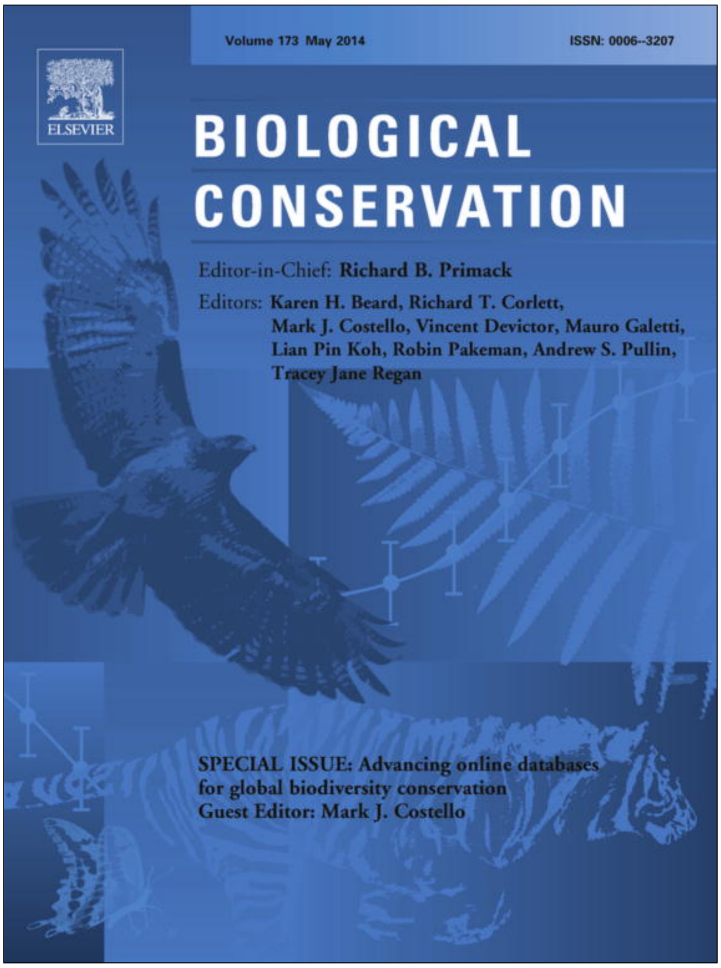
Volume 173 May 2014
ISSN: 0006-3207
ELSEVIER
BIOLOGICAL CONSERVATION
Editor-in-Chief: Richard B. Primack
Editors: Karen H. Beard, Richard T. Corlett, Mark J. Costello, Vincent Devictor, Mauro Galetti, Lian Pin Koh, Robin Pakeman, Andrew S. Pullin, Tracey Jane Regan
SPECIAL ISSUE: Advancing online databases for global biodiversity conservation
Guest Editor: Mark J. Costello

Biological Conservation 173 (2014) 155–165

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Strategies for the sustainability of online open-access biodiversity databases

Mark J. Costello [a,*], Ward Appeltans [c,l], Nicolas Bailly [j], Walter G. Berendsohn [b], Yde de Jong [d], Martin Edwards [h], Rainer Froese [g], Falk Huettmann [k], Wouter Los [d], Jan Mees [c,e], Hendrik Segers [f], Frank A. Bisby [i,1]

[a] Institute of Marine Science, Leigh Marine Laboratory, University of Auckland, New Zealand
[b] Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Koenigin-Luise-Str. 6-8, 14195 Berlin, Germany
[c] Flanders Marine Institute (VLIZ), Wandelaarkaai 7, B-8400 Ostend, Belgium
[d] Faculty of Science, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands
[e] Ghent University, Department of Biology, Marine Biology Research Group, Krijgslaan 281/S8, 9000 Gent, Belgium
[f] Belgian Biodiversity Platform, Freshwater Laboratory, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, B-1000 Brussels, Belgium
[g] Helmholtz-Centre for Ocean Research, Duesternbrooker Weg 20, 24105 Kiel, Germany
[h] Sir Alister Hardy Foundation for Ocean Science, Citadel Hill, The Hoe, Plymouth PL1 2PB, UK
[i] University of Reading, UK
[j] WorldFish – FishBase Information and Research Group, Aquatic Biodiversity Informatics Office, Los Baños, 4031 Laguna, Philippines
[k] EWHALE Lab, University of Alaska-Fairbanks (UAF), Fairbanks, AK 99775, United States
[l] Intergovernmental Oceanographic Commission of UNESCO, IOC Project Office for IODE, Wandelaarkaai 7/61, 8400 Oostende, Belgium

ARTICLE INFO

Article history:
Received 3 March 2013
Received in revised form 25 July 2013
Accepted 30 July 2013
Available online 11 November 2013

Keywords:
Species
Taxonomy
Publication
Citation
Funding
Intellectual property

ABSTRACT

Scientists should ensure that high quality research information is readily available on the Internet so society is not dependant on less authoritative sources. Many scientific projects and initiatives publish information on species and biodiversity on the World Wide Web without users needing to pay for it. However, these resources often stagnate when project funding expired. Based on a large pool of experiences worldwide, this article discusses what measures will help such data resources develop beyond the project lifetime.

Biodiversity data, just as data in many other disciplines, are often not generated automatically by machines or sensors. Data on for example species are based on human observations and interpretation. This requires continuous data curation to keep these up to date. Creators of online biodiversity databases should consider whether they have the resources to make their database of such value that other scientists and/or institutions would continue to finance its existence. To that end, it may be prudent to engage such partners in the development of the resource from an early stage. Managers of existing biodiversity databases should reflect on the factors being important for sustainability. These include the extent, scope, quality and uniqueness of database content; track record of development; support from scientists; support from institutions, and clarity of Intellectual Property Rights. Science funders should give special attention to the development of scholarly databases with expert-validated content. The science community has to appreciate the efforts of scientists in contributing to open-access databases, including by citing these resources in the Reference lists of publications that use them. Science culture must thus adapt its practices to support online databases as scholarly publications.

To sustain such databases, we recommend they should (a) become integrated into larger collaborative databases or information systems with a consequently larger user community and pool of funding opportunities, and (b) be owned and curated by a science organisation, society, or institution with a suitable mandate. Good governance and proactive communication with contributors is important to maintain the team enthusiasm that launched the resource. Experience shows that 'bigger is better' in terms of database size because the resource will have more content, more potential and known uses and users of its content, more contributors, be more prestigious to contribute to, and have more funding options. Furthermore, most successful biodiversity databases are managed by a partnership of individuals and organisations.

* Corresponding author. Tel.: +64 94145031.
E-mail address: m.costello@auckland.ac.nz (M.J. Costello).
[1] Authors are alphabetical except for the first (lead) author and the last (deceased).

0006-3207/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biocon.2013.07.042

1. Introduction

The Internet has rapidly become the first place most people look for information. They will find information on almost anything, but still in a somewhat anarchic form. It behoves the scientific community to ensure that authoritative information is available, can be easily distinguished from less scholarly or up-to-date resources, and is as comprehensive as possible (Costello et al., 2006). As the print media developed, so did peer-review and editorial control, where individual scientists published their work through resources whose quality is primarily controlled by the scientific community. A similar approach is desirable for online resources, including taxonomic and ecological databases (Costello and Vanden Berghe, 2006; Costello, 2009a; Huettmann, 2005, 2009; Costello et al., 2013a,b). Thus online resources should be led, edited and authored by well-qualified experts (Zuckerberg et al., 2011; Cushman and Heuttmann, 2010). As in the print media, scientists need to discipline their efforts to create authoritative, collaborative, online information resources and databases overseen by the scientific community. Print publications have the benefit of one-off production costs, revenue earned from sales, and archiving in libraries, but they also entail significant costs for scientific institutions to maintain subscriptions, library premises and staff. Particularly in areas of scientific research, they may become out of date quickly. In contrast, web resources can be regularly updated and upgraded at little cost, but they require regular input and quality control efforts that have continuous costs for maintenance. Furthermore, instead of earning revenue from sales or user fees, both contributors and users expect their access to be free and open. Criticisms from users who have unrealistic expectations that online resources should be even more comprehensive and accurate than the print literature can undermine the support for these same resources. Thus, once project funding to establish a web resource has expired, it can become difficult to maintain it (Graham et al., 2004; Merali and Giles, 2005; Costello, 2009b; Drew, 2011; Brewer et al., 2012). This is especially a problem for stand-alone web resources and short-term projects that are serving data based on human validation, and thus dependent on curation to maintain quality and up to date service. Although the lack of long-term funding support for such open-access databases is a great worry amongst scientists, the options for such funding, and the factors facilitating a sustainable business model, have not been discussed in the literature. Here we seek to address this from our experiences in developing a variety of biodiversity databases since the 1990s, and observations of related initiatives world-wide. We provide examples of how some leading global, online biodiversity databases have been developed, managed, and governed; describe the challenges and costs in their maintenance; and the importance of clarity in Intellectual Property Rights for database succession planning. We conclude with a summary of management options and recommendations to maximise the sustainability of biodiversity databases.

