
Biodiversity Informatics, Biota/FAPESP and the future a personal view

<http://www.biotaneotropica.org.br/v1n12/pt/abstract?article+BN00601122001>

Arthur D. Chapman

Received: September 05, 2001

Accepted: September 05, 2001

Chapman, Arthur D., 5 Sargeant Place, Wanniasa, ACT 2903, Australia

arthur@onaustralia.com.au

Arthur.chapman@ea.gov.au

<http://www.erin.gov.au>

Abstract

This paper explores the development of the FAPESP/Biota project from the view of an external adviser. It explores the implementation of advice received and the development of the project through the first three years, and provides some views on future directions for the project.

Keywords - *BIOTA/FAPESP Scientific Advisory Committee, Biodiversity Virtual Institute, State of São Paulo, SinBiota, SinBiota, databases, biodiversity, Metadata and data management, data modeling, relationship to international organizations*

Resumo

Este artigo analisa o desenvolvimento do Programa BIOTA/FAPESP do ponto de vista de um membro do Comitê Internacional de Avaliação (Scientific Advisory Committee). Discute a implementação das sugestões propostas no decorrer do processo anual de avaliação e o desenvolvimento do Programa nos últimos três anos, oferecendo algumas opiniões a respeito dos rumos que o Programa deverá seguir nos próximos anos. Este artigo analisa o Internacional de Avaliação decorrer do processo anual de opiniões a respeito dos rumos

Palavras Chave - *Comitê Internacional de Avaliação (Scientific Advisory Committee) do BIOTA/FAPESP, Instituto Virtual da Biodiversidade, Estado de São Paulo, SinBiota, bancos de dados, gerenciamento de dados e metadados, modelagem, relação com organizações internacionais*

Introduction

The United Nations Conference on Environment and Development (UNCED) meeting in Rio de Janeiro in March 1992 known as Agenda21 stressed the need to bridge the data gap (United Nations 1992). One area seen as a major cause of the ever-widening data gap, was the lack of standardization and accessibility to data. One of the key agreements adopted at Rio was the Convention on Biological Diversity (United Nations 1993), a core premise of which, was the exchange of information and technologies (Article 18, United Nations 1993).

For the past three years, I have had the privilege of being a member of the Scientific Advisory Committee of FAPESP/Biota (the Biodiversity Virtual Institute) for the São Paulo State Foundation for Research Support (FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo). The Biota program has the aim of studying all aspects botanical, zoological and microbial of the Biodiversity of São Paulo. My involvement with Brazil's biodiversity, however, stretches back to the beginnings of the 1990s.

In the late 1980s and early 1990s Geographic Information Systems (GIS) and Relational Database Management Systems (RDBMS) began to become more user friendly and be used more extensively in environmental science and environmental management. The idea of databasing large amounts of biodiversity and environmental data goes back far longer than this, however, and as early as 1974, Australian biologists developed an interchange standard for the exchange of museum and herbarium data (Busby 1979). These standards were developed to facilitate the transfer of information between institutional databases, but the reality of the time was that very few biological institutions in Australia had suitable databases, and indeed very little museum or herbarium data had been databased at this time (Richardson & McKenzie 1991). The standards were not wasted, however, as many institutions began to use the standards in the design of their databases.

Early in the 1990s, a number of biologists around the world saw the information explosion as a challenge and an opportunity. How could existing, and rapidly increasing stores of information be best organized; how could the information be made available in the most useable way; how could new technologies be used to harvest the data and information; and how could these new technologies be utilized to make the information available to users?

In 1989, ERIN (the Environmental Resources Information Network) was established within the Federal Department of the Environment by the Australian Government “... to draw together, upgrade and supplement information on the distribution of endangered species, vegetation types and heritage sites” (Hawke 1989). The ERIN information bases were designed to answer questions crucial to the management and conservation of the Australian environment.

In 1992, around the same time as the Agenda21 meeting in

Rio, meetings of the Taxonomic Databases Working Group (TWG) and the International Organisation for Plant Information (IOPI) were being held in Xalapa, Mexico. Both of these organizations had been set up to develop standards and to promote the databasing and exchange of biodiversity, and especially taxonomic, data around the world. Also at this time, the Mexican Government established the Mexican National Commission for the Knowledge and Use of Biodiversity (Conabio) to “coordinate conservation and research efforts designed to preserve biological resources, to conserve the nation's resources and to generate criteria for sustainable development” (<http://www.conabio.gob.mx/>).

