4**	**11.2**	**27.4**	**64.4**	
NGOs	**0.7**	**4.5**	**16.4**	**23.9**	**54.3**	
BRC	**3.2**	**12.8**	**30.4**	**30.6**	**23.3**	
COU	**0.0**	**0.0**	**0.2**	**8.8**	**91.0**	
OGD/DOE	**0.0**	**0.0**	**0.0**	**3.8**	**96.3**	
SNCA	**0.0**	**0.0**	**0.6**	**7.4**	**91.0**	
NRA	**0.0**	**0.0**	**3.8**	**36.4**	**63.4**	
NPA	**0.0**	**0.0**	**0.0**	**2.4**	**97.6**	

Key to abbreviations - see page 10

2.3.14 Organisations with a wider temporal spread of records are mostly those concerned with species biogeography and taxonomic research, *e.g.* research councils, BRC, BSBI, museums, museum-based record centres and national recording schemes. They also hold a larger percentage of the national taxon-based data resource. For example, about 45% of the records of the BRC pre-date 1970, *i.e.* around 5% of the national total, but it also holds nearly 7% of the national total of post-1960 records and this percentage would be even higher if birds were excluded from the national totals. Although bird organisations hold records covering a wide temporal span almost 65% have been collected since 1980. Nevertheless, the very large numbers of bird records (>40 million) means that their pre-1970 records account for almost 10% of all UK taxon-based records.

2.3.15 The middle of the spectrum is represented by the wildlife trusts and NRA regions even though more than 95% of their records are post-1970 due to their deliberate expansion of recording in recent years.

2.3.16 Biotope and land type datasets include an even greater proportion of post-1980 records than do taxon-based datasets (Tables 2.3.7 and 2.3.8). In a sample of 457 datasets analysed, 78.8 % of all records were post-1980 and 95% were post-1970. They show a near logarithmic rate of increase in record collection (Figure 2.3.1) reflecting the increasing demand for biotope and land type data for use in conservation and planning over the last twenty-five years.

Table 2.3.7 **Average percentage of biotope and land type datasets collected during selected periods, arranged by type of survey**

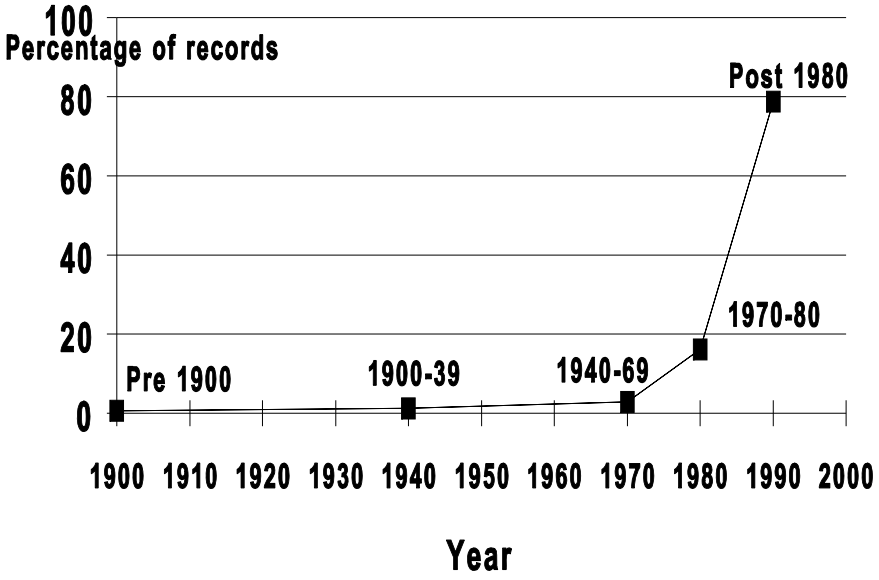
	N=460	Pre 1900	1900-39	1940-69	1970-80	Post 1980	Total %
General Landuse	**27**	**0.2**	**4.1**	**1.5**	**21.3**	**72.8**	**99.9**
Phase I	**30**	**0.1**	**0.1**	**0.3**	**4.8**	**94.7**	**100.0**
Other General	**20**	**0.0**	**0.0**	**5.0**	**27.5**	**67.5**	**100.0**
Woodland	**59**	**0.6**	**0.1**	**1.8**	**18.0**	**79.1**	**99.7**
Grassland	**49**	**0.0**	**0.0**	**1.0**	**19.7**	**79.1**	**99.7**
Heath & Moor	**23**	**0.0**	**0.0**	**1.4**	**18.6**	**79.3**	**99.3**
Wetland	**37**	**0.8**	**1.1**	**2.0**	**16.1**	**79.9**	**99.9**
Freshwater	**57**	**0.5**	**0.8**	**1.9**	**18.6**	**78.2**	**100.0**
Coastal & Est.	**58**	**0.3**	**0.5**	**0.9**	**16.4**	**81.8**	**100.0**
Inland Rock	**8**	**0.0**	**0.0**	**4.1**	**32.1**	**63.8**	**100.0**
Urban	**20**	**0.0**	**1.3**	**1.9**	**10.6**	**86.3**	**100.0**
Agriculture	**17**	**0.0**	**0.0**	**3.1**	**17.8**	**78.8**	**99.7**
Marginal Land	**29**	**0.1**	**0.2**	**1.7**	**17.2**	**80.5**	**99.8**
Quarries	**10**	**0.0**	**0.0**	**3.4**	**25.1**	**71.0**	**99.5**
Other	**16**	**0.1**	**0.2**	**2.1**	**20.1**	**77.0**	**99.7**
Averages:		**0.2**	**0.6**	**2.2**	**18.9**	**78.0**	**99.8**

Table 2.3.8 **Average percentage of holdings of biotope and land type records collected during selected periods, arranged by organisation type**

	N=457	Pre 1980	1900-39	1940-69	1970-80	Post 1980
WLT	**88**	**0.0**	**0.0**	**1.2**	**21.6**	**77.2**
LRC	**130**	**0.0**	**0.0**	**0.9**	**18.9**	**79.8**
MUS	**19**	**1.2**	**1.8**	**5.7**	**18.4**	**72.9**
NRS/BSBI	**6**	**5.0**	**10.0**	**10.0**	**10.0**	**60.0**
BIRD	**18**	**0.0**	**0.0**	**10.6**	**67.8**	**21.7**
NVCA	**8**	**0.0**	**0.0**	**0.0**	**7.4**	**92.5**
COU	**33**	**0.0**	**0.0**	**0.0**	**10.0**	**90.0**
SNCA	**58**	**0.9**	**0.4**	**3.2**	**9.1**	**86.5**
OGD/DOE	**21**	**0.0**	**1.2**	**1.2**	**2.6**	**95.0**
RC	**3**	**0.0**	**0.0**	**0.0**	**0.0**	**100.0**
NRA	**17**	**0.0**	**0.0**	**0.3**	**11.0**	**89.0**
NPA	**56**	**0.0**	**1.8**	**0.1**	**16.8**	**81.3**
Averages:		**0.6**	**1.3**	**2.8**	**16.1**	**78.8**

Key to abbreviations - see page 10

Figure 2.3.1 **Rate of increase in the collection of biotope and land type data (N=460 datasets)**



2.3.17 In the sample, only museums, the BSBI network and national recording schemes (combined as NRS/BSBI in Table 2.3.8) and statutory nature conservation agencies hold pre-1900 data, amounting to less than 1% of the overall total. Thereafter, data for later date ranges tend to reflect the main periods of activity of the organisations concerned. Thus local records centres and wildlife trusts, which have expanded steadily since the early 1970s provide a good percentage of records from the 1970s as well as the 1980s and 1990s. The high proportion of recent records for bird organisations (68%) reflects the number of datasets held by the BTO.

2.3.18 The date ranges of biotope, land-type and monitoring datasets reflects the increase in land-cover, habitat and landscape survey since the late 1970s. This stems from an increasing need for strategic overviews, at local, regional and national levels, particularly for use in nature conservation and local structure planning. Monitoring of landscape change has become increasingly important in recent years and new techniques, such as remote sensing, are being used to update existing maps and extending the collection of data to new areas. At the local level, much surveying has been aided by the availability of

surveyors through government employment schemes and grant-aid from the statutory conservation agencies.

2.3.19 Compared with taxon-based datasets, the highest proportion of the total resource of biotope and land-type records is in the hands of agencies responsible for planning and conservation. They are principally recent records and experience shows that there is a high degree of cooperation and sharing of data at the county level between wildlife trusts, local records centres, local authority planning departments and the statutory nature conservation agencies. However, the use of existing data is far less effective than it might be. Less than 10% of the data collected by each organisation is used by other organisations (see 2.6) and there are many examples of planning and conservation organisations contracting new surveys rather than investigating the resources of, and participating in, local networks.

2.3.20 Users of biological information, whether organisations affected by, or concerned with, planning or conservation legislation, or concerned with scientific issues, rely for historical information on national record schemes, local record centres or BRC. However, only 13% of taxon-based records and less than 1% of biotope and land-type records relate to the pre-1940 period in the samples surveyed here and for the pre-1900 period the figures are 4% and 0.2%, respectively. Most of these datasets do not include records which could be trawled from museum collections, major herbaria, local natural history society archives and publications. The resource of taxon-based records in museum collections is immense. Although the true size of this potential source of early biological records has never been quantified (Williams 1987), a recent estimate (Walley pers. comm.) as part of the work of the Federation for Natural Science Collections Research (FENSCORE) suggests that there may be 15-20 million biological specimens from the UK, held in UK museums, which are a potential source of useful, localised data. Unfortunately, such collections are difficult to utilise directly as they are rarely in a form that can be used to extract records easily. They are often poorly documented, and many are known to be poorly curated. With a few exceptions, mainly related to surveys of taxonomically difficult groups, museum collections provide only a very small input to the overall use of biological records, particularly by the conservation and planning organisations.

Taxa

2.3.21 Tables 2.3.9-12 summarise the holdings of taxon-based data identified from an overall total of 1385 datasets. In 991 cases quantitative estimates of the extent of these holdings were given by the respondents. Information in the remainder (394) was imprecise (e.g. 'many', 'few', 'thousands') or contained no indication of quantity. Quantified information has been summarised to give totals for appropriate taxonomic groupings (Table 2.3.9), but in some cases respondents gave information which was not differentiated even to these levels of taxonomic detail. Table 2.3.10 summarises the total numbers of taxon-based records (in excess of 1 million) held by the major organisations and organisational types. Table 2.3.11 summarises the distribution of holdings of the main groups of taxa records by organisational types and principal specialist organisations. The numerically larger data holdings are considered in more detail in 2.3.39 to 2.3.44 Almost without exception these datasets relate to the spatial and temporal occurrence of taxa; data on the ecological attributes of taxa are considered in 2.3.31. The Survey database contains considerably greater detail about data holdings than are summarised in Tables 2.3.10-12 This extra information is referred to throughout subsequent chapters.

Table 2.3.9 **Numbers of taxon based records sorted by organisation types and the percentage of the total number of records identified by the Survey**

Based on quantified information on 991 of the 1385 datasets reported to the Survey

Org Type	Number of records	% of total	Number of datasets in CCBR Survey
BTO	**28 288 655**	**44.5**	**9**
Bird Clubs	**9 062 020**	**14.3**	**46**
BRC	**6 128 663**	**9.6**	**44**

BSBI	**5 783 524**	**9.1**	**88**
LRC	**5 656 813**	**8.9**	**679**
SNCA	**1 998 290**	**3.1**	**79**
NRS(-BSBI)	**1 628 400**	**2.6**	**18**
MUS	**1 556 761**	**2.4**	**63**
WLT/UWG	**971 120**	**1.5**	**122**
NRA	**780 100**	**1.2**	**38**
RC	**586 000**	**0.9**	**20**
NVCA	**341 325**	**0.5**	**16**
CC/L.AUTH	**268 246**	**0.4**	**47**
NPA	**182 691**	**0.3**	**38**
OGD	**181 460**	**0.3**	**20**
DOE	**62 500**	**0.1**	**10**
Consultants	**54 100**	**0.1**	**3**
County Nat Hist Socs	**10 000**	**0.0**	**1**
Badger Groups	**1 400**	**0.0**	**44**
Total	**63 542 068**	**100.0**	**1385**

Key to abbreviations - see page 10

Table 2.3.10 **Taxon based data holdings sorted by taxonomic groups**

1. **Overall number of datasets known to contain records of that taxonomic group (N = 1385).**
2. **Number of datasets of that taxonomic group for which quantified information on the number of records was provided by respondents (N = 991).**
3. **Total number of records quantified for that taxonomic group (derived from the datasets in 2 only).**
4. **Total number of records (in 3) expressed as a percentage of the national data resource quantified by the CCBR survey (63 542 068 records).**

Taxonomic group	**1 Total datasets**	**2 Quantified datasets**	**3 Total records**	**4 % of total data**
All species (unspecified)	**4**	**2**	**120 000**	**0.2**
Cryptogamic plants		**130**	**104**	**2 152 700** **3.4**
Bryophytyes	**93**	**69**	**1 687 515**	**2.7**
Lichens & Fungi [(1)] **37**	**35**	**465 185**	**0.7**	
Vascular plants	**182**	**129**	**13 937 232**	**21.9**
Non-insect invertebrates	**314**	**239**	**1 928 156**	**3.0**
Invertebrates (unspecified)	**226**	**169**	**1 460 416**	**2.3**
Snails & slugs (Mollusca)	**39**	**30**	**354 182**	**0.6**
Spiders, mites (Arachnida)	**48**	**40**	**113 558**	**0.2**
Insects	**385**	**285**	**3 385 872**	**5.3**

Insects (unspecified)		**243**	**177**	**783 999**	**1.2**
Butterflies & Moths		**60**	**44**	**1 556 004**	**2.4**
Beetles	**44**	**33**	**739 705**	**1.2**	
Two-winged flies (Diptera)	**38**	**31**	**406 458**	**0.6**	
Vertebrates		**371**	**228**	**41 917 814**	**66.0**
Vertebrates (unspecified)		**15**	**8**	**146 135**	**0.2**
Fish	**32**	**23**	**98 157**	**0.1**	
Amphibians & Reptiles [(2)]		**58**	**43**	**30 693**	**>0.1**
Birds	**146**	**86**	**41 339 697**	**65.1**	
Mammals		**120**	**68**	**303 132**	**0.5**

(1) Almost certainly underestimated
(2) Since the CCBR survey was conducted, BRC has compiled a dataset of 49 000 records of amphibians and reptiles.

Table 2.3.11 **Organisations holding taxon based datasets in excess of 1 million records**

1) Number organisations which provided quantitative information on data holdings to the CCBR survey
2) Total number of records held in 1992/93
3) Total number of records (2) expressed as a percentage of the total number of taxa records (63 542 068)

Organisations/ organisational types	**1** **organisations**	**2** **Number of records**	**3** **Number of% of records**
British Trust for Ornithology	**1**	**28 288 655**	**44.4**
County bird clubs	**17**	**9 062 020**	**14.2**
Biological Records Centre	**1**	**6 128 663**	**9.8**
BSBI vice-county recorders	**59**	**5 783 524**	**9.1**
Local records centres	**29**	**5 656 813**	**8.9**
Statutory nature conservation agencies	**5**	**1 998 290**	**3.1**
National biological recording schemes	**23**	**1 628 400**	**2.6**
Museums (excluding local records centres)	**3 1 556 761**	**2.4**	
All others	**33**	**3 438 942**	**5.4**

Table 2.3.12 **Distribution of the main groups of taxon based records by organisation types and principal specialist organisations**

1) Total number of records. Only major organisations and organisational types are listed.
2) In each organisation or organisational type, the number of records for the respective taxonomic group is expressed as a percentage of the total records for the group.
3) Number of quantified datasets for each taxonomic group/organisation.

Organisational types and specialist organisations	1 Number of records held	2 % of total records	3 Number of datasets
Ornithological (total)	**41 339 697**	**-**	**86**
BTO	**28 288 655**	**68.4**	**9**
Bird clubs	**9 062 020**	**21.9**	**24**
Statutory nature cons. agencies	**1 215 000**	**2.9**	**11**
WWT [(1)]			
RSPB [(2)]			
Others	**2 774 022**	**6.8**	**42**
Vascular plants (total)	**13 937 232**	**-**	**129**
BSBI v-c recorders	**5 782 924**	**41.5**	**59**
BRC	**3 533 982**	**25.3**	**3**
Local records centres	**1 980 162**	**14.2**	**32**
Others	**2 639 564**	**18.9**	**35**
Cryptogamic plants	**2 152 700**	**-**	**94**
BRC	**898 151**	**41.7**	**7**
National recording schemes	**648 500**	**30.1**	**4**
Local records centres	**263 410**	**12.2**	**57**
Others	**342 639**	**15.9**	**26**
Insects (excluding moths/butterflies)	**1 930 162**	**-**	**241**
National recording schemes	**941 400**	**48.7**	**17**
BRC	**445 799**	**23.1**	**19**
Local records centres	**407 540**	**21.1**	**167**
Other government departments	**122 110**	**6.3**	**5**
Others	**13 313**	**0.7**	**33**
Moths & butterflies	**1 556 004**	**-**	**44**
BRC	**617 474**	**39.7**	**2**
National recording schemes	**503 000**	**32.3**	**2**
Local records centres	**419 283**	**26.9**	**26**
Others	**16 247**	**1.0**	**14**
Invertebrates [(3)]	**1 928 156**	**-**	
National Rivers Authority	**759 300**	**39.4**	**7**
Statutory nature cons. agencies	**346 991**	**18.0**	**17**
BRC	**232 222**	**12.0**	**8**
Local records centres	**211 175**	**10.9**	**160**
County councils	**105 235**	**5.5**	**11**
County wildlife trusts	**89 338**	**4.6**	**21**
Others	**183 895**	**9.5**	**15**
Vertebrates (excluding birds)	**578 117**	**-**	
Local records centres	**201 739**	**34.9**	**95**
BRC [(4)]	**117 004**	**20.2**	**2**
National voluntary cons. agencies	**54 125**	**9.4**	**9**
Environmental consultants	**54 100**	**9.4**	**3**
County wildlife trusts	**49 735**	**8.6**	**12**
Statutory nature cons. agencies	**25 685**	**4.4**	**7**

Others	**75 729**	**13.1**	**13**

Notes
(1) Information not supplied, estimated to be 1 million records
(2) Information not supplied, estimated to be 0.5 million records
(3) Includes datasets incorporating records of insects
(4) Since the CCBR survey BRC has compiled a dataset of 49 000 records of amphibians and reptiles (not included in these totals).

2.3.22 A majority of the data identified by the survey - 65% - relate to birds, with some 41 million records held by national and local ornithological groups, including 1.9 million held by JNCC. Despite this vast total, it does not include the RSPB's sites and species database or the data holdings of WWT, because neither organisation returned quantified information in the Survey. Some of their data may be replicated, for example, in the returns from BTO and JNCC. The principal national ornithological organisations (BTO, RSPB and WWT) and the statutory conservation agencies (which hold important collections of ornithological data of their own) collaborate on survey, monitoring and research and also with locally based groups, such as bird clubs and local groups of BTO, RSPB and WWT. The Irish Wildbird Conservancy (IWC)) is the main ornithological organisation in the Republic of Ireland. The components of ornithological 'networks' are summarised annually in The Birdwatcher's Handbook (Pemberton 1993). The complexities of these 'networks', and the abundance of data, reflect the strong national enthusiasm for birds: for example, BTO has over 9 000 members (most of whom are active field ornithologists), WWT has 70 000 members and supporters, and RSPB has 850 000 subscribing members. Local bird clubs are estimated to have a total of more than 40 000 members in the UK. Although bird clubs take part in BTO surveys, and therefore some of their data may be replicated in the BTO data holdings, some of the 9 million records known to be held by bird clubs are the result of purely recreational birdwatching ('listing' and 'twitching') rather than structured survey or monitoring. Most of these records are used locally, for example in annual county bird reports which log notable species and unusual sightings, and some are used to contribute to local or national nature conservation and environmental planning. The largest single dataset for birds is the 23 million ringing records held by BTO.

_____BOX_____

Examples of major national ornithological surveys and monitoring, with the organisers and start date or duration

Ringing Scheme - BTO (1) - 1909 onwards
Heronries Census - BTO - 1928 onwards
Nest Record Scheme - BTO (1) - 1939 onwards
Wetland Bird Survey - BTO/WWT/RSPB/JNCC (2) - 1993 onwards (3)
Common Birds Census - BTO (1) - 1962 onwards
Atlas of Breeding Birds - BTO/IWC - 1968-72
Waterways Birds Survey - BTO (1) - 1974 onwards
Wintering Birds Atlas - BTO/IWC - 1981-84
Seabird Monitoring Programme - JNCC/Seabirds Group - 1984 onwards
Constant Effort Sites Scheme - BTO (1) - 1981 onwards
New Breeding Birds Atlas - BTO/IWC/SOC (1) - 1988-91
Breeding Birds Survey - BTO/JNCC/RSPB (2) - 1994 onwards (4)

Notes
(1) Funded in part by the statutory nature conservation agencies
(2) Funded by the cooperating organisations
(3) Combined from:
National Waterfowl Counts - WWT (1) - 1947 to 1993
Birds of Estuaries Enquiry - BTO (1) - 1971 to 1993
(4) Will eventually replace the Common Birds Census

NB The acronmys are explained in the Glossary (page 152)

2.3.23 Vegetation provides the ecological matrix for most other terrestrial and freshwater biota. Although other botanical groups are active, the most important group concerned with vascular plants is BSBI. Through a membership of 2 700 and networks of voluntary local specialists (vice-county recorders) and taxonomic referees, BSBI is a self-contained and highly effective national survey group. Data are held by individual members and are collated by the vice-county recorders. Since BRC was set up in 1964, BSBI has worked in close collaboration with it on all major survey and monitoring projects, with BRC effectively acting as the national databank for BSBI. In most cases, BRC collates records from the vice-county recorders or the organisers of special surveys. At the present time, BSBI is seeking funding for a New Atlas project which will collate data on all species, to update the seminal Atlas of the British Flora published in 1962 (Perring & Walters 1962).

_____BOX_____

National botanical survey, monitoring and database projects, with the organisers and start date or duration

Atlas of the British Flora - BSBI/NC - 1954-62
Red Data Book - (NCC)JNCC (with ITE/RSNC/WWF) - 1977,1983, 1993 onwards
National Vegetation Classification - NCC 1975-1990
Update of BRC botanical database - ITE/(NCC)JNCC - 1980 onwards
BSBI Monitoring Scheme - NCC/ITE/BSBI - 1987-88
Scarce Plants in Britain - JNCC/BSBI/ITE - 1990-92
Atlas/Database of Aquatic Plants - ITE/JNCC/NRA - 1989-94

NB The acronmys are explained in the Glossary (page 152)

2.3.24 The national Biological Records Centre is based at the Institute of Terrestrial Ecology (ITE) at Monks Wood. Its database of 6.2 million records covers over 9 000 taxa. Because of its origins in the BSBI Atlas of the British Flora project and the subsequent close associations with BSBI and British Bryological Society (BBS), BRC holds over 4 million vascular plant and bryophyte records which have resulted in a succession of national atlases (e.g. Perring & Walters 1962, Stewart, Pearman & Preston 1994, Preston & Croft in press, Hill, Preston & Smith 1991, 1992, 1994). Other long standing surveys centred on BRC, which contribute records to the BRC database, include those for mammals (Arnold 1993), amphibians and reptiles (Arnold in press), butterflies (Heath, Pollard & Thomas 1984), dragonflies (Merritt, Moore & Eversham 1994) and non-marine molluscs (Kerney in prep), as well as many other, less well studied groups. The range of BRC recording schemes is summarised by Harding & Sheail (1992) and a list of published atlases is given by Harding (1989).

2.3.25 The statutory nature conservation agencies hold considerable numbers of datasets which contain information on taxa. Some datasets are taxa-oriented, such as the Invertebrate Site Register (ISR) (Ball 1994), but others are the results of biotope surveys, such as those of grasslands or wetland dykes conducted by the former NCC (reports listed by Palmer (1991)) or the surveys of lakes undertaken for the Environment Service of the Department of the Environment for Northern Ireland (DOENI). These biotope survey datasets have been shown to be a rewarding source of data on taxa, for example on scarce plants (Stewart, Pearman & Preston 1994) and aquatic plants (Preston & Croft in press). The Survey under-represents the importance of the statutory agency datasets, many of which are computerised, because few were quantified in their returns. Data on marine taxa held by the agencies are considered in 2.3.37 - 2.3.41. In addition to their own programmes of work on data on taxa, the agencies support the work of BRC through a contract administered by JNCC. Since 1993 there has been a formal agreement between ITE and JNCC over joint responsibility for the BRC database.

2.3.26 Most national biological recording schemes operate in association with BRC, the main exceptions being those for birds (see 2.3.22), lichens and fungi. In the case of the BRC schemes data are held by national or regional scheme organisers until such time as BRC has resources to undertake work on their data; this is especially true for data on insects for which the national schemes hold twice as many records as BRC (Table 2.3.12). With the growing use of personal computers and access to institutional computers,

the volume of computerised data available direct from such schemes (rather than from BRC) is increasing steadily. For example, the British Lichen Society (BLS) database is maintained at Bradford University and is well established with at least 400 000 records of the geographical distribution of taxa (Hawksworth & Seaward 1991, M R D Seaward pers.comm.). Since 1989 the British Mycological Society (BMyS) has established its own database (based at the C.A.B. International Mycological Institute (IMI)), which aims to collect ecological information as well as data on geographical distribution.

2.3.27 Vertebrates (other than birds) are somewhat under-represented in the Survey. This is particularly true of mammals because information on the data holdings of several important surveys was not made available. These include surveys by MAFF, Forestry Authority, Mammal Society, Bristol University, Vincent Wildlife Trust and most local bat and badger groups. All these organisations are known to hold significant quantities of data on selected species of mammals. The Mammal Society is currently preparing an updated list of projects on British mammals; the 1992 edition listed over 140 separate projects ranging from a long term study of one otter family to a national sightings scheme for whales and dolphins with 800 observers. Since the Survey, the BRC database for amphibians and reptiles has been completed, adding some 49 000 records to the total given in Table 2.3.10. Sources of information on freshwater fish are particularly diffuse but are believed to include water companies, water regulating agencies (NRA and the Scottish River Purification Boards), electricity companies (e.g. Turnpenny 1985), angling and sport-fishing groups and research and conservation organisations (see Maitland & Campbell 1992).

2.3.28 The voluntary nature conservation organisations hold considerable amounts of information on taxa, although only a small number of the wildlife trusts (other than those which also function as local records centres) were able to provide quantified information on their data holdings. As in the case of the statutory agencies, biotope survey datasets held by the voluntary organisations are potentially as important as their strictly taxon-based datasets. Many wildlife trusts rely on local experts, such as BSBI vice-county recorders, for the supply of site-based information. The NT and the National Trust for Scotland (NTS) hold large amounts of site-based data, including information on taxa.

2.3.29 Local biological records centres operate to a range of priorities and standards for the collection of data. The extent to which there is collaboration with other local data sources and data holdings (e.g. BSBI vice-county recorders, bird clubs or wildlife trusts) varies considerably. This variability affects the extent to which any given local records centre will act as a main local depository for data or stands independently of other data sources. Museum-based local records centres potentially have the benefit of data resources derived from their own collections, but only 17% of records were specified as coming from this source.

2.3.30 National (and some other) museums and herbaria contain important national collections of preserved specimens. These collections have not been covered by the Survey because, in most cases, data on taxa can be retrieved only by direct examination of the data labels and other information sources associated with the specimens. The information resource of biological collections in museums and herbaria generally in the UK was only obliquely recognised in the Museums Association's report on biological collections (Williams 1987). However, more recently the importance of collections as an information resource has been acknowledged in the UK Biodiversity Action Plan and has been actively promulgated on a global basis in Systematics Agenda 2000. Few UK collections are catalogued in forms which enable the retrieval of conventional biological records of taxa (a notable recent exception is Grayson 1994), but some progress is being made in publishing *metadata* about museum collections through regional initiatives coordinated through FENSCORE. Museums and herbaria are acknowledged as important, but largely inaccessible, sources of mainly historical data. Some national surveys have made extensive use of data from museum collections, especially for taxonomically difficult groups (e.g. pond weeds (Preston & Croft in press) and atomariine beetles (Johnson 1993)). Limited use has been made in the UK of museum collections to assess scales and rates of change in the occurrence of taxa (authors in Hawksworth 1974).

2.3.31 The importance of the published literature and archival material on taxa should not be overlooked, particularly in providing a historical perspective to modern information, for example in biogeographic research and in examining the long term effects of environmental changes (see for example Prendergast & Eversham in press). Despite this, there is a tendency to regard all information to have a finite 'shelf life',

particularly in relation to site selection and site protection. Even in these cases, historical information may provide a valuable guide, in the absence of recent data, to what should be looked for to evaluate a site. Access to published literature sources is facilitated by a range of abstracting and indexing journals and bibliographic publications. NCC developed ENTSCAPE, a computerised bibliography of the literature on British invertebrates published since 1930 (Penny & Key 1994), but it has not been updated since 1989 and it is not accessible outside the conservation agencies. Access to archival material, such as field survey records or personal notebooks (which often can provide more detailed information than databases, publications or museum collection data labels), is notoriously difficult in the virtual absence of effective archival systems. The Natural History Museum (NHM) is a potential archival source, but little use seems to have been made of this facility. BRC, in common with many local records centres, maintains an archive of record cards and other documents to support its computerised database but few, if any, such archives are managed to the accepted standards for curating documentary archives. A similar situation exists in relation to archival material held by university departments. The need for effective management of ecological archives has been considered in recent years as part of NCC's Great Britain Nature Conservation Review Survey and by a working party convened by the Linnean Society, but with no apparent outcome.

2.3.32 Information on the ecological attributes and requirements of taxa, often in anecdotal forms, has traditionally been available in publications such as handbooks, identification guides and atlases, and other publications which synthesise existing knowledge, for example Ellenberg (1988, 1991) for flowering plants and Emmet (1991) for larger moths. Some information of this type now exists in more collated forms (Table 2.3.13). The Ecological Flora Database (Fitter & Peat 1994) provides an excellent model for databases of ecological information, but the amount of detailed information which has been collated for flowering plants cannot be matched by that for any other taxonomic group. The importance of ecological information about taxa is now more widely recognised and increasing effort is being put into collating such data, particularly for use in relation to nature conservation and environmental assessment, and as an aid in research.

Table 2.3.13 **Sources of collated data on the ecological attributes and requirements of taxa**

Key to information:
1 Reference to descriptive account of the collation
2 Brief description
3 Source/availability

Currently available in public domain

Ecological Flora Database
1 Fitter & Peat (1994)
2 Detailed data on the attributes of 1777 native and naturalised and 280 introduced flowering plants
3 Available on line from Bath Information and Data Services

BUGS Entomological Database
1 Sadler, Buckland & Rains (1992)
2 Detailed text on the habitats, biology and distribution of over 5000 insect taxa including Holocene fossil occurrences
3 Available on disk from Department of Prehistory, University of Sheffield

Recorder
1 Ball (1992)
2 PC-based data management package for biological recording containing protection status and habitat information
3 English Nature

Life history and habits of the British Lepidoptera
1 Emmet (1991)

2 2500 taxa - tabulation of life history, status, distribution, habitats, flight times and foodplants
3 Published form (Emmet 1991)

Phytophagous insects/mites on trees
1 Winter (1983)
2 Indexed list of 1400 taxa with host trees
3 Forestry Commission booklet

English Nature - Habitat Fragmentation: Species at risk
1 Kirby (1994)
2 Annotated list of Red Data Book and Nationally Notable terrestrial and freshwater invertebrates, including status, habitat, distribution, recordability, mobility and population structure
3 English Nature Research Report

Currently not available in public domain

Phytophagous Insects Data Bank
1 Ward (1988)
2 Records of 45 000 linkages between phytophagous insect species and plants
3 ITE Environmental Information Centre

Biotopes Occupancy Database
1 Eversham *et al.* (1992)
2 2000 invertebrate taxa classified in CORINE biotopes
3 ITE Biological Records Centre

Invertebrate Site Register (ISR)
1 Ball (1994), Procter & Key (1994)
2 Information management system contain species, threat status, distribution and biology
3 JNCC

ENTSCAPE Invertebrate bibliography
1 Penny & Key (1994)
2 Computerised bibliography of British national and regional literature on non-marine invertebrates from 1930. Keyword index to taxonomic group, geographical area and subject
3 English Nature

Biotopes and land types

2.3.33 CCBR sought to assess the extent of information on biotopes and land types held by the range of organisations covered by the survey. It was not appropriate to repeat reviews of well documented collations of information, such as Phase 1 Habitat surveys in England (Wyatt 1991), the National Vegetation Classification (Rodwell 1991-), the Ancient Woodland Inventory (Spencer & Kirby 1992), the Countryside 1990 survey (Barr *et al*. 1993) and the ITE Land Cover Map (Fuller, Groom & Jones 1994, Barr *et al*. 1993). Also, national 'site registers' such as the BTO's Register of Ornithological Sites (Fuller 1982) and the JNCC's Invertebrate Site Register (Ball 1994) contain biotope information on more than 4000 and 9000 sites respectively. The statutory nature conservation agencies have conducted extensive surveys of coastal and marine biotopes (2.3.36 - 2.3.40). Information about data resulting from large scale landscape monitoring of ESAs by MAFF was not made available to the CCBR survey. Land use and ecological associations were characterised in the Countryside Commission's New Map of England pilot project using a limited range of data mainly from secondary sources (New Map Consortium 1993, Countryside Commission 1994).

2.3.34 Information on 522 datasets concerning biotopes and land types was provided to the CCBR survey by 87 organisations. However, only 303 of these datasets were quantified in the returns to the survey.

From the sample of 522 datasets, it is possible to examine the coverage of biotopes and land types, and to relate this to the types of organisations holding information (Tables 2.3.14).

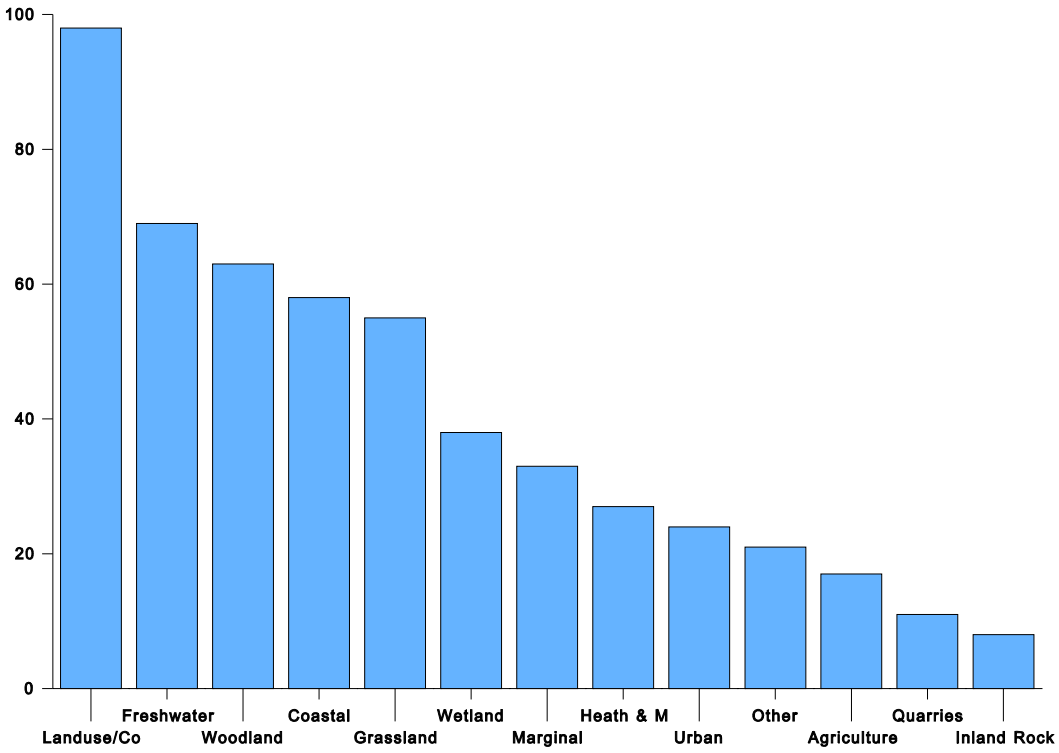
Table 2.3.14 **Summary of biotope and land type data holdings documented in the CCBR survey - Number of organisations responding, number of datasets and number of records quantified in responses**

Organisational type	Number of organisations responding	Number of datasets	Number of records quantified	
Local records centres	20	157	499 102	
Wildlife trusts	14	87	22 481	
Statutory nature cons. agencies	5	66	581 893	
National Park authorities		8	61	123 375
National Rivers Authority regions	7	25	112 648	
Urban wildlife groups	6	23	5 998	
National recording schemes	6	10	20 100	
County councils	5	36	409 125	
Other government departments	5	17	16 850	
National voluntary cons. agencies	5	8	13 975	
Department of the Environment	1	7	654 700	
Research councils	2	3	30 004	
British Trust for Ornithology	1	19	501 000	
Environmental consultancy	1	2	51 000	
Educational establishment		1	1	1 890

2.3.35 The quantified information summarised in Table 2.3.14 covers a wide range of data types, such as the maps and the target notes resulting from Phase 1 Habitat surveys, and the data on land use (at 1 km square definition) in the CIS. These totals provide little real assessment of the scale and complexity of existing data holdings. However, it is possible to estimate the potential national resources of some important types of data in dispersed holdings, based on the sample from the CCBR survey. For example, the results of Phase 1 surveys, as map-based information and target notes, may constitute a resource of recent data on more than 3 million land parcels, mainly in rural areas. Similarly, the information in site files held by wildlife trusts and local records centres can be expected to cover at least 150 000 'sites' which could range from roadside verges to extensive moorlands. It has already been noted that there is some overlap of information on biotopes and land types with information on taxa, such as lists of vascular plants, birds and butterflies, especially in the site files of trusts and records centres and the biotope surveys by the statutory nature conservation agencies.

2.3.36 The biotope and land type surveys summarised in Figure 2.3.2 show a predictable bias to general land use/land cover (18.8%), freshwater (13.2%), woodland (12.0%), coastal/estuarine (11.1%) and grassland (10.5%). The sample is too small to be able to analyse the geographical coverage of local surveys, and the coverage of national or country surveys is already well known (see 2.3.32).

Figure 2.3.2 **Numbers and types of land cover and biotope datasets**



Marine

2.3.37 Data on UK marine biota are being collected and collated through several initiatives, mainly led by the statutory conservation agencies (see 2.3.38). Other governmental organisations, such as NERC, MAFF and NRA, are involved with data on marine biota, as are some universities and biological societies. The Marine Conservation Society (MCS) is the leading voluntary conservation organisation in the UK specifically concerned with the marine environment.

2.3.38 A review of information on the coastal and marine habitats and communities in the UK, and their conservation, was published by MCS (Gubbay 1988) drawing on research undertaken by NCC since 1974. The Coastal Ecology Branch of NCC was established in 1979 and its work has been continued by the JNCC Coastal Conservation Branch since 1991. Part of this work has been the establishment of a Coastal Resource database. Marine conservation issues are covered by the Marine Conservation Branch of JNCC including the Marine Nature Conservation Review (MNCR) which was established in 1987. The MNCR has compiled databases for taxa (e.g. 70 000 records of seaweeds, 7 000 records of marine fishes) and the biotope databases also contain species data. In addition to the UK coverage of the MNCR, DOENI holds data from surveys of the coast and estuaries, and the intertidal and subtidal zones in the province. The eventual outcome of the MNCR will be a series of nearly 20 Theme Reports covering ecological or geographical units of Britain's coast and seas, which will synthesise over 100 MNCR survey reports, occasional reports and published and field data in the MNCR databases. The statutory agencies have commissioned work such as the Directory of the North Sea coastal margin (Doody *et al*. 1993). They have also prepared an Atlas of marine biological surveys in Britain (Mills *et al*. 1993) in digital form for use with the UK Digital Marine Atlas (UKDMAP) (see 2.3.39). JNCC also has a leading role in collating information on seabirds (see for example Walsh *et al.* 1993). JNCC is a partner in BioMar, a collaborative project partly funded by the EU, which is developing protocols for biotope mapping and survey in the marine environment, linked to the European CORINE classification.

2.3.39 NERC has an important role in marine sciences globally (NERC 1993), including some work in UK waters. The British Oceanographic Data Centre (BODC) is based at the Proudman Oceanographic Laboratory. BODC has been responsible for the development of the PC-based United Kingdom Digital Marine Atlas Project (UKDMAP) which contains over 400 charts covering a wide variety of marine themes (mainly physical and administrative), but including some of birds and mammals, and fishery statistics. The Plymouth Marine Laboratory (PML) is the home of the Marine Biological Association of the UK (MBA) and acts as a focal point for many of the marine data holdings of NERC relating to UK waters, including databases such as the Plymouth Marine Fauna Database and the British Marine Fishes Database. The Sir Arthur Hardy Foundation for Ocean Science, also based at PML, is responsible for the Continuous Plankton Recorder Survey (CPR). This is a 60-year time series survey in north Atlantic and European coastal waters, which was rescued through the Foundation, but the MBA's 70-year 'Russell' Cycle time series sampling in the English Channel has been stopped. The Sea Mammal Research Unit conducts survey and monitoring of UK seal populations, and advises government on seals, dolphins and whales (see also 2.3.27). BRC ceased to be involved with biological recording of marine taxa in the mid 1980s, due to lack of resources, but prior to this datasets were compiled and atlases published for marine dinoflagellates (Dodge 1981) and seaweeds (Norton 1985). Subsequently, atlases of the crabs of the north-east Atlantic (Clark 1986) and the marine molluscs of north-west Europe (Seaward 1990, 1993) have been published resulting from marine recording schemes originally associated with BRC.

2.3.40 Information on the data holdings of other organisations was not made available to the CCBR survey. Potential sources with interests in marine biological recording include MAFF and NRA (in particular, data relating to commercial fisheries of all types) and university departments. In the case of universities, data collected in association with NERC community projects such as LOIS (Land Ocean Interaction Study) are likely to be deposited with BODC.

2.3.41 Despite its importance to us as an island nation, and the obvious range of marine biotopes in the coastal waters of the UK, marine biodiversity has been largely neglected in many recent initiatives. The figure of 8 000 taxa of marine organisms occurring within UK waters (less than 10% of the estimated total for terrestrial and freshwater taxa), quoted in *Biodiversity, the UK Action Plan*, is almost certainly a considerable underestimate. Little detailed attention was afforded to marine biodiversity in *Biodiversity, the UK Action Plan* and in *Sustainable Development, the UK Strategy* (Cm.2426), it was considered mainly from the viewpoint of the exploitation of wild populations of commercially important species. The plan for action on biodiversity from the voluntary conservation sector (*Biodiversity Challenge*, RSPB 1993) was even more neglectful of marine biodiversity. Proposals for a Marine Biodiversity Network were formulated at a workshop convened by the MBA in April 1993, to address some of the consequences of this apparent neglect. The workshop was concerned with issues wider than just the UK's own biodiversity and involved representatives of the main academic institutions but, surprisingly, neither MAFF nor MCS were represented. Developments from this workshop are awaited, but further attention was focused on the problems of the UK coastal environment at The Coastline Conference in May 1994.

2.4 SURVEY, SURVEILLANCE AND MONITORING [2]

[2] For definitions of *survey*, *surveillance* and *monitoring* see the Glossary and Hellawell (1991), Baillie (1991) and Rowell (1993)

2.4.1 Although many of the data described in Chapter 2.3 have originated from voluntary sources, distinctions should be made regarding the broad methodological and strategic origins of data, particularly at the national or country scale. Government requirements for data derived from strategically planned surveys, surveillance and monitoring, following the Environment White Paper *This Common Inheritance*, are being reviewed by DOE, particularly for use in future editions of *The UK Environment* report (DOE 1992) and in implementing *Biodiversity, the UK Action Plan*. A brief review has been made by CCBR of the survey, surveillance and monitoring of taxa, biotopes and land types in the UK, elements of which were incorporated in the *Biodiversity, the UK Action Plan*. The practical methods used in surveys, surveillance and monitoring are beyond the scope of the Survey, but the present situation with regard to standards has been reviewed.