2. Biodiversity databases

Some biodiversity databases collect primary data and information, such as measurements from instruments, records of biological specimens and observations, and/or may contain expert judgements on species concepts and classifications, or interpretations of data and current knowledge. Much of the content may have been derived from the published literature, as is the case in many scientific papers and books. Other web sites, such as iSpecies and the Encyclopedia of Life (EoL), aggregate content from such databases. Here we are primarily concerned with primary resources validated by experts in the subject. These experts choose to contribute their time to making their accumulated data and knowledge publicly available because they see this as an important service to science and society (Fig. 1). These resources are thus scholarly publications and should be so recognised (Box 1). However, this is often not practised by their institutes, funders or publishers. For example, the manner of citation of online resources in journals is variable (Box 2).

Box 1 Databases as publications.

It is important that we consider these online resources as scientific publications (Huettmann, 2007a; Costello, 2009a). All the authors and editors involved should be named and their contributions citable. The scientist's names and standing lend authority to the quality of the content. Knowing the authors can also indicate their bias, and as they may support particular schools of thought or advocate particular approaches to issues related to their science. Citations provide the recognition that scientists require for their career development, and that their employer organisations and funding agencies may require to demonstrate their productivity. In science, authors and editors do not typically get paid for publishing their knowledge, but it is the practice to cite their work, and it influences their reputation, employability and promotion. As Latour (1987) explained in a still debated interpretation of what is science, the capital of the scientists is not money but rather recognition that is gained by citations of their work, which is measured today on traditional publications (printed or electronic). Citation also has implications for the permanency and archiving of versions of the database, because science requires accessibility to a resource as it was cited, not only the latest version.

At least in the biology and ecology disciplines, the scientific community appears still to be reluctant when it comes to citing online databases and generally prefer to cite traditional references. For example, some of the species databases from the FADA project have been made available as classical scholarly publications, in addition to as a web resource. Whereas both are equally sound scientifically, having passed peer review, the latter is being regularly updated and should therefore be the more useful resource. Nevertheless, whereas one of them (Segers, 2007) has been cited 71 and 102 times as a paper publication, not a single citation of any of the 14 FADA online databases could be traced (Science Citation Index Expanded 3rd February 2012 and Google Scholar search 7th February 2012 respectively). Similarly, the editors of WoRMS were publishing a series of synthesis papers in a journal, but of the first five papers so published, not one cited the WoRMS database in their references, and only two mentioned it in their text. Thus, even the scientists developing scholarly archived web resources need to be reminded to cite them; and the same applies to editors and publishing houses.

Box 2 Database citation.

One obstacle to online resources being supported and recognised in citation metrics is that some journals only cite web resources in the text or footnotes (Hardisty and Roberts, 2013). Thus their citations are not picked up by journal citation tracking systems that use DOIs or similar metrics. For example, FishBase, an online database with standardised information on all fishes, had 1666 citations of the different versions of ''Froese, R., Pauly, D. FishBase. World Wide Web Electronic Publication'' in the ISI Web of Science, under their ''Cited References Search,'' in March 2011. Yet, these citations are not counted in most citation tracking systems for either author because FishBase is not included among the ISI-screened journals. Also, neither authors' institutes count the bi-monthly online editions of FishBase in their compilations of staff publications. This is a typical situation and has implications for the annual budget assignments of their departments and the recognition of the FishBase scientists themselves, and digital databases in science overall.

A particular challenge of open-access databases is that their content may be re-published by other websites, sometimes without attribution of the source. Even when a recommended citation is provided by a website, users often fail to use it and at best only cite the link to the webpage (i.e. the universal resource locator or url) which is not a permanent address. The best-practice community norm that applies to citing print media is not well established for web-based publications yet, even when they are permanently archived (like print publications are in libraries). Many editors and journals are negligent about it, despite the fact that the online database was considered important enough to cite, and sometimes contained all the data used in the study. Authors may only cite the secondary or tertiary web source as if it was a primary source. Thus, the authority (and funders and institutional recognition) behind the original material is lost. This is also an issue in the published literature, namely that many users want synthesised information. Thus review papers and textbooks have a much wider readership than primary articles. Experience has shown that such publications have high citation rates and financial benefits which help motivate scientists to contribute. Open-access online resources, being so readably accessible, should have even higher citation rates. However, most are not even formally cited as scholarly references.

Recent efforts to establish ''data papers'' (Chavan and Penev, 2011) as a means to take advantage of the established way of editing, peer review and citation indexing for referencing of databases may provide a solution to this problem, but it remains to be seen if this can be established in the scientific community (Costello et al., 2012a, 2013b). The Thomson Reuters Data Citation Index has been established to parallel the citation indices for scholarly papers. It tracks citations of data in repositories that meet its criteria, including persistence, peer-review, and links to research literature (Thomson Reuters, 2012). In addition, Datacite statistics report which datasets DOI were the most accessed (http://stats.datacite.org). Both of these approaches provide metrics of data use, comparable to citations of literature.

Leading examples of biodiversity databases are the Global Biodiversity Information Facility (GBIF) and GenBank. Both have thousands of contributing scientists worldwide and the longer established GenBank has thousands of users in the research community. Both are open-access and receive direct financial support from governments. GBIF was devised by the OECD Megascience Forum, initiated by a meeting of the OECD country's research ministers, and has been recognised by the Convention on Biological Diversity (CBD) as the leading source of primary biodiversity data. Thus financial support for GBIF is provided by many countries because it supports their international agreements and policies (e.g. under the CBD) to make biodiversity data publicly available. However, if we consider a decade as sufficient time for a resource to be 'established' there are several other databases that have survived without direct government support, although most have had indirect support through grants from research funding organisations.

The Ocean Biogeographic Information System (OBIS) was initially developed as part of the Census of Marine Life (CoML) through funding from the A.P. Sloan Foundation such that it publishes over 30 million species distribution records and is one of the largest contributors of data to GBIF (Costello et al., 2007). It is now supported by governments and in-kind institutional commitments under the governance of the International Oceanographic Data and Information (IODE) programme of UNESCO's Intergovernmental Oceanographic Commission (IOC). The Continuous Plankton Recorder (CPR) survey has created a database of plankton distribution across the oceans and was built on many decades of government support. After a funding crisis where direct government support was abruptly ended, it became an independent charity. The CPR survey is now managed by the Sir Alister Hardy Foundation for Ocean Sciences (SAHFOS) with most of its financial support from several countries as well as research funding.

Examples of leading taxon-specific biodiversity resources include AlgaeBase (Guiry and Guiry, 2013), FishBase (Froese and Pauly, 2013), and Biogeoinformatics of the Hexacorals (Fautin and Buddemeier, 2008). Their development has been funded by a sequence of numerous grants and contracts won by their respective scientific champions, and FishBase is governed by a consortium of institutions. The Freshwater Animal Diversity Assessment (FADA), in comparison, started off as a project with a primarily scientific goal, the production of an authoritative review of freshwater biodiversity (Balian et al., 2008), and only in the second instance turned its attention towards the publication of the databases underpinning these reviews. The Catalogue of Life (Bisby et al., 2012) is a publication listing all species on Earth led by Species 2000, a not-for-profit non-governmental organisation based in the UK, and ITIS, a U.S. federal programme. The involvement of scientists from various institutions allowed it to find a new host organisation in 2013 (i.e. Naturalis, Leiden, Netherlands) when the previous host institution could not afford it anymore. A common feature to almost all of the databases is that they involve collaboration between scientists in different organisations, formal partnership agreements with host institutions and/or consortia agreements.

