22 Information systems Introduction

The purpose of an LRC is to collect, collate and disseminate wildlife information. This cannot be achieved successfully without using computerised information systems. In less than a decade the development of the internet and World Wide Web has revolutionised the way we can communicate and disseminate information. Current advances in linking database technology to web applications will further revolutionise how we think about data gathering and sharing information.

As the focus for local recording networks, an LRC should lead the way in information technology. Computers pervade every aspect of an LRC's work and staff should be well versed in their use. The management of data and the routines required to generate products are very demanding so the hardware and software employed by LRCs must be fit for purpose.

An LRC needs to be at the heart of the information flows within its local area (see section 12 *Data flows*). It also needs to promote and support the use of information technology amongst local recorders. The National Biodiversity Network is all about making wildlife information more widely and freely available; to achieve this aim, data must be computerised and held in a form that allows them to be readily transferred between their users. This means that LRCs must take compatibility and standards into account when establishing their computer systems and when working with recorders. LRCs must ensure that they have access to the technical expertise necessary to support the interchange of electronic data; this is especially true for those LRCs that have opted to use non-standard software or operating systems. LRCs are in a unique position to lead the way in demonstrating how the use of new internet technologies will vastly improve the sharing of data.

This chapter provides an overview of the information systems needed by LRCs. As well as an introduction to the main aspects of Recorder 2000, the principal biological data management software that incorporates NBN standards, other biological recording software is considered. The roles of geographical information systems (GIS) and Ordnance Survey data are also covered, as well as the important role that LRCs must play in developing and supporting the effective use of IT within local networks of recorders.

23 Recorder 2000

Policy & Principles

- **Recorder 2000 has been designed to implement NBN data standards and to improve the flow of information within the NBN.**
- It is available through a network of resellers, who can also provide training and support.
- Implementing Recorder 2000 takes considerable planning and preparation.
- At the time of publication, it does not manage all the metadata that LRCs may want it to.
- Recorder 2000 is extensible, so it can be added to in order to meet the needs of individual LRCs.
- Recorder 2000 relies on an MS Access 97 database. This has implications for its use.

23.1 Introduction to Recorder 2000

Recorder 2000 is a completely new piece of software that unifies species, habitat and site data using the following NBN standards:

The NBN data model

- The NBN dictionaries
- The NBN transfer format

These standards aim to make data more consistent and easier to exchange. Recorder 2000 is not a fixed product but can be customised to meet each user's needs.

23.2 Aims of the software

The software was designed to achieve three main aims:

23.2.1 Implement NBN standards

Recorder 2000 incorporates many NBN standards, so that anyone using it is implementing those standards, possibly without even knowing it. The standards include: standards for how metadata about data-sets are recorded; standard terms for recorders to use when entering their data, especially standard names for species, habitats and administrative areas; standards for checking both that the original data being entered into the software are accurate (record validation) and that they are being entered correctly (checking). Throughout this guidance, confirming that a taxon identification is accurate is referred to as 'verifying' that record. This term is not used in Recorder 2000, which instead uses the term 'validating'.

23.2.2 Increase the flow of data within the NBN

Recorder 2000 can import data-sets from other copies of Recorder 2000 and other recording software. The import routines check the incoming data to ensure that they fit the NBN data model, so they can be correctly assimilated into the database. They also check for duplicate records, to avoid a record being exported to someone else's machine and then being reimported onto the originating machine as a new record; Recorder 2000 can do this because each copy of the software has a unique number which is attached to every record entered onto that copy, so each record has a unique label (the globally unique identifier or GUI). Care should be taken not to try to import the same data more than once from software other than Recorder 2000, as the new copies will be assigned new GUIs.

23.2.3 Replace Recorder 3.3

Recorder 3.3 was written in a software language called AREV, which is no longer supported by its manufacturers. By contrast, Recorder 2000 uses a back-end database in Microsoft Access 97; this is well supported and integrates well with other Windows software such as word processors and spreadsheets. Moreover, Recorder 3.3 is very much a package focused on species recording; Recorder 2000 gives more flexibility in managing habitat and site data.

In summary, Recorder 2000 has been designed to retain nearly all of the features of Recorder 3.3 while implementing the NBN data model and other standards, and making data transfer easier and more reliable. It is *not* designed as an interface for the public to learn more about biodiversity data.

23.3 Key features of Recorder 2000

23.3.1 The NBN data model

The NBN data model was designed to allow as many dimensions of a record to be managed as possible. However, when it comes to turning the ideal logical model for managing data into a real piece of software, technical constraints limit the number of dimensions that can be managed. Thus the physical data model (ie the data management model that Recorder 2000 actually uses) is a compromise between the ideal and what is technically possible. The software stores data in about 120 tables, each with between two and 28 fields—a very complicated structure.

There are two main strands to how the NBN data model, and therefore Recorder 2000, manages data: one to handle hierarchical information about sites (called 'locations' in Recorder 2000), the other to handle hierarchical information about observations of species and habitats (ie records).

sites	surveys
sub-sites	survey events
features	samples
	occurrences

The user can link these two types of information by linking any site or sub-site with any survey event or sample.

23.3.2 Records

Records of biological observations are composed of several items in Recorder 2000: surveys, survey events, samples and occurrences; hereafter referred together as the observations hierarchy.

An *occurrence* simply says *what* was recorded. One or more occurrences can be recorded in a sample (eg a patch of MG5 grassland).

A recorder takes a *sample* whenever they records habitats and species using a particular method. The sample says *how* the record was made, which could be, for example, catching moths in a moth trap, recording plants in a quadrat or making a casual observation of a toad in the field.

One or more samples can be taken in a *survey event* which gives information about a visit to a locality, typically a day out in the field. The survey event explains *who* went *where* and *when*.

Finally, one or more survey events together comprise a *survey*, which gives metadata about the data-set such as *why* the data were collected and who ran the survey. This metadata can be passed on to the NBN Index to raise awareness of the data-set to which they relate.

Records can be marked as confidential; records so marked cannot be exported or viewed by users below the 'full edit' user level, nor will they appear in reports unless specifically requested. Recorder 2000 allows you to manage information about specimens collected in the course of making species records. Though it can record who *ran* a survey, Recorder 2000 does not currently manage information about the ownership of observations.

Whenever someone classifies an observation as belonging to a taxon or biotope, then they are making a determination. Determinations lie at the heart of biological recording and it is determinations that are managed in biological information management systems such as Recorder 2000. Determinations of an observation might include:

- the original identification
- subsequent confirmation of the original identification
- invalidation of the original identification
- alteration of the original identification
- further identification as an item on a different checklist (a synonym)

Recorder 2000 lets the user track the different determinations that might be made for an occurrence. For reporting purposes, LRCs will probably want to select their preferred determinations.

23.3.3 Sites

Sites are called *locations* in Recorder 2000. The program can manage information about any number of locations and each location can have sub-locations—which can in turn have their own sub-locations and so on *ad infinitum*.

Locations and sub-locations can also have innumerable *features* linked to them. These can be anything that is particularly interesting about a location (eg a pollard oak, a calcareous flush, an unusual association of marine algae). Threats to, and management of, features can be tracked. However, spatial references cannot be assigned to features.

23.3.4 Documents and other references

Recorder 2000 can be used to manage references to sources such as paper files, published reports, journals and other documents. References can also be made to electronic data (eg photographs, scanned maps, sound files) stored elsewhere on the PC or network on which Recorder 2000 resides; the user is able to link straight through to these other files.

23.3.5 Measurements

Both biological and non-biological measurements (eg number of animals in a school of porpoises and soil pH respectively) can be recorded in several different places in Recorder 2000 (ie in samples, occurrences and locations—including sub-locations).

23.3.6 People and organisations

Recorder 2000 has a sophisticated system for managing information about people and organisations related to the observations hierarchy and people who use the software itself. It can hold addresses and other contact information for individuals and the organisations they represent.

23.3.7 Viewing, adding and editing data

Recorder 2000 uses some familiar ways to present the data held within it. Users of Windows Explorer will be familiar with the tree views used in the data entry screen: the left-hand pane lists the available items and the right-hand pane gives more detail for the item selected in the left-hand pane. Items can be dragged and dropped within windows to move them, and between windows to copy them. Each window also has a button for finding related data in other windows.