2.4.2 Surveys of taxa take many forms, from listing the species of a taxonomic group that occur in a spatial unit such as a defined site, a grid square or a vice-county, to recording, on one occasion, the microsite and substrate occupied by individual specimens of, for example, centipedes. The methods used to acquire such information include transects, quadrats and simple searching as well as a variety of trapping and sampling techniques (especially for invertebrates). There is surprisingly little information on standard survey techniques for any taxonomic group, except birds, and especially for invertebrates although some aspects have been considered by the Freshwater Biological Association (Furse *et al.* 1981), the Field Studies Council (Disney *et al*. 1982, Disney 1987) and the Joint Committee for the Conservation of British Invertebrates (Brooks 1993). Individual national biological recording schemes usually provide instructions to recorders, but few cover the available or preferred methods for survey.

2.4.3 Few methods are designed to provide quantitative data and results capable of replication, and most of these are associated with major national surveys (e.g. the New Atlas of Breeding Birds) or with surveillance or monitoring projects (e.g. Common Bird Census, Butterfly Monitoring Scheme). However, many of the most effective practical methods for rapid surveys (e.g. for site inventories) are known only to experienced field workers and are difficult to document because they involve a complex range of skills. In the virtual absence of accepted methods for surveys of most groups, other than some for birds and vascular plants, the importance of the experience of the surveyor, and the amount of time spent on the survey, together with the time of year, are the most important factors in determining the effectiveness of the survey. Recent guidelines on invertebrate site surveys (Brooks 1993) fail to address the issue of the

competence of surveyors and the resources devoted to surveys. When data for taxa are aggregated from several sources, their use must make allowance for uncertainties about the survey methods and resources used in the acquisition of the data. Well documented datasets record the sampling methods used but these are generally the exception, especially where data originate from volunteers.

2.4.4 Surveys of biotopes and land types, being many fewer in number than surveys of taxa, are more likely to follow one of several standardised methods such as Phase 1 Habitat Survey (England Field Unit 1990) or the NVC (Rodwell 1991-). However, Wyatt *et al.* (1994) have demonstrated the variety of surveys and classifications of land cover and land use currently available in the UK. Some indirect information on basic approaches to biotope surveys was gathered as part of the Survey (Table 2.4.1). The large number of datasets derived from ground survey includes 36 based on Phase 1 Habitat Survey methods, 132 using the RSNC/NCC habitat classification (which is in many ways similar to Phase 1), 239 using *target notes* and 83 using the NVC.

Table 2.4.1 **Methods used to acquire biotope survey data in a sample of 392 datasets**
Some surveys used more than one method

Method	Number of datasets
Ground survey	352
Aerial photographs	88
Map interpretation	13
Satellite images	10
Other [1]	20

[1] Other includes a range of marine and freshwater surveys, literature searches and responses by the public.

2.4.5 Much surveillance and monitoring on a national scale is undertaken already and some baselines have been established by surveys. Recent reviews have examined the extent and potential of many surveillance and monitoring projects. DOE commissioned a review of potential sources of species and habitat statistics which covered 828 projects (Crawford *et al.* 1989, 1990) (Table 2.4.2). It concluded that only 10 projects could be used immediately to supply data appropriate to DOE's needs, although 30 or more projects would be capable of supplying appropriate data with additional work, resources, collation and analysis. Data from trial analyses of seven of the 10 projects (Banwell & Crawford 1992) were used to provide measures of change included in *The UK Environment* report (DOE 1992).

Table 2.4.2 **Wildlife monitoring projects identified in the DOE review (Crawford *et al.* 1989)**

Topic/taxa (Subsets are inset)	Number of projects	
General vegetation	84	
Terrestrial & freshwater habitats	305	
Terrestrial & freshwater plant taxa (total)	89	
Orchids		28
Other vascular plants		47
Invertebrate taxa (total)	105	
Butterflies		29
Vertebrates (total)	162	
Freshwater fish		26
Amphibians and reptiles	11	

Birds		**80**
Mammals		**43**
Marine taxa & habitats	**83**	

2.4.6 Monitoring of birds has undergone considerable modification in the 1990s (see 2.3.3) with the establishment of the Wetland Bird Survey in 1993 and the Breeding Birds Survey in 1994. The latter change resulted from reviews and evaluation of long-running projects such as the Common Birds Census and the Waterways Bird Survey (Baillie 1990, 1991 and Baillie *et al.* 1991). The New Atlas of Breeding Birds (Gibbons *et al.* 1993) demonstrated the value of repeated surveys, using comparable techniques, at intervals of a decade or two, in measuring changes in the distribution and breeding of species (DOE 1994).

2.4.7 BRC has undertaken two surveys of invertebrate monitoring projects. In 1988 some 200 independent butterfly monitoring transects were identified, operating on methods based on those of the Butterfly Monitoring Scheme (BMS), but since 1988 the number of independent transects is known to have increased considerably. Some of these independently operated butterfly monitoring transects have the potential to be brought into the national BMS as it develops and changes, for example to include more 'wider countryside' sites. In 1992 a survey of population monitoring of terrestrial & freshwater invertebrates identified 346 projects (excluding butterflies and the Rothamsted Insect Survey) and described the methods being used (Croucher 1992). Few of these projects were included in the DOE review (Crawford *et al.* 1989). As a development of the BRC recording scheme for dragonflies, a pilot project to monitor the abundance of dragonflies at 15 sites was begun by BRC in 1994.

2.4.8 Few nationwide or site-based projects are sufficiently well established to provide a framework for future monitoring. Table 2.4.3 lists a selection of such projects including 14 from which data have been incorporated in *The UK Environment* (DOE 1992). Monitoring of SSSIs in Great Britain, and Areas of Special Scientific interest (ASSI) in Northern Ireland is carried out by the statutory agencies and summaries of damage and loss are published annually in environmental statistical reports.

2.4.9 Not all of the surveillance and monitoring projects with potential to provide annual results were covered by the DOE review (2.4.5 above). Others operate on longer time scales or have established baselines from which future monitoring could be developed. Examples of the range of period frequency of some of the most important projects are given in Table 2.4.4.

2.4.10 Correlation between data from regular monitoring at selected sites (e.g. Butterfly Monitoring Scheme or Breeding Birds Survey) with longer time series survey/surveillance data (e.g. BRC or BTO national surveys) provides opportunities to examine the results of samples in a wider context. For example, changes in the northern edge of the range of the Hedge Brown butterfly in England and contemporaneous changes in its mean flight period have been examined by relating data from the BRC Butterfly Recording Scheme and the Butterfly Monitoring Scheme (Pollard 1991). Results from a number of long term population monitoring projects were utilised in the interpretation of survey data in *The New Atlas of Breeding Birds in Britain and Ireland 1988-1991* (Gibbons *et al.* 1993).

Table 2.4.3 **Principal nationwide and site based wildlife surveillance and monitoring projects**

Topic	**Extent**
Nationwide (UK or Great Britain)	
Total coverage	
Land Cover Map	**17 land cover types**
National Biological Recording Schemes	**15 000 taxa**
Breeding Bird Atlas	**250 taxa**
Wintering Bird Atlas	**200 taxa**
British Lichen Society Atlas	**1 880 taxa**
Sample coverage	

Countryside Surveys	
1978	**256 1km squares**
1984	**384 1km squares**
1990	**508 1km squares**
National badger survey	**700 1km squares**
Plant Monitoring Scheme (BSBI)	**350 10km sq/1000 tetrads**
Key Squares Survey (BTO)	**350 10km sq/1000 tetrads**
Regional coverage	
Phase 1 Habitat Survey	**89 surveys**
National Parks	**10 parks**
Environmentally Sensitive Areas	**19 ESAs**
Site based	
Environmental Change Network	**9 sites**
Butterfly Monitoring Scheme	**100 sites**
Constant Effort Sites	**90 sites**
Rothamsted Insect Survey (Moths)	**70 sites**
Rothamsted Insect Survey (Aphids)	**24 sites**
Seabird Monitoring Programme	**150 sites**
Common Bird Census	**250 sites**
Waterways Birds Survey	**100 sites**
National Bat Colony Survey	**350 sites**
Wintering wildfowl & waders	**50 taxa**
Nest Record Scheme	**30 000 nests**
National Amphibian Survey	**150 sites**
Invertebrate Sites Register	**10 000 sites**
National Otter Surveys	**7 000 sites**
Red & Grey Squirrels in State Forests	**1 000 10km sq**
Rare Plants in Great Britain	**300 taxa**
Rare Breeding Birds Panel	**100 taxa**

Table 2.4.4 **Examples of the period frequency of national wildlife monitoring**

Annual
Environmental Change Network
Butterfly Monitoring Scheme
Rothamsted Insect Survey (Moths)
Rothamsted Insect Survey (Aphids)
Seabird Monitoring Programme
Constant Effort Sites
Rare Breeding Birds Panel
Wetland Bird Survey
Breeding Birds Survey
Nest Record Scheme
Red & Grey Squirrels in Crown Forests
National Bat Colony Survey

5-10 Years
Countryside Survey (Countryside 1990)
Environmentally Sensitive Areas monitoring
National Otter Surveys

11+ Years
Plant Monitoring Scheme
Breeding Bird Atlases
Rare Plants in Britain

Baseline established
Land Cover Map
Phase 1 Habitat Surveys
National Biological Recording Schemes *
Scarce Plants in Britain
Lichen Mapping Scheme *
Invertebrate Site Register *
Lower Plants Biodiversity Register *
Key Squares Survey
Wintering Birds Atlas
National Badger Survey
National Parks monitoring

*** Surveys containing some time-series data**

2.5 METHODOLOGIES

2.5.1 This section examines the practical methods involved with the collation and management of biological records. It covers five main topic areas: media for recording data, data input methods, the validation of data (this in three sub-topics - spatial, terminological and other), data management standards and computer use.

Media for recording data

2.5.2 The media used by the primary 'field' recorders to store data for their own use, and in particular to transfer their data for collation with those of other recorders, are of five basic types:

* **Unstructured paper forms - such as notebooks and lists;**
* **Collections of specimens and/or graphical records of all kinds;**
* **Data labels, catalogues and local publications;**
* **Structured paper forms - such as recording cards and forms;**
* **Electronic forms - such as portable computers and data-loggers for use when collecting data, and software packages for storing data.**

2.5.3 Traditionally, field naturalists and ecologists have recorded their observations in notebooks or on sheets of paper, including whatever information seemed pertinent at the time. This important form of storage of historical information is almost the only form in which taxa data were stored until the 1950s. Many recorders still prefer to make initial field notes, from which they may transfer information to other media at a later stage, e.g. when identifying voucher specimens collected in the field. Small, portable, audio cassette recorders are used by some recorders in the field, prior to transfer of information to other media, but this is a particularly ephemeral medium for holding information.

2.5.4 The importance of historical information held in unstructured media, such as notebooks, is easily overlooked when considering the need for current information in planning and site and species protection. Such sources may contain information not reproduced in resultant publications, e.g. on the data labels of voucher specimens, or in systematised media (cards or electronic data) derived from them. Some local museums hold important collections of these types of manuscripts, some dating from the nineteenth century or even earlier. At present, there is no unified policy on the acquisition and retention of such sources. The Museums and Galleries Commission (Paine 1992) provides brief guidelines and refers to BS 5454 recommendations for storage and exhibition of archival documents. However, as noted earlier (2.3.31), the retention and curation of ecological archival material in the UK is, at best, haphazard.

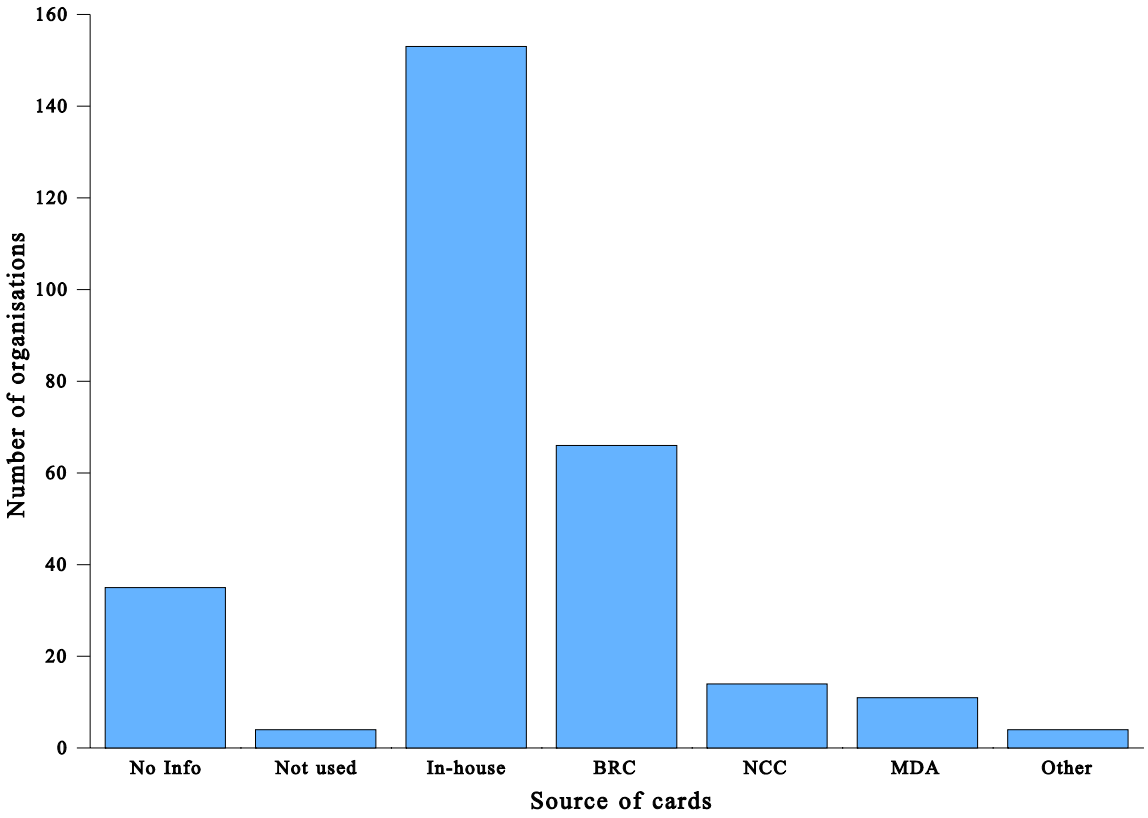
2.5.5 An important benefit of organised biological recording has been that steadily increasing numbers of specialists have compiled their records in more structured forms, using more accessible media. However, the potential resource of historical information, which is so important in measuring temporal changes at all spatial scales, and which often survives only in unstructured media, should not be overlooked. There is one notable example where the enthusiasm for having interesting records accepted by others, has imposed some structure on manuscript sources. This is in the submission of records to the British Birds Rarities Committee, which requires details from the observer's notebook including sketches, weather conditions and salient features seen, thus encouraging a more complete record than might otherwise be made.

2.5.6 Although collections and publications are considered later (2.3.30-32), their relevance as recording media should be noted. For many historical records, they are the only source of information. Collections also provide an implicit opportunity to validate taxonomic information, enabling identifications to be confirmed or corrected. Catalogues of collections, where they exist, rarely provide even the most basic information needed for a biological record and as a general rule, the older the specimen, the more likely it is that the data will be inadequate. Data may be obtained from the labels of collected specimens but often the associated spatial and ecological information on them is imprecise or inadequate in other ways for use as a biological record. Documentation of collections, both public and private, is at an early stage in the UK, but metadata about location and content of collections is being compiled on a voluntary basis (see 2.3.30). Graphical records may become separated from other material and some, such as photographs, can readily suffer from deterioration.

2.5.7 The use of publications as primary sources of data presents the familiar problems of taxonomic reliability and data quality, which can be overcome only by the subjective judgements of successive generations. Access to literature, save for that most widely used, remains difficult despite the growing number of abstracting and bibliographic publications, and databases such as BUGS and ENTSCAPE (see Table 2.3.13). Access to complete runs of the large number of local, often defunct, natural history publications is difficult for all but a privileged few able to use specialist and copyright libraries.

2.5.8 Record cards, forms and sheets are the most familiar recording media for use with organised surveys at both local and national levels. The purpose of these field-recording cards is to prompt for, and structure, the data being recorded and to provide for consistency of records between recorders and different, especially successive, visits or samples. Only 194 organisations in the Survey responded on the use of recording cards but almost 80% (155/194) used cards of some type, four did not use cards, and 35 did not specify if cards were used. Of the 155, 153 used cards of their own design, but 85 also used standard cards, of which 65 used those obtained from BRC (Figure 2.5.1), especially species-list cards, or variants on the BRC format.

Figure 2.5.1 **Sources of recording cards**



2.5.9 Examples of more than 160 different site and species record cards were received. They range from little more than a slip of paper with areas to record species, locality, recorder and date, to highly structured examples with multiple choice tick boxes for controlled terminology. The examples form a continuum in both design and intended use, but can be classified in four major groups covering taxa, site characteristics/taxa, data storage and summary, and a range of special applications.

2.5.10 Taxa recording cards are most frequently used for presence/absence surveys of species belonging to a single individual higher taxon (e.g. flowering plants, butterflies, dragonflies, molluscs). Three basic types are used widely and national versions are supplied by BRC for use by the various national recording schemes, examples of which are figured in Harding (1991). Species list cards list the taxa (usually including code numbers for each taxon) in the appropriate group (e.g. butterflies) known to occur in the area being surveyed (e.g. Britain, Scotland, Suffolk) and recording boxes for at least the grid reference, locality name, recorder name and date. Many include boxes for recording other information, such as abundance, breeding status and biotope. Single species cards are used to record basic information (grid reference, locality name, recorder name and date) for a number of records of one species. This type is used particularly for extracting data from collections, and especially by invertebrate zoologists when identifying preserved specimens (for example at the end of a collecting trip or at the end of the field season). Individual records cards are used to record the basic information on one species at one location on a single date, and also to record a range of additional information (e.g. breeding status, sex, development stage, land owner and site management). They are normally used for uncommon species and unusual records, but are also popular for compiling records from published and manuscript sources.

2.5.11 The proliferation of customised cards for use both nationally and regionally (e.g. for an individual county) has continued since the use of record cards was surveyed in 1983 (Whiteley 1983). Although most include the basic data fields (such as species, grid reference, locality, recorder and date) which are widely accepted as the minimum components of a biological record, their design varies greatly. The lack of both consistency and compatibility of information being sought on such cards is worse when additional data fields are considered. Of those that include such fields, few have used standardised biotope or 'habitat' classifications. This is often for pragmatic reasons: many predate popular classifications such as that used in Phase 1 Habitat surveys, whereas others aim to collect data on the habitats of species at a very fine spatial resolution (see for example the classification used for woodlice (Harding & Sutton 1985)). This lack of overall standardisation restricts the opportunities to compile compatible data, for anything more than the basic data fields on taxa.

2.5.12 An important limitation on the use of record cards as the means of managing biological records is their one dimensional nature. Species list cards can be filed only by spatial units, and not by species, single species cards can be filed only by taxa, only individual record cards can be filed by either spatial units or taxa. Cross-referencing data on record cards is impractical or, with card copying, time consuming and expensive. Record cards were originally introduced by BSBI and BRC as a means of capturing data for automated data manipulation, not as the main data storage medium. Unfortunately, this practice developed a couple of decades before the wide availability of facilities to manipulate data electronically. How to cope with the resultant backlog of unprocessed data, on structured forms such as record cards, is an important short term issue in establishing greater efficiency in the dissemination of biological records.

2.5.13 Some cards are used to record the principal characteristics of a site, but also have some provision to record species. Often these cards are designed to be dual purpose and used as the principal data storage medium. Some site, and habitat, based taxa cards could be classified with this type. Frequently, site description cards lacking any provision for recording taxa are associated with separate taxa recording cards. The Marine Conservation Review Littoral/Sublittoral Site Record Card is the most complex card of this design with tick boxes for a wide variety of habitats, substrates and taxa and other physical data (see Appendix 4). Some site description cards also cover land-use and threats to sites.

2.5.14 Cards designed for storing and summarising taxa data are used, or have been used until recently by most local records centres. Such cards frequently include a printed grid map of the county or district covered by the centre. The data transcribed on to these summary cards are basic and concerned mainly with documenting the distribution of a species in a particular recording area. Usually these cards are used as the manual index to taxa-based surveys, but are being steadily replaced by low cost personal computers to store and to map records using software such as Recorder or DMAP.

2.5.15 A variety of forms designed for specialised surveys (e.g. water quality sampling, bird ringing records, transect samples) were supplied in response to the Survey. Most of these forms include provision for recording basic information, such as species, site details, recorder and date, but many have other, more specialised fields (e.g. sampling methods used, sample/survey period, weather conditions, soil or water chemistry). Some of the most widely used specialist cards are those for biotopes surveys such as the NCC Grassland survey.

2.5.16 Only 5% (10 out of 194) of the respondents to the Survey reported using electronic recording media of any type, of which only 3% (6 out of 194) used hand held computers (Table 2.5.1). These returns almost certainly underestimate the extent to which electronic data capture is being examined, particularly by the larger and better funded organisations such as utility companies, NRA and BTO.

Table 2.5.1 **Use of electronic recording media**

Media type	Number of organisations	Types of organisation
Barcode reader for species codes	1	NRA
Optical character recognition	1	LRC
Optical mark reader for forms	2	NRA & Bird
Hand held computers (5 of which are Psion Organisers)	6	NRA, NP, NSCA, Bird & LRC

Key to abbreviations - see page 10

2.5.17 Some of the recently produced recording sheets have been designed to use Optical Character Recognition (OCR) and Optical Mark Recognition (OMR) technologies for automated data entry into computer systems. Although OMR was used by BRC as early as 1977, it was found at that time to be unreliable. Subsequently these technologies have been developed further and the BTO has used OMR successfully on the recording card for the Nest Record Scheme since 1990 as well as for recording some garden birds.

2.5.18 The main constraints on the use of computers for field data collection are the cost, size (especially the keyboard), battery life and availability of recording software. Small hand held computers currently do not have enough memory to run the database packages used on desktop computers or to hold data for validation. Their main potential advantage, at present, is the saving in time and risk of error in data transcription to computer databases when electronic transfer is used. The power of small hand held machines is increasing rapidly and the slightly larger versions ('palm top' or 'pocket' computers) can accommodate simple word-processing, spreadsheet and database software (such as MS Works) which can be programmed for data capture. The more powerful notebook and sub-notebook computers are capable of running a full range of software and have keyboards approaching normal size. However, most weigh more than 2kg and are not easily held whilst typing in data. Their battery lives may be as short as 2 hours although they can be run from car accessory sockets. They are suitable for only limited forms of field

recording, such as work in a vehicle but because they are now as powerful and have as much data storage as desktop machines, they can be used to run databases or specialist applications.

2.5.19 Pen-based notepad computers are becoming available but still suffer from unreliability and error in transcribing written data in to memory. Moreover, no commercial model is yet sufficiently robust for field use and battery life will continue to place severe temporal restrictions on them.

2.5.20 A notebook computer combined with a portable global positioning system (GPS) has been developed in various versions and has been used, as a backpack, with NAVSTAR satellites to plot geographical locations as longitude and latitude. Aspect and elevation are potentially capable of being determined by such systems. At present the cost is prohibitive for most biological users in the UK but the continued development of hand-held GPS technology for navigation - the principal use at present - is likely to reduce costs in time.

Data input methods

2.5.21 Data entry in to computerised systems can become a severe 'bottleneck' in data management. This occurs most often when a centre is converting to a computerised system from a manual one, where a decision has been made to 'computerise' large collections of old records, or during a special survey where recorders may return records on an annual basis, as in a local flora project.

2.5.22 The standard form of computer data entry is 'key-to-disk' where an operator types the data into a data-entry (e.g. MODES), or database, program (e.g. Recorder) using on-screen forms or data entry windows and the records produced are stored on the computer's hard disk. An increasingly common form of cumulative data entry is the direct incorporation of information supplied by recorders on floppy disks using software provided by a parent organisation (such as BTO).

2.5.23 'Double-keying', in particular the repunching of punched cards as a form of validation persisted into the earlier years of 'screen-based' computing but with the availability of sophisticated screen-based data validation, such as that used in the Recorder program, the practice has largely been abandoned. Only 0.8% (7 out of 897) taxon-based datasets in 5 organisations used double-keyed data validation (Table 2.5.2). In practice, double-keyed data entry was liable error because of the second operator repeating the same mistakes as the first, especially when copying from cards with closely spaced entries, e.g. large species checklists.

Table 2.5.2 **Data input validation methods used**

Totals	Manual	Computer	Dbl key	Archive	Voucher	N=897
ALL	**709**	**345**	**7**	**635**	**366**	**897**
WLT	**71**	**8**	**0**	**83**	**1**	**99**
MUS/LRC	**459**	**240**	**1**	**367**	**287**	**568**
NRS	**21**	**5**	**3**	**17**	**16**	**26**
BIRD	**16**	**12**	**1**	**16**	**0**	**16**
NVCA	**12**	**2**	**0**	**12**	**1**	**12**
COU	**8**	**0**	**0**	**5**	**0**	**11**
SNCA	**35**	**32**	**0**	**42**	**30**	**55**
OGD	**10**	**0**	**0**	**16**	**3**	**17**
NRA	**23**	**2**	**0**	**29**	**23**	**30**
NPA	**11**	**3**	**2**	**12**	**2**	**15**
Other	**3**	**0**	**0**	**3**	**3**	**3**

Key to abbreviations - see page 10

2.5.24 The use of optical scanners for entry of text and graphics into computer systems is now widespread. For some applications, document scanning, i.e. generation of a page image in the computer, is combined with Optical Character Recognition (OCR) to convert scanned documents into text files which can then be further processed using word processing software or imported into databases. This technique has been used by organisations including the British Library and Royal Botanic Gardens (RBG) at Kew for import of catalogues and indexes, whilst some museums have experimented with it as a technique for entering object records from registers.

2.5.25 Reliable OCR software has been relatively expensive until recently and most work has been contracted out to specialist companies. The cost of OCR software and hardware appropriate for PCs is now reducing in line with other PC software (£50-£500). Only one respondent to the Survey had used OCR for data entry.

2.5.26 OCR has the potential to be more widely used as a tool to import large quantities of text data (e.g. lengthy site descriptions or descriptive data on information networks) but these can suffer from a substantial number of read errors, particularly if the source document is not in a common print typeface such as Times, Helvetica or Courier. It could be used to import the simpler forms of biological record, such as species checklist cards for single grid squares or sites, but successful interpretation of written comments will be limited by the quality of the handwriting! However, usage is unlikely to increase substantially in the near future because of the difficulties of translating scanned information from cards and notebooks into the relational data structures used by the most widely used biological recording databases (e.g. Recorder, BIORECS, COBRA, LEVANA). Import routines would need 'intelligent' software to recognise key words or concepts in order to validate the scanned data and apply the necessary internal relational codes. Such developments are still in the research area.

Data Standards & Data Validation

2.5.27 Validation of records presupposes unambiguous and agreed standards of terminology but these do not always exist. Indeed, the naming of the various component units of biological records, whether biological taxa, biotopes, land cover types, people or places, and the management of these names as data, presents a range of problems in accessing information accurately and without ambiguity. Agreed standards in terminology help to avoid the situation where a unit name may have a range of meanings depending on the user, the date of use, and the context in which it was used. It is a common misconception that terminological control in the natural sciences is well developed and, compared with other disciplines, well coordinated.

2.5.28 Validation of biological records covers three main concepts; attribution, controlled terminology, and error trapping. The principal area of concern is the attribution of the species identification. This also applies, but to a lesser extent, to biotope and assemblage (e.g. NVC) identification. Here the problem is further confounded by the lack of general agreement on terminology (see 2.5.29 - 2.5.34, 2.5.43 - 2.5.45). Errors in spatial and geographical references are often detectable without too much difficulty depending on the system employed (see 2.5.68 - 2.5.78). One of the dangers of the widespread use of computer databases is that without due attention to validation they can give spurious authority to doubtful records and perpetuate the acceptance of incorrect records.

Taxa - Standards

2.5.29 Although the flora and fauna of the UK is often quoted as being the best documented in the world, there is no official register of taxa and no readily accessible source of checklists and it is difficult to obtain a precise, up-to-date figure for the number of taxa which occur in the UK and to differentiate between *native* and *non-native* species (e.g. see Cm. 2428, p 29). Traditionally, in the UK, national checklists of selected groups have been compiled by relevant experts, many of whom are amateurs, through national biological societies and in some cases by staff at institutions such as NHM, RBG (Kew and Edinburgh) and IMI. Publication of national lists has been undertaken mainly by voluntary groups such as BSBI, BBS, BMyS, BLS, the Royal Entomological Society of London (RESL) and the Linnean Society or by commercial publishers, either as dedicated checklists (name lists, e.g. Howson 1987, Kent 1992, Inskipp &

Sharrock 1992) or as effective checklists, usually in taxonomic synopses (e.g. Marshall & Haes 1988, Stace 1991, Plant 1994). It is only rarely that either type of checklist includes complete synonymies. Possibly the most complete sets of up-to-date checklists are at the Biological Records Centre and within the Recorder data management package, having been compiled from a wide variety of sources, often from the work of volunteers. The BSBI list of vascular plants, which is linked in to Recorder, is maintained as a database at Leicester University (Kent 1992) and a complete synonymy is being added to it gradually.

2.5.30 Compilation of international taxa checklists, including many UK taxa, is a rapidly developing activity. *Flora Europaea* is held in a computerised form both at RBG Edinburgh and the Botany Department of the University of Reading. International species mapping projects provide a framework for such lists also, and an increasing number of national and international lists in Europe are computerised (Harding 1990). Organisations such as the World Conservation Monitoring Centre (WCMC) and Birdlife maintain international checklists for some taxonomic groups. A world checklist of vascular plants is being developed by International Organisation for Plant Information (IOPI) (Burnett 1994).

2.5.31 Taxonomic study of the flora and fauna of the UK is a continuing process. Relatively few species are described each year as new to science from the UK and these are mainly segregates of previously aggregated taxa, species which were confused with previously known species, species which are in some way cryptic, or species in poorly studied groups. However, systematic revisions often result in the relocation of a species in a different genus or the allocation of a completely different scientific binomial to a species, usually on grounds of precedence established by international nomenclatural commissions. Both forms of apparent instability in nomenclature are often incomprehensible to non-biologists. Achieving nomenclatural stability is a long standing problem acknowledged by all users of organisms (authors in Hawksworth 1991). Arguments have been made frequently for greater stability and standardisation of both scientific names (for administrative purposes, e.g. Council of Europe 1986) and vernacular names (for international communication, e.g. Inskipp & Sharrock 1992). A radical and extreme solution to some of these problems has been proposed through use of coding systems, *biocodes*, which use sequences of letters or numbers, as an alternative to Linnaean binomial nomenclature (Heppell 1990).

2.5.32 At a national scale, the case for the development and greater use of vernacular names, ostensibly to make species nomenclature more accessible to non-specialists, has been proposed on many occasions. Vernacular names based on dialect names have a life of their own and often are colourful and descriptive, but frequently refer to aggregations of species. Formalised vernacular names have come to be accepted for relatively few groups, for example vascular plants, some fungi, vertebrates, macro-Lepidoptera, Odonata and Orthoptera, but even with these groups there is some dispute about 'official' vernacular names. The Wildlife and Countryside Act 1981 has led to the use of some novel vernacular names, for purely administrative purposes, and a large number of unfamiliar and ambiguous vernacular names have been proposed for agricultural and pest invertebrates (Seymour 1989). In general, vernacular names are more liable to ambiguity and uncertainty than scientific names, and vernacular names are better used as adjuncts to, rather than as the basis of, a biological record

2.5.33 The practice of applying codes (sequences of numbers or letters) to signify scientific binomials and other taxonomic units has been adopted widely in the computerisation of data. The use of coding systems has generated much discussion among those involved with data management, with several systems being advocated. Most numerical coding systems have been designed for convenience in maintaining lists of taxa and sorting within single file data structures. There are three basic types of coding systems: sequential, hierarchical and mnemonic. Despite their limitations, codes have a place in the management of digital data. There is growing recognition that codes should be a matter of concern, primarily, to the data manager and the computer, rather than the field biologist and the systematist although in many small organisations it is but a single individual who carries out all these tasks!

_____BOX_____

Examples of the three basic types of taxa coding systems

*** Simple sequential numbers - the first species in the list (whether sequenced alphabetically or systematically) is numbered 1 and subsequent species are numbered 2,3,4,etc. Variants of this type have a break in the numbering, between each genus, of one or more numerals or apply numeric or alphabetic prefixes.**

*** Hierarchical numbering - numbering sequences (usually related to systematic lists) which allow sorting of supra-specific taxa such as genera, families and higher taxa.**

*** Mnemonics - abbreviated species names are used in several systems. A simple, two-letter mnemonic code system has been used by BTO for national surveys (Lack 1986, Gibbons *et al.* 1993) and a 5-letter mnemonic for ringing where the number of species is greater.**

2.5.34 Early BRC recording cards used simple numerical systems but most BRC cards and their related database directories use hierarchical systems. A hierarchical system was used by Maitland (1977) for freshwater fauna and the taxonomic sorting codes in Recorder are arranged hierarchically (Ball 1992). Mnemonics are used in the data entry stage of the Recorder system (Ball 1992) but in combination with a full numerical species code, a *Soundex* search (i.e. *the name sounds like...*), and vernacular names where available. The 339 returns in the Survey containing information about the use of taxonomic coding systems, revealed that the most widely used were those in Recorder (35% of respondents), those developed by BRC (27%) and the Maitland (1977) codes (6%). The Recorder package incorporates other systems, such as BRC and EURING, as *synonyms*.

Taxa - Validation

2.5.35 With easily identified and common species there are few misidentifications and the occasional mis-recording does not fundamentally alter perceptions of the species' distribution or biology. With rare and critical groups mistakes are difficult to eliminate unless, for instance, a record falls well outside the normal geographical or habitat range. The inaccuracies generated by recorders can never be entirely eliminated although in the better organised surveys they are greatly reduced by judicious vetting of the recorders, or the existence of voucher material of various kinds - specimens, photographs, or detailed descriptive material..

2.5.36 In the case of newly segregated species, critical species and difficult groups, identification depends as much on the knowledge and experience of the identifier as on any other factor. Many expert recorders and record coordinators, e.g. some BSBI vice county recorders, have reservations about contributing records for wider use because they fear that their records will be devalued by other poorly attributes or validated records. In practice, the major compilers of taxon-based biological records (BTO, local records centres, national recording schemes and BRC) are fully aware of the sources and limitations of the records that they manage and limit their use accordingly. The weakest link in the system is that few datasets give details of attribution and reliability documented in a way which allows data to be transferred and merged into other data holdings with confidence.

2.5.37 Table 2.5.3 shows the kinds of taxonomic validation procedure applied by different organisation types. About half of all records are checked by staff in the originating organisation and a somewhat lower percentage are also checked by local specialists, the overlap suggests that just over 80% of determinations are checked in this way. The figures for the BSBI and national biological recording schemes emphasise the concentration of local and national expert voluntary recorders in these groups. Generally, about 10% are checked by national taxonomic specialists but this number is weighted heavily towards records of difficult 'critical' species. About half of the records in 36% of all datasets in the analysis were backed by voucher specimens and a third in 19% of datasets checked against collections. These figures seem unduly high but are biased by the large number of datasets held either by museums and local records centres based in museums, *e.g.* 81% of records in 70% of datasets are held by museums, or derived from recording schemes that rely heavily on the determination of collected specimens. Bird organisations do not collect voucher specimens, rarely check identifications against skin collections and are based largely

on sightings. Bird organisations make extensive use of local and national panels for vetting rarity and scarce migrant records. Since they account for at least two thirds of all biological records (see 2.3.22), the number of records nationally supported by vouchers or checked against collections is much lower than the figures suggest. Local records centres apply the widest range of validation techniques to the datasets they manage whilst local wildlife trusts and urban wildlife groups rely principally on their own in-house skills. These differences reflect different working priorities. The records held in wildlife trusts are generated and used internally for local wildlife conservation and their limitations are well understood by those concerned. Local records centres are collators and disseminators of information from a range of sources for a wide range of purposes and therefore require a more sophisticated range of options to maintain a reliable data resource. There is an overlap between these roles since some wildlife trusts run records centres and some records centres exist primarily as suppliers of 'site-based' conservation information to local authority planning departments.

2.5.38 Table 2.5.4 shows how taxonomic validation is applied in different taxonomic groups. The majority of validations for all groups, is carried out by the staff of the organisation responsible for collating the records. There are significant differences between groups where other forms of validation are used.

2.5.39 The use of reference collections for identification and validation is important for lichens, beetles, flies and insects other than lepidoptera. These groups include large numbers of similar species and well-ordered museum collections augment the often poorly illustrated or difficult to obtain, reference works on these groups. Museums contain extensive collections of Lepidoptera but these are consulted less frequently for identification/validation purposes because most lepidopteran records are of butterflies and macro moths, for which more expert recorders and better identification texts are available. Recorders are also less likely to collect Lepidoptera (other than micro-Lepidoptera) for identification than other insect groups. This is true also for plants where only critical species or difficult groups , including lower plants, need to be represented by vouchers or checked against specimen collections. The rather low number of arachnid records checked against collections may reflect the lack of available collections.

2.5.40 Very few vertebrate records are checked against collections and most of the vetting is carried out in-house or by local specialists who may be responsible also for collating records. Only rare or unusual vertebrate records tend to be vetted by local panels or national experts. National experts play a larger role in the vetting of invertebrate records largely through the collation of records by national recording scheme organisers and through the informal regional networks of some national schemes.

2.5.41 Most taxonomic attributions in biotope, site and monitoring datasets (about 80%) are validated in-house (Table 2.5.5) although wildlife trusts, local authorities and museums/records centres claim to refer a considerable percentage of records to local specialists. This may reflect the concentration of records from site surveys carried out by professional contractors such as free-lance botanists. The high percentage of validation by national specialists of datasets held by statutory conservation organisations, the research councils and government departments reflects the concentration in these organisations of site-related records derived from national recording schemes and from contracted experts.

Table 2.5.5 **Methods used by major organisational types for the validation of taxa identifications in land cover and biotope datasets**
Colls: Comparison with collection specimens. In-house: routine identification by in-house staff. Loc. spe.: local taxonomic authority. Nat. spe.: national taxonomic authority.

	Colls	Avg %	In-house	Avg %	Loc. spe	Avg %	Nat.spe	Avg %	N=325
LRC	3.5	11.7	80.2	15.2	80.2	33.7	2.3	15.0	86
Museums	31.3	20.0	100.0	90.6	31.3	10.0	0.0	0.0	16
WLT	1.6	100.0	58.7	85.7	63.5	64.9	7.9	80.2	62
BIRD	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	16
NRS	16.7	33.0	33.3	99.5	66.7	23.3	83.3	54.8	6
NVCA	16.7	4.0	33.3	91.0	0.0	0.0	83.3	80.2	6
COU	0.0	0.0	81.5	65.5	18.5	100.0	0.0	0.0	27
NPA	14.3	2.0	96.4	93.6	28.6	26.0	21.4	9.7	28
NRA	52.4	4.9	100.0	87.8	23.8	12.4	42.9	2.8	21
RC	0.0	0.0	66.7	100.0	0.0	0.0	33.3	100.0	3
SNCA	6.9	33.0	75.9	94.8	27.6	40.8	44.8	32.4	29
OGD/DOE	0.0	0.0	90.9	94.3	40.9	100.0	13.6	65.0	22
OTHER	0.0	0.0	100.0	10.0	100.0	10.0	0.0	0.0	2
EDU	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	1
All	8.6	14.3	79.4	68.6	47.5	45.7	15.0	38.9	
No. datasets	28		259		155		49		325

Key to abbreviations - see page 10

2.5.42 Early versions of the Recorder biological records management package had the facility for automatically checking a species record against the recorder's name and a measure of the recorder's known abilities for that taxonomic group. This has since been dropped partly because this grading of individuals could cause problems under the Data Protection Act and also because a recorder's abilities may change with time giving rise to anomalies within the database. Records in Recorder still indicate whether any record is confirmed, doubtful or known to be erroneous, but this information has to be set manually at the time of data entry.

Land Cover, Biotope and Vegetation Classification - Standards

2.5.43 There have been many attempts to develop land use and vegetation 'classification systems' for the UK; for example systematic accounts of plant communities (Tansley 1911, 1939), the ill-timed National Atlas project (Taylor 1939, 1940) and the Land Utilisation Survey (Stamp 1962). Animal communities also have been considered (Elton 1966, Elton & Miller 1954), but apart from those for birds (Fuller 1982), zoological habitat classifications are fragmentary.

2.5.44 The extent of use of land cover classifications systems, reported to the Survey, is summarised in Table 2.5.6. The most widely used environmental recording package, Recorder (Ball 1992), includes the RSNC/NCC habitat system as its basic land cover/habitat classification, but this is augmented with the National Vegetation Classification (Rodwell 1991-), the Peterken woodland stand types classification (Peterken 1981) and an urban vegetation classification (Shimwell 1983). A classification system for the habitats of birds developed by the BTO (Crick 1992) has been incorporated into some versions of Recorder. The close similarities between the classification used for Phase 1 Habitat Survey and the RSNC/NCC habitat classification (see England Field Unit 1990) mean that over 37% of the respondents' datasets use what is effectively a single classification. The surprisingly low uptake of both the CORINE Biotopes and ITE classifications is probably because they have come into the public domain only recently. The high used of 'in-house systems' raises doubts about the potential for integrating data using such classifications with data using more standard systems.

Table 2.5.6 **Use of land and habitat classifications**
Results from all respondents and datasets (N=421)

Classification system	% use
British Trust for Ornithology	1.4
CORINE Biotopes *	0.2
In-house systems	16.9
ITE land classification	0.7
MAFF Agricultural Land Classification	0.9
National Parks Monitoring Scheme	1.6
Phase 1 Habitat Survey	17.4
National Vegetation Classification *	12.8
Peterken Woodland Standtypes *	2.6
River Corridor Survey	0.5
RIVPACS	2.6
RSNC/NCC habitat classification *	20.2
Shimwell urban classification *	0.7
None	21.4

*** Indicates classification systems used in Recorder**

2.5.45 Land cover definitions in use or applicable to the UK have been reviewed and compared by ITE for DOE (Wyatt *et al.* 1994). The review covered 17 surveys and classifications although it did not include

some of the more specialised and selective classifications currently in use in biological recording (e.g. Crick 1992, Peterken 1981, Shimwell 1983). The review provides a standard framework for the classification of land cover categories of national importance. Software has been developed to make comparisons between selected pairs of classifications. The standard framework provides a system which is exhaustive (i.e. it comprises categories which are inclusive of the total population of land cover classes employed in nationally and internationally important surveys), exclusive (i.e. no category overlaps with any other) and structured as a hierarchy. Although it contains a few anomalies (e.g. parkland is classified under grassland rather than woodland, submerged macrophytes are omitted), it has the advantage over most other classifications of including agricultural use and the built environment as well as semi-natural vegetation types. The interrelationships between land cover classifications can be accessed interactively through CIS.

Land cover & Biotopes - Validation

2.5.46 No direct information was collected on the reliability of biotope attribution by site surveyors. The degree of reliability is likely to vary between survey types and upon the surveyors employed. Quality and standardisation are most likely to vary in general land use surveys including Phase 1 Habitat surveys, mainly through lack of experience and changes in surveyors over the period of the survey. The time of year in which such surveys are carried out is also important for the reliable identification of land types such as unimproved grasslands. With surveys such as Phase 2 Habitat or water quality surveys, the work is likely to be done by more experienced contractors or professional staff.

Spatial & Geographical Referencing - Standards

2.5.47 An accurate and unambiguous description of the precise location of a specimen, sample or site is one of the most valuable elements of any biological record. This apparently simple concept is not as straightforward as it seems. First, terminological problems associated with place names should be noted. Second, it is important to distinguish between three features:

- **the kinds of sampling unit employed in the field;**
- **the actual precision of the act of recording;**
- **the units used to present the record.**

For example, a survey may be made using a 1km grid as the sampling unit, records may be made by means of 6-figure grid references delineating 100 m squares (e.g. TL365636) and the data displayed on a map based on 10km squares. However, spurious accuracy can be introduced when a detailed *centre point* grid reference is cited to cover an entire site such as a wood or some other discrete landscape feature. Finally, as shown by the Survey, a considerable diversity of referencing practices are in current use. Although not all are equally valuable, complete uniformity is neither necessary nor desirable since some kinds of record are best represented by specialized descriptions, e.g. soil fungi.