Not long afterwards, Dan Janzen of the University of Pennsylvania, organized for a group of people from Conabio in Mexico, Inbio in Costa Rica, LIPI in Indonesia, and the Kenya National Museums, as well as a number of self-funded people from the USA and the Natural History Museum in the UK to spend two weeks exchanging information and ideas, and examining in detail, the database and GIS structures within the Environmental Resources Information Network in Australia. Although each of these agencies were established in different ways and under different administrative arrangements, this meeting formed the basis of a continuing exchange of information and methodologies and played a major part in shaping how each of these organizations developed through their formative years.

The internet at this time was largely restricted to research institutions and universities, but the development of the Gopher protocols saw the internet move into the wider community. Gopher was essentially a textual exchange format, but biologists quickly began to see the possibilities it presented not only for the exchange of information, but for the exchange of ideas. The Base de Dados Tropical in Campinas, Brazil was one organization that quickly latched on to the possibilities and set up a network of loosely connected technically-interested biodiversity-related institutions in the form of the Biodiversity Information Network, or Bin21 (Canhos *et al.* 1994, Green & Croft 1994).

Further development and expansion of the internet happened with the development of the http protocols and the World Wide Web. These developments provided the impetus and means for the enhanced exchange of biodiversity information. The Convention on Biological Diversity (CBD) recognized these possibilities with formation of a Clearing House Mechanism (CHM) for exchange of information among signatory nations. Unfortunately, despite an early flurry of activity, the Clearing House Mechanism has not been as effective as envisioned by many in its early days. In the mean-time, the expansion of the internet and of improved distributed systems and databases, has meant that the need for a formalized Clearing House Mechanism has declined.

First preliminary meeting Serra Negra 1997

In the lead up to 1997, a small group of dedicated individuals had seen the possibilities of using databases and the internet for conservation studies in Brazil, and had begun work toward setting up a biodiversity database system at the Base de Dados Tropical in Campinas. They, and others, began planning an interim BIOTASP (Biota - São Paulo) and organizing a number of small workshops.

In 1997, a workshop in Serra Negra in central São Paulo examined the use of databases and the internet for the conservation of biodiversity within that State. The workshop, 'Bases para a conservação da biodiversidade do estado de São Paulo', brought together biologists from a range of disciplines, along with a small number of Government officials and select international advisers. Prior to the Serra Negra meeting, the small group of international advisers were invited to briefing meetings in Campinas and invited to present information on their own systems and experiences. Discussions were held on the lessons learnt from these existing systems, and how those lessons could be applied in the State of São Paulo. It was interesting to note that a number of the organizations involved were the same as those that met in Canberra, Australia to examine the Environmental Resources Information Network some five years earlier.

At the Serra Negra meeting, it was fascinating to see biologists of all disciplines pouring over maps and getting excited as they realized that they weren't unique in their thinking and that many benefits could be gained from the exchange of information and in developing consistency in the databasing of biotic and environmental information. The experts exchanged information as they worked collaboratively together to begin determining biodiversity priorities for the State. The sense of cooperation and collaboration was encouraging and refreshing. Fish biologists, algologists, botanists, mammalogists, and others all sharing information and gaining new insights about their own information and where it fits into the broader conservation picture. It became obvious to many that there could be advantages to their own research in being able to exchange information on areas of mutual or common research interest. The sense of cooperation at that meeting led to the development of what we now know of as the FAPESP/Biota program. This could not have happened without the support of the political contingent at the meeting, and without a lot of lobbying by a dedicated group of key biologists and organisers. As one of the advisers, I left the meeting excited and encouraged for the future of the São Paulo State's conservation efforts. I am sure that those more intimately involved, and with a larger investment in the future of the biodiversity of the State, felt even more enthused than I. This is borne out by the hard work put in since to develop what is fast becoming one of the world's leading bioinformatic systems.

A key outcome of these meetings was an agreement on the importance of documenting biodiversity as a step towards management and conservation and in developing common goals and approaches.

In keeping with the aims of the workshop, the discussions and results are available via the internet and can be seen at <http://www.biota.org.br/info/workshop/>.

