Alpha e-taxonomy: responses from the systematics community to the biodiversity crisis

S. J. Mayo[1], R. Allkin[1], W. Baker[1], V. Blagoderov[2], I. Brake[2], B. Clark[3], R. Govaerts[1], C. Godfray[3], A. Haigh[1], R. Hand[4], K. Harman[1], M. Jackson[1], N. Kilian[4], D. W. Kirkup[1], I. Kitching[2], S. Knapp[2], G. P. Lewis[1], P. Malcolm[1], E. von Raab-Straube[4], D. M. Roberts[2], M. Scoble[2], D. A. Simpson[1], C. Smith[2], V. Smith[2], S. Villalba[1], L. Walley[2] & P. Wilkin[1]

Summary. The crisis facing the conservation of biodiversity is reflected in a parallel crisis in alpha taxonomy. On one hand, there is an acute need from government and non-government organisations for large-scale and relatively stable species inventories on which to build major biodiversity information systems. On the other, molecular information will have an increasingly important impact on the evidential basis for delimiting species and is likely to result in greater scientific debate and controversy on their circumscription. This paper argues that alpha-taxonomy built on the Internet (alpha e-taxonomy) can provide a key component of the solution. Two main themes are considered: (1) the potential of e-taxonomic revisions for engaging both the specialist taxonomic community and a wider public in gathering taxonomic knowledge and deepening understanding of it, and (2) why alpha-species will continue to play an essential role in the conventional definition of species and what kinds of methodological development this implies for descriptive species taxonomy. The challenges and requirements for sustaining e-taxonomic revisions in the long-term are discussed, with particular reference to models being developed by five initiatives with botanical exemplar websites: CATE (Creating a Taxonomic E-Science), *Solanaceae* Source, GrassBase and EDIT (European Distributed Institute of Taxonomy) exemplar groups and scratchpads. These projects give a clear indication of the crucially important role of the national and regional taxonomic organisations and their networks in providing both leadership and a fruitful and beneficial human and technical environment for taxonomists, both amateur and professional, to contribute their expertise towards a collective global enterprise.

Key Words. alpha e-taxonomy, biodiversity, systematics.

Introduction

Papers by Scoble (2004), Godfray (2002), Godfray *et al.* (2007) and Scoble *et al.* (2007) have reviewed and discussed in detail the current broad picture of e-taxonomy. Thus, in this paper we focus specifically on the potential of e-taxonomy for delivering descriptive, morphology-based species taxonomies of higher plant groups.

Our first main theme relates to the general public as a broad user community and the role that taxonomy fulfils in enhancing general understanding of biodiversity and the environment. We start from the premise that the most conspicuous plant and animal groups have the greatest impact on people's awareness and therefore the highest priority for e-taxonomy. A second premise is that people have a strong desire for a greater understanding of their world, so by making it easier to acquire a deeper first-hand knowledge of species, greater involvement in biodiversity conservation is likely to be stimulated. Achieving this goal requires that e-taxonomy will be not just a simple media change from paper to internet pages, but will involve a more significant transformation into a multi-directional interaction in which the roles of

Accepted for publication January 2008.
[1] Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AB, UK. E-mail: S.Mayo@rbgkew.org.uk
[2] The Natural History Museum, Cromwell Road, London, SW7 5BD, UK.
[3] Department of Zoology, University of Oxford, South Park Road, Oxford, OX1 3PS, UK.
[4] Botanischer Garten und Botanisches Museum Berlin—Dahlem, Freie Universität Berlin, Königin-Luise-Str. 6-8, 14191, Berlin, Germany.

producer and user become less polarised. With appropriate tools and information, a much wider public can become engaged directly in the taxonomic activities of identifying, recording and perhaps even describing species, especially in those groups with which people can interact most directly, i.e. the macrofloras and macrofaunas; descriptive taxonomy can move beyond herbaria and museums to include anyone with a deep interest in natural history.

Our second major theme is that in these same macro-organismal groups, although the revision and exploration of species limits will increasingly involve both morphological and molecular data, both data types will remain of fundamental importance to species taxonomy. We argue that morphological species concepts have a distinctive role as the major reification (literally: the transformation of an abstract idea into a material object) device within the biodiversity knowledge system that serves the general public and the broad scientific community, irrespective of the search for finer grain taxonomies required for more specialised purposes. In this context, a new task for taxonomists is to bring morphological and molecular species taxonomies into optimal alignment wherever possible. This in turn prompts the need to develop data-gathering and analytical tools that could offer richer and more accessible methods for comparing taxonomic patterns arising from the different data fields.

The full development of e-taxonomy holds out enticing prospects for an increased impact of taxonomy in the wider world, but it will need both individuals and institutions to contemplate significant changes. The experience so far gained from five international e-taxonomy projects is used as a basis for mapping out some of these challenges. We have emphasised plant examples in the present paper, using the term "e-taxonomy" to mean the creation of a web-based alpha-taxonomy that provides the full information content of a taxonomic revision in the flowering plants. Although much of what is discussed is being applied to the descriptive taxonomies of many other organisms, angiosperm taxonomy currently provides an especially favourable context for developing e-taxonomies because of the existence of a variety of common resources and standards, e.g. International Plant Names Index http://www.ipni.org; Angiosperm Phylogeny Website http://www.mobot.org/MOBOT/research/APweb/; Authors of Plant Names http://www.kew.org/data/authors.html/; Taxonomic Literature Ed. 2 http://tl2.idcpublishers.info. In addition, the Global Strategy for Plant Conservation (http://www.plantlife.org.uk/international/plantlife-policies-strategies-gspc.html), by prioritising the creation of a world checklist of plant species, has focussed international collaboration and resulted in rapid progress towards this goal, thus providing the backbone on which to construct rich e-revisions.

Definitions

The scientific area dealt with in this paper is species delimitation and description, i.e. the main subject matter of most taxonomic monographs, revisions, and floristic treatments. We use the term alpha-taxonomy (Davis & Heywood 1963: 4) and avoid others that have been used in recent literature such as "Linnaean taxonomy" (Godfray 2007), "Linnaean system" (Savolainen *et al.* 2005), "conventional morphological taxonomy" (Smith *et al.* 2005), "traditional taxonomy" (De Ley *et al.* 2005) and "classical taxonomy" (Davis & Heywood 1963), partly because we wish to be quite specific about the layer of taxonomy being discussed here, and partly because part of the argument we are making is that alpha-taxonomy is an essential part of modern taxonomic practice, and underpins all biodiversity research. "Alpha-taxonomy" reflects the pioneering role of descriptive species taxonomy in making the preliminary maps of biodiversity and providing the species names that are universally used. The business of alpha-taxonomy is to confront "the bewildering blizzard of wild biodiversity at the edge, middle and focus of society" (Janzen 2004) and distil from it, by delimitation and description, the set of units that can be used by the general and scientific public for other purposes. The prefix "alpha" seems entirely appropriate, with its sense of exploration and discovery and the implication that its products are rough diamonds that form the point of departure for further research from ecological and evolutionary viewpoints.

The Requirement for Alpha-Taxonomy

Is Current Alpha-Taxonomy Still Relevant?

In contemplating a wholesale migration of conventional species taxonomy to the Internet, it is inevitable that many people with an interest in taxonomy, whether as users, managers or practitioners, will ask themselves whether it is really worth the effort. This is not a trivial question. Godfray (2007) has recently challenged the view that morphology-based taxonomy will always be needed for many of its traditional key applications involving the identification of disease vectors, pest organisms and other economically significant groups, maintaining that for organisms in which taxonomic information is crucially important for social, medical or economic reasons, molecular approaches will be and already are being used as routine. Molecular methods are making a vast new field of taxonomic information available at species and infraspecific levels and along with it new or improved analytical tools to create taxonomies from these data. There can be no serious doubt that these techniques represent a key advance in species taxonomy, especially for identification, but also beyond.

Molecular data finds perhaps its most exciting application in the poorly known areas of the biosphere which alpha-taxonomy has hitherto barely touched. Consider soil meio- and micro-organisms (Blaxter 2004; Blaxter *et al.* 2005). The true extent of the biodiversity in these ecosystems has barely begun to be assessed by taxonomists, but molecular tools are beginning to revolutionise this (e.g. Yergeau *et al.* 2005). In this unexplored biodiversity territory, molecular alpha-taxonomy is creating an almost entirely new domain of knowledge. But in addition to pushing back these new frontiers, molecular methods are also seen to offer a new perspective for the general delimitation of alpha-species in higher plants and animals (e.g. Savolainen *et al.* 2005; Pons *et al.* 2006).

