

The politics and practice of accessibility in systematics

Christine Hine
Department of Sociology
University of Surrey
Guildford
Surrey, GU2 7XH, UK
c.hine@surrey.ac.uk

Draft paper prepared for presentation at Past, Present and Future of Research in the Information Society, 13-15 November 2005, Tunis, Tunisia. An official side event preceding Phase II of the World Summit on the Information Society (WSIS).

Abstract

This paper focuses on a case study exploring moves within the biological discipline of systematics to make the tools and the products of research accessible online to the wider community. The case study will focus on the drivers for accessibility and the constraints on realising them for this community, drawing out concerns that making material accessible raises and formulations of the audience for accessibility. The presentation will highlight some elements of good practice, and some caveats about the extent to which open access should be pursued as a goal in itself. The paper calls for attention to the specificities of accessibility for each community, as it develops a situated response to broader calls for open access.

Introduction

The desire to create new ways of sharing scientific information pervades many areas of contemporary policy and practice. Information and communication technologies provide the impetus to envision new infrastructures, and offer the possibility of transcending existing constraints and inequalities. We still, however, have little detailed understanding of the way in which these visions will impact on practice, nor of the factors which give shape to initiatives and the constraints which may in turn derail them. In this paper I will be exploring the way in which one scientific community has experienced and explored the visions offered by digital technologies for open access. I will be arguing that high level calls for open access do indeed resonate with many areas of practice, and provide

fruitful ways for practitioners to understand their work and extend it in new directions. On the ground, however, open access to scientific information turns out to be highly specific to the situations in which it is developed. It has become inevitable in many fields of science that digital solutions to the sharing of information will be developed, thanks to the prevailing cultural current that places hopes and expectations in these technologies. Specific solutions, however, are often carefully crafted for their context and in the process may explore other technologies and forms of interaction. It appears to be important, on this basis, to continue to hold discussions on the appropriate technologies for particular information ecologies, and to remain agnostic about the forms in which information should be communicated.

The community on which I focus in this paper is systematics, the branch of biology concerned with the classification and naming of organisms and with exploring the relationships between them. I have recently been conducting a historico-ethnographic project in which I have been exploring the place of information and communication technologies in the contemporary landscape of systematics. This research is based upon interviews and collaborations with systematists in the UK, together with broader ranging internet-based research exploring the development of mailing lists and web sites in the discipline, and considering the extent to which the emerging internet landscape reproduces or transforms the traditional institutional landscape. To understand the contemporary situation it has proved vital to understand the historical legacy of the discipline, to develop awareness of the political profile and multiple accountabilities which practitioners and institutions experience, and to understand the material culture which underpins the activities of systematics.

In the next section of the paper I will introduce the discipline of systematics and explore the context in which open access initiatives have arisen in the discipline. The following two sections then describe in more depth two pertinent aspects of the experience of systematics which may prove telling for other disciplines exploring open access. The first aspect is the pressure from cultural expectations and from policy and funding bodies to inaugurate digital solutions to information sharing. The second of these sections then

focuses on the processes through which appropriate solutions are crafted for specific circumstances. Finally, the paper concludes with an elaboration of some general messages for open access initiatives which might be drawn from these specific experiences.

Systematics: open access in a complex political geography

Systematics institutions hold the collections of plant and animal specimens which form the foundation of taxonomic practice in biology. Any systematists describing species or developing new classifications needs to consult these collections. The major institutions hold vast collections: the Royal Botanic Gardens at Kew has 7-8 million plant specimens, while at the Natural History Museum in London there are 28 million insects and spiders alone. Historical specimens are a particularly valued part of these collections, both for the window they offer on biodiversity over time, and because some are type specimens which are the formal representatives, in nomenclatural terms, of each species. A systematist working on a taxonomic revision would almost inevitably require access to the type specimens of the species concerned. There are long established conventions for acceptance of visiting taxonomists to view collections, plus arrangements for loan of specimens. The specimen collection thus forms a scientific instrument for the exploration of biological diversity over time, and through practices of loan and visits has long been a distributed instrument, accessible by systematists across national boundaries. Nonetheless, geographic and economic constraints have always been a factor in shaping access, and it is clear that working at a major institution will be a significant advantage for doing taxonomic work.