Though there are many fields in Recorder 2000 which can hold data, comparatively few of these are mandatory fields, which are colour highlighted (yellow by default). LRCs should think about encouraging recorders to collect more detailed information (so increasing its value), and about how this extra detail might be managed in the software (see section 10 *Minimum record standards*).

Because the data model in Recorder 2000 forces the user to adopt a certain amount of rigour in their data management, data entry can seem rather slow. This is compounded by the fact that the software's data validation routine validates each record as it is entered, which reduces entry speeds further.

However, the software includes some tools to make it faster. The rucksack is a convenient place to store items that you use frequently, be they taxa, biotopes, locations, people or documents. Think of it as a clip-board that gives you easy access to lots of bits of data.

If you often enter species lists for site visits, particularly when the species tend to be the same, then you should use recording cards to enter data quickly and you can base a recording card on a rucksack.

Another way to speed up data entry is to enter data quickly in another format and then import them into Recorder 2000, leaving validation and development of the observations hierarchy to be done automatically at the import stage. At the time of publication, there is a spreadsheet import wizard add-on, available from the NBN website, that LRCs might find useful.

23.3.8 Maps

The software is designed with some mapping functionality, but does not qualify as true GIS, since it cannot be used to analyse data. It can be used to obtain spatial (eg grid) references, digitise site boundaries (often called polygons) and to pinpoint where samples were made. It can also map the locations of taxon occurrences, biotope occurrences, samples or survey events. All these data can be exported to other users of Recorder 2000, such as satellite recorders.

Recorder 2000 comes with the following base maps:

- Great Britain
- Ireland
- UK

It is also supplied with the following boundary data:

- counties (post 1974)
- vice-counties (as supplied with the distribution mapping software, DMAP)

It is possible to use base maps other than those supplied with Recorder 2000 (eg those available from Ordnance Survey), but the cost of licences for using such data should be borne in mind.

23.3.9 Reporting

Recorder 2000 uses a complicated data model (see section 23.3.1 *The NBN data model*) which can make it hard to report back on the data contained in the database. The reporting tool is currently quite limited, and anyone wanting to ask sophisticated questions has to do so by other means, such as SQL, Crystal Reports or Microsoft Access 97's own reporting tools. Using these tools requires a thorough understanding of the NBN data model. Report templates can be saved in Recorder 2000 and then passed to other users.

LRCs should consider sharing custom-built reports, perhaps through forums like the NBN website. There may be a case for accrediting the reports that people share within the NBN.

A significant advantage of Recorder 2000 over Recorder 3.3 is that it is much easier to use other applications, such as word processors and spreadsheets, to report the data in clear, attractive products. Another advantage is the 'query construction kit', which is a set of prewritten queries that can be bolted together by the user to generate very specific queries. This allows people with a less exhaustive understanding of the data model to form the queries they need.

23.3.10 Exporting and importing data

Recorder 2000's reporting tools are used to create export files for passing data to other users. The standard export format between copies of Recorder 2000 is XML (extended mark-up language), which not only carries the data, but also says what type of data each item is.

When records are exported from a copy of Recorder 2000, the exporter can go on to edit or delete those records on that copy of the software. However, when a user *imports* data from another copy of Recorder 2000, they cannot edit the records they import; they can only augment them. Instead of editing a record directly, the user must contact the originator of the record and ask them, first, to correct the record (since only the originator can edit the record), and, second, to re-export the record to them. When the record is reimported, the import procedure warns the user that they are trying to import a duplicate record; at this point they should choose to overwrite the old, incorrect record with the new, correct one. This allows users of data in the NBN to follow an audit trail that reveals the history of a record and its subsequent interpretation; it also helps to avoid duplication of data. For example, if I send you a record for a natterjack toad, but you decide that the record is inaccurate because the animal was seen in entirely the wrong habitat by an unreliable observer, you cannot edit my record, but you can add to it to say that you have redetermined the record as invalid and that your determination is the preferred one. More importantly, you should contact me, as the originator of the record, and ask me to review the determination. I might then redetermine the record as a common toad and re-export the record to you; when you reimport the record, you would accept the new record in preference to the old one.

Because Recorder 2000 allows users to exchange data quite easily, it is well suited to use in LRCs which work with 'satellite' recorders. Each recorder should have their own copy of Recorder 2000 with its own site ID, and should exchange records with the LRC according to LRC policy. However, data custodianship (see below) becomes an acute issue in satellite systems.

Please see section 22.6 *Data custodianship* for issues to consider when exporting and importing data.

23.4 Recorder 2000 resellers

Recorder 2000 is available from a network of resellers administered by JNCC; the NBN website holds details. The resellers have a range of skills, from training to writing add-on modules; but all are able to give basic advice on installing Recorder 2000 and preparing to implement it in an LRC. The LRC manager should discuss these matters, and ongoing support and training, with their reseller.

23.5 Preparing to implement Recorder 2000

In preparing to implement Recorder 2000, the LRC should consider a number of factors, covered under the headings below.

23.5.1 The limits of the database

The back-end database of Recorder 2000 can hold up to 1 Gb of data. Some LRCs are approaching or have even passed this limit. However, there are likely to be other versions of the software released that make it possible to manage much larger data-sets. The latest details are available on the NBN website.

23.5.2 Access 2000

Recorder 2000's back-end database is in MS Access 97, which can be used to look at, manipulate and report on the data. However, if you have Access 2000, you should proceed with *extreme caution*: when you open an Access 97 database using Access 2000, you are given the opportunity to convert it into an Access 2000 database; you should *never* do this as it would make the data unreadable by Recorder 2000! Furthermore, it is advisable that you do not add, edit or delete any data in the back-end database using Access 2000 except after getting expert advice.

23.5.3 Hardware and networks

The software requires:

- Windows 95 or later (includes NT)
- a Pentium 233 with 32 Mb RAM (as a minimum)
- a CD-ROM drive if you want to access dictionaries from disc
- 800 x 600, 256 colour graphics
- 20 Mb hard disk space for the application, but 80 Mb for the back-end database if you load the taxon dictionary from CD-ROM (which is recommended)
- enough hard disk space to accommodate your data (1,000,000 records occupy roughly 1 Gb)

Recorder 2000 databases are large (an empty database without records, but with the taxon dictionary, occupies 80 Mb), so calling data and running queries in particular can be time-consuming. An LRC manages hundreds of thousands of records, and therefore needs at least 128 Mb RAM.

23.5.4 User levels

Each registered copy of Recorder 2000 recognises the following user levels, and extends the following privileges to each:

User level	System manager	Full edit	Add only	Record cards only	Read only
Can create other users?	✓	*	*	*	*
Can use database management tools?	✓	*	*	*	*
Can add data?	✓	✓	✓	Can only add records using record cards or single record entry screen	*
Can edit data?	✓	✓	Can edit own records	*	*
Can delete data?	✓	✓	Can delete own records	*	*
Can see confidential records?	✓	✓	*	*	*
Can see non-confidential records?	✓	✓	✓	✓	✓

Before installing Recorder 2000, the LRC should decide who will be the system manager for the software and therefore be responsible for installing and maintaining it and administering the user levels of others. Note that satellite recorders with their own copies of the software will be the system managers for their copies. As soon as the nominated person installs Recorder 2000 they should assign themselves as system manager and change their password; they should then check that they can access the software under their own name, and then delete the default user.

The LRC should develop a procedure to decide what privileges different users should have. Different user levels have different training needs. Whenever a new user starts, the system manager should use the procedure to decide what privileges that person can have, and then add this information to the system. All users should tell the system manager their password in case they forget it. The LRC should avoid having unnamed users on its system.

23.6 Data custodianship

The ease with which Recorder 2000 users can exchange data raises the issue of data custodianship. The central idea of data custodianship is that there should always be a single master copy for a data-set; many slave copies can be passed out from the master copy, and they can have existing records augmented and new ones added; but such changes should regularly be passed to the master copy for collation and distribution to all slaves. The aim is to clarify the ownership and management of data-sets, to ensure there is at least one full copy of each data-set available to the owner of that data-set, to simplify data flow between different copies of Recorder 2000 and to reduce duplication of records. Data custodians tend to have a role in passing

the data-set to other nodes in the NBN, including the NBN Gateway, and in managing the process of verifying the accuracy of records, if not actually verifying the records themselves.