2.5.48 The use of place names in biological recording is more important for the intelligibility of information to the user rather than for the accurate spatial referencing of information (which is better served by using coordinates or land parcel numbers). For example, it is easier to understand the name *Overhall Grove, Cambridgeshire* than *the wood at centroid TL 338 632* or *OS Parcel No 0023 (TL 3263-3363)*!

2.5.49 Though maps series and gazetteers, the Ordnance Survey (OS) and the Ordnance Survey of Northern Ireland (OSNI) provide a sound basis for the standardisation of place names, which is in the public domain. There are difficulties associated with names, for example, there are only two localities in Britain named Conington, but both are in the present administrative county of Cambridgeshire (thus, unless further qualified, the name Conington, Cambridgeshire is ambiguous) and some names are very common (e.g. the OS Gazetteer records over 120 localities in Britain with the name Castle Hill). There are some inconsistencies in the OS/OSNI data, particularly in the spelling of names in the Celtic languages (Cornish, Gaelic, Irish, Manx and Welsh), in anglicised versions of Celtic names and in perpetuating the misspelling of minor local names (e.g. Cowdell End for Cordell End in Elsworth, Cambridgeshire). Variation in the spelling or form of names between editions of maps is a notable problem in the use of historical information (e.g. the location shown on OS 1 inch sheet 90, published in 1955, as *Hawksheath*, is shown as *Hawks Heath Fm* on the OS Landranger sheet 98, published in 1989). In very sparsely

populated areas, with few features, the positioning of place names on maps can be ambiguous and may vary between editions.

2.5.50 The commonest geographical sampling units used in biological recording are either Ordnance Survey grid squares (from 100m to 10km,) or named sites, but many other formal and informal units are used, for example, *sectors* for coastal bird counts and *river corridor divisions*. The use to be made of a record also affects the kind of referencing system used.

2.5.51 Incompatibility between sampling units is a major barrier to merging or comparing data from different sources. Despite their wide applicability, grid references are not used universally. If they were so adopted a near ideal situation for terrestrial records would be to allocate a six or eight figure grid reference to every biological record, regardless of the type of survey whether site, transect, tetrad, etc., giving a resolution to 100m or 10m. Grid references can be related to all other geographical classifications in use. Such precision, especially 8-figure references (10m square), might not be desirable for particularly sensitive data, for example, in relation to the conservation of rare plants or animals, but this would depend on their accessibility. The only exceptions to grid designations would be historical and marine records. Most of the former are typically given as localities and accurate grid references are rarely available. In the marine situation, comparable longitudinal and latitudinal figures would give similar precision and could be related to grid references in coastal areas.

2.5.52 Early biological recording schemes, such as the *Atlas of the British Flora* project (Perring & Walters 1962) and, initially, BRC itself were concerned chiefly to record presence or absence of species at various levels of grid square, most commonly as 10km squares. Even at the level of county or vice-county, surveys may still record only to 2x2km (tetrad) level. This is acceptable for establishing a baseline of distributional knowledge, for planning further studies, and for making the best use of scarce manpower. This level of spatial location is used in national strategic datasets such as the Department of the Environment's Countryside Information System. Gridded data at these levels can be linked to other gridded datasets such as the ITE Land Cover Map and may thus be associated with much other implied or derivable data. Linking data sets in this way is a valuable technique for examining general distributions and changes in distribution at the national or regional scale.

2.5.53 The emphasis in taxon-based survey continues to be on the presence or absence of species at various scales of gridded unit down to 1km square, rarely to 1ha squares (Ely, pers comm). Grid-based data can be related to site-related records, depending on the relationship between grid resolution and site size, although an accurate and precise grid reference can sometimes allow the site to be inferred and vice versa. It is encouraging, therefore, that 55% of the taxonomic records in 1092 datasets in the survey could be directly related to site (Table 2.5.7). However, since two-thirds of all records held by bird organisations are site-based and since bird records account for 65% of the national total of biological records (see 2.3.5), the survey data are heavily biased by them. Nevertheless, the percentage of site-based records is higher for wildlife trusts (83%) and county planning departments (91%) because their work is strongly site-oriented.

Table 2.5.7 **Summary of the spatial units used in taxa-based datasets**

datasets	10 km	2 km	1 km	100 m	Site	VC	Lat/Long	UTM	Other	No.
ALL	**68.0**	**49.1**	**43.2**	**32.5**	**54.9**	**36.1**	**2.5**	**0.2**	**2.3**	**1092**
WLT	**26.9**	**26.0**	**22.3**	**22.3**	**83.2**	**11.6**	**0.0**	**0.0**	**0.1**	**95**
SNCA	**66.7**	**62.3**	**56.4**	**54.9**	**41.1**	**8.4**	**30.8**	**0.0**	**2.0**	**66**
RC/BRC	**93.2**	**40.2**	**39.5**	**21.7**	**66.2**	**80.0**	**0.0**	**0.0**	**0.0**	**46**
OGD	**10.0**	**5.0**	**5.0**	**0.0**	**85.0**	**0.0**	**0.0**	**0.0**	**5.0**	**20**
NGO	**81.8**	**81.8**	**81.8**	**81.8**	**63.6**	**0.0**	**0.0**	**0.0**	**18.2**	**11**
NRS	**86.7**	**33.7**	**31.0**	**30.4**	**11.4**	**15.4**	**0.9**	**2.6**	**2.9**	**96**
NRA	**67.1**	**67.1**	**67.1**	**67.1**	**18.5**	**11.8**	**8.8**	**0.0**	**5.9**	**34**

NPA	**61.4**	**29.2**	**21.7**	**21.7**	**63.2**	**32.7**	**0.0**	**0.0**	**4.5**	**33**
MUS	**70.1**	**54.0**	**48.2**	**25.0**	**56.3**	**42.4**	**0.3**	**0.0**	**4.2**	**347**
LRC	**80.5**	**57.3**	**52.0**	**35.6**	**54.4**	**51.0**	**0.2**	**0.0**	**0.8**	**601**
L.AUTH	**8.5**	**8.5**	**8.5**	**2.1**	**95.7**	**2.1**	**0.0**	**0.0**	**0.0**	**47**
BIRD	**38.7**	**31.7**	**29.4**	**25.6**	**67.6**	**3.3**	**6.7**	**0.0**	**5.0**	**30**
BSBI	**91.1**	**64.9**	**30.1**	**20.8**	**0.2**	**1.8**	**0.0**	**0.0**	**0.0**	**55**
OTHER	**14.0**	**14.0**	**14.0**	**14.0**	**52.7**	**0.3**	**0.0**	**0.0**	**0.0**	**15**

Key to abbreviations - see page 10

2.5.54 The BSBI network is organised around vice-county recorders who are responsible for collating records from members and passing on relevant survey results to the BSBI and BRC. Their records and national recording schemes are, with some individual exceptions, predominantly grid-based and, for the commoner species, are not directly site-related. Data for rare or otherwise interesting species, e.g. critical groups, taxa under special study, are normally recorded with both detailed grid references and site-related data. However, detailed consideration of returns from BSBI vice-county recorders shows that they are becoming more closely involved with their local biological records networks, e.g. wildlife trusts, records centres and planning departments, indeed, some are actually based in local biological records centres. They are, therefore, moving away from their traditional emphasis on grid-based recording.

2.5.55 The figures in Table 2.5.7 for national recording schemes and the data held by BRC reflect the earlier emphasis on grid-based recording but, for new records, BRC has a clear policy to incorporate both site and grid references wherever possible. This trend is likely to continue since much of the interest of BRC's users is concerned with sites, species of special conservation interest, or with linking distribution data for species with other environmental data using geographical information systems (GIS) (e.g. Ullyett *et al.* 1993). It is true of the majority of other national recording schemes. Some, including BMyS, have been site-oriented from the start and, indeed, this is the traditional practice amongst the older 'natural history' societies: for example, the YNU has such records going back for 125 years although these have been augmented with grid references only in recent decades.

2.5.56 Taxon-based datasets held by the statutory conservation agencies and the NRA regions are more often related to grid references than to site boundaries. In the case of the NRA regions, many of their surveys are carried out on river corridors, sample points, or catchments and in the statutory agencies, taxon datasets include marine benthos sample points, seabird-at-sea sightings, and other similar datasets which may use count areas, latitude and longitude, or other sampling units rather than point grid references. If such schemes are to contribute to a national dataset their data will need to be convertible to grid referencing. This is possible, indirectly, with data in some of the computerised datasets, *e.g.* Invertebrate Site Register, Marine Conservation Review Database and bird databases, which can be extracted or correlated with that held in a GIS using a variety of geographical units including boundary information.

2.5.57 The aggregated returns for 406 biotope, site and monitoring datasets which included details of the spatial units used are shown in Table 2.5.8. The total is strongly biased because of the large number of datasets held by two types of organisation, namely, 125 held by local records centres/museums and 96 by wildlife trusts/urban wildlife groups.

Table 2.5.8 **Summary of spatial units used in biotope and land type datasets**

datasets	10 km	2 km	1 km	100 m	Site	Parcel	Admin.	Lat/Lo	Other	No.
ALL	**47.4**	**44.1**	**43.8**	**34.1**	**58.9**	**13.1**	**9.1**	**4.4**	**7.0**	**406**
WLT	**55.4**	**47.9**	**47.9**	**40.9**	**40.7**	**6.6**	**4.1**	**0.0**	**7.0**	**96**
SNCA	**83.3**	**78.6**	**78.6**	**51.2**	**48.8**	**38.1**	**40.5**	**19.0**	**14.0**	**42**
RC	**66.7**	**66.7**	**66.7**	**66.7**	**33.3**	**0.0**	**0.0**	**0.0**	**0.0**	**3**

OGD	**41.7**	**33.3**	**33.3**	**16.7**	**50.0**	**0.0**	**0.0**	**8.3**	**0.0**	**12**
BTO	**0.0**	**0.0**	**0.0**	**0.0**	**84.2**	**0.0**	**0.0**	**0.0**	**16.0**	**19**
NGO	**0.0**	**0.0**	**0.0**	**0.0**	**50.0**	**25.0**	**25.0**	**0.0**	**0.0**	**4**
NRS	**95.7**	**74.3**	**72.9**	**72.9**	**14.0**	**12.9**	**0.0**	**0.0**	**4.0**	**7**
NRA	**30.9**	**30.9**	**30.9**	**35.5**	**41.8**	**0.0**	**0.0**	**4.5**	**18.0**	**0**
NPA	**43.2**	**42.3**	**39.4**	**32.9**	**72.6**	**3.3**	**22.6**	**0.0**	**16.0**	**31**
MUS & LRC	**41.1**	**40.2**	**40.2**	**25.6**	**66.0**	**15.2**	**0.8**	**0.8**	**2.0**	**125**
COU	**40.3**	**40.3**	**40.3**	**40.6**	**91.3**	**17.5**	**17.5**	**17.5**	**3.0**	**40**
OTHER	**33.3**	**33.3**	**33.3**	**33.3**	**33.3**	**0.0**	**0.0**	**0.0**	**0**	**3**

Key to abbreviations - see page 10

2.5.58 In Table 2.5.8, the percentage data recorded at resolutions of 1km, 2km and 10km are derived mainly from figures available from more detailed grid references and do not imply a primary interest on the part of the collecting organisation in data recorded to these levels. In the case of national recording schemes there is a definite emphasis for site/biotope data to be associated with larger grid units because of the use of the traditional 10km grid by surveys. Even so, 95% of national recording scheme site/biotope records are based on either six figure grid references, or named sites.

2.5.59 Although most site/habitat data are held either in order of site, or site and grid reference, e.g. keyed to 10km index maps, records normally include reference to county, district, parish and vice-county. Retrieval by administrative unit is important for those data centres serving local planning departments where, for example, maps from Phase I Habitat surveys, and target note data, may be supplied to local district authorities. A growing number of county and district authorities have digitised these maps for use in computer mapping and GIS (see 2.5.143-145). In the Survey, county planning departments held the highest percentage of data attributable to site (91%, mostly derived from local wildlife trusts and record centres) but with a substantial percentage (17.5%) linked to land-parcels and administrative boundaries (17.5%).

2.5.60 Some of the variation between types of organisation arises from the types of site/biotope data they hold. The statutory conservation agencies and research councils hold several datasets collected in relation to surveys which may be grid based, for example,sample-based land-use surveys such as the Countryside Survey 1990 (CS 1990) (Barr *et al.* 1993). Local records centres and wildlife trusts tend to hold information on named sites, such as sites of conservation interest, woodland surveys, Phase II grassland surveys, Phase I Habitat surveys and other surveys based on land- parcels. The spatial units used at country and regional level, or local levels differ. In the former, the main interests are in the correlation of biotope or taxon data with other kinds of data such as physical environmental data using GIS through the exchange or merging of data with those of other organisations. Gridded data are the most convenient for correlation because there are no multiple or overlapping boundaries to take account of. Grid-based aggregation of data is the most obvious way to collect a consistent dataset across a wide geographic area using limited manpower resources. However, a dataset can be based on administrative units, e.g. CORINE biotopes.At the local level the most important aspect is whether the spatial precision of the data is great enough to be used for conservation and planning purposes. At this level, such as a county, taxa surveys are usually carried out using 2x2km squares (tetrads) or 1km squares to produce important baseline distribution data which can be compared directly with repeat surveys, or with grid-based surveys in other areas. However, all local planning is related to land-parcels and administrative units. Planning enquiries, large or small are concerned only with the sites affected by the planning proposal and evidence needs to be available in this form. In the case of large sites, of which only a part may be affected by a development (such as a road or rail link passing through a large wood), information may be needed at sub-site or compartment level.

2.5.61 Local authorities may need to combine conservation data with social information, such as census data or other information held by postcode, ward or other 'non-compatible' units. This need has also been recognised by English Nature in its Natural Areas programme.

2.5.62 County wildlife trusts and local authority ecologists are principally concerned with discrete sites or protected areas. As more local records centres become part of, or act as agencies to, local authority planning departments, they have had to concentrate increasingly on site-based recording. Unlike

taxon-based recording, which is primarily carried out by teams of volunteers, site and biotope-based survey is more frequently carried out by contracted teams or, reactively, by in-house staff.

2.5.63 Designation of the boundaries of sites for different statutory and non statutory conservation status and the use of site names for favoured recording areas by local naturalists may not coincide, nor may they coincide with Ordnance Survey land parcels. Thus, even with SSSIs and favourite local recording areas (*honeypot sites*), data from different sources cannot always be correlated with, or be accurate to, the level needed for planning. A common solution to this problem is the adoption of a locally agreed list of sites, normally held by the local records centre, wildlife trust or planning department. These lists may serve a useful purpose for other requirements, but correlation with other sources of information such as landscape conservation boundaries, NRA river corridor and catchment areas, or land ownership, are not always reliable.

2.5.64 A potential disadvantage of site-based recording is that unless individual observations are accompanied by accurate grid references, they can only be related to the whole site, which may be large and, possibly, with ill-defined boundaries. The Recorder database, used in more than 120 biological recording organisations, allows the creation of any number of hierarchically related subsites which can be applied from the regional to the microsite level but this is not an effective procedure with sites which overlap. JNCC is recording the relationship of defined sites to other sites in several databases using concepts of adjoining, overlapping and enclosing boundaries, but manipulation of spatial data in text-based (rather than GIS) databases is arduous and unreliable.

2.5.65 A long-term solution to correlating site-based data could lie with the development of GIS. Their use has grown significantly in the last three years, especially in the statutory and local government sectors. Only 5.5% (18 out of 330) of respondents to the questionnaire claimed to use GIS in relation to biological records and, nationally, from indirect knowledge of other organisations (who did not reply or were not directly approached) this (50 out of over 1000) reflects the active use of GIS for the same purpose.

2.5.66 The limiting factor on the use of GIS is the availability of digital data on map units with the spatial precision required for the records employed. The OS has plans to provide large-scale digital maps but, at present, the cost implications are likely to make their use prohibitively expensive save for major national organisations or commercial organisations. A potentially more affordable alternative but only with a resolution at the 1km level, is the CIS which will become available shortly. Records, at this level, can be both compared and incorporated into CIS. CIS and, to a far greater degree of precision, GIS, would permit better retrieval of data in relation to differing site and administrative boundaries than can be achieved with text-based databases and manual map-based systems. They offer an efficient method for correlating biological data with other, potentially, boundaried datasets such as archaeological sites, monument records, and sites of geological interest (SSSIs and regionally important geological sites (RIGS)). There is also considerable potential for sites to be correlated with environmental data including geological base maps, soils and climate records.

2.5.67 Despite the wide uptake of GIS by local authorities and in higher education, even the technology available at present remains too expensive for the majority of organisations involved in biological recording. In the long run it is not so much the cost of hardware and availability of suitable software that is a block to development but the availability of base maps and national datasets. In particular, most local records centres and wildlife trusts cannot afford access to the available digitised boundaries of administrative areas, land parcels and statutory designated sites or regions (SSSIs, AONBs, ESAs, National Parks).

Spatial & Geographical referencing - Validation

2.5.68 Much of the geographical location data associated with biological records is derived from the grid reference or site name. Commonly derived geographic classes are parish, district, county, region and special conservation area (AONB, ESA, National Park, etc.). Recording cards may have boxes for the field recorder to enter this information or it may be added later by data centre staff. Some biological recording programmes use related tables to link administrative units (parish/community council, district,

county/region) to provide context sensitive popups for validation during data entry. Grid reference integrity is an important requirement when exchanging data.

2.5.69 The two most frequent errors checked for are misread or transposed grid references and misspelling or misapplication of site names. Tables 2.5.9 and 2.5.10 show the degree to which checks for these errors are carried out in different organisational types and on different types of dataset.

Table 2.5.9 **Methods used by major organisational types for the validation of spatial information in taxon-based datasets**
Gazetteer: use of a general gazetteer. List: check against list of delineated sites. OS chk: check of Ordnance Survey grid references.

	Gazetteer	Avg %	List	Avg %	OS chk	Avg %	Other	Avg %	N=738
LRC	62.3	95.6	41.6	99.7	84.4	94.2	1.4	94.0	358
Mus	84.5	40.3	2.4	100.0	85.7	58.8	10.7	100.0	84
WLT	3.6	47.7	64.3	93.5	40.5	91.9	0.0	0.0	84
BIRD	50.0	95.0	25.0	51.0	37.5	95.0	12.5	100.0	16
NRS	63.6	45.1	9.1	65.0	86.4	81.6	9.1	55.0	22
CNHS	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	1
NVCA	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	2
COU	25.0	100.0	75.0	100.0	50.0	100.0	0.0	0.0	4
NPA	0.0	0.0	81.3	77.3	37.5	57.5	18.8	83.3	16
NRA	65.4	84.7	69.2	85.3	100.0	93.1	7.7	90.0	26
BRC	81.6	84.5	0.0	0.0	100.0	98.2	100.0	98.2	38
RC	0.0	0.0	0.0	0.0	25.0	100.0	75.0	100.0	4
SNCA	5.7	66.7	45.3	96.7	37.7	83.5	20.8	100.0	53
DOE/OGD	0.0	0.0	33.3	94.4	51.9	96.4	18.5	100.0	27
OTHER	0.0	0.0	100.0	10.0	100.0	70.0	0.0	0.0	3
All	50.5	81.1	38.2	93.2	74.0	88.3	10.8	96.8	
No. datasets	373		282		546		80		738

Key to abbreviations - see page 10

Table 2.5.10 **Methods used for the validation of spatial information in taxon-based datasets, arranged by principle taxonomic groups**
Gazetteer: use of a general gazetteer. List: check against list of delineated sites. OS chk: check of Ordnance Survey grid references.

	Gazetteer	Avg %	List	Avg %	OS chk	Avg %	Other	Avg %	N=738

Not specified	**0.0**	**0.0**	**100.0**	**50.0**	**100.0**	**50.0**	**0.0**	**0.0**	**1**
Gen. Inverts	**50.7**	**85.1**	**50.0**	**97.3**	**72.8**	**89.9**	**7.4**	**99.0**	**136**
Insects	**65.7**	**84.8**	**42.5**	**97.9**	**83.6**	**92.6**	**9.7**	**100.0**	**134**
Beetles	**50.0**	**73.9**	**26.7**	**97.5**	**86.7**	**87.4**	**16.7**	**98.0**	**30**
Flies	**66.7**	**87.6**	**33.3**	**91.3**	**87.5**	**88.0**	**25.0**	**88.0**	**24**
Lepidoptera	**53.8**	**72.8**	**28.2**	**88.2**	**74.4**	**79.8**	**12.8**	**99.0**	**39**
Arachnids	**56.3**	**93.3**	**43.8**	**100.0**	**71.9**	**93.0**	**0.0**	**0.0**	**32**
Molluscs	**34.8**	**70.0**	**34.8**	**97.5**	**69.6**	**82.5**	**13.0**	**82.5**	**23**
Gen Verts	**16.7**	**70.0**	**58.3**	**100.0**	**50.0**	**88.3**	**8.3**	**88.3**	**12**
Fish	**35.0**	**82.9**	**50.0**	**73.0**	**75.0**	**78.3**	**5.0**	**90.0**	**20**
Birds	**44.1**	**78.0**	**29.4**	**77.8**	**50.0**	**81.3**	**11.8**	**93.8**	**68**
Amph. & Rept.	**38.5**	**85.3**	**17.9**	**100.0**	**76.9**	**93.9**	**5.1**	**100.0**	**39**
Mammals	**35.1**	**74.5**	**13.5**	**84.0**	**75.7**	**89.1**	**21.6**	**98.8**	**37**
Fungi & Lichs	**68.0**	**80.0**	**60.0**	**100.0**	**64.0**	**91.9**	**8.0**	**100.0**	**25**
Lower Plants	**46.2**	**72.4**	**40.4**	**96.4**	**73.1**	**86.8**	**15.4**	**90.0**	**52**
Higher Plants	**45.5**	**74.6**	**31.8**	**93.5**	**78.8**	**85.9**	**6.1**	**100.0**	**66**
All taxa (totals)	**50.5**	**81.1**	**38.2**	**94.2**	**74.0**	**88.3**	**10.8**	**96.8**	**738**

Key to abbreviations - see page 10

2.5.70 Three quarters of all datasets have their grid references checked but the spread is not uniform, for instance, less than half of bird organisation, wildlife trust and statutory nature conservation agency datasets are checked in this way. Some organisations including local authority planning departments and central government departments provided consistently low figures for grid reference checks (about 50%) but it is possible that they were using other techniques such as correlation in a GIS to check data integrity. Organisations using a high percentage of contracted survey appear to rely, without further verification, on the supplier of data to supply correct information.

2.5.71 Organisations which use Recorder get a degree of basic grid reference checking for format and possible transposition of eastings and northings. Local data centres (trusts, record centres and planning departments) are also likely to pick up gross errors if data is mapped either as species distribution maps, or site/biotope distribution maps.

2.5.72 Error trapping and validation of site names include spelling and format checks and checking against either national or local gazetteers. The use of site names without associated grid references, common in published works and with museum specimens, causes problems when relating data to other spatial units particularly when sites have several alternative names or names that refer to various overlapping administrative or conservation boundaries.

2.5.73 All organisations contacted carried out some form of site validation against either a gazetteer or a locally agreed set of sites and most used both national gazetteers and agreed site lists. Museums, BRC and NRA all check a high percentage of names associated with taxonomic datasets against national gazetteers which reflects the large geographic areas these organisations have to cover. Local authorities, national parks and wildlife trusts tend to use local gazetteers or agreed site lists to check taxonomic data which is in-line with the predominantly local, site-based nature of their work.

2.5.74 With site and biotope datasets, names are nearly three times more likely to be checked against local site lists (47%) than national gazetteers (19%) although both may be used, e.g. local authorities check 57% by gazetteer and 77% by local list. Bird organisations check surprisingly few of their site names and wildlife trusts appear to check names less for site datasets than for taxon-based ones.

2.5.75 Grid reference checks are used more frequently than name checks for site/habitat surveys in all organisations except national voluntary conservation organisations - the two respondents being predominantly site oriented. This could be related to the number of Phase I-style, land-cover surveys in which target notes may not be related to named locations. It was not clear from the Survey returns whether organisations using databases such as Recorder, BIORECS and COBRA, with internal site dictionaries, regard this as site name-validation in their returns.

2.5.76 Other methods of geographical validation reported were checks for sites marked on OS maps and checks against site boundaries held on other paper maps or GIS (2 organisations only). One of these, collecting marine data, used Decca navigation and latitude/longitude to pinpoint record sites and the other carried out ground-truthing surveys for a percentage of sites to check geographical and habitat data.
2.5.77 There are some powerful packages for PCs and work stations which allow automated entry of grid references directly from a map by pointing and clicking with a mouse. This ensures that grid references are accurate and can link records to multiple or overlapping site boundaries. The use of maps during data entry also allows for easy visual checking of names and grid references against formal and 'fuzzy' geographic entities such as the Mendips or the Flow Country .

2.5.78 Table 2.5.11 shows the returns from 191 datasets which provided information on validation of data by *ground-truthing*. The sample is biased towards certain organisational types but illustrate some trends. Wildlife trusts and museum record centres rely extensively on local expertise to validate this information, presumably among the people contracted to carry out surveys. The large number (25/32 datasets) ground-truthed by local authorities and the national park authorities reflects a growing use of aerial photography for land cover and landscape surveys. Ground-truthing in statutory and research council organisations reflects the use of satellite imagery in special projects within these organisations.

Table 2.5.11 **Methods used by major organisational types for the validation of other information in land cover and biotope datasets**

method	Ground truth		Local experts		National experts		Other		Total no datasets
	%	No.	%	No.	%	No.	%	No.	
WT	17.0	9	69.8	37	20.8	11	3.8	2	53
LRC	2.4	1	97.6	41	0.0	0	0.0	0	42
MUS	83.3	5	16.7	1	16.7	1	16.7	1	6
NRS	0.0	0	100.0	1	100.0	1	0.0	0	1
BTO	0.0	0	0.0	0	100.0	1	0.0	0	1
NVCA	100.0	2	50.0	1	50.0	1	0.0	0	2
RC	66.7	2	0.0	0	33.3	1	0.0	0	3
COU	78.1	25	21.9	7	0.0	0	3.1	1	32
NRA	0.0	0	40.0	2	40.0	2	60.0	3	5
NPA	75.0	9	75.0	9	16.7	2	0.0	0	12
SNCA	30.4	7	65.2	15	43.5	10	0.0	0	23
OGD/DOE	0.0	0	75.0	6	25.0	2	0.0	0	8
EDU	100.0	1	100.0	1	0.0	0	0.0	0	1
OTHER	100.0	2	100.0	2	0.0	0	0.0	0	2
All	33.0	63	64.4	123	16.8	32	3.7	7	191

Key to abbreviations - see page 10

2.5.79 The nature of biological recording has changed significantly during this century (Berry 1988). The two most important developments have been the move to organised recording projects and the availability of computers to process large quantities of data. These developments have brought about changes in both the data collected and the way they are structured for storage and retrieval.

2.5.80 County floras of the nineteenth and early years of the twentieth century and works such as Taylor's *Monograph of the Land & Freshwater Mollusca of the British Isles* (Taylor 1894-1921) tended to be the consuming work of highly motivated individuals in correspondence with a relatively small number of other contributing individuals. This is a tradition which in some ways continues with national recording schemes for the less popular invertebrate groups. Despite the introduction of innovations such as the vice-county unit for recording (Watson 1859, Praeger 1901), the work tended to be anecdotal in nature with little emphasis on standardised or quantitative coverage.

2.5.81 The interest in organised recording of the distribution of taxa and habitats that has been so much a feature of the last thirty years (e.g. since the publication of the *Atlas of the British Flora* - Perring & Walters 1962) has encouraged greater emphasis on the structure and standardisation of recorded data.

2.5.82 Surveys such as the national recording schemes set up in the wake of the *Atlas* require only a very basic data structure covering the classic *What? Where? When? Who?* combination of fields. Data other than the species name may be very limited particularly with distribution surveys based on 1km or larger grid squares. Managing the resultant data is relatively simple, although before the advent of personal computers this was still time consuming. The first edition of the *Atlas of the Lichens of the British Isles* (Seaward & Hitch 1982) took two years to prepare by hand whereas the second edition took a mere three weeks with the aid of a computer (Hawksworth & Seaward 1990). It should not be overlooked that there is a substantial cost in time and money to enter data into a computer to achieve such benefits!

2.5.83 The greater pressures on the use of the countryside in the latter part of the twentieth century, coupled with a wider appreciation of the need to document this diminishing resource, has required the development of new ways of recording and ways in which less skilled individuals can be involved. The rate of change also demands survey techniques directed at gathering data over a shorter period of time than the average, of about 10 years, taken to compile a county flora. This emphasis is itself presently undergoing change as there is a growing appreciation of the need for monitoring.

2.5.84 Recorded information has, therefore, steadily increased in volume, become more complex and requires ever greater levels of structure and standardisation. The involvement of increasing numbers of individuals, the extension of recording to cover more detailed levels of habitat classification and monitoring, and the consequential increase in complexity of the data imposed the need for more sophisticated methods of storing, handling and processing data to enable it to be interpreted. Fortunately these needs have coincided with the development of computers able to handle the data and could never have come about without them.

2.5.85 Computer databases can be used to edit, copy or retrieve information quickly and efficiently and also provide a more compact storage medium than paper. Management and processing of the data, however, depends crucially on its structure and the development of software programmes to control the required manipulations. There are, also, a number of new problems introduced by computers. Some, such as the ergonomic problems of using visual display units (VDUs) and keyboards will disappear with general advances in hardware design. In the case of software, there are still restrictions on the way that data can be manipulated which await applications better able to reflect the data models and data types associated with environmental records. The solution of other problems, such as the ownership of data in electronic circulation or the maintenance of quality in distributed databases require legal and procedural developments within the biological recording community itself.

The Use of Computers in Biological Recording

2.5.86 Some organisations, such as BRC, have been using computers since the days of punched cards in the 1960s. In the 1970s a few museums and records centres were able to take advantage of local authority

or university mainframe computers including the Hancock Museum in 1975 and the West Yorkshire Data Bank (now the West Yorkshire Ecological Advisory and Information Service) in 1977. The greatest increase in the number of organisations using computers has occurred since the development of PCs in the 1980's.

2.5.87 A survey of local records centres, museums and wildlife trusts carried out for NFBR in 1985 (Copp 1985) found 9 museums and records centres using mainframe computers and 28 museums, records centres and wildlife trusts using microcomputers of 12 different makes. The numbers included 12 wildlife trusts using Comart microcomputers and a simple site recording database, both supplied by the then Royal Society for Nature Conservation (RSNC) with grant aid from NCC and BP. The RSNC project was an early attempt to introduce standardisation and shared development effort into the use of computers by wildlife trusts but was overtaken by the speed of change in computers and software and many trusts went on to develop and follow their own strategies.

2.5.88 Table 2.5.12 shows that by 1992/93, 73% of Survey respondents used computers for some aspect of biological recording. In 19% of these organisations, computers were fully integrated into their work and in a further 34% they were important for some uses.

Table 2.5.12 **Computer use in biological recording**

Numbers are number of organisations responding - total Number of organisations answering this question = 221 of 355 returns (62%)

Are computers used for biological records:	Yes	No
	162	59

Frequency of use:	Not stated	Minimal	Infrequent	Frequent	Very frequent
	79	15	20	57	50

Extent of use:	Not stated	Isolated use	Important for some uses	Fully integrated
	78	24	75	43

Computers used:	Not stated	PC	LAN	Mini	Mainframe	Agency
	62	138	32	13	18	8

PC = stand-alone personal computer
LAN = Local area network

2.5.89 Comparison with the same organisations as in the 1985 NFBR Survey shows an increase from 37 organisations to 59 (Table 2.5.13), an increase of around 8%. While this indicates that many individuals had become aware of the value of computers to their work, the real change has been in the machines and software used. In 1985, few of the declared users were doing more than experimenting with technology and there was no reliable biological recording software capable of managing large datasets. By 1992/3 the

majority of users had powerful machines and were running advanced software such as the Recorder database.

Table 2.5.13 **Comparison of numbers of local records centres, wildlife trusts and museums using computers (Sources: 1985 (Copp 1985), 1992/93 - CCBR Survey)**

Year of Survey	Number of responses	% of orgs contacted	Number using computers	% of respondents
1985	75	56%	37	49%
1992/3	103	74%	59	57%

Data Management

2.5.90 Details of data management were available from the Survey for 1094 taxon based datasets (Table 2..5.14) and 458 biotope and land-type datasets (Table 2.5.15). About 42% of taxon-based datasets are fully computerised and c. 70% are managed manually, with some overlap in organisations that run dual systems. Fewer biotope and land-type datasets are fully computerised than taxon records but 29.5% of biotope and land type records are kept in summary form on computer compared with only 6.5% for taxon based records. This is, presumably because most biotope and land-type records vary in format and include long text descriptions, maps and photographs whereas species records generally consist of a few fields of concise information (e.g. taxon, identifier, date, grid reference). The original paper records for biotope and land-type datasets are more frequently kept as archives than taxon based datasets, possibly because they include a greater variety of items of long-term interest (e.g. maps and photographs) which are not readily computerised and may therefore have to be maintained as a combination working file and archive. Records for taxon-based datasets are also frequently transcribed from originals which may be returned to the originator.

Table 2.5.14 **Data management in taxon-based datasets**

Organisation type	Number of datasets	Fully computerised records		Summary records on computer		Manual system		Archive of originals	
		No.	%	No.	%	No.	%	No.	%
WLT	101	21	20.8	6	5.9	69	68.3	37	36.6
LRC	396	168	42.4	5	1.3	300	75.8	72	18.2
MUS	169	109	64.5	0	0.0	132	78.1	41	24.3
NRS	24	12	50.0	1	4.2	19	79.2	6	25.0
BSBI	79	13	16.5	4	5.1	72	91.1	5	6.3
BADG	7	4	57.1	0	0.0	6	85.7	0	0.0
BIRD	40	27	67.5	2	5.0	35	87.5	6	15.0
NVCA	15	9	60.0	0	0.0	7	46.7	0	0.0
BRC	42	32	76.2	8	19.0	5	11.9	26	61.9
RC	4	4	100.0	0	0.0	0	0.0	0	0.0
COU	46	2	4.3	0	0.0	44	95.7	1	2.2
NRA	36	12	33.3	5	13.9	12	33.3	27	75.0
NPA	35	0	0.0	3	8.6	35	100.0	0	0.0
SNCA	72	36	50.0	28	38.9	28	38.9	18	25.0
OGD/DOE	25	10	40.0	6	24.0	9	36.0	1	4.0
OTHER	3	3	100.0	3	100.0	3	100.0	3	100.0

	1094	**458**	**41.9**	**71**	**6.5**	**770**	**70.4**	**243**	**22.2**

Key to abbreviations - see page 10

Table 2.5.15 **Data management in biotope and land type datasets**

Organisation type	Number of datasets	Fully computerised records		Summary records on computer		Manual system		Archive of originals	
		No.	%	No.	%	No.	%	No.	%
WLT	**81**	**17**	**21.0**	**14**	**17.3**	**43**	**53.1**	**37**	**45.7**
LRC	**126**	**30**	**23.8**	**71**	**56.3**	**45**	**35.7**	**71**	**56.3**
MUS	**19**	**9**	**47.4**	**4**	**21.1**	**19**	**100.0**	**6**	**31.6**
NRS	**10**	**1**	**10.0**	**0**	**0.0**	**8**	**80.0**	**6**	**60.0**
BIRD	**19**	**3**	**15.8**	**1**	**5.3**	**0**	**0.0**	**18**	**94.7**
NVCA	**7**	**3**	**42.9**	**1**	**14.3**	**2**	**28.6**	**4**	**57.1**
RC	**3**	**3**	**100.0**	**0**	**0.0**	**0**	**0.0**	**2**	**66.7**
COU	**42**	**11**	**26.2**	**0**	**0.0**	**30**	**71.4**	**12**	**28.6**
NRA	**16**	**7**	**43.8**	**4**	**25.0**	**8**	**50.0**	**9**	**56.3**
NPA	**53**	**4**	**7.5**	**6**	**11.3**	**48**	**90.6**	**5**	**9.4**
SNCA	**56**	**35**	**64.5**	**21**	**37.5**	**17**	**30.4**	**32**	**57.1**
OGD/DOE	**24**	**10**	**41.7**	**11**	**45.8**	**3**	**12.5**	**7**	**29.2**
OTHER	**2**	**2**	**100.0**	**2**	**100.0**	**2**	**100.0**	**2**	**100.0**
EDU	**1**	**1**	**100.0**	**1**	**100.0**	**1**	**100.0**	**1**	**100.0**
	459	**136**	**20.6**	**135**	**29.4**	**225**	**49.0**	**212**	**46.2**

Key to abbreviations - see page 10

2.5.91 BRC and other units within NERC have high levels of computerisation of data. BRC manages records for many of the national recording schemes (which accounts for the 50% level of computerisation of records claimed for these schemes!). Some local records centres have invested heavily in computerised records management and at least one centre has approximately one million records held on computer with several others holding about half a million each.

2.5.92 Computer management of data is low within the National Park Authorities and few county planning departments have computerised species data (4.3% of datasets) although more have computerised site records (26.2% of datasets) which contain species records. Between 20% and 22% of biological datasets in wildlife trusts (excluding those with semi-independent records centres) have been computerised. More than 40% of taxon-based datasets in local records centres are computerised but biotope and land type records are more likely to be in summary form (56.3% compared to 23.8%). In the voluntary sector, the bird organisations are notable for the degree to which their species observations are kept on computer (67.5%) but the methods of storage vary from simple word processor files to the use of sophisticated databases such as COBRA. BSBI vice-county recorders give a good indication of the spread of computer use among individual amateur naturalists. Of the 79 BSBI recorders that gave details of data management, only 13 (16.5%) had fully computerised records and 4 (5.1%) had summary computer records. Management of data by BSBI recorders varied from simple word processor files to PCs running Recorder, BIORECS and DMAP programs. Some BSBI vice-county recorders are closely linked to local records centres and freely give their data to the centres for management. All BSBI centrally coordinated survey data are sent to BRC, where it is entered on computer and managed as part of the national BRC database.

2.5.93 The conclusion from the Survey is that although organisations such as BRC, JNCC and BTO are far advanced with computerisation of biological records and a number of records centres act as important foci for the computerisation of data, a substantial part of the potential national network (e.g.

about 55%) is still in the early stages and much information, particularly that related to biotopes and land-types remains in paper files or non-standard word processor documents.

2.5.94 There is also appreciable confusion due to the growth in numbers of surveys, their differing objectives and very different levels of data collection which has been further confounded by the lack of any generally agreed standard for biological recording.

2.5.95 One course of action would be to develop a full biological recording standard which could be used to define a series of minimum application standards to match the most common usages of biological records. These minimum standards would serve to guide collectors and managers of biological data on the fitness of their records for different purposes and help to establish a uniform approach to data validation. For the potential user of data this would give a degree of confidence in the likely content of 'products' offered by data suppliers and offer the opportunity of merging data from more than one source, both of which are notably absent at present.

2.5.96 Such a proposal need not restrict the scope of biological recording, as there will always be special purpose surveys with their own data requirements and surveys will vary according to local circumstances. A series of checklists of basic data concepts and guidance on data format which would make data usable for other purposes or enable data exchange could, however, enable survey designers to build in 'standards' at an early stage and possibly extend the potential use of their data. These concepts, together with the Museum Documentation Standard (MDS) (Museum Documentation Association 1991), which has been used for the development of forms for recording site descriptions, are discussed further in Chapter 6.

2.5.97 In the absence of a specific, theoretical *Biological Recording Standard* the most complete data model available, implicitly incorporating a recording standard and applicable to a wide range of biological records, is offered by the Recorder database package (Ball 1992). Despite the growing number of computerised datasets, easy and reliable electronic data exchange between organisations remains difficult to achieve. The principal problems encountered are those of data structure and data integrity. The problem of data structure can be overcome by using the same database as in the case of the concentration of BIORECS users centred around the Wildlife Trust in Dyfed. Data exchange between copies of Recorder is possible but requires appreciable technical expertise and is not yet automated. Although Recorder does not have a data exchange facility at present, an automated routine is proposed for inclusion in Version 4.

2.5.99 It is not practical or desirable to attempt to restrict the development of new or alternative database products, a better approach would be to develop a common data standard which can be used for translating data from the database table and field structure of one database into that used by another. The JNCC Environment Systems and Standards branch are promoting the use of a design technique known as *logical data modelling* to identify and agree data definitions for conservation databases in use in the country agencies. It is hoped that the use of logical data modelling will provide a framework for planning relationships between information systems and aid data exchange. This initiative could be extended to include the interests of other biological recording organisations.

2.5.100 The technical problems of data translation are not the greatest bar to enhanced data exchange in the future. Success depends upon the use of standardised terminology (e.g. habitat/biotope names) for indexing and retrieval, the development of techniques for maintaining non-networked, distributed databases and solving the practical problem of detecting duplication of sites in text-based databases.

2.5.101 Biological recording databases, such as Recorder and BIORECS, are tools which can be used to enter, validate, store and retrieve data from a wide variety of sources. It is not their function to give guidance to what level data should be collected for different purposes. Indeed, in the absence of other guidance, the long data entry screens which are typical of most packages may suggest to inexperienced users that they should be seeking to record all the possible data on those screens. The level to which data should be collected, for specific purposes, remains a frequently discussed problem and needs to be taken into account in any proposal to develop a biological recording standard.

Computers currently in use

2.5.102 There have been very important developments in computer hardware and software since 1985. Desktop machines now commonly have the processing power and storage capacity formerly associated with mainframe computers and the quality of graphics and user interfaces on personal computers far exceeds that traditionally available on large machines. More recently there has been a revolution in networking and communications such that many of the former barriers between machines and different operating systems are disappearing and the potential for building distributed databases across a range of 'platforms' is a reality. For the biological recording sector the most important development has been the declining cost and ready availability of PCs.

2.5.103 The IBM-compatible (DOS-based) PC is the most frequently used type of computer for managing biological records. Of 162 responses to the Survey, 138 (85%) used PCs. Of the 17 badger groups and 18 bird clubs that responded, all used PCs. A similar survey of bat groups carried out by the Bat Conservation Trust in 1992 indicated that 15 out of 29 responding groups (51%) used computers and of the 12 that provided details, all used PCs but no two used the same make or model of machine. Among the local records centres that responded to the questionnaire, 20 used stand-alone PCs, 2 used PCs linked to a local area network (LAN) and 3 were connected to local authority mainframes. Among local wildlife trusts, most if not all now have computers and they have long been supported, encouraged and supplied with software through The Wildlife Trusts. Sixteen trusts gave information on their computers, 15 used stand-alone systems and one used PCs on a LAN.

2.5.104 The details obtained of 204 PCs showed that they ranged from the most basic, using 8088 processors and obsolete operating systems (CPM and DOS below Version 3) to the most recently available 486 machines running under Windows 3.1. At least 43 different makes and many more models of machine were in use but only 3 machines were not using a compatible form of DOS operating system (MSDOS, PCDOS or DRDOS). Of the 148 respondents who supplied technical details of their machines, 55% had 386 processors and a further 24% used 486 processors. This indicates that most machines have been purchased since about 1990 and most would be capable of running or being upgraded to run, modern (i.e. WINDOWS-based) software. Only 2 respondents used Macintosh machines. Although widely used for graphics and desk-top publishing they lack appropriate software and were more expensive than PCs until recently which has discouraged their use in biological recording.

2.5.105 LANs linking PCs, either in a peer-to-peer relationship or in a client-server configuration, allow several users to access software and common resources (e.g. printers) within an organisation. LANs are increasingly, being linked to wide area networks by means of dedicated 'bridges' and 'dial-up' modem links, but linking PCs and LANs to mainframes or wide area networks (e.g. JANET) is technically difficult, generally expensive and confined largely to organisations with specialist computer support (e.g. NERC, universities and the statutory conservation agencies). Network links are becoming common in local authority organisations: 32 out of 221 (14.5%) respondents to the Survey used LANs including two local records centres and two major national recording schemes. Details of 25 network systems were supplied, 19 were DOS-based and 6 were UNIX-based: 16 of the DOS-based systems used Novell Netware in various versions.

