Assembling Nature

the social and political lives of biodiversity softwares

Rebecca Ellis, Maria Pacha, Claire Waterton

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
science in society
CSEC
LANCASTER UNIVERSITY

Designed and Printed by statusdesign.co.uk 01406 380620, 1.5M.

Foreword

The proliferation of information technologies for documenting and representing the world's biodiversity has recently increased dramatically, in pace with society's concerns about accelerating species loss. The Rio Convention on Global Biodiversity in 1992 marked the global acknowledgement of a planet-wide biodiversity crisis. Since that time databases have been seen as an important tool for understanding the extent of biodiversity loss and for accurate documentation of that which remains. In this sense, databases have taken up a role as an important underpinning to national and global policymaking.

This booklet is the result of an ESRC Science in Society funded project, 'Databases, Naturalists and the Global Biodiversity Convention' that aimed to examine 3 different IT-based data frameworks used to gather, collate, exchange and represent data about the distribution of plant and animal species within the UK:

'Recorder'
'MapMate'
'National Biodiversity Network' (NBN)

Coming from a background in the 'social studies of science and technology' (or STS), the authors carried out interviews and discussion groups with designers, users and contributors to UK biodiversity data frameworks to examine the **construction** and the **use** of these data frameworks from a **sociological,** rather than a technical or biodiversity perspective. In other words, it aimed to extend and broaden the way that we understand these technologies into the social domain. In doing so it aimed to think through and beyond biodiversity databases themselves to consider the relationships between nature, technologies and society.

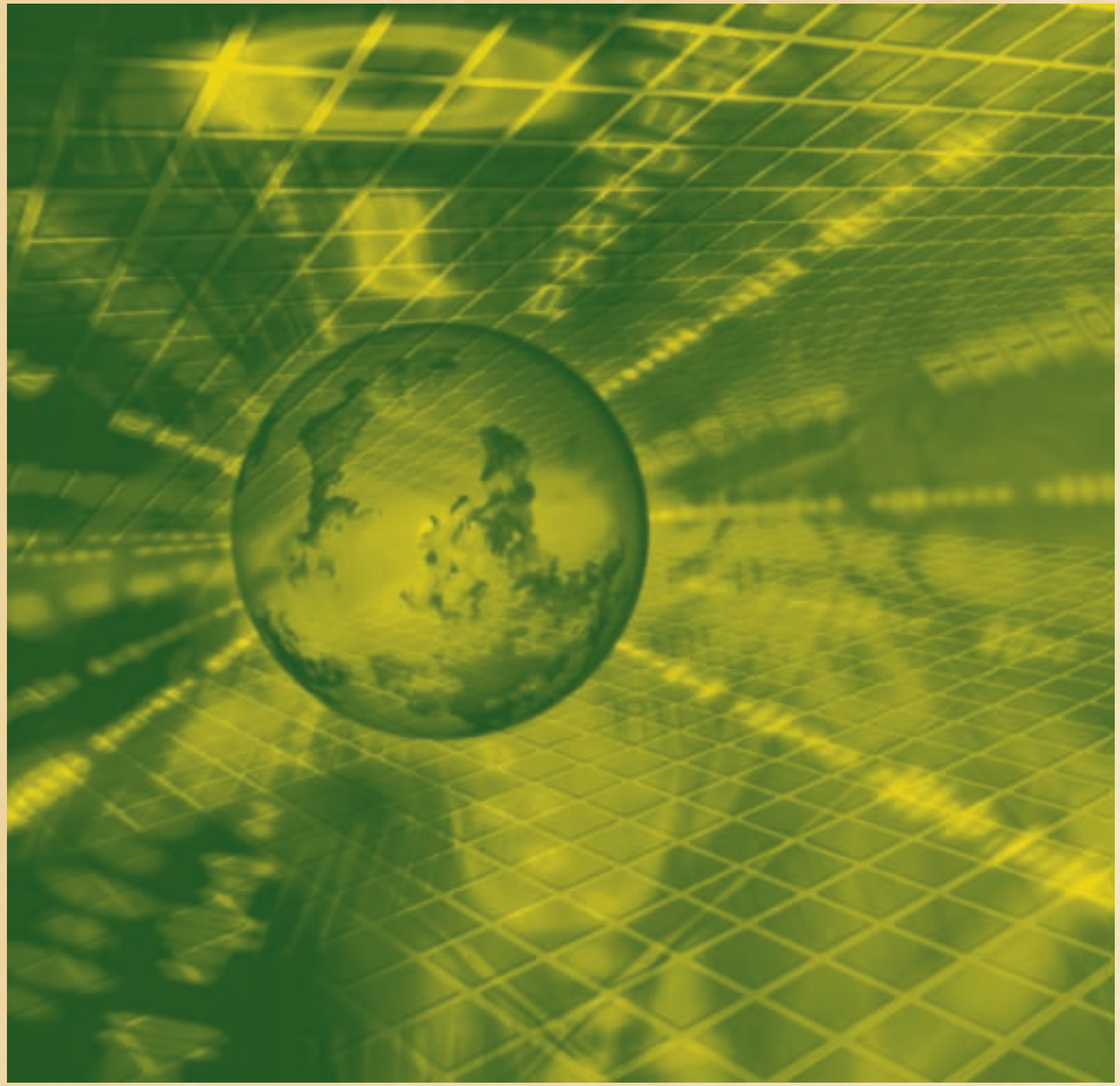
The booklet has three main aims:

First – to share reflections from the ESRC Science and Society project

Second – to invite the reader to look at the world of computerised biodiversity data frameworks from a social perspective on science and technology

Third – to raise some questions for future technology-human-nature relationships and for the construction of future biodiversity data frameworks.

Steve Rayner
Director of the ESRC Science and Society Programme



Contents

Biological recording: from sketch to digit

When we think today about early examples of a human fascination with recording the natural world, we could think about prehistoric rock art or Aristotle's meticulous observations of fish and other organisms. In many ways, the concerns of natural historians that emerged millennia ago are a far cry from those of today's producers and managers of biological data. Historical differences notwithstanding, biological recording at all times and in all places involves a close intertwining of human perception, nature and some form of technology (be it a pencil and notebook, a punch-card or a palm-top computer). The development of a variety of techniques and technologies for capturing, organising and displaying knowledge of nature is really quite extraordinary and is testament to human creativity in its encounter with the wonders of the natural world. In this context, it is important to consider how humans, representations of nature and technology each shape one another. Illustrations, species lists and digitised datasets, to name but some examples cannot be separated from the ways of knowing nature that they enable and the forms of human community they bring together.

In the biodiversity sciences, the use and expectations of computer-based information technologies have intensified dramatically over the last few years. Two obvious influencing forces driving these changes have been: first, the rapid development of computer memory, flexibility and interoperability; and second, the UK's signing, together with more than 150 nations, of the Convention of Biological Diversity in 1992. As part of this international agreement[1], signatories were required to devise conservation and sustainable development strategies, all of which demanded rigorous and comprehensive species mapping, listing and monitoring. In addition, the international framing of the CBD meant that value was given to biological records far beyond their local provenance and patterns of circulation. Naturalists, conservationists and biodiversity policy makers can now create, have access to, search and manipulate data about the natural environment for an unprecedented range of scales and purposes.

But this new global-local terrain of data production, curation and management is being forged on many fronts. Each new and developing biodiversity software also produces new and variable relationships between technologies, nature and society. We focus here on a dilemma discernable in three of the new data frameworks being developed in the UK.

1 See CBD Article 7: a) identify components for biological diversity important for its conservation and sustainable use; b) monitor, through sampling and other techniques, the components of biological diversity; c) maintain and organise, by any mechanism, data derived from identification and monitoring activities.