2.1. Role of host institutions

Many successful databases have been developed and maintained by institutions. For example, the Index Herbariorum is a traditional botanical resource (in print since 1952) that went electronic in 1997. It lists the herbaria of the world, including an abbreviation for each collection which is used everywhere when botanical specimens are cited. It is hosted by the New York Botanical Gardens (Thiers, 2012). The International Plant Name Index

(IPNI), probably the most important online index of plant names, is a collaboration of three traditional botanical nomenclators on-line, namely Index Kewensis, Harvard Gray Card Index, and the Australian Plant Name Index; respectively hosted by The Royal Botanic Gardens Kew, The Harvard University Herbaria, and the Australian National Herbarium (IPNI, 2012). Collaborative networks of institutional databases may become increasingly important as the Internet, and mechanisms of Internet data provision, become more stable. Indeed, GBIF set out (and in part still is) a network of institutional data providers. The mechanism to provide access to distributed databases through common protocols and data standards are well developed (e.g. Holetschek et al., 2009) and joint data provision may help to achieve impact and synergies that aid in long-term sustainability of the databases themselves (Gemeinholzer et al., 2011).

Other biodiversity databases are edited by members of the scientific community and hosted by institutions. For example, MycoBank, the database of and administered by the International Mycological Association for new fungal taxa (Robert et al., 2005) is hosted by the Centraalbureau voor Schimmelculturen (CBS) Fungal Biodiversity Centre – an institute of the Royal Netherlands Academy of Arts and Sciences. The Index Fungorum is supported by a community of scientific experts and a partnership of three custodian organisations: CABI, CBS and Landcare Research. Its Intellectual Property Rights (IPR) resides with the contributors and custodians. In these cases the databases are a valuable resource for their host organisation and staff who take a lead role in managing and contributing to the database as part of their employment. Other examples include ERMS (Costello, 2000; Costello et al., 2001), Fauna Europaea (de Jong, 2011), and Euro+Med PlantBase (Euro+Med, 2006), together forming the Pan-European Species Directories Infrastructure (PESI, Costello and de Jong, 2010a; de Jong et al., 2010). However, the IPR of the content of these databases is held by the Society for the Management of Electronic Biodiversity Data (SMEBD) (Costello and de Jong, 2010b). ERMS, initially hosted by a university, is now hosted by the Data Centre of the Flanders Marine Institute (VLIZ) in a similar arrangement to the examples above. Being hosted by VLIZ has enabled it to grow to become WoRMS (Costello and Appeltans, 2008; Appeltans et al., 2012a).

In other cases, the host institution may have been selected on the basis of funding obtained by their staff, but a long-term commitment by that institution, such as if those staff leave or retire from the organisation, may not necessarily have been made. Thus there are many small one-person operated databases that struggle to develop and may become orphaned (Table 1, Box 3). This situation is particularly an issue for universities when academics retire. In the case of CephBase, the world database on cephalopods (octopus, squid and their relatives), disputes over ownership led to its demise (Wood et al., 2001).

Box 3 The need for and benefits of shared database infrastructures.

Data producers and owners may be individuals, research groups, institutes and larger organisations. This approach can end up with a multitude of data owners who also keep the data (physically) with local servers. This is not an efficient solution, and comparable with the old times when people kept their money at home, rather than trusting a banking system with safe deposits and additional services for transactions. A collaborative data infrastructure could provide secure persistent storage, data identifiers, authenticity and workflow support for accessing and mining data (Los and Wood, 2011; Boyle, 2013). The implications of this approach is a move to a data infrastructure as a service, rather than a confused cottage industry or competitive silos of separate and competing databases with serious problems of connectivity and interoperability missing the context and the wider good. Such a collaborative data infrastructure would allow for an environment where ''data themselves are the infrastructure'' (Bachem et al., 2010). This assumes that produced data are easily deposited in an infrastructure remotely controlled and curated by the data producer. As it is probably easiest to control and transfer your money remotely through a bank, data producers and owners should be facilitated with hard- and soft-ware to provide data security and mobility (see Mordecai et al., 2010 for a real world example in Ornithology). This system needs to evolve as new data materialise and users needs become more demanding (Boyle, 2013).

Another benefit of collaborative databases is synergism (Huettman and Meyers, 2009). The sum of their data may reveal insights not obvious from the sub-datasets. For example, it was a surprise that the rate of discovery of new marine species to science in Europe was linear when this is the best studied region of the world (Costello and Wilson, 2011). But a greater surprise was when it was found that this was in the context of several times more people naming new species in recent decades than ever before (Costello et al., 2012b). Thus the rate of discovery was being maintained by more people, suggesting that it was getting relatively harder to discover new species (Costello et al., 2013c). Without such a global database such trends would have been dismissed as peculiar to particular taxa. Collaborative infrastructures also provide the opportunity for large scale collaboration between the contributors to the database (e.g. 122 authors in Appeltans et al. (2012b)).

3. Money matters

Resources to maintain and develop online biodiversity databases may be provided through contributors (and their employers') time, and/or host institution administrative, technical, scientific staff and infrastructure support. However, additional resources may be required to fill gaps in content, pay for time of external experts and contributors, improvements to or re-investments in the infrastructure, and engagement with the wider community through correspondence and attending scientific meetings. Research is also necessary to develop the software tools, and test and demonstrate the ability of the data to produce results of value to science and/or policy regarding biodiversity. A recent full-cost estimate of maintaining a resource in Germany of around 3.8 million high-resolution herbarium images on-line was about €80,000 annually – and this did not include work on selecting and validating the content itself (Täschner and Jaspersen, unpublished data). These resources may be provided through special project funding, but a revenue stream independent of project funding is also desirable to help bridge gaps between other funding sources. Long-term plans and stable funding needs to include coordination and communication costs (including staff time and travel to meetings), software and hardware maintenance, associated overheads and stochastic events. As in business enterprises, a diverse funding portfolio is preferable to a reliance on one or two revenue sources.

The funding sources may be grouped into host organisations, funding agencies, individual scientists, and users. For example, ERMS was initiated by a €385,000 research project in 1997 (Costello, 2000), moved to a new host institution in 2000 which had €250,000 of projects that built upon it, received six small grants from projects funded by EU and USA sources totalling €110,000, and then €300,000 and €400,000 projects in 2004 and 2008 respectively to develop the content, editorial board, and infrastructure further. It further benefitted from support through focused initiatives like the Register of Antarctic Marine Species that supported itself by government funding (SCAR-MarBIN and AntaBIF projects). Time contributed by editors to date has been worth an equal

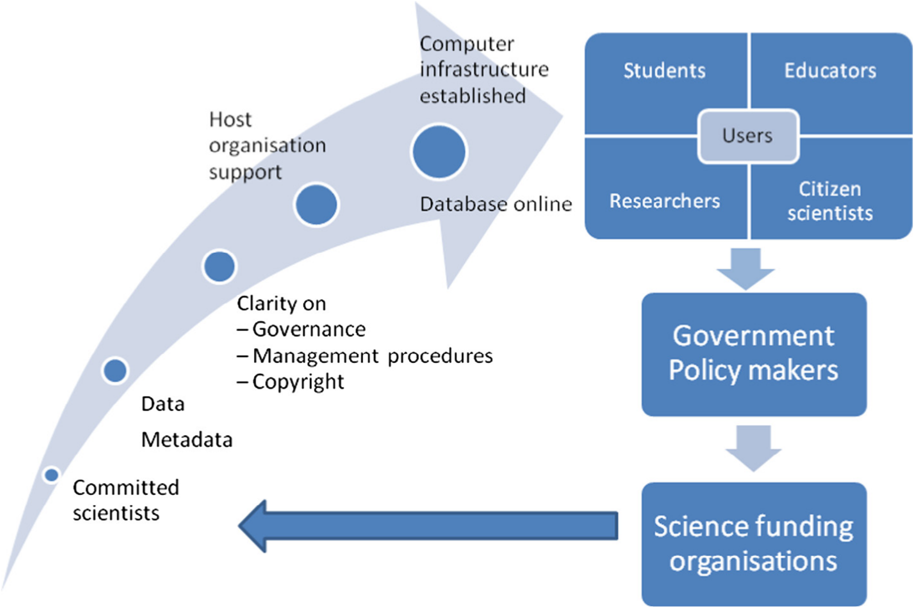

Fig. 1. The ingredients for the establishment of a scholarly and sustainable biodiversity database.