2.5.106 Mini-computers are widespread and intermediate between PCs and mainframe computers. Large, powerful minis are used in units of NERC (including BRC), the country statutory nature conservation agencies and BTO. Three of the 13 respondents to the Survey who used mini-computers were county records centres, one linked to a university and two to their local authorities. The advantage of these machines is that a large number of simple terminals can be linked up to one machine which under UNIX or its proprietary variants gives true multi-user, multi-tasking access without the complications of using separate DOS and network software. UNIX is the favoured operating system and development environment for the Open-Systems Initiative and is widely used in both the academic and commercial worlds. Multi-user Unix-based systems cost more and have higher hardware requirements than DOS/WINDOWS PCs and generally need support by specialist technical staff. Recent developments in work stations using graphical user interfaces similar to WINDOWS, although still costly are beginning to bring Unix-based systems into more organisations especially for use with GIS.

2.5.107 Of the 18 (8%) organisations that use mainframe computers for managing biological records, details of 16 were supplied. Three were local records centres linked to local authority computers, 2 were museums (1 linked to local authority and 1 to a university), 2 were local authority planning departments, 1 was a national recording scheme linked to a university, 1 a vice county recorder (linked to a university) and 12 were centrally funded organisations (Government departments , Water Purification Board, the statutory nature conservation agencies and a research council). About the same number of records centres and museums are using mainframes as 10 years ago (4 are the same) compared with the growth in use of PCs and networked PCs.

2.5.108 The advantages of using mainframe computers are that they are supported by technical staff and the running costs may be born by a parent organisation (e.g. local authority or university). Some users have been able to incorporate their data into a wider corporate data management strategy, thereby raising the profile of their work and gaining access to computer resources otherwise unattainable. The major gain for those local authority museums and records centres that have achieved this is access to sophisticated GIS software and a recognition of their biological and geological records as a valuable asset to planning throughout their authorities.

2.5.109 Eight respondents to the Survey said they used outside agencies to manage or input biological records. These included government departments that contracted out the work of creating and populating databases. These figures are misleading since four responses were from national recording scheme organisers who had sent data to BRC for computerisation. But this applies to all 43 schemes (out of 62) that send data to BRC as well as JNCC and the country agencies that contract BRC to supply them with biological records in computerised format. BRC, therefore, is an important provider of computer services to a broad spectrum of users.

Programs currently in use

2.5.110 The database software available at the time of the 1985 NFBR survey was very limited. Among the mainframe users, applications were developed in whatever database the installation supported including Famulus at Manchester Museum, Spires at the Hancock Museum and a specifically written cataloguing system at Leicestershire Museum. The growing number of PC users also had the problem that their machines had very little capacity to run complex programs or hold large data files. The most widely used database programming language was Dbase II, one of the most progressive at the time being the Passmore Edwards Museum sites, species and mapping application which ran on an Apricot microcomputer. The majority of organisations using PCs wrote their own simple filing programs or kept data in word processor files (usually in WordStar). These attempts to create biological records management systems were valuable for the experience they gave but none were sustainable, because none were satisfactory in terms of functionality or data validation. They were highly specific to single installations and resources were lacking to develop them for wider use. In particular, database languages were not powerful enough to cope with the complex data models and numbers of records associated with biological recording. Later, rapid developments in database technology outstripped the resources of organisations to redevelop applications.

2.5.111 Over the past fifteen years, a number of museums and some records centres, have used structured collection and locality recording cards developed by the MDA and used MDA agency services to provide them with printed indexes and catalogues. The MDA used an in-house database called GOS, later marketed as the MUSCAT database. Neither GOS nor MUSCAT achieved any wide uptake because they were technically difficult to work with and lacked reporting flexibility. GOS was based on the hierarchical data structures for museum information management, developed by the Information Retrieval Group of the Museums Association in the late 60's. It was refined into the MDS, MDA cards and MODES data entry program, still used widely for object cataloguing. MDA locality cards are still popular both for recording details of biological sites and their geological variant and were used in the National Scheme for Geological Site Documentation in the 1970's and 1980s. The problem with data recorded in MDS format is that it is very difficult to relate its hierarchical data structure to the relational model used by most modern database management systems.

2.5.112 Since 1985, both computers and database software have improved dramatically whilst falling costs have encouraged use within all organisation types. The pattern of use is very varied but current computerisation of biological records can be divided into five areas.

* **Paper records passed to an agency for management - This is still an important way of allowing small or unfunded organisations to gain some benefits of computerisation and for multiple organisations to collaborate on joint projects. Examples include the national recording schemes that pass records to BRC, museum collection summaries collated by FENSCORE, and individual museums that send records to MDA or processing.**
* **Non-database text files and spreadsheets - A growing number of individual recorders and small voluntary groups (e.g. local badger or bat groups) have personal computers but not the expertise or resources to create or acquire databases. In some instances data may be in the form of extensive text descriptions, not suited to the format required by currently available database applications. Many software packages are used including Ami-pro, Wordperfect, Excel & Lotus 1-2-3.**
* **In-house databases written in programming languages - Databases written 'from scratch' in computer programming languages started to disappear as third and fourth generation database languages became widely available although there is some resurgence of interest caused by WINDOWS-based object-oriented languages such as Visual BASIC which include powerful file handling facilities. Databases of this type are mostly written by computer enthusiasts and are found in only a few organisations. Most are small and highly specific but notable examples reported to the Survey included BIORECS which is used by over 50 individuals and the British Lichen Society Database, both written in PASCAL. The Survey also found databases written in Visual BASIC, COBOL, BASIC and C.**
* **In-house databases written in commercial database management systems - Database technology has advanced rapidly in recent years. The most important developments being fourth generation (4GL) relational database management systems that allow users to quickly prototype and create applications and more recently WINDOWS-based object-oriented front-end programs with expert help systems (*wizards* in Microsoft ACCESS) that allow users to create sophisticated database applications with a choice of underlying file formats. The availability of modern generic database packages has lead to an increase in the number of in-house databases. Examples are widespread amongst all organisation types and include those written by organisation staff as well as bespoke applications written by external consultants. Typical examples include the Marine Conservation Review Database, Seabird Register and Sand dune/Shingle databases in JNCC, RSPB Sites and Species Database, the Kent County Council Countryside Information System - all written in Advanced Revelation; the BRC database in ORACLE; the NRA Biologists BS System in CLIPPER and Gloucester Wildlife Trust's SITEBASE written in FoxPro. The British Fisheries Database and Plymouth Marine Fauna Database were created with Apple Hypercard and are among the few biological records applications on Apple Macintosh computers. Database packages used for biological records reported to the Survey included FoxPro, Smart, Microsoft Works, ACCESS, Q&A, PC File, Paradox, Oracle, Rbase, Superbase, Dbase III, Dbase IV, CLIPPER, Advanced Revelation, Famulus, Prime Information, RapidFile and Dataease.**
* **Databases developed for distribution or sale - This group includes both general site or species recording programs (e.g. Recorder, BIORECS, DRECS and WILDWATCH) and those written for specific markets (e.g. COBRA, BIRD Recorder, CLUB Recorder for birds and, LEVANA for butterflies). Some databases written for specific taxa have been modified for wider use (e.g. BRASSICA for plants and RECORDIT for molluscs) and museum cataloguing programs have been adapted by users for their biological records (e.g. MODES and MICROMUSEE). Some of the packages that are offered for sale have been developed with little or no attention to data standards or data transfer and can be highly idiosyncratic. The package that has had the greatest co-operative development effort and has the greatest technical and user support is Recorder. Recorder is also the most widely used of biological records packages in the UK and is described separately below.**

2.5.113 Recorder was developed for use by the NCC's Invertebrate Site Register (Ball 1994) but following interest from local records centres, wildlife trusts and recording scheme organisers was developed into a generalised biological recording package. Recorder is written for the Advanced Revelation database management system and uses a relational data model with entities including individual taxa and sites. It can be used to store data ranging from site descriptions and simple presence/absence checklists for sites to detailed records for single taxa which include habitat and abundance codes. Associated files allow the

recording of many further details including people, visits, events and bibliographic references. The development of Recorder has been a large scale cooperative effort with considerable input from naturalists who supplied information for checklists and authority files, computer specialists attending the technical working group and volunteer testers. The statutory nature conservation agencies have invested considerable funds, staff time and grant aid to develop and disseminate Recorder in association with The Wildlife Trusts and with further grant aid from outside sources.

2.5.114 Recorder has recently been upgraded to work with version 3 of Advanced Revelation, which has improved its functionality within the constraints of DOS-based text-mode operation. Recorder can be run from WINDOWS but is not a WINDOWS application and a true WINDOWS version is not yet available. The current version is complete with a run-time licence of Advanced Revelation, which will greatly increase its availability to smaller organizations and individuals, as previous versions required a separate and expensive licence for Advanced Revelation. Recorder remains a PC-based application which will work either as a stand-alone or on a network but it is not available for larger machines or UNIX-based operating systems. Data transfer between systems remains difficult using Recorder.

2.5.115 Some 200 copies of Recorder are licensed to organisations and individuals of whom 99 are members of the Recorder User Group. Of the 355 organisations that responded to the Survey, c.15% (51) use Recorder and a further c.8% (28) hoped to use it in the near future. The statutory nature conservation agencies and The Wildlife Trusts have declared their ongoing commitment to support Recorder and continue to provide resources for its development. The Recorder User Group, a self-help group that was formed when no help was guaranteed from elsewhere, operates through the National Federation for Biological Recording. It continues to assist but support for all users is now provided by the statutory nature conservation agencies, currently through The Wildlife Trusts, which also provide training.

2.5.116 Recorder does not suit all organisations and individuals. BIORECS is a simpler package, attractive to many individual recorders and has more than 50 registered users (S. Coker pers. comm.). In West Wales it has been used successfully for local data input by specialist recorders and data transfer to the local wildlife trust. Recorder and BIORECS use different data structures and file types and data cannot be transferred between the two without an intermediate phase of translation and validation. Special interest groups may also find that a generalised package such as Recorder does not suit their own needs and a number of databases are now available targeted at individual taxa including birds, butterflies and spiders. Some such as COBRA, a bird recording database used by several County Bird Societies, are written in Advanced Revelation like Recorder and are broadly compatible with it but others are not.

2.5.117 There are no biological recording packages such as Recorder for those managing biological records on corporate mini-computers and mainframes. The organisations using larger machines all use databases developed in-house and none have been made available on a wider basis. Among the more sophisticated is the Cornish Biological Records Unit's ERICA program which holds nearly one million records of Cornish plants, animals and fossils and which also provides a subset of maps and data for use on a computer network accessible by Cornish schools and other educational establishments.

2.5.118 Development of software for mini- and mainframe computers is considerably more expensive than its counterparts on PCs. In the museum world, however, a number of national museums, university and larger provincial museums have formed LASSI (Large Scale Systems Initiative) which is looking at the development of UNIX-based documentation software for multi-user systems, including documentation of geological and biological collections. There are no similar developments in the rest of the biological recording sector in the UK, principally because much of the collection and collation of biological records is concentrated in the voluntary and local museum sector which do not normally have access to the expensive hardware and software involved. ALICE software, developed at the Royal Botanic Gardens, Kew, is now used there and at the Natural History Museum in cataloguing herbaria.

2.5.119 Computers are widely used to record and monitor wildlife throughout the world. Two major integrated systems, the Nature Conservancy's Biological and Conservation Data System (BCD) in the USA and the Australian government's ERIN system are described in Chapter 6.

2.5.120 There are a growing number of databases established to provide national and international statistics, information for international projects or various extended checklists. Among taxonomists, the greatest progress has been made by botanists and microbiologists. One of the largest international projects, coordinated by IOPI, is to create in the first phase a distributed computerised world checklist of vascular plants. The database is part of a project to create a global plant species information system broadly similar to that being developed for Leguminous plants by ILDIS. The IOPI database will provide an international, standardised, taxonomic checklist and will, if freely available, be an important validation source for future biological recording programs. The IOPI database is interesting because its development has been subject to formal analysis and the data model and entity structures were prepared using CASE tools. This is in line with a growing emphasis in many of the larger organisations involved with environmental information (e.g. JNCC) towards logical data modelling and using formal methods in an attempt to improve software standards. A further technical feature of the IOPI model is its use of the INTERNET for communications and file transfer which could be a model for the UK biological records network. IOPI also holds a database of plant databases at the Royal Botanic Gardens, Edinburgh.

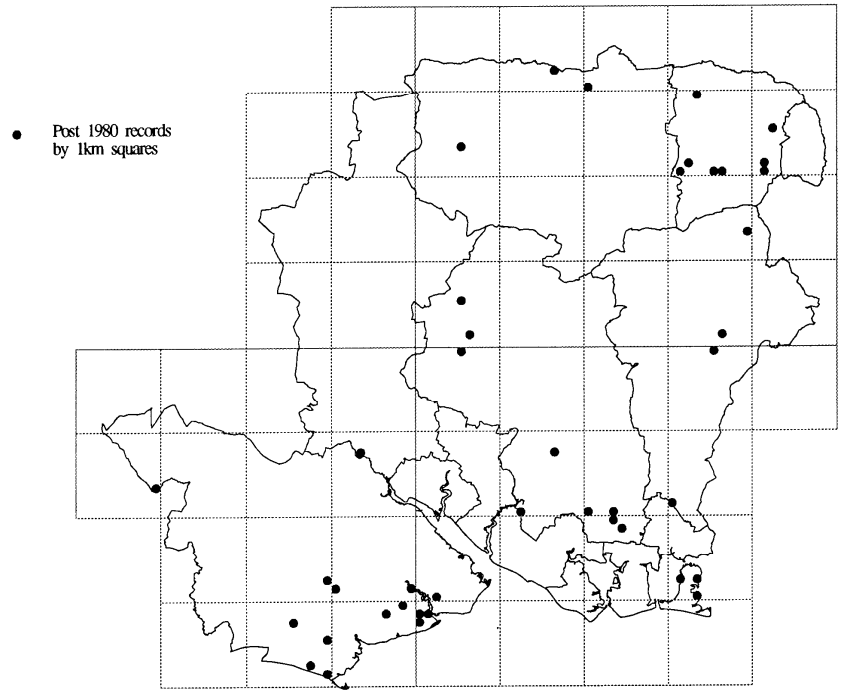
2.5.121 A local development of a plant information system is represented by the UK Ecological Flora Database, developed at York University and available over the JANET network from Bath University Information and Data Services (BIDS). The database holds extensive ecological information on 1777 British native and introduced species of higher plant. The data are held as ORACLE tables and can be accessed through a purpose-written user interface or through SQL Access to the data is by subscription only and is targeted primarily at academic users. Other datasets available through BIDS include the BSBI checklist of the Flora of the British Isles and Key Indicator Species for British Wildlife. Other datasets, including the ILDIS database are expected to follow.

Mapping Programs in use

2.5.122 Distribution maps are one of the commonest requirements from biological records. They illustrate the geographic distribution of species, they can be valuable as a form of validation (for example in identifying grid references that fall outside the recording area), and, during the course of surveys, used to monitor the degree of coverage achieved. When published as provisional atlases they encourage further recording to complete the geographical coverage and update old records.

2.5.123 Basic dot maps use selected grid square resolutions to summarise species distribution. At the national level the preferred resolution is the 10km square, whilst at the local level smaller scales are often used according to the size of the county and the geographic resolution of the data available. These maps can be made more meaningful by the inclusion of more detailed geo-political boundaries (Figure 2.5.2) and physical features such as river courses.

Figure 2.5.2 **Map of the distribution of Dyer's Greenweed (*Genista tinctoria*) in Hampshire showing local authority boundaries (by kind permission of Hampshire County Council Planning Department)**



2.5.124 A variety of mapping programs were found by the CCBR survey ranging from simple 'in-house' versions written by naturalists or organisational programmers to commercially available packages which include statistical plots using histograms and pie charts plotted on maps (e.g. PC.MAPICS). Only 4 examples of simple mapping applications on mini- and mainframe computers were reported to CCBR compared to 64 users of PC applications. The most widely used PC mapping programs are PLOT5 and

DMAP. Hampshire County Council use ARC/INFO to plot diagrams and maps of species and biotope records (Figure 2.5.2).

2.5.125 PLOT5 is the dot distribution mapping program written by Dr S.G. Ball of JNCC and included with the Recorder database, it is also used in association with a number of other in-house databases within the country statutory agencies and JNCC. PLOT5 uses an outline map of the UK which can be scaled dynamically during display to suit the user and over which a variety of grids can be displayed. It can plot a file of grid references with a choice of icons. PLOT5 uses up to 10 scalable symbol types to differentiate data points, can include quantitative data, and can also use colours and shading to cover areas of the map. The latest release of PLOT5 in Recorder 3.2 has the ability to select groups of dots from the displayed map and send the list back to Recorder for inspecting and printing details. Newer versions of PLOT5 have also had significant improvements made to the printed output including Laser printing and export of images to word processor documents. PLOT5 is very closely integrated with the reporting facilities in Recorder and it is therefore likely to continue to be improved and grow in use.

2.5.126 A second PC-based mapping program, UKDMAP, is used within NERC and the statutory agencies as a presentation tool for displaying area and distribution data. Originally written as an information program for marine data, it includes an outline map of the British Isles and surrounding sea areas. Various datasets such as distribution of fisheries can be selected and combined (e.g. with sea floor deposit maps). Copies of UKDMAP can be bought as a marine encyclopedia but the underlying display software has also been utilised by a number of users within the statutory agencies to display land-based as opposed to marine data. UKDMAP's advantages over PLOT5 are that it is a user-oriented presentation program which can be used to distribute spatial information but it was not designed to run as an integrated distribution mapping program.

2.5.127 The most widely used dot map program in the UK is DMAP written by Dr A. Morton of Imperial College. DMAP is a single-user, standalone PC application which is available as a DOS and a WINDOWS application. It can be configured to run in a variety of ways by inclusion of command line strings at run-time. This facility makes it easy to integrate DMAP into database applications and a number of the more widely used packages include DMAP drivers (for example, COBRA, BRASSICA, RECORDIT and the geological recording program GD2). A DMAP driver is available as an extra for Recorder. It is often used in this way because DMAP produces very high quality printed output to a wide range of printers including postscript lasers.

2.5.128 DMAP uses simple coordinate files to draw outline maps and reads grid references for plotting from an ASCII file. The outline map is usually at the county level but can be any size. Multiple boundaries are allowed, so maps can show other political boundaries or physical features such as rivers. Data can be resampled and plotted at a variety of grid levels from hectare to 10km square or plotted as the original grid reference. The Kent County Council countryside information database includes a DMAP driver which allows selection of individual district maps or a whole county map according to the data being plotted. In addition to the 'county-based' DMAP there is a special installation which allows data to be plotted for the whole of Great Britain and includes a rotated grid for the proper plotting of Irish grid references. This special installation was used in modified form by BTO to produce distribution maps for *The new atlas of breeding birds in Britain and Ireland 1988-1991* (Gibbons *et al*. 1994).

2.5.129 In August 1993 there were 255 registered DMAP users, including 123 individuals naturalists, 56 county and regional organisations, 29 national organisations, 28 research and education organisations and 19 consultants (A. Morton pers. comm.). The 255 users includes 71 who use the WINDOWS version and 36 using the special British Isles installation. Future development of DMAP will be concentrated on the WINDOWS version.

2.5.130 A number of commercially available systems are in use which can produce distribution maps but the cost is normally several hundred pounds to several thousand pounds compared to £45 - £75 for DMAP and the free distribution of PLOT5 to Recorder users. Among those in use were academically oriented packages such as PC.MAPICS and GIMMS which can carry out sophisticated statistical plots of data and simple business packages including MapBase and ATLAS Pro. MapBase, for instance, offers a detailed UK map including overlays of towns, villages, roads, features and a gazetteer. The commercial

map packages merge into low end GIS with the ability to do point-in-polygon retrieval of data and on-map access to the database. There are, as yet, very few users in the biological recording sector.

Geographical Information Systems in use

2.5.131 The last 5 years have seen a spectacular growth in the number, scope and availability of GIS. However, despite the development of a number of PC-based systems and the import of low cost 'academic' programmes from the United States, the take up by organisations involved in biological recording and planning has been very low. A survey carried out by the County Planning Officers Society in September 1992 found that less than 1 in 4 county authority planning departments (i.e. 11 organisations) used GIS in relation to environmental records. The CCBR survey found a further 18 environmental organisations regularly using GIS with biological records. These included Scottish Natural Heritage and English Nature, two government departments, two national park authorities, 2 NRA regions, 3 University departments, 3 water companies, 1 wildlife trust, 2 local records centre and NERC (including BRC). A number of organisations (including BTO, National Trust and most National Parks and NRA regions) are actively investigating or developing corporate GIS which will include biological records and many more intend to in the future. A number of organisations including the RSPB and the National Parks have used GIS in special projects both in-house and in co-operation with university departments.

2.5.132 The CCBR survey therefore, indicates that outside university departments there are a minimum of 30 and probably not more than 50 organisations which are regularly using GIS in relation to biological records in the UK. The precise number of individuals and departments in universities involved in GIS research and contract work which impinges on biological records is not known. The Regional Research Laboratories which are funded by the Economic and Social Research Council (ESRC), and at least twenty university departments and institutes actively engaged in GIS-based projects and education, are a major force in the development of GIS techniques applicable to environmental science and conservation. However, most projects are unique and GIS is still far from being integrated into the working life of most organisations involved in biological recording.

2.5.133 It is difficult to estimate the potential biological recording market for GIS, but between 350 to 400 organisations could have an interest, including local records centres, wildlife trusts, county planning departments, metropolitan planning departments, major non-governmental organisations, statutory conservation agencies, regulatory agencies, government departments and utilities. This implies that the number of biological records holders who use GIS is not more than 14% of the main potential market and is certainly less than 3% of the total of 2000 possible organisations and key individuals, such as vice-county recorders, involved. Despite the low number of active projects in the UK there are enough examples to demonstrate the varied use of GIS and its value at all levels from national policy makers to local records centres and conservation trusts. A selection of examples of the use of GIS at different levels follows.

GIS at the National Strategic level - Countryside Information System (CIS)

2.5.134 The ability to hold metadata on biological records and other environmental data is also an aspect of the DOE Countryside Information System (CIS)[3], which represents a new approach in providing wide and easy access to geographically based environmental data. CIS is a desktop computer package containing data in a standard format that allows direct comparisons to be made between data from different sources. It was designed to make available the results of Countryside Survey 1990 to decision makers in Government departments and in the statutory nature conservation agencies, but it is able to handle any information that can be summarised at the resolution of 1km squares. Thus biological records can be mapped (with tabular displays of additional data) for any chosen region of Great Britain. Data already used on the system include vegetation data from ITE's Countryside Surveys in 1978, 1984 and 1990, BTO Breeding Bird Survey records, plant and invertebrate records from BRC, ITE satellite imagery of land cover form CS1990 and the locations of designated areas.

[3] **Details of the Countryside Information System can be obtained from Yvonne Parkes, NERC Centre for Ecology and Hydrology, McLean Building, Crowmarsh Gifford, Wallingford, Oxon. OX10 8BB. Telephone 01491 838 800.**

GIS in the country statutory conservation agencies

2.5.135 English Nature (EN) inherited the Intergraph GIS and computer cartography system installed in NCC in 1989. One of the first priorities was the digitisation of SSSI boundaries by Tayward Data Graphics, who now handle sales of boundary data to other GIS users including county planning offices. The EN system has full GIS capability, digital mapping, satellite image and spatial analysis. Among recent products is a series of county maps based on the AA's topographical datasets overlain with SSSIs, National Nature Reserves and Local Nature Reserves. Mapping of biotopes for the Moorlands and Peatland Resource Databases involved the integrated use of air photo interpretation, ground survey and classified satellite images. The hardware used includes 11 terminals connected to an Intergraph I250 VMS vector-based system, 9 terminals attached to an Interpro vector and raster and a single Intergraph workstation running Intergraph and MapInfo vector and raster systems. Hardware and software costs were c. £150 000 and the installed mapbase costs were c. £300 000 excluding digitisation costs for the SSSI boundaries.

2.5.136 Scottish Natural Heritage (SNH) uses a number of GIS including an Intergraph System similar to that in English Nature, ARC/INFO and MapInfo and also uses ERDAS software for satellite and other raster image processing. Applications include SSSI boundary maps, general land cover and habitat mapping and a specific study of peatlands. SNH is part of the consortium including the Scottish Office Environment Department and Forestry commission that contracted the Macaulay Land Use Research Institute to interpret and digitise the 1988 Land Cover of Scotland (LCS88) aerial photographic survey. The land cover classification used had six principal categories sub-divided into 46 major categories and 124 cover types but analysis included over 1000 mosaics of these types. Interpreted boundaries were digitised on a ten by ten kilometre tile basis and transferred to a SPANS GIS to create a raster dataset which was analysed for basic statistics using ERDAS. More recently the data has been vectorised in an ARC/INFO system for more general release. The LCS88 data are being used for a wide range of applications within Scottish Office departments (e.g. looking at distribution of habitats in relation to the EC Birds and Habitats Directives) and in SNH LCS88 provides reconnaissance level information to target sensitive areas for detailed ground survey.

2.5.137 The Countryside Council for Wales (CCW) has given the development of GIS facilities a high priority and is undertaking a five year development plan to put GIS into its regional offices. The first stage includes trials using ARC/VIEW at the Berwyn Office. Projects at CCW Headquarters include a cartographic facility which is being used to improve the quality of digitised SSSI boundaries and various one-off GIS trials using ARC on Sun Sparc Stations. The developed system may use ARC linked to the corporate ORACLE database and digital plotting datasets downloaded from various Advanced Revelation applications.

GIS in regional planning and conservation - National Parks

2.5.138 . All the National Parks have been involved with a landscape change survey (MLCNP) carried out by the Cranfield Institute of Technology between 1988 and 1991 (Taylor *et al.*1991). The MLCNP prepared for each national park a SPANS GIS database holding information on landscape features of the parks in the 1970s and the mid-1980s. Data were derived from aerial photograph interpretation and converted into digital form to give area feature maps with 20 metre resolution. Area features were aggregated into 38 land cover classes and held as quadtree maps in the SPANS GIS. Point and line data were classified into 12 classes and held as lengths or counts per kilometre square in a Foxbase relational database. The MLCNP database has been used to produce maps and tables analysing change in each park for county, district, parish and 10km square over a ten year period. The task was an arduous and lengthy one, for instance, digitising the hand-drawn 1:10,000 overlays for the Lake District Park took six man months.

2.5.139 Not all Park Authorities have adopted the MLCNP GIS but the Lake District National Park Authority is evaluating a development project jointly with the Countryside Commission comprising case studies including distribution mapping, planning constraint mapping, thematic mapping, digital terrain mapping and spatial analysis. These projects are being carried out using two PC-based vector/raster

software packages, SPANSMAP and MAPDATA. Digital elevation data obtained from the OS has been used by the Lake District National Park Authority to carry out a forest design study for Whinlatter Forest in the north-west of the Park. Landscape data was draped over the digital terrain model which could then be rotated and viewed from various angle to assess the impact of boundary changes. Other parks have been assessing the use of PC-based GIS in their work. The Peak District National Park Authority use WINGS, which runs under WINDOWS, for species distribution mapping, planning constraint mapping, maintaining site boundary data, utility and asset mapping and ecological analysis. The species mapping is carried out by direct access to a database written in Superbase 4. The system is at present still under evaluation prior to extensive digitisation of Phase I Habitat survey maps. The Brecon Beacons National Park Authority is using MapInfo to maintain site boundaries in vector format.

GIS in national and international research - ITE

2.5.140 ITE has been developing the use of GIS for many years, mainly within the Environmental Information Centre (EIC), of which BRC is a component unit. BRC has only recently begun to assess the use of GIS in its work and among other projects, it is currently evaluating the use of IDRISI , a GIS available at low cost to academic organisations. One recent EIC project carried out in collaboration with the Environmental Resources Unit of Salford University has developed live links between the BRC ORACLE database and a Laser-Scan HORIZON GIS (Ullyett, *et al.* 1993). The BRC datasets comprise both species observation records and a database of ecological preferences of major groups. These data have been linked in the GIS with physical variables (e.g. altitude and rainfall) and land cover information derived from the National Land Cover Map project (see Barr *et al.* 1993). Tests have given promising results particularly with the preliminary study of species in moorland areas. This work will be extended in future projects looking at patterns of biodiversity in Britain and attempting to detect regional 'hotspots' (see for example Prendergast *et al.* 1993). This project is among the first of its kind within Europe and is an important pointer to future uses of biological records.

GIS in national non-governmental organisations

2.5.141 . A number of major national non-governmental conservation agencies are actively evaluating GIS at present although none use GIS yet for managing and analysing biological records. The National Trust is evaluating a vector-based GIS with property records and the BTO has recently purchased the ARC/INFO package. The RSPB has used GIS in collaborative research as in the study of wader populations in the Flow Country of Scotland (Avery & Haines-Young 1992) where ground observation was linked to Landsat TM images to produce prediction maps of likely breeding concentrations.

2.5.142 RSPB is also using a PC-based raster GIS (Datascape) linked to an Advanced Revelation database to create an estuaries GIS for the UK. Basemaps are scanned in as tiled raster images and areas of interest are digitised on-screen as copies from field survey maps. The database holds information for a wide variety of entities ranging from protected sites to marinas, power stations and car parks. These data can be plotted as icons on the map as can icons representing wildfowl counts from the Wetland Birds Survey. The system in use lacks the power and accuracy of larger GIS, such as ARC/INFO, but demonstrates what can be achieved rapidly and economically using relatively simple PC software.

GIS in county planning

2.5.143 The use of GIS within local authorities is growing rapidly and is likely to become the single most important market for GIS in the UK (Rix 1993). There are, however, few if any, successful fully corporate systems in place and few planning departments (about 11 counties and possibly twice that in some of the larger districts) have access to GIS for environmental information.

2.5.144 A number of district and county authorities have undertaken projects to computerise planning alert maps and Phase I type landcover surveys. These include Avon County Council using a raster-based PC GIS (Datascape), Kent County Council using a PC vector-based GIS (ARC/INFO) and East Sussex County Council using a PC CAD package (AutoCad). The Berkshire County Council Department of Highways and Planning has installed an ARC/INFO Atlas of digital data on a network of PCs in three of

its divisions. This application uses OS vector maps which can be selected by area of interest with overlays, which include archaeological sites and the Berkshire Habitat Survey. Hampshire County Council had an early experiment with GIS using IBM software on its mainframe but now uses an ORACLE-based system with a data structure/file system modelled on Recorder: the system is linked to ARC/INFO. Some 4 000 site records and 160 000 species records are included although the boundaries of the biological sites are not yet digitised. Among the proposed uses are distribution mapping, planning constraint mapping and displaying protected site boundaries.

2.5.145 The Kent County Council Countryside Information System is another system using PC-based ARC/INFO software. This project is nearing completion and has generated information on more than 23 000 land parcels in the county. Target notes with summaries of habitats, species, damage and management are entered onto an Advanced Revelation database and fair copies of field maps are digitised into the ARC GIS. The survey has been funded by a partnership of the county council, English Nature, Kent Trust for Nature Conservation, Southern NRA, British Rail and the Kent District Councils. Partners receive printed reports listing sites by habitat together with a set of coloured alert maps showing landcover distributions from the Phase I Habitat survey of the county. This system would form an excellent nucleus for an environmental information system for the county although unfortunately there is at present no link with the Kent Biological Records Centre held by Maidstone Museum.

GIS in wildlife trusts and local records centres

2.5.146 The CCBR Survey found only 4 county wildlife trusts and local records centres using GIS out of a sample of 69 who responded to the questionnaire, a further two were evaluating systems and two other users were known to the authors. Two museum-based centres had access to GIS systems on local authority mainframes, one local record centre used ARC/INFO on a mini-computer and two used a PC-based raster GIS. The typical interest of users is the spatial retrieval of 'site' data to answer such questions as "what sites of wildlife and geological importance lie within this area?", the area of interest being defined by a user-drawn polygon, radius or corridor surrounding a drawn line. These searches would normally be initiated when scanning planning proposals or in direct response to enquiries. The Gloucester Wildlife Trust has its PC-based GIS fitted with a touch screen and has been programmed as a wildlife information system available to the public. The Cornish Wildlife Trust in conjunction with Cornwall County Council has recently obtained a large grant from the EC LIFE fund to set up a Cornish Environmental GIS, which demonstrates the value of co-operation in undertaking expensive GIS projects although in this case it does not include the already firmly established Cornish Biological Records Unit which holds a computerised record approaching one million records.

Communications

2.5.147 A genuine revolution in communications is at present under way. This involves not only computer data but all forms of electronic information including voice and pictures. The development of integrated services digital networks (ISDN) and the exponential growth in use of world-wide computer networks such as the INTERNET have already changed the way that business is conducted and scientific information disseminated. It is now standard procedure for microbiologists and geneticists to place research results on the INTERNET or in the Microbial Strain Network and some journals make this a prerequisite of publication. Current estimates (July 1994), which increase almost monthly, indicate that the number of INTERNET users, currently around two million world-wide, is growing at 15% per month and services such as E-mail, hitherto mainly restricted to the mainframe-using, academic community on the JANET network are becoming common-place means of communication.

2.5.148 The revolution is somewhat slower coming to the biological recording sector. Only 3 out of 355 (<1%) respondents to the CCBR survey gave an E-mail address and 11 out of 317 (3.5%) said that they exchanged or provided data over a computer network. The principal E-mail and file transfer used by respondents to the CCBR survey were JANET and, in a wider context, INTERNET. Unfortunately use of JANET is restricted at present to higher education, research council and governmental users and costs could not be born by the many smaller organisations involved in biological recording. There are,

however, a number of readily available commercial services, for example COMPUSERVE and Demon.Co.UK, which provide bulletin boards, conference services, file exchange and E-mail with gateways to other networks. A sustained effort is now being made in the UK to make INTERNET more widely accessible.

2.6 DATA EXCHANGE AND TRANSFER

2.6.1 From the results of the Survey, the exchange of data between organisations appears to be low. Three quarters of all data are used only within the original collecting/collating organisation. Nevertheless, there are a number of formal and informal networks of organisations that share or exchange information. At the national level the most extensive exchange is that between ornithological organisations (BTO, WWT and RSPB) and JNCC, where counts of wetland birds, in particular, are widely exchanged. Data are collected and collated mainly in paper format but are computerised by BTO and WWT and shared in computer readable formats. At the local level many counties now have biological or environmental recording forums which coordinate the collection and sharing of survey data. These forums usually include local records centres, wildlife trusts and county planning departments, together with representatives from other organisations such as local natural history societies, badger groups, bat groups and the regional staff of the statutory nature conservation agencies. Data are usually exchanged in the form of photocopied record cards or as site registers and alert maps (i.e. maps which show sites of importance for wildlife). There are, as yet, no fully computerised biological data exchange networks in the UK.

2.6.2 The majority of the 317 organisations which provided information on the provision and exchange of data could quantify the levels of data flow and quantity of records only in very broad terms (e.g. 'frequent', 'occasional', 'don't know'), presumably because most data transfer or exchange was not carried out on a formal basis and, as result, was poorly documented. There are some exceptions among most of the organisational types, chiefly those which are highly computerised, those which have formalised policies on data acquisition, or those contracted to other organisations to supply data (such as BRC to JNCC).

2.6.3 Information from the Survey on the exchange of data between organisations inevitably is biased by the number of organisations in each organisational type that responded, but it shows clearly (Table 2.6.1) that local records centres, museums, BSBI vice-county recorders, national recording scheme organisers, statutory nature conservation agencies and wildlife trusts are the principal suppliers of information to the widest range of organisations. Wildlife trusts receive information from the widest range of sources. The figures also show that 72% of the local records centres responding exchanged data with wildlife trusts and 82% provided information to (or received it from) individual members of the public.

Table 2.6.1 **Examples of data exchange between organisations, numbers of 'transactions' reported to the Survey**

	WLT	SNCA	NRS	BSBI	MUS	LRC	Overall Total	
BTO	**4**	**8**	**1**	**0**	**8**	**8**	**66**	
NRA	**11**	**7**	**3**	**1**	**11**	**15**	**70**	
NT	**5**	**8**	**19**	**3**	**9**	**1372**		
EDU	**7**	**9**	**5**	**1**	**13**	**14**	**74**	
MUS	**8**	**6**	**15**	**7**	**7**	**11**	**75**	
JNCC	**3**	**11**	**14**	**2**	**10**	**12**	**78**	
SNCA Regions	**10**	**7**	**7**	**1**	**16**	**22**	**89**	
LRC	**0**	**7**	**14**	**24**	**4**	**8**	**90**	
In-house staff	**9**	**11**	**3**	**0**	**17**	**22**	**91**	
Local biological societies	**5**	**4**	**4**	**19**	**17**	**22**	**93**	
BRC	**4**	**6**	**20**	**13**	**20**	**25**	**103**	
EN/CCW/SNH	**10**	**9**	**13**	**26**	**7**	**9**	**103**	
L.AUTH Planning Depts	**16**	**8**	**5**	**26**	**17**	**23**	**130**	
Individual recorders	**17**	**8**	**24**	**13**	**25**	**34**	**153**	
NRS	**7**	**5**	**9**	**70**	**20**	**27**	**155**	
WLT	**4**	**13**	**14**	**39**	**25**	**29**	**176**	

Key to abbreviations - see page 10 and Glossary

2.6.4 The main flow of biological records in the UK is summarised in Figure 2.6.1 as a simplified network diagram, based on the Survey and on interviews. The movement of biological records is not limited solely to the routes depicted because it is possible for any individual organisation to develop links with any other organisation in the network. Most of the links are not formalised and many are based on the role of key individuals who may be involved with more than one type of organisation. For example, the organiser of a national recording scheme may also work for a national non-governmental conservation organisation or a museum and also have close links with a local natural history society. Many of the links between organisations are dependent on close, but informal, personal cooperation between individuals.

2.6.5 Formal links between organisations are being based increasingly on policy agreements and the development of special interest networks, and on financial contracts and service agreements. At the local level, some local records centres are now funded principally through grants or service agreements with local government planning departments. Nationally, many bodies which had developed with full or substantial funding from central government have experienced a progressive loss of such 'core-funding' and are now increasingly funded through service contracts, in many cases through competitive tendering. Centrally funded organisations, such as universities, national museums and research councils now operate in ways which have many similarities with companies in the private sector, but with few of the financial advantages. These developments are leading to changes in previously well established and open routes for data flow, especially where financial considerations restrict previously free movement of data between organisations.

2.6.6 Figure 2.6.1 shows the relationship between the voluntary, local government, central government and private sectors. At the national level, the main interface between the voluntary sector and the centrally funded sector is through JNCC and BRC. JNCC is the main link to the ornithological and marine recording networks which are well defined and include negotiated agreements between the main organisations with established protocols for the collection and transfer of survey data. BRC has a role in the supply of data on other forms of wildlife for which it is contracted to JNCC, but there is some overlap in responsibilities. For example, BSBI records are handled by BRC but information concerning rare and scarce species may go direct to JNCC. BRC is the central focus for a large number of national invertebrate recording schemes but individuals also provide data to the Invertebrate Site Register which is operated by JNCC. Both BRC and JNCC provide data to the country agencies as requested.

2.6.7 BRC has a formal agreement for the exchange of data with the statutory nature conservation agencies, which is linked to the contractual relationship between JNCC and ITE to support the work of BRC. In recent years BRC has become part of ITE's Environmental Information Centre and like other agencies is diversifying its work in response to the need to support work through contract funding. BRC's original role was to map the fauna and flora in a British and European context, achieved mainly by encouragement, coordination and publication of recording carried out mainly by voluntary groups. These links with the national recording schemes and local recording, conservation and planning networks remain strong and make BRC a major gateway between the voluntary and statutory sectors. However, there is no formal agreement for data exchange between BRC and its main sources of data, the voluntarily organised national recording schemes.

2.6.8 At the level of individual counties, data flow is well-established although the details vary between counties and in some it is poorly developed. The main nodes of the local network are the wildlife trusts, local records centres, county and the district planning departments. Local natural history societies are often an important source of records and expertise, and some aspects of recording and conservation are handled by special interest groups including bat groups, badger groups and RIGS groups (which cover geological and geomorphological conservation). The individual organisations involved are coordinated either through informal or, increasingly, formal wildlife forums. In many counties the record centre is based at a local authority museum, in others it is part of the wildlife trust and elsewhere it is part of the county planning system.

2.6.9 In most record centres, other than some based at museums, the principal sources of funding are county and district planning departments and much of their activity is, of necessity, directed to providing

planners with alert maps and interpreted information on sites of interest for wildlife. The main data flow in local networks is, therefore almost one-way but each of the organisations involved may generate biological records and these are often freely exchanged, so there is significant feedback. Most local record networks are characterised by free-flow of data and this extends outside the immediate network to the use of records in local education, advice to the general public and exchange of records with national non-governmental conservation organisations, national recording schemes and BRC. Wildlife trusts and records centres frequently provide data to regional staff of the conservation agencies, usually through contract surveying, or in return for grant aid. Sale of information to commercial environmental consultants, developers and utilities, although often cited as a source of income, normally account for only a small fraction of the overall flow of data in the system, and for a small part of the total budget.

2.6.10 At the governmental level, very little of the data flow comes directly from voluntary sources. Much comes indirectly through the country agencies and JNCC but most is generated within departments, especially MAFF (e.g. ESA monitoring), or contracted out either through the research councils (NERC and the Biotechnology and Biological Sciences Research Council (BBSRC)), or by tender to staff in higher education establishments, the national museums or individual units of the research councils. Of the centrally funded organisations, the statutory nature conservation agencies and the NRA regions have the closest links to the local recording networks, providing and receiving data and, as a consequence, they are becoming an important source of supplementary funding into those networks.

2.6.11 There is no single route for information to enter the system. New biological records can be collected or commissioned by any of the organisations or organisation types. The principal source of records, particularly taxon-based records, remains the voluntary sector although biotope, site-based and monitoring data, in particular, are collected by contracted surveyors and in-house staff in the statutory funded organisations (e.g. ESA monitoring by MAFF, river corridor surveys by the NRA regions, and biotope surveys commissioned by the statutory nature conservation agencies). There are well defined routes for data flow and data exchange agreements in place, but it is often seen to be easier for organisations to commission their own recording programmes rather than attempt to locate and interpret existing data. One reason for this is that there is little consistency between 'networks' either in the structure of data or accessibility of data. It is also hindered by the lack of a *metadatabase* describing data sources, content and quality.

Data Transfer Formats

2.6.12 Where data provision and transfer take place, they are normally in paper format although a growing number of organisations use computer databases to manage their data. Of 169 organisations which responded to the Survey on what media they used to provide information, 135 (82.3%) supply straight photocopies of original data, 89 (54.3%) provide interpreted or transcribed data in paper format, 93 (56.7%) provide maps. Only 45 (27.4%) provide data on floppy disk, 3 (1.8%) on microfiche, 4 (2.4%) on magnetic tape and 11 (6.7%) were able to provide data over a computer network. Apart from BTO, the network users were all large publicly funded bodies or units within these bodies including ITE, JNCC and EN. The list of network users included only one county planning department, but many local authorities use computer networks and have the technical resources potentially to transfer data internally and externally by this means. However, in the majority of local authorities, the records centres, museums and conservation sections of planning departments have traditionally had low priority for access to these systems.

2.6.13 Among the 57 organisations providing information on computerised databases to the Survey, 27 (45.7%) reported that they could export data in ASCII format, 26 (45.6%) as Dbase files (.DBF format), 7 (12.3%) by SQL, 3 (5.3%) in Lotus 1-2-3 (.WKS) format, 2 (3.5%) in AREV and 2 (3.5%) in other formats. Only 19 organisations gave details of actually transferring data. Of these 10 used ASCII, 7 used Dbase, 1 SQL and 1 AREV. All used custom data structures. The main computer exchange of data was among the larger and centrally funded organisations. NRA regions transfer survey data in Dbase format to a central database held by a NRA region (Thames). In JNCC, both the Marine Conservation and Vertebrate Ecology and Conservation branches import computerised data from outside sources such as BTO. Some museum-based records centres exchange data between museum cataloguing systems and the records centre database.

2.6.14 The Survey showed that, at present, data exchange and transfer still rely heavily on manual methods and are generally inefficient. The poor number of responses to questions on data transfer formats and actual transfers implies either that this area of data management is poorly understood by most of those involved in the collection and management of biological records, or that they have had little opportunity to develop effective methods. However, manual methods such as maps and field notes, may be more appropriate for transferring data in some circumstances, for example in the absence of GIS facilities and when detailed information about an individual site is required. Access to information in paper forms can be greatly increased by computer cataloguing and indexing.