Another much-discussed trend is the waning number of taxonomists who have the specialist knowledge (usually based on morphology, ecology and anatomy) required to name and classify organisms (Gaston & May 1992). One factor that discourages young taxonomists from pursuing a career in descriptive taxonomy is the poor reward they receive from journal impact-factor measures that ignore a large sector of publication outputs which have always had higher value within the discipline, such as major flora and monograph series and books. Basic questions thus arise for practitioners and their institutions: where does current alpha-taxonomy fit into this rapidly changing landscape, why does it matter, and how should it be reformulating itself?

The advent of e-taxonomy seems to crystallise these issues. The Internet offers opportunities for species taxonomy to become more attractive, more accessible and achieve a greater impact, but it also brings challenges; to change how taxonomists are evaluated professionally, collaborate and receive rewards, to alter drastically the traditional styles of presenting the information. Ultimately e-taxonomy offers an opportunity to standardise the taxonomic process itself so that it becomes accessible as a scientific activity to a much wider community around the world. It is a moment for progressive and reforming attitudes that are timely in the quickening global crisis of biodiversity conservation. As more and more people become concerned for the future of biodiversity in the face of global change, there is an increasing demand for knowledge and an anxiety to contribute — conservation is a collective enterprise. E-taxonomy offers a realistic way to empower a larger global community, not only to identify the organisms that interest them, but to contribute directly to taxonomic science as a mass activity and hence to a major advance in biodiversity knowledge. The future role of professional taxonomists can be seen in this context as the coordinators and curators of a vastly greater information landscape. It is thus conceivable that e-taxonomy could become the most significant advance in descriptive taxonomy since Linnaeus solved the first bioinformatics crisis (Godfray 2007).

Stakeholder Demand

The demand for taxonomy coming from the global community (Samper 2004) is not so much for reformulation of the species concepts themselves, which for the most part are workable as they currently exist, even if in constant need of repair and refinement in the light of new information. Instead, the primary need, increasingly urgent because of biodiversity loss, is for identification tools and knowledge of the geographical distribution of the known species of the macroflora and macrofauna of the world. The relationships of species are crucial for the taxonomist and many other scientific stakeholders, but less so for the general user community, who by and large are content to take the species as they come and accept the simple and pragmatic ranking system within families and genera. What seems clear is that the morphology-based alpha-taxonomy of the more conspicuous organisms of global ecosystems can provide an adequate framework for setting major conservation priorities (Mace 2004). In particular, morphological taxonomy of larger organisms (as distinct from micro-organisms) is also a viable area of activity for the wider community who are looking for a way to engage constructively with mapping, monitoring and other aspects of the environmental sciences, thereby acquiring a deeper knowledge and understanding of their own biodiversity.

It is illuminating in this regard to consider some current broader initiatives in which morphological alpha-taxonomy plays the key "backbone" role. A recent high-profile conservation programme (Millenium Seed Bank, http://www.kew.org/msbp/) contracted taxonomists to produce field identification guides (unpublished) for specified geographical regions in partner countries to target endangered, endemic or potentially useful plant species for *ex situ* conservation by seed banking. Using these specially designed tools, it was possible to prioritise, multiply and upscale taxonomic field surveys for the purpose of conserving vital species germplasm for the future, and it empowered the developing world project partners who carried out the surveys. The data were based on existing alpha-taxonomic resources in the botanical literature and herbaria.

Another example of the way demand for species taxonomy is shaping up can be seen in worldwide initiatives such as the Catalogue of Life (http://www.catalogueoflife.org/), Global Biodiversity Information Facility (http://www.gbif.org), the Encyclopedia of Life (http://www.eol.org/home.html/) and national projects such as Mexico's CONABIO (http://www.conabio.gob.mx/), Costa Rica's INBIO (http://www.inbio.ac.cr/), the Canadian University Biodiversity

Consortium (http://www.canadianbiodiversity.org/) and the new *Atlas of Living Australia* being created by the Australian government. These show that there is a general demand for broad-based information resources to tackle biodiversity conservation and their importance is demonstrated by significant government and private funding for such projects. Since these large-scale biodiversity information architectures are being erected on the foundation of current alpha-species taxonomy, it is clear that a total rebuild of species taxonomy and nomenclature from the ground up, using quite different approaches, is neither necessary nor desirable. However, what certainly is needed is a fast, effective interface with the current species taxonomies of organisms and this creates an immediate and urgent demand for e-taxonomy. E-taxonomy is the medium needed to connect users with the taxonomic community, its knowledge base and the coming generation of rich biodiversity information portals that are being set up around the world.

In botany, the production of Floras and associated outputs such as floristic atlases, have long been a major priority for both institutions and individuals. These projects are vital for providing taxonomic knowledge to the public and may include major participation by non-professional botanists. Integrating floristic treatments and global e-taxonomy revisions at the species level is still in its infancy, but internet technology has the potential to synergise the effort invested in monographic and regional approaches so that both feed into a single global output.

Taxonomy as Knowledge

Taxonomy, irrespective of its more utilitarian applications, has fundamental importance as a knowledge base, as a language for interpreting the biosphere. In this sense it has a position in the public mind somewhat similar to that of cosmology, for which there is a huge public appetite. Taxonomy provides important keys to understanding the biological world of which we are part. And everyone is a taxonomist — our brains have clearly evolved this ability as a crucial faculty for survival (Berlin 1973). Intuitively we classify phenomena and reify the classes into objects that we use to explore relationships and generate systems of knowledge. Humans have an eager appetite for classifying information — we detest chaos. To tap more effectively into this basic human drive and release its latent energy, e-taxonomy needs a broader goal, not just to construct a taxonomic system and deliver it for general consumption, but something richer and more ambitious: to equip people with the means to understand and investigate biodiversity for themselves. This is a powerful raison d'être for alpha e-taxonomy.

The Mutual Dependency of Morphological and Molecular Alpha-Taxonomy

Although e-taxonomy is crucial for engaging public interest in biodiversity information, the arguments of the previous section alone are not completely convincing as a long-term scientific justification for morphology-based alpha-taxonomy. Molecular genetic knowledge of organisms is building up at a very rapid pace, and although molecular studies are currently dependent on alpha-taxonomy to provide a sampling space (DeSalle 2006), it seems quite possible that this may be only the first step to more radical change. Some taxonomists (e.g. Tautz *et al.* 2003; Pons *et al.* 2006) foresee that after this initial period, a parallel molecular alpha-taxonomy might arise; that is, a system of alpha-taxa discovered, circumscribed and described from molecular data alone. This is already the case with micro-organisms and in a more minor way with some difficult species complexes of huge importance to people (e.g. mosquitoes) or as model systems to science (e.g. *Helianthus* and *Senecio*) where molecular analysis far extends the work of alpha-taxonomy. Although this vision of the future has been much contested for higher animal and plant groups (e.g. Scotland *et al.* 2003), it has had, nevertheless, an impact on the general debate regarding taxonomy's future and could have further impacts on strategic planning for funding agencies and the major taxonomic institutions themselves. It is necessary to ask and to find answers to key questions. Where will morphological alpha-taxonomy find its niche in this new scientific landscape? What is its future and its long-term importance?

A fully integrated taxonomic system, at least for macro-organisms, needs to map morphological and genetic taxa onto one another. Because phylogenetic analysis based on DNA sequence data is now the dominant method for resolving the limits of higher level taxa, recognising a higher plant family from its external morphology is no longer seen as an essential criterion for its delimitation. Nevertheless, taxonomists make strenuous efforts to reify the newly refined molecular-phylogenetic concepts of major groups using phenotypic information, as can be seen in recent major general treatments of Angiosperms (Stevens 2007; Judd *et al.* 2002; Soltis *et al.* 2005). Intuitively we all understand why this is highly desirable. Science must communicate its concepts, not simply create them. If genetic data and genetic taxa are to become the machine language of comparative biodiversity science, as seems quite possible (e.g. Markmann & Tautz 2005), then morphology-based taxonomy, at least for macro-organisms, will be the high-level language which interfaces with the world and provides the conceptual vocabulary with which most people think and communicate about biodiversity. Again, it has to be emphasised that where phenotypes fail to provide this, as in many groups of micro-organisms, genetic-molecular taxonomy will stand alone. But where morphology and other phenotypic characters offer a rich and meaningful footprint of evolutionary and ecological processes, it is foolhardy not to deploy them.