Following the Serra Negra meeting, the international advisers were asked to write brief articles on their impressions of the meeting and to list ideas that they had for the way forward. It is pleasing to note that many of those suggestions have been followed up and implemented in the development of Biota. Many of the suggestions still apply, and I have grouped them below and commented on their implementation.

Metadata and data management

A number of advisers at the Serra Negra meeting stressed the need for structured metadata to allow better access to São Paulo's vast data resources. As stated by Colwell and others (Colwell 1997), like any country databases in Brazil exist in a variety of formats in many locations and institutions. The only way for users to know what databases are available is to record information about each of them, who is responsible for them (ie the custodian), what access conditions apply to the data they contain, as well as a range of other characteristics of the database. With this information, users are able to determine the fitness of the data for their particular use and determine who they need to contact to access the information therein (Chapman & Busby 1994, Freeman *et al.* 1999).

The development of metadata by any organization should be regarded as a priority. The Biota program organisers were aware of this from the start and as early as the Serra Negra workshop were asking participants to fill out survey forms about existing datasets that participants were responsible for. These have since formed the basis of the SinBiota 'metadatabase', whereby all new data must be accompanied by a metadata record before being included in the SinBiota system.

Consistent and appropriate standards for data management go hand in hand with the metadata and one cannot properly exist without the other. Data management standards need to be flexible and be able to accommodate continuing changes in technology and in the nature of data itself. These standards form a part of the metadata, and supply linking protocols between the various databases described in the metadata.

In addition, update mechanisms need to be developed and recorded as part of the metadata. This is an often neglected part of data management. Negotiations for data need to include procedures for updating the data at regular intervals if appropriate. Also, mechanisms need to be implemented for tracking changes and for archiving data that is no longer relevant and/or is being replaced or updated.

Standard taxonomies

Species do not recognize political boundaries. While it is necessary to work within political units, taxonomic identifications and evaluations require knowledge of the species and related taxa throughout their entire biogeographical ranges. Where possible, centralized

taxonomies, preferably including synonymies, can be used to link disparate databases. This allows for information on the same species to be accessible, even if collected under different names in different projects. It also reduces errors due to mis-spelt names and outdated synonymies.

While it is often difficult to justify funding for activities beyond political boundaries, the scientific importance of pursuing them should be built into all aspects of the program (Potts 1997). For example, where species are widespread, such as with birds, it may be more productive to create field guides for all (or a large part) of Brazil or South America, and later subset into State or regional guides, rather than attempt to create separate guides for each region or State.

The use of a standard taxonomy, or species names list, to link species-related databases (i.e. any database that includes species names as a key) is highly recommended. To date the Biota program has not included a standard names list and have accepted uncritically all names submitted to them in associated databases. The Species 2000 Programme (<http://www.species2000.org/>) is developing dynamic checklists of taxa which include an annualized checklist, available either via the internet or on CD-ROM. Such a checklist could form the basis of a State-wide list. Names in taxonomic groups not yet covered in Species 2000, could be added to the list, gradually building a State-wide checklist. The inclusion of a standard taxonomy where possible, allows for names in different databases that may be synonyms, to be linked. In this way a São Paulo-wide checklist of standard taxonomies can be gradually built up.

Protocols also need to be developed for the handling of unnamed and unidentified taxa that arise in Biota projects. Protocols for handling these need strengthening, to ensure that data remain referenced to organisms in the future.

Data quality and validation

Data quality is an often neglected aspect of data and database management. It is important that databases are constructed with data quality recording in mind. This includes the inclusion of accuracy fields wherever possible, especially for all spatially-related fields such as latitude, longitude and altitude. Often these accuracy fields are included as categorical fields with values of 1 to 5, one end of the scale being more accurate than the other. The use of categories, however, leads to a loss of valuable information. If categories go from 1 for 100 metres, 2 for 1000 metres, 3 for 5000 metres, etc., then a record that is accurate to 2 km has to be recorded as a 3 with the resultant loss of information. It is far better, wherever possible, to record accuracy as raw data, for example, in metres (Chapman 1992, 2000). Data quality and validation need to be considered at an early stage and should be considered at the design stage of any database development or collection strategy (Chapman & Busby 1994).