For each data-set, the LRC should agree with the owner and other holders of copies of the data-set who the data custodian will be and how the data will be managed. The data custodian can use Recorder 2000 to manage the data-set. It is particularly important to agree data custodianship when an LRC or one of the other copy holders first implements Recorder 2000 (or any other incarnation of the NBN data model) and starts to transfer data from other software into Recorder 2000. This is because the LRC and the other copy holder may transfer the data-set to Recorder 2000 without reference to one another, thereby creating two copies of the data-set with completely different GUIs (see section 23.2.2 *Increase the flow of data within the NBN*). Once these data have been passed on to other copies of NBN compatible software, it can be very hard to track down and remove the duplicated set.

23.7 Installing the software and replacing the back-end database

When the software is installed, the user has to enter a site ID (identification code) and a security key; these must agree before the software can be used. The site ID contributes the first eight digits of each record's GUI.

The data that is stored in Recorder 2000 is actually kept in a back-end database called nbndata.mdb. It is possible to substitute one copy of the file for another. However, you should not do so, because the GUIs for records are created in the Recorder 2000 front-end application, rather than the back-end database, so you would run the risk of creating duplicate GUIs in the two different copies of the back-end database. All the user configuration settings would also be lost. The import tools should be used to alter the contents of the back-end database.

23.8 Migrating existing data to Recorder 2000

There are several issues to consider when moving data-sets from the existing system to Recorder 2000.

23.8.1 Ongoing data entry

When preparing to implement the software, the LRC probably does not need to stop entering new data into its existing systems, particularly if these systems already hold large amounts of data. Recorder 2000's slow data entry speeds mean that there will usually be a backlog of data waiting to be transferred. Furthermore, the process of importing data from existing systems is unlikely to be significantly slowed down or complicated by any new data that you add to them after you have made the decision to adopt Recorder 2000, provided that new data entry is sufficiently rigorous to ensure a smooth import into Recorder 2000 (see section 23.8.2 *Planning and preparation*).

23.8.2 Planning and preparation

It is likely that the change-over to Recorder 2000 will be a time-consuming and complicated process. The LRC should conduct a trial before it undertakes the change-over proper; seek expert help (eg from its Recorder 2000 reseller); keep an eye on the NBN website for the latest news and advice; and *always* back up its existing data before attempting the change-over in earnest.

Data needs to be prepared before they are imported. The system manager should:

- decide on the observations and locations hierarchies for records, document it and prepare to implement it as the data are imported
- repair breaches of referential integrity in the existing systems
- check that data required by Recorder 2000 are present
- check that the data comply with Recorder 2000's validation routines

23.8.3 Upgrading from Recorder 3.3

The Recorder 2000 CD-ROM has an update for Recorder 3.3 (to update it to version 3.4) that includes both an export routine to allow export of data from Recorder 3, and an import routine to allow these data to be imported into Recorder 2000. Note that Recorder 2000 will import records for species created by users in Recorder 3.3, however once imported these records cannot be exported to another copy of Recorder 2000.

23.9 Mapping and GIS

When installing the software, you should plan with great care which base map you will use in Recorder 2000: if you want to change it, you will have to delete it along with all the polygon data that you have generated using that base map.

Recorder 2000 is capable of some of the data management tasks that would more usually be accomplished through GIS (eg managing boundaries of locations and information about the administrative areas in which they lie). However, GIS should be used for all but the most simple mapping. This can be done by:

- exchanging data between the two by batch imports
- building links between Recorder 2000's back-end database and the GIS
- developing a dedicated add-on module to act as an intermediary between the two

It is for users to decide whether they should make live links to GIS or rely on periodic batch updates, depending on their software and hardware and the frequency with which the data change. Making live links between the two applications raises the question of whether to allow them only to read from one another, or to write to one another too. Resellers of Recorder 2000 can provide advice.

GIS users should be aware that Recorder 2000 can import data from MapInfo or ArcView. However, it uses an unusual file format (GSF files) to manage polygons, which cannot be exported to these packages.

23.10 Satellite recorders

Many LRCs work with a network of satellite recorders. Each LRC needs to agree and document the details of its relationships with satellite recorders, including:

- ownership of software (eg does the LRC give Recorder 2000 to a recorder or just loan it to them?)
- ownership and custodianship of the database stored in the software
- ownership and custodianship of the original data-sets and their constituent records
- data exchange systems
- access to the data and information derived from it

23.11 Training and support

An LRC should have an ongoing programme of training to develop its staff and volunteers' IT skills. However, when first implementing the software, the LRC should budget not only for the cost of the software itself, but also for the one-off capital cost of training users.

In planning training, the LRC should consider the differing needs of staff, satellite recorders, volunteers and other users, and should try to integrate Recorder 2000 training with other training on LRC systems (including both IT systems and other policies and procedures). It may be a good idea to base each user's training needs on their user level. The following table makes some suggestions.

User level	Possible training and support needs
System manager	Network administration Database tools (eg Access 97, SQL, Visual Basic, links to GIS) The NBN data model Data custodianship NBN data exchange standards Data entry and validation Viewing and reporting data
Full edit	The NBN data model Data custodianship NBN data exchange standards Data entry and validation Viewing and reporting data
Add only	Data custodianship Data entry and validation Viewing and reporting data
Record cards only	Data entry and validation Viewing and reporting data
Read only	Viewing and reporting data

It should be possible to organise training through a Recorder 2000 reseller, but the LRC may be able to deliver some in-house, particularly on basic topics like viewing data, simple reporting, and using recording cards.

Once users are trained, they still need support. The LRC needs to anticipate the level of support that both new and seasoned users will need, and plan and budget to deliver this. As with training, it may be possible to deliver much of this support in-house, depending on the expertise and staff time available within the LRC. Otherwise, support time must be bought in from a reseller or from another individual or organisation using the software.

23.12 Using Recorder 2000 in an LRC

Recorder 2000 is a sophisticated application, and there are a number of issues relating to its use to be considered by LRCs.

23.12.1 Validation and checking

Validation. Anyone entering data into Recorder 2000 might enter inaccurate information, either because the record they are entering is incorrect, or because they have made a mistake transcribing the data onto the computer.

Recorder 2000 automatically checks data as it is entered to see if either kind of mistake has been made; this process is called 'validation'. When dates and spatial references are validated, it makes simple comparisons between the data being entered and the data already on the machine to see if they are consistent. For example, the software does not allow you to enter a record if its date is earlier than the date the recorder was born.

See section 6 *Species identification—verification, Case study three* for more information about how Recorder 2000 validates taxon records. Note that the same technique can be used to validate biotope records, except that this has not yet been implemented in Recorder 2000.

Items that have failed validation will not be exported or imported by Recorder 2000, and, unless the user specifies otherwise, they will not appear in reports. They can have other items attached to them, but any such dependent items fail validation because their parent item fails.

Checking. Recorder 2000 has a facility for tracking whether records have been faithfully transcribed during data entry: each taxon or biotope occurrence has a flag to say whether it has been checked for correct transcription. Only checked records can be exported.

23.12.2 Dictionaries

Recorder 2000's taxon dictionaries have been designed to be easy to use, to reduce the temptation for people to write their own lists and so lose the benefits of shared taxonomic lists. New facts can be added to checklists, as can new facts about checklist items, such as their local status, but these will not be exported. Creating new checklists is primarily the duty of NBN partners (notably JNCC, the Natural History Museum and national recording schemes and societies) and requires a thorough understanding of the NBN data model (as implemented in the dictionaries), of the design of the Recorder 2000 application, and of how Recorder 2000 populates the back-end database.

The dictionaries have been designed to allow items in checklists to be translated from one to another (eg translating bluebell to *Endymion non-scriptus* to *Hyacinthoides non-scripta)*. However, at the time of publication, this facility is not yet in place.

The taxon dictionary assigns each taxon to a taxonomic rank, from kingdom down to hybrid, sub-species, form, race or variety.

23.12.3 Data not managed by Recorder 2000

There are several data items that LRCs might need to add to Recorder 2000 before being able to get the most out of the software. It is possible for a person skilled in the use of Microsoft Access and who understands the NBN data model to add tables to the database, and the means to populate those tables. However, users should be aware that not all of the NBN data model has been implemented in Recorder 2000 and that missing data items may be included in the model, in which case they should follow the model instead of creating a new way to tackle the same problem.