Imagine a scenario in which molecular samples are taken from natural populations, without normal voucher specimens, sequences are obtained and made accessible over the Internet with their metadata, and the purified total DNA deposited in a recognised archive. By an appropriate clustering procedure using these and other data, it is established that a new cluster based on these data meets, let us say, a set of criteria that allow its recognition as a species. It is then formally named and an appropriate molecular diagnosis is published along with the name. Such an alpha-taxon would be a valid unit of biodiversity and might later be found to have interesting or ecologically significant properties. But to make it straightforward to relate it back to field populations some other kinds of data would be required. If the plant is to be easily located again, even with the originally collected geographical information, the metadata would need to include information on the phenotype. And this information should take account of phenotypic variation to be precise enough to allow recognition of individuals in all the locations where it occurs, or indeed anywhere else where it might occur. If a handheld DNA "barcorder" were used to identify the plant, it would still be necessary in the field to locate and choose the right plant to sample, unless one had the time and resources to re-sequence everything in sight. Phenotypic information is therefore needed, in a form sufficiently rich to express the taxon's variation and how to diagnose it. Even if the molecular-genetic alpha-taxon eventually assumes primacy as the fundamental reference, it needs to be reified by formulating a concept based on the physical macro-structure of the organism, whenever this is possible.

We can thus discern three distinct scenarios in the interaction between species concepts based on genetic and phenotypic data domains. In the first, molecular concepts are dependent on the current morphological alpha-species that are being used as a sampling space (Chase *et al.* 2005). Second, parallel molecular alpha-species concepts develop, in which the molecular information provides the basis for finding species-level groups independently (Tautz *et al.* 2003; Blaxter *et al.* 2005; Pons *et al.* 2006). In the third scenario, the two classes of species concept are optimised (Janzen *et al.* 2005), since it is certain that DNA alpha-species will not always be congruent with morphological ones; e.g. Pillon *et al.* (2006). In this paper we argue that the third phase, analogous to Hennig's "reciprocal illumination" (Hennig 1966) is essential, that it is a two-way process and that it is here where an important aspect of the scientific development of alpha-taxonomy probably lies (see De Ley *et al.* 2005). The optimisation process requires advances in methodology so that alpha-taxonomy becomes a discipline in which analysis, reanalysis and model-building become as routine as they are for genetic, ecological and phylogenetic science (Godfray & Knapp 2004).

Alpha-Taxonomy and Species Concepts

One concern about current alpha-taxonomy, important for both the taxonomic and conservation communities, is the scientific significance of the species concepts employed. Most of the approximately 380,000 Angiosperm species currently recognised (Paton *et al.* in press) are based only on a combination of morphological characters and geographic distribution. Are they appropriate for use in modern evolutionary biology or should species be defined by phylogenetic methods (e.g. Davis & Nixon 1992)? Can we obtain a universal definition of the species category (de Queiroz 1999) and how would this help delimit species concepts in practice? Many more such questions are under discussion, most turning on the nature of species concepts and there is much genuine doubt about these issues (Zander 2007).

One relevant theme that needs further exploration is the validity of taking an empirical approach to species delimitation in alpha-taxonomy. As many authors have discussed (e.g. Davis & Heywood 1963), building the concept of a species of macro-organisms is itself a process, not a single act. It is easy to forget, in the focus on inventory and checklists required for practical needs, that species concepts vary enormously from rich and sophisticated theories at one end (e.g. Lexer *et al.* 2005 on *Helianthus*) to the most threadbare hypotheses of species described from single cultivated collections of uncertain origin at the other. Alpha-taxonomy is surely a process of building initial, shallow empirical hypotheses from an easily observed set of characteristics (including molecular characters). Therefore any alpha-species concept has, at least to start with, a substantially arbitrary nature, being highly dependent on the data employed, the characters selected, the individuals sampled and the thought processes, often unspecified, used to derive it. A raw species hypothesis evolves into a more substantial theory by bringing into play other evidence dimensions such as geography, ecology, reproductive biology, population biology and genetics (Hennig 1966; Lipscomb *et al.* 2003). None of this, however, alters the empirical nature and the inescapable requirement for an initial triage of the "blizzard of biodiversity". Three lines of development become apparent when alpha-species are considered in this way, purely on their own merits. First, it might be worthwhile to develop a range of more explicit descriptions of how alpha-species hypotheses can progress to fully fledged theories of individual species — there might be various solutions: for example, species theories for conservation (Mace 2004) might emphasise different evidence dimensions from palaeontological ones (Forey *et al.* 2004). Second, how could the triage techniques themselves be improved? If genetic species concepts are to be reified by phenotypic ones, new techniques for optimising the data patterns are needed that are not predicated on the notion that only one data set provides the "right" answer. Buried in this issue are tensions concerning the ontology of species concepts on the one hand and their detection by data on the other, which are beyond the scope of this article but are significant for both theory and practice (Davis & Nixon 1992; de Queiroz 1999; Boyd 1999). Thirdly, testing emergent alpha-species by other evidence dimensions such as ecology and genetics should in turn create a feedback response that re-adjusts the alpha definitions of a species concept. This feedback loop is usually incomplete in taxonomic practice, perhaps because there are as yet no mechanisms for storing, viewing and synthesising the much more complex data of "omega-species".

Making E-Taxonomy Work for Users

New Opportunities

Bringing alpha-taxonomy to a broader community of users and contributors, as outlined above, will have a major impact on how the interface is designed. What changes are needed in the methodology and organisation of alpha-taxonomy to make it work for a wider use group? Some areas are considered in what follows. But as a guide, let us consider that this public will consist of anybody with a serious enough interest in biology and natural history that they are willing to use a website and its functionality to help them interact with the organisms on which they are focussed.

Consensus Taxonomies

The central theme of the view of e-taxonomy presented in this paper is the notion of a consensus taxonomy for a plant family or genus. As previously mentioned, our focus is on the delimitation of species rather than their scheme of relationships (although this is also a legitimate objective of a consensus taxonomy). Consensus taxonomy, in this restricted sense, means that at least a critical mass of the taxonomic community working on a family (for example) agrees to collaborate in the production of a single provisional statement of the current species delimitations that can be used by less specialist stakeholders. This taxonomy will be versioned to a specific date and new versions will be produced in a regular cycle, analogous to journal publication. Species concepts in different editions will be electronically linked to make it straightforward to understand what workers mean by a particular name at a particular time, and what it means now. The content of the consensus will evolve by means of contributions and proposals made by anyone, but which pass through a peer-review process built into the software system. The review process will be carried out by the taxonomic community, which will aggregate itself by means of a transparent and equitable social organisation that provides a platform for discussion and debate aimed at facilitating the necessary consensus decisions. In our view, all contributions that do not become part of the consensus should remain on the e-taxonomy website accessible for the more specialist user. Should the facts or consensus opinions change they might later become part of the consensus.

Producing a consensus taxonomy is a necessary response to a public demand already widely recognised in the taxonomic community. The development of major "currently accepted species" checklists over the past ten years (e.g. the World Checklist for Selected Plant Families, http://www.kew.org/wcsp/) testifies to this concrete need as expressed in the Global Strategy for Plant Conservation (http://www.plantlife.org.uk/international/plantlife-policies-strategies-gspc.html). E-taxonomy, in the sense adopted here, is a further extension of the same trend. To a greater extent than checklists, it requires negotiation, agreement and adoption of conventions by taxonomists, for example "conventionalising" species taxonomy between regular updates. Broadly speaking, this is not so different from the traditional acceptance of a major monograph as "standard". Consensus taxonomies generated by e-taxonomy can improve significantly on this model by making the "standard" more genuinely a reflection of collective judgement and by allowing better visibility for alternatives wherever these are unavoidable. It is widely understood that alternative taxonomies should be recognised, and that a proactive spirit is required in implementing this in web portals. This is not as radical as it may seem because taxonomic revisions and monographs typically consider alternative views (Scoble *et al.* 2007).