The use of standard taxonomies to check names for correctness is another form of validation. Only by using valid names, can checklists can be merged and exchanged

between databases and linked to additional information. As mentioned above, collaboration with Species 2000 may be an option for developing checklists of valid taxonomic names.

Data sharing

There is a vast amount of biodiversity information already in electronic form in Brazilian institutions outside the State of São Paulo and in other countries. Working out cooperative arrangements for data-sharing with these institutions can begin with queries to their databases for relevant electronic information, followed by arrangements either to download this information to a central data bank somewhere in São Paulo, or perhaps long-term agreements for online access over the internet. Wherever possible, a distributed system is preferable, whereby data is maintained at the source by the custodians and owners of the data. Alternatively, the data may be stored centrally with the custodians maintaining and updating the information on line. Any successful system will adopt a range of ways of maintaining the data, but if the system is too centralized, data can quickly become outdated and of little value.

Existing non-databased information

There is also a vast amount of information on historic specimen labels and in field notes from specimens deposited in São Paulo institutions such as museums and herbaria. The task of bringing this information into usable electronic form is enormous, but can be scheduled to proceed in parallel with other objectives. For example, type specimens could be made high priority, or groups of specimens of particular conservation interest, etc. There is an advantage in prioritising the databasing of taxonomic groups across the State, as this will lead to a complete coverage of one group at an early stage rather than lots of incomplete parts arising from collections from different data custodians, each with their own priorities.

At present there is interest from a number of large developed nation collection institutions in beginning to database their collections, and many are looking at priorities for that databasing. Developing countries that are embarking on a process such as the FAPESP/Biota program need to inform and influence those institutions as to their priorities. As most collections are stored taxonomically and not geographically, taxonomic groups that are of importance to developing countries need to be identified and publicized. As priorities are developed for the State of São Paulo northern hemisphere collecting institutions may be persuaded to make those groups their priorities as well.

In most cases, it is far cheaper to digitise existing data than to collect new data (Chapman & Busby 1994). It is important therefore, to utilise existing data wherever possible, before attempting the collection of new data.

Data integration

The integration of data is a key to the success of any

conservation operation. Once the data is collected, it needs to be useable through its integration with other data sets. Individual databases can be good for some purposes, but the true value for use in environmental decision making is in their integration with other data through GIS systems, etc. This integration is already becoming a strength of the FAPESP/Biota program.

Gap filling

As existing information is databased, the information can be used to identify gaps in the information - not only geographically, but also environmentally. In this way, future collecting activities can be prioritized to concentrate on filling those gaps. These gaps, when identified, can be used for setting funding priorities and for setting targets for further data collection. It should be recognized that many gaps will take a long time to fill, and those gaps should not be an excuse for a slow down in other parts of the program. The vast data resources being databased as part of the FAPESP/Biota Program will form a basis for these types of analysis.

Tools and standards

Many electronic tools are available for storing, processing, handling, manipulating and disseminating biodiversity information. It is important to make use of existing tools where possible, rather than losing time by devising (and maintaining) new ones. It is also important, not to get tied to proprietary, outdated or non-updateable technology for data storage and dissemination, publication or analysis. Dissemination technologies are rapidly changing as evidenced by the changes occurring on the World Wide Web, and any system must be able to take advantage of any new developments in this area.

As far as is possible, it is preferable to adapt existing standards (global, national, etc.) than develop new ones. There is also an advantage in involving stakeholders in any modifications of existing standards and/or development of new ones as this is more likely to provide a better base for acceptance and compliance by the users.

Modelling and statistics

It has been evident that few Biota projects are using predictive approaches for filling knowledge gaps. Predictive modeling can be very powerful for assessing adequacy of understanding of the environment. For example, predictions of species, distributions, abundances or habitats, made before sampling a new area, may indicate where to concentrate survey effort. In areas as large and as diverse as the State of São Paulo, with scarce and uneven collecting, it is important to use predictive methods to help fill in those gaps. A number of good bioclimatic and environmental modeling tools and systems exist which allow for environmental surrogates to fill in for areas of lacking knowledge (Chapman & Busby 1994). An effort needs to be made, however, to develop the underlying environmental parameters that these models require, so that scientists can then easily use them to begin filling in the

knowledge gaps. To date, few Biota projects have made use of sample-based biodiversity statistics or multivariate approaches, and almost no one has used bioclimatic modeling, or other techniques from the growing range of biodiversity analytical tools. Biota may consider using international experts and workshops to guide the setting up of biodiversity informatics tools and in training São Paulo scientists in their use.