23.12.4 Add-on modules

Recorder 2000 has been designed so that it can be extended in a number of ways. Users can:

- design new data entry forms (eg to simplify data entry for quadrat data or to match data entry screens more closely to recording cards used in the field)
- add new import and export methods (eg to import data from a spreadsheet)
- add new tables to the underlying database (eg to manage simple geological data)

Some add-on modules, such as one to import data from spreadsheets, are available to download free from the NBN website. It is expected that there will be a proliferation of add-on modules as more and more people start to use Recorder 2000, and it may be that an accreditation system is introduced so that users can be sure that add-ons will function as advertised.

LRCs should consider using the NBN website and other forums to share ideas about what add-on modules they want to develop; in this way they can share the costs and the risks of development.

23.12.5 Other databases

Many people already use a different proprietary database or have created their own; some will want to convert to Recorder 2000 and others will stick to their current software. Either way, this leaves the problem of how to export data to Recorder 2000; solutions include using a generic spreadsheet import tool (eg the add-on module available from the NBN website) or custom-building data export tools for the database in question (eg Biobase will export data in NBN transfer format using such a tool).

LRCs should encourage data collectors to convert to Recorder 2000. The software has the following advantages for collectors:

- Recorder 2000 users will find it easy to import, and so view and use data from the LRC and others using the NBN transfer format.
- LRCs can supply Recorder 2000 at low prices or even for free.
- LRCs can supply site maps and habitat details to Recorder 2000 users, making recording easier and more interesting for them.
- Statutory agencies such as English Nature plan to insist that all raw data they use comply with the NBN data model.

24 Other biological recording software

Summary

- LRCs should use NBN data standards in their data management. This should include promoting the standards with recorders, which can often be done effectively by encouraging recorders to use compatible recording software.
- There is a range of biological recording software, other than Recorder 2000, that recorders may find useful and which can export data to Recorder 2000.

24.1 Background

A database is simply an ordered collection of data that can be queried to provide information. A single table listing the 'who, what, where and when' of observations can be regarded as a database, but normally databases are collections of many tables of data that relate to one another (eg list of sites, list of recorders, list of species names, list of observations). If we wish to share information between databases, we must adopt consistent ways of storing and representing our data, and use names and terms that are widely understood.

These standards include both operating standards, that provide quality control, and data standards, concerned with the content and format of data. Many of the standards can be supplied or supported through the use of suitable software and should inform the LRC's choice of database software.

Problems often arise through the lack of common data standards or the lack of computers and suitable software to deliver those standards. The problem of processing large quantities of record cards and paper files is made worse through the use of poorly designed forms and lack of control of recording terminologies. LRCs should work with their recorders to seek ways in which the workload of data transcription can be reduced. The LRC partnership should consider providing, or helping with the purchase of, computers and software for key recorders or scheme organisers to do their own data entry and validation.

The NBN has done much to define and consolidate data standards for biological recording. LRCs must observe and comply with NBN standards if they are to play a full part in the national network and if they wish to achieve NBN accreditation. To achieve this, LRCs must work with their key data suppliers to ensure that the use of standards is disseminated through the local network. A good way of achieving this is through promoting the use of appropriate software.

24.2 NBN data standards

The NBN relies heavily on the adoption of commonly recognised standards for the format, content and terminology of biological records. The NBN, and its partners, have been active in defining data standards and supporting work to create new software that uses them. An explanation and details of all the NBN data standards can be downloaded from the NBN website (www.nbn.org.uk). The NBN data standards are based around:

- The NBN data model
- NBN dictionaries
- The NBN metadata model
- The NBN data transfer standard

It is important to understand that the task of defining and agreeing standards is an ongoing one. Unfortunately, with data standards it is not possible to create a single, unchanging data model or a fixed 'approved' list of terms. Requirements and usage change constantly, and we need a sophisticated mechanism for dealing with them.

24.2.1 The NBN data model

The NBN data model is a formal description of the relationships between items of information (the logical model), which provides a framework that can be used to describe the relationships between all aspects of biological data.

The NBN data model can be expressed in many ways, and many very different database applications are likely to be developed on the basis of the model. The important feature is being able to 'map' the data between applications and to know how the terms and codes used relate to each other. Recorder 2000 is based on a large sub-set of the NBN data model and employs both dictionary and data transfer standards.

24.2.2 NBN dictionaries

One of the problems that arises when we attempt to exchange or merge biological records is that different surveys or recorders often use different recording terms, taxon names or habitat classifications. A frequent problem has been that, because of the lack of freely available, up-to-date lists, recorders have often resorted to making their own.

The lack of standard dictionaries and terms can lead to:

- problems with quality control in recording projects
- potential confusion over meaning
- lack of comparability of terms between databases and, therefore, difficulties with data exchange
- inconsistent data retrieval

The following NBN dictionaries have therefore been developed:

- **Taxon dictionary**. The taxon dictionary is not simply a single list of currently accepted names, but includes many cross-related lists including all names used in UK and European conservation legislation.
- **Biotope dictionary**. The biotope dictionary includes all of the commonly used habitat and biotope classifications, including the NVC, NCC/RSNC Phase 1 Survey and Europe-wide classifications.
- **Administrative areas dictionary**. The administrative areas dictionary includes hierarchical lists of recognised administrative and geographical areas, including old and new counties, unitaries, districts and parish councils. It also includes other area names such as national parks, sea areas and conservation agency natural areas.
- **Protected sites dictionary**. This dictionary (currently under development) includes all UK sites designated under UK and European legislation and conventions, including NNRs, SSSIs, SPAs and SACs.

These dictionaries should be used as the basis for recording, whatever software is used.

24.2.3 The NBN metadata model

The NBN metadata model provides a means of describing the origin, availability, content, ownership and constraints on use of individual data-sets. It provides a template for holding this information, and also provides the means for creating indexes to the LRC's data holdings. See section 18 *Documenting data*.

24.2.4 The NBN data transfer standard

The NBN data transfer standard is derived from the NBN data model. It provides a means to exchange information between databases or between databases and other applications such as websites. To achieve this, items of information are enclosed by 'tags' that describe what the items are and how they relate to each other. The format used for the tags is called Extensible Mark-up Language (XML). XML, which is similar to the Hypertext Mark-up Language (HTML) that is used in World Wide Web pages, is rapidly becoming the main way in which heterogeneous information is moved in electronic systems. XML export and import using the NBN data transfer format is used in Recorder 2000 and has also been adopted by other commonly used recording software, including Biobase and AditSite.

24.2.5 Recorder 2000

Recorder 2000 is a database application, developed by the NBN, that embodies NBN standards and provides the functionality required for collating and managing biological records from heterogeneous sources. Recorder 2000 is based on the NBN data model and incorporates NBN dictionaries, data validation and transfer standards. In its present configuration, Recorder 2000 is particularly suitable for LRCs but is also suitable for use by volunteer recorders. However, Recorder 2000 may not be suitable for all naturalists, and even when it is, there is a need for support and training. See section 23 *Recorder 2000.*

24.3 Other biological recording software

Although many LRCs will use Recorder 2000 and will be able to support and supply copies to satellite (outposted, associated) users, Recorder was not written as a competitor to other biological recording software. The principle should be for people to use what is most suitable and convenient for them, but LRCs should always promote the dual objectives of common standards and mobility of data. Recorders supplying data to the LRC should be encouraged to use other applications when these are more appropriate (eg either simpler or more specialised to meet their needs); these other packages should preferably use NBN standards, particularly with regard to data transfer.

There are numerous PC-based biological records databases on the market either as freeware or at low cost. Details and availability of programs change regularly, so the LRC should keep track of websites and talk directly to recording groups if it wishes to advise individual naturalists on what is suitable for them. Recorder 3.x is still widely used by amateur and professional naturalists and version 3.4 has an export module to Recorder 2000. It is likely that some users may choose to stay with 'old' Recorder for some time and, as official support for the package ceased in June 2001, it is necessary for the LRC to retain some knowledge of this software or to identify other sources of help for the future.

Of the other applications available, some were written for general recording but there are also many written for specific taxonomic groups such as birds (eg COBRA, Bird Recorder, Avianstore), butterflies (eg Transect Walker) and amphibians and reptiles. A number of programs that were written initially for one recording group (eg MapMate was written for moths) have been gradually extended by the addition of new dictionaries and facilities. BioBase can cover general recording but has also been modified to provide more specific versions for individual taxon groups (eg mammals).