Table 1
Examples of how a database may address some common problems in sustaining a taxonomic biodiversity database based on the authors experience. VLIZ = Flanders Marine Institute; WoRMS = World Register of Marine Species; CBS = Centraalbureau voor Schimmelculturen (Fungal Biodiversity Centre); BGBM = Botanic Garden and Botanical Museum Berlin-Dahlem; EoL = Encyclopedia of Life; SMEBD = Society for the Management of Biodiversity Data; CETAF = Consortium of European Taxonomic Facilities; GBIF = Global Biodiversity Information Facility.

Problem	Solution	Example
How to keep database infrastructure (hard- & software) updated	Find a host institution with a mandate for data management that can maintain a system regardless of whether special project funding exists	VLIZ hosts WoRMS, which now incorporates (rescued) previously isolated (orphaned) databases CBS hosts MycoBank BGBM host Euro + Med PlantBase
How keep databases content updated	Motivate individual scientists to edit it	SMEBD members edit WoRMS and associated marine databases
No clear leadership to maintain databases	Find a professional society with appropriate mandate	SMEBD members elect Council to provide this leadership, primarily by delegation to committees and host institutions to take responsibility particular databases
Need incentives for experts to develop content	Make being a database editor prestigious by being part of a larger database and personally invited by peers as being the best person to do the task	Easier to find editors for a global (WoRMS) than regional (ERMS) database as the former is more prestigious and influential
	Facilitate editor recognition through highly visible conventional citations	Citations placed at foot of web pages and at database level
	Editors keep control of data through a scientific society or formal agreement	SMEBD holds IPR for ERMS, WoRMS and other databases
	Provide tools that aid editors work and research (e.g. checklists, monograph content, images, etc.)	Tools for automatically matching lists of species names to synonyms, to upload images, literature, etc. enable resource to be extended for editors and other users
	Report on usage	Web use is tracked (e.g. unique visitors, web page, downloads)
	Get formal recognition from scientific societies and institutions (CETAF, GBIF) of importance of the database for science	CETAF 2004 statement regarding European taxonomic databases
	Being an editor increases opportunities for getting funding through grants and contracts obtained by SMEBD community	WoRMS editors received grants obtained from European Commission research projects, GBIF and Census of Marine Life
	Involve editors in conventional publication opportunities	Co-authored publications, special issues or collections of papers in international journals
Competing databases in the same domain	Exchange content between resources so experts only need to do their work once	FishBase obtains taxonomy and nomenclature from the Catalog of Fishes GBIF, EoL, Species 2000, get marine species data from WoRMS
Too many small databases run by one or few people.	Consolidate content into one database, and coordinate people through a society, committee, or partnership	WoRMS integrates global, regional and thematic databases

amount (ca. €1.4 million). The incremental extensions of the content resulted in a new goal to produce WoRMS, which doubled the size of its editorial board while making it a more prestigious and valuable resource (Costello et al., 2013a) (Table 1). Between projects, the host institution can maintain the online services and address user needs, while the editors can keep it updated with modest time input. However, collectively this time input is very valuable and may exceed technical support costs.

Most scientific publications are not open-access and require the user or their institution to pay, either by subscription or purchase of the article or book. Over 80% of science journals are funded by subscriptions (Ware and Mabe, 2009). Open-access publishing typically requires the author to pay, often thousands of dollars per article. Some authors may be able to access institutional funds or research grants to cover these costs; but most authors and students cannot. Open-access publishing is the lower-cost model, and it results in far greater accessibility of scientific knowledge and consequent benefits to society (SQW Ltd., 2004; Houghton et al., 2009). In contrast, in the commercial world authors (e.g. journalists) are paid for the articles they write but have to compete for that access and must write for a wide audience so that the publisher can sell the magazine or book. This model may exclude the publication of good science because only the most popular authors would get published.

Online biodiversity databases have the problem that their readers and contributors expect them to be open access, so that the two conventional methods of funding to support the publication costs, namely reader-pays or author-pays, are closed. An additional problem is that such databases' costs are comprised entirely of 'fixed costs' and overheads, which may increase as the resource grows in size and complexity. In contrast, the relative unit cost of e-resources from subscriber and author-pay models become less as the resources grow (SQW Ltd., 2004). Unless this financial dilemma for open-access biodiversity databases is addressed the resource is unlikely to develop and continue. Much like libraries and museums, which are infrastructures to insure the long-term maintenance and accessibility of print publications and specimen collections, databases are information infrastructures that depreciate over time unless they are funded and maintained. A project-based funding model, as in most of the examples cited, is therefore *a priori* suboptimal for the long-term maintenance of the resource. Indeed, Boyle (2013) argues that a good data management system should cost more to maintain and adapt than to be established. This is because it must continually adapt to new data, opportunities provided by new technologies, and user needs over time and not be trapped by an expensive inflexible infrastructure.

Other successful models for supporting databases that are non-profit are subscriber funded and thus not open-access. Aquatic Sciences and Fisheries Abstracts (ASFA) is over 40-years old. Science libraries in many countries compile its content and its abstracts are sold through a commercial publishing company. The funds returned are used to provide grants to content providers. JSTOR and BioOne, both non-profit organisations, archive back issues of journals and re-package them so subscribers can get many journals for less than the cost of subscribing to each journal individually. Their journals get more subscribers even if each receives less revenue than it may from a direct subscription. Related models might see open-access databases provide only limited content open-access, limit downloads, and require heavy users (e.g. research projects, or organisations) to negotiate a special licence that helps cover their maintenance costs.

It seems unlikely that scholarly biodiversity databases could operate in a subscriber pay-model. To readers, it appears that 'adequate' information is available from many sources, in print and online, even un-authored, but that appear scientifically credible. So most readers are unlikely to pay for special access to a particular online resource, and if resources establish pay-per-use charges they will lose potential readers. Instead, readers would turn to other websites for information, even if these are less comprehensive, permanent, authoritative (i.e. authored by known experts) nor peer-reviewed. However, this does not exclude the possibility that some institutions, libraries, government agencies, and companies might pay sufficiently high subscriptions to support a resource of exceptional quality, perhaps for greater visibility, user services and privileges. Thus most journal abstracting services are subscriber-pay databases even though abstracts of most publications can now be found by internet search engines. Examples of subscriber-pay biodiversity databases include Zoological Record (now part of the Thomson Index of Organisms) and the Index of Fungi.

WoRMS licenses complete copies of its databases upon request to organisations and individuals at no cost. However, WoRMS incurs costs associated with its organisation (SMEBD) (e.g. annual financial and governance audit, administrative support, travel for committee members to represent WoRMS) and has content gaps that have not received project funding. In a recent survey, users of WoRMS were polled about making financial contributions. All users that responded felt this to be a reasonable request and suggested a charge of two to three thousand dollars or Euros a year was reasonable, equivalent to a journal subscription or paying for open-access publication of one paper; several have since voluntarily contributed this amount. Other users suggested they could more easily sponsor meetings or targeted services (e.g. filling a particular gap in data content). Thus applications for the database now include a request for a contribution, but this is waived where the user cannot pay. Whether these volunteered contributions will be sufficient remains to be seen, but a similar approach is bearing fruit for the open-access journal *PLoS ONE.*

There is overwhelming support amongst international organisations and publishers of scientific literature, for policies that data should be made publicly available, at least to enable reproducibility of research studies (Costello, 2009a; Huettmann, 2010). Metadata policies are in full support of this effort (Huettmann, 2005, 2009; Zuckerberg et al., 2011). The European Commission (2011) has promoted a policy of Open Data for public information because of its economic benefits. The same principle of open data for scientific information can similarly benefit society by enabling innovative use of the data that may not have been predicted by its creators (Costello, 2009a; Huettmann, 2011). While enforcement of data publication policies by funding agencies and publishers is variable, the increased availability of professional data centres to store and re-distribute data while citing its source is changing the science landscape. Soon, authors who do not release data for independent analysis will be criticised by the scientific community, including funders, readers, referees of papers, and editors (Huettmann, 2007b; Costello, 2009a; Huettmann et al., 2011, for reviews). This movement will increase government funding for data centres that in turn may help authors organise and publish their data. Such centres may include national governmental, institutional and thematic data centres, the ICSU (International Council for Science) World Data System, the IOC/IODE network of Oceanographic Data Centres and Associate Data Units (including OBIS), GBIF, data infrastructures like LifeWatch and European Marine Observation Data Network (EMODNET) in Europe, and others contributing to the Group on Earth Observations – Biodiversity Observation Network (GEO BON). However, not all online biodiversity resources may come within the scope of this system and they will require special government, institutional and/or community support.