The focus around access to material collections forms a significant challenge for contemporary systematics. Systematics has a complex political geography, shaped by the legacy of the discipline in a colonial past. Former colonies tend to find that much of their biological heritage, in the form of specimen collections, remains in major institutions based in the countries which formerly colonized them. Since they often, too, have limited resources to send systematists abroad, they may find themselves significantly hampered from conducting research on their own biological diversity. Capacity for taxonomic work

is therefore hard to build and sustain. Alongside this enduring geographic imbalance a new awareness of the practical and political significance of taxonomic information has recently emerged. At the Rio Earth Summit in 1992 the Convention on Biological Diversity was signed, committing its signatories to share access to taxonomic expertise (<http://www.biodiv.org>). The term “taxonomic impediment” was used to express the notion that not knowing about taxonomy, and not sharing what was known, was a significant obstacle to meaningful efforts to conserve biological diversity (Hoagland, 1996). The complex political geography of systematics was noted, and processes of data repatriation (latterly more often termed data sharing) were proposed to address the prevailing inequalities.

Data sharing has thus been instituted as a way of providing access to the knowledge resources for conservation of biological diversity and of building capacity in countries hitherto hampered by lack of access. Within this context it has been important that such digital resources as are created are seen to be openly accessible. It is thus becoming routine for systematics institutions to make their collections databases (albeit often provisional and incomplete) openly available via the Internet to an indeterminate audience. In contrast with many scientific disciplines which are experimenting with closed e-science communities, systematics has been positioned by a consciousness of the political dimensions of exclusion to opt for open data sharing practices.

In addition to the political geography of the discipline, other factors predispose systematics to explore digital solutions to information sharing. The existing culture of distributed work via loans and visits has relatively readily enabled systematists to see sense in placing catalogues online for colleagues elsewhere to consult. There is also a strong tradition of indexing in the culture, initially developed as a means of providing access to the legacy literature. This, again, has translated relatively easily into a digital environment. The upshot is a rich and diverse landscape of digital initiatives in systematics. Much of this is built around the existing institutions, but cross-institutional and transnational initiatives are becoming more common. Concerns faced by the institutions include the growing recognition of collections data as intellectual property,

and the difficulty of negotiating open access for indeterminate audiences with the need to preserve the interests and long term viability of the institutions. In terms of systematic practices, interest in digital resources is high, particularly as they relate to means of accessing existing literature and data, but concerns persist about the extent to which digital resources can reproduce the important qualities of material specimens and about the extent to which they might disturb established nomenclatural practices.

The political geography of the discipline, and its existing material culture, have played a large part in shaping the current internet landscape in systematics. In the next section I will focus in more detail on some of the expectations which have made this activity meaningful and driven it forwards. Particular sets of expectations around digital technologies have played an important part in shaping the discipline's response to the internet, playing out both on a level of individual practice and in high level policy forums where the discipline's activities are evaluated.

If you build it they will come: the inevitability of digital solutions

An examination of the contemporary situation in systematics reveals some fundamental expectations built in at policy and funding levels, and through diverse strands of accountability and cultural connection, that solutions will be digital. In this section I will briefly enumerate some sources of pressure on the discipline to develop digital solutions, and some of the consequences of these pressures for the resources which systematists have developed.

In the previous section I elaborated the complex political geography of systematics, and its manifestation in the Convention on Biological Diversity. Many current initiatives make sense within this context: or, to put it another way, digital resources have become a way for national governments (officially the signatories to the convention, and responsible for over-seeing responses) and for systematics institutions to understand their obligations under the convention. As signatories to the convention national governments are required to take account of their taxonomic expertise and resources, and to enumerate the ways in which they have developed appropriate sharing arrangements. Part of the

UK's activity in this regard was a review of the state of systematics in Britain instituted by the House of Lords Select Committee on Science and Technology published in 2002. Examining that review offers the opportunity to examine some of the expectations being placed upon digital technologies. From my perspective it provides a window on the cultural processes through which digital solutions come to be inevitable.