2.6.15 The software in use is one of the hindrances to changing this situation. Although many records centres, trusts and individual recorders are using the Recorder biological recording package, the structure of the data within Recorder and the complications caused by Advanced Revelation's internal filing system make reliable data transfer between Recorder systems technically difficult to achieve. Data are routinely transferred between some Recorder users but this relies on the technical skills of the managers involved. A data transfer module for Recorder is to be developed but no date has been fixed for its release. Some users of the COBRA bird recording package (also written in Advanced Revelation) are routinely importing data from separate data input modules and transferring records to other COBRA users within a county network. In Dyfed a large number of recorders use the BIORECS recording package and routinely transfer data to the wildlife trust database. Successful examples such as these rely on a high degree of cooperation between recorders to ensure compatibility of records, for instance, using an agreed list of sites so that problems of site overlap or hierarchical mismatches are avoided, or at least reduced.

2.7 NON-BIOLOGICAL DATA

2.7.1 In most cases, reference to non-biological data is essential for the interpretation of biological records. For example, species ranges may be related to climate, soils or surface geology, and site-based biological characteristics may be determined by ownership, management and the protection status of sites. All these types of data are spatially referenced and most are also temporally referenced. Most non-biological data, other than those collected contemporaneously, have to be acquired from secondary sources. The agencies responsible for many of these types of data are described in the invaluable but neglected Chorley Report on geographical information (DOE 1987).

_____BOX_____

Examples of non-biological data for use with biological data

Basic cartographic information:
grids, administrative boundaries, settlements, communications and services, landscape features
Topography:
altitude and relief, slope, aspect, land use
Soils:
types, chemical composition, quality, sensitivities
Drift, surface and solid geology:
types, exposures, Regionally Important Geological Sites, aggregate and mineral workings
Meteorology:
precipitation, humidity, temperature, sunshine, wind speed/direction, evapo-transpiration, visibility
Freshwater:
types of water bodies, catchments, water use, environmental attributes, water quality, pollution
Marine:
marine charts, water quality, flow, tides, erosion, pollution
Air quality and pollution:
trace gases, particulate deposition, ultraviolet radiation
Built environment and development:
transport routes and usage, derelict land, contaminated land
Land ownership, access and site protection:

land registers, public rights of way, site protection and designation (e.g. SSSI/ASSI, National Park), Tree Preservation Order

Historic heritage:

Sites and Monuments Record, Scheduled Ancient Monuments, Historic Parks and Gardens, listed buildings

Socio-economic:

population census

2.7.2 Current use of non-biological data was covered in 104 questionnaire returns, suggesting that this topic was under-represented in responses to the CCBR survey. The distribution of these returns among organisational types showed a bias to use by organisations involved with nature conservation, planning and land use regulation. Local government and almost all the National Park authorities and NRA regions, use non-biological data, as do the statutory conservation agencies, NERC and DOE. Other significant users are local records centres/museums and wildlife trusts. Table 2.7.1 summarises the types of non-biological data held by these organisations: some hold several types of data. The methods used to correlate biological data with these non-biological data are summarised in Table 2.7.2. More than one method of correlation is possible for any one biological dataset. Further details of the use of GIS are given in 2.5.131 to 2.5.146 but it is notable that use of GIS is still far from universal. Planning constraint maps were used by those organisations with active involvement in the local planning process, for example, 9 national park authorities, 6 local records centres, 6 local authorities and 5 wildlife trusts/urban wildlife groups. Information on the geographical units used to correlate non-biological data with biological data was provided for only 72 returns, and many of these used more than one unit: sites 39, site boundaries 33, grid squares 37, parishes 11. Also listed were river and catchment boundaries, common land boundaries and vice-counties. The terms *sites* and *site boundaries* were used as alternatives by respondents, so that almost all 72 organisations correlated data by defined sites.

Table 2.7.1 **Types of non-biological data held by organisations and the number of users (N = 102 organisation)**

Type of data	Number of users
Soils	34
Geology	58
Climate	21
Land use	43
Land ownership	45
Protected areas	52
Tree preservation orders	28
Footpaths	37
Sites and monuments record	41
Scheduled Ancient Monuments	42
Listed buildings	37
Historic landscapes or gardens	28
Water quality	21
Chemical data	16
Pollution	26
Other	11

Table 2.7.2 **Methods used for correlating biological data with non-biological data (N = 103)**

Method used	Number of respondents

Geographical information systems (GIS)	**17**	
Other computer systems	**39**	
All paper maps	**47**	
planning constraint maps		**23**
other maps		**38**

2.7.3 Many of the non-biological datasets which could be considered to be of key national importance for the use and interpretation of biological data are held by governmental organisations with 'agency' status, or by commercial companies. Basic, essential national datasets in digitised forms, such as those for soils (Soil Survey and Land Research Centre), geology base maps (NERC), weather records (Meteorological Office), English SSSI boundaries (Taywood Data Graphics), Scheduled Ancient Monuments (English Heritage/Cadw/Historic Scotland) and, especially topographic and basic cartographic data (Ordnance Survey), have been compiled with funds originally provided by central government. Because these organisations are now obliged, as a result of Government policies on tradeable information (Department of Trade and Industry 1986), to recover their costs and in some cases to charge full commercial rates, such datasets are available only at very high costs to the user. Where the user is another government agency, local records centre or voluntary organisation, these costs are frequently beyond their budgets. It is worth noting that this situation in the UK compares unfavourably with that elsewhere, .eg. in the USA, where such information is supplied either at the cost of providing access or free of charge.

2.8 USES AND USERS OF DATA

2.8.1 The purposes for which biological recording is undertaken have changed with time and have gained in complexity since its formalised beginnings in the 1950s. The modern phase of mapping species distributions, which began in the 1950s, was driven by simple scientific enquiry, by the need for information to conserve wildlife at national and regional levels and by interest in the environmental factors which affect distribution (Perring 1960, 1976). Similarly, the Phase 1 Habitat Survey, initiated by NCC in the 1970s, was intended to provide knowledge of natural and semi-natural biotopes, and their location, extent and distribution, to aid in wildlife conservation (Wyatt 1991). In the early days of biological recording, user demand was assumed rather than defined. Although NC, the voluntary conservation movement and planning authorities were, as early as 1973, defined as *consumers* of biological records (Stansfield 1973), the first attempts to define the true range of uses and users were by authors in Copp & Harding (1985) and Stansfield & Harding (1988) and in the Linnean Society report (Berry 1988).

2.8.2 A range of more focused approaches to recording has developed since the late 1970s. Increasing awareness of the potential uses of data, and of the real costs of acquiring and managing them, have brought about the need to define objectives and establish priorities. However, such changes have been piecemeal due to lack of coordination or clearly defined policies beyond those of individual organisations. At the same time, biological recording has continued to develop in response to changing demands for data, brought about, for example, by new legislation and changing public perceptions. In its Statement of Intent in 1990, CCBR summarised four major uses for biological records:
*** Biological research;**
*** Environmental assessment;**
*** Planning;**
*** Land management.**
However, these and most other types of use are inextricably linked, thereby underlining the multi-purpose nature of most biological recording.

2.8.3 Legal requirements for biological records are described in Chapter 3, but most are implied rather than explicitly defined. Most international, EU or national legislation assume an information resource on the number and distribution of species and habitats, both within designated sites and more widely

(regionally, nationally or EU-wide). Almost all of the main functions of the statutory nature conservation agencies rely on the existence of such information (Table 2.8.1).

Table 2.8.1 **Functions of the statutory nature conservation agencies supported by biological recording**

Function	Contribution by biological recording	
Determine Species Status and Threat		
1. Establish criteria	✓	
2. Evaluate species distribution	✓	
3. Evaluate change in distribution	✓	
4. Evaluate species population	✓	
5. Evaluate change in population	✓	
6. Determine threats	✓	
7. Evaluate degree of present protection		✓
8. Declare status	-	
9. Advise on legislation	-	
10. Record legislative action	-	
Interpret Species Ecology		
1. Establish criteria	✓	
2. Evaluate species/habitat relationships		✓
3. Evaluate impacts	✓	
4. Determine relationships to threats	✓	
5. Determine threats	✓	
Maintain and Enhance Species		
1. Protect species	✓	
2. Evaluate sites	✓	
3. Designate sites	✓	
4. Maintain and manage sites	✓	
5. Protect sites	✓	
6. Assign grants	-	
7. Licensing	✓	
Inform and Educate		
1. Identify customers	✓	
2. Identify how customers' needs will be met	✓	

3. Disseminate and advocate (including publication of books)	✓
4. Advise Government/other government departments	✓
5. Evaluate effectiveness of advice product	-
6. Set priorities for information collection	✓
7. Develop strategies	✓

2.8.4 National and local voluntary bodies have a self-appointed, but nevertheless important, role in monitoring the implementation of both international and national legislation on the environment. They are users of information of all types, including biological information, but rarely have sufficient resources to collect their own data and are, therefore, often dependent on the availability of information in the public domain. Access to such information often has proved difficult, despite recent legislation on access to environmental information (see Chapter 3).

2.8.5 The Survey described and quantified the range and rates of data usage. In Question 5.2 (see Appendix 2), 13 key purposes for which data are collected and used within organisations were listed. To these 13, a further 4 purposes which were recorded by several respondents under *Other* have been added. Table 2.8.2 summarises the relationship between types of organisations and these 17 main purposes for which data are collected and used.

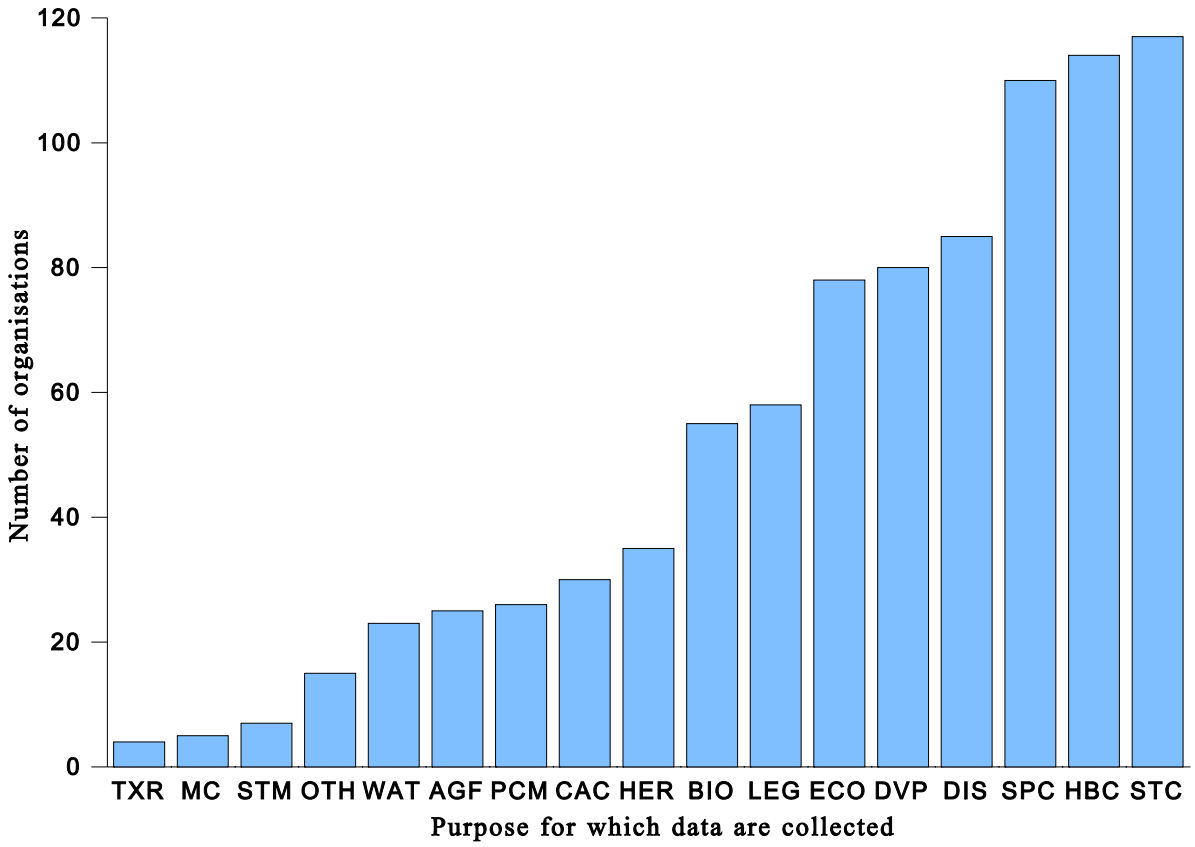
Table 2.8.2 **Organisational types and purposes for which data are collected and used**

	No.	Agf	Biog	Cac	Disu	Dvp	Eco	Hbc	Her	Leg	Mc	Pcm	Spc	Stc	Stm	Txre	Wat	Oth
BIRD	3	1	1	0	2	2	2	1	0	0	0	0	3	2	0	0	0	1
CNHS	2	0	2	0	1	0	1	1	0	0	0	1	2	2	0	0	0	0
NRS	30	2	24	2	13	4	16	15	3	4	0	1	18	18	0	1	1	5
WLT	30	4	3	2	14	26	12	28	4	16	0	2	23	29	1	0	3	1
LRC/MUS	36	3	15	10	30	19	17	25	16	10	0	4	24	26	0	1	4	3
NVCA	4	1	1	1	1	1	2	4	1	2	1	1	2	3	0	0	1	0
L.AUTH	9	3	0	3	3	6	1	8	2	1	0	1	7	7	1	0	1	2
RC	5	0	1	0	4	1	4	1	1	4	0	2	3	3	0	0	1	0
SNCA	9	3	2	4	7	2	4	8	3	7	0	3	7	7	1	0	1	0
OGD	6	2	1	1	5	3	1	4	4	3	2	2	4	2	2	0	1	3
NPA	9	7	1	7	4	8	9	9	4	4	0	2	7	9	0	0	3	1
NRA	8	0	0	0	0	3	2	4	0	4	0	8	3	3	0	0	5	0
OTHER	3	0	1	0	0	1	1	2	0	2	0	0	2	2	0	0	2	0
Totals	154	26	52	30	84	76	72	110	38	57	3	27	105	113	5	2	23	16

Key to abbreviations - see page 10

2.8.6 This analysis demonstrates both the wide range of applications of biological records and the large number of organisational types holding data for any one application type. Of the 154 organisations, the majority use biological records in relation to site, habitat and species conservation, but with a wide range of other uses (Figure 2.8.1). The results are clearly biased by the high number of returns from wildlife trusts and local records centres/museums. Some surprisingly low figures are shown for biogeography and, in particular, for taxonomic research, which possibly reflects the general neglect of taxonomy in the UK (Advisory Board for the Research Councils 1977, House of Lords Select Committee on Science and Technology 1992, Krebs 1992). However, the target of the CCBR survey was those organisations which collect and hold data rather than secondary users.

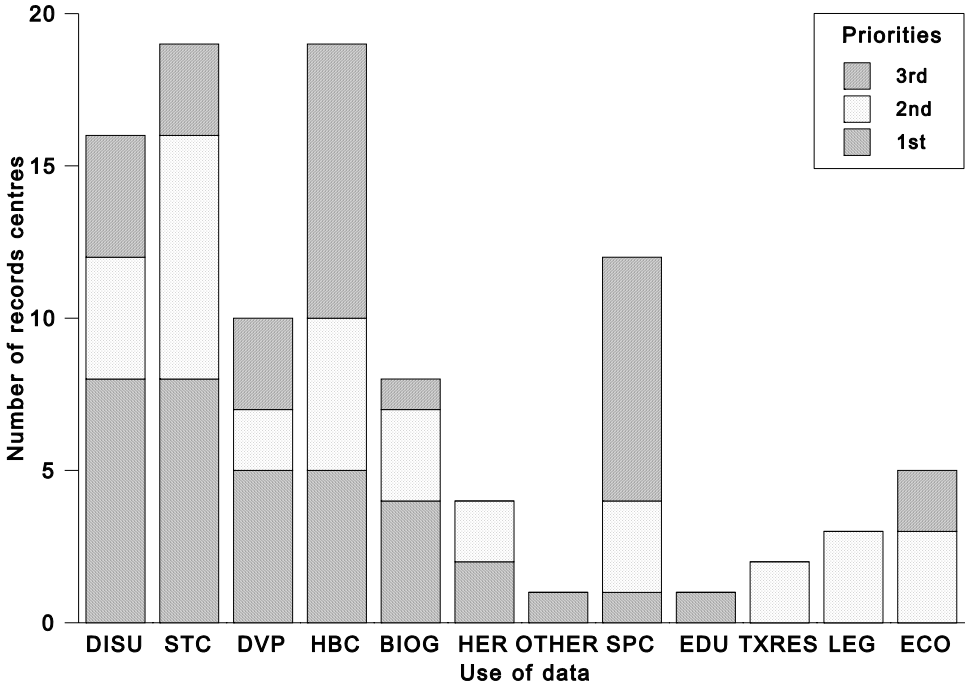
Figure 2.8.1 **Purposes for which data are collected and used - all organisational types**



Key to abbreviations - see page 10

2.8.7 The pattern of use and dissemination demonstrated in Figure 2.8.2 reflects the broad range of responsibilities perceived by organisations as being within their remit. For local records centres/museums, the priority ranking from the analysis of returns (Figure 2.8.2) suggests that use in development and strategic planning is lower than expected, but this is possibly a matter of interpretation by the respondents who may classify work related to planning applications as being site or habitat conservation.

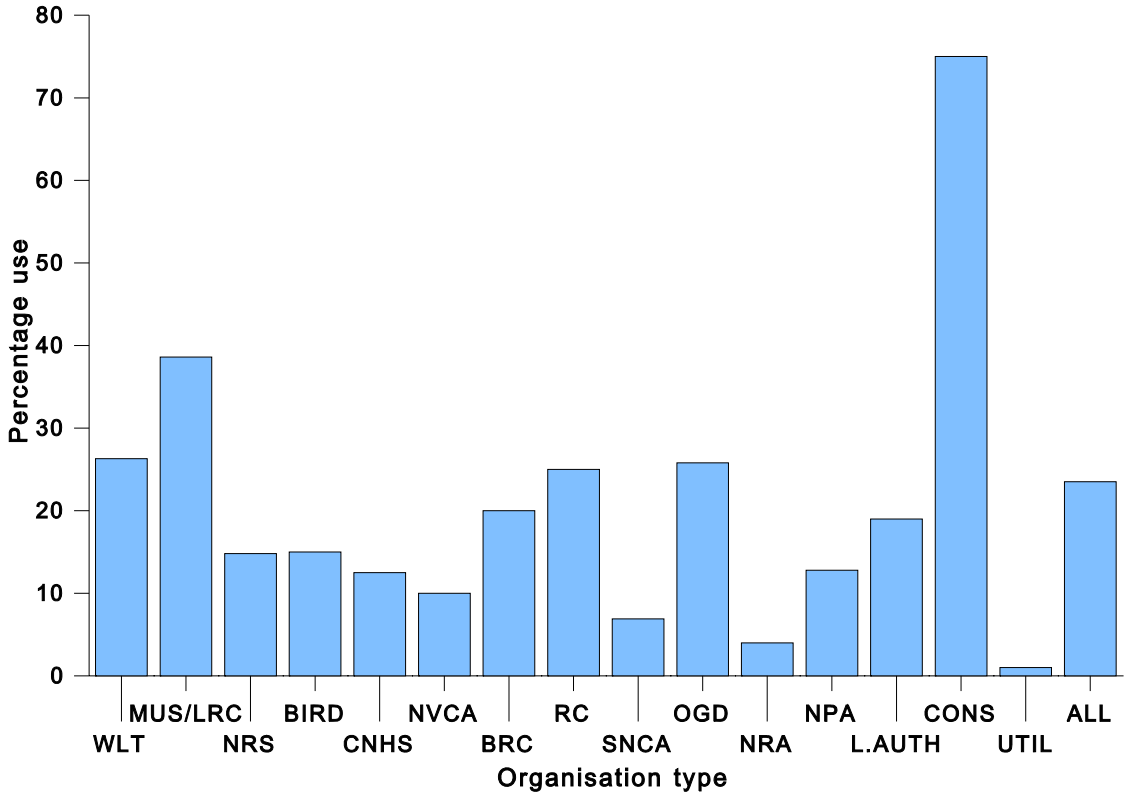
Figure 2.8.2 **Priority of purposes for which data are collated by local records centres (N=135)**



Key to abbreviations - see page 10

2.8.8 The primary users of data from biological recording are summarised in Table 2.8.2, showing a bias to use of data within the immediate organisations responding to the survey. Use of data by organisations other than immediately in-house or by the funding body (parent) is only 23.5% overall. Particularly low figures for external use of data are apparent for the NRA regions and the statutory nature conservation agencies (Figure 2.8.3), but this is not unexpected in view of their advisory and regulatory roles. The high percentage (38.6%) of external use of data at local records centre/museums shows their role in the supply of information, which fits the accepted role of local centres.

Figure 2.8.3 **Primary users of data (N=135)**



Key to abbreviations - see page 10

3 BIOLOGICAL RECORDS AND THE LAW

ABSTRACT

Legal obligations to make, compile and maintain records in the UK.
Ownership and intellectual property rights: copyright and moral rights, ownership of specimens and records, obligations on transferring intellectual property rights, Crown copyright, duration of copyright.
Compilations of records: ownership, copyright, potential liabilities.
Access to records: principals, application of Environmental Information Regulations.

3.1 INTRODUCTION

3.1.1 Knowledge of the law concerning the ownership, compilation and use of biological records does not appear to be widespread amongst the recording community. Moreover, recent European and consequential UK legislation concerning copyright and environmental information have changed the position somewhat.

3.1.2 A formal legal opinion prepared by Messrs Morrell, Peel & Gamlen is provided in Appendix 4, of which this chapter is a simpler précis, together with some additional material.

3.1.3 All the statements in this chapter and in Appendix 4 apply throughout the UK (i.e. England, Scotland, Wales and Northern Ireland) and up to 31st July 1994, unless indicated to the contrary. At present, the Isle of Man follows the UK in these matters and so do the Channel Islands but both have the legal capabilities to adopt different laws if they so wished.

3.2 OBLIGATIONS TO MAKE, COMPILE OR MAINTAIN BIOLOGICAL RECORDS

3.2.1 There appear to be no obligations under present national or European legislation, or through international agreements, which require any individual, organization or agency in UK to make, compile and maintain biological records. The nearest to such a requirement is in the Wildlife and Countryside Act 1981, where Section 43 requires county planning authorities covering National Parks to prepare maps of areas of moorland and heath of important natural beauty, and in the Environmental Protection Act 1990, which requires the JNCC to establish *"common standards throughout Great Britain for the monitoring of nature conservation and for research into nature conservation and the analysis of the resulting information*". However, the newly published Planning Policy Guidance note on nature conservation (PPG 9, para. 24) interprets the Town and Country Planning Act 1990 (sections 11 and 30) to state that local authorities have an obligation "*to ensure that the [local] plans are based on fully adequate information about local species, habitats, geology and landforms*" (DOE 1994b). Nevertheless, virtually all national and European legislation and international agreements imply that biological records should be made and maintained (see 1.3). Indeed, it is hard to see how some of the obligations could be met without so doing. It would not be unexpected if a more specific requirement were to come from such responsibilities under international conservation obligations, especially the Biodiversity Convention. Even so, this would not become binding until incorporated into national legislation.

3.2.2 The remarkable fact remains that, despite the widespread use of, and reliance on, biological records, whether for conservation, planning or other purposes, there is no explicit legal obligation for such records to be made, stored and maintained in UK for any purpose!

3.3 OWNERSHIP AND ASSOCIATED INTELLECTUAL PROPERTY RIGHTS

The nature of intellectual property rights

3.3.1 Intellectual property rights (IPR) are of two types - copyright and moral rights.

3.3.2 Copyright accords protection to a record in its permanent form, whatever that may take - as a written record ('literary work', however formulated - prose, coded, tabular, or spoken), an illustration of any kind ('artistic work', as drawings, paintings, diagrams or photographs), a sound recording, or broadcast, an electronic recording of any kind (including digital recordings) or a film. The Copyright Designs and Patents Act 1988, which applies to the whole of the UK (but not the Republic of Ireland), requires also that the work shall be original (i.e. originate from its author and be a product of his/her skill and judgement) and that the author meets a UK qualifying requirement, in effect, that the originator, whether an individual or organisation, is a British citizen, subject, domiciled, resident, or coming in some way under the jurisdiction of the Act and the publication of the work is in the UK, or as defined in the Act.

3.3.3 Copyright ensures that a substantial part of any record cannot be reproduced without the permission of the owner - 'substantial' meaning the essential part of the work. Prior to July 1st 1994, copyright lasted in the UK for 50 years from the moment when the record came into existence, or 50 years from the end of the year in which its owner dies. From July 1st 1995, the Copyright Protection Directive harmonises this period in the EU to 70 years from the death of the author or after the work is lawfully made available to the public.

3.3.4 Moral rights arise from the identification of the originator of the record as its author, i.e. who made the record, if it is published (in any form), broadcast or included in a cable programme service. They are of two kinds, paternity (i.e. the right to be identified as the owner) and integrity (i.e. the right not to have the record subjected to derogatory treatment).

Ownership of a specimen and ownership of the intellectual property rights of a record of the specimen

3.3.5 Care needs to be taken to distinguish between the rights of ownership of a specimen, as such, and the ownership of IPR in virtue of a record describing the specimen or giving information about it.

3.3.6 A specimen does not of itself constitute a biological record. However, if written or graphical information is attached to the specimen, although not necessarily its name, the actual recorder of such information is, normally, the owner of its accrued IPR under the law. The owner of the land does not own the IPR of the record unless he/she made the record, or unless this has been the subject of a written agreement between the owner and the record maker before the latter made the record, provided the record maker was not on the site illegally. Ownership of the specimen may lie with the owner of the land or, if it has been removed legally, with the individual who has removed it. However, the owner of a specimen *in situ* (e.g. a landowner) controls access to it, for which a charge or conditions may be made, including conditions concerning ownership of copyright information.

3.3.7 Ownership of a record, therefore, lies in the first instance, with the recorder and he/she, by virtue of the accrued copyright and hence IPR, controls the right to its reproduction. A widespread situation where it is not always understood that this condition must be met, is where the records made by members of a society are submitted by them and published in the society's publication(s). Copyright of each record is still vested in its originator.

3.3.8 Note that records made by an individual *as part of commissioned work* are owned by that individual not the commissioning agency unless written agreement otherwise has been reached prior to the

commission being undertaken. In contrast, if an employee makes a record in the course of his/her employment, then the employer is the owner, unless agreed otherwise.

3.3.9 In any case where the owner of the specimen has laid down written conditions concerning the ownership of the copyright of any records made of the specimen in his/her ownership, those conditions will prevail.

Transfer of IPR of a record

3.3.10 Copyright of a record can be assigned or licensed *only through a written agreement* signed by, or on behalf of the assignor, to another party. In the case of assignment, ownership and, therefore, IPR are transferred to the assignee; with a licence, ownership continues to lie with the record's originator although the licensee may be permitted, by the terms of the licence, to reproduce a record.

3.3.11 Moral rights can be waived, in whole or in part, in writing but can be neither assigned nor licensed. At present, it is not clear whether waivers should be sought from originators by compilers of databases but it would be a wise precaution to so act until the law has been tested.

3.3.12 Moral rights are important if a record is *reproduced in a form other than that in which it was originally published*. It makes no difference whether or not reproduction of a record has been agreed either through assignment or by a licence. *Unless further specifically agreed*, it is assumed the reproduction will be in the same form as the original. This does not absolutely preclude alterations being made but, whether or not a reproduction in some form other than the original is a derogation is, ultimately, a matter for a court to decide. Normally, common sense alterations will not cause such problems (e.g. replacement of a descriptive location by a grid reference, or redetermination of incorrectly identified material) provided they are done for a generally acceptable reason but, even so, failure to indicate who was the originator of the record, when this is important - say for reasons of priority - might be a derogation.

Crown copyright - a special case

3.3.13 Copyright of works, coming within the normal categories (i.e. made by the Crown or an officer or servant of the Crown) is vested in the Crown and lasts for 125 years unless published before the end of 75 years from the end of the year in which it originated. In this case, Crown copyright extends only 50 years from the end of the year of origination of the work.

3.3.14 It would appear that Crown copyright will apply to records made by MAFF, DOE, or other government departments *and by their contractors, unless otherwise specified*, but probably not to non-governmental public bodies (NDPBs) such as JNCC or the nature conservation agencies and probably not the research councils and their institutes. However, some reports of these non-governmental agencies are published by HMSO when they fall under the rules of Crown copyright.

3.3.15 Crown copyright can be assigned or licensed. Moral rights do not apply to works in which Crown copyright subsists.

3.4 CONSTRAINTS ON THE COMPILATION OF BIOLOGICAL RECORDS

3.4.1 A constraint on compilation may arise when a biological record is made which involves a graphical record, for example, on an Ordnance Survey, or similarly copyrighted, map, although citation of grid references is quite allowable. In the former case the approval of the originator of the map should be obtained and due acknowledgement made if the record is published or transferred beyond the record makers' possession.

3.4.2 Ownership of a collection or compilation of records may reside in the compiling person or agency but, unless permission has been given for their reproduction by the originator(s), the IPR continues to

reside with the originator(s). This is even the case when records are donated voluntarily. A verbal expression of intention is not enough, a written indication that the IPR have been assigned or licensed must accompany the donation. Hence, if it is intended to reproduce the compilation in any form for public use, a compiler will have to negotiate either an assignment or licence with each original owner as well as a waiver. If reproduction will involve a change in form, it is best to include this in any agreement. This requirement applies to all the original records whether in whole or in part. This is of particular importance when the data involved come from multiple sources some, or all, of which may themselves be compiled data and the compiler then creates a composite or synthesised product such as incorporating the data in to a GIS, or map. Note that not only are the originators of the biological data involved but, if part of the records are graphical, the kinds of consideration set out in 3.4.1 will also apply. Moreover, if a computer program is used to load, retrieve, or manipulate any of the data at any stage, it too will almost certainly have attracted copyright. In general, an acknowledgement of the use of such programs is sufficient. In case of doubt it is advisable to seek legal advice.

3.4.3 Copyright of a compilation, obtained and reproduced legally, may reside under UK law in the compiler simply in virtue of its compilation. However, in many European countries and the USA, a compilation must show true originality and creativity before it can acquire copyright as a literary work. Legally copyrighted compilations extend for the same duration as do original records. The EU has now proposed that originality shall be a requirement for copyright of all electronic databases and that these will be protected from extraction, i.e. removal and incorporation in another database, for 15 years. This is likely to become the European law and would then have to be assimilated into UK law. During the protected period, a compiler may licence extraction or utilisation of the compiled database under terms laid down by the compiler. It would be wise to maintain a watching brief on this, potentially, rapidly changing area.

3.4.4 It should be remembered also that there may be an international dimension to a data compilation if any part has been created in an industrialised country; the position of developing countries is variable. The Berne Copyright Convention and Universal Copyright Convention ensure that works created in industrialised countries attract copyright in the UK. Exposition of international copyright law goes beyond the remit of this report but it should be born in mind that GIS and computer software employed in the UK may well have been derived from such countries. Advice on international copyright law should be sought, for safety.

3.5 POTENTIAL LIABILITIES INCURRED IN THE USE AND COMPILATION OF COMPILED RECORDS

3.5.1 Compilers acquire a range of potentially serious liabilities in making data publicly available if no protective action is taken.

- **Owners' rights of copyright and moral rights will not be infringed if assignments, licences and waivers, including permission to change the form of the record if necessary, have been obtained from every originator, or legal intermediate owner(s).**
- **In addition, if records include identifiable, personal references (e.g. name and address) to living individuals the compilation will need to be registered satisfactorily under the Data Protection Act 1984. In the event of doubt, advice should be sought from the Data Protection Registrar.**
- **A compiler will need also to take preventative action against claims for negligence, i.e. an acceptable minimum standard will be expected. A situation of negligence could arise if the compiler had not taken a 'duty of care' to ensure the accuracy of the records, both in their compilation and in ensuring, so far as possible, the accuracy of the data received, whether from the originator or through an intermediate owner. Unfortunately this can be defined only as a result of case law. What is clear, however, is that an especially high degree of skill and care will be expected of a compiler who sells information and this will also apply if the compiling agency is recognised publicly as a national agency. Therefore, the best possible monitoring and verification procedures should be employed and publicly promulgated. In any event, a compiler should prepare appropriate disclaimers of liability and take out suitable protective insurance!**
- **Care needs to be taken also that confidentiality is not broken if unpublished records are included. An appropriate screening procedure should be incorporated in the compilation procedure.**

*** Liability extends to hardware and software suppliers as well as the data supplier and any one may incur part of the overall responsibility for the satisfactory operation of a system involving data compilation, interpretation and supply.**

*** Lastly, contractual liability needs to be assured. This is relatively simple, requiring strict adherence to the terms of any contract. Especial care is necessary when contracts include clauses specifying action to be taken to avoid negligence or to assure quality.**

3.6 ACCESS TO BIOLOGICAL RECORDS

3.6.1 The EC Directive on the Freedom of Access to Information on the Environment has now been implemented in this country by the Environmental Information Regulations (EIR) (SI 1992, No. 3240). The purpose of the EIR is to enable the public to obtain access to information concerning obligations under the Environmental Protection Act 1990 affecting pollution control, waste and litter, There is no doubt, however, that it will, *in principle*, be applicable to biological records since the phrase "the state of any flora and fauna , the state of any soil or the state of any natural site or other land" is included in the description of information relating to the environment.

3.6.2 The Crown, Government departments, local authorities and other persons carrying out functions of public administration at any level, provided they have responsibilities in relation to the environment amongst their functions, and any other body with public responsibilities for the environment provided it is controlled by a person in the categories previously defined are specially identified in the EIR. *Anything* in the *accessible* environmental records held by such bodies (including the country conservation agencies, JNCC and probably the research councils) must be publicly available on request. A reasonable charge may be made for such information. It is regrettable that neither a definitive list of all the relevant bodies is included in the EIR, nor are the terms 'accessible environmental records' or 'reasonable charge' defined: each organisation is advised to decide for itself on all these matters!

3.6.3 If information has been received in confidence, or if its disclosure would increase the likelihood of damage to the environment, then public access to it can be denied. These requirements apply both to original records and compilations held by the bodies described in 3.5.2. Other categories of confidential information are defined in Regulation 4(2) but most are unlikely to be relevant to biological records. However, it should be noted that information relating to, or the subject matter of, legal proceedings, and material in draft or working documents can be treated as confidential.

3.6.4 An important proviso is made in Regulation 4(3)(d) and (e). Information is confidential if supplied by a person who is under no legal obligation to supply it to any of the bodies described in 3.5.3 and who does not consent to its disclosure. Data supplied on a voluntary basis to an organisation such as a country conservation agency or local records centre can, therefore, only be made publicly available provided the record maker has no objection. It might, therefore, be necessary to obtain a waiver from the originator(s) if records are disclosed, to ensure that there has been no breach of duty of confidence.

3.6.5 Lastly, it should be noted that the EU has proposed that in the event of a dispute, it will lie with the supplier of information to prove that access should be denied.

4 FUTURE NEEDS IN BIOLOGICAL RECORDING

ABSTRACT

Needs for data implied by Government policies and legislation, needs identified in* Biodiversity, the UK Action Plan*, minimum information required about national and local resources of species and biotopes, assessing threatened species, measuring changes over time, evaluating the success of legislation in protecting species and biotopes, information for research, education and public information.

4.1 INTRODUCTION

4.1.1 Future needs in biological recording should be examined strategically, prompted by two basic questions.

*** What are the requirements for data?**

*** How can the requirements best be met?**

This chapter describes the requirements for biological records, based on current practices and predicted needs, and examines the problems and means for meeting those requirements.

4.1.2 Several uncertainties need to be considered in defining what data may be required and how the requirements should be met. These factors are relevant to all the main areas of use: nature conservation, development planning, environmental monitoring and research.

- **The optimum types of data, needed to carry out these and other functions, are not clearly defined.**
- **None of the present users of biological records has statutory obligations to collect or hold data, although some policies have an implicit need for biological records (see Chapter 3).**
- **None of the principal sources of records has statutory obligations to hold data.**
- **The legal constraints of copyright and liability, in relation to the collation, management and supply of biological records, are poorly understood by almost all concerned with biological recording (see Chapter 3).**
- **International, national and provincial legislation in planning, environmental monitoring and nature conservation will continue to develop in range and complexity and can be expected to increase the demand for data.**
- **The implementation of this legislation and its successful delivery by all types of agencies will develop and change and the agencies responsible for implementing legislation and monitoring its delivery will also continue to develop and change.**

4.1.3 Future trends in some aspects of legislation, particularly in its implementation and delivery, may be expected, for example by reference to *Biodiversity: the UK Action Plan* and to the developing policies of the statutory agencies concerned with nature conservation and other forms of land use and environmental auditing. It is essential not to focus inflexibly on current needs, but to consider also the potential uses for data beyond the applications presently recognised.

4.1.4 The different types of environmental legislation (e.g. planning, monitoring and conservation) are often interrelated in terms of data requirements, but these interrelationships are not necessarily understood by either the users or suppliers of data. The potential financial benefits of shared responsibility for data collation and management have been recognised by relatively few of the organisations involved.

4.1.5 When considering predicted needs, these must be considered in relation to weaknesses identified in the present system. It is not practical to meet many of the current and predicted requirements for data,

due to lack of resources, and unless a complete set of objectives is defined now, the resources needed for effective biological recording will continue to be underestimated. Although much biological recording has been dependent hitherto on specialists working in a voluntary capacity, future requirements for data should not be unduly influenced by current perceptions of the availability of data.

4.1.6 Data need to be structured and enhanced in ways that facilitate their interpretation and use. Data which need detailed interpretation by experts for every enquiry will be less accessible and less useful than those which can be used readily by non-experts in at least some circumstances. Data should be capable of being accessed in ways which permit users to undertake some interpretation themselves and to know the limits of the interpretation they themselves may be able to put on the data. This has important implications for the range and types of data required and the data management methods used. Also, data need to be structured and managed in ways which will facilitate access and exchange and, where appropriate, made available through use of public access computer networks.

4.2 NEEDS IMPLIED BY LEGISLATION AND POLICIES

4.2.1 All major forms of use of the environment have impacts on species and biotopes. These uses range from agriculture, forestry, military training and nature conservation to transport, energy production, manufacturing industries and housing development. The impacts may be adverse, for example by destroying sites or degrading environmental conditions through pollution, or they may be beneficial, for example by creating improved conditions through changes in land management.

4.2.2 Information about species and biotopes is essential to enable everyone concerned with the use of the environment to implement Government policy and to conform with legislation whilst undertaking their own legitimate activities. It has been noted previously (Chapters 1.2 and 3.3) that although UK Government policies and legislation on environmental matters, such as the Wildlife and Countryside Act 1981 (WCA 1981), the Environmental Protection Act 1990 and the Environmental Impact Assessment regulations, imply the need for data, including some types of biological records, there is no legal obligation for biological records to be made or maintained. The first time that a need for biological records was explicitly recognised in a statement of Government policies was in *Biodiversity, the UK Action Plan*.

4.2.3 Despite this absence of statutory obligations, the governmental nature conservation agencies, local authorities and others are charged with responsibilities which cannot be carried out without access to biological records. Without appropriate information, agencies will be unable to deliver to the nation as a whole, or to local communities, the results which, under UK and international legislation, they have been charged to deliver. Examples of the types of information are described in following sections (4.3 to 4.5).

BOX

***Biodiversity: the UK Action Plan* - the need for information on species and biotopes**

- **'The Government is formally committed under WCA 1981 to review every five years the status of wild plants and animals'**
- **'SSSIs will continue to be used as the basis for securing the conservation and enhancement of the best sites for wildlife. Further sites will be added where significant gaps in representation are identified'**
- **'Government policies support the conservation of sites through the requirements under the planning system and statutory consultation processes'**
- **Biogeographic areas should be identified as part of the development of strategic approaches to conserving biodiversity**
- **Implementation of several statements of intent in 'Progress towards objectives' will be possible only if the continued supply of up to date data on species and biotopes is secured**

4.2.4 *Biodiversity: the UK Action Plan* highlights current and planned Government policies, for example to protect and conserve species, to maintain a network of protected areas, and to foster biodiversity and sustainable use of the environment in agriculture and forestry, and in the management of fresh water and the marine environment. It also lists examples of specific targets for key habitats and species. Several paragraphs in the chapter on Targets and Monitoring include statements which reinforce the need for

information on species and biotopes. The *Action Plan* proposes actions (10.32) which have potential to address the issue; for example through the work of CCBR, the formation of the Biodiversity Action Plan Steering Group (10.39) and the working group established to improve the accessibility and coordination of biological datasets (10.40).

4.2.5 The implication of the present situation is that usable data on species and biotopes must be available for those concerned with environmental management, whether at the level of Government departments or individual land owners. Those data should be:
*** collected and structured for definable uses;**
*** derived from reliable sources;**
*** validated to agreed standards;**
*** stored using secure methods;**
*** made accessible to users.**

4.2.6 At present, the acknowledged sources of data on species and biotopes are not systemised and other potential sources are poorly recognised. Very few of these sources are capable of making data readily available to potential users. Considerable resources of data exist but they are dispersed widely, both geographically and administratively. Many of these sources are already used in support of environmental management, but they are accessed through *ad hoc* methods, and many exist on an extremely fragile basis, often only through personal commitment and goodwill.

4.2.7 In summary:
*** there are identifiable needs for information, but there are no assured means of supplying those needs;**
*** there are potential sources of information, but they lack an organisational framework and most are under resourced or are completely unfunded;**
*** this dislocated dichotomy is the central issue of this review.**

4.3 DEFINING THE RESOURCES OF SPECIES AND BIOTOPES

4.3.1 A primary need for biological recording is to help to define the national resources of species and biotopes. This basic information, an inventory of spatially referenced data on the national, natural biotic resources, is a fundamental and unavoidable element in providing the context in which national environmental policies are developed, implemented and delivered. Such national policies may be developed in wider international contexts, where again a baseline of national data is essential to enable the UK position to be expressed authoritatively. Similar data should be available at lower administrative levels, for example to enable national policies to be implemented and delivered at a local level and for local policies to be developed.

Species

4.3.2 For species, the minimum data about the resource could include:
*** a complete inventory of the species occurring in the UK;**
*** summaries of the geographic range of each species in the UK and in Europe;**
*** a summary of the biotopes with which each species is associated in the UK;**
*** an indication of the frequency of occurrence and threat status of each species;**
*** time series measurement of the resource.**
Inevitably, because organisms respond to environmental changes (often rapidly to quite subtle changes), these data cannot be collected once and thereafter used indefinitely as a static measure. Mechanisms to update the data and to monitor key elements are essential if these responses are to be identified.

4.3.3 It has been estimated that there are about 88 000 terrestrial and freshwater species and 8 000 marine species in the UK (*Biodiversity - the UK Action Plan*, p29). These totals exclude bacteria and viruses and underestimate the numbers of protozoa, algae and fungi. The Action Plan provides numbers which underestimate some other, better known groups, such as flowering plants, for which the current, national taxonomic guide (Stace 1991) lists twice as many (nearly 3 000) 'native, naturalised, recurrent casual and crop plants'. Most available 'national' lists of species cover Britain and all of the island of Ireland, rather than just the UK, but proportionately there are few species which occur in the Republic of

Ireland and not the UK, even if marine taxa are included. As Britain and Ireland represent a coherent biogeographic unit (in global or even European terms), there is a strong case for close cooperation in biological recording between the UK and the Republic of Ireland.

4.3.4 The flora and fauna of the UK is probably the best documented in the world (May 1993), but species are being added (and removed) from the UK lists as a result of ongoing survey and taxonomic and nomenclatural research; this is a continuing process which affects almost every taxonomic group. There is no national inventory of species, and no organisation is responsible for cataloguing the biological taxa present in the UK to uniform standards, although both BRC and the Recorder data management package have extensive lists compiled from a variety of sources. This uncoordinated approach means that there must be doubts about the integrity of the systematics and nomenclature used in the resultant lists. There is no official national centre for information on the occurrence of species. The Biological Records Centre at ITE Monks Wood holds the most comprehensive biogeographical database in the UK with over 6 million records of more than 9 000 species, but even this includes information on only about 10% of the taxa known to occur in the UK. Similarly, such information about the ecological requirements of species as exists is dispersed and patchy (see 2.3.31).