Some programs are written and supported by a single individual; others are supported by groups, especially for the addition of up-to-date taxon lists. Many volunteer recorders keep their data on home-written databases and spreadsheets.

None of these programs is suitable for use by LRCs as their main data management tool, but they certainly do have a role in data collection and local networking of volunteer recorders.

A growing number of biological recording applications are adopting NBN standards or have mapped their data structure to the NBN data model so that data can be transferred. Some have adopted the NBN XML data transfer standard. Applications of this type should be easy for LRCs to import data from into Recorder 2000; they may also be more suitable for general or non-technical recorders than Recorder 2000 is. Even simpler applications or home-written programmes should be able to produce data in flat spreadsheet-type forms that can be imported using the Recorder spreadsheet import wizard.

The number of applications that advertise an ability to interface with or export data to Recorder 2000 is increasing steadily. Some of those of particular relevance are summarised below; where costings are given, they are correct at the date of publication.

24.3.1 BioBase

Thurner Automation, Winton, Lower Ham Lane, Elstead, Surrey GU8 6HQ

01252 702563

MikeTAuto@aol.com

£75 plus VAT

BioBase is a simple-to-use yet powerful Windows-based computer package for taxon-specialised biological recording. It has been adopted as the standard for county recorders by The Mammal Society, the Bat Conservation Trust, the British Dragonfly Society, the British Lichen Society and the Bees, Wasps and Ants Recording Scheme (BWARS). BioBase is approved and used by BSBI and by the Herp Groups of Great Britain and Ireland. Versions are also available for bryophyte and beetle recording. A general version is available for users to configure with their own species checklist(s), stage, abundance, habitat, status and other reference data. More information is given in the attached case study.

24.3.2 MapMate

Teknica Limited, The White House, Montacute Road, Stoke Sub Hamdon, Somerset TA14 6UQ

www.mapmate.co.uk

Freeware but a registration fee of £10 is payable

MapMate is a not-for-profit biological recording application written *by* recorders *for* recorders. After two years of development and one year in 'real use' trials by the Somerset Moth Group, MapMate is now available for wider circulation. Originally developed for moths, it is now finding use in many other orders. MapMate is compatible with the NBN data model as realised in Recorder 2000. It uses a simplified 'sub-set' of Recorder data fields and the same (MS Access) database system.

24.3.3 AditSite version 3.4

Adit Limited, Tyn Rardd, Dwyran, Anglesey, LL61 6AJ

01248 430075

A wildlife recording system that runs under Windows and can be used to record sightings of birds, butterflies, moths, mammals, reptiles, wild plants etc, using built-in species lists. It can be used to plot records using an associated new mapping program, Aditmap.

24.3.4 Bird Recorder and Pocket Bird Recorder

Bird Recorder 32

Standard edition £70

World edition including world species distribution £97.50

Professional edition including world species distribution £130

Pocket Bird Recorder for Palm Computing Platform (Palm Pilot)

£35 (£25 if purchased with Bird Recorder 32)

Pocket Bird Recorder for Handheld PCs

£35 (£25 if purchased with Bird Recorder 32)

Pocket Bird Recorder for Palm-sized PCs

£35 (£25 if purchased with Bird Recorder 32)

Bird Recorder 32 comes with a complete and up-to-date world species database or optionally a local database (Western Palearctic or Nearctic). World version contains distribution information for all the countries of the Western Palearctic and all the states of the United States. Distribution for the rest of the world is available as an extra option.

24.3.5 Avianstore

Wes Halton, 5 Westland Ave, Farnworth, Bolton, Lancs. BL4 9SR

01204 709302

Wes@cygnus.airtime.co.uk

£50

Avianstore is a bird recording database species which uses the Masterfile Professional (MPRO) database system from Campbell Systems. It comes with a with full Western Palearctic list, divided into common and rare.

24.3.6 DMAP for Windows

Dr Alan Morton, Blackthorn Cottage, Chawridge Lane, Winkfield, Windsor, Berkshire, SL4 4QR

01344 883929

www.dmap.co.uk

info@dmap.co.uk

DMAP is a widely-used computer program which produces distribution maps, multi-species maps and coincidence maps, which are displayed in colour on-screen, with multiple zoom capability. The maps can be printed at high resolution in colour or black-and-white, or they can easily be exported as high-resolution publication-quality images to other software for incorporating into reports or publications. The distribution data for mapping can be entered into a text file, a spreadsheet, a database, or a biological recording package (including AditSite, BioBase, IPMR and Recorder 2000). The distribution data can consist of grid references, a variety of other metric co-ordinates, or latitude/longitude. A web page specifically for people who are interested in using DMAP with Recorder 2000 has been set up on the DMAP website.

Case study

BioBase

BioBase is a simple-to-use yet powerful Windows-based computer package for taxon-specialised biological recording. It has been adopted as the standard for county recorders by The Mammal Society, the Bat Conservation Trust, the British Dragonfly Society and the British Lichen Society. BioBase is approved and in use for vascular plant recording by BSBI (fully meeting their data transfer standard), by the Herp Groups of Great Britain and Ireland and by the Bees, Wasps and Ants Recording Scheme (BWARS). Versions are also released and in use for bryophyte and beetle recording. A general version is available for users to configure with their own species checklist(s), stage, abundance, habitat, status and other reference data.

BioBase facilities include:

- examination, editing and new entry of sites, people, literature and museum record sources
- examination, browsing, editing and new entry of recording cards (allowing entries for recorder, date, location details and species details)
- single species entry from species recording cards
- reports (to screen, printer, word processor file and Excel spreadsheet file) of: records for a species; records for a site/master site; a species list for a site/master site; a species list for a 10 km/1 km square
- distribution/coincidence mapping via a link to the DMAP mapping package
- import/export with other BioBase systems
- import/export with Recorder 3 centres
- export to national recording scheme (eg BRC)
- export to NBN (eg Recorder 2000) using the NBN data exchange standard

Records used for reports and distribution mapping may be limited by date range and can be marked as confidential with the option to be excluded from reports, maps and exports.

Codes (eg BRC codes) are used for species and, if allocated, may be used for people, but Recorder 3 codes are also stored and these are used for export to Recorder 3. Recorder 3 codes are used for other information (eg habitat, abundance and site status).

The NBN export is supported by the inclusion of NBN (Recorder 2000) codes for standard items including species (taxa), habitats (biotopes), counties and vice counties (administrative areas) and by the generation of term lists for some items with user-defined codes. In addition to the observer, BioBase has provision to record the determiner who has validated the records, and a facility to mark records as acceptable, in need of confirmation, or known to be incorrect. The XML export file can be imported by any NBN-compatible system (eg Recorder 2000), and the data are then placed within the structure of the NBN data model.

BioBase is designed for stand-alone use by taxon-specialised field observers and by county recorders undertaking collation. By virtue of the inclusion of external codes, BioBase can also interchange data with Recorder 3, and can be used as an efficient data capture system for Recorder 2000. A suite of ARev import/export procedures is available to support the Recorder 3 end of the data exchange link. Recorder 2000 has a standard NBN XML import facility.

The BioBase to BioBase export/import facility permits groups of recorders (eg members of a county bat group) to share their data.

Data entry is performed largely by mouse action from drop-down lists. Species can be entered by selection from a checklist of scientific names or common names (if available), or by species code. Checks are made on the validity of numeric, date and grid reference formats. Records may be related to a defined site or to a location which is unrelated to a site.

Additional information appropriate to the taxon version can be added to species records, such as abundance, search method, breeding signs and activity observed. BioBase can thus be used very simply as an occurrence recording system or, with extra data entry, as a more detailed research recording system.

BioBase is an application under the run-time version of the MS Access database supplied with the package. If the user has a developer version of Access 2, Access 7, Access 97 or Access 2000, then the open system interface (BioExt) can be used to attach to the BioBase data tables (read-only to protect the integrity of the data) so that any required queries can be implemented, custom reports generated and custom analysis performed on the data. Since the underlying database is MS Access, many other programs such as spreadsheets, word processors, DTP and GIS can readily use the data.

BioBase is available at an end user price of £75 plus VAT. Organisations (eg national recording schemes) can

benefit from a 20 per cent discount (details on application) and can then pass the package on to their members at the discounted price of £60, or with a further subsidy at their discretion.