The report of the House of Lords Select Committee on Science and Technology (2002) makes clear that digital technologies are valued both for the practical purchase that they offer and for their symbolic qualities. The report summary notes:

We highlight the importance of digitizing the systematic biology collections, which will both increase accessibility of these data and help to update the archaic image of systematic biology. We also suggest that the systematic biology community should consider exploring new ways of presenting taxonomic information, in particular through increasing the amount of information available in digital form via the world-wide web, and should consider updating the system of naming previously undocumented species. (Select Committee on Science and Technology, 2002: 5)

The Committee echoed concerns from systematists that their discipline had a low status amongst other biologists and users of taxonomic information. Digital technologies were seen as a way of attracting more funding and more personnel into the discipline. In particular the report seized upon recent calls for systematics to make itself more attractive to its funders and its users by adopting radical digital solutions (Godfray, 2002a; ; Godfray, 2002b). Digital technologies were seen as a route to progress for the discipline, in a form highly resonant with the computerization movements that Kling and Iacono observed in other spheres of life (Kling and Iacono, 1988; ; Kling and Iacono, 1996; ; Iacono and Kling, 2001). As noted in these other instances of belief in the transformative capacities of computing technology, the self-evident benefits of the technology are accepted to such an extent that resistance would appear to be willful stubbornness.

I would argue that the prevailing expectations that digital resources will be developed, both as specifically manifested in the Convention on Biological Diversity and the national responses to it, and as a broader cultural current, mean that detailed research on who will use digital resources and how has been almost completely sidelined. There is little requirement, in this climate, to justify an innovation by reference to an identified group of users actually calling for it, to run carefully evaluated pilot projects demonstrating enhanced access and the meeting of needs, or to conduct usability studies and detailed requirements analysis. The requirement for a digital resource is taken to be self-evident. Users, and the practices to make use meaningful are expected to emerge once the resource is in place rather than preceding and informing resources development. This “if we build it they will come” expectation is built into the current political climate which systematics faces, such that institutions cannot afford not to develop and promote these resources. Users are key figures in the rhetoric of the discipline and of its funders, but the detail of usability remains a low priority.

Crafting solutions for situations: exploring appropriate technologies for open access

The analysis presented above seems somewhat bleak in terms of the ability of open access initiatives in systematists to deliver what their users want and need. It should be countered, however, by a recognition of the wide diversity of initiatives that are being developed, and the extent to which solutions appropriate for use are being crafted within the overall climate of belief in digital solutions. In this section I will explore a few examples of the technological and social innovations which characterize many initiatives in contemporary systematics. Top level expectations that the discipline will digitize provide the environment within which these initiatives make sense, but systematists are able creatively to work with that environment to develop solutions appropriate for use.

Developing digital technologies for systematics has involved an array of creative technological solutions: little is available “off the shelf” to meet the very specific needs arising from systematics’ material culture and existing communication ecology. The production of digital images of plant specimens requires some special attention in order to preserve the fragile and precious herbarium sheets. The solution employed at the Royal

Botanic Gardens at Kew is to turn the scanner upside down in a cradle, allowing the specimen to stay untouched and undisturbed beneath the technology which is, for once, not the precious part of the operation (Royal Botanic Gardens Kew, n.d.). Production of images of tiny insects at the Natural History Museum in London is a careful craft, involving use of software to stitch together a patchwork of images at different depths of field in order to achieve a crisp single image (Pitkin, 2002). Legacy systems are a particular issue: a precious index card system of Lepidoptera names, originally typed and then annotated in varying handwriting by curators over the decades required a highly specific and carefully crafted solution to produce a digital equivalent. Computer scientists and taxonomists worked together, a cheque scanner was carefully adapted, and systems set in place for parsing images, checking data and ensuring that original images were kept to be consulted alongside parsed data (Beccaloni, Scoble et al., 2003).