The discussion of stability of alpha-species delimitations ("Taxonomic inflation") in the context of conservation biology (e.g. Isaac *et al.* 2004; Harris & Froufe 2005; Knapp *et al.* 2005; Mallet *et al.* 2005; Pillon & Chase 2006) is pertinent to the notion of conventionalising alpha-species and hence to consensus taxonomy. Developments in measuring and monitoring biodiversity and its change through time and space need some level of stability in alpha-species taxonomies, but the introduction of new data fields that can also be used to erect species concepts will certainly perturb existing species delimitations. Reconciling these contrary trends needs conventions and solutions reached by agreement, since the goal of an absolute criterion for operationally delimiting taxonomic species across all groups of organisms is unattainable. A consensus taxonomy established for a specified time period constitutes a convention that also embodies a mechanism for controlled change to which a range of stakeholders can contribute. Versioning consensus taxonomies at specific times points the way to developing a correction technology whereby existing conservation data can be automatically remapped to new versions of alpha taxonomies. The increasingly collective approach envisaged for e-taxonomy could provide a much more responsive context for such advances than the present more fragmented working environment.

The problem with the idea of formalising consensus taxonomies is that it could be seen as imposing a limit on the independent thought and expression essential for any healthy scientific activity. If contributing to a consensus taxonomy comes to be seen as crucial for a taxonomist in that field (see below), then publishing one's own opinion about issues will, nevertheless, always remain equally important. This is especially true for scientists in the earlier phases of their careers. It seems necessary then, for there to be a parallel forum to the consensus taxonomy, a debating chamber. Although conventional journal publication might seem the obvious choice for this, the disadvantages are fragmentation and the problem of impact factor. In a more flexible reward environment (see below), an e-taxonomy debating chamber could reflect diverse opinions, provide a dynamic and less formal form of electronic publication and achieve significant impact rewards. At the earlier stages of launching a consensus taxonomy, a parallel debating e-space would also provide a mechanism for gathering taxonomists together with the explicit purpose of working out consensus positions, e.g. for descriptive terminology. This e-space could also play a key role in generating alternatives to the consensus taxonomy that would eventually be posted (equivalent to formal publication) on the main consensus site, an option that forms part of the CATE model's design (see below).

The fact is that taxonomists are not so much the quarrelsome atavists beloved of scientific journalists, but mostly enthusiasts who see themselves as part of a collective enterprise, namely the construction of a great edifice that maps and describes the world's biodiversity. Most taxonomists search almost by instinct for a piece of the pyramid where there is no existing worker and begin to prepare their bricks to add to the whole. Concrete evidence of the capacity of the taxonomic community to work together for the common good is to be found very near at hand, for example, in the Codes of Nomenclature (ICZN 1999; McNeill *et al.* 2006). Here we see a labour of generations stretching back over 143 years that has been sustained by dedication to a general cause, both by many individuals and many institutions. These achievements are highly apposite to the present prospects for e-taxonomy in regard to collective action at the personal and institutional levels.

Standardising Terminology

Building the first upload of descriptive taxonomy to a website highlights an opportunity to make two major gains in regard to descriptive terminology, neither immediate but both realistic. Firstly there is the prospect of standardising the terminology by agreement within the taxonomic community of a major plant group, and second that of developing a non-verbal analogue set of terminological definitions using diagrams and illustrations — in effect replacing very general handbooks such as that of Stearn (1992) with a glossary specific to the group and developed by that community. We envisage that achieving this goal will require a process of negotiation and debate over some time, perhaps several years, and will require meetings to resolve disputed issues.

Standardising Descriptions

Once a standard, illustrated terminology has been made available over the Internet, the way is open to standardise the composition of descriptions, with the goal of replacing the first generation of species descriptions (salvaged from the scattered current literature) by new standardised versions. The major objective of this phase would be to develop an agreed atomised descriptive structure down to character-state level that is as compliant as possible with the predominant standard within taxonomy as a whole (see TDWG, http://www.tdwg.org). Again, this activity would be driven and moderated by the relevant taxonomic community.

Taxon Description Kits

With the terminology standardised and illustrated and a database structure in place for storing, searching and analysing descriptive data, it would then be possible for database compliant descriptions to be prepared by anyone with access to the plants, by providing a kit interactively from the site. This is an essential step towards providing tools and guidance to a much wider community in order to involve interested people in the process of taxonomic description. The long-term aim would be to upscale and accelerate taxonomic information gathering, including description of field specimens (see below). This could be a fruitful means of involving students in descriptive taxonomy, especially in biodiversity-rich or poorly surveyed parts of the world, providing the incentive of recognition through web-publication.

Automatic Key Updates

Interactive identification keys to all species and genera are important tools for web taxonomy. One clear requirement that has emerged from experience gained in the projects described below is that addition of new character information to species descriptions needs to trigger, automatically if possible, changes in the matrix that drives the key. Currently available key-writing software still needs further development to make this kind of updating possible.

Global Collaboration

Standardised tools and protocols for describing the species of a plant family provide a firmer basis for global collaboration and strengthening ties between expert amateur taxonomists and professionals, in effect empowering the wider community who are genuinely interested in descriptive taxonomy. This is a necessary but not a sufficient basis for a thriving collective effort. In addition to the tools, there must also be rewards, equality of opportunity and transparent evaluation of the merit of the contributions of participants.

Morphological Repositories: The Raw Data of Descriptive Taxonomy

Following this line of argument further, field botanists could not just make herbarium specimens for their own institutional collections but could also collect standardised descriptive data and deposit it in morphological repositories of observational data associated with the plant family portal. Instead of field collection being only a preliminary to publication, the provision of good quality observational data sets from field collections could become an activity that attracts a credit comparable to publication. Such an activity would be similar to depositing molecular data in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/). As anyone knows who has authored a standard taxonomic treatment, a description is only as good as the minority of specimens that are well-documented and well-prepared. The increasing practicability of delivering very high quality digital images through websites will also make detailed visual documentation of morphology increasingly important.

Morphological repositories will also be important in initiating a move towards the universal availability of the raw data of descriptive taxonomy, currently almost always unavailable because of the traditional paradigm of the species description as the base information. The Internet makes it practicable to now circumvent what is in effect a major barrier to testing alpha-species concepts by re-analysis of and accumulation to existing datasets, a procedure that is one of the key innovations brought into systematics by molecular taxonomists. So far, morphological repositories remain merely a hopeful aspiration, as none of the existing e-taxonomy projects are yet implementing them, though some authors have begun to make raw data sets available over the Internet for their published revisions (e.g. Henderson 2004).

Graphic Tools to Aid Alpha-Species Formation

If taxonomists used specimen data matrices rather than descriptions as the basis for routinely circumscribing species, it would be possible to envisage online analysis tools to explore character distinctions between alpha-species and to justify or test the circumscriptions of newly proposed species. The constantly changing datasets of specimens mean that species descriptions and species distinctions can go quickly out-of-date, but many years can pass before a thorough re-assessment is made by a taxonomist undertaking a new revision. Specimen-level morphological data provide a way to accumulate a standard dataset of morphological character information, which can be explored by graphical tools underpinned by statistical analysis. This can be visualised as a kind of virtual console in which the taxonomist, wherever they might be located, can download a large existing dataset, add their own new data and re-assess species boundaries using standardised techniques, giving a more transparent basis to their proposals for taxonomic changes.

Contributions, Rewards and Publication

Although to begin with, an e-taxonomy site will act as a parallel and subsidiary resource to the mainline published work on the plant group, it is likely that the momentum of generating a portal for web revisions will push taxonomic communities towards the view that new work made available on the sites constitutes publication. For the enterprise to work effectively, publishing new taxonomy on websites must offer authors a real and attractive professional advantage. While the Codes still specify paper publication for descriptions of taxa with new names, there is nothing to prevent new descriptive treatments from being so published. Without the commitment of the key taxonomists, however, the websites stand no chance of being kept up-to-date. Their commitment will only come if the career rewards are at least as great as those gained from conventional publication. If the websites are not kept up-to-date they will fail to meet the increasing global demand for a standard, web-accessible expert view of the taxonomy of any group (Godfray 2002). If the sites do not reflect the considered views of the experts, users will obtain their information from less critical sources; among these are the possibilities offered by internet engines which forage for information associated with the species names supplied by the user, e.g. “mash-up” technology (Scoble *et al.* 2007).