BioBase is supplied by Thurner Automation

Winton, Lower Ham Lane, Elstead, Surrey GU8 6HQ

01252-702563

MikeTAuto@aol.com

25 Geographical information systems

Summary

- LRCs should use GIS to help them capture, manage, manipulate, analyse and display data with geographical or spatial elements.

25.1 Background

The future of biodiversity data management and product supply will increasingly involve the use of geographical information systems (GIS). If LRCs do not already own or have access to GIS, it should feature in their development plans.

The scale and scope of GIS packages vary considerably. Geographical information systems are able to capture, manage, manipulate, analyse and display spatially referenced data (ie any data that can be related to a grid reference, latitude/longitude or a boundary that can be drawn on a map). Some are geared more towards the capture and manipulation of maps (digital mapping) whilst others have specific analytical capabilities. There are specialised GIS for the manipulation and analysis of remote sensing data (eg satellite images).

Most GIS display data in layers that can be individually selected or merged, usually against a map backdrop. Data in GIS may be either vector (ie made up of lines and points) or raster (essentially spatially referenced images). Specially 'ortho-rectified' aerial photographs can also be used as raster backgrounds. GIS displays may be two-dimensional, but there are also a growing number of resources for three-dimensional visualisation, including the ability to 'drape' aerial and satellite images over digital elevation models.

The larger GIS applications, such as those produced by Intergraph, LaserScan and ESRI, are modular, and a full set of modules for use on a network can be hugely expensive, but there are now many smaller packages available that are suitable for stand-alone or small network use. Prices for desktop GIS range from around £1,000 downwards. Price is not necessarily a guide to functionality; some very powerful systems can be obtained freely or very cheaply, and many are available on the World Wide Web. Some GIS functionality is now often available within other applications (eg within some computer-aided drawing packages, such as AutoCAD), and most route planners allow the display of user-defined data. Recorder 2000 includes some GIS functionality in its map-viewer, and this may prove an attractive and valuable tool for users. The two desktop GIS most widely used in LRCs are ArcView and MapInfo.

25.2 What are GIS used for?

GIS is primarily a computerised tool for managing data and solving problems that have spatial components. The typical uses are:

- computerised cartography: digitising and displaying maps and boundaries
- database querying (eg what sites of importance lie within two kilometres of point X or corridor Y? What is at this point?)
- data visualisation (eg presenting distribution data against a map back-drop). Most GIS can display quantitative data in symbolic form (pie-charts or histograms) or as shaded areas.
- spatial analysis: analysing data in a geographical context (eg how the known distribution of a plant community relates to soil, geology, slope and drainage, and how much of the resource falls within designated sites; finding the optimum route for a new road, balancing cost against environmental impact)
- predictive modelling (eg what effect a building development would have on the wildlife resource of a defined area or what the effect of a one metre rise in sea level would be)
- analysing remote sensing data: display and analysis of aerial and satellite imagery (eg identification of habitat types from false colour images)
- resource management: applications range from local authority housing stocks to management of woodlands
- delivery of multi-media information linked to maps (eg map-based query and display on websites, including hyperlinks to other sites)

Some of these uses are very specialised and require expensive equipment, software, images and training to achieve.

25.3 LRCs and GIS

Every LRC should have access to GIS facilities, and for many GIS will become the main tool for querying the database and for the development of products. GIS applications are by their nature complex pieces of software, and some of the products on the market are arcane in their layout and controls. Wherever LRC staff are using GIS, it is essential that adequate resources are invested in good quality training. Of equal importance is the need for appropriate hardware to ensure that software, data and staff time are used effectively.

A hosted LRCs might not have its own software but be linked to a larger system maintained by its host organisation.

25.3.1 GIS use in an LRC

Normally, the first use that an LRC will make of GIS is for presentational purposes such as displaying distribution maps and, later, the printing of Phase 1 Survey maps or planning control alert maps.

Digitising facilities may be used to capture site boundaries or habitat boundaries within sites, which can be used in display products, area analysis and for dissemination to naturalists to aid recording.

Once there are sufficient data in the database, including habitat boundaries and observational records, the GIS can become a valuable tool for making data queries (eg providing alert maps or a means of extracting species records in user-defined polygons).

GIS can be used to analyse data in many ways, including coincidence modelling where species distributions can be mapped against other species, habitats and environmental variables. Such analyses can suggest new locations to search for uncommon species or associations. Temporal data can be combined with environmental variables to map and analyse species spread or decline. More sophisticated analysis can include using non-biological data to predict environmental impact from developments (eg building a new car park adjacent to ecologically fragile areas such as sand dunes).

It should be noted that even in LRCs that do not yet have GIS, much can be achieved with the map functionality included in Recorder 2000. Although not an analytical tool, Recorder's maps can display data and polygons and be used to query the database in spatial terms. They are also fully integrated for data entry.

25.3.2 Establishing GIS

For a new LRC or one that does not yet have GIS facilities, planning for the establishment of GIS should be a high priority. In considering how to establish GIS capability, considerable thought must be given to a range of issues, notably resource constraints that often limit the use of GIS by LRCs, including:

- cost of purchasing and installing the software (probably the cheapest bit!)
- cost of purchasing suitable computers (GIS software often has many menus, control bars and information sub-windows. This means it can be difficult to see the maps unless you have a big screen such as a 19" monitor—and these are costly)
- cost of obtaining a suitable set of digitised maps (see section 26 *Ordnance Survey data*)
- need for staff training
- lack of data of suitable quality—you often do not discover this until you start trying to analyse your existing data!
- lack of digitised boundaries
- time—some GIS take a considerable amount of time to set up (eg installing and registering maps, defining and setting up layers, linking to data tables). Time is needed to identify and obtain items such as essential boundary data. Even with training, there can also be a significant 'learning curve' if you are getting to grips with GIS on your own.

Fortunately, all of these limiting factors can be overcome, and many established LRCs now see GIS as an everyday working tool which is indispensable to their function. The LRC might be able to tap into the GIS resources of a host organisation or one of the key partners and operate under an extension to their licence. Even if direct access to a partner's GIS is not possible or convenient, it should be possible for the LRC to take advantage of partner resources such as digitised administrative boundaries and training. Otherwise, LRCs should aim to purchase a desktop GIS such as MapInfo or ArcView. In the short term, good use can be made of the freely available mapping facilities within Recorder 2000, and distribution maps can also be produced using Dmap (see section 24 *Other biological recording software*).

25.4 References and sources of further information

There are many books and websites which provide a good introduction to GIS.

Burrough, P, *Principles of Geographical Information Systems for Land Resources Assessment.* Clarendon, Oxford, 1986 (since revised).

Langley, Goodchild & Maguire*, Geographical Information Systems, Principles,Techniques, Applications and Management*. 2 volumes, John Wiley & Sons, 2nd edition 1999. (highly priced)

A useful GIS bibliography is obtainable at:

www.agi.org.uk/pages/freepubs/bibliog

There are many sites on the web that give access to GIS resources and details of available products; see, for example:

www.geo.ed.ac.uk/home/giswww.html

www.freegis.org

MapInfo

www..mapinfo.com

ArcView

www.esri.com

Case study

Geographical information systems

Warwickshire Habitat Biodiversity Audit

Background

The Warwickshire Habitat Biodiversity Audit (HBA) is a project that is surveying habitats and managing habitat data for the county. The data generated will support the project partners work on identifying Wildlife Sites and on the LBAP. In the longer term, a data management system will be established that will enable partner organisations to monitor changes in habitats and land use.

What does the HBA use GIS for?

The project makes extensive use of geographic information systems (GIS) in editing existing data, adding new data, viewing and demonstrating GIS and analysing data. The nature of the data the HBA handles summarises the full range of data types that GIS creates, and includes:

- line (or arc) data (eg hedgerows)
- shape (or polygon) data (eg fields)
- spot (or point) data (eg an individual species record)

GIS can produce coded maps showing habitats and other features. More importantly, it can analyse the data it manipulates (eg by automatically and accurately measuring distances and areas). Moreover, when you enter information on to GIS you also create attributes that describe the thing that you have created. Therefore, any piece of information (eg a field) holds information about itself, including the:

- area of the field
- perimeter of the field
- type of habitat the field represents
- Phase 1 classification code
- type of label needed (eg 'I for Improved Grassland)
- colour or shade required for that habitat (only applicable in ARC/INFO)
- source of data (eg aerial photograph, field survey, survey review)
- date of source data

This means that different maps can be produced to show different things. Normally, Phase 1 habitat maps are used, but maps can be generated depicting the various sizes of fields or the ratio of field area to field perimeter, or a map showing the location of specific habitats of a particular area, or even a map showing the extent of data taken from aerial photographs compared to field surveys. For example, a recent task involved using habitats, such as rich grasslands and woodlands, to produce point data on a one-kilometre grid that shows the extent of each Broad Habitat type as a distribution map, where the size of the point corresponds to the extent of that habitat in any one-kilometre square.