New data frameworks: a tricky dilemma

Whilst nationally/globally oriented biodiversity data frameworks enable powerful connections to policy institutions and responsibilities, their connection to local, dispersed communities of data producers is seen as relatively weak

On the other hand, whilst *localised* networks of data contributors enable strong internal social and data connections, these networks lack connections to policy-level global data bases

This dilemma is probably recognisable to many in the biological recording communities in the UK. But from an STS perspective, they suggest that we need to acknowledge that these computer based data frameworks are not **just** data holding technologies. They also enable, sustain, and are themselves sustained by:

- **social** relationships
- relationships between the social and the **natural**
- institutional **and decision-making responsibilities and powers**

There are 2 points to note here:

First, these often hidden aspects of computer based data frameworks have real **consequences for biodiversity protection;**

Second, because they have consequences, they raise questions of **responsibility** in relation to the development of these kinds of technologies.

The social and political lives of biodiversity software

The controversial idea that technical artefacts like computer softwares and data interfaces can enable and sustain social, natural and political relationships and dynamics has been captured by the more generalised idea that technological artefacts of all kinds **'have politics'**. This idea in fact has a long history in sociological, philosophical and political thought, from Plato's Republic to the writings of Marx and Engels and more recently STS scholars such as Langdon Winner[2].

One important aspect of this is the idea that the adoption of a technical system assumes the simultaneous adoption, or maintenance, of a particular set of social conditions as the operating environment of that system. In 1872, for example, Engels analysed the way that the steam-powered machinery of a large cotton factory implied the inevitable subordination of workers to the particular rhythms and requirements of 'the steam' - in other words, this technology brought along with it new social realities of shiftworking and 'factory time'.

An example from today's perspective might be the development of e-mail technologies. In similar ways, we can see how e-mail contains in-built assumptions about the way we work, the way we communicate, our expectancy of speed and efficiency in human interactions and communication, and so on.

Adrian Mackenzie, Lancaster University

2 Winner, Langdon (1986) *The Whale and the Reactor: A search for limits in an age of technology*, Chicago: University of Chicago Press.

Before looking at what these ideas might mean for biodiversity data frameworks, we sketch below three brief histories

Recorder

In the early 1980s, most biological data was recorded and stored on record cards and exchanged within specific communities of recorders with a shared interest in certain groups of organisms. The first example of software designed to cater specifically for naturalists' needs was created with backing from the NFBR (National Federation of Biological Recorders) by an invertebrate specialist with a flair for computer programming. In the early 1980s he designed the 'Invertebrate Site Register Database' using the software 'Advanced Revelation'. This later became Recorder 3.

During the 1990s, Recorder 3 underwent significant changes, driven in particular by raised user-expectations with the advent of 'Windows' for user-friendly, graphic design interface and the signing by the UK of the global Convention on Biological Diversity in Rio in 1992. This led to a perceivable shift away from the needs of local recording towards more complex software uses and a need to standardise and centralise diverse and large datasets. The Joint Nature Conservancy Council (JNCC) and the NBN Trust (see below), both bodies with responsibilities under the Rio Convention, provided institutional and financial support for the development of a new version, Recorder 2000.

Recorder 2000 was more complex than many naturalists needed and it lost popularity within their communities. The observation has often been made of Recorder that technological developments have proceeded apace without a socially sensitive and robust training or 'roll-out' framework. This suggests a neglect of social infrastructure in favour of technological fixes and is a tendency which has historically had knock-on effects for Recorder up-take. However, the organisations developing the software have always been aware of the possible risk of marginalizing locally-rooted naturalist communities. Considerable thought has recently gone into elaborating new versions with naturalists' needs in mind (Recorder 2002 and 2006). The most recent development has been the design of a simple recorder interface, 'Recorder Web', which 'talks to' recorders using browsers and offers tactile, graphic, user-oriented interfaces for data capture and reporting whilst still aiming to be a 'definitive deliverer of standards'.

