In Practice

INSTITUTE of ECOLOGY and ENVIRONMENTAL MANAGEMENT

BIODIVERSITY BEYOND 2010

Where is All the Data Going?

Roy Tapping
Manager, Cofnod – North Wales Environmental Information Service

Wales is lucky in many respects; fantastic scenery, a clean environment, albeit a bit wet at times, and a strong cultural heritage. It also boasts the first complete network of Local Records Centres (LRCs) anywhere in the UK. This is probably not as impressive as the scenery or culture, but for those who need access to high quality biodiversity information the Welsh LRC network, provides seamless coverage across Wales. For nearly a decade Welsh LRCs have offered data services to environmental consultants and over the past year and a half we have been working with IEEM's Welsh Section to better understand and cement our working relationship.

Initial engagement with IEEM was through workshops held in North and South Wales during 2009. This was followed earlier this year by a questionnaire sent out to 50 environmental consultancies and more recently a meeting held with government and private sector ecologists in North Wales. As well as generating better understanding between the complementary communities of ecological and data managers, these efforts helped us identify the key issue of data flow between consultants and LRCs from our discussions. This is the focus for this article and I hope to be able to put forward a perspective borne in Wales, but which is equally applicable throughout the UK.

The Local Records Centre I work for, Cofnod in North Wales, is relatively new to providing data to consultants. However in just three years we have built our customer base from zero to over 120 consultancies, many of which are repeat customers. Consultants appear to understand the charging model established through LRCs and in most cases their costs are recouped through their clients. I hear feedback from consultants saying they like the

quality and speed of our service, the way the data are presented and they find it useful background to planning their own surveys and desktop assessment. Yet for the hundreds of requests for information we have dealt with over the past three years, I can count the amount of data we have received from consultancies on one hand. This is not entirely surprising as there is little compulsion to supply LRCs with data. However, the same cannot be said for data that comes from volunteers and public sector organisations, which continues to increase annually.

I hear with increasing regularity that the greatest source of new data, especially for protected species, is through the thousands of privately funded surveys carried out each year by environmental consultants. So why are we worried about this? You could say it is in our interest, as we are concerned with gathering as much relevant data as possible. However LRCs are not the end users of data, we merely manage it; try to improve its completeness, its uniformity and its quality. So it should be the end users who are most interested in making sure they have access to the most comprehensive relevant data. Environmental consultants are major end users, making it important that others, and for that matter themselves, can potentially access improving data. In most cases valuable studies end up forming part of paper archives. Take documents prepared for the planning process, there seems to be a common misconception that once a document reaches the 'public domain', it becomes automatically available for all to use and consume, including LRCs. Yet in truth much of this data ends up in various paper or electronic filing systems. These systems tend to be cumbersome, making the data virtually impossible to cross reference. Eventually the data gets lost under the weight of more recent information and all the valuable time and effort used to create it goes to waste.

LRCs offer an alternative to this, but why is data not being shared more fully? From the engagement with IEEM we have had in Wales some common reasons have started to emerge:

- Concerns over ownership of the data and in particular whether there would be a breach of client confidentiality.

- A lack of clarity about what data are required and how to share it.

- The extra time and possible expense involved in sharing data.

Data Ownership

The issue of data ownership seems to be the most misunderstood; I for one have wrongly believed that the owner of the data was whoever paid for its collection. Although this holds true under certain circumstances, the owner of the data is generally the one who created it. This is a clear principle under copyright law. There are two exceptions to this. Firstly, data collected whilst in someone's direct employ becomes the property of the employer. Secondly, where there is a legal transfer of ownership, usually through a signed contract between the collector and another party. So what does this mean? Well for most freelance consultants it means that, unless specified by the client, any data collected is theirs freely to share. For consultancies with employees, the data is owned by the consultancy and unless specified by the client, they in turn can freely share it. Given the legal position there seems to be two main reasons for not sharing data. Firstly, procedural difficulties in, for example, persuading a larger consultancy it can freely share its data. Secondly, concerns over whether the sharing of such data will damage customer relations and be a breach of client confidentiality. Neither of these issues seem insurmountable, however the first relies on organisational change, the second could be just a matter of clarifying intentions to share data. For some time we have been promoting the use of the following data sharing clause by consultants asking them to use these when setting up contracts with their clients:

'Unless you request otherwise, we intend to share all relevant biological survey data with the appropriate Local Records Centre.'

This clause and other similar ones we have discovered are widely used by environmental consultants. By doing

so they not only demonstrate their commitment to share data, but also make their intension clear to clients, thus potentially mitigating any damage it may have on consultant/client relations.

In certain cases the nature of the project is so sensitive it makes it very difficult to share data. As an LRC manager I would say only share data with us if you feel empowered to do so. Furthermore if you wish to share data but want to restrict access to that data, we can look at the best way of achieving this. For example some data may only be sensitive until it is in the public domain, at this point the data could be shared or the restrictions lifted.

What to Share and How to Share It

If environmental consultants are saying that they do not know what and how they should share their data, this must be an issue for LRCs to address, by providing better guidance on how and what they should receive. However once the guidance is prepared it may be difficult to promote and turn it into accepted practice. There may therefore be a role for collaborative guidance developed between IEEM

and the body representing LRCs, the Association of Local Environmental Records Centres (ALERC). Such guidance could offer some practical assistance when following the IEEM's Code of Professional Conduct section 5, paragraph 5.7, which states:

'Wherever possible, make scientific data collected during the course of their professional duties available to others such as records centres.'