Training of young scientists

Specialized workshops bringing together young scientists to discuss the Biota projects they are working on, began at the second annual review of the Program, in the forests of Intervales in 2000, following recommendations of the Scientific Advisory Committee. The exchange of ideas, and the development of collaborative arrangements between the young scientists was indeed encouraging. It was evident from discussions with many of these young scientists that they saw benefits in not having the senior professors present for much of the symposium, however, it was also evident that there would be further benefit by having some overlap with the Principal Investigators toward the end of the symposia.

The development of training courses on specialized topics of interest across a range of projects may also be worth considering. These could cover such topics as collecting methodologies, analysis and modeling techniques, data storage and presentation, etc. Not only would the training itself be valuable, but the bringing together of junior scientists at these courses would provide additional opportunities for collaboration and interchange.

Education and outreach

There needs to be full interaction at all stages, between providers of information, the data analysts and managers, and end users to ensure that the data being collected and analysed is in a format that will provide the information required by users. This involvement of stakeholders also applies to the development of standards at all levels of the process. Apart from the obvious users of biodiversity data, such as for science, nature conservation, sustainable use, policy development and general ecosystem management, it is important not to forget outputs that may be used by the general public. The development of a good school's program has already been considered by FAPESP/Biota and this needs to be enhanced.

The Biota goal of providing internet access to biodiversity information for the State has already been implemented most effectively. The time is ripe for the Biota program to develop a long-term plan for increasing public awareness and appreciation of the habitats and organisms of the São Paulo State, not only within the State, but also nationally and internationally. Public interest and involvement in the conservation of biodiversity needs to be increased by providing appropriate levels of information, for example, in Visitor's Centers in natural areas, through development of local field guides to groups of organisms (e.g. birds,

butterflies), through natural history programs in elementary and secondary schools, and possibly through the development of television series on the biodiversity of São Paulo. The creation of awareness of biodiversity and its importance, in all levels of society, is essential for its long-term survival and conservation.

Long-term monitoring and the use of common reference sites

The selection of specific sites in specific environments such as Coastal, Mata Atlantica and Cerrado, and where interdisciplinary studies may take place, would provide significant added value to the program. The Biota project has already instituted funding of some common sites and is encouraging interdisciplinary studies where possible. The goal of intensive surveys and inventory throughout all areas of the State is obviously unrealistic, and the overall level of knowledge will prove greater through a well-designed mix of intensive and less intensive studies. By using select sites in these different environments it is likely that a better understanding of each of those environments will ensue, and lead to a better understanding of the overall biology of the State.

The selected common sites could also be used as long-term monitoring sites for detecting change as well as providing references for comparison between sites. The development of a long-term monitoring strategy needs to be considered at an early stage and integrated into new collecting activities and strategies (Chapman 1997).

Mapping of information

Maps showing existing knowledge, regardless of how preliminary, will be an important step in developing a large-scale conservation strategy for the State of São Paulo. They can point to important regions for which knowledge is lacking as well as to areas where immediate conservation actions may be a priority. In this way they can be an extremely valuable tool in priority setting. It is critical, however, that no matter how preliminary the nature of the maps, that they be fully documented. Only in this way can users be confident in making decisions based on them. Maps are an essential tool for presenting information to users, including policy makers and environmental managers, in a way that is easily understood and assimilated.

A well developed GIS unit may help improve the standard and quality of maps across the project and lead to greater standardization between projects. The development of an expanded GIS and analysis unit, in parallel with SinBiota, may be an element that the FAPESP/Biota program will need to consider.

Collections and collections infrastructure

The maintenance of collections and collections infrastructure should not be neglected in the pursuit of databasing. It is important that administrators do not view the maintenance of collections and the databasing of the

information they contain as alternatives. Some administrators believe that the database and maps, are end product and once completed, that the collections themselves become superfluous. But, through time, as the quality of the information associated with the collections (identifications, vouchers, etc.) increases then the quality of the database also increases (Chernoff 1997).

Programs such as FAPESP/Biota generally have an unexpected spin-off, in that they often generate an increased rate of collection. It is important that this be borne in mind in the funding of infrastructure support so collections and infrastructure can be expanded to meet the increased demand.