4. Intellectual property

Inherent in managing a scientific-purpose database are issues of ownership, governance, and Intellectual Property Rights (IPR)

(Fig. 1). All parties involved must be clear on who owns the database content, database software, hardware, and IPR, and agree on the policy for its development, metadata and redistribution (Van den Eynden et al., 2011; Hagedorn et al., 2011; Zuckerberg et al., 2011). These issues should be documented and transparent to all parties. Unfortunately this is often not the case.

To clarify who holds the copyright, individual contributors and/or their employers should sign an agreement with each other (i.e. a partnership or consortium agreement) or another body to determine who will hold the copyright and thus license the use of the collective database. Such copyright agreements are analogous to when authors sign copyright clearance forms when publishing in scientific journals. This other body could be a learned society, or the host institution of the database (i.e. its publisher). It may be desirable to incorporate a body to own the content and/or entire database. Being a legal entity is a requirement of many funding organisations, although their host or other institutes or organisations may act in this capacity. Examples of organisations established to manage biodiversity data on behalf of their contributors include Species 2000 and the Society for the Management of Electronic Biodiversity Data (SMEBD), both established as limited companies (Costello, 2000) (Table 1). Incorporation may also be desirable for limiting any liability of the contributors should people misuse the information, and for financial liability should the database have debts or other claims made against it (e.g. an employment dispute). It also clarifies which body is responsible for the distribution of the database and taking action should it be plagiarised, misrepresented or its conditions of use be contravened. For example, FishBase is governed by a consortium agreement but has now, in addition, formed a legal entity to employ the programmers and encoders and facilitate participation in projects. A disadvantage in setting up a separate organisation, whether a scientific society, charitable trust, foundation, and/or a limited company, is that its operation requires a secretariat, financial accounting and audits. However, this management discipline can also promote good practice for the enterprise. These costs must be recovered by some revenue, such as donations from users, subscriptions to publications, sales, royalties, or surpluses from hosting conferences. To a large extent the management of these issues is simplified when one institution takes ownership and responsibility for a database. However, should that institution's priorities and/or staff change, or it have a funding crisis, that the database could be compromised. Thus partnerships between individuals and/or institutions aid sustainability.

Perhaps most importantly, legal incorporation hands over governance to a well-defined body where the individual scientists and organisations involved may change over time without compromising the continuity of the database and its management. This can provide clarity and transparency in how the resource is managed, and allow for the contributors to have ownership and control the resource's development by democratically electing the organisations' directors and officers. It may involve the founding scientists relinquishing control over their creation to this organisation. If they do not do this it is unlikely that other scientists will be motivated to contribute when the fruits of their efforts appear to benefit the reputations of other individuals at the expense of the actual contributors (so called 'free riding'). Thus the scientific community takes responsibility for both the quality and comprehensiveness of the database, and its continuation. It should be noted here that individuals who are privileged as employees of the government's 'public service', are supposed to work for the public good (even if not to their individual gain). However, these initiatives still need their champions, whether the founder or their successor, especially in rapidly emerging biodiversity infrastructures that need to adapt to new technologies.

Editors and authors of the database content can be selected by their peers as the best-qualified and available experts (Table 1). Scientists who are users of the resource should provide feedback to improve it. These roles are not fulfilled automatically but require proactive attention by leaders within the science community. A scientific society could endorse leadership roles, such as being committee members and chairs. Examples include the International Mycological Society that oversees MycoBank, and SMEBD that oversees several taxonomic databases. Parallels may be drawn with many scientific societies that engage commercial, institutional, or government organisations to publish their journals and/or host their conferences. However, in almost all these cases the journals are not open-access and are run on a subscription based model that returns a surplus (or profit). In such cases, a close partnership may develop between the society and the publisher. In the case of SMEBD databases, the content is copyrighted to SMEBD, but the software and hardware infrastructure is the property of the database's host institution. As part of the Consortium of European Taxonomic Facilities (CETAF), museums, herbariums and collections of microorganisms have agreed that it is part of the work of their employees to contribute to online open-access databases, in particular ERMS, Fauna Europaea and Euro+Med PlantBase (CETAF, 2004) (Table 1). Other scientific societies should consider endorsing selected online biodiversity resources to give their members and the wider community an indication of the scholarly value of such resources, and encourage experts to develop them further.

5. Succession planning

A challenge in establishing biodiversity databases is that they are often little used until they reach some critical size and reputation. For example, over the past decade the Global Biodiversity Information Facility (GBIF) and its data providers (e.g. aggregator networks like VertNet, MANIS, ORNIS, OBIS or the hundreds of institutions involved) have published hundreds of millions of distribution records of hundreds of thousands of species (Costello et al., 2013b). However, the unique scientific insights possible from such massive global databases are only now beginning to emerge (Huettmann, 2007c; Appeltans et al. (2012b); Costello et al., 2012a,b, 2013a,c; Nemitz et al., 2012). Examples of the use of databases should be planned to demonstrate their value and potential; this may require close collaboration between data managers and researchers. FishBase is a good example of how its development mutually benefited from research conducted by its scientific leaders (Froese and Pauly, 2000).

A typical taxonomic database may have been developed by one, or a few scientists who work for a university, museum, NGO or other science organisation (Table 2). It may also involve self-employed and retired scientists, citizen scientists, and graduate students (Fig. 1). It may have been created initially to satisfy the needs of a funding agency for research results to be made publicly available; and its creators may or may not wish to continue to develop it. The scientist(s) may be very successful in winning regular funding to develop the database over some years, and perhaps a decade or more. When the project funding ends, the resource will become out-of-date and may go off-line when the host institute reorganises its web site, or the scientists move positions or retire. For these reasons, there should be a succession plan for the resource to be continued. This may involve transfer of the resource to a new host institute, mirroring the database, deposition of the database and web interface into a secure data-centre for documentation and archival, integration into another database that has long-term support, additional project funding, finding a sponsor, transferring control to other scientists, and any combinations of these options (Table 2). Having several collaborators involved of different ages will facilitate a smooth transition to new leadership. We propose prerequisites for ensuring that it lives on:

1. The resource must
   a. Have a significant amount of unique content.
   b. Be regarded by scientific peers as of the best available quality.
   c. Have scientists willing to spend time maintaining or developing its content.
   d. Have a good track record in its development.
   e. Contain a clear description of its data fields and their relationships (metadata) and data management practices.
   f. Be clear on its Intellectual Property Rights (IPR), copyright and ownership; i.e. who authorises licences for the use of the resource; and thus on what conditions users may use the resource, such as one of the Creative Commons licences or a similar statement.
2. If the above is the case the resource should seek recognition by
   a. Having its value to science endorsed by large scientific and governmental organisations.
   b. Having a large user community, especially amongst scientists, government organisations, and appropriate industry, but also public (e.g. student, teacher) use.
   c. Documenting the use of the resource, such as pages viewed and content downloaded over time, citation in scientific papers, use as a standard reference, links to it from other websites, and uses in research, management and/or policy arenas.

If these conditions are satisfied then it truly has become part of the science infrastructure. Other scientists or organisations should feel proud to continue it. The factors that motivate scientists and institutions to support such resources are similar, and include its prestigious nature, direct benefits to their work, and relevance to their career advancement or organisation's mission. These characteristics and recognition will facilitate the finding of an institution to host the database. A larger and more widely used resource will be easier to obtain funding for because (a) it will be more prestigious and useful for an organisation to sponsor or host it, (b) it will be more attractive to scientists to be its editors or authors, and (c) it will have more potential funding sources, perhaps globally, including users and project funding. However, although large, a database may have many entry pages that provide contributors with a local identity. WoRMS for instance has 100 global, 12 regional, and 4 thematic overlapping sub-databases fully integrated into the same core database with a common taxonomy (Costello et al., 2013a). Having the database management shared by several scientists and organisations, further reduces the risk of crisis due to changing personnel, funding losses, and/or changing institutional support.