MapMate®

The first version of Mapmate appeared in the mid 1980s as a result of the combined recording and computing interests of a moth enthusiast. The software became immediately popular within local natural history societies and demand for the package gradually spread throughout the UK. To deal with unprecedented uptake of Mapmate, a company, 'Teknica', was created. In 2006, 7000 licenses had been issued and the community of active users now numbers over 6000. The relative popularity of Mapmate amongst the amateur naturalist community has spread in recent years to organisations such as the Royal Society for the Protection of Birds and the Botanical Society of the British Isles and at least two of the country's Local Record Centres have opted to use Mapmate.

In the designer of Mapmate's words, this is a technology 'for the people' and is maintained by and further supports a peer-to-peer network of data contributors and users:

'Mapmate is like a community of users that have a distributed database that they all share in….What is creating it and behind it is a community of users and all the data is all over the place. That is how we always thought about it, rather than just like a programme: … a tool that allows you to do this with other people' (Mapmate designer).

Mapmate is perceived by statutory conservation institutions to be somewhat limited in potential, given that it is largely used within closed circuits of naturalists managing small data sets. Indeed, until recently, for a variety of technical and institutional reasons, it has been difficult to translate Mapmate datasets into a form compatible with policy needs. It is worth noting however, that most Mapmate users are more interested in maintaining a known, trustworthy community of naturalists and reliable and relevant datasets for their own recording purposes; many Mapmate users do not aspire to large scale biodiversity mapping, global reporting and preservation.

National Biodiversity Network

A perceived need to centralise the UK's biological records in a digitised accessible form for multiple global users inspired the design of the NBN in the early 1990s. Whilst recognised naturalist-based organisations had supported and promoted such a vision, it required government-level financial and institutional backing. This was forthcoming and was shaped and supported by the Convention on Biological Diversity. By 2000 the NBN had become a trust. One of the most prominent of its many objectives was: 'to enable central and local government policies that might affect our biodiversity to draw upon the widest possible sources of information'. In 2004 the main delivery mechanism of NBN - the 'NBN Gateway' - was launched on the World Wide Web. Its main job was to 'simplify the job of sharing and using information on biodiversity'. By early 2006 it hosted 19, 817027 million species records which had come from 150 datasets.

The establishment of this database and the associated system for data sharing has met with considerable challenges. Although the NBN Trust is still optimistic about the future relevance of the NBN, it is perceived by users that there is a very unclear target audience for the NBN gateway. The gateway presents itself as 'all things to all people'. NBN is seen by some as a centralising, bureaucratic system aimed at national government, far removed from the needs of volunteers collecting data. Some users feel that it has focussed on refining standards and tools at the expense of understanding its users and developing trust in the system. NBN has also unfortunately been seen by some organisations as a threat – undermining rather than supporting their own local data archiving and dissemination roles.

NBN designers (and some users) have recently begun to find ways of facilitating the entry of data from alternative data software technologies (e.g. Mapmate and Biobase). It is clear from this observation that NBN designers are, like Recorder designers, creatively considering ways of rebuilding trust between recording communities and available technologies.

We have suggested so far that the inherent properties and design of a technology often assume, support or require the adoption of particular kinds of natural and social realities and relationships.

In the next few pages we explore what this might mean, and why it matters.

Representations of the Natural in the 3 different data frameworks

Nature as 'biodiversity'

One assumption behind biodiversity auditing and data collation work from the mid 1990s onwards has been that nature protection requires the representation of nature as mobile, standardised and as digital data sets. For **Recorder, NBN** (and their developing variations), this translated into a need to create and mobilise standardised data for national and global policy access.

From an STS perspective, this kind of representation of the natural is difficult to disaggregate from the shaping power of data-holding technologies themselves. We also need to think about the idea that this particular version of the natural – as digitised, audited biodiversity - might **only selectively** encapsulate and represent nature. A typical STS question that is worth asking here is: what might be the future of those parts of nature which, for whatever reason, simply do not make it into a biodiversity database?