Building a website as a collective enterprise changes the status and role of authors quite significantly. While the initial content can be attributed to a single contributor (and web editor, if different), once the process of change and modification begins, it is easy to envisage that a species treatment will become a mosaic of contributions of different sizes and from potentially many people. It will no longer be possible to attribute such treatments to a single author, nor will the current convention of senior and junior authors be easily adapted to this end. The crucial importance of maintaining commitment from the contributing community means that new solutions are required. Like a medieval cathedral constructed over generations, a consensus taxonomy, by definition, must avoid the insoluble question to which individual should be granted greatest credit. The point is that a consensus is brought into existence by the labours of a broad collective on behalf of everyone; it will have no conventional authorship. This raises two questions.

First, if impact measures continue to exercise their current importance in career evaluation, would or should taxonomists be judged more by the extent of their contributions to this common global resource, than those they have made specifically to the research community via conventional scientific publication? From this standpoint, measures for contributions to the consensus are critically important and will require a quite different form compared to conventional authorship, although the principle of peer-review must remain fundamental to the process.

Second, are there grounds for grading the kinds of contributions made to the site, or is this an opportunity to abolish the implicit intellectual hierarchy of the conventional publication model? It is not so long ago that artists were not specifically acknowledged in the text of an illustrated botanical description. Still today botanical collectors who provide the named material for research are often vouchsafed an acknowledgment that seems, to them at least, quite inadequate given the effort required to make good collections and the overriding importance of authorship to career evaluation. The convention of the senior author as the intellectual prime mover of a piece of research reflects the importance of its creative and synthetic aspects and indeed has a noble history, harking back to times when the act of publication by free-thinking individuals was a courageous and even revolutionary act (see chapter on Galileo in Hawking 2002). But in modern times, this tradition can distort and undervalue the contributions of the many individuals who are almost always involved in any scientific publication; expertly documenting and naming a plant collection is also an intellectual exercise.

E-taxonomic websites offer an opportunity to atomise contributions and make possible the specialisation of information provision without losing the ability to gain meaningful career recognition. What this could mean in practice is that a person contributes those particular subsets of information about which they can realistically be most expert, given their own context and conditions.

Within the taxonomic community, discussions have focused on a new kind of report output from websites that could form the basis of career evaluation. Websites lend themselves to the automatic detection of contributions and user-views and so it would not be a difficult technical problem to provide a statement detailing, in some attractive quantified form, the contributions of individuals and the use that has been made of them in terms of page views, which could in turn form the basis for impact measures. It may be more challenging to shift the well-entrenched habits of evaluation used by research scientists in their own communities, but it seems probable that clear evidence of impact on the wider world will serve the broader interests of scientists and their institutions. In our view, alpha-taxonomists are consistently disadvantaged by current science metrics and have so far been unable to influence science administrators and managers to adopt more appropriately focused methods of evaluating their outputs.

Models of Social Organisation for E-Taxonomy

From what has been outlined above, it seems clear that developing an effective global portal for a taxonomic group involves not just data and technology but also "virtual institution"-building. The commitment and cooperation of the taxonomic community relevant to the taxonomic group in question is fundamental to the success of the enterprise. The concept of a website hosting alternative hypotheses as well as a consensus taxonomy is subtly different from the conventional scientific paradigm of the fruitful interaction among competing hypotheses. The debate and clash of scientific views are essential but these can be seen as important primarily to the scientific community involved; the wider world requires of the scientists that they produce a provisional common view, fit-for-purpose and regularly updated. This is the service and production aspect of taxonomic science and it is here, ironically perhaps, that taxonomists can gain their greatest recognition and make their most significant impact and, hopefully, find new sources of support and funding. But how is this to be achieved?

What kind of organisational model is needed for the taxonomic community, or at least that part which will volunteer to drive and maintain the website? The guiding principle must be to attract those taxonomists who have most to contribute, irrespective of other considerations. This implies avoiding personal and institutional branding issues, whilst accurately apportioning credit where this is due. It means that the human organisation that sustains the site should be, at least in theory, capable of self-transformation irrespective of any particular individual or institute. In reality, experience suggests that from a purely technical and financial point of view, the server management will probably need to be assumed by an institution with appropriate interests and infrastructure. But from the standpoint of human resources, the steering group or the editorial body, which is likely to be geographically widely distributed, could be subject to the same kind of periodic renewal processes normal for a scientific society, with elections and specified terms of office. Active members would select themselves by volunteering their services as taxon editors, reviewers, content providers and office-holders. The organisational structures of scientific societies and journals offer bases for developing appropriate models, respectively, for web-mediated self-renewing and open social organisations and peer-review systems. Despite the facility for communication which the Internet provides, experience suggests that to launch these human organisations effectively, face-to-face meetings of the communities are essential, both to begin the discussion process and to build the institutional structures. Some of the e-taxonomy website communities discussed below have already taken the first steps in this process.

Some Examples of Plant E-Taxonomy Websites

Despite the exciting prospects that it holds for the immediate future, creating e-taxonomy websites in practice, with the kind of functionality we have discussed, necessitates overcoming some difficult hurdles and solving some complex problems, both technical, financial and organisational. The major initial problem in setting up e-taxonomies is the requirement for suitable software and the availability of web-server resources and the technical staff to host and maintain it. As there is currently no off-the-shelf software, IT developers are needed. The five plant projects discussed below are currently tackling e-taxonomy using somewhat different approaches. Three of the projects, CATE, *Solanaceae* Source and GrassBase, have selected "guinea-pig" taxa and then developed custom-made software designed to trial different aspects of the e-taxonomy "vision". The EDIT Exemplar Groups each target a different family or large tribe of organisms to provide a broad range of requirements that feed the development of EDIT's Cybertaxonomic Platform http://wp5.e-taxonomy.eu/EDIT-Architecture.html). The focus of EDIT's plan for providing e-taxonomy functionality to taxonomists is driven primarily by pragmatism: to facilitate taxonomists' current practice by providing tools that make it easier to aggregate and integrate information and expert human resources. The combination of content teams and developers is gradually producing convergence on a more general model for various components such as the data model, the editing software and the web portals. Finally the EDIT Scratchpads project adopts the approach of making available to the broad taxonomic community a relatively simple generic content management system and encouraging taxonomists to come forward and try out e-taxonomy for themselves.

These projects currently constitute a dynamic international and inter-institutional environment for developing e-taxonomy and there are many significant interactions between them. Each tends to emphasise different aspects and priorities of the alpha-taxonomic process as a whole and they are thus broadly complementary to one another. Together with other web-based systematics projects, these initiatives are playing a prominent role in developing ideas and forward thinking for non-commercial scientific community organisations like Biodiversity Information Standards (TDWG, http://www.tdwg.org) which meet regularly to establish software standards for taxonomy.

An interesting feature of these projects is the way resources are used for personnel. All rely to a greater or lesser extent on investment in at least one software developer. In the case of EDIT Exemplar Groups, this support is mainly provided through an inter-institutional network rather than specifically for the taxonomic group in question. But all also rely on a botanist or zoologist who provides an interface between the established specialists and the software developers. What is emerging clearly is that these "interface e-taxonomists" represent a new layer of expertise and skills within taxonomy that could soon become crucial. This is a role for which the current younger generation will have a particular importance, through greater openness and adaptability to web technology and the emerging culture of web-based science. Less constrained by traditional methods of presenting taxonomic information, and with a more positive attitude to quantitative and modelling techniques for data analysis, this generation is needed to bridge between the databases, web portals and their designers, the encyclopaedic taxonomic data repositories in museums and herbaria, and the broad communities of taxonomic research scientists, both professional and non-professional.