Technical adaptations are thus common in order to address the specificities of material objects. Creative responses are also needed to fit new technologies into existing communication ecologies. The example of web publishing serves to illustrate the complexity of responses. Whilst some systematists have been keen to seize the advantages of the web as a fast publishing medium with wide accessibility, there have been concerns that it is too ephemeral a medium for a discipline that amasses a cumulative archive. Whilst web publishing has to some extent been accepted in the official nomenclatural codes that regulate the discipline, safeguards against its perceived ephemerality have been put in place requiring lodging of identical copies in multiple libraries (International Commission on Zoological Nomenclature, 2000: article 8.6). The discipline of systematics has made its own judgments about the qualities of the medium and adapted them to suit its situation.

If technologies are crafted for specific situations, so too are the social arrangements to make particular projects work. Whilst in the grander visions of digital futures for systematics it is possible to assume that users will flock to solutions, specific projects often involves “users” and “providers” working together in roles which often blur over time. The Darwin Initiative in the UK funds projects under a rubric of promoting

biodiversity conservation and the sustainable use of resources

(<http://www.darwin.gov.uk>). Many systematics projects under this initiative have produced web site resources as a part of their work, offering an accessible public face to make products widely available. Arguably, however, as important as the final project is the work that goes on bringing people together across the provider/user divide, enabling them to cooperatively craft and appropriate solution. A project described within the Select Committee report as exemplary highlights the web resource as a product of collaborative effort:

The Natural History Museum, the Plymouth Marine Laboratory and the Kasetsart University in Bangkok and the Ministry of Fisheries, Thailand collaborated to explore the potential of the world-wide-web as a tool for exchange of taxonomic information between biodiversity researchers. The project has enabled researchers to share information about polychaete worms, in order to determine whether specimens found in various places are the same or different species. (Select Committee on Science and Technology, 2002: 15).

Darwin Initiative projects have no standing remit to develop digital resources. That many do testifies to the belief in the web as a means of providing a way to share data and link researchers together. The web is often seen as a way of offering long term access to resources, often in forms that provide for different modes of use than the printed paper product. The polychaete worm project described above developed an interactive web-based identification key, as a means to promote accurate identification by non-taxonomists. As much as it was a repository of data, however, the project found much of its value by linking researchers together both in workshops and training events and by providing a web site to allow information on new species to be shared rapidly.

Whilst digital resources abound in systematics, practitioners have a highly nuanced understanding of the capacities of the technology for them, of its political significance, and of the importance of alternatives. Working with others in a wide variety of circumstances offers a clear view of the importance of alternatives to digital resources for

specific circumstances of use. Whilst laptops and global positioning technologies have become routine companions on collecting expeditions, any systematist who has conducted field studies will have an awareness of the unreliability of technologies when confronted with non-ideal conditions. Whilst a digital interactive identification key has a seductive appeal, many will be aware of the circumstances when an A4 laminated sheet is a far more reliable and usable technology for an identification in the (wet and muddy) field. Whilst high quality digital images available on demand over the internet may seem to be the herbarium technology of the future, there is an awareness that there may be a considerable fondness for the older technologies. Instead of a digital image, with all that implies in display equipment, bandwidth and storage capacity, a high quality colour print that can be stored in a cupboard with related specimens and consulted time and again without additional cost can be very appealing.

In the face of the pressure to develop digital infrastructures for data sharing in systematics, a wide diversity of solutions have arisen. Many are carefully crafted in the face of specific circumstances, developing appropriate technologies and social relations for the task at hand. Often, digital technologies will be only part of an array of solutions, with the mobility of contributors, work across the provider/user divide and sharing of information and perspectives as important as developing an end product. Some solutions for “data sharing” will be digital, whilst others may be hybrid forms (a capacity to produce an alternative form of hard copy identification aid from a digital resource) and others may bypass the digital altogether. Digital solutions catch the headlines in contemporary systematics, but a diverse array of work on heterogeneous solutions goes on under the surface.