Nature as human-natural interaction

The vision behind the design of **Mapmate** was one of enhanced human interaction through the sharing of data (e.g. dot maps) about the natural world. As such, nature in **Mapmate** remains closely attached to naturalists both in terms of their relationships with their local patch, or passion for a particular group of organisms, and in terms of their relationships and control over the sharing of data with other naturalists. Without the pleasure of gaining and exchanging records, no nature would be made visible in Mapmate data frameworks. With this technology, nature transforms into a constellation of human-natural interactions. The obvious question here is whether the kind of institutions responsible for nature protection in the UK can practically use these more complex human-natural representations of nature in their day-to-day planning, policy making and methods of protection.

Nature aggregated

In policy terms, to talk of nature, is to quantify it and report on percentages of existing, increasing and declining populations of species and habitats. The **NBN data portal** represents nature in the form of aggregated lists and distribution maps. From an STS perspective, however, the aggregation of data raises questions about the danger of erasing the provenance and recording histories of datasets. This may be an especially important point given the increasing interoperability of data interfaces within the more powerful 'webby' frameworks of the NBN. How far is it possible to keep intact the recording histories of data whilst simultaneously increasing data interoperability?

Representations of the Social and Political in the 3 different data frameworks

User Communities

Recorder, Mapmate and the NBN were each developed with quite different assumptions about the interests, needs and concerns of their human data contributor and user communities. From an STS perspective, such assumptions have consequences for the ways in which technologies represent the natural world and ultimately have an impact upon biodiversity protection. These assumptions, therefore, also imply that the data frameworks 'have politics'; they make a difference to what is known, what is protected, how and by whom.

It is commonplace in biodiversity policy circles nowadays to think that 'we need to know what we've got before we know what we should be protecting'. The design of frameworks like **NBN and Recorder**, although stimulated in part by a desire to democratise UK biodiversity information, has been largely oriented towards biodiversity targets and policy visions such as those held within JNCC and defra, where a major preoccupation is to report at national and global levels on the status of UK biodiversity.

Mapmate evolved as a technology closely aligned to the interests of local communities of naturalists with a passion either for a particular patch or group of organisms. **Mapmate's** vision of society is quite different to a policy-oriented, pan-optical, globalising vision. Rather it is one in which the aim of knowledge making is to enrich the visions and lives of its users. **Mapmate** is described as being a technology 'for the people'. Accordingly the software is also largely co-designed by its users and is able to flexibly adapt to changing user needs.

Assumptions of user behaviours

In its 'mid-phase' (from **Recorder 3 to Recorder 2006**), **Recorder** was designed explicitly as a 'definitive deliverer of standards'. This implied not only the production of standard records but an implicit disciplining of any user who wanted to use the software to order their own data. Users were expected to be computer literate, time-rich, dedicated recorders who would readily adopt this data technology. Examples of unruly and techno-phobic human behaviour did not fit with the software's expectations. Such expectations had the effect of turning many people away from later versions of **Recorder.** Strenuous efforts have recently been made to re-align the software more closely to the practices and needs of naturalists and other non-institutional users through the creation of more flexible interfaces.

Mapmate, on the other hand, requires less discipline and technical know-how of its users, but works from the assumption that they work in close relationship with each other and actually create social groups through the sharing of biological records. In a sense, we could say that this takes precedent over the desire to comprehensively document the natural world.

Finally, through the **NBN,** the main assumption about users is that they accept their role as participants in a larger network. **NBN** is becoming an accessible data portal for global society at large, and works from the assumption that society wants and needs maps and lists as forms of information about nature. Without this relationship between society and information, the data framework runs the risk of becoming redundant.