_____BOX_____

Officially recognised Red Data books/lists covering all Britain and/or all Ireland*

Taxonomic group	Date	Coverage
Mosses, liverworts and lichens	In preparation	Britain and Ireland
Stoneworts (Characeae)	1992	Britain and Ireland
Flowering plants	1977, 1983 & in preparation	Britain
	1988	Ireland
Insects (some orders only)	1987	Britain
Invertebrates other than insects (some groups only)	1991	Britain
Birds	1990	Britain
Vertebrates	1993	Ireland

*** Unofficial and regional Red Data lists/books are not included**

4.3.5 Methods to measure the frequency of occurrence of species differ for many taxonomic groups and authoritative information exists for only a small percentage of our rarest and most threatened species. Estimates of the extent to which species are under threat of extinction regionally, nationally or internationally (Red Data List status), or are scarce and therefore potentially under threat, are based on knowledge of factors such as geographical range, temporal changes in range, ecological requirements, frequency of occurrence and perceptions of the causes of decline. Official British Red Data lists have been compiled mainly for the more popular groups (birds, flowering plants, and some insect and other invertebrate groups), most of which are based on standard criteria. Recently JNCC has collaborated with the statutory bodies in Northern Ireland and the Republic of Ireland in the preparation of Red Data books to cover all of the British Isles. Other national and local Red Data lists/books have been prepared using a range of criteria. Reliable assessment of the *threat* status of species also requires measurement of temporal changes of at least the range and/or frequency of species and long term data to assess natural variability. These data can be derived from surveillance and monitoring or, because few long term datasets exist, by reference to sources of historical data.

4.3.6 The need for information on the national or local resource of species can be expressed as six simple questions:

*** What is it? (Taxon)**

*** Where does it occur? (Geographical range)**

*** Where does it live? (Ecological requirements)**
*** How many are there? (Estimate of population)**
*** Is it threatened and, if so, how? (Measurement of threat and causal factors)**
*** Is it changing in any respect? (Time series information)**
Due to their uncoordinated and dispersed nature, it is difficult to obtain these types of information efficiently, even for those organisations with statutory responsibilities to protect species. Failure to develop appropriate policies and the inability to deliver detailed and summarised information into the public domain are direct consequences of this lack of efficient access to basic information.

Biotopes

4.3.7 Biotopes are the ecological matrix in which species occur. Some biotopes, such as woodland dominated by oak, ash and hazel, are widespread, others, such as inland mobile sand dunes, are small, localised and inherently important. Thorough assessment and regular measurement of changes in the resources of biotopes will provide essential indications of changes which have affected or may affect species. For biotopes, information about the resource should include:
*** an inventory of the biotopes occurring in the UK (based minimally on a nationally agreed classification);**
*** a summary of the geographic range, area and frequency of each biotope in the UK and in Europe;**
*** an assessment of the key threats to biotopes and their capacity to resist threats without loss of quality or range;**
*** time series measurement of the resource to monitor the impact of threats.**

4.3.8 The biotope resource of Britain has been described using many different surveys and classification systems based mainly on types of land cover, land use or vegetation. The results of some major surveys in Great Britain have been published recently, for example the Countryside Survey 1990 (Barr *et al.* 1993), the satellite Land Cover Map (Fuller *et al.* 1994) and the National Vegetation Classification (Rodwell 1991-). DOE commissioned a review examining 17 land cover surveys and classifications, which proposes a standard framework of reference allowing them to be compared (Wyatt *et al.* 1994). DOE and the statutory conservation agencies are supporting the development of the Countryside Information System (CIS) (see 2.5.134) and NERC has developed UKDMAP (see 2.3.39) which provides an analogous data resource on the marine environment.

4.4 MEASURING CHANGE AND EVALUATING THE SUCCESS OF LEGISLATION

4.4.1 Implicit in present legislation on the environment, where it affects species and biotopes, is the need to measure changes over time and to evaluate the success of legislation and consequent actions in delivering results. These measurements are being made at a national (UK or GB) or country scale. Similar measurements are required at a more local scale. A small but growing number of local authorities, principally at the county level, are conducting environmental audits which include the sustainable use of the countryside and the need to conserve and monitor wildlife. In some of the most recent audits the role of local records centres in supplying reliable wildlife data has been recognised. In recent structure plans, many local authorities (using a variety of terms) have recognised *sites of importance for nature conservation* (SINCs) and the need to obtain data and to monitor and review these locally designated sites (see Collis & Tyldesley 1993). The principle of designation and management of SINCs is supported by EN (Cranbrook in Collis & Tyldesley 1993) and this informal type of site designation is recognised in PPG 9 (DOE 1994b).

4.4.2 At present this type of information is summarised annually in the Digest of Environmental Protection and Water Statistics, The Scottish Environment and the Environmental Digest for Wales. Measurements of longer term changes were included in *The UK Environment,* published in 1992, drawing on a very limited range of pre-existing datasets, mainly for birds. There is a clear need to develop means for collecting and analysing data to provide annual or periodic wildlife statistics. Such a wildlife reporting procedure should provide the following:
*** an unbiased overview of the UK wildlife resource;**

*** assessments of change and stability in the resource;**
*** information relevant to current and future policy issues;**
*** information and interpretations of data which can be understood by, or which can be interpreted for the general reader.**

4.4.3 When assessing change and stability of species, assemblages and biotopes it is essential to prioritise work using units which can be confidently predicted to respond sensitively to environmental variables on short time scales. It is equally important to maintain the assessment of long term changes, for example to differentiate natural variations from progressive trends identified from short term data. It is uncertain whether adequate information of either type is available at present. However, as a result of the large investment of resources in collecting data on birds (and the size of the existing ornithological datasets), it is inevitable that birds will feature significantly in any future wildlife reporting procedure. However, the temptation to regard birds or rare vascular plants as paradigms for most other taxonomic groups and assemblages must be resisted because their ecological requirements differ markedly from, for example, mycorrhizal fungi, epiphytic lichens, grassland butterflies, saproxylic insects and freshwater molluscs.

4.4.4 The assessment of change at the scale of biotopes is already being addressed by DOE, through recent support for the Countryside Survey 1990 and the ITE Land Cover Map, and through active consideration of follow up surveys at suitable intervals. It is at the detailed scales of local authorities, that more consideration needs to be given to the assessment and monitoring of local resources. There are opportunities to promote and develop collaboration between the statutory conservation agencies, local authorities and the voluntary conservation organisations to initiate and undertake biotope resource surveys, to establish local priorities in relation to national surveys and partly to 'ground truth' the national surveys. This is particularly important in local planning where a biotope which is not threatened locally may be declining nationally, or vice versa. Much local information already exists, such as Phase 1 habitat surveys (Wyatt 1991), and can be updated, but it is not necessarily available to other agencies with potential interests.

4.5 RESEARCH, EDUCATION AND PUBLIC INFORMATION

4.5.1 The Government has a policy to support both environmental research (in addition to that related to nature conservation) and environmental education (see for example the 1990 White Paper *This Common Inheritance* and *Biodiversity: the UK Action Plan*). At the tertiary level, research may form part of the education process. Research, in particular, may result both in the acquisition and the use of biological records. The research applications of biological records range from very detailed site or population based studies (e.g. Warren 1994), through to the use of highly summarised data, for example in biogeographic studies (e.g. Prendergast *et al*. 1993).

4.5.2 NERC has maintained a national Biological Records Centre since 1964, with additional support from successive statutory conservation agencies since NCC became separately funded in 1973 (Harding & Sheail 1992). Since 1973, NERC's primary interest in the BRC database has been in the research applications of the data, in baseline documentation of the biodiversity of the UK (including the production of distribution atlases) and in examining environmental changes, such as the potential effects of global warming (e.g. Elmes & Free 1994) and agricultural policies (e.g. Firbank *et al.* 1994).

4.5.3 Because at least 70% of taxa records (and some 20% of biotope and land type records) are derived from voluntary sources, there is an identifiable need to provide the suppliers with results from their work. Quite apart from the legal obligations of users to their suppliers (see Chapter 3 and Appendix 4), there is the moral obligation to provided volunteers with feedback and reward for their contribution of knowledge and expertise.

4.5.4 It is essential to recognise that the majority of data in biological recording is supplied by volunteers and non-governmental organisations, who have justified expectations of something in return for their work, although this need not necessarily be a financial return. It could include rights of ownership to the data they have contributed, regular feedback (proportionate to their contribution) and, in particular, rights of access to collated and summarised, non-confidential information in forms and by methods that are readily available. Some volunteers are active in publishing results from their work, for example in

specialist journals, where the information enters the public domain. However, most volunteers appear to contribute data for a perceived common good. For example, they are interested in safeguarding species and sites, and in publishing local and national summaries, such as atlases and handbooks, thereby continuing the well recognised traditions on which biological recording was established and has developed in the UK.

4.5.5 The use of biological records in education is mostly indirect, mainly through publications such as handbooks and atlases and in scientific journals. Direct use of data in education is limited mainly to higher education, where inevitably there is overlap with use in research.

4.5.6 Interest in the flora and fauna of a county or district is increasingly being fostered through local museums and wildlife trusts. In addition, a variety of *ad hoc* schemes has arisen to further the documentation of the biodiversity of local regions or biotopes. Biological records form essential contextual information for exhibits, publications and education programmes (for all age groups and specialisms) based on these local centres.

4.5.7 Despite these applications of biological records, there is no explicit statutory need for their use in research and education, but without them much work related to biodiversity would be either impossible or based on inadequate information.

4.5.8 *Biodiversity: the UK Action Plan* proposes that there should be greater public access to information on biodiversity. Consideration is given to a charging policy for derived data, the provision of summary statistics and information to a wide user community through information centres such as museums, libraries and schools.

4.5.9 There is an important educational issue in relation to the supply and the use of biological records. Supplying data on species and sites to developers is still perceived by many volunteers as favouring 'the enemy' of conservation, whereas planning legislation and practice are intended to protect species and sites in a context of sustainable use. Greater access to, and use of, reliable and up to date data should result in more reasoned decisions in both planning and conservation, and better research. Volunteers must be involved, through education, in the broad range of uses of their data, the need for active participation in recording and the need to continuously improve the data resource. Although public involvement in environmental decision making is a wider issue, the potential role of these experts in biological recording should not be overlooked.

5 THE ESSENTIALS OF, AND POTENTIAL FOR, A NATIONAL SYSTEM FOR BIOLOGICAL RECORDING

ABSTRACT
Summary of thinking behind the need for a national system and of the current state of biological recording in the UK; hierarchical relationships between existing organisations.
Potential models for a national system, particularly in the USA and Australia, and their relevance to the UK situation.
Government policies and statutory needs for biological recording.
Strengths and weaknesses of the present situation for developing a national system: potential functional units, data interchange, coordination and regulation, metadata about components, phased implementation.

5.1 INTRODUCTION

5.1.1 In the light of the evidence presented in the preceding chapters, this chapter examines the opportunities which exist to develop biological recording in ways which will provide for the information requirements of the 1990s and beyond. Do existing organisations and *de facto* information networks provide a basis for establishing a national system or systems? Alternatively, it might be preferable or necessary to establish a national system *de novo* and, if so, do existing systems outside the UK provide appropriate models?

5.1.2 There has been discussion of the concept of a national network in biological recording since at least 1970 (Perring 1971) but the first detailed examination of the need for, and feasibility of, a national system was made in 1987/88 by the Linnean Society working party (Berry 1988). The working party report suggested that NCC or DOE should take a lead in setting up a coordinating commission and that BRC at ITE Monks Wood and the Rural Areas Database at Essex University should act as central agencies in the transfer of data. These proposals underestimated the need for information to flow in many directions, the presence of existing networks, local demands for data and the ephemeral nature of some organisations.

5.1.3 For a viable national system to be established successfully not only must the needs, referred to in Chapter 4, be met but there must be clear benefits for all the organisations and individuals likely to be involved, whether as suppliers, managers or users of data. This implies the development of more effective networking to supply information both for local use at the local level and to supply information for other uses at country and national levels, often in synoptic forms. The supply of some detailed data, for example on internationally threatened species, will have to be effective from the local level through to a European or wider international level.

5.1.4 Proposals for a national system must be seen in the context of the 1992 Convention on Biological Diversity and the UK response in *Biodiversity, the UK Action Plan* (Cm 2428) and *Sustainable*

Development, the UK Strategy **(Cm 2426). Particularly relevant is the Government's commitment (Cm 2428 para 9.50) to arrange a feasibility study on the development of a UK Biota Database to advise on data requirements, accessibility standards and protocols, data management, technical options and costs. This review by CCBR, initiated before the Biodiversity Convention, can serve as an important contribution to the proposed feasibility study by Government. In preceding chapters, it has provided the most complete account available of the present situation and, looking to the future, the strategic steps necessary to ensure an effective national system are set out in the rest of this chapter. The specific requirements to establish a national system are then developed in Chapter 6 and, finally, recommendations for action, where possible, are listed in Chapter 7.**

5.2 SUMMARY OF THE CURRENT STATE OF BIOLOGICAL RECORDING IN THE UK

5.2.1 The UK has a good but far from complete record of its national biological diversity. Good time-series data exist for some taxonomic groups and for land cover, which enable changes to be measured meaningfully, in few cases over a period of 50 years or more. However, little of this resource of spatially and temporally referenced information has arisen as the result of conscious national policies or commitments to document biodiversity and measure changes. Inevitable consequences of the diverse origins of this information are that the data are of variable quality and in a wide range of formats. Nevertheless, a national strategy for biological recording does not start from a position of minimal knowledge and could be developed from the present situation, as revealed by the Survey and described in detail in Chapter 2.

5.2.2 Some key points of relevance to a national scheme are described below.

- **Agreed basic methodologies or standards exist for very few types of biological recording.**
- **The existing national data resource of geographically referenced data on taxa is well in excess of the 63 million individual records documented in the Survey and in reality may be in the region of 80 to 100 million records.**
- **The number of records of biotopes and land types which are referable to defined sites probably exceeds 5 million records.**
- **The resources of data on taxa are held mainly by non-governmental bird organisations (44% of the total of 63 million taxa records), local bird clubs (14%), NERC and national biological recording schemes (12%), local records centres and museums (11%), BSBI (9%) and statutory conservation agencies (3%), but these figures are distorted by the very large number of ornithological records. Similar organisations are responsible for most data on biotopes and land types.**
- **The present resources of records of taxa are substantially dependent on the work of amateur naturalists rather than paid surveyors. Records of habitats and land types are compiled mainly by paid surveyors.**
- **The contribution made by volunteers and amateurs to the collection of biological records is a notable strength in the UK. However, under present conditions, dependence on these sources must be regarded as insecure because demand is now exceeding supply, while there is no obligation on volunteers to provide data and few resources to enable them to do so efficiently.**
- **The contribution made by volunteers allows scarce government resources to be directed to the collation, validation, storage and interpretation of the data and subsequently to the dissemination and use of the results. However, on present evidence, these resources are insufficient for these tasks to be adequately and effectively performed and many important datasets remain inaccessible to potential users. In addition to their contributions of data, in some cases (e.g. BTO) volunteers underwrite the costs of work commissioned by governmental agencies through their subscriptions and donations to societies.**
- **The existing data resource is dispersed between as many as 2000 functional units in the UK which hold or use data at local, regional, country or national levels.**
- **Local and regional units include:**
  local records centres;
  wildlife trusts;
  museums;
  local offices of the statutory conservation agencies;
  local authority planning departments;
  county/regional biological specialists (e.g. BSBI vice-county recorders, birds clubs and natural history

societies);
individual naturalists.

* **Country and national units include:**
  statutory nature conservation agencies;
  other governmental bodies (e.g. DOE, NRA, MAFF, Forestry Authority, Ministry of Defence);
  non-governmental environmental conservation and research organisations (e.g. WT, BTO, WWT, RSPB,
  NT, NTS,);
  national biological societies and recording schemes;
  governmental research organisations (e.g. NERC including ITE, IFE, BRC, BBSRC).
* **There is no designated national biological data centre in the UK.**
* **Data units and the users of biological records form a loosely structured, stratified hierarchy connected, in part, by several existing, *de facto* networks which already convey data from suppliers to a range of users, but most operate almost entirely independently of each other (Figure 5.1).**
* **Technological advances, in particular the use of relational databases to manage and disseminate data, and GIS to manage and interrelate spatially referenced datasets, are under-utilised in many areas of biological recording. This is due in part to the costs of this technology and lack of technical expertise, but also to a lack of coordination. Similarly, computer networks are under-utilised, especially by the voluntary sector but also by government agencies, partly due to the costs and difficulties of access to national networks.**
* **Agencies which obtain at least some of their funding from central governmental sources already support much of the infrastructure of biological recording, apart from data collection by volunteers. However, funding for biological recording comes via several departments and its final use is determined in a large number of governmental and non-governmental organisations.**
* **The wide ranging legal aspects of biological recording are poorly understood by most of the individuals and agencies involved.**

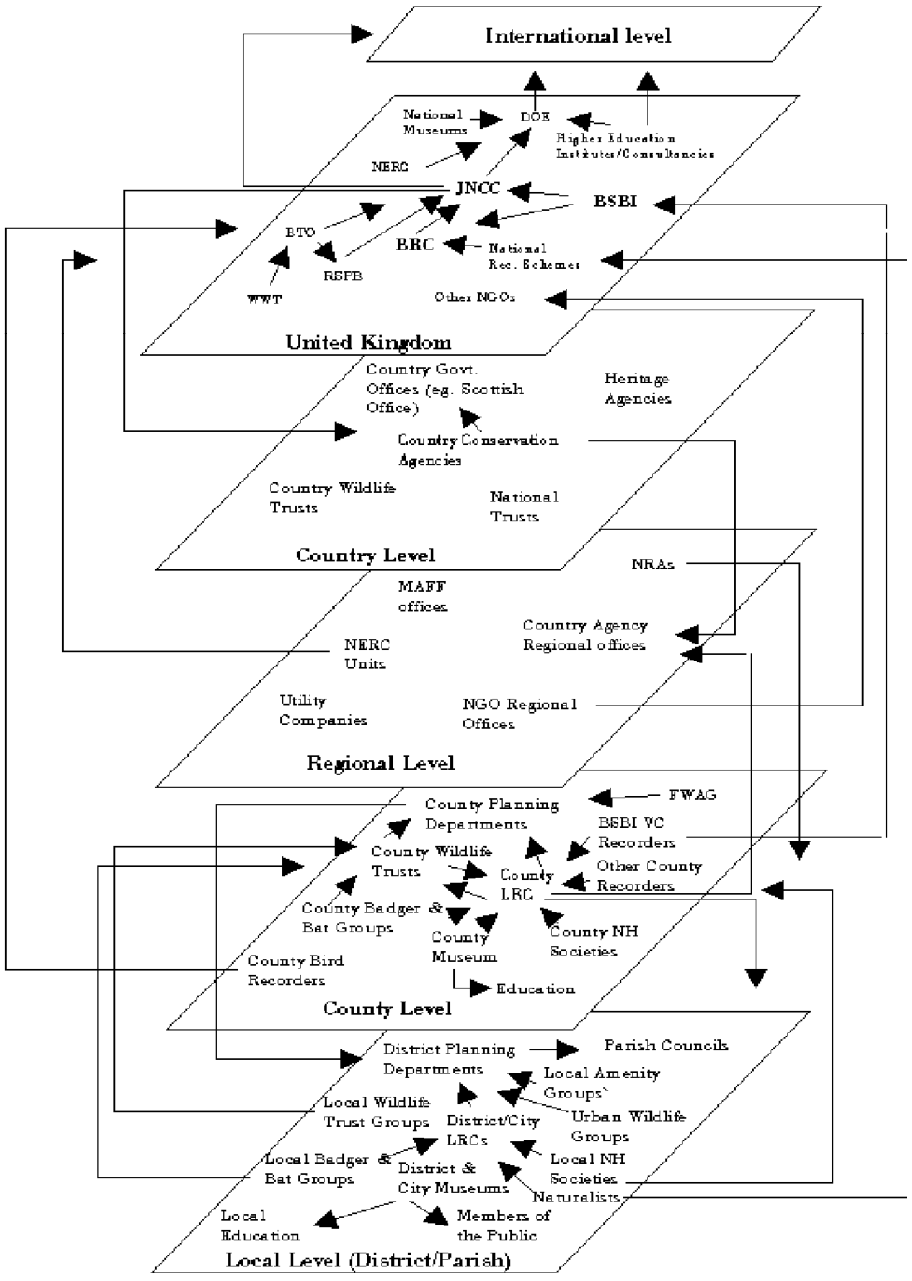
5.3 MODELS FROM OTHER COUNTRIES

5.3.1 Biological data centres have been set up in several European countries and in some there is more than one centre. In 1985 the Council of Europe published a list of European 'data banks in the field of nature conservation', which was to be updated with a survey in 1988, but the update was never completed. Designated national biological data centres have been set up by government departments in Denmark, France, Netherlands, Sweden and Finland. The majority of other data centres are based within museums or universities and are often confined to records of actual specimens held in collections. The extent to which any of these data centres have been established, or are now maintained, as a direct result of legislation or governmental policies is unclear. There are few comprehensive descriptions of the operation of most of these data centres, but none operate as part of a dispersed national information network established through national legislation. A working system of public access to environmental information has been set up in the Republic of Ireland with an information 'shop' in central Dublin (ENFO) which has a strong educational role. However, it seems to have few biological data and, presumably due to inadequate resources, has difficulty in providing access to interpreted information based on complex data. Only in Belgium is there a situation which is similar to that in the UK, where a small federation of biogeographic data banks has been formed on an entirely voluntary basis (Dufrene, Lebrun & Rasmont 1992). However, the federation is mainly of individual specialists, academics and small specialist groups, not local governmental organisations.

5.3.2 In the USA, a national biological survey was proposed as a result of a conference in 1985 organised by the Association of Systematic Collections (Kim & Knutson 1986), but this proposal has progressed slowly. In his Earth Day speech in April 1993, President Clinton announced the creation of a new national biological survey within the Interior Department. The survey was set up in November 1993 and early proposals suggested an annual budget of $180 million with more than $100 million coming from the US Fish and Wildlife Service, drawing on 1600 employees seconded from within the Interior Department. Its duties include an inventory and monitoring of biological resources, and the preparation of complete inventories of all US plants and animals is envisaged.

5.3.3 Two important national computerised systems do exist and represent different approaches. These are The Nature Conservancy's Biological and Conservation Data System (BCD) originating in N. Carolina in 1974 and now covering the USA and beyond, and the Australian government's ERIN system.
5.3.4 The private US conservation organisation, The Nature Conservancy (TNC), has established a Natural Heritage Data Center Network. In 1993 there were 85 data centres throughout the western hemisphere including

Figure 5.1 **Schematic diagram of the hierarchical relationships between organisations involved in the collection, management and use of biological records.**
Owing to the complexity of data flows, it has been possible to show only a selection of the data flow links to illustrate the relationships within and between the various levels.



all 50 US states, several US Bioreserves, some National Parks and National Forests, Puerto Rico, three Canadian provinces and 13 countries in Latin America and the Caribbean, with the long term objective of extending to the whole American continent. A collaborating centre has been set up recently in New Zealand. In the USA, the Network is based on the Natural Heritage Programs and Conservation Data Centers, most of which are based on agencies and departments sourced in the individual states, National Parks or National Forests. The operation of the Network, is based on partnership between voluntary, State and other official organisations, with over 400 biologists and computer technicians dedicated to the combined effort. Although voluntary bodies collaborate, mostly operating within individual states, the balance between their contributions and professional contributions is heavily biased to the latter. The Network and its component parts concentrate their efforts 'on species and natural communities of special concern' (Anon 1993a). Although the system has developed piecemeal over several decades and is uneven in its effectiveness, its has established standards for the activities of the component parts which are widely accepted and applied. Access to the data held on the network is determined locally by each centre and at national and international level by TNC headquarters. Consideration is now being given to the global public release of some information through INTERNET.

5.3.5 The data centres collect a wide range of environmental information including species biogeography, communities, population trends and ecology together with non-biological information on the distribution of parks, reserves and other managed areas, land ownership and socio-economic data. All the centres use a common PC-based program (BCD) written in Advanced Revelation which ensures a common data structure between programs allowing for relatively easy exchange of information as well as data aggregation and analysis across political boundaries (Anon 1993b).

5.3.6 The BCD , starting from a single programme in 1974, is written and supported by TNC headquarters staff and is a major element in the strategy for promoting common standards throughout the network. The file structure of the BCD reflects the types of information that the centres collect. These include: sources, elements (e.g. species), element occurrences (a population at a given locality tracked by observations on populations, records from museum collections etc.), sites, managed areas, tracts (land ownership), projects, transactions (land sales etc.), taxes, actions and contacts. Not all parts of the database need be used at any installation and there are further optional files available for those centres that wish to customise their systems for other data. The system is therefore flexible whilst maintaining an overall level of standardisation. The BCD is kept under review by an Operating Procedures Group which responds to users comments and assesses changes in software and technology, one current area of interest is improving the interface to GIS. Maintenance of the distributed network of centres and provision of training is a major undertaking and requires support both centrally and from regional information managers and project coordinators.

5.3.7 BCD uses both menu-driven and command line controls and users need to develop a good understanding of the underlying data structure and the operation of Advanced Revelation, to control both the input and extraction of information. This approach can be intimidating for novice users especially because it lacks clear user specifications. TNC defend the approach adopted for BCD with the argument that biodiversity and nature conservation is a complex business and its practitioners should understand the data they are handling. Nevertheless, because it has developed over several years, it is cumbersome and can be very slow, and is not particularly 'user-friendly'.

5.3.8 BCD is a very powerful tool which is successful because it has been created as part of a clear strategy within an influential organisation that is committed to its professional long-term support and development. The most significant difference from the UK situation is that the BCD has been set up independently of the Federal Government, by a well funded non-governmental organisation.

5.3.9 The Environmental Resources Information Network (ERIN) was set up as a result of decision by the federal government in 1989 (Slater & Noble 1991). It is a programme within the Commonwealth of Australia Government Department of Arts, Sport, the Environment, Tourism and Territories and is administered by the Australian National Parks and Wildlife Service. It is intended to provide an environmental knowledge base to aid planning and conservation at a governmental level, drawing together and upgrading information on the distribution of endangered species, vegetation types and

heritage sites. In its first year (1990/91), it had a budget of $2.1 million. ERIN has two distinctive features: it is simple, having been set up by and within a single government department, and its network covers only the agencies within the portfolio of that department, both as suppliers and users of data. The ERIN network has links to other agencies, such as the National Resources Information Centre, specimen-based data from State and Federal herbaria, museums and official expeditions, the Australian Surveying and Land Information Group, as well as other State, Commonwealth, NGO or other organisations. A key feature of the ERIN system is that it does not manage original data itself, but accesses data managed by others to provide environmental information. These features reflect a very different situation to that which exists in the UK for biological recording.

5.3.10 From the outset the use of 'state-of-the-art' networked computer technology together with GIS was central to the strategy. The technology chosen had to comply with the Australian Government's Open Systems policy and therefore had to be UNIX-based. Software includes ARC/INFO GIS and ERDAS remote sensing together with other statistical and modelling programs and the main database software used is ORACLE with information held as a series of ORACLE tables. The datasets being assembled include a data dictionary and catalogue to keep track of datasets with details of attributes, codes used and sources. There is also a taxon module holding standard names and descriptions of Australian fauna and flora and further modules of specimen observations, managed areas and a management information system. The managed areas module has details of national parks and nature reserves classified into 45 different classes. A directory of environmental information, experts and references (FINDAR) has been established which gives network users access to 130 environmental datasets, more than 1000 taxonomic experts and a comprehensive bibliography of publications and maps.

5.3.11 The ERIN project is an exciting and powerful demonstration of an integrated environmental information network. It differs from the American TNC model in that it is exclusive to a close knit group of government funded organisations and the technology chosen to power the system is technically more difficult to manage and more expensive to run.

5.3.12 Neither the USA nor the Australian situations have close similarities with that in the UK. That in the USA provides a powerful demonstration of how a system to provide well designed information can act as a catalyst for cooperation and networking and therefore provides a possible working model for a national system based on voluntary partnership. However, BCD would not be readily applicable in the UK because the whole structure of conservation, biological recording and relationship between agencies is very different in the UK from that in North America. It is important, however, as an example of bringing together information from a wide variety of sources (e.g. 20% of extant BCD biological records are from museum collections) and as a practical way of introducing standards that simplify data exchange and sharing. The ERIN model, with its top-down approach based in a single Government department, would be applicable to the UK only as a *de novo* system which could possibly embrace only Government departments and agencies, research councils and higher education establishments. Such a model would effectively exclude most of the potential voluntary suppliers of data and most of the local users, both of which are essential to comprehensive UK network. Thus the single largest source of records and recorders would not be used, or at best, the potential to make existing activities more efficient and to benefit the whole concerned community would be lost. An important lesson from both BCD and ERIN is the use made of data from museum and other collections, an aspect which is neglected in the UK, despite the very real potential resources of data. An important difference between BCD and the situation in the UK is that the former is properly centrally funded and supported as part of government policy.

5.3.13 A number of undeniable factors mean that the UK is facing a more complex situation than elsewhere. These factors include:

- **The number of agencies involved with biological recording;**
- **The lack of funding or the tortuous routes for funding these agencies (although much of the funding ultimately comes from central government);**
- **The absence of legislation to establish biological survey and inventory at a national level as the responsibility of any agency.**

A national system should exploit the positive features of the situation of which, the most important is that the UK is probably far richer in information and sources of information than elsewhere. Moreover, Figure 5.1 demonstrates that the potential exists for knitting together existing activities more effectively to

cover all levels of information than in any existing system. Over time, with a clear policy, agreed standards, improved methods of storage and exchange of data, and building on existing positive features, an opportunity exists to create a national system at least as good as the two just described and, potentially, providing more and better information. Such an approach seems likely to be more acceptable to all participants, statutory and voluntary, than importing novel systems which would require extensive and costly modification to adapt them to the UK situation. Evolution rather than revolution would seem the most rational course to adopt in the UK.

5.4 A BASIS IN POLICIES AND LEGISLATION

5.4.1 Both the BCD and ERIN have a firm basis in established policy and, in the case of ERIN, legislation. A national system based on explicit Government policies, such as those outlined in *Biodiversity, the UK Action Plan*, if necessary reinforced by legislation, would enable biological recording to develop most effectively the integration and coordination which is presently lacking. However, the actual form of the legislation required lies outside the scope of this review. The present dispersed and uncoordinated situation might be able to be adapted by negotiated agreement, but work towards such an agreement would need leadership and facilitation by a central organisation, such as a government department, with a commitment to involve all types of organisations involved with biological recording in the UK.

5.4.2 An important issue related to policies, negotiated agreements and legislation is the absence of formalised status for biological data centres (both local and national), which must be considered as fundamental components of a national system. Local records centres already exist in many areas, but the potential viability of each centre as a component of a national system, would have to be assessed against established criteria (see 5.7) There are some similarities to the management of archaeological information in England and Wales in that locally based Sites and Monuments Records (SMRs) were established in England and Wales from the 1960s onwards. Their authority and utility were enshrined retrospectively in the England and Wales General Development Order (Statutory Instrument No. 1813, 1988). Although SMRs are not yet a statutory function of local authorities, the Department of the Environment's Planning Policy Guidance note on Archaeology and Planning (PPG 16) describes the preferred content on an SMR and promotes consultation of SMRs as a first step in early consultations between developers and planning authorities.

5.4.3 Any agreement or facilitating legislation should be based on national policies for all aspects of information on the biodiversity of the UK and should address the following:

- **The need to acquire, manage and disseminate data;**
- **The development and auditing of national minimum standards for the operation of data centres;**
- **The establishment and maintenance of a national system of accredited biological data centres;**
- **The development of a common medium for access to and the exchange of data;**
- **The establishment of freedom for the interchange of non-confidential, and non-interpreted data throughout the system;**
- **The establishment and maintenance of metadata about the national system, with open public access.**

5.4.4 The National Heritage Data Center Network in the USA provides a basic model for a partnership system once a negotiated agreement or facilitating legislation has been established. As in the USA, it would be essential to bring together the majority of relevant agencies, to work co-operatively, not competitively, and to follow agreed standards. In particular, the active cooperation of local or regional authorities with voluntary and national bodies would be essential. In present circumstances there is no adequately funded agency in the UK, public or private, which could take on the initiating and coordinating role assumed by The Nature Conservancy in the USA. However, this function could, in principle, be carried out by an existing organisation, such as the statutory nature conservation agencies or BRC, but this would require additional funding and powers, but neither is concerned with biological records, in planning and land use, in the ways that DOE is in the UK.

5.4.5 An important aspect of the debate about statutory needs for biological recording should be the types of data needed for a range of applications at a national or regional level. Data are needed to provide national and regional statistics of change, annually or at longer intervals, as well as, for example, to provide broad overviews and means of assessing the threat and scarcity status of taxa and assemblages.

There is a clear time frame for the development of a reporting system for the EU Habitat and Species Directive. The need for data on species and biotopes is fundamental for evaluating sites for development control and environmental assessments and statements.

5.4.6 The distinction between survey, surveillance and monitoring has been made earlier (2.4), and in *Biodiversity, the UK Action Plan,* the need for monitoring (or at least a programme of planned surveillance) was stressed as a means of auditing the success of measures to preserve biodiversity. It is, and is always likely to be, impossible to monitor closely either the whole fauna and flora, or even representative species of all major taxa in the UK. Monitoring of restricted groups can be valuable and the very limited amount undertaken in the last 25 years has concentrated on birds, and to a lesser extent on butterflies and moths and some vascular plants. The results of this very small amount of monitoring provide clear measurements of the extent and rates of changes in these groups, which are repeatedly cited. However, the real differences in response to environmental changes, which are known to exist for individual taxa, must be accommodated in establishing future priorities for monitoring. For example, the different geographical concentrations of biodiversity (biodiversity hotspots) examined by Prendergast *et al.* (1993), and the probable effects of climate change on rare species, discussed by Elmes and Free (1994), demonstrate the variety of responses by taxa. A strategically planned and more integrated programme of monitoring is needed to provide a wider ranging and more reliable measurement of assessment of change. In particular monitoring of taxa should be integrated with monitoring changes at a landscape scale (as was begun in CS1990) and with a broad spectrum of other environmental variables (as has been initiated at the dozen or so sites in the Environmental Change Network and facilitated by the Countryside Information System (CIS)).

5.5 THE FUNCTIONAL UNITS OF A NATIONAL SYSTEM

5.5.1 The range of existing units is listed in 5.2.2 and their hierarchical relationships are shown in Figure 5.1. With the development of the EU's European Environment Agency, and the consequent requirement for a national network defined in the EEA regulations (see *Official Journal of the European Communities* No. L 120 of 11 May 1990), the UK focal point is the DOE. However, contact between the EEA and the actual sources of data is through the topic centre network structure currently being developed across the whole EU. Some units (e.g. JNCC, BTO and RSPB) are already concerned with biological information at an international level, for example in relation to the Berne Convention and EU Habitats Directive. Most national biological societies and recording schemes include the UK and the Republic of Ireland, as a natural biogeographic region. BRC and BTO collate and manage data from the Republic of Ireland, in collaboration with the National Parks and Wildlife Service and relevant voluntary bodies, and also contribute to Europe-wide species distribution studies. The Natural History Museum (London), in common with other national and provincial museums and botanic gardens, houses international collections and has expertise in aspects of global biodiversity.

5.5.2 At a local level, most individual records centres operate within an informal network with their main local suppliers of data (volunteer specialists, representatives of national societies and local groups) and, sometimes in a more formalised relationship, with local users, such planning departments, wildlife trusts and local staff of the statutory nature conservation agencies.

5.5.3 With so many active units, most of which recognise and value their own perception of their unique role in biological recording, the establishment of a national system should aim to accommodate the appropriate existing units. Regulation (through negotiated agreement or legislation) could be used to impose a system on some of these disparate units, in particular those supported (wholly or in part) by public funds. However, voluntary partnerships would be essential to incorporate the existing data holding/handling units and thereby the currently available data. A combination of regulation and voluntary partnership is seen as the most practicable and potentially successful method.

5.5.4 The principal biological recording agencies are described in Chapter 2.2. and 2.3. These agencies should form the nucleus of the system. In some cases they already form and, to some extent, operate as *de facto* networks. However, most of the links between agencies are not formalised, except in some cases through contractual agreements.

5.5.5 A nodal structure for a national system should be based on local data being managed for use at the local level, but with links to a national system. This structure should allow users to obtain access to data or interpreted information (appropriate to their needs) at any level. It should not be necessary for a user to enquire for data through a central agency, other than by initial use of a metadatabase. The proposed UK Biota Database may have a role to play in maintaining such a metadatabase which would be an essential component of a national system.

5.6 INTERCHANGE OF DATA THROUGHOUT THE SYSTEM

5.6.1 An essential feature of a national system must be access to, and interchange of, data so that those held locally are accessible to country and national agencies and *vice versa*. Such access and interchange of data cuts across the present trend to regard data as having a saleable value to the agency holding them, although 'metering' the use of data could enable wider access and cost recovery. Although open access to data is in the spirit of the 1990 EU Directive on the Freedom of Access to Information on the Environment (90/313/EEC), this Directive does not preclude the levying of *reasonable charges* for data. The UK interpretation of the Directive in the Environmental Information Regulations (SI 1992 No. 3240) and current practice in, for example, the Soil Survey and the British Geological Survey all reflect this view. However, the UK interpretation of the Directive is disputed by some NGOs where access to governmental data is sometimes priced beyond the budget of voluntary bodies: in some cases these are the same NGOs that form partnerships with governmental agencies, for example in nature conservation and ecological research.

5.6.2 A guiding principle behind a system of data interchange is that any commercial value associated with non-interpreted data, especially data collected or collated at public expense, should be removed by the open exchange of such data within the system. A prerequisite of this approach is that the management of data is funded securely. Ease of access to more complete datasets would enhance the ability of participating nodes in such a system to provide commercial services, based on the creation of value-added products, by applying their interpretational skills. Other options must be considered for those organisations which depend largely or solely on the use of their data for funding, for example 'metered' or subscriber access to datasets (e.g. the Ecological Flora database). Government sponsored access systems such as CIS and UKDMAP also provide access to datasets.

5.6.3 Any existing overlap of responsibilities and interests which determine the use of biological records should not inhibit access to relevant data and interpreted information. For example, a planning authority, a developer and those opposing the development should have access to the same information, although each may put a different interpretation on it.

5.6.4 Open access to data raises some potential problems apart from the issue of the commercial value of data and the need to support data management and further survey and research:
*** The complexities of copyright and the ownership of data (see Chapter 3 and Appendix 4);**
*** Potential misuse of sensitive information (e.g. on threatened species);**
*** Misinterpretation of information (e.g. due to lack of background knowledge or awareness of the limitations of the data).**

5.7 COORDINATING AND REGULATING THE SYSTEM

5.7.1 A national system, composed of such a wide variety of agencies, cannot be expected to operate without coordination and some degree of regulation. A coordinating and regulating structure for biological recording in the UK, whether undertaken by a single agency or a by consortium of agencies, could be expected to cover a range of essential activities:
*** Advising government on the availability and reliability of UK biological data;**
*** Developing and administering nationally agreed standards for data and the technical standards for computing;**
*** Accrediting the component agencies of the system;**
*** Evaluating and overseeing relevant training and the production of manuals for staff throughout the system;**

*** Operating the national metadatabase of information on the system.**

5.7.2 There are few directly comparable situations, but there are some similarities with the coordination and regulation of museums in the UK through the work of the Museums and Galleries Commission (MGC), the Museums Association (MA) and MDA. There are also technical associations of museum staff, such as the Biology and Geology Curators Groups. At present there is no organisation in biological recording comparable to MGC, MA or MDA, although NFBR and BRISC are technical associations and WT takes a coordinating role with wildlife trusts and urban wildlife groups. As important users of data and sources of funding, albeit principally in the area of nature conservation, the statutory nature conservation agencies, occupy an important, but often poorly defined, position in relation to many potential components of a national system. However, they currently do not have the resources to develop such a coordinating role, even in their own specialist field.

5.7.3 A model for a coordinating system which exists in the USA has been described (5.3.4 - 5.3.8). There, The Nature Conservancy establishes, through collaboration with state governments, autonomous data centres which it operates by providing central technical, scientific and administrative support and training. This approach would almost certainly be seen to be too centralist for the UK, but close examination of the workings of The Nature Conservancy partnerships is relevant in the context of a national system for the UK. In particular, the commitment of the individual States to the system has implications, by analogy, for a possible role for local authorities in the UK to act in a similar role.

5.8 ACCESS TO A SYSTEM: DIRECTORY OF DATASETS, METADATA AND BASIC DATA

5.8.1 If awareness of the availability of data is to be increased then it is essential that metadata about the national system, or at least a basic directory to the system, its components, their data holdings and their operating standards would have to be compiled and made accessible. It has been recognised already, in *Biodiversity, the UK Action Plan*, that information on existing data collection and monitoring systems must be compiled, maintained and updated.

_____BOX_____

Metadata relating to biodiversity (From *Biodiversity, The UK Action Plan* , p 148)

1 What data or data sets are held in each institution or centre?
2 Where are datasets held and in what form?
3 How can they be accessed?
4 What is the quality of the data?
5 How were those data collected?
6 How can these data be processed and what have they been used for?
7 Who collected the data and who has used them?
8 Why data were collected and used?

_____

5.8.2 A database of metadata could be based on and developed from the CCBR Survey database (see Chapter 2 and Appendices 2 and 3), subject to the agreement of those who contributed information. Some of the information it contains would be inappropriate for a metadatabase, for example information on funding, data exchange frequency or some of the more detailed information about species or habitat/land type datasets. Its coverage would need to be enlarged appreciably to include, for example, all of the local records centres and wildlife trusts and greater information on the data holdings of statutory organisations and other governmental agencies.

5.8.3 Such a modified, enlarged metadatabase could be compiled, maintained and updated within core funding, for example by the coordinating agency(ies) or as part of the UK Biota Database. It would act as the 'shop window' of the system and in particular for the individual nodes of the system.

5.8.4 Other, similar data directories are being developed in the research and environmental policy areas (e.g. European Environment Agency, GENIE), and the national Data Archive at Essex University already

has metadata about some environmental datasets, such as those assembled as part of the defunct Rural Areas Database.

5.8.5 An essential ingredient in the success of both the BCD and ERIN is that both systems impose uniform standards on the data which they access. Indeed, the establishment of an agreed biological recording data standard is fundamental to the establishment of any system that aims to increase the availability of data. At present, there is no agreed data standard in the UK common to all biological recording. In fact, most of the main collections of data have been compiled using predetermined parameters that have defined the standard for each particular dataset or database. To define and impose a mandatory data standard on all participants in recording activities is likely to be resisted and unlikely to be successful. Successful adoption is most likely if the development of an agreed data standard within the present framework of biological recording in the UK is built on existing practices, known requirements and legal obligations relating to data. It should be noted that although the development of common standards (of all types) is a role set out for JNCC in the Environmental Protection Act, inevitably this role is concerned primarily with aspects related to nature conservation.

5.9 THE NECESSITY FOR PHASED IMPLEMENTATION OF A SYSTEM

5.9.1 The remit of biological recording has been defined (Chapter 1) and examined in detail (Chapter 2). The necessity for a national system to have a firm basis in policy and negotiated agreement or legislation has been noted (5.4). A national system should operate from the local to the international levels and data should be exchangeable throughout the system (5.5 & 5.6). It should be coordinated by an appropriate regulatory body (5.7) responsible for establishing and maintaining a biological records standard and controlling access to the system through the development of a metadatabase (5.8). The necessary standards are discussed in Chapter 6. Clearly, all these developments can not be achieved at once.

5.9.2 The implementation of the proposed system must be expected to be phased over a number of years and is likely to develop piecemeal. Nevertheless, it is important to plan its development. Some aspects are readily achievable, but others would almost certainly require some form of facilitating legislation, or lengthy negotiation. Currently unknown factors, such as the proposed feasibility study on a UK Biota Database, the present restructuring of the Natural Environment Research Council and the development of the European Environment Agency could be expected to have direct bearing on the means and speed of implementation of a national system.