Frequent engagement with users is desirable to ensure their needs can be planned for. This will ensure that the resource is closely aligned with key scientific and policy requirements and evolves as required. This will involve communication by correspondence and interactions at scientific meetings. Special workshops and web-based tools and services can also aid user engagement. The resource will need to provide a service that is unique in terms of quality and/or comprehensiveness as compared to alternatives. For example, GenBank is now an integral part of the world science e-infrastructure with (partly competing) host institutions in three different countries, and a large global network of scientists who use it in their research. This has been aided by the large financial resources for human, animal, plant and microbial genetic sciences, and the fact that genetic data is more amenable to data management than text based information. Other examples of well-established databases focused on biodiversity content show they are all significant in size and have a large international user community. Other species-based biodiversity databases should consider how they could achieve such critical-mass of users and consequent interest from national and other funding sources.

In addition to efforts to make resources useful, potential data providers should also listen to the demands of user communities. For example, the Conference of the Parties (COP) of the United Nations Convention on Biological Diversity adopted the Global Strategy for Plant Conservation ("GSPC"), calling for the development of a global checklist of plant taxa – this request was headed by the institutions coming together to create The Plant List (2010). The COP has now called for a consolidated update of the GSPC 2011–2020 (COP, 2010) and its target 1 calls for the achievement of 'An online Flora of all known plants' by 2020 (COP, 2010). The implementation has been discussed and was endorsed by the Convention's Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA, 2012), which clearly calls for cohesion of the existing flora and checklist databases in botany to achieve this aim. Efforts are under way under the auspices of the Global Partnership for Plant Conservation to organise this endeavour, which hopefully will bring funding to primary data providers and their databases.

6. Management options

Should the database be recognised by other scientists and organisations as being sufficiently unique, large, and authoritative, the following management options are then available for its succession (Table 2):

1. One or more organisations agree to host the database from within their own budget.
2. One or more scientists agree to take responsibility for its content, quality and development on their own time.
3. Other scientists continue to find funding on a project by project basis.

Table 2
An ideal scenario in the development of a biodiversity database.

1. A gap is identified in the availability of a scholarly, expert validated, online resource or database. This may be added to an existing resource, or be a new one
2. Proponents communicate plans to related initiatives to avoid duplication of effort, and ensure the new resource will add to what is available to users
3. A leader and team of experts are willing to commit time to the resource development over several years
4. The proponents involve biodiversity informatics experts in the design of the database so it conforms to existing standards and best practice
5. A partnership of organisations agrees to host the database infrastructure and make it accessible online under the management of the proponents for at least five years. This may be the employer of one or more of the proponents
6. The proponents agree on issues of database governance, management procedures, and copyright (e.g. what legal entity holds copyright, copyright transfer agreements)
7. The proponents obtain funding to develop the database rapidly, including their time and/or hiring data entry staff, so it becomes significant in quantity and quality
8. The database develops a regular user community, and is so recognised by external organisations. This is simplest if the database is added onto an existing database rather than stand-alone
9. Experts continue to update and expand the database, including new experts replacing others as they retire
10. The value of the database to the host organisation(s) is such that it includes its maintenance in its regular operational funding. Alternatively, the proponents may find a host organisation that is able to provide longer-term support

4. A sponsor provides annual funding, or an endowment that provides annual funding.
5. Users pay subscription fees for operational costs, which may include hardware, informatics, and personnel time.
6. The resource raises funds through donations, advertising, publications, CD sales, or other products.
7. Funds are raised through special services built on the resources (e.g. data analysis, reports).
8. Additional content and services are available to users who pay a subscription fee.
9. Mixture of above.

If a resource is entirely the responsibility of one organisation or institution this exposes it to risk should that institute have a change in policy, funding crisis, or changes in key personnel who may champion it. A preferable model is for the database to be the responsibility of a consortium, partnership and several scientists (Hardisty and Roberts, 2013). The database could still be hosted by one organisation but with its development co-managed with others.

It may be useful to distinguish the resource into four components, namely: (1) overall management and governance (including network coordination, assisting experts, responding to enquiries, engagement with users); (2) the expert community who contribute, validate and quality control its content; (3) the software, including the relational database (storage or information system), web interface and supporting technical services to experts and users; and (4) the hardware, including server maintenance and connectivity to the internet, archiving, back-ups, 24/7 international online access, and response time. These components may be managed by different people and have different funding streams. If the resource is a stand-alone facility, with its own hardware, software, information-technology support and scientific staff, it will have a significant budget, probably in hundreds of thousands of euro or dollar per year. For example, FishBase has an annual budget of about €300,000 per year. However, most taxonomic databases are modest in their demands for resources and may be more cost-effectively maintained if they are integrated into larger information systems. If this is planned at an early stage, and the database follows common standards, the difficulties in extracting data due to idiosyncratic formats will be minimised.

A shared resource management and infrastructure has more advantages than disadvantages (Box 3). A management team enables continuity when personnel change and is good succession planning. A shared technical infrastructure with similar databases can save costs in maintenance of hardware and software, data archiving, and system management time. An infrastructure shared with other databases and scientists, will be larger and of greater significance for data management, and thus less likely to be abandoned in a funding crisis, and more likely to get support from science organisations and users. However, such a shared infrastructure will probably require compromise and patience between the people involved. They may also have less flexibility in terms of policy and database development then if they had complete control. However, these accommodations are in the best interest in continuing the database. Thus funders of databases should require proponents to clarify issues of intellectual property, governance, management and how it will be sustained beyond the term of the project funding. For example, it can be unclear what legal entity owns and has copyright on a database that has had contributions from many people from different organisations and countries.

7. Recommendations

The participants involved in sustaining biodiversity databases may be classified as their initial creators, current managers, funding agencies, and the wider scientific community (Fig. 1). We suggest that:

- The creators of online biodiversity databases should consider whether they have the time and resources to make their database of such unique significance that other scientists and/or institutions would continue it. If this is not certain then they may be wise to engage these partners in the development of the resource from an early stage so they feel responsible for it (Table 2).
- Managers of existing biodiversity databases should reflect on how their resource addresses the factors here identified as being important for sustainability, including the extent, scope, quality and uniqueness of database content; track record of development; support from individual scientists; support from institutions, and clarity of IPR, and how well they are maintained and delivered.
- Science funders should focus on the development of scholarly databases with primary, expert-validated content, rather than secondary databases that harvest this content.
- Science funders should require proposals to explain how a project will deal with issues of copyright, data ownership, governance, management procedures, succession planning, and long-term infrastructure support.
- The science community needs to appreciate and support the efforts of scientists in contributing to open-access databases, including by citing these resources in the Reference lists of publications that use them. Tenure and promotion criteria should recognise such publications in individual performance measures.

8. Conclusions

The ideal approach to sustain biodiversity databases is for them to (a) become integrated into larger databases with a consequently larger user community and pool of funding opportunities, and (b) be owned and curated by a collaborative partnership including a science organisation, society, or institution with a suitable mandate. In this regard, bigger is better because the resource will have more content, more potential and more known uses and users of its content, more contributors, be more prestigious to contribute to, and have more funding options. The organisational model should be designed to ensure sufficient resources for its development, in terms of both money and people's time. While developing in this way, it is important to maintain the collegiality and team spirit that is often key to the success of such initiatives. This may be achieved through good governance, formal agreements, including transparency of management, democracy and meritocracy, fairness, engagement and proactive communication with past, present and potential contributors. Of course other models may also work and the ultimate measure of success is their longevity.

Acknowledgements

This paper acknowledges the leading role that the recently deceased Frank Bisby played in the establishment of Species 2000, the European Register of Marine Species, and the Society for the Management of Electronic Biodiversity Data. His ideas, discussions, networking, and hard-work in obtaining funding, have inspired and developed online taxonomic databases (Heywood, 2011).

We also thank Peter Davis (University of Auckland), and the NCEAS, IPY and U.S. GAP projects for helpful discussion and experience. This study was supported by the Pan European Species-directories Infrastructure (PESI) project funded by the European Commission 7th Framework Programme for research.

References

Appeltans, W., Bouchet, P., Boxshall, G.A., Fauchald, K., Gordon, D.P., Hoeksema, B.W., Poore, G.C.B., van Soest, R.W.M., Stöhr, S., Walter, T.C., Costello, M.J. (Eds.), 2012a. World Register of Marine Species. <www.marinespecies.org> (accessed 15.02.12).