The Additional Time and Cost of Sharing

Very much linked to issue two, the final issue for not sharing data, the additional time and cost could also be looked at jointly by IEEM and ALERC. From the feedback that we have received it remains a major barrier to sharing, especially where contracts have tight financial margins. It should be our aim to make the sharing of biological data as painless as possible. In LRC circles we talk of collecting data once and using it many times. It makes little sense therefore that information which is already collected, copied and summarised is then reformatted for submission to LRCs. I think therefore we should initially look at ways to better share the data in its raw or summarised form, thus keeping extra time and cost

to a minimum. There may be technical solutions to this and ones which I believe are in the interest of LRCs and IEEM to discuss.

The improved sharing of biological data gathered as part of privately funded surveys should not be outside our realms of possibility, especially in an age which relies so heavily on information exchange. It does however require some effort from both environmental and data managers. It is in our interest to encourage a much improved data flow, for until we do there is always the potential to duplicate effort and draw conclusions from incomplete data. I hope this article has stimulated some thoughts on how the sharing of biological data could be progressed with LRCs, I also hope that it will encourage IEEM and ALERC to look jointly at tackling this issue. I know there are some practical solutions, we just need a little help to make them happen.

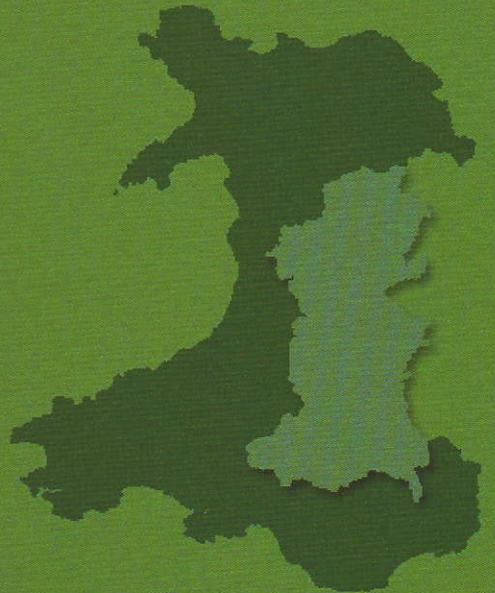Correspondence:
roy.tapping@cofnod.org.uk

Welsh Local Records Centres web portal (www.lrcwales.org.uk) acts as a gateway to their individual websites. Its useful 'Find a Local Records Centre in Wales' tool, searches for a location then displays details of the appropriate LRC.

Where is All the Data Going? An Extra View

Andy Tasker CEnv MIEEM
Development Director, Middlemarch Environmental, and Consultant, Priority People

'Where is all the data going?' was a challenge that I raised at IEEM's Nottingham Conference in November 2007, so it is good to see from Roy Tapping's article that some progress has been made. I estimated then that somewhere between £100 million and £200 million was being spent annually on acquiring ecological data that was mostly lost in consultancy reports sitting on client's shelves[1] - and in the intervening three years, despite the recession, the figure must be about the same now. That is an awful lot of potentially useful ecological information that we continue to throw away.

IEEM has been actively involved in looking for solutions, from meetings that I and the Secretariat had with both NBN and ALERC in 2008 through to the Professional Affairs Committee more recently. And of the three headings that Roy describes – data ownership, methods of sharing, and paying for it – I think that the main problem remains the detail of formatting and sending data.

To me it seems slightly strange that a succession of Government-funded initiatives have failed to solve this apparently simple issue, but I suspect it is because until recently the custodians of the data, from Local and Regional Records Centres to the NBN, have not been talking to the potential providers of data in the private sector. So some of the more bizarre ideas have included all manner of complex forms that providers have been expected to fill in, then send to the appropriate local centre. In reality this has not happened simply because of the time and effort of doing it.

Surely what is needed is a simple nationally-agreed format for collecting data on both species and habitats. This needs to be designed not just by data custodians – who may want undue levels of complexity – but by data providers too. If data were collected to this agreed standard in the field, it would then be a simple process for consultants to copy and paste it into reports, and at the same time e-mail it to a common data entry gateway. The need for this central gateway, rather than the local, regional and national systems that currently exist, is to make the process achievable for consultants and other providers. Clearly the data recorder will know the gird reference and/or geo-reference for their data – but which vice-county is it in? To be honest, who cares, apart from the data custodians? Given that all the data will be electronic, an automated reposting to the appropriate LRC would make everything achievable.

Imagine a future where all the ecological information that is collected can be simply stored and accessed! Imagine the huge amounts of current data that would then be available for everyone to use! Clearly there would need to be some safeguards where commercial clients required total confidentiality for a period of time, but even here the issues are not insurmountable.

I hope that IEEM can continue working in partnership with data custodians, data providers, agencies and governments to make this happen. Soon would be good!

Notes

[1] See a summary report in the Amphibian and Reptile Group newsletter of Autumn 2007: http://www.arg-uk.org.uk/Downloads/ARG%20Today%203.pdf

Correspondence: andy@priority-people.co.uk

Freshwater survey