5.9.3 Discussion and negotiation between key agencies and groups can continue and develop at minimal cost. The CCBR review provides a starting point for such dialogue. Much of the dialogue to date has been in relation to nature conservation and biogeographic research and this axis is further reinforced in the sub-group on data formed by the Biodiversity Action Plan Steering Group. More active involvement of those concerned with development planning at the local level and all forms of environmental monitoring is required, both at an official level and through special interest groups such as NFBR, BRISC and ALGE.

_____BOX_____

Phased priorities for starting the implementation of a national system

1 Agree remit of data
2 Negotiate agreement or legislate to establish a national system
3 Promote formalised links at local levels between data centres, planning authorities and voluntary and statutory conservation agencies
4 Develop and promote technical standards for all stages of data management (from collection to networks
5 Define and secure the resources for a national system
6 Compile and maintain metadatabase
7 Establish voluntary accreditation systems for data centres and promote accredited centres to user community

5.9.4 Once a *business case* for biological recording and the remit of data have been established, a first stage in the development of a national system should be to begin to establish minimum standards for the operation of data centres. A voluntary accreditation system has been developed by BRISC for local data centres in Scotland which operates at a simple level and does not cover topics such as quality control and legal rights over data. Although it provides a useful model for a more extended system of accreditation in the UK, it would require considerable modification to be suitable for a national system which involved more than local data centres.

5.9.5 The relationship between data centres and potential users with *de facto* needs for data should be formalised to enable the centres to operate with security and continuity of funding. In some cases formalised relationships already exist (e.g. BRC and JNCC, BTO and JNCC, several local centres and their respective planning authorities or conservation agencies), but these relationships are subject to constant review and often provide only limited security or scope for long-term planning.

5.9.6 Consideration of the needs for technical standards is developed further in the next chapter but more detailed work will be required, particularly to relate the standards for centres to those of their potential user communities.

5.9.7 The data transfer system will develop from the above (5.9.5 and 5.9.6), but the detailed specifications systems will need further detailed examination.

5.9.10 While many aspects of these developments can proceed without changes in policy, legislation or the establishment of an accepted coordinating and regulating body, these issues should be pursued in parallel with other developments if a national system is to established effectively and as soon as practicable. However, unless a mechanism exists which enables local data centres to operate on a formal and recognised basis (such as a nationally negotiated agreement or legislation) and with security of funding, it is unlikely that a national system could be established to operate in the ways envisaged, with long term security for the functional units of the system.

6 IMPLEMENTING A NATIONAL SYSTEM FOR BIOLOGICAL RECORDING

ABSTRACT

Justification for a national system based on the present situation and resources.
Standards: a biological records data standard, data quality.
Database content and compilation protocols: metarecords, data attributes, validation protocols, data transfer and legal obligations.
Networking: national metadatabase, multilayered network, nodal structure, physical network, computer network.
Management: coordinating and regulating the system, general policies, quality assurance and operating policies for data centres, accreditation.
Access to data in a national system: interchange of data, accessing the network, administration of the system, technological developments, quality control, format and protocols for the mobility of data.
Securing and funding a system: official recognition of biological recording, regulating the system, starting the system, coordinating agency, present funding and potential costs, income generation.

6.1 INTRODUCTION

6.1.1 The preceding chapters of this report have examined three central issues:

a) the acquisition of biological data has been and continues on an *ad hoc* basis;
b) as a result, there is lack of compatibility and effective integration of data, and there are deficiencies in the coverage achieved for the range of complexity and diversity of information required;
c) without a more coordinated and strategically planned approach, many contemporary data are of only limited use. The final part of the report considers and makes recommendations for the establishment and operation of an integrated computerised national system for biological records and recording. A national system is a direct result of the wide discussions which preceded the establishment of CCBR and the commissioning of this study.

6.1.2 Support for biological recording can be justified only if there are identifiable requirements for data, although a long term view needs to be taken in specifying such requirements. It would be foolhardy to dismantle or abandon much of the present effort and commitment to biological recording in an effort to fulfil only short-term objectives. Furthermore a broad view of requirements is needed because of the many different users and the equally diverse uses to which they apply data. Also, much of the *ad hoc* biological recording which is currently undertaken by volunteers has the potential to continue in current ways, often with narrow objectives and poor coordination, if the current situation is not changed. The potential of voluntary sources to supply better, more focused data to *external* users (i.e. other than to themselves and to their own special interest groups) can be realised only by changing the current situation.

6.1.3 Many potential users, in particular local government, developers and industries which impact heavily on the environment, have not yet recognised the full potential of the use of biological records to enable them to fulfil statutory obligations resulting from legislation on wildlife conservation and planning. The anomalous situation with local authorities is particularly noticeable: some support the operation of

local data centres through core funding or service agreements, whereas others eschew the need for the types of services associated with local centres. Many potential users are unaware that data and information of relevance to their needs are available or could be compiled. Many do not know how to access them. Unless this situation changes, potentially important data will continue not to be recognised and the lack of any integrated system will perpetuate inefficiency and poor use of resources. It will hinder a wide range of policy decisions which should be made only with appropriate, accurate and comprehensive information on organisms and the environment.

6.1.4 It is, therefore, axiomatic that some form of integrated national system is essential if the types and quantities of information needed to fulfil many local, national and international obligations and aspirations relating to flora, fauna and biotopes throughout the UK are to be met. Vested interests in systems employed by particular organisations and poor understanding of what is needed and how best it could be achieved are obstacles which must be overcome, but at the same time, the specificity of what particular organisations may need for their own purposes has to be recognised. There are many examples where governmental agencies have failed to use existing data and potential data sources to target publicly funded programmes. The problem to be addressed is not whether an integrated, nationally agreed system is needed, but what form or forms it should take and how it should operate. This report considers how it may be possible to develop towards a national system from the present situation.

6.1.5 Three general considerations underlie the conclusions and proposals developed in the following six sections (6.2 to 6.7).
*** A national system should be able to deliver consistent and reliable products to users at costs which reflect the true value of the information when compared with other relevant costs incurred, for example, in development planning or site protection. Currently, biological records are undervalued as an essential component in development planning or site protection and also there are no agreed standards for quality control on data. Thus, the present *market* for data often exploits the goodwill of voluntary specialists and local data centres and sometimes draws on information (such as from some short contract surveys), which is poorly controlled for quality.**
*** Adequate levels of funding are important for the effective development and maintenance of the system. Existing funding, in some respects, could provide improved results if it was deployed to achieve better integration, but better overall resourcing, particularly for data management is essential if quality products are to be delivered. Inevitably, consideration of funding is closely linked to the vested interests of organisations, but the potential for substantial improvements in the delivery of information to users is a powerful argument in favour of the rectifying inefficiencies in the present use of resources and in justifying the need for additional resources.**
*** The essential role of volunteer specialists who supply data, particularly taxa data, cannot and should not be overlooked. These specialists are a crucial resource responsible for the supply of most such data. They will need to be convinced of the wisdom of changes and the potential benefits for themselves (for example, better protection of sites and species, greater security for the units to which they submit their data, more and better summarising publications such as atlases and taxonomic guides, and better feedback), if they are to cooperate fully with changes in the system.**

6.1.6 In subsequent sections of this report, the key issues in the implementation of a national system are discussed, based upon examination of known and predicted user demands, the existing situation, the legal constraints on recording and CCBR's unique experience in all aspects of biological recording and biological data management. Rather than describe a series of theoretical options (which could range from acceptance of the *status quo* to a multi-million pound, national, public-access data network), proposals are made for pragmatic changes and developments from which a scheme, based initially on the existing situation and current resources, could be developed. A preferred sequence of essential steps, each stage of which could be separately costed, is set out in 6.1.9. This approach is fully consistent with the objectives defined in 1.4.1 and 1.4.2 and with the need to adopt a course of evolution rather than revolution.

6.1.7 This study originated before the publication of two recent statements of Government policy - *Biodiversity, the UK Action Plan* (Cm 2428) and *Sustainable Development, the UK Strategy* (Cm 2426), and was not originally intended to provide a response to relevant issues raised in those publications. The subject area of the study is apposite to the current debate prompted by these publications, in particular, when relating to the supply and management of data on biota. There is clear recognition in *Biodiversity,*

the UK Action Plan of the need to coordinate appropriate data 'to provide a library of data and information sources as well as standard summary data on biodiversity' (para 9.37). Subsequent paragraphs of the _Plan_ (9.38 to 9.44) describe some of the complexities of the UK situation and the varying levels of involvement and interest in biodiversity data, and a 'UK Biota Database' is proposed. The feasibility of such a database is now the subject of discussions by the Biodiversity Action Plan Steering Group and its Data Subgroup.

6.1.8 Information technology is developing very rapidly, and national policies and priorities for the use of this technology, both in the service of government (at all levels) and in public access, are continuing to grow. In this dynamic, even volatile, situation, where many important developments are anticipated, including administrative changes (see for example 5.9.2), any proposal for a system should address standards and protocols rather than the precise methods of implementation. It might be possible to implement some aspects of a system almost immediately. For example, where data holding agencies may be operating in ignorance of their legal obligations compliance should be addressed as a matter of urgency, once legally valid procedures have been defined. Other aspects, such as those which involve closer cooperation between autonomous agencies, may require several years of negotiation to bring into operation.

6.1.9 Action on the main components of a national system could be undertaken in the following sequence, but mechanisms and suitable agencies to carry out these actions have not been identified at this stage.

- **Promote greater coordination between suppliers and users of data to ensure the quality of data, to reduce duplication of effort and to promote closer collaboration and partnership.**
- **Develop and promote a biological recording data standard (see 6.2).**
- **Promote the preparation of a metarecord (based on the biological recording data standard) for every relevant existing dataset and database.**
- **Catalogue the present resource - Establish a national metadatabase.**
- **Develop and promote quality assurance policies and protocols.**
- **Develop and promote legally valid data transfer policies and protocols.**
- **Promote the use of information technology and the computerisation of relevant data.**

When these stages have been initiated and have begun to become accepted in the biological recording community it will be possible to establish a national system under some form of centralised leadership, guidance or regulation. There is already some progress in most of these aspects, through convergent thinking within the community, as a result of initiatives such as Recorder, and through the existence of NFBR, BRISC and CCBR which act as _foci_ for discussion and dissemination of information.

6.2 STANDARDS FOR RECORDING BIOLOGICAL DATA

A biological record data standard

6.2.1 As proposed in 5.9, the development of a data standard within the present framework of biological recording in the UK should build on existing practices, known requirements and legal obligations relating to data. Early attempts to define _a biological record,_ in terms of the design of record cards (e.g. Heath & Scott 1977) or more theoretically (e.g. Harding 1985), were confined by narrow views of the levels of standardisation necessary for data. However, these earlier attempts did not really address the need for adaptability in the use of data. In particular, they were concerned mainly with the data fields and did not address matters such as terminology and validation procedures. At present there is no agreed data standard for biological recording, although in reality most of the main collections of data have been compiled using predetermined parameters that have defined the standard for the particular dataset or database.

6.2.2 The development of a biological record data standard must be considered in the wider context of a general data model. Some progress has been made already in this area, by JNCC and the country agencies, using logical data modelling techniques to identify common data requirements. This work should be extended to cover the wider context of the biological recording network in which these agencies operate and to allow for the developing requirements related to biodiversity monitoring and the need for a national metadatabase. The Standard will be also need to be flexible enough to accommodate additions as changes occur and new categories arise or concepts are developed.

6.2.3 Three main factors influence the efficiency with which data are stored and subsequently retrieved, transferred to other data management systems, or merged with other datasets:
*** The structure of the information recorded;**
*** Use of standard terminologies;**
*** Use of controlled syntax.**
The aim of a data standard would be to increase efficiency through defining preferred options in these and related areas.

6.2.4 The advantages of setting a specific data standard for biological recording are :
*** Consistency in the design of recording media;**
*** More efficient recording practices;**
*** A mechanism for assessing comparability between datasets and database applications;**
*** Greater potential to share software and to develop communal software;**
*** Greater ability to share or exchange information;**
*** Better quality control applied to records;**
*** Improved data management.**

6.2.5 Figure 5.1 outlined the current multi-level nature of the partial network for collection and use of biological records in the UK. The international level (not developed in Figure 5.1) should include at least two further levels (European continental and global). The data requirement changes at each level of the network, with less detail and broader categories being used as data are aggregated or compared at progressively higher levels. Thus at the scale of an English county, it is appropriate and practical to record subdivisions of defined sites and to describe them using 'local' terminology such as the RSNC/NCC Phase I Habitats or NVC categories. At a European scale, these local classifications are too narrow, so that sites recorded in relation to the EC Habitats Directive would be described using CORINE Biotopes terminology. A 'local' data standard for the UK should be relevant to the specific needs of organisations at the various levels described in Figure 5.1, but it should also be able to provide a sound basis for interfacing with European and global systems. For example, the overall range of land cover definitions, and the potential to cross reference them at different scales, are discussed by Wyatt *et al.* (1994).

6.2.6 The development of a Biological Records Data Standard for the UK is perhaps the most important issue in the establishment and operation of an integrated, computerised national system of biological records and recording network. For data to be retrieved from this national resource, they would have to be structured consistently and conform to conventions for terminology and syntax. To achieve this, the Data Standard should consist of two parts. The first would define the characteristics of the individual datasets (the *metarecord*) and would be used to compile a national register, or metadatabase, to provide an index to the national resources of data, the types and quality of data and their availability. The second part would define the structure of individual records in the datasets and should cover all types of 'biological records' (e.g. taxa, sites and biotopes)

_____BOX_____

The basis of a data standard for biological records

The following are examples of main headings of information in a biological record, which would need to be incorporated in a generalised model and for which appropriate termlists and syntax may need to be developed.

Record identifier: Unique identifier, recorder, date and time.
Record management: Confidentiality, status (correct, needs checking etc.), date entered.
Recording detail: *Recording method* - survey type, method, quality, extent.
_____*Recording conditions* - weather, period of observation.
Site-related information: *Site/location identification* - name, site/sub-site relations.
_____*Geographic references* - geopolitical, informal regions, cartographic references, land parcel numbers.
_____*Biotope/land cover descriptors* - biotope/landcover unit, measurements (e.g. area), biodiversity assessment.
_____*Landform* - altitude, depth, aspect, exposure, landform, micro-relief, water, soil, sediment, geology.
_____*Physical descriptors* - microclimate, chemical measurements.
_____*Biotic descriptors* - biotope keywords.
_____*Human factors* - present use, potential use, events.
_____*Monitoring details* - visits, condition assessments, appraisal dates.
_____*Conservation and management* - status, threats, damage, management methods, agreements.
_____*Tenure* - owner, tenant details.
_____*Access* - access route, restrictions, facilities.
Taxon-related information: *Identification* - taxon, determiner, date, validation, voucher, links to collections and literature references.
_____*Location* - (link to details held under site-related information).
_____*Population* - sex, stage, number.
_____*Behaviour* - associated species, nature of association, activity (feeding, flying, nesting).
_____*Ecology*: survey specific observations.
General (can be linked to any other attributes)
_____*Associated people* - names addresses, organisations.
_____*References* - literature, photographic.

_____

6.2.7 A Biological Records Data Standard is implicit in theRecorder database package. Over some seven years, a data structure has been developed which can be used for a range of species-based and location-based biological records. The database is further enhanced by extensive format and syntax control of data fields and reference to standard termlists for taxonomy and biotope/land cover classifications. However, no database could cover all possible types of biological record or suit every purpose, so that Recorder is not a full Data Standard. Specialist applications, such as the MNCR database always will be needed to suit specific requirements. For many potential users, Recorder is too all embracing for their purposes (for example, data input for only a few taxonomic groups at a local scale), and for others it is not available on their computer 'platforms' (e.g. organisations which have standardised on ORACLE). In all these cases, and because new and different programs will continue to be developed (for example, to run under WINDOWS, or with improved or customised mapping or multimedia extensions), a full Biological

Record Data Standard should provide both a guide during development and a mechanism for comparing the structure and scope of different applications.

6.2.8 To develop a comprehensive Biological Records Data Standard for the UK, progress must be made with the three factors outlined above (6.2.2). Much already has been achieved: for example, the consolidation of many existing taxonomic checklists within Recorder and the recently published *Comparison of land cover definitions* (Wyatt *et al.* 1994). The Museum Documentation Standard (MDA 1991a&b) provides a good starting point for the consideration of the general information categories relevant to localities, ownership and identifications and, in particular, serves as a good guide to syntax control (standard ways of dealing with names, dates, grid references). Similarly, there is considerable relevant experience in data standards among staff at the Royal Botanic Gardens (Kew and Edinburgh) and among members of the Systematics Association. The Recorder database is probably the best source of models for biological records syntax control and conversion routines (such as latitude/longitude to grid reference conversion) in a strictly UK context.

Data Quality

6.2.9 Not only must data comply with a common data structure but it must be made more mobile within a national system. This will depend on the control of terminology used in each discrete dataset and the methods employed to maintain the integrity of the data during and after incorporation into the dataset (regardless of whether or not data are computerised). It is almost certainly inevitable that guidelines for data quality standards will need to be developed as part of the process of regulating or accrediting data centres within a national system.

6.2.10 The terminology used to record information is as important as the data structure to which it is applied. There are two parts to terminology standards (MDA 1991a&b); *syntax* conventions (e.g. the sequence in which information such as a date or grid reference is written down) and *vocabulary* conventions (the words used, as defined in a thesaurus or termlist). Both aspects must be applied to information which is to be indexed for retrieval for wider dissemination. In biological recording, terminology control is essential for the names applied to taxa and biotopes and also for place names and personal names. The present situation, and in particular the problems associated with maintaining national terminological standards for taxa, are discussed in 2.5. No organisation has responsibility for maintaining a national list of taxa and its related terminology control and this is a serious absence in the scientific resources of the UK. The cooperative efforts to compile taxa lists for use in Recorder, although they demonstrate that there is scope for greater partnership in this area, in fact led to delays and inefficiency because of the absence of an agreed central source of information. The establishment of standard terminology for land cover definitions has been promoted by DOE (Wyatt *et al.* 1994).

6.2.11 There are no national and only a few internationally agreed terminologies or syntax rules which are applicable to biological records, although some aspects are covered in international library and cataloguing standards. The Taxonomic Database Working Group of the International Union of Biological Sciences has endorsed some existing international standards, created others and is actively involved in establishing additional standards, eg habitat, soil and landscape descriptors (Bisby 1994). Some of the standards in museum documentation apply to biological records: for example, Light (1992) lists several international standards including ISO 2709 (framework for data exchange), ISO 2014 (writing of calendar dates in numeric format) and ISO 3166 (country codes). Other relevant standards exist for the citation of names (as guidelines in the *Museum Documentation Standard*) and for the creation of thesauri (ISO 2788 & BS 5723, Establishment of monolingual thesauri British Standards Institution 1987). The *Museum Documentation Standard* (MDA 1991a) includes guidelines on vocabulary and syntax in a detailed field-by-field dictionary, which have been followed by some records centres based at museums.

6.2.12 Validation of data is essential in quality control and should occur at all stages, from the collection of data, to their incorporation in a computerised database or a manual archive.

6.2.13 The most basic form of data validation is the correct identification of taxa or biotopes. This level of validation relies largely on peer appraisal of the competence of those contributing data, which may be

through, for example, a self-regulating national recording scheme or a specific quality control exercise such as that undertaken for CS1990 (Barr *et al.* 1993). The taxonomic competence of specialists is almost impossible to regulate in terms of conventional quality assurance, although some formal training is available, for example in use of the National Vegetation Classification. Competence in plant and animal identification is being tested through a certification system (IDQs) run by the Natural History Museum. Professional organisations, such as the Institute of Ecology and Environmental Management, the Institute of Environmental Assessment and the Institute of Biology, are particularly concerned to ensure the scientific competence of their members, mainly through professional qualifications and peer appraisal. Academic training in ecological taxonomy has never been an important source of experienced field workers and most have gained their practical experience through informal routes, such as local natural history societies, bird clubs and Field Studies Council (and similar) non-vocational courses. Although the inexorable drive towards performance measurement is already impinging on biological recording (e.g. at the level of data collation), it is doubtful that there will ever be the resources to impose a regulating system on the majority of active volunteers in biological recording. In taxa recording, the existing system of informal peer appraisal, and tuition and support for less experienced workers, operates mainly through national and local societies and recording schemes, although the validation of identifications is sometimes a very personal and sensitive matter for individual specialists. Although it is a remarkably robust and effective system, it needs to be supported and extended, particularly to encourage greater coverage of neglected groups.

6.2.14 The format or structure of data collected or contributed in biological recording traditionally has been guided by specially designed recording cards and forms (see 2.5) and by printed instructions. Although generalised *Instructions for Recorders* were published by BRC (Heath & Scott 1977), the most comprehensive instructions have been produced for many national projects (e.g. Butterfly Monitoring Scheme, BSBI Monitoring Scheme, New Atlas of Breeding Birds, British Mycological Society surveys, National Amphibian and Reptile Recording Schemes). Similar instruction manuals have been prepared for biotope and land cover surveys, such as Phase 1 Habitat Survey (England Field Unit 1990) and CS1990 (Barr 1990). However, experience shows that even with the use of structured forms and detailed instructions, the format of records submitted by recorders is often highly individual and somewhat at variance with the preferred form of the guidelines. Nevertheless, it will be important for the future that written instructions are compatible with national defined standards. Data from surveys and monitoring projects which have not benefitted from even this level of format planning will, inevitably, be more difficult to reconcile with a data standard or to accommodate in structured database. Of particular concern are the amounts of data which have been and are still being collected, which follow no model for the minimum content of data or their format. Much of this results from commissions, by statutory authorities and voluntary conservation organisations, of specialist surveys on local, small budgets. Often the preferred product of such surveys is a written report, without provision for the long term security of primary data. In this regard, the introduction of the proposed standards will provide such bodies with a much-needed tool for the improvement of the overall quality and comparability of work which they commission, and is likely to lead to its rapid and widespread application.

6.2.15 The error trapping and validation techniques employed differ depending on whether data are computerised or are held in manual systems. The type of computerised system used influences the extent to which such validation is possible: data held in other than fully relational databases are more difficult to validate using automated techniques.

6.2.16 Where entirely manual systems are in use, the record cards, forms, lists, maps or photographs which constitute the biological records are also the main storage format for data. These documents are usually filed in sequences which relate to taxa, defined sites or geographic units (e.g. grid squares). Little terminology control or format validation is possible with data in these formats, and indexing is possible only through the transcription of names, dates and grid references to index cards. An analogous situation exists with uncatalogued collection of specimens, where the specimens and data labels exist in an 'archival' form. However, the original quality of the data on these documents and collections may be as good as, if not better than, some computerised data in comparable institutions. To dismiss these uncomputerised data as worthless because they are at present largely inaccessible and unvalidated, would seriously under estimate their potential value as long term data resources. To incorporate the numerous repositories of data of this type into computerised form would be both time-consuming and expensive, but

a programme of updating is necessary and needs to be planned for the foreseeable future if invaluable historical data are not to be lost.

6.2.17 Terminology and format validation are common in database applications, where it is essential for effective retrieval of data. Organisations using computers for data management are capable of validating data using the database software at the time of data entry. For example, data entry can be restricted by comparison with standard terminological lists, such as for taxa or biotopes, by *parsing* to check for format (as in dates or grid references)and by matching data to existing authorised files, such as for place names or people. In most cases the routines double as error checking routines to trap misspellings and other mistakes by recorders or mistyping during data entry. The sophistication of checking and the degree to which these techniques are used varies greatly between database applications, nor are they infallible. Not unexpectedly, as a comprehensive, purpose-built, biological record database for the UK, the Recorder package has the most wide-ranging and sophisticated validation facilities of any of the databases examined.

6.2.18 The spatial referencing of survey data is at present very variable, with too much effort, especially by volunteer specialists, being put into recording at spatial scales which are appropriate for distribution mapping but for little else. Unfortunately, the long standing confusion of *recording* (the collection of data) and *mapping* (one method for summarising data) as being synonymous or interchangeable terms has still not come to be understood by many of the more traditional recorders. The potential usefulness of all data types, but particularly taxa-related observations, could be increased greatly by the use of more precise spatial referencing in all types of biological recording. Data recorded using six figure (100 m square) or eight figure (10 m square) grid references can be re-aggregated for use at any one of the spatial resolutions favoured by users, such as 1 km square, tetrad, 10 km square or 50 km square. Also, precise references may be used in a GIS to plot coincidence with non-grid spatial units such as the boundaries of biotopes, defined sites, land holdings or administrative units. The use of *sites* as recording units presents problems because the delineation of an individual site (which may vary greatly in size) is seldom defined uniquely and unambiguously.

6.2.19 Use of more precise spatial referencing could be adopted even when the overall survey strategy is to list species at a site or to map regional or national distributions. When entering a new sampling area, a preferred procedure should be to record a six figure grid reference for each species observation, even if this resulted in only a single occurrence of common species being recorded in the whole sampling area. This practice should become standard for taxa regarded as being scarce, threatened or indicative of a particular biotope. Perceptions among recorders that surveys cannot be rapid and detailed, and that 'recorder fatigue' will result, need to be revised in the light of the demand for data, the requirements of users and the potential long-term benefits to recorders and their special interest groups through the availability of such data. It is not many years before such detailed recording may be considerably easier as hand-held global positioning system devices become more accurate and they and field data-logging computers become widely and inexpensively available. The issue for now is to begin to educate recorders about the need for more detailed data, in advance of the predicted technological developments.

6.2.20 In establishing a national data standard and promoting its use, it is important to recognise that many data, which may not conform fully with these exacting standards, may have a potentially important role as part of the national resources of data. In particular, historical records cannot be expected to conform with many of the standards that have been developed retrospectively. The importance of historical data has been stressed earlier, and the difficulties associated with accessing many key sources of historical data have been described. The development of a national data standard will provide a better framework for establishing the priorities and best methods for providing access to this often hidden resource of information. However, some old material could be left in paper formats, and not computerised, as long as they are properly archived. The existence of such sources of data would need to be catalogued in the proposed metadatabase to enable potential users to trace them.

6.3 DATABASE CONTENT AND COMPILATION PROTOCOLS

6.3.1 A national system for biological recording, regardless of whether it is based on a nationally accepted data standard or on wide inter-agency collaboration, must be able to provide consistent and reliable data

to users. The content of databases in the system and the protocols used in their compilation will determine the overall utility of the system. The potential to use data for more than one purpose, through the establishment of a national system, will be determined by four main aspects that are common to all datasets and databases:
*** Information about the data (*metarecord*);**
*** The form and content of the data (*data attributes*);**
*** The quality of the data (*validation protocols*);**
*** The availability of the data and associated legal issues (*data transfer and legal obligations*).**
A metarecord for each dataset

6.3.2 A metarecord is a descriptive index of the information contained in each dataset or database. Such indexes or directories are essential, not only for the basic documentation of the data by their managers, but also to enable potential users to assess whether the data may be of use to them. It has become apparent from the Survey that few datasets are adequately indexed: indeed the information provided by some respondents to the Survey suggests that many have given no consideration to indexing their data. The statutory nature conservation agencies are known to have data catalogues for internal use. NERC is compiling its Corporate Data Catalogue, a central computerised index to the principal data holdings of all the NERC institutes, which is already accessible through the Joint Academic Network (JANET)and will eventually be accessible through international networks.

6.3.3 The metarecord of each dataset or database should serve five main purposes:
*** A simple description of the content;**
*** A control document which states the origin, ownership and management of the data;**
*** A reference guide to the terminological standards used;**
*** A reference guide to the validation and quality control procedures used;**
*** A standard entry in the national metadatabase.**
The model for a metarecord is essentially what has been outlined as part 1 of the proposed biological records data standard (see 6.2).

_____BOX_____

Proposed contents of a standard metarecord for individual datasets or databases

1 Descriptive name of the dataset
2 Owner of the data: individual or organisation with title to the dataset *
3 Manager: current post and individual with responsibility for the maintenance and quality of the data
4 Addresses of manager: postal and electronic
5 Content of dataset
5.1 Taxon-based giving the taxonomic group at the highest level to cover all groups
5.2 Biotope or land cover datasets giving the type of survey
6 General geographic area covered by data
7 Date range of records
8 Origin of the records in the dataset
9 Survey methods
10 Volumetric information:
10.1 Number (or estimate) of the records, land parcels or target notes
10.2 Static or ongoing (if ongoing with rate of growth)
10.3 Number of taxa in taxon-based datasets
11 Validation and quality controls procedure used
12 Classifications and term lists used
13 Indication of the presence of sensitive or confidential information
14 Current data management methods
15 Data retrieval charges
16 Access agreements and levels of access

*** Where the dataset has been copied or transferred from another source metarecord should indicate that a signed and dated data transfer agreement exists which shows:**

1 Details of the original owner;
2 Agreement to the transfer of title or right of use;
3 Any restrictions on use of the data defined as part of the transfer;
4 Format of data.

6.3.4 It is inevitable that there may be a variety of interpretations of what constitutes a dataset or a database. For example, in the case of taxon-based records, this may range from treating all records as a single dataset, to differentiation to the level of individual families, genera or species, or to the original survey that was the source of the data. This inherent variability could be accommodated in guidelines suggesting that a metarecord should be prepared for datasets based on a single survey (e.g. a grassland survey using NVC), an organisation or major recorder (e.g. a BSBI vice-county recorder), a single taxonomic group (e.g. dragonfly records) or a biotope study (e.g. landscape changes). The necessary important points are that the details of source, ownership, data structure and management are comparable between datasets and the classification and volumetric information are sufficient to enable indexing of these aspects. Earlier experience with FENSCORE, a comparable project to document natural history collections in the UK, shows the need to be practical; over-refinement leads to unachievable workloads, whilst insufficient detail results in imprecise records of limited use.

6.3.5 The preparation of a metarecord for each dataset or database is an essential first step in the development of a comprehensive national biodiversity metadatabase or directory and to increasing opportunities for mobility of data for new applications. The creation of this metarecord should be built in to the specifications of all new surveys and database compilation, with little or no extra cost to the individual project. However, the resource implications of retrospectively creating metarecords for existing datasets are considerable but must be recognised. The database which was compiled as part of the CCBR's study and which is the source of much of the quantitative information in this report, represents a significant step towards the compilation of this metadatabase.

Data attributes

6.3.6 The range of actual and potential forms of biological records is almost infinite and past discussions on what constitutes 'a minimum biological record' have achieved little other than in narrowly defined applications. The ability to reuse data in secondary applications makes practical and economic sense, but it is probably inevitable that future data collection will be determined by both the clearly defined, primary requirement for data and the resources available to carry out the work. It is unlikely that data collected without a clear end-use will have an important place to play in biological recording, other than in a broad contextual setting. Opportunities for the reuse of data would increase through planning the recording activity and structuring the recorded data with reference to a Biological Records Data Standard. Use of a non-exclusive Data Standard, such as that proposed here (see 6.2), would ensure that data could be assembled appropriately for use in secondary contexts. Thus the first aim of a Biological Record Data Standard would be to identify and accommodate attributes that are common to the majority of biological records.

6.3.7 Four factors are particularly important in developing database applications in which the data holdings may be used in secondary applications:
*** Documentation of the validation, levels of accuracy and quality controls applied to each attribute;**
*** Identification of the ownership of data and any constraints on their re-use in other applications.**
*** Identification of the terminological standards used, so that relevant transfer protocols can be identified (or developed) to enable use of the data with other, similar data.**
*** Anticipation of additional requirements for the potential use of data within geographical information systems (e.g. standardised spatial referencing).**

Validation protocols

6.3.8 The need for and types of validation of data are discussed in 6.2 and the need for agreed standards and methods for quality control are discussed in 6.6. For a national system to operate to agreed standards in all aspects, but particularly with regard to quality control, will require acceptance of validation

standards and protocols at all stages throughout the system. It is in this area that the need for some agreed procedure or agency, to define standards and to coordinate and regulate their application, becomes most apparent.

Data transfer and legal obligations

6.3.9 The legal position relating to data ownership and copyright is clearly described in Chapter 3. There are many instances where organisations have been unaware of the relevant law and may inadvertently be operating illegally. These problems will be compounded further if data holding agencies provide access to data over electronic networks, with the potential loss of control over data that this implies. To ensure effective legal control of data is a further reason why proper documentation, such as a metarecord for each dataset or database, is essential. In particular, those organisations which collate biological records should draw up formal data transfer agreements with their suppliers of data, and also negotiate formal agreements over ownership and right of use when they are commissioned to carry out surveys or supply data. The precise form of such agreements will require input by legal experts to ensure their legal validity.

6.3.10 Most, if not all, of those responsible for the collation of biological records will need to reassess their rights and obligations in relation to the data they hold. Few organisations involved in biological recording in the UK have formalised systems for establishing their rights, and those of others, in the use of biological data when transferred between systems, particularly where this involves transfer of data into the organisation from volunteers.

6.3.11 The ownership of newly acquired data, and any rights pertaining to their use, could be treated in an analogous way that museums use when accepting donated items, *w*hen a *transfer of title* is recorded on standard forms. In many cases, especially at record centres or wildlife trusts, individual arrangements may be made with donors or other suppliers of data which include restrictions on use of the data. For example, a donor may not wish their data to be copied for use in the databases of a third party or they may place restrictions on the release of information on rare species. It should be the responsibility of the collating organisation to negotiate the most reasonable terms for subsequent access to data. Adherence to professional standards and principles of quality assurance would enable collating organisations to convince potential data suppliers to grant discretion to the centre for use of their data. All standard record forms used by volunteers should carry a legally valid statement regarding the transfer of copyright or permission (by assignment or license) to use donated data (see Chapter 3 and Appendix 4). The precise form of such statements will require input by legal experts to ensure their legal validity.

6.3.12 The very large holdings of data that have been collated already by organisations must be managed in accordance with the legal advice summarised in Chapter 3. A large scale retrospective 'legalisation' of existing data holdings would seem to be essential for many of the present activities that involve the use and dissemination of data to be brought technically within the law; further advice on the legal obligations of existing organisations must be sought. The resource implications of complying with these obligations need to be examined and a phased programme towards 'legalisation' developed by each organisation with its normal constituency of suppliers and users of data. In particular, data centres should be certain of their legal obligations to the original recorders or owners of data and act to establish terms of use for the data that are already held at the centre. Many large scale data gathering exercise, such as public outreach projects by local records centres and some national recording schemes, present a special problem. The resources required to establish legal change of ownership of contributed data of this type, if existing laws were interpreted literally, would effectively disable many of the existing centres. Wherever possible, use should be made of existing groupings of volunteers, such as societies and recording schemes, to speed up the process of establishing rights over the use of donated data. It will be beholden on data centres to maintain adequate documentation of the transfer by volunteers of their rights over data in this retrospective exercise.

6.4 METADATABASE, NODAL STRUCTURE AND NETWORKING

6.4.1 The availability of data within a national system must be made apparent to potential users. The inability to gain an overview of the national resource is one of the main failings of the present situation

and an important contributory factor in the overall difficulties which have been identified in this review. In order to develop a UK Biota Database, as proposed in *Biodiversity, the UK Action Plan,* it is essential that information on data collections and monitoring systems is compiled, maintained and updated regularly. The metarecord for each dataset or database proposed here (see 6.3) would provide basic information for inclusion in a national metadatabase, covering the generalised topics described in the *Plan* (see 5.8), and could form part of the UK Biota Database.

Establishing a metadatabase of biological recording in the UK

6.4.2 The resultant database should be in the public domain and act as the 'shop window' of the system and, in particular, of the individual nodes of the system. Because of the wide range of potential users of data likely to wish to consult the metadatabase, the system used for public access must be widely and inexpensively available. The metadatabase should be available on-line to subscribers to national computer networks and/or by modem direct dialling and should be accessible without charge. On-line access to metadata is preferred because the metadata for individual units in a national system will require to be updated frequently and modern technologies should be embraced from the outset if there is to be investment of resources in a national system. For those who may not need to consult the metadatabase frequently, it could be published annually on CD-ROM and in printed form, but this should be seen as supplementary to on-line access, not as an alternative. Technological developments over the next 5 to 10 years will be crucial in determining how best to provide access to metadata to the widest possible potential user community, including internationally.

6.4.3 Public access to the UK Biota Database has already been considered in Chapter 9 of *Biodiversity,the UK Action Plan* and similar options could be considered for metadata on biological recording. Publicity about the metadatabase should be continuous and widely based to ensure that the full potential user community is reached, including environmental consultancies and educational establishments. That community is expected to expand with time, particularly in the areas of development planning, countryside access and recreation, and in primary, secondary and undergraduate education. Public access to the UK Biota Database has already been considered in Chapter 9 of *Biodiversity, the UK Action Plan* and similar options could be considered for metadata on biological recording.

6.4.4 As suggested earlier (5.8.2), a database of metadata could be based on and developed from the CCBR Survey database. Although the Survey provides only sample coverage of the overall biological recording resource in the UK, it is estimated that the Survey database includes information about some 50% of the known and potentially available biological data in the UK. This information was collected between September 1992 and September 1993 and inevitably is going out-of-date, at an estimated rate of about 10% per year. The Survey database contains basic information on more than 900 organisations and key individuals (such as BSBI vice-county recorders), together with details of 2000 datasets based on taxa, sites or biotopes.

6.4.5 The Survey database includes only a small sample of information relating to potential sources of data in the physical collections at museums or derived from undergraduate and postgraduate research projects. Also, other surveys of the existence of data resources have been undertaken recently, for example on behalf of the NRA and SNH, and several selective and incomplete directories of data sources have been published in the last few years (e.g. Barlow *et al*. 1992, Donn & Wade 1994). *The Green Index* (Milner 1994) lists the majority of local records centres and natural history societies in Britain and Ireland. Nevertheless, the CCBR database is the largest and most thorough examination of the data resource and would provide a valuable starting point for a national metadatabase. However, any such database will be of long term use only if it is accessible and has an assured mechanism for updating and maintaining the existing information and adding new information.

6.4.6 Contributors of information to the metadatabase would not need extra equipment or software to do so because their information could be received in either digital or written formats. However, in many cases there will be a real cost in extracting information from manual or poorly automated data storage systems. The main reason stated by organisations for failing to respond in detail (or at all) to the Survey was their lack of resources to prepare quantified answers to questions, particularly those about data holdings. If information to be held in the metadatabase is to be comprehensive, some incentive (such as financial aid or other resources) will have to be made available to data centres and data holders to enable

them to prepare and maintain their individual metarecords. To compile the initial metadatabase on an entirely voluntary basis is likely to be slightly more successful than the CCBR Survey in acquiring information from poorly resourced data centres and data holders because the level of detail required would be considerably less.

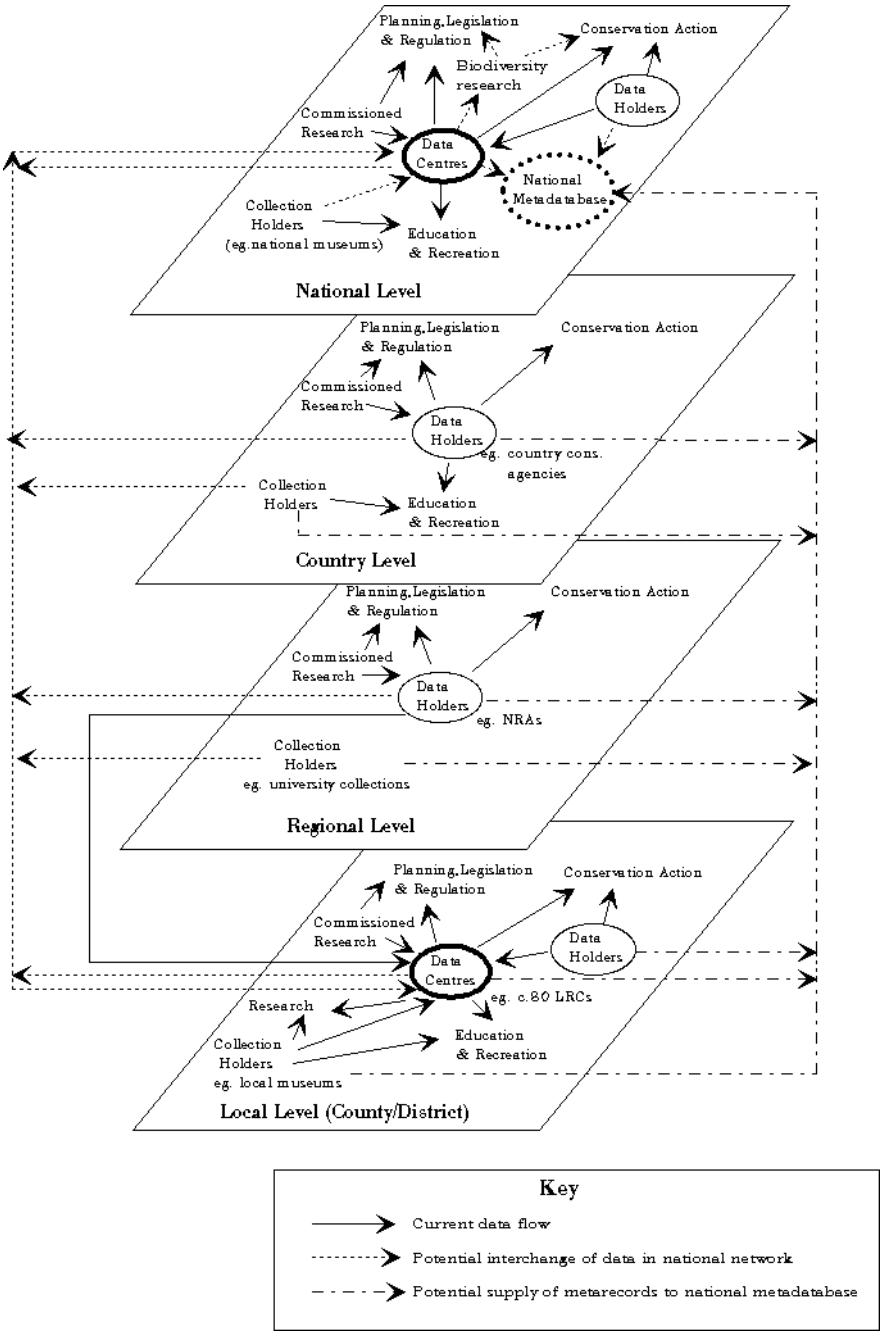
6.4.7 The opportunity to interrelate with GENIE, the European Environment Agency, the national Data Archive at Essex University and other initiatives must not be overlooked. Also, the potential to interrelate with other environmental data should be considered, for example the earth science conservation datasets such as county RIGS data (Harley 1994) and archaeological sites and monuments records (Fraser 1993).

6.4.8 The detailed structure of the metadatabase should be the subject of a separate analysis and design project. It will need to take account of both domestic and international needs, including the ways in which organisations are able, or may wish, to access data over computer networks.

A multi-layered network

6.4.9 The physical network in biological recording in the UK includes collectors, holders and users of data, and is composed of five distinct layers based mainly on geographical/administrative units (Figure 5.1). These units range from the local scale (e.g. parish, city or district) to the whole UK as a nation. Within each layer there are individuals or organisations concerned with some or all of the functions of biological recording at that level.
These functions include:
*** Survey;**
*** Collation of data;**
*** Environmental and conservation legislation;**
*** Development and structure planning;**
*** Action on wildlife conservation;**
*** Environmental consultancy;**
*** Education;**
*** Recreation;**
*** Research.**

Figure 6.1 **The current and potential national biological records network**


Planning, Legislation & Regulation
Conservation Action
Biodiversity research
Data Holders
Commissioned Research
Data Centres
National Metadatabase
Collection Holders (eg. national museums)
Education & Recreation
National Level
Planning, Legislation & Regulation
Conservation Action
Commissioned Research
Data Holders
eg. country cons. agencies
Collection Holders
Education & Recreation
Country Level
Planning, Legislation & Regulation
Conservation Action
Commissioned Research
Data Holders
eg. NRAs
Collection Holders
eg. university collections
Regional Level
Planning, Legislation & Regulation
Conservation Action
Commissioned Research
Data Holders
Data Centres
eg. c.80 LRCs
Research
Collection Holders
eg. local museums
Education & Recreation
Local Level (County/District)
Key
Current data flow
Potential interchange of data in national network
Potential supply of metarecords to national metadatabase

6.4.10 Within each level, responsibility for these functions may be assumed by different types of organisation. At the district and county levels there is no single model for how these functions are funded or who assumes responsibility for them. At the regional, country and national levels, the numbers of organisations involved are smaller and their individual roles are often more clearly defined, although there is at least potential conflict for the role of principal data collator at each of these levels. The model presented in Figure 5.1 can now be re-stated as a functional model (Figure 6.1).