Appeltans, W., Ahyong, S.T., Anderson, G., Angel, M.V., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., Berta, A., Błażewicz-Paszkowycz, M., Bock, P., Boxshall, G., Boyko, C.B., Brandão, S.N., Bray, R.A., Bruce, N.L., Cairns, S.D., Chan, T.Y., Chan, L., Collins, A.G., Cribb, T., Curini-Galletti, M., Dahdouh-Guebas, F., Davie, P.J.F., Dawson, M.N., De Clerck, O., Decock, W., De Grave, S., de Voogd, N.J., Domning, D.P., Emig, C.C., Erséus, C., Eschmeyer, W., Fauchald, K., Fautin, D.G., Feist, S.W., Fransen, C.H.J.M., Furuya, H., Garcia-Alvarez, O., Gerken, S., Gibson, D., Gittenberger, A., Gofas, S., Gómez-Daglio, L., Gordon, D.P., Guiry, M.D., Hoeksema, B.W., Hopcroft, R., Jaume, D., Kirk, P., Koedam, N., Koenemann, S., Kolb, J.B., Kristensen, R.M., Kroh, A., Lambert, G., Lazarus, D.B., Lemaitre, R., Longshaw, M., Lowry, J., Macpherson, E., Madin, L.P., Mah, C., Mapstone, G., McLaughlin, P., Meland, K.L., Messing, C.G., Mills, C.E., Molodtsova, T.N., Mooi, R., Neuhaus, B., Ng, P.K.L., Nielsen, C., Norenburg, J., Opresko, D.M., Osawa, M., Paulay, G., Perrin, W., Pilger, J.F., Poore, G.C.B., Pugh, P., Read, G.B., Reimer, J.D., Rius, M., Rocha, R.M., Rosenberg, G., Saiz-Salinas, J.I., Scarabino, V., Schierwater, B., Schmidt-Rhaesa, A., Schnabel, K.E., Schotte, M., Schuchert, P., Schwabe, E., Segers, H., Self-Sullivan, C., Shenkar, N., Siegel, V., Sterrer, W., Stöhr, S., Swalla, B., Tasker, M.L., Thuesen, E.V., Timm, T., Todaro, A., Turon, X., Tyler, S., Uetz, P., van der Land, J., van Ofwegen, L.P., van Soest, R.W.M., Vanaverbeke, J., Vanhoorne, B., Walker-Smith, G., Walter, T.C., Warren, A., Williams, G., Wilson, S.P., Hernandez, F., Mees, J., Costello, M.J., 2012b. The magnitude of global marine species diversity. Current Biology 22, 1–14.

Bachem, A., Wittenburg, P., Giaretta, D., Romary, L., Lopez, D.R., Andersson, T., Genova, F., Best, C., Los, W., van den Stompel, H., Wood, J., Marinucci, M., Vigen, J., 2010. Riding the Wave – How Europe can Gain from the Rising Tide of Scientific Data. High level Expert Group on Scientific Data. European Commission. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf> (accessed 03.03.13).

Balian, E.V., Lévêque, C., Segers, H., Martens, K. (Eds.), 2008. Freshwater Animal Diversity Assessment. Developments in Hydrobiology 198 (reprinted from Hydrobiologia volume 595), 637pp.

Bisby, F., Roskov, Y., Culham, A., Orrell, T., Nicolson, D., Paglinawan, L., Bailly, N., Appeltans, W., Kirk, P., Bourgoin, T., Baillargeon, G., Ouvrard, D. (Eds.), 2012. Species 2000 & ITIS Catalogue of Life, 2012 Annual Checklist. DVD, Species 2000, Reading, UK.

Boyle, J., 2013. Biology must develop its own big-data systems. Nature 499, 7.

Brewer, S., Jackson, S.T., Williams, J.W., 2012. Palaeoecoinformatics: applying geohistorical data to ecological questions. Trends Ecol. Evol. 27, 104–112.

Chavan, V., Penev, L., 2011. The data paper: a mechanism to incentivize data publishing in biodiversity science. BMC Bioinform. 12 (Suppl. 15), S2. http://dx.doi.org/10.1186/1471-2105-12-S15-S2.

Consortium of European Taxonomic Facilities, 2004. CETAF Expresses its Support for European Directories of Species Names. <www.cetaf.org and www.smebd.eu> (accessed 27.02.13).

COP, 2010. Decision Adopted by the Conference of the Parties to the Convention on Biological Diversity at its tenth Meeting. X/17. Consolidated update of the Global Strategy for Plant Conservation 2011–2020. <http://www.cbd.int/doc/decisions/COP-10/cop-10-dec-17-en.pdf> (accessed 29.07.13).

Costello, M.J., 2000. Developing species information systems: the European Register of Marine Species. Oceanography 13 (3), 48–55.

Costello, M.J., 2009a. Motivation of online data publication. Bioscience 59 (5), 418–427.

Costello, M.J., 2009b. Sustainability of online taxonomic information systems. In: Grant, F., de Jong, Y., Kirschner, J., Petrik, P., Segers, H., Sharman, M., Tillier, S., Watt, A., Young, J. (Eds.), Strategies in Taxonomy: Research in a Changing World. Report of an e-conference, pp. 54–55.

Costello, M.J., Appeltans, W., 2008. Taxonomic editors plan a World Register of Marine Species (WoRMS). MarBEF Newsletter No. 8, pp. 36–38.

Costello, M., de Jong, Y., 2010a. Developing taxonomically authoritative online species databases in Europe: the PESI project. EDIT Newsletter No. 19, pp. 11–12.

Costello, M., de Jong, Y., 2010b. SMEBD: a society aiming to make biodiversity data available for the benefit of science and environmental management. EDIT Newsletter No. 19, pp. 6–7.

Costello, M.J., Vanden Berghe, E., 2006. "Ocean Biodiversity Informatics" enabling a new era in marine biology research and management. Mar. Ecol. Prog. Ser. 316, 203–214.

Costello, M.J., Wilson, S.P., 2011. Predicting the number of known and unknown species in European seas using rates of description. Glob. Ecol. Biogeogr. 20, 319–330.

Costello, M.J., Emblow, C., White, R. (Eds.), 2001. European Register of Marine Species. A check-list of marine species in Europe and a bibliography of guides to their identification. Patrimoines Nat. 50, 1–463.

Costello, M.J., Vanden Berghe, E., Browman, H.I. (Eds.), 2006. Introduction: ocean biodiversity informatics. Mar. Ecol. Prog. Ser. 316, 201–202.

Costello, M.J., Stocks, K., Zhang, Y., Grassle, J.F., Fautin, D.G., 2007. About the Ocean Biogeographic Information System. <http://hdl.handle.net/2292/5236>.

Costello, M.J., Michener, W.K., Gahegan, M., Zhang, Z.-Q., Bourne, P., Chavan, V., 2012a. Quality Assurance and Intellectual Property Rights in Advancing Biodiversity Data Publications. ver. 1.0. Global Biodiversity Information Facility, Copenhagen, pp. 33.

Costello, M.J., Wilson, S.P., Houlding, B., 2012b. Predicting total global species richness using rates of species description and estimates of taxonomic effort. Syst. Biol. 61 (5), 871–883.

Costello, M.J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D.P., Hoeksema, B.W., Poore, G.C.B., van Soest, R.W.M., Stöhr, S., Walter, T.C., Vanhoorne, B., Decock, W., Appeltans, W., 2013a. Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. PLoS ONE 8 (1), e51629.

Costello, M.J., Michener, W.K., Gahegan, M., Zhang, Z.-Q., Bourne, P., 2013b. Data should be published, cited and peer-reviewed. Trends Ecol. Evol. 28 (8), 454–461.

Costello, M.J., Wilson, S., Houlding, B., 2013c. More taxonomists but a declining catch of species discovered per unit effort. Syst. Biol. 62 (4), 616–624.

Cushman, S., Huettmann, F., 2010. Spatial Complexity, Informatics and Wildlife Conservation. Springer, Tokyo, Japan, 448 p.

de Jong, Y. (Ed.), 2011 Fauna Europaea Version 2.4. Web Service: <http://www.faunaeur.org>.

de Jong, Y., Boumans, L., Kouwenberg, J., Enghoff, H., Boegh, P., Aktaç, N., Yurtsever, S., Hussey, C., Hyam, R., Costello, M., Bourgoin, T., Berendsohn, W.G., von Raab-Straube, E., Güntsch, A., Appeltans, W., Vanhoorne, B., 2010. A Pan-European Species-directories Infrastructure (PESI). In: Atlas of Biodiversity Risk, Pensoft Publishers, 6–7pp.

Drew, J.A., 2011. The role of natural history institutions and bioinformatics in conservation biology. Conserv. Biol. 25, 1250–1252.