1. *CATE* (http://www.cate-project.org, www.cate-araceae.org, http://www.cate-sphingidae.org)

The CATE (Creating a Taxonomic E-science) project, funded by the UK's Natural Environment Research Council (NERC) under their e-science programme, began in December 2005 and is a consortium of the Royal

Botanic Gardens Kew, Natural History Museum London and the University of Oxford Department of Zoology. CATE has its origins in a paper by Godfray (2002) which described the challenges facing taxonomy in the internet age, and suggested some avenues for innovation, including peer-reviewed consensus taxonomies. The CATE project was set up to test the feasibility of these ideas using pilot studies on two families, one botanical (*Araceae*, the aroids: http://www.cate-araceae.org) and one zoological (*Sphingidae*, the hawk moths: http://www.cate-sphingidae.org). In each case the goal is to build a site covering about 1,300 species using existing literature and collections to extract taxonomic and nomenclatural information, descriptions and keys, type specimen images, other images, literature references and links.

The design of the software is based on existing taxonomic information models such as Taxon Concept Schema (TCS, http://www.tdwg.org/activities/tnc/tcs-schema-repository/), Structure of Descriptive Data (SDD, http://www.diversitycampus.net/Projects/TDWG-SDD/index.html), Access to Biological Collections Data (ABCD, http://www.bgbm.org/TDWG/CODATA/Schema/) and the Common Data Model (CDM, http://wp5.e-taxonomy.eu/EDIT-Architecture.html). The aim is that the CATE software application will be generic, i.e. usable for other plant and animal groups. The CATE software architecture is database agnostic, i.e. it will work over a variety of different database formats and structures and the data is decoupled from on-screen delivery, which allows very flexible presentation. The application is being built to facilitate integration with other frameworks such as GBIF (Global Biodiversity Information Facility, http://www.gbif.org/) and Encyclopedia of Life (EoL, http://www.eol.org/) through the use of LSIDs (Life Science Identifiers) which are unique electronic tags for taxon concepts.

The target audience consists of two broad user types; taxonomists who both produce and need access to information, and non-taxonomists whose primary interest is to obtain access to information (although the system is open to anyone to make contributions). The management of the e-taxonomy website (the consensus taxonomy and any alternative hypotheses) can be divided into two main areas: content management, to be handled by the taxonomic community through an editorial body, and application management, which because of the costs of hosting and maintenance is probably only viable for appropriate major institutions such as natural history museums and herbaria.

Any user can propose changes to the revision, and the decision whether to include them or not in the consensus classification is made by review and moderation carried out by members of the interested community. Accepted proposals are incorporated into the next version of the consensus, but those that are rejected are nevertheless hosted on the site and visible as alternative hypotheses, attributed clearly to their authors.

Contributions of non-moderated content are also an important component and already in the process of being implemented on the two taxon sites. This provides a way for people to contribute information, images or ideas that do not require community approval, but which are nevertheless of interest and will be credited to their authors. This type of content can potentially be incorporated into future versions of the consensus.

We believe that the major innovation of CATE is its ambitious goal to provide software to implement a fully operational editing and peer-review system for proposals from the taxonomic community, making possible continuous update of consensus revisions. It focuses, more strongly than the other projects discussed here, on consensus taxonomies, peer-review and the integration of the relevant taxonomic community. The project also emphasises the provision of certain new features: illustrated interactive keys to all species and genera in the *Araceae* revision and a comprehensive set of standardised whole insect and genitalia images (the most critical morphological characters) in the case of the *Sphingidae*.

2. *Solanaceae Source* (http://www.nhm.ac.uk/solanaceaesource)

This major project, funded by the Planetary Biodiversity Inventory initiative of the U.S. National Science Foundation (NSF) as "PBI Solanum: a worldwide treatment", has since 2004 been creating a new worldwide monograph of the large and economically important genus *Solanum*. The project is part of a much larger effort by the NSF to fund groups of taxonomists working on alpha-taxonomy and making it available electronically on the web — the PBI initiative began in 2002. PBI Solanum has brought together an international team of taxonomic specialists of the family *Solanaceae* along with students and collaborators, and the website is currently managed at the Natural History Museum, London. Although existing modern monographs and revisions are incorporated into the internet monograph, a key feature of this project is the coordination of specialists to write entirely new revisions where these are needed. The project is designed to strengthen both integration of the specialist community by joint studies in the field and herbarium, and joint presentation of results at scientific meetings and through public outreach activities. The student training programme includes postdoctoral, postgraduate and undergraduate levels and is crucial to building the next generation of specialists. Another notable feature of the PBI Solanum project is the provision of small grants to recruit taxonomic collaborators who are not part of the core team, both established specialists and students, around the world. This mechanism helps to enlarge the critical mass of the support community for the website, which currently involves more than 30 taxonomists from 12 countries.

Another important feature is the inclusion of specimen data on the website in order to show how the species concepts were arrived at — all specimens determined by the author of a PBI treatment before its posting on the site are the gold standard for that species concept — new concepts can have new data. There is also a peer-review system where new descriptions are posted on the website for all partners to review and comment upon.

Solanaceae Source is a pioneering model of the broad functionality and range of information that taxonomists expect to provide through e-taxonomy websites. The strengths of the project are built on the provision of a funding base that is sufficient to ensure collaboration from the key scientists and their institutions, particularly engagement with two of the world's most important taxonomic organisations (The Natural History Museum and the New York Botanical Garden) and the project is linked with EDIT (see below). Another strength is the linkage with the genomics community through the SOL initiative, who are currently sequencing the tomato (*Solanum lycopersicum*) and potato (*Solanum tuberosum*) genomes in an international consortium (see http://www.cornell.sgn.org).

Unlike the Monocot websites discussed here, there was no prior species checklist for *Solanum* and its creation was an important and time-consuming first step. The core information provided through species pages is the standard data for monographs but enormously enhanced by a variety of links. These include: nomenclature, species descriptions and identification keys (currently provided as pdf or hyperlinked text, interactive keys are being developed), images of the plants (both in the field and herbarium specimens, including links to major image libraries), details of voucher specimens in herbarium collections (including links to major online specimen databases at New York Botanical Garden and Missouri Botanical Garden) and taxonomic bibliography. There is also access to phylogenetic hypotheses through interactive cladograms, links to *Solanaceae* genomic project websites and links from species pages to molecular sequence and chromosome databases. In addition, there is a wide range of other information about the genus *Solanum* and the family *Solanaceae*, and an attractive homepage with current news items and links.

3. *GrassBase*

The online world grass Flora predates many of the other initiatives described here. It is available through the Kew website at http://www.kew.org/data/grassbase/ and currently consists of two separate databases, namely a nomenclature/synonymy database and a descriptions database. Until now, these databases have been designed with a single primary use case in mind — making the production of Floras and monographs easier by giving the taxonomist easy access to the current state of knowledge on the family. A full description of the content of these databases is available in Harman & Clayton (2007), but they are summarised here for convenience. The nomenclature database contains full synonymy for the family, regional distribution of accepted names, type information and publication details. It is available for download as a MS Access application, and has a number of query forms built in. These include the ability to search for accepted names and synonyms, list species within a genus/area, list type and publication information and list geographic distribution for a species. Results are linked to the descriptions database via a hyperlink which opens up the description within a browser. Distribution information follows the TDWG standard (Brummitt 2001). The descriptions database consists of 11,000 species and 700 generic descriptions using 1090 different morphological characters. It is currently stored within the DELTA format (Dallwitz 1980; Dallwitz *et al.* 1993 onwards) which allows for the data matrix to be translated into natural language (as has been done for the species descriptions available online), into an interactive key to be run by INTKEY (Dallwitz *et al.* 1995 onwards), or into various data analysis formats such as the NEXUS format (Maddison *et al.* 1997). It is also possible to produce conventional binary identification keys and alternative language descriptions from the DELTA data, though this has not been done for the data within GrassBase.

Species descriptions are coded from species accounts within the various regional Floras or from the original publication coupled with examination of specimens when no Flora account is available. This is the only project described here in which the data is fully atomised. Having a standardised character set (http://www.kew.org/data/grasses-db/chars.htm) makes possible considerable uniformity in descriptive content, but the drawback is that they are very dull to read. Generic descriptions are derived from those of the constituent species by automatically summarising the characters and character states using INTKEY's functionality. This gives an indication of the frequency of individual character states within a genus and a framework upon which a taxonomist can base their final treatment.