Prioritization and program balance

With a project like Biota, it soon becomes apparent that one cannot focus on everything, so priorities need to be set - on environments, on particular species (endangered, keystone), on priority areas or "hot spots", etc. These need to be flexible to be able to take advantage of opportunistic data sets, and designed around conservation priorities (Chapman 1997). In the mid 1990s, Australia developed a bioregional approach to conservation (Thackway & Creswell 1995). Although not providing all the answers, the resultant regionalization provides an ongoing basis for conservation priority setting and funding across the country. Since then, a complimentary regionalization has been developed for the marine and coastal areas of Australia (IMCRA Technical Group 1997). The Biota Program may find it valuable in using the increasing information base it is producing to develop a similar bioregional approaches for the State of São Paulo.

At the same time, in order to get a consistent conservation coverage for the State, a good coverage of taxa and environments should not be neglected. In particular, key taxa (indicators, soil microbiology, etc.) need to be incorporated into the coverage of the program. Areas where added emphasis needs to be given, include studies on soil biota other than mites (bacteria, nematodes, earthworms and arthropods); expanding geographic coverage within the SinBiota databases to cater for data from outside the State where appropriate; giving consideration to cultural and social aspects of the environment; and giving greater emphasis to research in secondary and regenerating forests, disturbed, agricultural and urban areas.

The marine environment is an important and often neglected part of biodiversity investigations, even though 95% of animal phyla are mainly or exclusively marine. Marine research can be more expensive to conduct than terrestrial research because of equipment and sampling requirements, however, this environment should be an integral part of the Biota program because of São Paulo's strategic location on the Atlantic Coast.

Relationship with international organizations with similar goals

Already, Biota has shown a keenness to be involved with relevant international organizations, and has taken a lead role in a number of them. One of the unique aspects of the Biota project has been its willingness to use international experts and advisers at all stages of the project, from its first initial steps through to their use on the Scientific Advisory Committee for annual evaluation. This has been a particular strength of the Biota project compared with similar projects in other countries. Some of the more important international projects that Biota could profitably be associated with are:

The Global Biodiversity Information Facility (GBIF) (<http://www.gbif.org>) is a product of the OECD (Organisation for Economic Cooperation and Development) Megascience Forum and was fully established in February, 2001. GBIF was established to design, implement, coordinate, and promote the compilation, linking, standardization, digitization and global dissemination of the world's biodiversity data, within an appropriate framework for property rights and due attribution (OECD 2000). It is designed to work in close cooperation with established programs and organizations that compile, maintain and use biological resources and information.

The Clearing House Mechanism (CHM) of the Convention on Biological Diversity (CBD) was established under Article 18, paragraph 3 of the Convention on Biological Diversity (United Nations 1993) (<http://www.biodiv.org/chm>). The CHM is intended to become the primary global cooperative information network on the conservation and sustainable use of biological diversity.

The Inter-American Biodiversity Information Network (IABIN) (<http://www.nbii.gov/iabin/index.html>) is an initiative of the Summit of the Americas established to provide greater access to biodiversity information throughout the hemisphere. Established in 1996, its primary objective is to promote the collection, access to, and exchange of technical, scientific, and supporting information on biodiversity in the Americas. The Biota project is in an ideal position to take a leading role in IABIN, and already has been instrumental in that organization.

The Species 2000 Programme (<http://www.species2000.org/>) is developing a dynamic species' checklist on the internet through a common access gateway, linking an array of on-line taxonomic databases. This gateway links directly to the much more detailed information in the distributed databases. Users can locate species by name using the dynamic on-line gateway or by using an annual checklist, also made available on the Internet and on CD-ROM.

The International Working Group on Taxonomic Databases (TDWG) (<http://www.tdwg.org/>) was started

in 1985 as an international working group to explore ideas on standardization and collaboration between major plant taxonomic database projects. TDWG has since expanded its scope to include taxonomic databases from all biological disciplines. There may be significant advantages to Biota in being a part of TDWG in assisting with the development and use of standards and interoperability of biodiversity databases.

The present state of Biota

The Scientific Advisory Committee for the FAPESP/Biota program have now made two annual evaluations of the program and on each occasion, have praised the efforts of Biota in reaching its present stage of development in just three years.