Euro+Med, 2006. Euro+Med PlantBase – The Information Resource for Euro-Mediterranean Plant Diversity. Published on the Internet <http://ww2.bgbm.org/EuroPlusMed/>.

European Commission, 2011. Digital Agenda: Commission's Open Data strategy, Questions & Answers, MEMO/11/891. <http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/11/891&format=HTML&aged=0&language=EN&guiLanguage=en> (accessed 31.01.12).

Fautin, D.G., Buddemeier, R.W., 2008. Biogeoinformatics of the Hexacorals. <http://www.kgs.ku.edu/Hexacoral> (accessed 16.07.13).

Froese, R., Pauly, D. (Eds.), 2000. FishBase 2000, Concepts, Design and Data Sources. xvii+344pp. Los Banos, Laguna, Philippines, ICLARM.

Froese, R., Pauly, D. (Eds.), 2013. FishBase. World Wide Web Electronic Publication. Version 04/2013. <www.fishbase.org>.

Gemeinholzer, B., Dröge, G., Zetzsche, H., Haszprunar, G., Klenk, H.-P., Güntsch, A., Berendsohn, W.G., Wägele, J.-W., 2011. The DNA bank network: the start from a German initiative. Biopreserv. Biobank. 9 (1), 51–55.

Graham, C.H., Ferrier, S., Huettmann, F., Moritz, C., Peterson, A.T., 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends Ecol. Evol. 19, 497–503.

Guiry, M.D., Guiry, G.M., 2013. AlgaeBase. World-wide Electronic Publication, National University of Ireland, Galway. <http://www.algaebase.org>.

Hagedorn, G., Mietchen, D., Morris, R.A., Agosti, D., Penev, L., Berendsohn, W.G., Hobern, D., 2011. Creative commons licenses and the non-commercial condition: implications for the re-use of biodiversity information. ZooKeys 150, 127–149. http://dx.doi.org/10.3897/zookeys.150.2189.

Hardisty, A., Roberts, D., 2013. A decadal view of biodiversity informatics: challenges and priorities. BMC Ecol. 13, 16. http://dx.doi.org/10.1186/1472-6785-13-16.

Heywood, V.H., 2011. Professor Frank Ainley Bisby (24 April 1945–25 October 2011). Flora Mediterranean 21, 345–353.

Holetschek, J., Kelbert, P., Müller, A., Ciardelli, P., Güntsch, A., Berendsohn, W.G., 2009. International networking of large amounts of primary biodiversity data. Lecture Notes Inform. (LNI) 154, 552–564.

Houghton, J., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M., Gourlay, A., 2009. Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits. Victoria and Loughborough Universities, UK, 256 pp.

Huettmann, F., 2005. Databases and science-based management in the context of wildlife and habitat: towards a certified ISO standard for objective decision-making for the global community by using the internet. J. Wildlife Manage. 69 (20), 466–472.

Huettmann, F., 2007a. The digital teaching legacy of the International Polar Year (IPY): details of a present to the global village for achieving sustainability. In: Tjoa, M., Wagner, R.R. (Eds.), Proceedings 18th International Workshop on Database and Expert Systems Applications (DEXA), 3–7 September 2007, Regensburg, Germany. IEEE Computer Society, Los Alamitos, CA, pp. 673–677.

Huettmann, F., 2007b. Constraints, suggested solutions and an outlook towards a new digital culture for the oceans and beyond: experiences from five predictive GIS models that contribute to global management, conservation and study of marine wildlife and habitat. In: Vanden Berghe, E. et al. (Eds.), Proceedings of 'Ocean Biodiversity Informatics': An International Conference on Marine Biodiversity Data Management Hamburg, Germany, 29 November–1 December, 2004. IOC Workshop Report, 202, VLIZ Special Publication 37, pp. 49–61. <http://www.vliz.be/vmdcdata/imis2/imis.php?module=ref&refid=10720>.

Huettmann, F., 2007c. Modern adaptive management: adding digital opportunities towards a sustainable world with new values. Forum Public Policy: Clim. Change Sustain. Dev. 3, 337–342.

Huettmann, F., 2009. The global need for, and appreciation of, high-quality metadata in biodiversity work. In: Spehn, E., Koerner, C. (Eds.), Data Mining for Global Trends in Mountain Biodiversity. CRC Press, Taylor & Francis, pp. 25–28.

Huettmann, F., 2011. Serving the global village through public data sharing as a mandatory paradigm for seabird biologists and managers: why, what, how, and a call for an efficient action plan. Open Ornithol. J. 4, 1–11.

Huettmann, F., Meyers, Z., 2009. The Data Sharing Toolbox. The Wildlife Professional. The Wildlife Society, Fall, pp. 30–32.

Huettmann, F., Artukhin, Yu, Gilg, O., Humphries, G., 2011. Predictions of 27 Arctic pelagic seabird distributions using public environmental variables, assessed with colony data: a first digital IPY and GBIF open access synthesis platform. Mar. Biodivers. 41, 141–179.

IPNI, 2012. The International Plant Names Index. <http://www.ipni.org> (accessed 15.01.12).

Latour, B., 1987. Science in Action: How to Follow Scientists and Engineers through Society. Harvard University Press, Cambridge, MA, 274 pp.

Los, W., Wood, J., 2011. Taming the data wave. Science 331, 1515–1516.

Merali, Z., Giles, J., 2005. Databases in peril. Nature 435, 1010–1011.

Mordecai, R., Laurent, E., Moore-Barnhill, L., Huettmann, F., Miller, D., Sachs, E., Tirpak, J., 2010. A Field Guide to Web Technology. Southeast Partners in Flight (SEPIF). <http://sepif.org/content/view/62/1/>.

Nemitz, D.F., Huettmann, E., Spehn, M., Dickoré, W.B., 2012. Mining the Himalayan uplands plant database for a conservation baseline using the public GMBA webportal. In: Huettmann, F. (Ed.), Protection of the Three Poles, Springer, Tokyo, Japan, pp. 135–158.

Robert, V., Stegehuis, G., Stalpers, J., 2005. The MycoBank Engine and Related Databases. <http://www.mycobank.org> (accessed 15.01.12).

SBSTTA, 2012: Global Strategy for Plant Conservation: World Flora Online by 2020. <http://www.cbd.int/doc/meetings/sbstta/sbstta-16/information/sbstta-16-inf-38-en.pdf> (accessed 29.07.13).

Segers, H., 2007. Annotated checklist of the rotifers (Phylum Rotifera) with notes on nomenclature, taxonomy and distribution. Zootaxa 1564, 1–104.

SQW Ltd., 2004. Costs and Business Models in Scientific Research Publishing. A Report Commissioned for the Wellcome Trust. The Wellcome Trust, London, 24 pp.

The Plant List, 2010. Version 1. <http://www.theplantlist.org/> (accessed 29.07.13).

Thiers, B., 2012. Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff. New York Botanical Garden's Virtual Herbarium. <http://sweetgum.nybg.org/ih> (accessed 15.01.12).

Thomson Reuters, 2012. Data Citation Index. <http://wokinfo.com/products_tools/multidisciplinary/dci/> (accessed 06.07.13).

Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., Horton, l., 2011. Managing and Sharing Data. UK Data Archive, University of Essex, Colchester, 36 pp.

Ware, M., Mabe, M., 2009. The STM Report: An Overview of Scientific and Scholarly Journal Publishing. International Association of Scientific, Technical and Medical Publishers, Oxford, 68 pp.

Wood, J.B., O'Dor, R.K., Piatkowski, U., 2001. Introduction to CephBase www.cephbase.utmb.edu. In: Feoli, E., Nauen, C. (Eds.), Proceedings of the INCO-DEV International Workshop on Information Systems for Policy and Technical Support in Fisheries and Aquaculture, Los Baños, Philippines, 5–7 June 2000, ACP – EU Fisheries Research Report 8 [CephBase formerly at www.cephbase.utmb.edu, now archived at http://nano.nstl.gov.cn/sea/MirrorResources/6448/index.html but frozen].

Zuckerberg, B., Huettmann, F., Friar, J., 2011. Proper data management as a scientific foundation for reliable species distribution modeling. In: Drew, C.A., Wiersma, Y., Huettmann, F. (Eds.), Predictive Modeling in Landscape Ecology, Springer, New York, pp. 45–70 (Chapter 3).