Users are invited to provide feedback or corrections to both databases. However, it is difficult for users to see the impact of their comments and corrections. One aspect that is missing from the current GrassBase framework is the main community-driven e-taxonomic framework described earlier, and towards this end the data from GrassBase has been ported into a prototype collaborative web portal running the Drupal content management system (CMS) as part of a recent M.Sc. research project (Harman 2007). This portal allows registered users to provide feedback on individual pages, has built in versioning of individual pages, and allows pages to be edited online. This dataset could provide a useful test case for any future developments within this emerging field.

4. *EDIT Exemplar Groups: Palmweb* (http://www.palmweb.org/species.php) and *International Cichorieae Network* (http://wp6-cichorieae.e-taxonomy.eu/blog/home/)

Palmweb and the International *Cichorieae* Network (together with the Insect Exemplar Group, http://www.edit-insects.info/) are Exemplar Groups of the EDIT project, a five-year (begun March 2006) European Union-funded Network of Excellence initiative designed to promote integration between taxonomic institutes, particularly, but not exclusively, those in Europe.

Palmweb grew out of the European Network of Palm Scientists, which meets regularly and brings together scientists from various disciplines (EUNOPS, http://eunops.org). Members of this working group had already initiated a website (Palm base, http://www.palmbase.org) and Palmweb is a development of this on a larger scale. Other broader initiatives in the global community of palm taxonomists designed to aggregate an even larger critical mass of taxonomists are already looking to Palmweb as a major disseminating resource for taxonomic research results. In this project, the mobilisation of the taxonomic community is running in parallel with technological developments to upload current palm taxonomy onto the site.

Like the CATE *Araceae* site, Palmweb uses as its foundation the checklist of accepted species published electronically by the World Checklist of Monocotyledons (http://www.kew.org/wcsp/monocots). One of its mid- to long-term goals is to promote the development of an interactive technology so that the expert community can gradually take over the responsibility of updating the checklist by inter-operability with the Palmweb software. Palmweb is focused more on providing a very broad range of source materials to taxonomists, for example, multiple previously published descriptions, rather than developing a peer-reviewed consensus taxonomy.

The International *Cichorieae* Network site, live since February 2007, is based at the Botanic Garden and Botanical Museum Berlin-Dahlem, which has a research focus on the family *Compositae*. The site has already had a significant impact on the large global community of biologists interested in the huge family of the *Compositae*, with participants from 14 countries already registered. The International *Cichorieae* Network uses data from and interacts with Euro+Med Plantbase (http://ww2.bgbm.org/EuroPlusMed/query.asp), which provides a regional on-line checklist of *Compositae*, and closely cooperates with the project of a Global *Compositae* Checklist (coordinated by Landcare Research, New Zealand and Wageningen University, The Netherlands). The *Cichorieae* website team is currently working on two major tasks: data collecting and data display. For data collecting, cooperation is being sought with other institutions as data contributors (e.g. herbaria). Integration of the considerable amount of already extant data is crucial here. Work on data display currently involves the taxonomic backbone, with nomenclatural information including scanned protologues and type specimen data with images, and descriptions; this is a major task in a highly diverse plant group with a long history of taxonomic study by European botanists.

The characteristic feature of these EDIT Exemplar Groups is that they are an integral part of a larger project. These taxonomic content sites form part of Work Package 6 of the EDIT project. They have been created not only to mobilise their respective taxonomic communities and put taxonomic content onto the Internet, but particularly to provide specific test cases and information requirements for the development of the EDIT Internet Platform for Cybertaxonomy team (http://wp5.e-taxonomy.eu/EDIT-Architecture.html). The latter is setting up a broad integrative environment within which taxonomists can work more efficiently using both new and existing tools. These will be freely available and designed to be compatible with a wide range of software systems.

5. *EDIT Scratchpads*

The EDIT Scratchpads (http://www.editwebrevisions.info/scratchpads) are a complementary approach to the Exemplar Groups and focus on the critical challenge of attracting taxonomists into “e-taxonomy space”. Under this project, the Natural History Museum in London hosts and provides support for websites about any kind of organism, built on the generic content management software Drupal (http://drupal.org). Once the support staff has set up the project site, users can then create and manage their own content using step-by-step instructions online, with technical support from the Museum (http://www.editwebrevisions.info/application). The Scratchpads have already been very successful; at the time of writing 56 different websites have been created, ranging from Malesian mosses (http://malesianmosshandbook.myspecies.info/) to flies (http://diptera.myspecies.info/).

The Scratchpads project is especially interesting from two viewpoints. (1) It shows the crucial role that the major institutes need to play in hosting technically expensive infrastructure and providing expert technical support, beyond the reach of the individual taxonomists. (2) It demonstrates that there is a significant demand in the broad taxonomic community for this functionality. There is thus an excellent prospect for a "snowball effect" in e-taxonomy, given the availability of user-friendly technology in stable, sustainable form; this could be highly significant evidence for future funding prospects.

In conclusion, experience gained so far emphasises the crucial role of the major regional and national taxonomic organisations in implementing alpha e-taxonomy as a global resource. In this paper we have emphasised the value of the Internet for expanding the sources of expertise and knowledge for alpha taxonomy, in effect widening direct participation in alpha-taxonomy to include everyone who has a serious interest in the subject. But it is also clear that there is a vital need for flexible and collective leadership of the global network by the institutions and scientific societies that represent the international taxonomic community. The establishment of an effective global system based on reliable common standards and its long term management and development lies surely with these organisations.

Epilogue

Janzen (2004) envisioned an extra-terrestrial being landing on Earth, pointing a hand-held barcorder at plants and animals and getting instant DNA identifications. What about something a little more familiar? A piece of vegetation that represents the last bit of forest in a particular area. A local college or university biology department feels impelled to take action; they need to train their students, become the local experts, influence local and national politicians and landowners, schools and agriculturists, newspapers and TV stations, internet communications. They need to make the case, explain what is there and why it is important, recruit communities in the defence of these remaining precious stands of biodiversity-rich vegetation. They need information.

The Toyota stops, they step out, they point "it" around, satellite communications snap them seamlessly into a network of global DNA data facilities…and back in the lab/herbarium/museum, they download the identifications. A total of 1,985 plant taxa. Now what do they do?

To get any further down this road, these names must open the door to a variety of information sources, but especially to the rich and well-organised landscape of alpha e-taxonomy websites. In this paper we have described e-taxonomy as a two-way interaction which puts the experts directly in contact with the conservation frontline and the global perspective of species at the disposal of a local project. Local information can feed a global database and those who undertake this can obtain in return globally visible and measurable credit rewards. The alpha-taxonomists, morphological and molecular, are now coordinating and managing this as a fully functioning system, with a key role in the daily challenge to preserve our biodiversity and environment. There is no longer a barrier between local and global, provincial and metropolitan, amateur and professional, us and them. Instead there is a shared growth of knowledge and a deepening of understanding.

Acknowledgements

The authors are grateful to two reviewers for their comments and suggestions. We are also very grateful for the financial support received from the UK Natural Environment Research Council, grant numbers NE// C001532/1; NE/C51588X/2; NE/C515871/1.

References

Berlin, B. (1973). Folk systematics in relation to biological classification and nomenclature. Ann. Rev. Ecol. Syst. 4: 259 – 271.

Blaxter, M. L. (2004). The promise of a DNA taxonomy. Philos. Trans., Ser. B 359: 669 – 679.

____, Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., Abebe, E. (2005). Defining operational taxonomic units using DNA barcode data. Philos. Trans., Ser. B 360: 1935 – 1943.

Boyd, R. (1999). Homeostasis, species and higher taxa. In: R. A. Wilson (ed.), Species: new interdisciplinary essays. Pp. 141 – 185. MIT Press, Cambridge MA.

Brummitt, R. K. (2001). World Geographical Scheme for Recording Plant Distributions, Edition 2. Plant Taxonomic Database Standards No. 2. Hunt Institute for Botanical Documentation, Carnegie Mellon University, Pittsburgh.

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N. & Savolainen, V. (2005). Land plants and DNA barcodes: short-term and long-term goals. Philos. Trans., Ser. B 360: 1889 – 1895.

Dallwitz, M. J. (1980). A general system for coding taxonomic descriptions. Taxon 29: 41 – 46.