At each meeting, the Committee has attended verbal presentations by Project Leaders and held discussions with participants. It has made recommendations on the future of Biota, the overall aims of the Biota Program and on perceived gaps in the Program.

Following each meeting, members of the Committee have spent some time with key projects in order to get a better understanding of the details of those program, to work with participants in an interactive way and to discuss the projects in detail with them. As part of this process, I have spent some examining the databases and GIS systems that are now housed at the **Centro de Referência em Informação Ambiental (CRIA) created by the former BDT staff**. The SinBiota databases and other systems being used appear to be highly efficient and to some extent world-leading in their methodologies. There are many similarities and overlaps with the systems being used by **CRIA** and the Environmental Resources Information Network in Australia, and reflect the informal collaboration between these two institutions since the early 1990s. There are lessons to be learned from both organisations.

The future

The future of the Biota program is a bright one. The multidisciplinary nature of the Biota Program, the proactive nature of the Biota coordinating committee in actively seeking to identify and fill gaps in the program and the security of funding offered through FAPESP all contribute to what should be a long and successful future.

The multidisciplinary research that brings together scientists from different backgrounds and expertise is a key aspect of Biota's future. Many exciting findings are sure to arise from work at the interfaces between different biological disciplines and between these and other fields such as geology, geochemistry, climatology, modeling, etc. The diversity of projects being funded, and the varied emphases on different levels of biodiversity (genomic, species, ecosystem structure and function), as well as the positive and active involvement of young scientists and the involvement of the wider community through education and are sure to provide lasting benefits to the State's conservation needs.

Conclusion

The explosion in desktop computing, the introduction of PCs and the expansion of the Internet in the late 1980s and early 1990s, made possible the availability and distribution of information in a way that had previously only been dreamed about. It gave the wherewithal for massive amounts of biological information to be made available to researchers and the public alike in their own offices and homes.

Much has been said about the effect of the information age, the Internet and the World Wide Web on people's lives, education and business, but perhaps the greatest benefits has been in the area of biodiversity, and on the study of the environment and through the development of what has been termed 'biological informatics'.

The São Paulo FAPESP/Biota program is a world-leading program and is likely to prove a major element in the long-term conservation of the State's, and indeed the whole of Brazil's, natural resources. The program may even provide a template for similar programs across the continent. The Biota program does have the advantage of the backing of FAPESP, and it is pleasing to see biodiversity being recognized by FAPESP as a key program in its broader agenda.

The science in Biota is, in most projects, of high quality and equivalent to that in other countries, and in several projects is of outstanding quality and at the cutting edge. In many respects, the Biota program sets an example that many countries would be proud to follow.

I would like to praise the efforts of Biota in reaching its present stage of development in such a short time. It is a well-structured and well-coordinated program. The Coordenação Biota team, led by Professor Carlos Joly need to be congratulated.

ACKNOWLEDGMENTS

I wish to thank the director of FAPESP, Professor José Fernandez Perez and his staff, especially Carmen Weingrill, firstly for inviting me to be a member of the Scientific Advisory Committee for the BIOTA/FAPESP Program, but also for making me most welcome on my several visits to the State of São Paulo. Professor Carlos A. Joly, Dr Vanderlei Perez Canhos and Dr Dora Ann Lange Canhos have been most supportive of me at all times, not only during meetings of the BIOTA/FAPESP Program, but at many other times and in many places around the World. I regard them as true and trusted friends. The other members of the Coordenação Biota and all the Coordinators of BIOTA/FAPESP projects have made my visits to São Paulo most memorable and enjoyable, enough now for me to regard the State as a home away from home. Lastly, I thank Karl Bossard and Jeff Tranter of the Environmental Resources Information Network in Canberra for their comments on the manuscript.