6.4.11 Within this structure there are two organisational types that hold or collate data at each level although there can be some overlap between the roles of these types. *Data holders* collect, collate or hold data for their own internal use. *Data centres* collect, collate or hold data as part of an environmental information service, primarily for use outside their organisation. They are poles in a continuum in which overlap between these two roles is increased where a *data holder* (such as some local records centres) has to support its operation by trading in services which are largely dependent on the data holdings. As a consequence data become a resource and the role of this type of *data holder* shifts towards that of a *data centre*. Increasingly, most of the *data holders* which are not entirely voluntary are becoming *data centres* as they set up data exchange agreements or provide services based on their data holdings for some form of remuneration.

A nodal structure for the network

6.4.12 The development of a national system based on the present situation must take into account the very real concern of all members of the biological recording community to safeguard their own unique roles. One of the strengths of much biological recording is the extent to which there are already several

types of *de facto* network, locally and nationally. These will provide a framework for future developments towards a national system.

6.4.13 The present amount of coordination and communication between organisations within local *de facto* networks (e.g. within a county) is variable, but in many cases is well developed to ensure a supply of data from local specialists to local data centres. Good links also exist between specialists at a local level and record collators at the national level, such as BSBI, national recording schemes, BRC and the ISR. Nevertheless, the model (Figure 6.1) highlights some duplication in the present system. For example, the records of an individual specialist may be passed directly to a national recording scheme and thence to BRC, they may also go directly to a local records centre and to a local or regional natural history society, and then be copied to a national centre. Even if the present system was to remain unchanged, this emphasises the need for transfers of data to be documented and all records to be identifiable to source. However, it may be difficult to eliminate unnecessary duplication of data and effort until there is sufficient awareness and confidence in the system.

6.4.14 National networks for taxon-based data, such as those described in Chapter 2.3, should play an essential part in all stages of the development of a national system. It is through national societies and specialist groups that the all-important taxonomic expertise is regulated and promoted. Their existing systems for mobility of data may need to be regularised and updated, but every effort should be made to involve these groups in the development of proposals and not to impose decisions upon them.

6.4.15 Greater problems occur in relating the diversity of networks at the lower levels with those of the regional and national levels. There is little coordination between networks at the county and district levels while potential users at higher levels can experience difficulty in tracing information relevant to their needs or in obtaining consistent responses to requests from a number of data centres. This is especially so in the situation where there is no real index to potential sources, and such indexes as have been compiled (see 6.4.5) are highly selective in their coverage and detail. Specific proposals are not possible without further, detailed consideration.

6.4.16 The establishment of a national system should be based on:
*** development of a single, publicly accessible metadatabase;**
*** agreed standards and accreditation systems;**
*** encouragement to data centres to become *nodes* within a national system;**
*** encouragement to data holders to contribute information about their data holdings to the metadatabase.**
This pre-supposes mechanisms to compile and maintain the metadatabase, a regulating system to agree standards and accreditation systems, and a focal point to provide guidance, advice and encouragement at all levels. Data holders and data centres would have different roles in a national system, but all publicly funded data holders would be expected to provide controlled access to their data, as a result of the legislation covered in Chapter 3.5.
6.4.17 Acceptance of the role as an official *node* within the system would require positive action by each data centre:
*** to supply metarecords about data holdings to the metadatabase;**
*** to undertake to adopt agreed standards and quality assurance methods**
*** to develop and publish formal policies on key aspects of data management and dissemination;**
*** to accept the jurisdiction of a regulating system;**
*** to encourage data holders to use the professional services of their local data centre (or national centre, as appropriate) as an agency for providing users with access to their data.**

6.4.18 The number of recognised nodes within a local network might vary but, in view of the low potential of supplying biological records to generate income, the existence of competing data centres would represent a dissipation of funding and manpower resources, rather than 'healthy competition'. Data centres with an exclusive monopoly on local data should not arise because of the community structure of most biological recording at the local level.

The Physical Network

6.4.19 The apportioning of roles to existing organisations should lead, ideally, to the presence of one data centre node for each local, regional and country sub-network. It is probable that some revision of the location and geographical responsibilities of the existing data centres funded by local government will become necessary when new unitary local authorities are established in England and Wales. In England, there should be at least one data centre per county or appropriately sized unitary authority, so that coverage is entire but without overlap. At present some areas of England have no effective data centre, but in others there is some overlap between centres. In Wales the present role of local data centre has been taken on by wildlife trusts but their geographical coverage is patchy. The function of a data centre for land cover data rests with the Countryside Council for Wales. In Scotland the geographic coverage of data centres is both varied and patchy. The identification of Scottish data centres should be coordinated by BRISC, based on its present accreditation system, but accreditation criteria will be necessary to cover the whole of the UK. The proposed Northern Ireland Biological and Geological Records Centre, at the Ulster Museum, could serve as the data centre for the entire province. In addition to these data centres, the network would include a greater, but irregularly distributed, number of participating data holders. National scale data centres concerned with species, such as BRC and BTO, should be seen as integral parts of this network, but with roles which are quite separate from those of the local centres. They will have well defined routes to acquire and disseminate data which should be developed from their existing *de facto* networks as described in Chapter 2.3. An important outcome of the national system should be more effective interchange of data between local and national centres (in both directions) and the avoidance of duplicated work. The work of such national data centres, as sources of national scale collated data, is fundamental to the statutory responsibilities of JNCC, the country nature conservation agencies and DOE.

The Computer Network

6.4.20 Initially, mobility of data can be achieved by simple means such as transferring data on disk. As has been noted (2.8.9), use of disk transfer is already growing although few data transfers are currently made using networks. A second phase of development would be the creation of a computerised network based on connection of local data centres to the metadatabase and the provision of on-line handling of enquiries. The introduction of computer network links would have the following advantages:

- **the potential for electronic access to data, to increase their mobility;**
- **on-line access to the metadatabase;**
- **broadening of the potential user community;**
- **the potential to route enquiries to data centres by electronic methods (e.g. Email).**

To achieve this will almost certainly require some funding for equipment and software, for example through grants and partnership agreements. Funding for training and technical support for the system will also be essential.

6.4.21 At its simplest, the network could be developed through direct connection to a metadatabase host machine by dial-up modem. However, the alternative of setting up a public domain database and messaging system through one or more of the existing networks (e.g. JANET, COMPUSERVE, GreenNet or INTERNET), should be examined and costed in detail. Considerable progress has been made in recent years in the development of international biodiversity databases and access through public domain networks (e.g. INTERNET)(Green 1994, Canhos *et al.* 1992), which could be examined in relation to the UK situation. The choice of a network platform will need to be decided by a technical working group representing, and reporting to, all levels and interests, to ensure that the requirements and constraints at all levels are taken into consideration. A network which excludes half of the potential nodes, for example as a result of impatience to establish a system or the protection of vested interests, must be avoided at all costs.

6.5 MANAGEMENT: REGISTRATION, ACCREDITATION AND COORDINATION

Coordinating and regulating the system

6.5.1 Participation by organisations in a national system, including the contribution of information to a metadatabase, should be on a voluntary basis. However, a national system, embracing a wide variety of organisations, each with unique responsibilities to their user communities and sources of funding, cannot

operate without some form of coordination and regulation. This type of management of the system could be undertaken only if biological recording was afforded recognition as a discrete activity (e.g. comparable to the management of archives) rather than as an adjunct to well recognised activities such as planning, nature conservation, museum curation or research.

6.5.2 Responsibility for this management would best be undertaken by a single agency or by a consortium of agencies. The involvement of NGOs would be essential to be able to involve the broad community concerned with biological recording. There are few comparable situations to provide models of such management, either in the UK or abroad, although there are some similarities with the coordination and regulation of museums (see Box) and the Sites and Monuments Record system (see 5.4.3) in the UK, and with elements of both the Nature Conservancy in the USA and ERIN in Australia (see 5.3.4 - 5.3.12).

6.5.3 At present there is no *official* side in biological recording comparable to MGC, MA or MDA. For museums, the direct path to government is through a single department (DNH), but in biological recording access to government is diffuse, involving many departments, although DOE has overall responsibility for planning and nature conservation, and it also commissions much related research. NFBR and BRISC fulfil the role of technical associations, bringing together representatives of many local data centres, especially through the Recorder Users Group within NFBR.

_____BOX_____

The coordination and regulation of museums in the UK

The co-ordination, regulation and development of standards for museums in the UK is managed by the *official* side of museum administration: Museums and Galleries Commission (MGC), Museums Association (MA) and the Museum Documentation Association (MDA). The MGC is funded by the Department of National Heritage (DNH).
MGC: promotes co-operation between national and provincial museums, and has executive functions such as grant aid to museums through the nine Area Museums Councils, a museum registration scheme, co-ordinating the funding of the work of the MDA and publishing standards in the work of museums (e.g. the care of biological collections (Paine 1992)).
MA: the professional association for museum staff, which promotes standards, training and communication within the profession.
MDA: grant aided by MGC and others for standards development and outreach work. Also provides a range of self-financing services.
Technical associations of museum staff, such as the Biology and Geology Curators' Groups, provide active forums for information exchange on practical matters.

_____

6.5.4 A coordinating and regulating structure for biological recording in the UK should undertake the following.

- **Advise government and other users on the availability and reliability of UK biological data (compile and maintain a national metadatabase).**
- **Develop, administer and monitor a national data standard.**
- **Oversee the accreditation of the components of the system.**
- **Negotiate relevant training for staff throughout the system and the production of manuals.**

6.5.5 A formalised unit to oversee the operation of the national system, and to carry out the roles outlined in 6.5.4, is essential if the system is to improve significantly. Gradual 'osmosis' of ideas, standards and policies can be expected to continue within the present largely *ad hoc* system: for example Recorder provides a useful basis for structuring the data holdings of a data centre and is becoming a widely used standard package for local records centres. However, gradual and *ad hoc* acceptance of some degree of standardisation will be too slow, and prey to too many uncertainties, to be able to develop into a national system in the foreseeable future. A national system should be aimed to deliver consistently reliable data to a range of users throughout the country. The UK is very well placed to develop a comprehensive national system, so the opportunity should not be missed to develop a working example which may provide a model for other countries, especially elsewhere in the European Union.

6.5.6 New resources for implementing and managing a national system will have to be found because none of the potential 'members' of the system have resources to contribute either funds or manpower, unless a radical decision was made to divert existing funding to such work at the expense of existing work. The small numbers of individuals responsible for the operation of data centres in the UK are already fully committed to existing work. Even the important role of representation on specialist groups and professional committees has become increasingly difficult for many of them. For this reason, the management of a national system could not be undertaken within the existing resources allocated to biological recording, regardless of where responsibility for developing the management role was placed.

6.5.7 Because the metadatabase would be central to the development and success of a national system, it will require careful regulating and there should be close liaison with a formally constituted consultative and advisory group representing the participating data centres, data holders and users from each of the 'levels' of the network. Partnership between statutory agencies, local government and voluntary organisations will be central to the success of the system. The voluntary organisations, in particular, should receive assurance that their role will not being taken for granted, should get support and encouragement in adopting changes, and will have every right to expect to get something in return for their efforts.

6.5.8 The development of the metadatabase, and subsequent developments towards computerised network connections between data centres and others, would be helped by the formation of a technical working group, in particular where the developing information policies of key organisations can be discussed openly before irrevocable decision are made. These longer term policy issues could become part of the remit of the Data Sub-group of the Biodiversity Action Plan Steering Group, but pressure groups such as the Wildlife Link Joint Information Policy Group and NFBR could be expected to take an active interest in progress. However, it is essential that implementation of proposals for the coordination and regulation of the system must be by a group with a clear and non-partisan remit, which is capable of embracing the whole biological recording community.

Written policies for data centres

6.5.9 The existing national biological data resource exists in the form of individuals and organisations which relate to each other as a multi-dimensional 'network'. To increase the general availability of data, and especially to improved the flow of data between levels of this network, it is essential that the key issues of *data quality* and *data ownership* are resolved. A first step in this process should be through the development and agreement of formal policies and procedures to be adopted by the potential nodes of a national system. Working by example, it would be essential for national organisations, such as JNCC, BRC and BTO, to adopt policies (in reality this will involve minor adjustments of existing policies and formalisation of good practices as policies) and to further promote their adoption at the level of local data centres. At present the approach to policy development is incomplete and patchy so that this aspect of the operation of biological recording should be given high priority by funding organisations, to ensure that data centres operate both within the law and to an acceptable and demonstrable level of competence. This must be at an early stage in the development of a formal accreditation system for data centres.

Policy requirement for the collation, management and dissemination of data

6.5.10 Despite a widely acknowledged need (Copp & Harding 1985, Way 1986), there is no up-to-date handbook or set of guidelines for the operation of local data centres or any other organisation with a primary interest in collating biological records. Nevertheless, the present absence of formalised procedures for maintaining the quality of data and services should be set against a general move towards better quality assurance and accountability. Some wildlife trusts, local records centres and others are already examining the applicability of Quality Assurance and BS5750 to their work, though none of the respondents to the Survey had yet adopted it. During 1994, the Gloucestershire Wildlife Trust, with financial support from English Nature, has prepared formally documented policies and procedure statements for the operation of the Gloucestershire Environmental Data Unit, which could provide useful models for many other data centres.

_____BOX_____

Topics which require development of formal policies

Quality Assurance
- **Policy statement**
- **Quality system**
- **Organisation and management**
- **Contract review and control**
- **Product identification and document control**
- **Product process and testing**
- **Non-conformance and corrective action**
- **Quality audit and review**

Operating policies
- **Data exchange**
- **Data collection/collation**
- **Data supply**
- **Data management**
- **Data entry, update and maintenance**
- **Document and data control**
- **Data ownership and confidentiality**
- **Data security and backup**
- **Charging**

_____

Accreditation

6.5.11 The accreditation of data centres and data holders will depend on the adoption of the agreed standards and policies proposed earlier in this chapter, and on the establishment of a regulating body. A national system based on such standards and policies will lend itself easily to an accreditation system where the incentive to become an accredited data centre or data holder will be an important step in establishing credibility in the biological recording community. The success of this approach, albeit applied at a simple level, has been demonstrated by the BRISC Accreditation Scheme for local records centres in Scotland. By October 1994 this scheme had been taken up by 20 centres, complying with a range of criteria for accreditation at one of 5 levels (*BRISC Recorder News* No. 22). The process towards accreditation will be accelerated if the funding agencies recognise and use the policies and standards in specifying data acquisition by contractors.

6.6 ACCESS TO BIOLOGICAL RECORDS AND THEIR EXCHANGE

Interchange of data

6.6.1 An essential feature of the national system should be an ability to interchange data between data centres and data holders at all levels. Thus, appropriate summaries of local data could be made accessible to country and national organisations, while subsets of data held by country and national organisations could be made accessible to relevant organisations operating at local levels. Interchange would be dependent on issues described earlier, such as ownership of data, data quality and accreditation of data centres, having been resolved at each data centre. The legal ownership of many data holdings should be re-examined as a matter of urgency, in the light of the legal issues covered in Chapter 3. Protocols for the transmission of data should be established. Access to data throughout the system should include control mechanisms on their use to avoid their being applied inappropriately and to safeguard sensitive

information from potential misuse. Areas where the uncontrolled use of data would potentially present problems include:

* **data on rare and threatened species;**
* **release of data by those who do not own or have rights over them;**
* **exploitation of the research potential of data;**
* **exploitation of data for commercial gain.**

6.6.2 The government introduced the Environmental Information Regulations 1992 (1992 No.3240) to implement the EC Directive on the Freedom of Access to Information on the Environment (90/313/EEC). Interchange of data between organisations in a national system for biological records would almost certainly conform with the spirit of the Directive. Although the Regulations are directed towards formalising rights of access to environmental information by the public, the approach adopted could be applied to promote the interchange of data through a national system.

6.6.3 A guiding principle behind a national system to promote and facilitate data interchange is that any commercial value associated with non-interpreted data (raw data) should be removed by open access to such data within the system. Such access would, of course, respect any constraints imposed by the originators on the use of their data. Conversely, the ease of access to more complete datasets would enhance the ability of participating nodes to provide commercial services based on the creation of value-added products, by applying their professional interpretational skills. In some circumstances, such use might conflict with the original intentions of the contributors of data, but this is a legal issue best dealt with in the context of assigning rights to the uses and ownership of data. An approach based on the open interchange of non-interpreted data implies the need for funding (not necessarily wholly from the public sector) to support the infra-structural costs of data compilation and management. The sensitivity of volunteers to the apparent 'sale' of 'their' data is an issue that cannot be ignored, but this is largely a matter of gradual education of volunteers about how and why data are collated, managed and used in the best interests of wildlife. The wider issues of the commercial value of data are highlighted by the Government's Tradeable Data Initiative (DTI 1986). At present there is little competition for contracts between local records centres but a situation could develop (and in a few cases has already occurred) in which centres could take the opportunity to compete for work outside their immediate area. However, the proposal for a national network of accredited data centres could reduce this risk by seeking to avoid unnecessary overlap in coverage by local data centres. Any overlap of operational responsibilities is potentially wasteful.

Accessing the network

6.6.4 The establishment of a metadatabase, describing key attributes of datasets and databases, which would be accessible through a national system, has been discussed in 6.4. The present-day facilities for providing access to a metadatabase are considered in 6.4.21. Although there will be a need, in the short-term, to provide access to metadata, the long-term future of access systems must be considered so that the metadatabase is not tied into a system which will become obsolete in a few years. Due regard must be given to developing technologies in the adoption or development of an access system. There are two important issues in relation to access systems - the administration of a national system for biological recording and the technology of access systems.

Administration of the system

6.6.5 Administrative responsibility for developing and facilitating a national system, and in particular the metadatabase, will depend on the allocation of sufficient resources (see 6.7) to establish a coordinating and regulating body. Although recognition of, and funding for, this body is essential, the obvious need for the system to develop as a partnership between many types of organisation, means that a 'top down' administration is unlikely to be successful. The autonomy and specialisms of existing organisations have to be accommodated in a national system. However, incentives to partnership must be present, to encourage organisations to sign up to community-wide standards, quality control and accreditation, and to encourage user organisations to respect the assured quality provided by accredited data suppliers.

6.6.6 Accredited data centres must be afforded an advantage over non-accredited units and individuals in the supply of data and services to users. This is not to advocate local or national monopolies in the supply of data, any more than local public records offices have a monopoly on providing access to local archives. There are opportunities for Government to guide developers, planners and other users concerning their obligations to take full account of biological and related environmental factors in the planning process and also to refer them to sources of data and advice. Although the recent *Planning Policy Guidance* note on nature conservation (PPG 9) identified the need for *"fully adequate information about local species, habitats, geology and landform*" (see 3.2.1), it is unfortunate that the model of PPG 16 (Archaeology and Planning) was not followed. An opportunity has been missed, in PPG9, to direct developers and local government planners to potential sources of such information in the way that PPG16 listed SMRs. However, it is to be hoped that subsequent revisions of PPG9 could include lists of contact addresses of potential sources of collated information.

6.6.7 The administrative structure of a national system, a metadatabase or a focal point for the coordination of effort and involvement must have long-term security. This is essential for it to be able to develop and promote a national system based on partnership and commitment to a common approach throughout the biological recording community. A prerequisite for the harmonious and effective development of a national system is the representation, if not direct involvement, of participating organisations throughout the administrative structure. Over the last decade, many of the Government funded organisations (e.g. the Rural Areas Database, the former Nature Conservancy Council and the Natural Environment Research Council) working in related areas have been subject to radical changes of priorities, resources and funding systems. Therefore, it may be unrealistic to expect any one non-departmental governmental organisation to take on the long-term responsibilities of administering or acting as a focal point for the system.

Technological developments

6.6.8 The metadatabase, as a basic index to the network, should, in the long term, be accessible to as wide a user community as possible. For example, in Chapter 9 of *Biodiversity; the UK Action Plan* a proposal is made to establish 'one-stop information centres or data shops on biodiversity'. Such a distributed information network should form part of the wide range of indexing and access systems which are likely to develop in the near future for public and specialist information. Local information centres on a county-wide computerised information system, to serve public demand, have been set up recently by Hampshire County Council. Possible options for public access to a biodiversity information system should not be considered in isolation from other developments in information networks. Information technology is developing very rapidly, as are the commercial opportunities for those developing public access information systems. Detailed consideration of these issues is beyond the scope of this report, but biodiversity data and biological records should be included as one of the many fields of potential interest to the public, in relation to the Government's commitment to ensure public access to environmental information, preferably through development of information technology and public telecommunications. Communication at international levels will increase and prospects for direct access to and from the European Environment Agency and the growing number of data centres in other European countries will provide new opportunities for international collaboration, for example in measuring changes affecting biotopes and species.

Quality control in a national system

6.6.9 A system of accredited data centres and data holders, indexed through a national metadatabase, cannot be developed overnight. There are many fundamental practical issues (e.g. on common policies and standards) to be resolved, areas of responsibility to be defined and funding mechanisms to be secured. The credibility of a national system must be established early on if the present, effectively unregulated, situation is to be improved upon. The lack of understanding, among much of the present and potential user community, about the realities of the supply of data, in particular their sources and quality, has to be overcome in the early stages of developing a national system. A main factor in promoting a national system will be access to data and data dependent services which are consistent and reliable; if they are not both accessible and of assured quality then the system will fall rapidly into disrepute. Basic requirements of the system are that:

*** The metadatabase must be comprehensive, kept up-to-date and easily accessed;**
*** Enquiries directed to data centres and to data holders, at all levels, must be handled professionally and quickly;**
*** A service which is comparable throughout the UK should be available to users, particularly with regard to quality, range of services, speed of service and charges.**

6.6.10 Each node in the national network must provide a professional service for the collation, management and dissemination of biological records. To fulfil this role many, especially at the local level, will rely on their partnership with individual specialists, recording groups and environmental organisations throughout the various levels of the network. For data to flow effectively within the system common standards and technical protocols must be adopted.

Format and mobility of data

6.6.11 The format of data being collated or collected by data centres and data holders will be determined by practices which result from the development of a data standard and formal policies, and will be moulded by improved access to technologies. The fact that more than 50% of the data identified in the Survey are not computerised demonstrates that considerable investment will be necessary to bring most parts of a potential national system up to a basic technological standard, such as access to, and understanding of, computerised databases for the storage and retrieval of data. English Nature has already begun to explore opportunities for key datasets in its possession to be computerised by local records centres under service agreements. It is inevitable that there will be no resources for some types of data to be computerised even in the medium term (in particular historical records), but it is essential that a national system should be based on units which are capable of delivering at least some of their data in electronic forms.

6.6.12 The range of products provided to users by data centres and data holders will depend on user requirements, the resources of data, the technology resources and the manpower available to suppliers and users to enable them to apply current technologies (such as GIS or on-line access) and newer technologies (such as interactive multi-media and hypertext information systems). However, the reality of the present situation is that the majority of data products provided by data centres of all kinds are in paper forms, such as written reports of interpreted data or as publications such as distribution atlases. Very few centres provide raw data in any format, except as a result of data exchange agreements (e.g. between national recording schemes and BRC or between some local records centres and wildlife trusts). At present raw data are usually transferred as duplicate paper copies of original documents or, increasingly, on floppy disks.

6.6.13 Development of standard data management packages, such as Recorder, will facilitate the mobility of data by electronic means. This mobility will increase as data transfer facilities in data management systems are improved and as increasing numbers of data centres and data holders begin to use E-mail and acquire access to computer networks such as JANET, GreenNET and COMPUSERVE. The development of a computerised national network of biological data centres and data holders will augment the existing formats for the supply of data and will open up exciting possibilities, using modern technologies, to provide greater public access to summarised data for educational and recreational purposes.

6.6.14 Data transfers throughout the system should be documented in the metarecord, together with details of the origin of each discrete dataset and any restrictions on use of the data. The merging of datasets from different sources is not without its problems (see 2.5), but the majority can be overcome by the application of strict quality control on data in their care (e.g. by attributing identifications, complying with a biological records data standard and maintaining metarecords). Nationally accredited organisations should have to demonstrate publicly their adherence to agreed standards for all types of quality control. The matter of who should regulate taxonomic experts is best left to peer review and should not be the province of a central regulating body, although a centrally available list of experts should be maintained. It is essential to take into account legal implications of mobility of data, which, for example, are covered in the Data Protection Act (1984) and the Environmental Information Regulations (1992).

Protocols for the mobility of data

6.6.15 Data transfer includes both the movement of whole files or extracts from files of information within a computer network (*File transfer*) and the incorporation of records from one database into another (*Data merging*). File transfer protocols refer to the movement of files between computers over an electronic network (not to the process of loading data into a new database). Protocols are required to ensure that files may be transported successfully and without corruption using network, telephone or satellite links. Data merging protocols are the procedures required to translate the data from one database into the form used by another and then check the data for errors or record duplication before import. Successful import or merging requires complete understanding of the structure of both the incoming data and the target database, and knowledge of the classifications and coding systems used by both. Even where the two sets of data are identical in format (e.g. in the case of transfer between two users of Recorder) care has to be taken to ensure that record keys are not duplicated and data mistakenly overwritten.

6.6.16 At present, the majority of transfers take place by exporting ASCII or .DBF files from one database for import into another. Where the two databases have different structures or have used different classifications (e.g. for land cover classes), an intermediate translation phase is necessary and software specifically for the purpose must be written. The present situation will be improved through the promotion and adoption of a data standard with formats and termlists. Wider use of Recorder and the preparation of metarecords, which describe the classification, terminology and validation used, will be important steps towards protocols for merging data. Within the next few years the availability of data translation tools for common applications will increase: for example, a data transfer module for Recorder is planned, which would greatly aid the development of distributed data capture by records centres.

6.7 SECURING AND FUNDING A UK SYSTEM

6.7.1 The products of biological recording are needed and used as a result of Government policies, national and international legislation and scientific enquiry. Earlier chapters have described how the current situation in biological recording is not sufficiently organised to be able to meet present demands, does not make the best use of resources of all kinds and is inappropriate for meeting an anticipated increase in, and greater complexity of, future needs unless some form of official recognition of biological recording is established and a regulatory system introduced. Despite the apparent weaknesses of the present situation, biological recording in the UK is probably developed to a greater extent than in most other countries. Not only is the recording activity considerable but it is, potentially, capable of enormous improvement if common standards are established and activities integrated more effectively. It has been argued, therefore, that a national system for biological recording should not be developed *de novo* but should take full advantage of existing strengths, should be based on the present organisations and sources of funding and should utilise existing resources to best effect. How can such a development be secured and adequately funded?

Official recognition of biological recording

6.7.2 Official recognition is required to regularise the basis on which the system would be established. Formal recognition of key aspects is essential:

- **Biological records constitute an information resource which is essential to effective nature conservation, development planning, environmental monitoring, scientific research, education and public information;**
- **The community of specialist volunteers, organised through local and national societies and special interest groups, and working in partnership with local and national data holding organisations, is an invaluable primary source of biological records;**
- **Accredited sources and repositories of biological records should be established and maintained as part of a national system;**
- **Accredited sources and repositories should operate to a set of standards agreed or regulated through the system;**

*** National and local governmental and official agencies should promote the use of accredited sources and repositories of data in their own work, in the work of their agents and in work in their purview;**
*** Information about access to sources of data and their services should be collated centrally and disseminated through public information systems;**
*** The flow of data to, within and from the national system should follow the principles of the EEC Directive and UK Regulations on access to information on the environment.**

6.7.3 It may be possible to make some progress with the development of a national system by voluntary agreement, which would inevitably need some form of pump-priming and support for the process to be initiated. Although it is improbable that primary legislation is either likely or necessary to bring about the changes to the present situation which are seen to be necessary, subordinate legislation, for example linked to the Wildlife and Countryside Act 1981 and/or the Environmental Protection Act 1990, may be necessary to bring about some of the more prescriptive parts of the proposals for a national system. For example, legislation may be necessary to establish and regulate an accredited system with many nodes throughout the UK. It would establish biological recording in a workable administrative and cultural framework, on which a national network could be developed using modern information technology. The precise form of subordinate legislation which would be most appropriate in these particular circumstances is a matter for further discussion, possibly though the medium of the Biodiversity Action Plan Steering Group. Official recognition of this type has many similarities with that which already exists for archaeological sites and artifacts through Statutory Instrument No. 1813, 1988.

Regulating the system

6.7.4 Coordination of a national system will be possible only if the policies and practices of the operational units of the system can be regulated. There are two basic options for coordinating and regulating the system.
*** A coordinating agency with powers to influence the funding of units within the system, for example through a formal accreditation scheme. This type most commonly operates with quasi-official organisations, such as museums.**
*** A voluntary regulating scheme administered by a body set up for that purpose by the community being regulated. This type is more commonly found in commercial activities, such as for travel companies and outdoor recreation establishments.**

6.7.5 An added complication to options for regulation in biological recording is the variety of types of organisation which potentially should come within a national system, and their range of activities. They could range from registered charities, such as wildlife trusts and some biological societies, to public companies, local government funded museums, universities and Government departments. For most of these organisations, biological recording is only one of their responsibilities. Even where recording may be likely to become a higher priority, for example with local museums or wildlife trusts, it will remain only part of their overall activities. However, it might be appropriate for the biological recording activity of organisations (especially at a local level) to be set apart from their other activities as semi-independent data centres. In this way, each local centre could be overseen by a consortium of local users and funding bodies, as has occurred already, for example, in Cornwall and Somerset. The legal complexities of coordinating and regulating a system for biological recording in the UK have not been examined in detail as part of this review. It is essential that the legal aspects of the options for a coordinating/regulating system are considered in detail before action is taken to initiate other aspects of the CCBR report.

Starting the system

6.7.6 Development of a national system capable of delivering a consistent type of product throughout the UK will inevitably require some redistribution of existing resources and some new resources, at almost all levels. This important factor is crucial to the proposal to establish formal recognition of biological recording through subordinate legislation. The system will operate effectively only if organisations with statutory duties regarding the environment are obliged to take appropriate measures to obtain and use reputable information on species and biotopes. The purpose of the system will be to ensure that information is delivered but, without appropriate resources throughout the system, opportunities for improvement on the present situation are few.

6.7.7 A new, small organisation dedicated to the coordination of the system is likely to be needed, although it may be associated with an existing organisation. Most of the other elements of the system would develop from existing organisations, such as specialist societies and groups, local records centres, wildlife trusts, BTO, BRC and the statutory nature conservation agencies. It would be naive to expect that, as a result of establishing the system, significant changes in the operation of all these elements could be introduced without a need for some additional funding, in particular to bring about the necessary restructuring of the present situation and also to complete the coverage of local records centres.

6.7.8 The probable costs of the type of national system outlined earlier in this chapter, can be estimated only when clear options for establishment of the system have been determined. Progressive development of the system will be inevitable, so that predictions of costs should allow for a steady increase in work and costs over the initial years until the system is fully operational throughout the UK. In the first few years the system should become established on an interim basis, after which, progress should be reviewed and further stages planned, costed and funded as appropriate.

6.7.9 Before the national system can be properly constituted, with full coverage of the UK at a local level, some form of *shadow* coordinating group will be essential, to be responsible for overseeing progress from the present situation towards establishment of a recognised national system maintaining effective contact and dialogue with the recording community as a whole, and for liaison with those responsible for preparation of any subordinate legislation.

6.7.10 The early stages of setting up the system are likely to include the following, all of which should proceed with close collaboration between all the potential parts of the system:
*** develop and implement a national data standard for use throughout the system;**
*** prepare a candidate inventory of data centres and data holders to become 'nodes' in the system;**
*** prepare interim criteria for accreditation of 'nodes';**
*** accredit key 'nodes' on an interim basis;**
*** prepare and disseminate the interim metadatabase of the system;**
*** review progress and consult with 'nodes' and users to decide priorities and plan subsequent work;**
*** develop funding policies for new data centres and secure funds to establish and maintain them.**

6.7.11 At the level of an individual node in the system, the amount of work necessary to be considered for interim accreditation will vary greatly. It must be expected that some potential nodes will not fulfil even the interim criteria for accreditation. This interim stage should deal sympathetically with potential nodes, to ensure that the system includes as many appropriately organised units as possible and to exclude only those which either are too poorly resourced or actively do not wish to participate. Phased introduction of the system, say over five years, would enable the inevitable problems of an accreditation scheme involving autonomous organisations, to be resolved.

Funding a national system

a) Present funding
6.7.12 A national system, based largely on the present organisations, should expect to utilise the existing resources although, for reasons alluded to earlier (2.2), it has proved impossible to calculate the financial resources currently being used in biological recording in the UK. Several examples of expenditure on individual projects in the last few years are available, but they do not provide a clear view of the costs of biological recording as presently constituted. Many local records centres, together with national data centres such as BRC and BTO, depend on funding from a variety of sources, some of which cannot be guaranteed to continue or are subject to annually renewed contracts. It is probably evident that funding is, at best, uneven and, in many cases, it is clear that under-funding and insecurity of funding are widespread.

6.7.13 The ways in which many local and national institutions operate, are financed or charge for their services have changed in the last decade or will change soon. For example:
*** many organisations that formerly were funded by central Government have acquired agency status and are now fully or partly dependent on contracts for funding;**

*** the structure of local government is likely to change from 1995 onwards;**
*** governmental organisations, such as the NRA and the research councils, are being restructured;**
*** a market for data has developed, although the market generally provides low financial returns.**
Not all these changes have yet become apparent in biological recording, but all are likely to have significant effects on the operation of a national system and the ways in which it is funded in future.

b) Potential costs of local units in a national system
6.7.14 The basis for a national system, as a whole, already exists in the local records centres, most of which are operated within museums funded by local authorities or by wildlife trusts.
To establish this core, positive action is need to:
*** rationalise the present situation in some areas, for example where coverage is duplicated;**
*** develop new centres to cover areas of the UK not presently covered;**
*** equip the units and train staff to the accreditation standards;**
*** ensure appropriate and continuing funding.**

6.7.15 Complete coverage of the UK by local data centres is central to the successful operation of a coordinated system. Estimation of the costs of operating a local data centre, for example to cover a present-day county, should include the following activities:
*** The collation/collection of field data[4];**
*** Support and outreach schemes for local specialists, groups and societies involved with the collection of data;**
*** Staff for collation/collection of data and ongoing surveys;**
*** Staff for data capture and data management;**
*** Managerial and administrative staff (duties to include the interpretation and presentation of data);**
*** Computing and GIS hardware and software;**
*** External computer network costs;**
*** Staff training;**
*** Accommodation, overheads, travel and consumables.**

6.7.16 The Survey showed that estimated average cost for basic funding of a local records centre, with 2 or 3 salaried staff, was about £58 000 (at 1992/93 prices). The range of work undertaken within this scale of budget was almost certainly less than that listed above. The figure underestimates the probable real costs of a fully equipped and operational local data centre and should, as a minimum, be doubled to match the known costs of some of the more effective, better equipped and better staffed local centres. The size and staffing of a local data centre will be determined by the scale of operation and the total area and population covered. A local records centre capable of meeting the essential requirements in 6.7.15 is likely to need a salaried staff of 5 full-time equivalents, suitably equipped with computer hardware and software, network links and routine office facilities. It could expect to incur annual running costs (staff and full overheads) of about £150 000 at 1994/95 prices, although this figure would have to be increased considerably if the staff of the centre were to undertake significant amounts of field survey work. These estimates are almost certainly a significant increase on present expenditure on local records centres because there would be more centres (to complete coverage of the UK) and, in many cases, more and better equipped staff at each centre. These extra costs would be counteracted by greater efficiency in the system as a whole, but especially in the provision of data to users. The present-day funding for local data centres varies from centre to centre, but includes direct core-funding from local authority budgets (including planning, museums, arts and leisure, recreation, education), service agreements with local authorities, contracts and grants. The start-up costs for many local centres would include purchase of more suitable computing equipment (excluding GIS) - £10-15 000 in the first year and an annual budget of £3-5 000 for maintenance, upgrading and replacement (at 1994/95 prices). All local centres should eventually aspire to have PC-based GIS facilities, a developmental step which would require a larger computing budget for each centre (say £15-20 000 for start up and £5-8 000 for annual costs).

[4] **Delimitation of the scope of data to be collated and collected will have to be decided in the context of both local and national needs. The extent to which national needs would help determine work at a local level cannot be predicted here.**

6.7.17 The estimates of £1 million for start up costs and £10.8 million p.a. for running costs (including staff and overheads) are for a minimum of 70 local records centres. It should be realised that these sums are, in part, already being expended by existing centres, albeit without being *ring fenced*. Both security of funding and the need for additional funding within these estimates should be assessed in detail. By the time the use of GIS and electronic communications is introduced the costs of establishing many local records centres (i.e those which already exist in some form) are likely to be appreciably less than the estimates given above.

6.7.18 Although its is outside the scope of this review, there is a strong case for advocating the development of local data centres which cover environmental data additional to biological records, in particular for geology and archaeology. Examples of this broader type of data centre already exist in some areas, in particular several which include geological data. Broadly based local environmental data centres are a logical extension of the concept of local focal points for compiling and accessing environmental information. It is a stated policy of the Museums Association to promote county-wide environmental record centres, located in the appropriate museum service. Furthermore the MA Policy Statements (Anon 1990) state that" *the Museums Association should be responsible for the coordination of such interdisciplinary recording in liaison with existing national organisations responsible for coordinating data collection in their fields of interest*".

6.7.19 Prior to local government restructuring, most local data centres have been set up to operate at the level of administrative counties or, and in a few cases, for individual metropolitan boroughs. Whether it will still be appropriate, after local government restructuring, for most local centres to operate at the level of (former) counties will depend on the extent of that restructuring. A data centre covering the area of a mainly rural unitary authority will be less likely to be viable than one which operates over a wider area, possibly serving several authorities (as already occurs in London and some shire counties). If existing local centres are effective, it is to be hoped that they will continue to serve areas broadly similar to those covered at present - little benefit would derive from drastic upheavals in the coverage of such centres for purely administrative purposes. Assuming that the overall number of data centres follows broadly the present number of English and Welsh counties and Scottish districts, with one in each of Northern Ireland, the Isle of Man, Jersey and Guernsey, there would be a total of approximately 70 local data centres in the UK. However, the importance of providing access to local users (of all types) may mean that the density of local centres should be as uniform as possible throughout Britain, which might result in as many as 80 or 90 centres.

c) Potential costs of national agencies in a national system

6.7.20 The extent to which the types of national agencies would, or could, become part of a national system depends on the strength and effectiveness the system itself and the potential benefits that would derive from being part of the system. It is impossible to predict accurately the contribution of country and national organisations in terms of funds and resources, either as nodes in a network or as contributors to funding needed to operate the system through a coordinating agency. Organisations such as DOE, the statutory nature conservation agencies, NERC, BTO, WWT, The Wildlife Trusts and some national societies would be essential components of the system and, as has been noted earlier, national data centres such as BRC and BTO are not securely funded. It would be hoped that important compilers and users of data, such as MAFF, NRA, CC, NT and NTS, would wish to become part of the national system because of the obvious benefits to them that would be derived from access to wider resource of data.

d) Potential costs of coordinating and regulating a national system

6.7.21 An agency to coordinate and regulate the system, such as that considered earlier (6.7.4 - 5), must be proportionate to the scale of its responsibilities. The resources necessary to initiate this work must be sufficient, at the start, to be able to deal with the main duties of the *shadow* coordinating group described in 6.7.9. Development of the data standard could be undertaken within the agency or contracted out. The other duties should be carried out by the staff of the coordinating agency, to establish a collaborative relationship with the potential 'nodes' of the system and to ensure that the agency was seen as the focal point of the system. The number of staff and costs of the coordinating agency would be small, initially no larger than an individual local data centre. A staff of 5 and a staff/overheads budget of £150 000 to £200 000 per year would ensure rapid development of the system. The initial start-up costs for all types of

computing equipment and training, and possibly for contracted out work, would be additional. Travel costs, to visit data centres in the first few years, could be expected to be high. A smaller overall budget would slow down the development of the system, whereas the first year or two will be the busiest, with the recruitment and accreditation of the large initial core of data centres. It might be appropriate to channel funding through the coordinating group to grant aid existing and new local centres as part of the national system. This would of course increase the administrative work of the group and would require at least one member of staff to manage grant aiding.

6.7.22 Recorder is likely to be the most widely used data management package within the national system, at least in the initial years. It is essential that financial provision is made for the continued support and maintenance of Recorder but, equally, support for other widely used packages should be considered where necessary.

6.7.23 The costs of setting up the metadatabase, initially as a pilot project as part of the duties of the *shadow* coordinating group, would have to be costed into the budget for establishing the system. A substantial part of the original CCBR project was the questionnaire Survey which gathered information on over 350 organisations and individuals, the overall costs of which was about £50 000 (including the preparation of the questionnaire and the development of the database). A similar or larger sum would be needed for at least two years to establish the metadatabase, using the metadata compiled by CCBR as a basis. Once established, the running costs of the metadatabase (including maintenance and updating) would depend on the size, amount of remote access (including remote updating of individual metarecords) and any net income should the metadata be made available in hard copy and CD-ROM forms. It would be impractical to consider the metadatabase as having any real potential to become fully self-financing.

The practicalities of income generation

6.7.24 The common assumption that biological records have a tradeable value must be questioned. Even in the USA only about 5% of income comes from the sale of data (see 2.2.26). An inevitable development of this assumption has led to a view that the collection of raw data and the supply of interpreted information are, at least potentially, self-financing. Whilst this may be a valid assumption with some types of data, such as the small number of site or biotope datasets collected using venture capital, it ignores the origins and the full range of users of most biological records. The income that it is possible to generate from trading in interpreted information rarely, if ever, covers the costs of the collection, collation and management of data.

6.7.25 Although the origins and users of biological records are described in detail in Chapter 2, it is important to reiterate that, for most taxon data and many other types of data, the main sources are volunteers and the main users, other than the voluntary sector themselves, are planning departments and agencies funded by central government (e.g. the statutory nature conservation agencies and the NRA regions). These local and central governmental agencies have had the benefit of open access to free, or certainly inexpensive, data for over 40 years. Biological recording has been consistently under valued and under resourced. The present unsatisfactory situation reflects the low priority given to biological recording by its primary users. It is in their own interests, for local and central governmental agencies to develop their existing roles in recording through close collaboration in the national system so that they match their funding to the system as a whole to reflect more realistically the use they make of biological records.

6.7.26 The private, commercial sector is not a large generator or, at present, a large user of data. Even since the introduction of legislation requiring environmental impact assessments for some types of development, there has been only a slight increase in the income to record centres from commercial sources. This may reflect more about the interpretation of planning legislation than about commercial demands for data. Although the commercial value of biological records is never likely to be able to support the level of work necessary to maintain the data resource which is needed, the ability to supply good quality data efficiently is likely to increase demands and income from the commercial sector.

6.7.27 The complexities of personal and corporate rights of ownership and copyright of data have been described earlier. These are not complexities which should be allowed to inhibit the use of data, but the

use of data to generate income presents a new suite of legal problems. The legal issues surrounding tradeable data have to be resolved quickly and, as far as possible, simply if they are not to stifle the supply of data in a tide of bureaucracy which could affect all form of use.

6.7.28 Many data are collected by volunteers or under small contracts which often cover little more than travel expenses and subsistence costs. Key land cover and some site surveys tend to be rather better funded, using contract and professional staff. If a national system is established then there will be moves towards greater imposition of standards and more planned and focused surveys. Volunteers who record principally as a form of recreation are likely to respond sympathetically to increased demands on their time and patience, if there is a prospect of more and better information going into the nature conservation and planning processes, and if research is being done and published in accessible forms. The vast majority of volunteers want their data to be put to good use and understandably want some kind of recognition or acknowledgement for their work. The cost effectiveness of using experienced and well directed volunteers (or volunteers working on an expenses only basis), particularly in projects related to nature conservation, should not be overlooked.

6.7.29 The argument for open access to data applies also to the proposed national metadatabase and to local data centres in a national network. Although it is feasible to develop on-line databases, accessed by subscription or direct charging (as is common in the medical and bibliographic fields), there is unlikely to be a commercial market for UK biological records which could sustain the cost of operating either the metadatabase or the local nodes. The metadatabase should ultimately be openly accessible to all users as an on-line information source. Further detailed enquiries should be routed to the relevant local data centre or specialist data holder who would negotiate access to their data and charging rates in line with their access and charging policies formulated as part of the accreditation process.