Conclusions: finding sense and opportunity in grand initiatives

Systematics represents a community within which it has come to be understood that data and resources must be shared. They are trying to find appropriate and feasible ways to make that happen, in the face of concerted pressure from policy makers and funders to deploy the digital technologies which are seen as efficient, effective and symbolically potent. Systematics is an already hard-pressed community which has seen its practitioners

ageing and its funding base dwindling. It is also a thoroughly reflexive community: there are well established traditions of talking on a disciplinary level about the status of taxonomic work, and about appropriate directions for development. Systematics has thus been able to find its own sense and its own opportunity in the grand initiatives. Current concerns about status and funding and a tradition of reflexivity position the discipline to look for solutions and to explore the potential that digital technologies offer. I have presented the specific crafted solutions and grand visions as polar opposites, but it is not really appropriate to oppose them in scale in this way. Grand visions are populated and kept alive through the crafted solutions that populate them and through the ability of practitioners to interpret their activities through those visions. Whilst users may be sidelined in the more of policy talk that makes digital solutions inevitable, they come back in as systematists try to make sense of their work on a daily basis.

There are some messages here for an open access movement that places belief in the ability of digital solutions to realise access to information. The experience of systematics suggests that too great a focus on the movement, and too much emphasis on the ability of particular technologies to realise a desired effect can be counter productive. A belief in the inevitability of digital solutions can sideline consideration of potential users and transform it into a simple belief that they will come. From this perspective open access looks like a low cost technical fix to issues of inequality, and of course nothing is that simple. However, we can expect that within the “open access movement” a wide diversity of initiatives may proliferate, and these will make sense to those most directly involved in a variety of ways which cross and blur the distinction between providers and users of information. There will be a need to remain open to non-digital solutions, and to respect the capacity of practitioners to craft their own appropriate technologies, even whilst we celebrate the ability of grand visions of open access to inspire, stimulate and offer a way of making sense of diverse experience.

References

- Beccaloni, G. W., M. J. Scoble, et al. (2003). Computerising unit-level data in natural history card archives. *ENHSIN: The European Natural History Specimen Information Network*. M. J. Scoble. London, Natural History Museum: 165 - 176.
- Godfray, C. (2002a). How might more systematics be funded? *Antenna* **26**(1): 11-17
http://www.cpb.bio.ic.ac.uk/staff/godfray/antenna_article.pdf.
- Godfray, H. C. J. (2002b). Challenges for taxonomy - The discipline will have to reinvent itself if it is to survive and flourish. *Nature* **417**(6884): 17-19.
- Hoagland, K. E. (1996). The Taxonomic Impediment and the Convention on Biodiversity. *Association of Systematics Collections Newsletter* **24**: 62,66-67
<http://www.sasb.org.au/TaxImp.html>.
- Iacono, S. and R. Kling (2001). Computerization movements: the rise of the Internet and distant forms of work. *Information Technology and Organizational Transformation: history, rhetoric and practice*. J. Yates and J. van Maanen. Thousand Oaks, CA, Sage: 93-135.
- International Commission on Zoological Nomenclature (2000). International Code of Zoological Nomenclature. c/o The Natural History Museum, London, International Trust for Zoological Nomenclature.
<http://www.iczn.org/iczn/index.jsp>.
- Kling, R. and S. Iacono (1988). The mobilization of support for computerization: the role of computerization movements. *Social Problems* **35**(3): 226 - 243
- Kling, R. and S. Iacono (1996). Computerization movements and tales of technological utopianism. *Computerization and Controversy: value conflicts and social choices*. R. Kling. San Diego, Academic Press: 85-105.
- Pitkin, B. (2002) Virtual collections. Darwin Centre Live presentation
http://127.0.0.1:800/Default/www.nhm.ac.uk/darwincentre/live/presentations/presentation_virtual.html.
- Royal Botanic Gardens Kew (n.d.) Herbarium digitisation
http://www.rbgekew.org.uk/data/herb_digitisation.html.
- Select Committee on Science and Technology (2002). What on Earth? The threat to the science underpinning conservation. London, House of Lords.