LITERATURE

- Busby, J.R. (1979). *Australian Biotaxonomic Information System. Introduction and data interchange standards*. 25 pp. Canberra: Australian Government Publishing Service.
- Canhos, D.A.L., Canhos, V. and Kirsop, B. (1994) (eds). *Linking Mechanisms for Biodiversity Information*. Proceedings of the International Workshop held at the Tropical Database, Campinas, São Paulo, Brazil 23-25 February, 1994 [Also available electronically at <http://www.bdt.org.br/bin21/proceed94/procl.html>].
- Chapman, A.D. (1992). Quality Control and Validation of Environmental Resource Data, in *Data Quality and Standards: Proceedings of a Seminar Organised by the Commonwealth Land Information Forum, Canberra 5 December 1991*. Canberra: Commonwealth Land Information Forum.
- Chapman, A.D. (1997). *Workshop: Bases para a conservação da biodiversidade do estado de São Paulo*. View of Arthur D. Chapman [published electronically at: <http://www.biota.org.br/info/workshop/result/chapman>]
- Chapman, A.D. (2000). Quality Control and Validation of Point-Sourced Environmental Resource Data. pp. 409-418 in K.Lowell (ed.). *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*. Chelsea, MI: Ann Arbor Press.
- Chapman, A.D. and Busby, J.R. (1994). Linking Plant Species Information to Continental Biodiversity Inventory, Climate Modeling and Environmental Monitoring, pp. 179-195 in R.I.Miller (ed.). *Mapping the Diversity of Nature*. London: Chapman & Hall.
- Chernoff, B. (1997). *Workshop: Bases para a conservação da biodiversidade do estado de São Paulo*. Comments by Barry Chernoff [published electronically at: <http://www.biota.org.br/info/workshop/result/barry>]
- Colwell, R.K. (1997). *Workshop: Bases para a conservação da biodiversidade do estado de São Paulo*. Comments by Robert Colwell [published electronically at: <http://www.biota.org.br/info/workshop/result/colwell>]
- Freeman, N., Boston, T. and Chapman, A.D. (1999) Integrating Internal, Intranet and Internet Access to Spatial Datasets via ERIN's Environmental Data Directory in *Proceedings of the 26th Annual Conference of AURISA, Perth, Western Australia, 23-27 November 1998*. AURISA
- Green, D.G. and Croft, J.R. (1994). *Proposal for*

- Implementing a Biodiversity Information Network*. In *Linking Mechanisms for Biodiversity Information*. in Canhos, D.A.L., Canhos, V. and Kirsop, B. (eds), Proceedings of the Workshop for the Biodiversity Information Network, pp. 5-17. Fundacao Tropical de Pesquisas e Tecnologia "Andre Tosello", Campinas, São Paulo, Brazil.
- Hawke, R.J.L. (1989). *Our Country, Our Future*. 62 pp. Canberra: Australian Government Publishing Service.
- IMCRA Technical Group (1997). *Interim Marine and Coastal Regionalisation for Australia: an ecosystem-based classification for marine and coastal environments Versions 3.2*. Canberra: Environment Australia.
- OECD (2000). *Business Plan for the Global Biodiversity Information Facility*. [published electronically at <http://www.gbif.org/infofram.htm>].
- Potts, D.C. (1997). *Workshop: Bases para a conservação da biodiversidade do estado de São Paulo*. Comments by Donald C. Potts [published electronically at: <http://www.biota.org.br/info/workshop/result/potts>]
- Richardson, B.J. and McKenzie, A.M. (1991). *Australia's Taxonomists and Taxonomic Collections*. 44 pp. Canberra: Australian National Parks & Wildlife Service.
- Thackway, R. and Cresswell, I.D. (1995). *An Interim Biogeographic Regionalisation for Australia: A framework for setting priorities in the National Reserve Systems Cooperative Program*. Canberra: Australian Nature Conservation Agency.
- United Nations (1992). *Rio declaration on environment and development* : United Nations Conference on Environment and Development, 3-14 June 1992, Rio de Janeiro 8 pp. Rio de Janeiro; United Nations. [Also published electronically at <http://www.bdt.org.br/bin21/ws92/rio.declaration.html>]
- United Nations (1993). *Convention on Biological Diversity. Text and Annexes*. 34 pp. Châtelaine, Switzerland: Interim Secretariat for the Convention on Biological Diversity.

Title: Biodiversity Informatics, Biota/FAPESP and the future - a personal view

Author: Chapman, Arthur D. - ISSN 1676-0611

Revista Biota Neotropica, v 1, números 1 e 2 - 2001
<http://www.biotaneotropica.org.br/v1n12/pt/abstract?article+BN00601122001>

le+BN00601122001

Date Received: Sept 05, 2001 Accepted: Sept 05, 2001