The project is now heavily involved in producing strategic documents that promote issues of nature conservation in the Warwickshire area. The HBA has used GIS in conjunction with its data to provide sustainability indicators for the county and its unitary authorities.

What GIS does the HBA use?

All data entry is done using ARC/INFO (which is very expensive and quite slow at doing simple tasks). MapInfo is used for all arc editing (eg removing unwanted lines, such as buildings, from Ordnance Survey landline data, and adding vegetation or habitat boundaries), and is very quick at doing simple things and can run on relatively low-specification computers. ArcView is used to provide live GIS demonstrations and to undertake simple analysis. Complex analysis is done using ARC/INFO, with routines programmed in so that they can run overnight.

In summary:

System Type	Advantages	Disadvantages	Use by HBA
ARC/INFO	Extremely flexible and powerful with superior programming and data management capabilities.	Very expensive and requires high-specification hardware. Really needs dedicated staff to run the system.	All data entry, data management and the more complex analysis. The basis for translation of data to other GIS.
ArcView	Relatively affordable and quite easy to use. Fairly comprehensive but opportunities for add-ons. Windows-based and compatible with Microsoft products. Good analytical capabilities.	Cannot analyse aerial photographs or other images. Requires training and development to maximise use.	Providing demonstrations, map production, summary analysis.
MapInfo	Relatively affordable and easy to use. Very good and quick at arc-editing. Can display summary analysis well.	Cannot create and save feature code palettes. Advanced capabilities come as extra (such as programming).	Arc-editing, some analysis and report writing.

Contact

Ian Tanner
Habitat Biodiversity Audit, Warwickshire Field Services, The Butts, Warwick CV34 4SS
01926 412197
iantanner@warwickshire.gov.uk

26 Ordnance Survey data

Summary

- LRCs need access to digital geographical data to enable them to manipulate, analyse and present data effectively.
- An LRC can improve the accuracy of the geo-referencing of records collected by recorders by providing access to accurate maps.
- LRCs must conform to Ordnance Survey licensing terms.

26.1 Background

A key element of biodiversity data is geo-referencing. One of the major concerns of many users of biodiversity data is accuracy of geo-references; lack of precision can limit the potential uses of the data. This is particularly true of much historical data, which, for example, might have been collected with only a 10 km^2 grid reference. Use of accurate geo-referencing systems can increase the applicability of data to a wider range of uses.

With the advent of GIS (see section 25 *Geographical information systems*), there is an increased ability to manipulate data with spatial components, especially boundary data such as habitat polygons and site boundaries. GIS thus makes LRCs able to supply better products to users, and in a more efficient way.

The source of geo-references, whether they are required during data collection or for data manipulation and analysis, is normally Ordnance Survey (OS) data (in Northern Ireland, the Ordnance Survey of Northern Ireland (OSNI)).

Vector data are created by digitally capturing topographical features, such as buildings, roads, and rivers, as a series of lines and points. Each feature is given a unique 'code' and geographical co-ordinate; information about these features (eg the name of a river) can also be recorded. Using GIS allows these data to be manipulated, either to enhance presentations (eg choosing whether or not to show certain types of features) or to analyse data (eg searching for records within x km of a river).

Raster data are normally produced by scanning paper maps. The image produced is a series of pixels (a bitmap image) that cannot easily be manipulated. Hence raster data are of most use as backdrops.

26.2 OS data

OS paper maps are well known to many people. They are produced at a wide range of scales. OS also produces a range of digitised map products and associated data, including OS map 'tiles' and landline data, place name gazetteers and postcode spatial references. Maps are available in either vector or raster format. The smaller-scale maps are available as single files for the UK whilst the larger-scale maps (eg 1:50,000 and 1:10,000) are available as map tiles. The 1:50,000 map tiles cover 20 km x 20 km areas and the 1:10,000 tiles cover 5 km x 5 km.

The availability of different products changes rapidly and information should always be sought direct from OS or via their website.

26.2.1 OS data for LRCs

To be able to deliver the range of services that its users and recorders require, the LRC should use GIS to manipulate and present data and to supply map-based information to recorders. The LRC must ensure it has good systems in place for increasing the precision and accuracy of geo-references of biological data for its area.

To achieve this, the LRC needs a range of OS products.

Most LRCs, and many individual recorders, will have invested in paper maps at a range of different scales. These are commonly used for recording, and in many LRCs paper maps are still used to cross-reference manually, and sometimes to store data such as habitat parcels.

An LRC requires suitable backdrops to allow it to use GIS. As a minimum, this should be 1:50,000 raster data. This will allow the LRC to present data with a relevant backdrop, which provides contextual information and aids users with interpretation. However, to enable the LRC to use its GIS to increase the accuracy of data entry and provide more complex analyses of data, more detailed data would be required (eg data at 1:10,000 scale, and preferably in vector form).

26.2.2 Licences for OS data

OS data is Crown copyright and you need licensing to use them, and in particular to duplicate them, whether in paper or digital format. Paper maps may be drawn on and viewed, but may not be photocopied or published (either in whole or in part) without a licence. Digital data are licensed by OS for use over a given period of time (usually one year); licences are renewable and normally include arrangements for a number of hard copies to be made.

Local authorities and the statutory conservation agencies usually hold OS licences that enable them to provide OS information to third parties for specific, time-limited projects carried out on behalf of the licensed agency. The third party cannot retain any OS data at the end of the project. This means that, with the exception of LRCs which are hosted by organisations which have licences, LRCs cannot get by in the long term by borrowing a local partner's data; they need to hold their own licences.

The cost of licences is usually the critical factor that prevents LRCs from making full use of OS data. Charges are high and ongoing, and the wide range of products and variations according to type and number of users make for a rather complicated scale of charges. LRCs that wish to use OS maps on their websites or to provide GIS output containing OS data in their products must check copyright and charging issues.

The NBN Trust is seeking to negotiate a special rate for OS 1:50,000 raster map tiles for volunteer recorders to use with Recorder 2000. It is hoped that this will provide volunteer recorders with reasonable scale maps for use with Recorder 2000 and thus improve the geographical accuracy of their recording.

26.3 Other sources of digital maps

There are providers of digital maps other than OS; these include Bartholomews, the Automobile Association, Europa Technologies and Philips. Details of a number of suppliers of UK digitised maps can be found on the Graticule UK site (at www.graticule.com).

There are also a number of sources of interesting free maps and some satellite imagery. See, for instance, www.old-maps.co.uk for early OS maps and www.terraserver.com for interesting satellite images. See also www.mapdigger.com.

26.4 References and sources of further information

Ordnance Survey

Romsey Road, Southampton, SO16 4GU

08456 05 05 05

www.ordsvy.gov.uk

Ordnance Survey (Northern Ireland)

Colby House, Stranmillis Court,Belfast,BT9 5BJ

028 90 255755

osni@doeni.gov.uk

www.osni.gov.uk

27 Recorders and IT

Summary

- The LRC should provide support and advice to recorders on the use of relevant IT systems that will enable them to contribute data to the LRC effectively.

27.1 Background

LRCs need data from volunteer recorders, as they are a major source of taxon-based observations, and of much general site survey data. LRCs must, therefore, build close relationships with their recorders and have a clear understanding of recorders' individual needs (see volume 1, section *17 Relationships with volunteer recorders and recording schemes* and volume 2, section 13 *Working with data-providing individuals*).

There is often a range of attitudes and abilities amongst recorders in relation to their use of information technology. Some are expert and may even be a source of support to the LRC; others have progressed as far as using databases such as Recorder 3.3, Biobase or MapMate to manage their data; and others use completely manual indexing or rely totally on notebooks and recording cards.