Wired-in mechanisms of trust and exchange

The contribution and sharing of records, whether for global biodiversity protection or for local specialist community interests, depends upon establishing mechanisms for trusting data and data providers. From an STS perspective, it is interesting to note how trust relationships are embedded in the three different data frameworks. Through the technology of **Mapmate**, for example, data are exchanged through specific giving and receiving relationships and are therefore firmly personalised: the trustworthiness of the contribution is assessed through peer-to-peer judgement of data that is built into the hardware of the data framework. **Recorder** and **NBN** designers have also worked hard to incorporate naturalist 'practices of trust' into the hardware through mechanisms designed to make data ownership, quality and metadata transparent. Although these mechanisms are being continually improved, the extent to which such mechanisms **are** trusted by data contributors remains a question.

Assumptions of use for biodiversity governance

Whilst **Mapmate's** ability to adapt flexibly to user needs has been widely acknowledged, the peer-to-peer network of knowledge production may be overly self-reinforcing, limiting its links to other potential users and its scope and ability to scale up. The assumption is that users are active and committed at the local 'patch' level, but that they do not need to, or perhaps should be wary of, links to national or global frameworks. This has implications for the use of **Mapmate's** data for biodiversity protection.

NBN is now starting to be used to provide policy level statistics from within the JNCC for defra. This entails working with data at a very broad-brush level of detail.

There may be dangers ahead, however. Because of a history of a lack of trust in **NBN** networks and data, professional conservationists may still by-pass **NBN** as a data source, preferring instead to contact individuals whom they can trust to give them the latest up-to-date information on a given species or location. By-passed in this way, the **NBN** is in danger of losing authority. Its impressive capabilities to collate data and create 'reports' for biodiversity governance at national and international level, will become impotent without either the input of reliable data from those on the ground monitoring and recoding biodiversity, or the understanding of what that data means (its history and meta-data framings).

Paradoxically the two data frameworks - **NBN** and **Recorder** - that are specifically designed for improved governance of biodiversity seem vulnerable both to emptiness and to impotence. Both present problems. The lack of complete up-to-date data in the frameworks ('emptiness') may be a serious problem if the database is used as a bonafide representation of biodiversity in national or international contexts. On the other hand, the lack of use of the frameworks by practicing conservationists may lead to the frameworks' impotence as persuasive and legitimate tools of biodiversity governance. A possible and not unlikely scenario, and one of interest to STS, may develop whereby, *symbolically*, the data frameworks perform well at national and global levels of biodiversity governance. An underlying circuit of actors 'in-the-know' about species trends and decline may however view this symbolic show of data as shallow, alienating them from the official data gathering technologies and driving them further into circuits of knowledge production and exchange that have no purchase on biodiversity policy or decision making.

Questions for debate

A recurrent observation from Science and Technology Studies is that technological development often tends to focus too narrowly on the technology at stake and to lose sight of the social, cultural and political worlds that are being built into it. In the spirit of opening up a space for debate, and for giving life to the **implicit politics** of these technologies, we end this booklet, not with our own recommendations for their future design, but with some questions – meant for all designers, users and contributors.

There are two things that need consideration. First, to take seriously and to anticipate **the consequences** of the way that databases portray users, nature and governance. Second, to think about how the **responsibility** for designing data frameworks (including their implicit politics) can be shared. We offer the following questions as a way of provoking debate on these issues:

- If we accept that new and evolving data technologies are not just technical artefacts then who (besides technical experts) should be involved in, and responsible for, their development?
- How do the representations of the user built into data technologies enfranchise or alienate potential users of, and contributors to, biodiversity frameworks?
- Do the representations of nature contained within current biodiversity frameworks actually facilitate the protection of biodiversity? If not, what do such representations achieve?
- Do the visions of biodiversity policy and management (symbolic and otherwise) built into the designs of data frameworks have legitimacy within the biodiversity protection and recording communities? If not, what should be done?
- If greater emphasis is now being put into the interoperability of diverse biodiversity data frameworks, how can users and designers ensure that the specific histories and understandings of different data systems are not lost?
- What value lies in the diversity of ways in which nature can be documented and protected?