____, Paine, T. A. & Zurcher, E. J. (1993 onwards). User's Guide to the DELTA System: a General System for Processing Taxonomic Descriptions. 4th edition. http://delta-intkey.com/

____, ____ & ____ (1995 onwards). User's Guide to INTKEY: a Program for Interactive Identification and Information Retrieval. 1st edition. http://delta-intkey.com/

Davis, P. H. & Heywood, V. H. (1963). Principles of Angiosperm Taxonomy. Oliver & Boyd, Edinburgh.
Davis, J. I. & Nixon, K. (1992). Populations, genetic variation, and the delimitation of phylogenetic species. Syst. Zool. 41: 421 – 435.
De Ley, P., De Ley, I. T., Morris, K., Abebe, E., Mundo-Ocampo, M., Yoder, M., Heras, J., Waumann, D., Rocha-Olivares, A., Burr, A. H. J., Baldwin, J. G. & Thomas, W. K. (2005). An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. Philos. Trans., Ser. B 360: 1945 – 1958.
de Queiroz, K. (1999). The general lineage concept of species and the defining properties of the species category. In: R. A. Wilson (ed.), Species: new interdisciplinary essays. Pp. 49 – 89. MIT Press, Cambridge MA.
DeSalle, R. (2006). Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. Conserv. Biol. 20(5): 1545 – 1547.
Forey, P. L., Fortey, R. A., Kenrick, P. & Smith, A. B. (2004). Taxonomy and fossils: a critical appraisal. Philos. Trans., Ser. B 359: 639 – 653.
Gaston, K. J. & May R. M. (1992). The taxonomy of taxonomists. Nature 356: 281.
Godfray, H. C. J. (2002). Challenges for taxonomy. Nature 417: 17 – 19.
____ (2007). Linnaeus in the information age. Nature 446 (15 March 2007): 259 – 269.
____ & Knapp, S. (2004). Introduction. Theme Issue "Taxonomy for the twenty-first century". Philos. Trans., Ser. B 359: 559 – 569.
____, Clark, B. R., Kitching, I. J., Mayo, S. J. & Scoble, M. J. (2007). The web and the structure of taxonomy. Syst. Biol. 56: 943 – 955.
Harman, K. T. (2007). GrassBase — Steps towards integrating the online world grass Flora. MSc Thesis. Cranfield University.
____ & Clayton, W. D. (2007). Recent developments in Kew's grass databases (GrassBase). Kew Bull. 62: 375 – 379.
Harris, D. J. & Froufe, E. (2005). Taxonomic inflation: species concept or historical geopolitical bias? Trends Ecol. Evol. 20: 6 – 7.
Hawking, S. (2002). On the shoulders of giants. Pp. 391 – 626. Running Press, Philadelphia.
Henderson, A. (2004). A multivariate analysis of *Hyospathe* (*Palmae*). Amer. J. Bot. 91: 953 – 965.
Hennig, W. (1966). Phylogenetic Systematics. University of Illinois Press, Urbana.
ICZN (International Commission on Zoological Nomenclature) (1999). International Code of Zoological Nomenclature (Fourth Edition). [http://www.iczn.org/iczn/index.jsp]
Isaac, N. J. B., Mallet, J. & Mace G. M. (2004). Taxonomic inflation: its influence on macroecology and conservation. Trends Ecol. Evol. 19: 464 – 469.
Janzen, D. H. (2004). Now is the time. Philos. Trans., Ser. B 359: 731 – 732.
____, Hajibabaei, M., Burns, J. M., Hallwachs, W., Remigio, E. & Hebert, P. D. N. (2005). Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Philos. Trans., Ser. B 360: 1835 – 1845.
Judd, W. S., Campbell, C. S., Kellogg, E. A, Stevens, P. F. & Donoghue, M. J. (2002). Plant systematics: a phylogenetic approach. 2nd edition. Sinauer Associates, Sunderland, Massachusetts.
Knapp, S., Nic Lughadha, E. & Paton, A. (2005). Taxonomic inflation, species concepts and global species lists. Trends Ecol. Evol. 20: 7 – 8.
Lexer, C., Rosenthal, D. M., Raymond, O., Donovan, L. A. & Rieseberg, L. H. (2005). Genetics of species differences in the wild annual sunflowers, *Helianthus annuus* and *H. petiolaris*. Genetics 169: 2225 – 2239.
Lipscomb, D., Platnick, N. & Wheeler, Q. (2003). The intellectual content of taxonomy: a comment on DNA taxonomy. Trends Ecol. Evol. 18(2): 65 – 66.
Mace, G. M. (2004). The role of taxonomy in species conservation. Philos. Trans., Ser. B 359: 711 – 719.
Maddison, D. R., Swofford, D. L. & Maddison, W. P. (1997). NEXUS: An Extensible File Format for Systematic Information. Syst. Biol. 46: 590 – 621.
Mallet, J., Isaac, N. J. B. & Mace, G. M. (2005). Response to Harris and Froufe, and Knapp *et al.*: taxonomic inflation. Trends Ecol. Evol. 20: 8 – 9.
Markmann, M. & Tautz, D. (2005). Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences. Philos. Trans., Ser. B 360: 1917 – 1924.
McNeill, J., Barrie, F. R., Burdet, H. M., Demoulin, V., Hawksworth, D. L., Marhold, K., Nicolson, D. H., Prado, J., Silva, P. C., Skog, J. E., Wiersema, J. H., Turland, N. J. (2006). International Code of Botanical Nomenclature (Vienna Code). ARG Gantner Verlag, Ruggell, Liechtenstein.
Paton, A. J., Brummitt, N., Govaerts, R., Harman, K., Hinchcliffe, S., Allkin, R. & Nic Lughadha, E. (in press). Towards Target 1 of the Global Strategy for Plant Conservation: a working list of all known plant species — progress and prospects. Taxon.
Pillon, Y. & Chase, M. W. (2006). Taxonomic exaggeration and its effects on orchid conservation. Conserv. Biol. 21: 263 – 265.
____, Fay, M. F., Shipunov, A. B. & Chase, M. W. (2006). Species diversity versus phylogenetic diversity: a practical study in the taxonomically difficult genus *Dactylorhiza* (*Orchidaceae*). Biol. Conserv. 129: 4 – 13.
Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamouon, S., Sumlin, W. & Vogler, A. (2006). Sequence-based species delim-

itation for the DNA taxonomy of undescribed insects. Syst. Biol. 55: 595 – 609.
Samper, C. (2004). Taxonomy and environmental policy. Philos. Trans., Ser. B 359: 721 – 728.
Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K. & Lane, R. (2005). Towards writing the encyclopaedia of life: an introduction to DNA barcoding. Philos. Trans., Ser. B 360: 1805 – 1811.
Scoble, M. J. (2004). Unitary or unified taxonomy? Philos. Trans., Ser. B 359: 699 – 710.
——, Clark, B. R., Godfray, H. C. J., Kitching, I. J. & Mayo, S. J. (2007). Revisionary taxonomy in a changing e-landscape. Tijdschr. Entomol. 150: 305 – 317.
Scotland, R., Hughes, C., Bailey, D. & Wortley, A. (2003). The *Big Machine* and the much-maligned taxonomist. Syst. Biodivers. 1 (2): 139 – 143.
Smith, M. A, Fisher, B. L. & Hebert, P. D. N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos. Trans., Ser. B 360: 1825 – 1834.
Soltis, D. E., Soltis, P. S., Endress, P. K. & Chase, M. W. (2005). Phylogeny and evolution of Angiosperms. Sinauer Associates, Sunderland, Massachusetts.
Stearn, W. T. (1992). Botanical Latin. Fourth Edition. Timber Press, Portland.
Stevens, P. F. (2007 onwards). Angiosperm Phylogeny Website. Version 8, June 2007 [and more or less continuously updated since]. http://www.mobot.org/MOBOT/research/APweb/.
Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. (2003). A plea for DNA taxonomy. Trends Ecol. Evol. 18 (2): 70 – 74.
Yergeau, E., Filion, M., Vujanovic, V. & St-Arnaud, M. (2005). A PCR-denaturing gradient gel electrophoresis approach to assess *Fusarium* diversity in asparagus. J. Microbiol. Meth. 60: 143 – 154.
Zander, R. H. (2007). When biodiversity study and systematics diverge. Biodiversity 8: 43 – 48.