Many recorders do not have a clear idea of what relevance or value IT could have for them. The main considerations in encouraging the take-up of IT are:

- how can IT help the individual recorder?
- how can IT help local collators?
- how can IT improve the flow and better use of data?
- the cost and time implications for the recorder

27.2 Encouraging recorders to use IT

Care needs to be exercised when encouraging individual recorders to start using computers or specific pieces of software. Most recorders want to be out recording or checking determinations, and to many, data management is a chore. For some, the learning curve will be very steep, and failure to understand the purposes and limitations of software can soon lead to negative attitudes. The benefits of computerising data need to be apparent, and there must be a support system which is sensitive to individual needs. The key messages for the recorder are that:

- good software can help to avoid errors. By entering their own data, recorders can lower the number of errors and reduce the amount of checking associated with providing record cards to scheme recorders or LRCs for data entry. Most database applications have checking and validation routines that pick up common errors, and most use dictionaries to check terms (eg site or species names).
- good recording applications help to introduce and maintain standards. Standards are wide-ranging and include: minimum record standards for data to be re-used or passed to others; controlled terminology; and formats for dates and grid references.
- data in electronic form can be easily copied and backed up, making them more secure
- data can be shared more easily, especially between compatible database systems. Recorders no longer have to transcribe or photocopy paper records in response to requests for information.
- use of good software makes it easier to receive data from others
- computerised data can be more easily manipulated, searched and analysed. Mapping and report programs make it easy to see what information you have in whatever format or combination you need.
- the increasing availability of digital maps, especially OS raster coverage and the ability to use digitised aerial photographs, is of great value both in locating records (checking grid references) and in visualising data
- access to the internet opens many new possibilities. Email simplifies the task of communicating with other recorders, with specialists or with the LRC, especially when attaching or receiving data files. Access to the NBN Gateway will provide an important source of information, including national distribution maps and portals to specialist recording networks.
- looking ahead, IT can help with data capture in the field. There is a growing market for handheld and 'pocket' PCs, small hand-held devices whose functions range from simple notes and diary entries up to versions of standard office software (eg Word, Excel, Access). Many include email facilities and programs for linking to other computers for data transfer. Biological recording

applications are already becoming available for these machines (eg Bird Recorder; see section 24 *Other biological recording software*), and their use is likely to increase in the future, especially when linked to low-cost, accurate GPS (global positioning system, which can calculate location to within one metre by triangulation of co-ordinates from satellites) for recording locations.

For local collators such as vice-county recorders or the organisers of schemes with wider geographical coverage, IT can provide a vital means of sharing data. Recorders entering their own data cuts down on the time collators must spend on transcription. Electronic data can now be processed more easily, particularly with the use of new software such as Recorder 2000, which shows whether records have been checked and validated and allows referees to record determinations and comments against records. Mapping software allows the coverage of records to be easily plotted and passed back to recorders to encourage further effort. Some programs inform the user when they have entered a new site or county record. Electronic data that conform to common standards can also easily be sent to the LRC or national scheme for incorporation into information products.

27.3 Recorders as data custodians

A data custodian is the person or organisation that holds and manages the primary record. Records may be copied to others for integration into bigger data-sets or for creating specific products (eg an atlas), but only the custodian has the editing right for the record, and all copies should be marked as such. The data custodian may be the original recorder, or the recorder may pass this role on to a local or national scheme organiser. Some recorders and schemes may pass the role to an LRC. There may be local and national data custodians with different rights to the same data. The aim is that it will always be possible to trace the original record and to be clear about the copyright and intellectual property rights associated with it.

The increase in the use of IT is making it possible for more people to be data custodians and to keep control of their data whilst still enabling free exchange and wider use of data. One of the key attributes of Recorder 2000 is the recording of the ownership of records and control of editing rights. It is also no longer necessary for the LRC to physically hold data for these data to be accessible and used in LRC products. A system where data are widely distributed can only work well if people observe common standards and if communications are properly maintained. LRCs have a key role to play in promoting standards and supporting the network of local recorders. See section 12 *Data flows.*

27.4 Supporting recorders in using IT

It is important that the LRC, as the focus and support for a local network of recorders and data custodians, is able to promote good data management and the application of consistent data management standards. It is also important that the LRC can provide both technical and 'moral' support for recorders, especially those new to computers or to specific pieces of software. The LRC should work with recorders to identify their individual needs and decide how these needs can be met.

The LRC can provide support to recorders in a number of ways. These might include:

- encouragement and education in the value of IT to themselves, the LRC partnership and the NBN
- guidance in the purchase of computer equipment
- guidance in the selection of the most appropriate software to use
- provision of software (eg as a distributor of satellite licences of Recorder 2000 or other applications)
- individual training in the use of software or organisation of joint training sessions
- ongoing contact and support. It is important for recorders to have someone to turn to for help with technical or software problems, or just for encouragement.
- provision of support data (eg lists and boundaries of designated sites and agreed recording sites for importing into their databases)
- provision of digital maps for users with software capable of displaying them
- provision of context data (eg habitat maps to use in conjunction with species distributions)
- financial support for the purchase of computers or software for recorders to do their own data entry
- data input services for recorders who do not have the facilities, time or inclination to do their own

Case study

Supporting recorders in developing the LRCs network

Centre for Environmental Data and Recording

Background

The Centre for Environmental Data and Recording (CEDaR) was established at the Ulster Museum in 1995 through grant aid from the Environment and Heritage Service (EHS; part of the Department of the Environment for Northern Ireland).

The CEDaR Project is a partnership between the Ulster Museum, EHS and the environmental recording community of Northern Ireland. CEDaR has a broad remit focused on collating, storing and disseminating information relating to the geology and distribution of the flora and fauna of Northern Ireland and its coastal waters. Biological records are currently stored on Recorder 3.3 and the Marine Nature Conservation Review (MNCR) database.

CEDaR and recorders

CEDaR co-ordinates recording within a formal and recognisable framework. In the last five years CEDaR has addressed the needs of the local recording community with considerable vigour. More than a million records are currently stored on the Recorder database. During the initial months of the CEDaR project, it became apparent that a mechanism for identifying available and relevant data-sets was required. To assist with this, a metadatabase of all appropriate data-sets was constructed using Microsoft Access. This database provided CEDaR staff with an effective mechanism for prioritising the computerisation of records.

Support for recorders

Significant effort has been expended in developing mechanisms for the transfer of data to CEDaR. Electronic transfer of data-sets has become the norm. To assist with the flow of terrestrial and freshwater data to CEDaR, an external network of Recorder-based satellite sites has been established. There are currently nineteen satellite sites being supported by CEDaR. Protocols to maximise the potential of the satellite system include:

- EHS grant aid purchase of computer hardware for appointed satellites
- EHS provision of assistance with purchase of Recorder software
- CEDaR staff providing (free) Recorder training and support for all Northern Ireland satellite sites
- the development of standardised *SITES*, *PERSONAL* and *LITERATURE* tables
- having satellite-specific default numbers for Recorders tables
- having a single individual responsible for the collation of records for a given group
- policies to show that ownership and copyright remain with data suppliers

Key elements of support provided by EHS are financial assistance and the provision of local software training. When coupled with the realisation that a satellite can have its data immediately accessible on a PC, and the understanding amongst recorders that their data will be given extra value by being combined with those of other satellites, the incentives for the creation of a strong local network was generated.

The formation of a network of satellite sites has also provided CEDaR with a solution to the major difficulties associated with the validation and verification of records and the problem of duplication of data entry. This assemblage of individuals, groups and societies constituted the Northern Ireland Recorder Users Group (NI RUG). The RUG has recently evolved into the NI Environmental Recorders Group. This evolution was an acknowledgement of the growing number of data-sets made available to CEDaR by non-Recorder database users. The Group meets regularly and is the main platform through which recorders are kept informed of local developments. Furthermore, it also provides recorders with an opportunity to express their views on CEDaR's activities and direction.

Money generated through charging consultants for time spent in handling their enquiries is used by CEDaR to help support the local recording community, and has assisted in the funding of several local initiatives.

Contact

Dr Damian McFerran, manager
+ CEDaR, Ulster Museum, Botanic Gardens, Belfast BT9 5AB, Northern Ireland
‘ 028 9038 3154
: damien.mcferran.um@nics.gov.uk